

Надежность и эффективность В ТЕХНИКЕ

Справочник в десяти томах

Надежность и эффективность в технике

Справочник в десяти томах

Редакционный совет:

*В.С. АВДУЕВСКИЙ, акад. АН СССР
(председатель);*

В.И. КУЗНЕЦОВ, акад. АН СССР;

Н.Д. КУЗНЕЦОВ, акад. АН СССР;

В.А. МЕЛЬНИКОВ, акад. АН СССР;

В.П. МИШИН, акад. АН СССР;

В.Ф. УТКИН, акад. АН СССР;

К.В. ФРОЛОВ, акад. АН СССР;

Б.В. ГНЕДЕНКО, акад. АН УССР;

И.Н. КОВАЛЕНКО, акад. АН УССР;

Б.Ф. ЛОМОВ, чл.-корр. АН СССР

Том
3

Эффективность технических систем

*Под редакцией акад. АН СССР
В. Ф. УТКИНА,
д-ра техн. наук проф.
Ю. В. КРЮЧКОВА*



Москва
«МАШИНОСТРОЕНИЕ»
1988

ББК 30.14
Н17
УДК 621.192 (035)

Авторы тома: В. У. Торбин, Г. Н. Охотников, Е. С. Егоров,
Ю. В. Крючков, С. Н. Воробьев, А. Б. Назаров, В. В. Раздорский,
А. В. Ильичев, В. И. Кузнецов, В. Н. Харитонов, А. И. Яковлев

Рецензент д-р техн. наук проф. А. В. Журавлев

Надежность и эффективность в технике: Справочник:
Н17 В 10 т./Ред. совет: В. С. Авдуевский (пред.) и др. — М.:
Машиностроение, 1988. — (В пер.).

Т. 3. Эффективность технических систем/Под общ. ред.
В. Ф. Уткина, Ю. В. Крючкова. — 328 с.: ил. — 1 р. 80 к.

Изложены методологические основы исследования эффективности операций в технике. Приведены основные показатели и критерии эффективности и правила их выбора, принципы и методы имитационного моделирования при исследовании эффективности технических систем. Особое внимание уделено вопросам выработки решений на основе анализа эффективности операций. Рассмотрены методы обоснования решений при исследовании эффективности сложных технических систем.

Н 2702000000-611
038 (01)-88 — Подписное

ББК 30.14

СПРАВОЧНИК СПЕЦИАЛИСТА

Валерий Ульянович Торбин, Герман Николаевич Охотников,
Евгений Сергеевич Егоров и др.

НАДЕЖНОСТЬ И ЭФФЕКТИВНОСТЬ В ТЕХНИКЕ

Том 3

ЭФФЕКТИВНОСТЬ ТЕХНИЧЕСКИХ СИСТЕМ

Редактор И. И. Лесниченко. Художественный редактор С. С. Водчиц.

Переплет художника А. Я. Михайлова. Технические редакторы:

Н. Н. Скотникова, И. Н. Раченкова.

Корректоры: Т. В. Багдасарян, О. Ю. Садыкова.

ИБ № 4674

Сдано в набор 16.06.87. Подписано в печать 29.09.87. Т-14989.

Формат 60×90^{1/16}. Бумага офсетная № 2. Гарнитура литературная.

Печать офсетная. Усл. печ. л. 20,5. Усл. кр.-отт. 20,5. Уч.-изд. л. 27,25.

Тираж 12712. Заказ 162. Цена 1 р. 80 к.

Ордена Трудового Красного Знамени издательство «Машиностроение»,
107076, Москва, Стромынский пер., 4

Ленинградская типография № 6 ордена Трудового Красного Знамени
Ленинградского объединения «Техническая книга» им. Евгении Соколовой
Союзполиграфпрома при Государственном комитете СССР
по делам издательств, полиграфии и книжной торговли.
193144, г. Ленинград, ул. Монсеенко, 10.

Оглавление

Предисловие	9	2. Принципы системных исследований эффективности . . .	45
часть I		3. Общая схема исследований эффективности операций .	47
МЕТОДОЛОГИЧЕСКИЕ ОСНОВЫ ИССЛЕДОВАНИЯ ЭФФЕКТИВНОСТИ В ТЕХНИКЕ	10	Глава 4. ОБЩЕЕ ОПИСАНИЕ ПРОБЛЕМЫ ПРИНЯТИЯ РЕШЕНИЙ ПРИ ИССЛЕДОВАНИИ ЭФФЕКТИВНОСТИ ТЕХНИЧЕСКИХ СИСТЕМ (С. Н. Воробьев, Е. С. Егоров, В. У. Торбин)	51
Глава 1. ЭФФЕКТИВНОСТЬ ОПЕРАЦИЙ В ТЕХНИКЕ (Г. Н. Охотников)	10	1. Проблема принятия решений	51
1. Основные положения теории эффективности . . .	10	2. Модель проблемной ситуации	52
2. Факторы, определяющие эффективность операций в технике . . .	17	3. Задачи принятия решений	54
3. Шкалы и допустимые преобразования показателей (В. У. Торбин)	24	4. Классификация задач принятия решений	59
Глава 2. ПОКАЗАТЕЛИ И КРИТЕРИИ ЭФФЕКТИВНОСТИ (Г. Н. Охотников)	30	Глава 5. ОСОБЕННОСТИ ИССЛЕДОВАНИЯ ЭФФЕКТИВНОСТИ ТЕХНИЧЕСКИХ СИСТЕМ НА ЭТАПАХ ЖИЗНЕННОГО ЦИКЛА (Г. Н. Охотников, А. В. Ильичев).	62
1. Показатель эффективности	30	1. Жизненный цикл технической системы . . .	62
2. Формы показателя эффективности . . .	31	2. Схема обобщенной операции	64
3. Критерий эффективности (В. Н. Харитонов).	37	3. Оценивание эффективности технической системы в период ее создания	69
Глава 3. ОСНОВНЫЕ ПРИНЦИПЫ ИССЛЕДОВАНИЯ ЭФФЕКТИВНОСТИ В ТЕХНИКЕ (Г. Н. Охотников, В. И. Кузнецов, В. У. Торбин)	42		
1. Методологические уровни исследования эффективности технических систем . . .	42		

ЧАСТЬ II

МОДЕЛИРОВАНИЕ И ОЦЕНИВАНИЕ ЭФФЕКТИВНОСТИ ТЕХНИЧЕСКИХ СИСТЕМ

Глава 6. ОЦЕНИВАНИЕ ЭФФЕКТИВНОСТИ ОПЕРАЦИЙ (Г. Н. Охотников, В. У. Торбин)

1. Содержание задачи оценивания эффективности . . . 72
2. Классификация методов моделирования . . . 76
3. Принципы моделирования операций в технике . . . 79
4. Классификация моделей оценивания эффективности 80

Глава 7. МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ОПЕРАЦИЙ (Г. Н. Охотников, Ю. В. Крючков)

1. Общая схема операции . . . 83
2. Математическое описание операции . . . 84
3. Частные описания операций 88
4. Математические модели технических систем (А. Б. Назаров) 93

Глава 8. ИМИТАЦИОННЫЕ МОДЕЛИ ИССЛЕДОВАНИЯ ЭФФЕКТИВНОСТИ (В. В. Раздорский, Ю. В. Крючков, Г. Н. Охотников)

1. Классификация имитационных моделей . . . 99
2. Моделирование случайных факторов (А. Б. Назаров) 101
3. Программная реализация имитационных моделей . . . 109
 - Требования к программному обеспечению имитационных исследований . . . 109
 - Виды программной реализации имитационных моделей . . . 112

4. Проблемно-ориентированные имитационные системы исследования эффективности 116
 - Состав и структура имитационной системы . . . 116
 - Программное обеспечение имитационной системы . . . 117
 - Информационное обеспечение имитационной системы 121
 - Диалоговое обеспечение имитационной системы (А. И. Яковлев) . . . 123
 - Технологическая схема решения задач средствами имитационной системы 124

Глава 9. ОЦЕНКА КАЧЕСТВА МОДЕЛЕЙ И ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТОВ (Г. Н. Охотников, А. Б. Назаров)

1. Оценка качества моделей 127
2. Методы повышения качества оценок показателей эффективности . . . 129
 - Пассивные методы повышения качества оценивания показателя эффективности 131
 - Активные методы повышения качества оценивания показателя эффективности . 133
 - Косвенные методы повышения качества оценивания показателя эффективности . 134
3. Планирование имитационных экспериментов 136

ЧАСТЬ III

МЕТОДЫ ВЫРАБОТКИ РЕШЕНИЙ (С. Н. Воробьев, Е. С. Егоров, В. У. Торбин)

Глава 10. ПРОЦЕСС ВЫРАБОТКИ РЕШЕНИЙ ПРИ ИССЛЕДОВАНИИ ЭФФЕКТИВНОСТИ И КАЧЕСТВА ТЕХНИЧЕСКИХ СИСТЕМ.

1. Вертикальная декомпозиция процесса выработки решений при исследовании эффективности и качества . . . 148

2. Особенности выработки решений при проведении исследований по уровням «вертикальной» декомпозиции	151	Приложение. Элементы теории нечетких множеств	220
3. Требования к процессу выработки решения при проведении исследований по уровням «вертикальной» декомпозиции	156	Глава 12. ВЫРАБОТКА РЕШЕНИЙ В УСЛОВИЯХ ОПРЕДЕЛЕННОСТИ	222
4. Общая схема процесса принятия решений на заданном уровне декомпозиции исследования	162	1. Задачи принятия решений по скалярному показателю	222
Глава 11. ВЫЯВЛЕНИЕ И ИЗМЕРЕНИЕ ПРЕДПОЧТЕНИЙ	167	Общая постановка задачи.	222
1. Предпочтения и их свойства	168	Формализация задачи выбора	224
Понятие предпочтения	168	Типы задач принятия решений по скалярному показателю в условиях определенности	225
Способы выражения предпочтений	170	2. Задачи принятия решений по векторному показателю	230
Отношения	173	Общая постановка задачи	230
Нечеткие отношения	177	Эффективные стратегии	232
2. Оценивание предпочтений	179	Омега-эффективные стратегии	237
Отношения предпочтения и их свойства	179	Особые случаи омега-информации о предпочтениях	241
Нечеткие отношения предпочтения и их свойства	180	Одношаговые методы решения задач с векторным показателем эффективности	252
Функция выбора и ядро отношения	184	Многошаговые методы решения задач с векторным показателем эффективности	256
Функция эффективности	186	Глава 13. ВЫРАБОТКА РЕШЕНИЙ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ	266
3. Выявление и измерение предпочтений методами экспертного оценивания	187	1. Задачи принятия решений в условиях стохастической неопределенности	266
Общая характеристика и классификация методов экспертного оценивания	187	Общая постановка задачи	266
Организация и проведение экспертного оценивания	189	Задача принятия решений в условиях стохастической неопределенности по скалярному показателю	267
Обработка и анализ ранжировок	196	Задачи принятия решений в условиях стохастической неопределенности по векторному показателю	287
Обработка и анализ балльных и точечных оценок	203	2. Задачи принятия решений в условиях нестохастической неопределенности	296
Обработка и анализ попарных сравнений	206		
Определение коэффициентов относительной важности	207		
Комплексный метод выявления и измерения предпочтений	211		
Формирование и формализация экспертной информации в форме нечетких суждений	213		

Общая постановка задачи	296	в условиях поведенческой	
Задачи принятия решений		неопределенности	311
в условиях «природной» не-		Матричные игры	313
определенности	297	Биматричные игры	318
Задачи принятия решений		Игры в нечеткой постанов-	
в условиях стохастической		ке	320
и «природной» неопределен-		Список литературы	322
ности	307	Предметный указатель	325
3. Задачи принятия решений			

Предисловие

Третий том справочника посвящен вопросам исследования эффективности технических систем и содержит основные сведения по теории эффективности операций, необходимые в первую очередь для практической реализации системного подхода к исследованию и обеспечению надежности техники, а также для повышения эффективности организационно-технических решений в процессе создания и применения технических систем.

В первой части тома в систематизированном виде изложены важнейшие методологические положения и принципы исследований эффективности операций в технике. Последовательно раскрыты основные понятия теории эффективности операций, факторы, определяющие эффективность, и способы их измерения. Приведены формы показателей эффективности, применяемые в технике, и критерии эффективности, формируемые при различных концепциях рационального поведения системы в операции. Дана общая схема анализа эффективности операций в технике, в основу которой положены методологические уровни и принципы системных исследований эффективности. Рассмотрены общие проблемы принятия решений и особенности оценивания эффективности технических систем на этапах жизненного цикла.

Вторая часть тома содержит методические материалы по моделированию и оцениванию эффективности технических систем. В ней дано описание задачи оценивания эффективности, основных принципов и классификационных признаков моделирования. Приведена общая математическая модель операции как кибернетической системы, из которой в качестве частных случаев следуют различные модели конкретных операций. Изложены основы имитационного моделирования операций в технике и принципы построения проблемно-ориентированных имитационных систем, являющихся основным средством исследования эффективности сложных объектов.

Даны рекомендации по оценке качества моделей и планированию имитационных экспериментов.

В третьей части тома рассмотрены методические вопросы количественного обоснования организационно-технических решений на основе анализа результатов оценивания эффективности операций. Показана структура процесса выработки решений при исследовании эффективности технических систем. Значительное внимание уделено вопросам формализации слабо структуризованных процедур выявления предпочтений и принятия решений. Приведены методы и алгоритмы экспертного оценивания и даны рекомендации по их практическому использованию. Изложены методические подходы к выработке решений по скалярному и векторному показателям эффективности в условиях различной степени неопределенности стохастической и нестохастической природы, а также в условиях неопределенности поведенческого характера. Важное место отведено вопросам использования качественной информации при формировании составных критериев эффективности с целью выработки рациональных организационно-технических решений.

Все примеры, приведенные в третьем томе, являются гипотетическими и носят иллюстративный характер.

Теория эффективности операций, завершая первый этап своего становления, продолжает интенсивно развиваться: совершенствуется методологическое основание этой теории, ее математический аппарат, расширяется сфера прикладных исследований эффективности с целью научного обоснования решений в различных сферах народнохозяйственной деятельности. Авторы тома систематизировали и изложили лишь те сведения из теории эффективности операций, которые необходимы в первую очередь специалистам в области надежности, но недостаточно полно раскрыты в специальной литературе.

Часть I

МЕТОДОЛОГИЧЕСКИЕ ОСНОВЫ ИССЛЕДОВАНИЯ ЭФФЕКТИВНОСТИ В ТЕХНИКЕ

Глава 1. Эффективность операций в технике

1. ОСНОВНЫЕ ПОЛОЖЕНИЯ ТЕОРИИ ЭФФЕКТИВНОСТИ

Техническая система создается для удовлетворения определенных потребностей человека и служит активным средством в его целенаправленной деятельности.

В процессе трудовой деятельности человека постоянно возникают проблемы различной сложности. Причиной возникновения проблемы является расхождение между желаемым и действительным результатом при известных путях преодоления этого расхождения (несоответствия). Для решения проблемы необходимо выделить и достаточно четко сформулировать цели деятельности, осуществление которых существенно снижает или устраняет различие между желаемым и действительным результатом, т. е. решает проблему.

Глобальная цель, как правило, допускает декомпозицию, в результате которой формируются взаимосвязанные частные цели, которые в общем случае могут быть подвергнуты дальнейшему членению на более простые составляющие (подцели, задачи).

Исследователь, анализируя возникшую проблему, среди объектов материального мира выделяет некоторую целостность — систему, предназначенную для решения этой про-

блемы. Система представляет собой совокупность элементов (предметов любой природы), находящихся в отношениях и связях друг с другом. При объединении элементов в систему последняя приобретает специфические системные свойства, не присущие ни одному из элементов. Эти свойства называют *интегративными* или *эмерджентными*. Свойства системы шире суммы свойств составляющих ее элементов.

Простые проблемы решаются в рамках простых систем; для решения сложных, крупных проблем требуется создание сложных, больших систем. Деление систем на простые и сложные условно. Четких граней между уровнями сложности систем нет. И тем не менее введением некоторых признаков сложности удается обнаружить качественное различие между простыми и сложными системами. Одним из таких признаков является предсказуемость поведения, свойств и реакций на внешние воздействия. Системы со слабо предсказуемым поведением относят к сложным. Сложным системам присуща способность принимать решения. Это определяющее отличие сложных систем от простых. Последние не обладают такой возможностью.

Проектирование, строительство и эксплуатация крупных автоматизиро-

ванных производственных, энергетических, гидротехнических, транспортных, информационных и других больших систем видвинули актуальные проблемы, которые можно решить лишь при комплексном учете различных по своей природе факторов, многообразия отношений и связей между ними, внешних условий и т. д. Методологическим основанием подготовки и обоснования решений по сложным проблемам научного, экономического и технического характера является системный анализ. Использование методологии системного анализа при проектировании больших технических систем привело к появлению *системотехники*. Привлечение методологических средств системного анализа для решения крупных научно-технических проблем обусловлено прежде всего тем, что приходится осуществлять выбор в условиях неопределенности, вызванной наличием факторов, не поддающихся строгой количественной оценке. Приемы и методы системного анализа направлены на выдвижение альтернативных вариантов решения проблемы, выявление масштабов неопределенности по каждому варианту и сопоставление вариантов по их эффективности. Системный анализ рекомендует начинать процесс принятия решений с выявления и четкого формулирования конечных целей; рассматривать проблему как единую систему и выявлять все последствия и взаимосвязи каждого частного решения; согласовывать цели подсистем с общей целью системы; выявлять и анализировать возможные альтернативные пути достижения цели и выбирать из них наиболее эффективные.

Исходя из рекомендаций системного анализа, на начальной стадии исследований следует всесторонне изучить проблему, выявив ее масштабы, актуальность, источник возникновения, связи с другими проблемами, наличные (или потребные) ресурсы для ее решения, возможность решения в приемлемые сроки. Все эти вопросы составляют предмет проблемного анализа, результатом которого является формирование системы согласованных или альтернативных целей деятельности, направленной на решение проблемы.

Для достижения поставленной цели необходима целенаправленная деятельность — операция. Операция есть система целенаправленных действий, объединенных общим замыслом и единой целью. Понятие операции включает по меньшей мере три определяющих момента:

1) управляющая деятельность человека (органа управления, распорядительного центра, лица, принимающего решения), организующего операцию на основе выбора рационального способа использования активных средств для достижения цели операции;

2) активные средства (технические системы, ресурсы), находящиеся в распоряжении управляющего органа и используемые в операции в соответствии с выбранным способом (стратегией) управления;

3) другие средства (системы), непосредственно взаимодействующие с активными средствами, к которым обычно относят объекты воздействия активных средств, средства, находящиеся в распоряжении других распорядителей в операции (их активные средства).

В общем плане эти три момента отражают ответы на вопросы, как действовать, чем действовать и на что воздействовать для успешного достижения поставленной цели операции. Операция формируется в рамках S_0 системы, которая в качестве основных компонентов содержит орган управления операцией, активные средства и объект операции (объект воздействия). По отношению к S_0 -системе цель A_0 операции выступает основным системообразующим фактором как способ интеграции различных действий в единую последовательность (принцип целеобусловленности). Цель есть идеальное представление (предвосхищение) в сознании руководителя желаемого результата операции. Она определяет способы и формы действий, их характер и системную упорядоченность, а также средства достижения и выступает как определенный механизм интеграции различных действий в систему «цель — средство — результат». В формализованном виде цель выражается набором некоторых параметров целеполагания $Y_{тр}$ (в частном случае всего один параметр может отражать

цель операции). Функцию целеполагания может осуществлять старшая система, в состав которой входит исследуемая. В этом случае старшая система ставит задачу, т. е. осуществляет руководство исследуемой системой. Реальный результат Y операции (фактический или ожидаемый ее конечный итог) может не совпадать с желаемым. Под эффективностью операции понимают степень различия между реальным ее результатом Y и желаемым результатом $Y_{тр}$.

Возможные альтернативные пути достижения цели в общем случае обладают разной эффективностью. Системный анализ рекомендует сравнивать эти пути между собой и выбирать из них лучший на основе эффективности, к которой приводит тот или иной путь (вариант действий). Поэтому эффективность операции является важной категорией системного анализа.

К активным средствам операции относятся технические средства (орудия, машины, аппаратура управления, технические комплексы, большие технические системы), а также ресурсы (вещественные, энергетические, информационные, временные, денежные и т. п.). Технические средства входят в состав S_0 -системы в качестве ее подсистем (агрегатов) или элементов. Эти подсистемы, как правило, управляемы. Управляемая подсистема представляет собой динамический объект, который изменяет свои состояния при наличии некоторых внутренних закономерностей его функционирования, под воздействием внешней среды, а также команд (управляющих воздействий), поступающих от управляющей системы (системы управления) по каналам прямой связи. По каналам обратной связи управляющая система (орган управления) получает сообщение о состоянии управляемой подсистемы. Орган управления управляет динамическими объектами, соотносясь с общей целью S_0 -системы, с командами от внешней среды и с информацией, получаемой от управляемых подсистем, руководствуясь определенными принципами управления.

Система управления может содержать несколько органов управления (распорядительных центров) и иметь

иерархическую структуру. В состав распорядительных центров включаются люди (распорядители или субъекты системы), принимающие решения и реализующие на основе этих решений управляющие воздействия на объекты управления. В общем случае управлять объектами могут несколько распорядителей. Более того, эти распорядители могут ставить перед собой разные цели. Их взаимодействия в процессе управления операцией могут быть также различны — от содействия до противодействия.

Исследование операции проводится всегда с точки зрения интересов одного распорядителя (основного субъекта системы), которого называют лицом, принимающим решение (ЛПР). Цель A_0 , стоящая перед ЛПР, является той единственной целью, для достижения которой проводится операция. Цели, стоящие перед другими субъектами системы, могут быть согласованы с основной целью операции, способствовать ее достижению, а могут и противоречить основной цели, противодействовать ее осуществлению. Набор всевозможных взаимоисключающих (альтернативных) способов использования активных средств составляет множество допустимых стратегий U . Выбор стратегии из множества допустимых стратегий U является основным этапом принятия решения.

Процесс подготовки и принятия решений ЛПР, организация их выполнения и контроль составляют сущность управления операцией. *Управление* есть процесс формирования рационального (разумно обоснованного) поведения системы в операции. Основная *цель управления* состоит в том, чтобы обеспечить максимальную эффективность использования активных средств в операции при решении поставленной задачи (при достижении цели). ЛПР управляет операцией в доступных ему пределах, выбирая стратегии из допустимого множества стратегий U .

Целью исследования эффективности операции может явиться выработка рекомендаций лицу, принимающему решения, для рационального выбора стратегий, обеспечивающих с его

точки зрения успешный исход операции. Остальные субъекты системы также могут принимать решения, выбирая стратегии из соответствующих множеств допустимых стратегий, и тем самым могут управлять в доступных им пределах операцией, преследуя свои цели. Исследователь отражает их действия в виде соответствующих гипотез поведения.

Таким образом, формируя S_0 -систему для решения определенной проблемы, исследователь включает в ее состав все то, что непосредственно влияет на ход и исход операции, т. е. на достижение желаемого результата. В качестве активных средств в состав S_0 -системы входят технические системы, находящиеся в непосредственном взаимодействии между собой и внешней средой. Управлять этими системами могут распорядители (субъекты системы), среди которых выделяют основного распорядителя (ЛПР). Успешность операции (ее эффективность) рассматривается только с позиции ЛПР. Цель, стоящая перед ним, является целью операции.

Операция, как процесс функционирования S_0 -системы, описывается набором определенных параметров. Совокупность конкретных значений этих параметров в фиксированный момент времени называют *состоянием системы*. Функционирование системы есть процесс смены состояний. В каждый момент времени система может находиться только в одном состоянии. Состояние S_0 -системы и внешней среды в определенный момент времени называют *ситуацией или обстановкой операции*. Оценка обстановки перед началом операции и в процессе ее проведения, прогноз в изменениях обстановки являются важными этапами принятия решений. Ситуация (обстановка), сложившаяся к исходу операции, обычно определяет фактический (реальный) результат операции. Совокупность факторов, существенно влияющих на изменение обстановки (обуславливающих ее), называют *условиями обстановки*.

Наряду с качеством активных средств и способами их использования в операции условия обстановки операции являются основными составляющими

(слагаемыми) эффективности операции. Высокая эффективность операции обусловлена рациональным способом использования высококачественных активных средств в благоприятно сложившихся условиях обстановки.

Операция представляет собой обмен, в результате которого S_0 -система за приобретенную для себя пользу (полезный или целевой эффект) расплачивается некоторым количеством ресурсов C , т. е. терпит определенные издержки, выражаемые, например, стоимостью этих ресурсов или затраченным временем T на проведение операции. Применительно к этому эффективность операции характеризуется выгодностью для S_0 -системы такого обмена. Результат операции, таким образом, характеризуется не только полезным эффектом q (получением некоторой прибыли, достижением или сохранением желаемого состояния и т. д.), но и связанными с этим действием затратами ресурсов, времени и другими потерями, понесенными при проведении операции. Очевидно, чем ближе результат обмена к предельно выгодному, тем эффективнее операция. Предельно выгодный обмен характеризует потенциальную эффективность операции.

Результат Y операции ставят в зависимость от основных результирующих факторов — полезного эффекта q , затраченных ресурсов C и времени T . В свою очередь, результирующие факторы зависят от выбранной стратегии. Следовательно, результат операции также будет зависеть от стратегии u : $Y(u) = Y(q(u), C(u), T(u))$.

В общем случае функция $Y(u)$ может быть вектором, компоненты которого характеризуют результаты частных операций, направленных на достижение частных целей, на которые расчленяется общая цель операции.

Эффективность операции есть обобщенное — определяющее функциональное свойство S_0 -системы, реализующей операцию, которое с гносеологической (познавательной) точки зрения раскрывается через категорию цели (желаемый результат) и объективно выражается степенью достижения цели с учетом затрат ресурсов

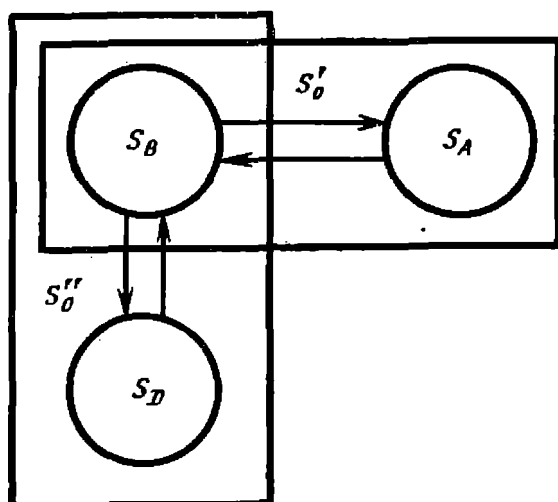


Рис. 1. Схема многоцелевой системы

и времени на реализацию операции. Любую операцию (S_0 -систему) можно рассматривать как составную часть более широкой операции, проводимой надсистемой, в которую входит S_0 -система. Надсистема стремится достичь определенную глобальную цель, организовав свою деятельность в соответствии с некоторой политикой, выбранной из множества допустимых политик (альтернатив, вариантов действий надсистемы). Эффективность исследуемой операции (в S_0 -системе) предопределяется ее полезностью в проводимой надсистемой политике, направленной на достижение глобальной цели. Эффективность S_0 -системы не выводится полностью из свойств S_0 -системы. Необходимо учитывать свойства надсистемы [21].

Потенциальная эффективность операции, характеризующая предельно выгодным обменом [67], определяется как эффективность операции при идеальном способе использования активных средств, т. е. выборе лучшей стратегии. Следовательно, потенциальная эффективность операции зависит лишь от качества активных средств. Поэтому, когда говорят об эффективности технической системы, которая используется как активное средство в некоторой операции, то имеют в виду потенциальную эффективность этой операции.

Операция как целенаправленная деятельность имеет единственную цель. Следовательно, S_0 -система,

в которой проводится операция, одноцелевая. Исследуемая техническая система как активное средство операции может быть многоцелевой. Однако относительно каждой цели, в достижении которой может участвовать многоцелевая техническая система, необходимо вводить S_0 -систему и в последней устанавливать эффективность операции. На рис. 1 приведена схема многоцелевой системы.

Пусть техническая система S_B является активным средством операции, проводимой в системе S'_0 , объектом воздействия в которой служит система S_A . Кроме того, техническая система S_B может быть использована в качестве активного средства в другой операции, проводимой в системе S''_0 , объектом воздействия в которой служит техническая система S_D .

Это разные операции. Их цели не согласованы. В таком случае рассчитывают эффективность двух разных операций. Может оказаться, что в первой операции исследуемая техническая система обладает высокой эффективностью, т. е. является высокоэффективным средством достижения первой цели, тогда как во второй операции эффективность низкая и эта техническая система является недостаточно эффективным средством достижения второй цели. Говоря об эффективности технической системы, имеют в виду эффективность использования ее в качестве активного средства в типовой операции, для проведения которой предназначена эта система. Если многоцелевая техническая система по каждой цели из их перечня имеет достаточно высокую эффективность, то это свидетельствует о том, что система имеет широкие функциональные возможности, т. е. обладает качествами, отражающими способность системы решать задачи, связанные с достижением каждой цели.

В прикладных исследованиях эффективности технических систем обычно приходится решать проблемы:

1) оценки эффективности операции с использованием технической системы (проблема оценки);

2) выбора рационального способа (стратегии) использования техниче

ских средств (систем) в операции (проблема выбора).

Оценка эффективности операции заключается в выработке так называемого оценочного суждения относительно пригодности заданного способа действий или приспособленности технических средств к решению определенных задач на основе измерения (оценивания) уровня эффективности операции.

Эффективность операции в технике оценивается с целью решения следующих задач:

принятия решения относительно допустимости практического использования оцениваемого способа действий в той или иной ситуации;

выявления вкладов (эффектов) различных факторов в общую эффективность операции, влияния взаимодействий факторов на эффективность;

установления путей повышения эффективности операции (выявления резервов эффективности);

выявления функциональных возможностей технических средств, используемых в операции;

сопоставления (сравнения) нескольких альтернативных вариантов действий или технических средств, их ранжирование по уровням эффективности (установление отношения предпочтения на множестве возможных вариантов).

Последняя задача часто выступает как подпроблема оценки эффективности. Ее называют *задачей сравнительной оценки вариантов* (способов действий, систем).

В общей проблеме выбора рационального способа использования активных средств в операции или путей развития технических средств выделяют задачи:

выбора целесообразного способа управления техническими средствами с заданными функциональными характеристиками;

определения рационального режима эксплуатации технических систем;

выработки оптимального плана операции;

оптимального распределения ресурсов между подсистемами в операции;

выбора рационального варианта проектируемой технической системы;

формирования программы развития больших технических систем;

выработки технико-экономических требований к создаваемым средствам (системам);

выдвижения гипотез рационального поведения подсистем, находящихся во взаимодействии с исследуемой.

Практическая направленность исследований эффективности в технике заключается в выработке решений на рациональное использование технических средств при достижении цели в разных условиях обстановки или на целесообразный вариант проектируемой технической системы.

В технике проблема эффективности тесно связана с проблемой надежности технических средств. Возрастание сложности технических систем приводит к снижению их надежности, а следовательно, к уменьшению их эффективности. Ненадежная система не может быть эффективным средством достижения поставленной цели. Анализ эффективности операции позволяет подойти к определению требований к параметрам надежности технических систем, используемых в качестве активных средств в операции. Это один из важных аспектов связи проблем надежности и эффективности в технике. Другим аспектом связи этих проблем является определение рациональных способов использования имеющихся средств для повышения надежности создаваемых технических систем. Недостаточная надежность проектируемой или существующей системы может явиться проблемой, для решения которой выдвигаются альтернативные цели (например, отказ от производства системы или замена ее новой, усовершенствованной; повышение надежности существующей системы до требуемого уровня; улучшение условий эксплуатации существующей системы и т. д.), осуществляется выбор одной из этих целей и организуется операция для достижения выбранной цели. Эффективность этой операции может быть оценена, например, вероятностью безотказной работы системы в течение заданного времени, если в качестве цели операции

выбрано повышение надежности системы. В таком случае показатель надежности системы выступает в качестве показателя эффективности операции по повышению надежности существующей или проектируемой технической системы.

Цель A_0 является исходным пунктом исследований эффективности операций. Для выбора рационального пути ее достижения необходимо иметь способ измерения эффективности операции. Показатель эффективности измеряет степень соответствия реального результата операции желаемому. На основе показателя эффективности формируется определенное правило выбора рационального способа использования активных средств в операции (стратегий). В качестве такого правила используется критерий эффективности. Критерий выбирается на основе определенных принципов рационального целенаправленного поведения. В зависимости от сложности S_0 -системы, реализующей операцию, и целей исследований вводятся различные принципы рационального поведения. Приходится устанавливать разные методологические уровни исследования эффективности операций для систем различной степени сложности.

При исследовании эффективности операций используют методологические подходы, основанные на эксперименте и моделировании. Экспериментальный подход предполагает проведение серии натурных экспериментов с реальной системой. Результаты экспериментов обрабатываются с использованием методов математической статистики (см. т. 2 справочника). Эксперимент обычно рационально планируется в целях получения наибольшего количества информации об исследуемом явлении. Математическая теория эксперимента [45, 66] позволяет выработать оптимальные планы экспериментов.

Экспериментальный подход к исследованию эффективности крупномасштабных операций ограничен. Сложность и масштабы современных технических систем, как правило, не позволяют проводить эксперименты с ними, а эксперимент с элементами

системы или с ее агрегатами не позволяет получить представление об эмерджентных свойствах системы. Основной исследовательской концепцией анализа эффективности операций является моделирование.

Эксперимент с самой системой заменяется экспериментом с ее моделью. Операция, реализуемая S_0 -системой, имитируется в рамках модели S_0 -системы.

При моделировании систем, уровень сложности которых не слишком высок, часто удается ввести математическую зависимость показателя эффективности от управляемых переменных (построить модель) в форме достаточно простой функции с системой ограничений в виде равенств или неравенств и задачу выбора рациональной стратегии поведения свести к одному из классов задач математического программирования. Выбор лучшей стратегии в этом случае обычно сводится к решению оптимизационной задачи одним из методов математического программирования. Некоторые распространенные методы математического программирования приведены в т. 2 справочника. В т. 3 главное внимание уделено исследованию эффективности сложных технических систем, используемых в крупномасштабных операциях. Основной формой изучения этих систем по общему признаку является имитационное исследование [43], в основе которого лежит имитационное моделирование. Развитие электронно-вычислительной техники открывает широкие возможности имитационного исследования крупномасштабных операций в различных областях техники.

Процесс выработки решений в сложных ситуациях изучается современной теорией решений. В рамках этой теории формируются принципы принятия решений, на основе которых вырабатываются подходы к отбору совокупности допустимых стратегий, к формальному описанию предпочтений лица, принимающего решения, к выбору принципов компромисса при наличии противоречивых интересов субъектов системы в условиях различной степени неопределенности обстановки и, наконец, к выбору рации-

ональных способов использования активных средств в операции на основе выдвинутых критериев эффективности операций.

Таким образом, в качестве предмета теории эффективности операций выступают закономерности, связывающие эффективность операции с качеством активных средств, способами их использования в операции и условиями обстановки. Объектом этой теории является операция.

2. ФАКТОРЫ, ОПРЕДЕЛЯЮЩИЕ ЭФФЕКТИВНОСТЬ ОПЕРАЦИЙ В ТЕХНИКЕ

Эффективность технических систем (ТС) определяется множеством различных по своей природе факторов. Под *фактором* понимают движущую силу какого-либо процесса (явления) или условие, которое влияет на тот или иной процесс (явление).

При исследовании эффективности технических систем, как правило, выделяют три группы факторов: качество ТС, условия функционирования ТС, способы использования (применения) ТС.

Для достижения определенной единственной цели создают систему. Рассматривают операцию, в которой анализируемая система является активным средством достижения поставленной цели. Определяют потенциальную эффективность операции при идеальном способе использования технической системы. Эту потенциальную эффективность и принимают за характеристику качества технической системы. Очевидно, показатель эффективности введенной операции является показателем качества технической системы. Таким образом, положительными (полезными) свойствами технической системы являются свойства, способствующие достижению заданной цели.

Для наглядного объяснения понятия качества технической системы удобно воспользоваться упрощенной геометрической моделью. На рис. 2 изображена схема направленности свойств S_B -системы, на которой показана совокупность векторов в некотором про-

странстве свойств. Эта совокупность свойств характеризует качество системы в общем смысле. Предположим, что цель A_0^1 определяет назначение системы, которая является активным средством в операции по достижении этой цели. Тогда цель A_0^1 задает определенную направленность свойств, т. е. определяет направление в пространстве свойств (линию цели $S_0 A_0^1$). Полезными в смысле достижения системой цели A_0^1 назовем свойства, отображаемые векторами, проекции которых на заданную линию цели направлены на цель. Противоположно направленные проекции имеют векторы, отображающие вредные в рассматриваемом смысле свойства. Нейтральные свойства отображаются векторами, ортогональными к линии цели.

Предположим, что техническая система является системой многоцелевого назначения. В этом случае необходимо последовательно рассмотреть все цели, для достижения которых предназначена техническая система. По каждой из них ввести соответствующую операцию (S_0^i -систему для достижения цели A_0^i), в которой исследуемая система S_B используется как активное средство, а затем по каждой операции определять ее потенциальную эффективность. Вектор показателей потенциальных эффективностей введенных операций является векторным показателем качества технической системы.

На рис. 2 вторая цель A_0^2 , для достижения которой также предназначена S_B -система, формирует в пространстве свойств другое направление (линия цели $S_B A_0^2$). Совокупность свойств S_B -системы, отображаемых векторами, проекции которых на линию цели $S_B A_0^2$ направлены на цель A_0^2 , является множеством полезных свойств системы S_B в смысле достижения цели A_0^2 . Эти свойства обуславливают потенциальную эффективность операции, направленной на достижение цели A_0^2 . Таким образом, качество технической системы в данном случае определяется двумерным показателем, компонентами которого являются показатели

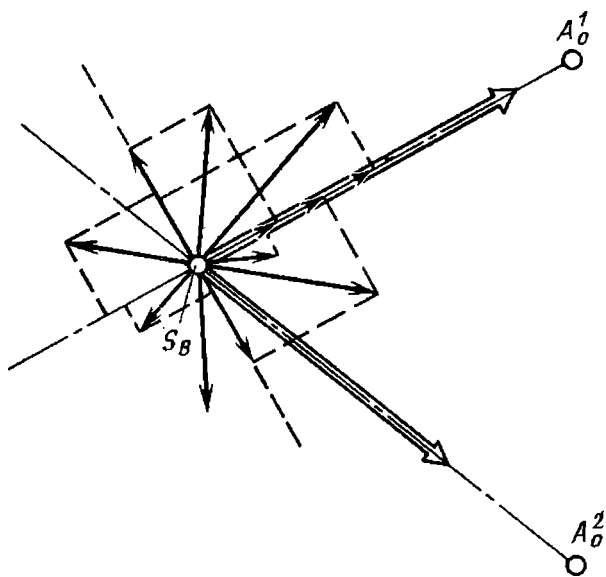


Рис. 2. Схема направленности свойств системы

потенциальной эффективности в операции по достижению цели A_0^1 и в операции, направленной на достижение цели A_0^2 .

Описанный подход к определению качества ТС не является всеобщим. При таком подходе полезность системы целиком определяется ее способностью достичь определенную цель. Но полезность в этом смысле не может охватить все возможные подходы к определению качества объекта. В практике довольно часто качество объектов определяют с позиции его ценности, которую далеко не всегда связывают с целевым назначением рассматриваемого объекта. Однако качество технических средств достаточно полно определяется как совокупность полезных, с точки зрения целевого назначения, свойств этих средств. Поэтому, когда говорят об эффективности технической системы, понимают ее потенциальную эффективность как основную характеристику качества системы.

В практике оценивания качества технических систем не всегда используют в полном объеме подход, основанный на определении потенциальной эффективности операции. В зависимости от сложности системы, цели исследования признают целесообразным введение нескольких уровней качества. Эмпирически установлены

уровни качества, получившие названия: устойчивость, помехоустойчивость, управляемость, способность, самоорганизация.

Первичным качеством любой системы является ее *устойчивость* (R — качество) [67]. Системы, не обладающие этим качеством, не могут существовать. Для простых систем устойчивость объединяет такие их свойства, как прочность, стойкость к воздействию внешних факторов, сбалансированность, стабильность, гомеостазис (способность системы возвращаться в равновесное состояние при выводе из него внешними воздействиями). Для сложных систем характерны различные формы структурной устойчивости, такие, как надежность, живучесть и т. д.

Другим качеством систем, более сложным, чем устойчивость, является *помехоустойчивость* (I — качество), понимаемая как способность системы без искажений воспринимать и передавать по каналам сообщений информационные потоки. Помехоустойчивость, другими словами, есть способность системы противодействовать помехам. Помехоустойчивость объединяет ряд свойств, присущих в основном системам управления. К таким свойствам относятся надежность систем связи и систем переработки информации, их пропускная способность, возможность эффективного кодирования и декодирования информации, электромагнитная совместимость радиоэлектронных средств и т. д.

Следующим качеством системы является *управляемость* (C — качество). Под *управляемостью* понимают способность системы переходить за конечное (заданное) время из одного состояния в другое (требуемое) под влиянием управляющих воздействий. В общем смысле управляемость есть способность системы выполнять команды управления, оперативно реагировать на них. Очевидно, управлять можно лишь помехоустойчивой системой. Если управляющие сигналы искажаются в недопустимых пределах при их передаче к объектам управления, то управление лишается смысла.

Управляемость обеспечивается прежде всего наличием прямой и обратной связей, которые служат для передачи управляемой системе команд (сигналов) управления, получения от нее сообщений о рассогласовании реального (фактического) и требуемого состояний управляемого объекта и исполнении команд управления. Управляемость объединяет такие свойства системы, как гибкость управления, его оперативность, точность, быстродействие, инерционность и др. Для сложных систем управляемость включает и способность выработки решений, на основе которых формируются управляющие воздействия.

В ряду усложнения качеств следующее место после управляемости занимает *способность системы* (A — качество). Речь идет о качестве системы, которое определяет ее возможности решать те или иные задачи, достигать тех или иных результатов в своей деятельности (производить в соответствующие сроки определенную продукцию, осуществлять определенный объем транспортных перевозок и т. д.). Данное качество объединяет совокупность свойств системы, определяющих ее функциональное назначение (производственное, транспортное, информационное и т. д.). К таким свойствам могут быть отнесены производительность, мощность, обеспеченность различного рода ресурсами и т. д. A — качество является определяющим при введении понятия эффективности системы (операции). Бессмысленно привлекать в качестве активного средства в операции систему, неспособную достичь цели операции. Операция, очевидно, будет неэффективной. Способность системы является необходимым (но не достаточным) условием эффективности операции, ее важнейшим фактором. Однако высокая эффективность операции обусловлена не только способностью системы, но и рациональным ее использованием.

Способность системы отражает потенциальную возможность решить поставленную задачу, т. е. ее потенциальную эффективность. Последняя, как уже отмечалось, определяется эффективностью операции при идеаль-

ных способах использования рассматриваемой системы.

Наиболее сложным качеством системы является самоорганизация (L — качество). Этим качеством обладают системы большой сложности, способные изменять свою структуру, параметры, ориентацию поведения в целях повышения эффективности выполнения своих функций. Самоорганизующая система [25] обнаруживает свойства, принципиально важными из которых являются: свобода выбора решений, способность к адаптации, самообучению, распознаванию ситуаций и т. п.

Принцип свободы выбора решений предусматривает выбор не единственного лучшего решения, а нескольких приемлемых решений. В зависимости от складывающейся ситуации при выборе решения на определенном шаге остается возможность вернуться к предыдущему шагу и изменить ранее принятое решение так, чтобы оставалась свобода выбора решения на последующих шагах процесса. Этим обеспечивается значительная гибкость управления и существенно повышается его эффективность.

С целью повышения эффективности операции самоорганизующиеся системы могут переподчинять свои подсистемы, перераспределять задачи между ними и ресурсы для решения этих задач.

На рис. 3 схематически показаны усложняющиеся качества систем. Перечисленные в порядке усложнения R , I , C , A , L — качества, условно изображены в виде системы вложенных прямоугольников. Данное графическое представление рассматриваемых качеств систем логически означает следующее: система, обладающая данным качеством (например, C — качеством), имеет и все другие более простые качества (I и R — качества), но не имеет качеств более высокого порядка (A и L — качеств).

Введение уровней качества технических средств позволяет ограничивать исследования качества одним из перечисленных уровней. Так, качества простых систем часто сводятся к устойчивости этих систем. Например, качество инженерных сооружений, таких,

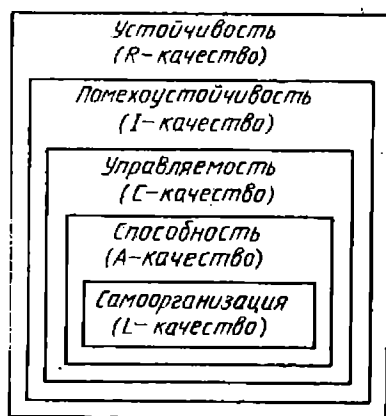


Рис. 3. Усложняющиеся качества систем

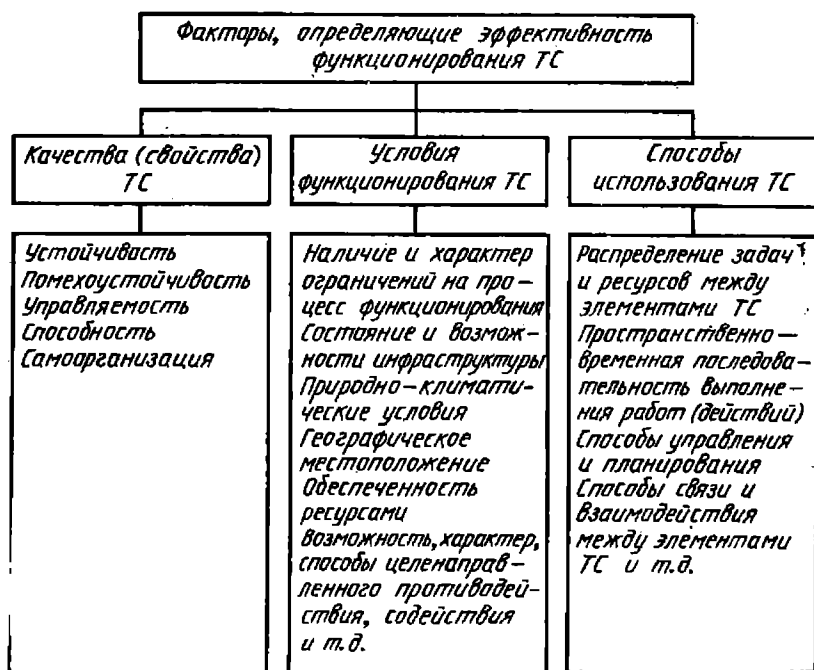


Рис. 4. Схема факторов, определяющих эффективность технических систем

как железнодорожный мост, определяется его устойчивостью. Здесь нет смысла подниматься на более высокие уровни при оценке качества. Очевидно, при анализе качества телевизионного приемника уровень устойчивости, определяемый в этом случае надежностью его работы, недостаточен. Целесообразно при оценке качества приемника перейти на уровень *I* — качества (помехоустойчивость), так как на этом уровне можно определить качество изображения. Однако при оценивании качества, например, стартового космического комплекса необходимо перейти на уровень *A* — качества и оценить потенциальную эффективность системы как ее способность вывести на орбиту искусственный спутник Земли (ИСЗ) с заданными параметрами. В этом случае приходится исследовать операцию, целью которой является вывод на орбиту ИСЗ. Потенциальную эффективность этой операции можно принять как характеристику качества стартового космического комплекса.

Уровень качества выбирает исследователь и в зависимости от сложности объекта определяет цели исследования, наличие информации о свойствах объекта, его целевом назначении и условиях применения.

Следующая важная группа факторов, оказывающих существенное

влияние на эффективность операции, характеризует способы применения ТС в операции. Эти факторы определяют (задают) порядок и приемы использования ТС для решения поставленных задач в операции. Каждой ТС присущи свои, специфические способы применения, которые характеризуются совокупностью соответствующих факторов. К ним относятся: распределение частных задач и выделенных ресурсов между элементами системы, пространственно-временная последовательность выполнения отдельных видов работ (действий), способы управления и планирования, способы связи и взаимодействия между элементами системы, режимы и регулярность использования и т. д.

К факторам, формирующим условия функционирования систем, относятся природные факторы (природно-климатические условия, географическое местоположение и т. д.), факторы, являющиеся следствием активных действий конкурентов или партнеров (например, возможность, характер, способы целенаправленного противодействия), а также факторы, характеризующие наличие и виды различного рода ограничений (экономических, социальных, экологических и др.).

Схема факторов, определяющих эффективность технических систем, приведена на рис. 4.

Среди факторов, которые учитываются при исследовании эффективности, большая доля факторов не контролируется ЛПР, например, объективные законы природы и общественного развития, которые необходимо учитывать в любом случае, а при их познании — использовать для достижения поставленных целей и учитывать их возможное негативное влияние. Неконтролируемыми являются и другие факторы; например, погодные условия, действия конкурентов относят к условиям функционирования (использования) ТС. Все эти факторы принято называть *неуправляемыми*.

К управляемым факторам относят такие, на которые ЛПР может влиять по своему усмотрению, т. е. которыми он может оперировать в процессе планирования и проведения операции. К ним относят, например, факторы, характеризующие способы применения ТС.

Стратегии U формируются из множества управляемых факторов. Однако в моделях операций в зависимости от характера задачи исследования эффективности множество стратегий U может быть сформировано из факторов, характеризующих не только способы применения, но и условия применения и качества ТС. Так, например, если необходимо выбрать лучший проект системы для определенных условий и способов ее применения, то тогда в качестве стратегии $u \in U$ выступают совокупности характеристик того или иного проекта. Или другой пример. В исследовательских задачах часто ставится задача выбора (создания) такого комплекса условий проведения операции (функционирования ТС) u^* , при котором эффективность операции (ТС) наибольшая. Очевидно, что в этом случае множество стратегий U формируется из множества факторов, характеризующих условия проведения операции (функционирования ТС).

По отношению к исследуемой системе факторы могут быть внешними и внутренними. Внешние факторы отражают влияние внешней среды, способствуя успешному проведению операции (полезные факторы), или противодействуя успеху операции (вред-

ные факторы). Внутренние факторы отражают взаимовлияние движущих сил внутри системы на ход и исход операции.

При исследовании систем факторы отображают в виде переменных (числовых и нечисловых). Классификация факторов приведена на рис. 5. С точки зрения информированности исследователя об этих переменных факторы делят на определенные Δ_F и неопределенные Δ_R . К определенным относят переменные, значения которых известны исследователю с требуемой точностью. Это различного рода заданные параметры, известные (регулярные) функции определенных аргументов и т. п. К определенным факторам также относят контролируемые входные воздействия, в том числе и управляемые переменные.

К неопределенным относят переменные, о значениях которых в реальном процессе исследования исследователь осведомлен не полностью. Природа неопределенности этих переменных (факторов) может быть различной. Обычно неопределенные переменные делят на две группы: случайные переменные Δ_F и неопределенные переменные нестохастической природы Δ_L .

Если распределение случайной переменной (например, в виде функции распределения) известно, то в этом случае говорят, что переменная статистически определена. Случайные переменные с неизвестными распределениями делят на два вида: с известными параметрами (характеристиками) распределения и с неизвестными параметрами. При исследовании систем со случайными факторами широко используются вероятностно-статистические методы. Например, методами параметрического статистического оценивания можно определить параметры распределения случайных переменных на основе статистических испытаний (если таковые возможны). Непараметрическое оценивание позволяет установить распределения случайных переменных.

Неопределенные факторы нестохастической природы Δ_L можно условно разделить на две группы: с известными функциями принадлежности (диапа-

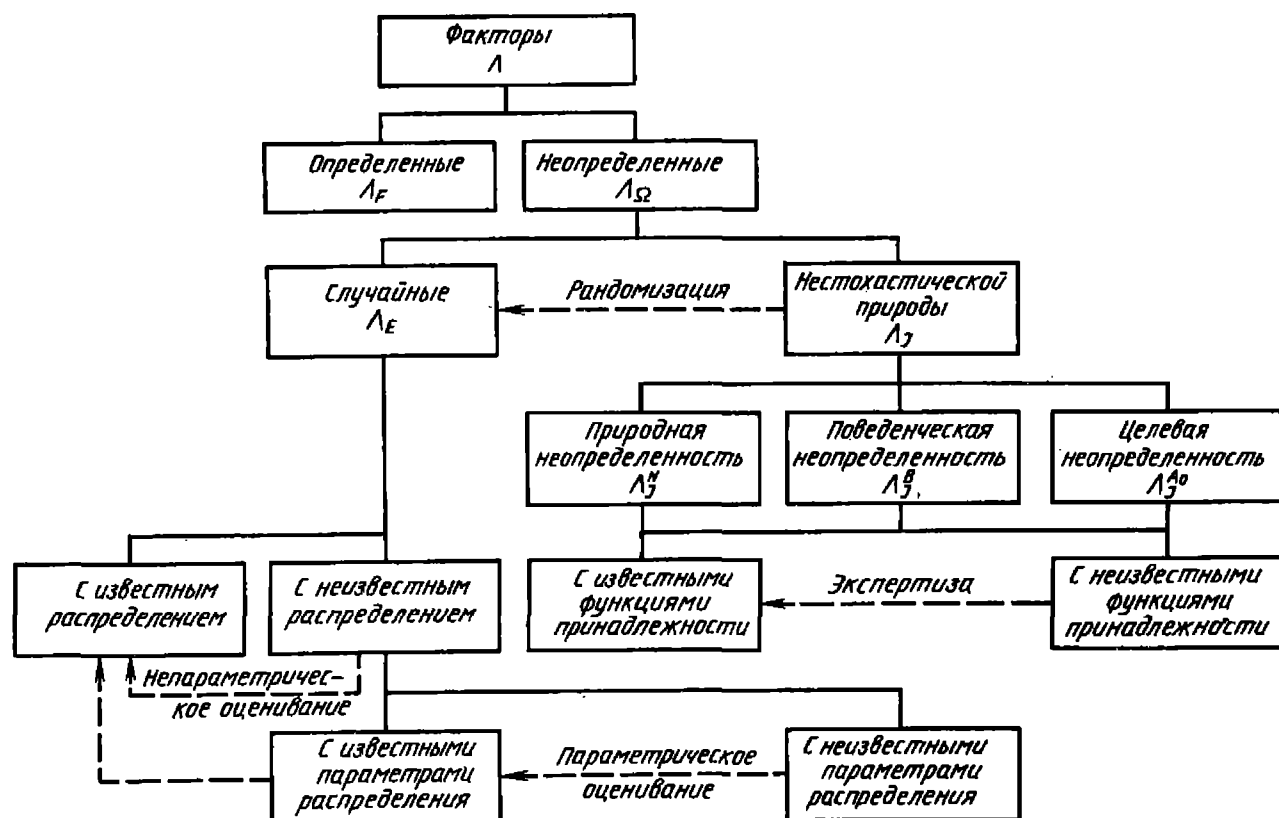


Рис. 5. Классификация факторов

зонами изменения переменных), с неизвестными функциями принадлежности.

Функция принадлежности задает некоторое подмножество (подобласть) общей допустимой области изменения фактора, определяемой, например, физической природой соответствующего фактора. Очевидно, подобласть, определяемая функцией принадлежности, в некотором смысле отражает степень неопределенности фактора. Разумеется, чем меньше подобласть, определяемая функцией принадлежности, тем меньше степень неопределенности фактора. В пределе функция принадлежности, выделяющая всего одно значение фактора, переводит его в разряд определенных факторов.

Наибольшей степенью неопределенности обладают факторы с неизвестными функциями принадлежности. Обычно к ним применяют процедуру экспертного оценивания диапазонов изменений их значений.

Часто для описания факторов нестохастической природы используется аппарат теории нечетких множеств в смысле Л. А. Заде [24]. Нечеткое

подмножество Ω некоторого универсального множества D ($\Omega \subset D$) характеризуется функцией принадлежности $\mu_{\Omega}(d)$, значение которой есть степень принадлежности элемента d подмножеству Ω . При этом $\mu_{\Omega}(d)$ может принимать любое значение в числовом интервале $[0; 1]$, т. е. $0 \leq \mu_{\Omega}(d) \leq 1$.

Подмножество в обычном («четком») смысле характеризуется функцией принадлежности, принимающей лишь два значения:

$$\mu_{\Omega}(d) = \begin{cases} 1, & \text{если } d \in \Omega; \\ 0, & \text{если } d \notin \Omega. \end{cases}$$

Таким образом, понятие функции принадлежности нечеткого множества есть довольно широкое обобщение функции принадлежности обычного множества.

В нашем случае универсальное множество D есть общая область значений фактора. $\mu_{\Omega}(D)$ выделяет нечеткое подмножество его реальных значений в исследуемых явлениях. Обычно функцию принадлежности вводят в виде так называемой лингвистической переменной, т. е. переменной, значения

которой определяются предложением в естественном языке. Например, категория «сложность» применительно к системе является лингвистической переменной, а ее значения выражаются словами: не сложная, не очень сложная, сложная, довольно сложная, очень сложная. В некотором классе систем указанная лингвистическая переменная задает нечеткое множество, например, сложных систем. На рис. 6 показана функция принадлежности $\mu_{\Omega}(x)$ для значения лингвистической переменной «сложная система», задающая нечеткое подмножество сложных систем на множестве систем с различным числом связей x между элементами.

В практике для построения функции принадлежности нередко прибегают к экспертному опросу с целью выявления доли экспертов, которые при фиксированном x отнесли систему к классу сложных. Эта доля экспертов и принимается в качестве значения функции принадлежности $\mu_{\Omega}(x)$. Например, если четыре эксперта из десяти систему, имеющую 20 связей, отнесли к классу сложных, то $\mu_{\Omega}(20) = 0,4$.

Нередко для описания неопределенных факторов нестохастической природы используют субъективные вероятности. В этом случае при анализе систем применяется теория вероятностей. Однако при введении субъективных вероятностей закон больших чисел может перестать действовать. Субъективные вероятности вводят обычно с помощью экспертного оценивания.

Некоторые неопределенности нестохастической природы иногда удается перевести в разряд случайных факторов с помощью рандомизации (см. рис. 5). Под *рандомизацией* понимают искусственное введение случайности в ситуацию, где она отсутствует. Например, при анализе эффективности технических систем поиска объектов исследователю может быть известен лишь район расположения некоторого объекта. Положение объекта неслучайно. Однако исследователь может предположить, что в пределах известного района положение объекта

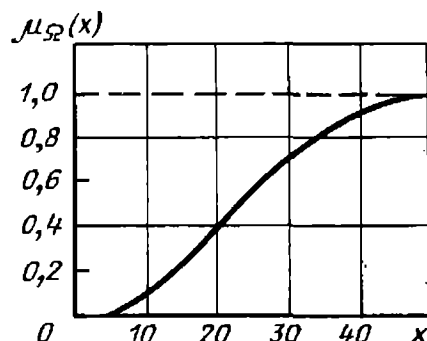


Рис. 6. Функция принадлежности нечеткого множества сложных систем

распределено с постоянной плотностью вероятности. Этим он рандомизировал положение объекта, т. е. ввел искусственно вероятностное распределение. Далее ситуацию с рандомизацией можно исследовать методами теории вероятностей и математической статистики.

Неопределенность нестохастического характера возникает обычно в силу следующих обстоятельств:

1) наличия целенаправленного противодействия со стороны конкурирующей системы, способы действий которой неизвестны исследователю; эту неопределенность поведения конкурента называют поведенческой неопределенностью Λ_J^B ;

2) недостаточной изученности некоторых явлений, сопровождающих процесс функционирования системы; неопределенность этого типа называют *природной* Λ_J^N ;

3) нечеткого представления цели операции, приводящего к неоднозначной трактовке соответствия реального результата операции желаемому; такую неопределенность называют *целевой* Λ_{Jo}^A .

Исследование эффективности технических систем с учетом неопределенных факторов нестохастической природы в значительной мере осложняется отсутствием достаточно общей теории (подобно теории вероятностей для исследования случайных явлений), формирующей методологические основания изучения явлений с неопределенными факторами. Тем не менее использование теории нечетких мно-

жеств, теории игр и теории решений позволяет найти некоторые пути решения задач исследования эффективности систем при наличии существенной неопределенности нестохастического характера.

В отдельную группу (по степени общности) выделяют результирующие факторы, т. е. факторы, непосредственно формирующие результат операции. К результирующим факторам, как уже отмечалось, относят полезный эффект q , достигнутый в операции, затраченные ресурсы (C) и сроки (T) проведения операции. Результат операции представляют в виде вектора, компонентами которого являются результирующие факторы, т. е.

$$Y = \| q, C, T \|,$$

или описывают функцией от результирующих факторов. Достаточно общим, например, является представление результата операции в виде степенной функции от результирующих факторов:

$$Y = \alpha_0 q^{\alpha_1} C^{\alpha_2} T^{\alpha_3}, \quad (1.1)$$

где $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ — параметры функции результата.

При этом результат будет нулевым, если хотя бы один из результирующих факторов положить равным нулю. Для описания результата в практике часто используют различные частные случаи приведенной функции. Так, если положить $\alpha_0 = \alpha_1 = 1$ и $\alpha_2 = \alpha_3 = 0$, то $y = q$. Если $\alpha_0 = \alpha_2 = 1$ и $\alpha_1 = \alpha_3 = 0$, то $y = C$. При $\alpha_0 = \alpha_3 = 1$ и $\alpha_1 = \alpha_2 = 0$ $y = T$. В этих частных случаях результат операции описывается лишь результирующим фактором (при этом на остальные результирующие факторы накладывают обычно ограничения в виде неравенств).

Если положить $\alpha_0 = \alpha_1 = 1$, $\alpha_2 = -1$, $\alpha_3 = 0$, то $y = q/C$. Это выражение иногда используют при анализе систем по методу «эффект — стоимость».

Используются и другие формы описания зависимости результата операции от результирующих факторов. Удобно, например, результат представлять в виде полинома от этих факторов.

При исследовании стохастических систем, т. е. систем, на поведение которых существенное влияние оказывают случайные факторы, результат операции Y будет случайной переменной.

3. ШКАЛЫ И ДОПУСТИМЫЕ ПРЕОБРАЗОВАНИЯ ПОКАЗАТЕЛЕЙ

Любая задача исследования эффективности предполагает измерение всех учитываемых факторов. Без этой формальной процедуры нельзя создать модель, адекватно отображающую реальную ситуацию, а следовательно, и получить сколь-нибудь значимые для практики результаты. И тем не менее, при разработке моделей исследования эффективности операций (технических систем) этим вопросам уделяется незаслуженно малое внимание. Нередко считается, что введенные в модель факторы уже измерены неизвестным и нерассматриваемым способом. В результате даже самые совершенные модели неспособны нормально, а зачастую и вообще функционировать из-за отсутствия взаимоувязанных процедур объективных и субъективных измерений показателей, характеризующих разнообразие по своей природе факторы, подлежащие учету при исследовании эффективности реальных операций.

Методы теории эффективности, и это закономерно, ориентированы на широкое использование всего арсенала возможностей современного математического аппарата и вычислительной техники. Реализация этих возможностей немыслима без числового представления всех видов информации о факторах, существенно влияющих на результаты исследований. Поэтому одна из главных проблем измерения в теории эффективности — унификация измерений всевозможных объектов с использованием числовой системы, позволяющей создать общую формальную схему как объективных, так и субъективных измерений. В решении этой важной задачи ведущая роль принадлежит теории измерений [56, 62].

При исследовании эффективности потребность в измерении обусловлена необходимостью выразить в числовой мере различные отношения между объектами с тем, чтобы при определении отношений между ними оперировать не с самими объектами, а с соответствующими этим отношениям числовыми мерами. Например, для установления отношения массы между телами нет необходимости сопоставлять сами тела, а достаточно сопоставить найденные путем измерений значения числовой меры для этих тел, поскольку масса, как физическая характеристика тела, выражена в определенной числовой мере.

В процессе исследования эффективности ЛПР, эксперты, исследователи формируют различные объекты (ситуации, цели, ограничения, стратегии, предпочтения, события, предметы, явления и т. д.) и измеряют их. Эти измерения могут быть субъективными или объективными, носить качественный или количественный характер. Субъективные измерения осуществляет человек, поэтому на их результат влияет психология мышления индивида, его знания, опыт и многие другие свойства личности. Объективные качественные или количественные измерения осуществляются измерительными приборами, действие которых основано на использовании физических законов.

Под *измерением* понимается процедура сравнения объектов по определенным показателям (признакам). В качестве показателей сравнения объектов используют пространственные, временные, физические, функциональные, экономические, социологические и другие свойства и характеристики объектов (например, показатели надежности технической системы, показатели эффективности операции и др.). Процедура сравнения включает определение отношений между объектами и способ их сравнения. Если введены конкретные показатели сравнения, то можно установить отношения между объектами (например, больше, меньше, равны, предпочтительнее и т. д.). Способы сравнения объектов между собой могут быть различными, например, последова-

тельно с одним объектом, принимаемым за эталон, или друг с другом в упорядоченной или произвольной последовательности.

Отношение является математической категорией, поэтому понятию измерения можно придать строгую форму. Для формального описания множества объектов и отношений между ними вводится понятие эмпирической системы

$$S^e = \langle D, R^e \rangle,$$

где $D = (d_1, d_2, \dots, d_n)$ — множество объектов, в качестве которых могут рассматриваться, например, факторы, стратегии, системы и т. п.; $R^e = (R_1^e, R_2^e, \dots, R_m^e)$ — множество отношений между объектами.

Отношение является самой общей формой описания связей между объектами, частным случаем которого является функция. Запись вида $d_i R_k^e d_j$ или $(d_i, d_j) \in R_k^e$ означает, что объекты d_i и d_j находятся между собой в отношении R_k^e . Такое отношение называется *бинарным* (двухместным), поскольку оно связывает между собой два объекта. Если отношение имеет место одновременно между тремя объектами, то оно называется *тернарным* (трехместным) а если между n объектами, то *n -нарным* (n -местным).

При измерении широкое применение находят некоторые типы бинарных отношений (эквивалентность, строгий порядок, квазипорядок, нестрогий порядок и др.), которые выделяются на основе рассмотрения таких свойств бинарных отношений, как рефлексивность, антирефлексивность, симметричность, антисимметричность, транзитивность (подробнее об отношениях и их свойствах см. гл. 12).

Эмпирической системе $S^e = \langle D, R^e \rangle$ ставится в соответствие числовая система

$$S^H = \langle R_{e_1}, R^H \rangle,$$

где R_{e_1} — множество действительных чисел; $R^H = (R_1^H, R_2^H, \dots, R_e^H)$ — множество отношений между числами.

Числовая система называется *полной*, если R_{e_1} есть множество всех действительных чисел. Отношениям строгого и нестрогого порядка между объектами соответствуют отношения строгого и нестрогого неравенства между числами, а отношению эквивалентности между объектами — отношение равенства между числами.

Числовая система $S^H = \langle R_{e_1}, R^H \rangle$ используется для унификации измерений. Измерение заключается в отображении объектов эмпирической системы (причем объекты, показатели сравнения и виды отношений могут быть самыми разнообразными) на множество чисел таким образом, чтобы отношения между числами, отображающими объекты, сохраняли отношения между объектами.

Числовая система сохраняет свойства и отношения объектов, если она изоморфна или по крайней мере гомоморфна эмпирической системе. Числовая система *изоморфна* эмпирической системе, если они подобны и между ними существует взаимнооднозначное отображение (функция) f объектов на множество чисел, такое, что отношение R^e между объектами имеет место тогда и только тогда, когда имеет место отношение между числами, отображающими объекты на числовой оси. Подобие двух систем с отношениями означает, что количество отношений и их местность в обеих системах одинаковы.

В некоторых случаях условие взаимной однозначности отображения f является слишком жестким и не всегда необходимым. Заменяя это условие в данном выше определении изоморфизма на условие однозначного соответствия, приходим к понятию *гоморфизма*.

Таким образом, измерение представляет собой отображение объектов эмпирической системы на множество чисел в числовой системе. При этом каждому объекту эмпирической системы с помощью отображения (функции) f приписывается число

$$c_i = f(d_i).$$

Отношения между числами при таком отображении должны сохранять от-

ношения между объектами. Например, если объект d_i предпочитается или равноценен объекту d_j ($d_i \succeq d_j$), то $c_i = f(d_i) \geq c_j = f(d_j)$.

Основными проблемами теории измерений являются проблемы представления и единственности. Первая из этих проблем состоит в доказательстве возможности представления эмпирической системы с помощью числовой системы, сохраняющей отношения между объектами, т. е. изоморфной или гомоморфной. В теории измерений доказано существование числовых систем для описания множества объектов, связанных отношениями эквивалентности, строгого и нестрогого порядка [62].

Проблема единственности состоит в определении всех возможных способов представления данной эмпирической системы различными числовыми системами и установления связи между ними. Эту проблему можно сформулировать как проблему определения типа шкалы.

Шкалой называется совокупность эмпирической системы S^e , числовой системы S^H и отображения f , т. е. шкала задается тройкой

$$\langle S^e, S^H, f \rangle.$$

Пусть W — некоторый показатель для выражения отношения между объектами D (например, W — показатель эффективности технической системы; D — некоторое множество технических систем); $\langle S^e, S^H, f_1 \rangle$ и $\langle S^e, S^H, f_2 \rangle$ — две шкалы, различающиеся отображениями f_1 и f_2 . Тогда $W_1 = f_1(d_i)$ и $W_2 = f_2(d_i)$ — числовые значения показателя W , полученные для объекта d_i эмпирической системы соответственно с помощью отображений f_1 и f_2 .

Возникает вопрос о взаимосвязи числовых значений W_1 и W_2 , полученных при отображении одного и того же объекта d_i , но в разных шкалах. Обозначим через $\varphi(W)$ функцию, устанавливающую связь между значениями W_1 и W_2 показателя W , т. е. $W_1 = \varphi(W_2)$. Функцию $\varphi(W)$ называют допустимым преобразованием

шкалы показателя W , если эта функция служит для выражения того же отношения между объектами, что и показатель W .

С помощью функции допустимого преобразования $\varphi(W)$ можно описать связь между любыми числовыми системами, выбираемыми для одной и той же эмпирической системы по показателю W . Поскольку свойства этой функции характеризуют взаимосвязь между числовыми системами, то ее можно использовать для описания понятия единственности отображения.

Множество Φ всех допустимых преобразований шкалы показателя W определяет тип этой шкалы. В этом случае говорят, что показатель W имеет шкалу типа Φ , или что показатель W измеряется в шкале типа Φ . Шкала считается тем совершенней, чем уже множество допустимых ее преобразований.

Каждый тип шкалы имеет свою информативность и свой класс допустимых преобразований, за пределы которого нельзя выходить без риска получить ошибочные или бессмысленные результаты. Особое значение это замечание имеет при исследованиях эффективности операций, которая зависит от большого числа различных по характеру и природе факторов, сложным образом связанных и взаимодействующих между собой.

При измерении показателей наибольшее распространение получили метрические, порядковые и номинальные шкалы. Среди метрических шкал выделяют абсолютные шкалы, шкалы отношений и интервальные шкалы. Рассмотрим перечисленные типы шкал в порядке возрастания их совершенства.

Номинальная (классификационная) шкала, или *шкала наименований*, — это наименее совершенная, простейшая, по существу, качественная шкала. Ее применяют для описания принадлежности объектов к определенным классам; она сохраняет отношение эквивалентности и различия между объектами. В этой шкале число используют лишь для обозначения и выделения объекта: всем объектам одного и того же класса присваивают одно и то же число, а объектам разных

классов — разные числа. Однако предпочтение между объектами не устанавливается.

Поскольку существует большое число вариантов присвоения чисел классам эквивалентных объектов, то понятие единственности отображения f состоит для шкалы наименований во взаимнооднозначности допустимого преобразования φ . Поэтому множеством допустимых преобразований показателя, имеющего номинальную шкалу, являются все виды функциональных преобразований, обладающие свойством однозначности:

$$\Phi_n = \{\varphi(W) \mid W_1 \neq W_2 \rightarrow \varphi(W_1) \neq \varphi(W_2)\}.$$

Таким образом, номинальная шкала единственна с точностью до взаимнооднозначности соответствия. Это означает, что в данной шкале отсутствуют понятия масштаба и начала отсчета.

Класс показателей, которые имеет смысл измерять в номинальной шкале, весьма узок. К таким показателям относятся, например, различные решающие правила распознавания образов, используемые для того, чтобы определить, соответствует ли данный образец эталону или нет. В этой шкале также измеримы любые пороговые функции, назначение которых — установить в бинарном коде «да — нет» принадлежность объекта заданному классу (типу, разновидности). Обычно шкалу наименований используют для индексации видов продукции, типов машин, номенклатуры изделий (спецификация изделий), нумерации подразделений в организации и т. п.

Порядковая (ранговая) шкала более совершенна, чем номинальная. Ее применяют для измерения упорядоченности объектов по одному или совокупности признаков. Числа в ней определяют только порядок следования объектов, но не позволяют утверждать, во сколько или на сколько один объект предпочтительнее другого. Для порядковой шкалы допустимое преобразование φ является любым монотонным преобразованием; поэтому эта шкала единственна до монотонного преобразования. Множество допустимых преобразований показателя, имею

щего порядковую шкалу, состоит из всех монотонно возрастающих функций:

$$\Phi_{\Pi} = \{\varphi(W) \mid W_1 > W_2 \rightarrow \varphi(W_1) > \varphi(W_2)\}.$$

Показатели, измеримые в порядковой (ранговой) шкале, значительно информативнее показателей, измеримых в номинальной шкале, так как позволяют судить об отношениях типа «лучше — хуже», «больше — меньше», существующих между объектами. Однако в этой шкале также отсутствуют понятия масштаба и начала отсчета, поэтому значения показателей W_1 и W_2 , измеренные в порядковой шкале, не позволяют ответить на вопросы типа: «на сколько (во сколько раз) один объект лучше, важнее другого?».

Наибольшее распространение порядковые шкалы получили при методах обработки экспертной информации о сравнительных оценках качественных свойств объектов. Оценки такого рода даются в баллах, а сами шкалы часто называют *балльными*. Балльным оценкам всегда присущ субъективизм, но его доля может быть различной. С этой точки зрения различают два вида балльных оценок: при наличии общепринятого эталона и при отсутствии такового (см. гл. 12).

Шкалу интервалов применяют для отображения различия между свойствами объектов. Эта шкала может иметь произвольные точки отсчета и масштаб. Множество допустимых преобразований показателя, имеющего шкалу интервалов, состоит только из всех линейных функций вида $\varphi(W) = aW + b$, где $a > 0$, b — произвольное число, т. е.

$$\Phi_{\Pi} = \{\varphi(W) = aW + b\}.$$

Следовательно, шкала интервалов единственна с точностью до линейного преобразования.

Конкретное измерение показателя в шкале интервалов осуществляется при фиксированных величинах a (масштаб, задающий единицу измерения) и b (начало отсчета). Основным свойством шкалы интервалов, определяющим ее название, является сохранение отношения интервалов при допу-

стимом преобразовании шкалы. В этой шкале отношение разности чисел в двух числовых системах определяется масштабом измерения.

В шкале интервалов измеряются, например, сроки выполнения различных работ, гарантийные сроки службы узлов, агрегатов, машин и т. п., дата выпуска изделия, начало операции и все другие показатели, для измерения которых необходимо фиксировать масштаб и начало отсчета.

Шкала отношений является частным случаем шкалы интервалов при выборе нулевой точки отсчета ($b = 0$). Однако эта шкала более совершенна, чем шкала интервалов, так как для нее соответствующее множество Φ_0 всех допустимых преобразований состоит только из преобразований подобия: $\Phi_0 = \{\varphi(W) = aW\}$.

В шкале отношений числа отражают отношения свойств объектов, т. е. во сколько раз свойство одного объекта превосходит это же свойство другого объекта. Измерения по этой шкале допускают сравнение не только самих значений показателей или их разностей, но и различных арифметических комбинаций этих значений, если, конечно, они имеют физический смысл.

Показатели, измеряемые в шкале отношений, наиболее распространены в технике, математике и, конечно, в теории эффективности. В качестве примеров таких показателей можно указать длину, массу, напряжение, стоимость и другие, для которых существует естественное начало отсчета (нулевая точка).

Абсолютная шкала — самая совершенная. В этой шкале принимается нулевая точка отсчета ($b = 0$) и единственный масштаб ($a = 1$). Для абсолютной шкалы соответствующее множество Φ_a допустимых преобразований состоит всего из одного элемента, представляющего собой тождественное преобразование, т. е.

$$\Phi_a = \{\varphi(W) \equiv W\}.$$

Это означает, что существует одно и только одно отображение объектов в числовую систему. Отсюда следует и название шкалы, так как для нее единственность отображения пони-

мается в буквальном, абсолютном смысле.

В абсолютной шкале определяется, например, количество объектов (предметов, событий и т. п.), которое может быть измерено единственным образом с помощью ряда натуральных чисел. Примером абсолютных шкал являются шкала температур по Кельвину, шкала значений вероятности события и др.

Таким образом, показатели могут иметь шкалы различных типов. Это дает возможность дать более точное определение количественного и качественного показателя.

Количественными показателями называют такие показатели, которые имеют шкалу не менее совершенную, чем интервальная. *Качественными показателями* называются такие показатели, которые имеют шкалу менее совершенную, чем интервальная.

Тип шкалы необходимо учитывать всегда, когда речь идет о том, какие действия имеет смысл осуществлять с числовыми значениями рассматриваемого показателя. В теории измерений эту проблему называют *проблемой адекватности* (осмысленности) числовых утверждений. Суть ее состоит в том, что числовое утверждение полагают адекватным тогда и только тогда, когда его значение истинности инвариантно относительно допустимых преобразований шкалы.

Пусть, например, W — некоторый показатель, а W_1, W_2, W_3 — его числовые значения, полученные в шкале интервалов, т. е. $\varphi(W) = aW + b$. Рассмотрим числовое утверждение $W_1 + W_2 > W_3$ и определим его адекватность (осмысленность). В этом случае

$$(aW_1 + b) + (aW_2 + b) > (aW_3 + b).$$

Очевидно, что это утверждение не эквивалентно исходному, и поэтому при измерении W в шкале интервалов утверждение $W_1 + W_2 > W_3$ не имеет смысла.

В то же время при измерении показателя W в шкале отношений, т. е. $\varphi(W) = aW$, утверждение $W_1 + W_2 >$

W_3 , как можно видеть, является осмысленным.

Выбор той или иной шкалы для измерения определяется характером отношений между объектами эмпирической системы, наличием информации и целями исследования. Применение количественных шкал требует значительно более полной информации об объектах по сравнению с применением качественных шкал, а получение такой информации связано с затратами времени и ресурсов. Поэтому прежде всего необходимо обращать внимание на правильное согласование выбираемой шкалы с целями исследования. Если, например, эта цель состоит в упорядочении объектов, то нет необходимости измерять количественные характеристики объектов, достаточно определить только качественные характеристики.

Законы функционирования исследуемых сложных систем могут быть изучены не столь тщательно и полно, чтобы по ним можно было легко определить класс допустимых преобразований и соответствующую ему шкалу того или иного показателя. В этих случаях целесообразней вначале использовать шкалу с максимально широким классом допустимых преобразований, при котором значение показателей еще содержит информацию, необходимую для начального этапа решения поставленной задачи. В дальнейшем, по мере получения информации, можно уточнять класс допустимых преобразований и переходить к более совершенным шкалам.

Необходимо различать *первичные* и *производные от первичных* измерения показателей. Последние называют еще *косвенными измерениями*. Они не зависят от эмпирической системы, а реализуются с помощью специальных вычислительных процедур, позволяющих получать числовое значение показателя на основе преобразования по определенным алгоритмам или формулам числовых значений величин, полученных прямыми измерениями. При исследовании эффективности операций (технических систем) этот прием используется очень широко.

Глава 2. Показатели и критерии эффективности

1. ПОКАЗАТЕЛЬ ЭФФЕКТИВНОСТИ

Показатель эффективности $W(u)$ операции есть мера степени соответствия реального результата операции требуемому (см. т. 1 справочника).

Показатель эффективности может измеряться как в метрической, так и порядковой шкалах.

Основным требованием при выборе показателя эффективности является соответствие показателя цели операции, которая отображается требуемым результатом Y^{TP} .

Для описания соответствия реального результата Y операции требуемому формально вводят числовую функцию на множестве результатов операции:

$$\rho = \rho(Y(u), Y^{\text{TP}}), \quad (2.1)$$

которую называют функцией соответствия.

Эта функция в некоторой шкале показывает степень достижения цели операции, а конкретный вид функции соответствия зависит от цели операции, задачи исследования и других факторов. В силу того, что $Y(u)$ может быть случайной переменной (числовой или нечисловой), функция соответствия также может быть случайной величиной (как числовая функция случайного аргумента). В некоторых задачах и Y^{TP} приходится вводить как случайную переменную.

Если результат выражается случайной переменной, то запись $Y(u)$ означает, что распределение Y зависит от стратегии $u \in U$. В этом случае функцию распределения $F_u(y)$ будем записывать с индексом u , так как вид ее зависит от u .

Введение в рассмотрение функции соответствия (2.1) позволяет принять математическое ожидание этой функции в качестве показателя эффективности $W(u)$, т. е.

$$W(u) = M[\rho(Y(u), Y^{\text{TP}})]. \quad (2.2)$$

Если $Y(u)$ и Y^{TP} — неслучайные переменные, то

$$W(u) = \rho(Y(u), Y^{\text{TP}}),$$

т. е. в детерминированном случае функция соответствия служит показателем эффективности операции.

Для того чтобы числовая функция (2.2), определенная на множестве стратегий U , могла рассматриваться в качестве показателя эффективности, помимо требования соответствия цели A_0 операции она должна удовлетворять следующим требованиям: содержательности и интерпретируемости, измеримости, соответствия системе предпочтений ЛПР.

Последнее из этих требований означает, что показатель эффективности должен учитывать психологические особенности лица, принимающего решение (ЛПР), отражающие его отношение к различным ситуациям в условиях неопределенности (например, склонность, несклонность или безразличие к риску).

Формально психологические особенности ЛПР можно учесть введением специальной оценочной функции $f^{\theta_c}(\rho)$, отражающей отношение ЛПР к риску.

Методы построения функции f^{θ_c} по информации θ_c рассматриваются в гл. 13.

С учетом этого показатель эффективности W есть математическое ожидание оценочной функции:

$$W(u) = M[f^{\theta c}(\rho(Y(u), Y^{TP}))]. \quad (2.3)$$

Показатели, построенные по правилу (2.2), часто называют «объективными», а по правилу (2.3) — «субъективными».

Если результат Y операции может быть описан единственной величиной y , то (2.2) и (2.3) определяют скалярные показатели эффективности. В противном случае приходится вводить векторный показатель эффективности

$$W(u) = \|W_1(u), W_2(u), \dots, W_m(u)\|^T, \quad (2.4)$$

где $W_i(u)$, $i = \overline{1, m}$ определяется по правилу (2.2) с подстановкой вместо $Y(u)$, Y^{TP} величин $y_i(u)$, y_i^{TP} частных характеристик исхода, т. е.

$$W_i(u) = M[\rho_i(y_i(u), y_i^{TP})], \\ i = \overline{1, m}.$$

Введение векторного показателя эффективности накладывает дополнительные требования: минимальности числа частных показателей и полноты. Подробнее существо требований, предъявляемых к показателю эффективности, рассматривается при описании специальной задачи моделирования цели операции.

Обычно векторный показатель вводят в случаях, когда единственная цель операции достигается решением нескольких задач, эффективность решения каждой из которых оценивается соответствующим частным показателем $W_i(u)$, $i = \overline{1, m}$, но свернуть эти показатели в один обобщенный показатель не удается. Эти частные задачи могут решаться отдельными подсистемами, входящими в общую систему S_0 , и тогда $W_i(u)$ есть показатель эффективности частной операции, проводимой i -й подсистемой. Кроме того, эти задачи могут решаться одной системой, но на разных этапах операции, и тогда $W_i(u)$ есть показатель эффективности решения задачи на i -м этапе операции.

Показатель эффективности $W(u)$ зависит от стратегии u . Он определяется на множестве допустимых стратегий U . В общем виде эта зависимость задается отображением

$$\Psi: U \rightarrow W,$$

т. е. отображением множества допустимых стратегий U во множество значений показателя эффективности W . Обычно отображение Ψ задается в форме определенной математической модели операции.

2. ФОРМЫ ПОКАЗАТЕЛЯ ЭФФЕКТИВНОСТИ

Показатель эффективности в форме (2.3) является наиболее общим. В зависимости от вида оценочной функции $f^{\theta c}(\cdot)$ и функции соответствия $\rho(Y(u), Y^{TP})$ из (2.3) можно получить различные показатели эффективности.

Покажем это на примере «объективных» скалярных показателей, часто используемых при исследованиях эффективности технических систем.

С целью отличия случайной величины от ее возможного значения, когда это не ясно из контекста, над соответствующей буквой будем ставить символ $\wedge$. Например, $\hat{\rho}$ — случайная величина, ρ — ее возможное значение.

Пусть цель операции описывается случайным событием A , наступление которого является желательным результатом операции. Комплекс условий, а следовательно, и вероятность $P_u(A)$ наступления этого события зависят от стратегии $u \in U$. Функцию соответствия ρ в этом случае вводят как бернуллиеву переменную, которая может принять лишь два значения: 0 или 1, т. е.

$$\rho(y(u), y^{TP}) = \begin{cases} 1, & \text{если событие } A \text{ наступило;} \\ 0 & \text{в противном случае.} \end{cases} \quad (2.5)$$

Очевидно, при таком введении функции соответствия $y^{TP} = 1$. Вероятность события A есть математическое ожидание бернуллиевой переменной или функции соответствия (2.5).

Действительно,

$$M[\rho(y(u), y^{TP})] = 1 \cdot P_u(A) + 0 \cdot P_u(\bar{A}) = P_u(A)$$

и, следовательно, показатель эффективности в форме (2.2) есть вероятность наступления события A :

$$W(u) = P_u(A).$$

Часто событие A выражается отношением между реальным результатом $y(u)$ и требуемым y^{TP} . Например, $A_1 = \{g(u) \geq y^{TP}\}$ или $A_2 = \{y_1^{TP} \leq g(u) < y_2^{TP}\}$. Функции соответствия для этих событий вводятся следующим образом:

для события A_1

$$\rho(y(u), y^{TP}) = \begin{cases} 1, & \text{если } g(u) \geq y^{TP}; \\ 0, & \text{если } g(u) < y^{TP}; \end{cases} \quad (2.6)$$

для события A_2

$$\rho(g(u), y^{TP}) = \begin{cases} 1, & \text{если } y_1^{TP} \leq g(u) < y_2^{TP}; \\ 0 & \text{в противном случае.} \end{cases}$$

Функцию соответствия (2.6) употребляют в случаях, когда требуемый результат задачи и его достижение являются неперенным условием выполнения поставленной задачи. При этом показатель эффективности

$$W(u) = P\{g(u) \geq y^{TP}\} \quad (2.7)$$

трактуется как вероятностная гарантия (или степень гарантии) выполнения поставленной задачи. Например, если цель операции заключается в обеспечении повышения срока службы

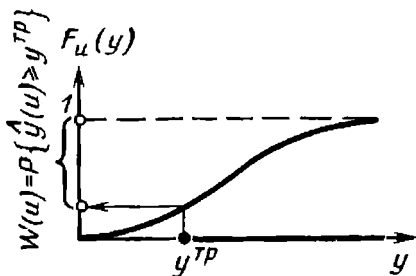


Рис. 1. Функция распределения результата операции. Вероятностная гарантия

изделия до уровня, не ниже требуемого y^{TP} , то показатель эффективности операции (2.7) есть степень гарантии или вероятность того, что срок службы изделия будет не менее требуемого.

При известной функции распределения реального результата $F_u(y)$ (2.7) записывают в следующем виде:

$$W(u) = 1 - F_u(y^{TP}).$$

На рис. 1 изображена функция распределения результата операции и показана вероятностная гарантия (вероятность $P\{g(u) \geq y^{TP}\}$).

Допустим, что имеет место некоторая неопределенность при установлении требуемого результата операций y^{TP} . Если эта неопределенность нестохастического характера, то можно ввести функцию принадлежности $\mu_A(y)$ для нечеткого случайного события $A = \{g(u) \geq y^{TP}\}$. В записи события A переменная $g(u)$ является случайной величиной с функцией распределения $F_u(y)$, но y^{TP} есть неопределенная переменная нестохастического характера с функцией принадлежности $\mu_A(y)$. Нечеткую (лингвистическую) переменную будем сверху снабжать символом «~».

Нечеткое случайное событие A введем следующим образом. Как известно из теории вероятностей, случайное событие A есть подмножество пространства элементарных событий E , т. е. $A \subset E$. Предположим теперь, что A есть нечеткое подмножество E (т. е.

$A \subset E$), заданное функцией принадлежности μ_A (для простоты будем пока рассматривать случай, когда E не более чем счетно). Теперь каждому элементарному событию $e_i \in E$ следует поставить в соответствие не только вероятность его наступления $P(e_i)$, но и степень принадлежности e_i подмножеству A , т. е. $\mu_A(e_i)$, ($0 \leq \mu_A(e_i) \leq 1$). Чтобы найти вероятность наступления нечеткого случайного события A , следует по всем $e_i \in E$ просуммировать произведения $\mu_A(e_i) P(e_i)$, т. е.

$$P(A) = \sum_i \mu_A(e_i) P(e_i).$$

Эта запись есть математическое ожидание дискретной случайной величины

μ_A . Таким образом, вероятность наступления нечеткого случайного события есть математическое ожидание функции принадлежности этого нечеткого события, т. е.

$$P(A) = M[\mu_A]. \quad (2.8)$$

Если целью операции является достижение результата $y(u)$ не ниже требуемого уровня при нечетком задании последнего, то функцию соответствия можно ввести по аналогии с (2.6):

$$\rho(g(u), \bar{y}^{TP}) = \begin{cases} 1, & \text{если } g(u) \geq \bar{y}^{TP}; \\ 0, & \text{если } g(u) < \bar{y}^{TP} \end{cases} \quad (2.9)$$

и тогда $M[\rho(g(u), \bar{y}^{TP})] = P(g(u) \geq \bar{y}^{TP}) = M[\mu_A(\bar{y}^{TP})]$

или в общем виде

$$P(g(u) \geq \bar{y}^{TP}) = \int_E \mu_A(y) dF(y).$$

Однако в качестве функции соответствия в этих условиях может быть использована и сама функция принадлежности, т. е.

$$\rho(g(u), y^{TP}) = \mu_A(y^{TP}).$$

Она в определенном смысле эквивалентна (2.6).

В практике оценки эффективности операций приходится иногда учитывать случайный характер требуемого результата. Например, в качестве результата операции по повышению надежности технической системы можно потребовать, чтобы срок ее безаварийной работы был не менее периода времени активных работ с использованием этой системы.

Здесь срок безаварийной работы и период времени активных работ могут иметь случайный характер, т. е. требуемый результат в этих условиях описывается случайной переменной g^{TP} , которую обычно называют минимально необходимым результатом. Если g^{TP} — случайная величина (числовая случайная переменная), то ее распределение вводят обычно с помощью функции распределения $F_H(y)$.

В случаях, когда целью операции является достижение результата $g(u)$

не ниже требуемого уровня g^{TP} при случайном характере последнего, функцию соответствия можно ввести по аналогии с (2.6):

$$\rho(g(u), g^{TP}) = \begin{cases} 1, & \text{если } g(u) \geq g^{TP}; \\ 0, & \text{если } g(u) < g^{TP}. \end{cases} \quad (2.10)$$

Показатель эффективности

$$W(u) = M[\rho(g(u), g^{TP})] = P(g(u) \geq g^{TP})$$

в рассматриваемых условиях представляет собой математическое ожидание функции соответствия (2.10) и является вероятностной гарантией того, что реальный результат будет не менее требуемого при случайном характере как реального, так и требуемого результатов.

Функция распределения $F_H(y)$ необходимого результата представляет собой условную вероятность события $\{g^{TP} < y\}$, вычисленную в предположении, что реальный результат принял значение y (т. е. наступило событие $\{y \leq g < y + dy\}$).

Таким образом,

$$F_H(y) = P\{g_H < y / y \leq g < y + dy\},$$

но, очевидно,

$$P\{y \leq g < y + dy\} = dF(y),$$

где $F(y)$ — функция распределения реального результата. Используя формулу полной вероятности, получим

$$W(u) = P(g(u) \geq g_H) = \int_E F_H(y) dF(y). \quad (2.11)$$

Здесь интегрирование проводится по всему пространству элементарных событий E , на котором определены случайные величины g и g_H .

При исследовании эффективности операций широко распространен показатель среднего результата, т. е.

$$W(u) = M[y(u)]. \quad (2.12)$$

Этот показатель используется в тех случаях, когда цель операции выражается числовой переменной. Очевидно, что (2.12) является частным случаем показателя (2.2), при котором функ-

ция соответствия равна реальному результату

$$\rho(y(u), y^{TP}) = y(u).$$

Вводя показатель среднего результата и зная, например, диапазон изменения результата, исследователь может сравнить его значение с предельно большим значением.

Важным свойством показателя среднего результата является его аддитивность, т. е.

$$M\left[\sum_i g_i\right] = \sum_i M[g_i]. \quad (2.13)$$

Если результат операции $g(u)$ представим в виде суммы результатов действий подсистем ($g = \sum_i g_i$), то средний результат операции равен сумме средних частных результатов, несмотря на возможную их стохастическую зависимость.

Если цель операции носит количественный характер, то в качестве показателя эффективности операции наряду с (2.7) может быть принят минимальный результат y_α , получаемый с заданной вероятностью α , т. е.

$$\alpha = P(g(u) \geq y_\alpha). \quad (2.14)$$

Очевидно,

$$\alpha = 1 - F(y_\alpha),$$

где $F(y)$ — функция распределения реального результата операции (случайной величины $g(u)$).

Решив это уравнение относительно y_α , получим

$$y_\alpha = F^{-1}(1 - \alpha). \quad (2.15)$$

Здесь y_α есть обратная функция к функции распределения $F(y)$ при

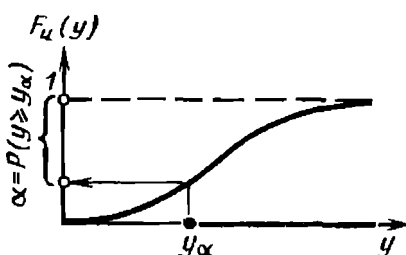


Рис. 2. Функция распределения результата операции. Гарантированный результат

значении аргумента $(1 - \alpha)$ (квантиль распределения $F(y)$). На рис. 2 изображена функция распределения результата операции и показан гарантированный результат (переход от $(1 - \alpha)$ к y_α).

При этом функция соответствия

$$\rho = F^{-1}(1 - \alpha) \quad (2.16)$$

есть величина неслучайная и ее математическое ожидание, следовательно, равно ρ .

Таким образом, показатель эффективности в форме (2.2) имеет вид

$$W(u) = M[\rho] = y_\alpha. \quad (2.17)$$

Этот показатель обычно называют *вероятностно-гарантированным результатом*. Требуемый результат косвенно отражает заданный (требуемый) уровень вероятности α (степень гарантии).

Иногда в качестве показателя эффективности в условиях, когда результат описывается случайной величиной, приходится использовать характеристику рассеяния (кучности) реального результата относительно требуемого значения или относительно своего математического ожидания. Функции соответствия в этом случае имеют вид

$$\rho(y_1, y^{TP}) = (g(u) - y^{TP})^2;$$

$$\rho(y_1, y^{TP}) = (g(u) - M[g(u)])^2.$$

В первом случае показатель эффективности есть математическое ожидание квадрата отклонения реального результата от требуемого:

$$\begin{aligned} W(u) &= M[\rho] = \\ &= M[(g(u) - y^{TP})^2], \end{aligned} \quad (2.18)$$

а во втором случае — дисперсия реального результата:

$$\begin{aligned} W(u) &= M[(g(u) - M[g(u)])^2] = \\ &= D[g(u)]. \end{aligned} \quad (2.19)$$

В практике исследований эффективности технических систем показатели (2.18) и (2.19) обычно используют как вспомогательные.

В качестве показателя эффективности операций, связанных, например, с распознаванием некоторых ситуаций, образов, объектов и т. п., теория статистических решений рекомендует выбирать средние потери (средний риск),

1. Формы функции соответствия и показателя эффективности

Требуемый результат y^{TP}	Функция соответствия $\rho(y(u), y^{TP})$	Показатель эффективности $W(u) = M\{\rho(y(u), y^{TP})\}$	Название показателя эффективности	Примечание
Случайное событие A	$\rho = \begin{cases} 1, & \text{если наступило } A \\ 0, & \text{если не наступило } A \end{cases}$	$P_u(A)$	Вероятность случайного события	—
Задан параметром y^{TP}	$\rho = \begin{cases} 1, & \text{если } y(u) \geq y^{TP} \\ 0, & \text{если } y(u) < y^{TP} \end{cases}$	$P(y(u) \geq y^{TP}) = 1 - F(y^{TP})$	Степень вероятностной гарантии достижения результата не ниже требуемого уровня	—
y^{TP} — нечеткая переменная с функцией принадлежности μ		$P(y(u) \geq \tilde{y}^{TP}) = M[\mu(y^{TP})]$		
y^{TP} — случайная переменная ϑ_H с функцией распределения $F_H(y)$		$P(y(u) \geq \vartheta_H) = \int F_H(y) dF(y)$		

Требуемый результат y^{TP}	Функция соответствия $\rho(y(u), y^{TP})$	Показатель эффективности $W(u) = M\{\rho(y(u), y^{TP})\}$	Название показателя эффективности	Примечание
Не определен (или $y^{TP} = \max(\min) y(u)$)	$\rho = y(u)$	$M[y(u)]$	Средний резуль- тат	Если $y(u)$ неслу- чайная перемен- ная, то $W(u) =$ $= y(u)$
Задан параметром y^{TP}	$\rho = (y(u) - y^{TP})^2$	$M[(y(u) - y^{TP})^2]$	Средний квадрат уклонения резуль- тата от требуемого	—
Не определен	$\rho = (y(u) - M[y(u)])^2$	$D[y(u)]$	Дисперсия резуль- тата	Результат $y(u)$ — случайная вели- чина
Задана степень га- рантии α	$\rho = F^{-1}(1 - \alpha) = y_\alpha$	y_α	Вероятностно-га- рантированный ре- зультат	—
Не определен	$\rho = \Pi_{ij}$	$\sum_i \sum_j \Pi_{ij} P(H_i) P(H_j/H_i)$	Средний (байесов- ский) риск (сред- ние потери)	Π_{ij} — функция потерь

которые появляются при неправильном (ошибочном) распознавании.

В этих задачах в качестве функции соответствия принимаются потери P_{ij} , связанные с тем, что объект, принадлежащий множеству H_i , был ошибочно (при $i \neq j$) отнесен к множеству H_j . Здесь, следовательно, функция соответствия имеет вид

$$r(y(u), y^{TP}) = P_{ij}.$$

Пусть $P(H_i)$ — вероятность того, что объект принадлежит множеству H_i ; $P_u(H_j/H_i)$ — условная вероятность отнесения объекта из множества H_i к множеству H_j (вероятность ошибки при $i \neq j$).

Показатель эффективности операции по распознаванию объекта записывается в виде средних потерь (байесовский средний риск):

$$W(u) = M[r] = \sum_i \sum_j P_{ij} P(H_i) P(H_j/H_i). \quad (2.20)$$

И в этом случае показатель эффективности имеет форму (2.2), т. е. является математическим ожиданием функции соответствия $r = P_{ij}$.

Индекс u в записи условной вероятности $P_u(H_j/H_i)$ показывает, что она зависит от стратегии распознавания $u \in U$.

Рассмотренные формы функции соответствия и показателя эффективности приведены в табл. 1.

Рассмотрим один из способов введения показателя эффективности для случая конфликтных ситуаций при наличии неопределенности в поведении оппонента. Пусть имеет место конфликтная ситуация и на результат операции влияет не только выбор стратегии $u \in U$ лицом, принимающим решение, но и выбор стратегии v оппонентом из известного множества его допустимых стратегий V . Таким образом, результат операции $y(u, v)$ зависит от $u \in U$ и $v \in V$. Функция соответствия, как и в предыдущих примерах, будет измерять степень соответствия реального результата операции требуемому:

$$r = r(y(u, v), y^{TP}).$$

При наличии случайных факторов введем математическое ожидание функции соответствия (условный показатель эффективности):

$$W(u, v) = M[r(y(u, v), y^{TP})] \quad (2.21)$$

и выдвинем гипотезу поведения оппонента: он выбирает свои стратегии $v \in V$ так, чтобы при любой стратегии ЛПР $u \in U$ минимизировать степень соответствия реального результата операции желаемому для ЛПР (т. е. интересы оппонента противоположны интересам ЛПР). В этих условиях в качестве показателя эффективности для ЛПР можно выбрать минимальное значение $W(u, v)$ для каждой стратегии $u \in U$, т. е.

$$W(u) = \min_{v \in V} M[r(y(u, v), y^{TP})]. \quad (2.22)$$

Эта форма показателя эффективности операции уже отличается несколько от (2.2). Однако функция соответствия $r(y(u, v), y^{TP})$ в выражении (2.22) может принимать любой вид из тех, которые были рассмотрены в предыдущих примерах. Смысл показателя $W(u)$ заключается в том, что он является гарантированным уровнем условного показателя $W(u, v)$. При любой стратегии оппонента и фиксированной стратегии ЛПР $W(u) \leq W(u, v)$, т. е. $W(u)$ является нижней границей среднего значения функции соответствия реального результата операции требуемому (для ЛПР).

3. КРИТЕРИЙ ЭФФЕКТИВНОСТИ

Критерий эффективности K есть правило, позволяющее сопоставлять стратегии, характеризующиеся различной степенью достижения цели, и осуществлять направленный выбор стратегий из множества допустимых.

Критерий эффективности вводится на основе определенной концепции рационального поведения.

Существуют три концепции рационального поведения систем (выработки решений): пригодности, оптимизации, адаптивизации.

Согласно концепции пригодности рациональна любая стратегия u , при

которой выбранный показатель эффективности принимает значение не ниже некоторого приемлемого уровня W^{TP} , т. е.

$$W(u) \geq W^{TP}, u \in U, \quad (2.23)$$

где U — множество допустимых стратегий.

Если показатель эффективности векторный, то неравенство (2.23) записывается для каждого частного показателя $W_i(u) \geq W_i^{TP}$, входящего в состав векторного показателя эффективности. Таким образом, уровень удовлетворения W^{TP} делит множество допустимых стратегий U на два непересекающихся подмножества: U^{SAT} — множество приемлемых (пригодных) стратегий и множество неприемлемых стратегий $U \setminus U^{SAT}$. Все приемлемые стратегии $u \in U^{SAT}$ равноценны (одинаково удовлетворительны), как и все неприемлемые стратегии из множества $U \setminus U^{SAT}$ одинаково неудовлетворительны. Подобная концепция приводит к негибкой и нецелеустремленной системе действий.

Концепция оптимизации считает рациональными те стратегии $u \in U$, которые обеспечивают максимальный эффект в операции, т. е.

$$W(u^*) = \max_{u \in U} W(u). \quad (2.24)$$

Оптимальная стратегия может быть неединственной, т. е. решение задачи (2.24) может дать множество равноценных оптимальных стратегий $U^* \in U$.

Использовать концепцию оптимизации можно в том случае, если комплекс условий проведения операции строго фиксирован, а показатель эффективности $W(u)$ — скаляр. Эта концепция приводит к целеустремленной, но не гибкой системе действий, так как не учитывается текущая информация об изменениях различного рода, происходящих в системе и во внешней среде при реализации решения u^* .

Концепция адаптивизации предполагает возможность оперативного реагирования в ходе операции на поступающую текущую информацию об изменении комплекса условий проведе-

ния операции. Суть концепции адаптивизации заключается в изменении стратегий управления u (стратегия понимается в более широком смысле и включает как способы изменения параметров системы, так и способы изменения ее структуры) на основе не только априорной, но и текущей и прогнозной информации с целью достижения или сохранения определенного состояния системы при изменяющемся комплексе условий проведения операции.

Множество допустимых стратегий U может видоизменяться в процессе поступления текущей информации. В качестве реакции на поступающую информацию и прогноза развития операции система может изменять цель своего функционирования.

В этом случае согласно концепции адаптивизации рациональной следует считать такую адаптивную стратегию $u(t)$ из множества $U(t, \tau)$, которая обеспечивает, например, выполнение условия

$$W_t(u^*(t), \tau) \geq W_t^{TP}(u(t), \tau), u(t) \in U(t, \tau), \quad (2.25)$$

где t — время; τ — упреждение прогноза.

Запись W_t означает, что показатель эффективности может меняться во времени.

Концепция адаптивизации приводит к целеустремленной и гибкой системе действий.

В рамках концепции пригодности рекомендуется выбор стратегии из условия (2.23). Приведем основные критерии пригодности.

1. Критерий приемлемого результата. В условиях определенности в качестве показателя эффективности может быть выбран результат операции $y(u)$, цель которой носит количественный характер. В этом случае $p(y(u), y^{TP}) = y(u)$. Критерий приемлемого результата рекомендует выбирать стратегию u из множества допустимых, удовлетворяющую условию:

$$y(u) \geq y^{TP}, \quad (2.26)$$

где y^{TP} — приемлемый уровень результата.

2. Критерий допустимой гарантии. Согласно этому критерию функцию соответствия выбирают в виде

$$\rho(y(u), y^{TP}) = \begin{cases} 1, & \text{если } y \geq y^{TP}; \\ 0, & \text{если } y < y^{TP}; \end{cases} \quad (2.27)$$

и в виде (2.9) при нечетком определении требуемого результата; в виде (2.10) при случайном характере последнего. Принцип допустимой гарантии рекомендует выбирать стратегии u из условия

$$u^* : P(y(u) \geq y^{TP}) \geq P^{TP}, \quad (2.28)$$

где P^{TP} — приемлемый уровень вероятностной гарантии.

3. Критерий допустимого гарантированного результата. Согласно этому критерию функция соответствия имеет вид

$$\rho = F_u^{-1}(1 - \alpha) = y_\alpha(u).$$

Пригодную стратегию выбирают из условия

$$u^* : y_\alpha(u) \geq y^{TP}, \quad (2.29)$$

где y^{TP} — допустимый уровень гарантированного (с вероятностью α) результата.

В зависимости от вида функции соответствия в рамках концепции оптимизации выделяют следующие критерии оптимальности.

1. Критерий наибольшего результата. По аналогии с принципом приемлемого результата здесь $\rho(y(u), y^{TP}) = y(u)$. Оптимальную стратегию выбирают из условия (2.24), т. е.

$$u^* : \max_{u \in U} y(u). \quad (2.30)$$

2. Критерий наибольшего среднего результата. При существенном влиянии случайных факторов часто сам результат $y(u)$ выбирают в качестве функции соответствия. Тогда показатель эффективности есть математическое ожидание результата (средний результат), т. е.

$$W(u) = M[y(u)]. \quad (2.31)$$

Критерий наибольшего среднего результата рекомендует выбирать в ка-

честве оптимальной стратегию u^* , для которой

$$u^* : \max_{u \in U} M[y(u)]. \quad (2.32)$$

3. Критерий наибольшей вероятностной гарантии результата. Если результат операции выражается случайной переменной $y(u)$ и четко определен требуемый результат y^{TP} , то функцию соответствия вводят в следующем виде:

$$\rho(y(u), y^{TP}) = \begin{cases} 1, & \text{если } y(u) \geq y^{TP}; \\ 0, & \text{если } y(u) < y^{TP}. \end{cases} \quad (2.33)$$

Показатель эффективности в форме (2.2) в этом случае есть вероятность того, что реальный результат операции примет значение не ниже требуемого уровня, т. е.

$$W(u) = P(y(u) \geq y^{TP}). \quad (2.34)$$

Критерий наибольшей вероятностной гарантии рекомендует в качестве оптимальной выбирать стратегию $u^* \in U$ из условия

$$u^* : \max_{u \in U} P(y(u) \geq y^{TP}). \quad (2.35)$$

Это условие сохраняется при нечетком определении y^{TP} с помощью функции принадлежности $\mu_A(y^{TP})$. В этом случае функцию соответствия следует отождествить с функцией принадлежности, т. е.

$$\rho(y(u), y^{TP}) = \mu_A(y^{TP}). \quad (2.36)$$

Условие (2.35) примет вид $u^* = \max_{u \in U} P(y(u) \geq \bar{y}^{TP})$. При случай-

ном характере y^{TP} условие (2.35) сохраняется в следующем виде:

$$u^* : \max_{u \in U} P(y(u) \geq y^{TP}). \quad (2.37)$$

4. Критерий наибольшего гарантированного результата. При случайном характере результата $y(u)$ операции гарантированным результатом (вероятностно-гарантированным результатом) называют уровень $y_\alpha(u)$, не ниже которого будет получен реальный результат с заданной вероятностью α , т. е.

$$\alpha = P(y(u) \geq y_\alpha(u)). \quad (2.38)$$

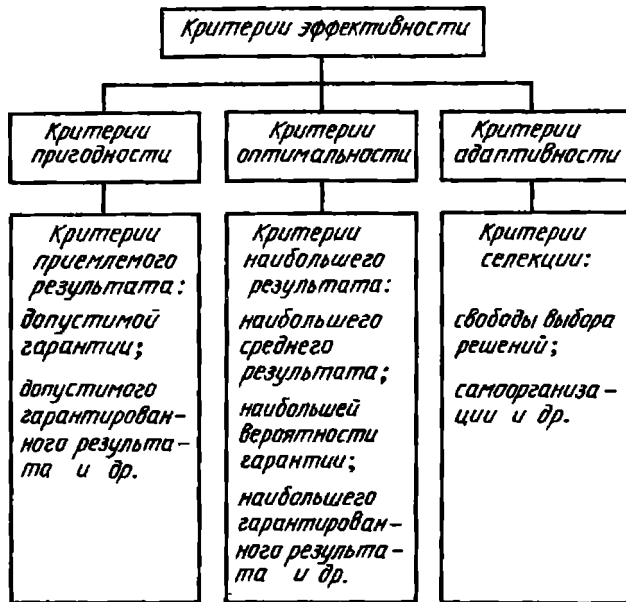


Рис. 3. Классификация критериев эффективности

Отсюда $y_{\alpha}(u) = F^{-1}(1 - \alpha)$, и функция соответствия есть $\rho = y_{\alpha}(u)$.

Следовательно, показатель эффективности в силу того, что $y_{\alpha}(u)$ не является случайной переменной (2.17), принимает вид

$$W(u) = y_{\alpha}(u). \quad (2.39)$$

Согласно критерию наибольшего гарантированного (вероятностно-гарантированного) результата оптимальную стратегию выбирают из условия

$$u^* : \max_{u \in U} y_{\alpha}(u) \quad (2.40)$$

при фиксированной вероятности α .

Пусть в условиях конфликта показатель эффективности $W(u)$ имеет вид (2.22). Критерий наибольшего гарантированного результата в этом случае в качестве оптимальной рекомендует выбирать стратегию u^* из условия

$$u^* : \max_{u \in U} \min_{v \in V} M[\rho(y(u, v), y^{TP})]. \quad (2.41)$$

Это условие отражает принцип максимина. При этом стратегию u^* называют максиминной.

В рамках концепции адаптивизации рациональное поведение сложной системы организуется в соответствии со следующими основными принципами выбора критериев эффективности.

1. Принцип селекции. На каждом шаге многоэтапного процесса принятия решений этот принцип предусматривает отбор (селектирование) нескольких решений, близких к лучшим. Затем из этих решений следует сформировать ряд комбинаций и на следующем этапе принятия решений отобрать несколько комбинаций, близких к лучшим, пользуясь критерием более высокого порядка, чем на предыдущем шаге. Таким образом формируется ряд усложняющихся комбинаций решений, из которого отбираются не только самые лучшие решения, но и близкие к ним.

2. Принцип свободы выбора решений. Этот принцип может быть реализован при организации поведения систем, способных к самоорганизации (обладающих L — качеством). Он рекомендует не принимать на основе априорной информации решения во всех деталях на всю обозримую перспективу (иногда его называют принципом неокончательных решений). Свобода выбора здесь заключается в возможности пересмотра, уточнения ранее принятого решения в зависимости от текущей информации.

3. Принцип самообучения. Согласно этому принципу адаптивное поведение системы обеспечивается в процессе многократных внешних воздействий на систему, запоминанием реакций на эти воздействия и результатов реагирования, а также корректировкой реакций, направленной на повышение эффективности поведения системы. Накапливаемая информация в процессе самообучения используется для совершенствования критерия эффективности.

Классификация критериев эффективности приведена на рис. 3.

Важнейшим положением, которого всегда следует придерживаться при выборе критерия эффективности операции, является согласование цели операции и критерия эффективности.

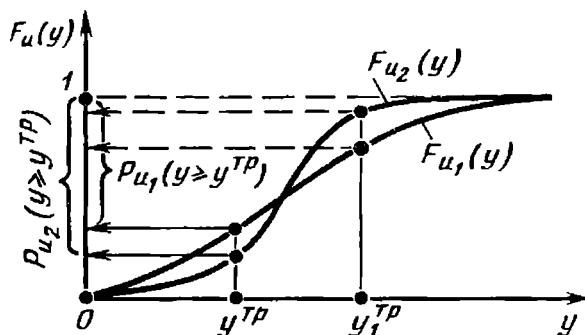


Рис. 4. Функции распределения реального результата операции при использовании стратегий u_1 и u_2

В случаях, когда определен требуемый результат операции, критерий наибольшей вероятностной гарантии лучше согласован с целью операции, чем критерий наибольшего среднего результата. Последний не учитывает дисперсии, характеризующей рассеяние реального результата операции. На рис. 4 изображены функции распределения реального результата операции $F_{u_1}(y)$ и $F_{u_2}(y)$ при использовании стратегий u_1 и u_2 . Предположим, что математические ожидания результатов для стратегий u_1 и u_2 одинаковы, т. е. $M[y(u_1)] = M[y(u_2)]$. При фиксированном значении требуемого результата, как видно из чертежа, $P_{u_1}(y \geq y^{TP}) < P_{u_2}(y \geq y^{TP})$. Следовательно, предпочтение следует отдать стратегии u_2 . Однако, если значение требуемого результата выбрать достаточно большим (y_1^{TP}), то преимуществом будет обладать стратегия u_1 ; но при этом уровень вероятностной гарантии неприемлемо низок. Таким образом, решения, принимаемые на основе критерия наибольшей вероятностной гарантии, могут существенно зависеть от уровня требуемого результата. Аналогичный вывод можно сделать в отношении критерия наибольшего гарантированного результата, при котором решения зависят от уровня гарантии α . На рис. 5 показана схема выбора стратегии по критерию наибольшего гарантированного результата

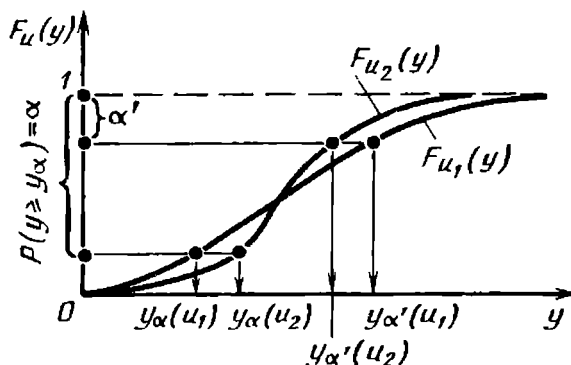


Рис. 5. Схема выбора стратегии по критерию наибольшего гарантированного результата

для случая, когда из двух допустимых стратегий u_1 и u_2 при уровне вероятностной гарантии α предпочтение отдается стратегии u_2 ($y_{\alpha}(u_1) < y_{\alpha}(u_2)$). Однако при уровне гарантии α' следует выбрать стратегию u_1 , так как $y_{\alpha'}(u_2) < y_{\alpha'}(u_1)$, но этот уровень гарантии, как показано на рисунке, неприемлемо мал. Таким образом, критерий наибольшей вероятностной гарантии следует использовать при не слишком больших значениях требуемого результата y^{TP} (обычно меньших среднего результата, т. е. $y^{TP} < M[y(u)]$). Критерий наибольшего гарантированного результата используют при достаточно высоком уровне степени гарантии ($\alpha > 0,6 \div 0,7$).

Критерий наибольшего среднего результата в практике исследования операций получил наибольшее распространение. Это обусловлено аддитивностью показателя среднего результата (2.13), что в значительной мере в некоторых случаях облегчает его расчет. Однако показатель среднего результата (2.12), лежащий в основе этого критерия, не учитывает в явном виде требуемый результат. Кроме того, ориентация на средний результат оправдана при массовом повторении операции. При однократном проведении операции (уникальные операции, системы) нецелесообразно использовать критерий наибольшего среднего результата.

Глава 3. Основные принципы исследования эффективности в технике

1. МЕТОДОЛОГИЧЕСКИЕ УРОВНИ ИССЛЕДОВАНИЯ ЭФФЕКТИВНОСТИ ТЕХНИЧЕСКИХ СИСТЕМ

Практическое использование методологии системного анализа обусловлено целью и глубиной исследования, сложностью изучаемого объекта, степенью неопределенности его поведения и другими обстоятельствами. Системный подход к исследованию эффективности технических систем реализуется в различных формах и проявлениях. Это связано с большим разнообразием систем по их составу, структуре, целевому предназначению, способам использования, условиям функционирования, а также по принципам усложняющегося поведения.

Основой изучения сложных систем служат принципы усложняющегося поведения. Эти принципы позволяют установить существенные моменты, исходные положения, лежащие в основе поведения той или иной системы в зависимости от ее сложности.

В системологии (теории сложных систем) [67] установлен ряд принципов усложняющегося поведения систем.

Принцип вещественно-энергетического баланса. Поведение системы в любых условиях не приводит к нарушению законов сохранения вещества, энергии. Этот принцип лежит в основе поведения всех материальных систем — от простейших до предельно сложных. В практике исследования систем он важен при моделировании систем. Часто сама модель системы записывается как уравнение, отражающее вещественно-энергетический баланс системы. Для простейших систем рассмотрен-

ный принцип является основным, определяющим их поведение. Эти системы назовем *e*-системами.

Принцип гомеостазиса. Система, реализующая в своем поведении этот принцип, должна обладать свойством гомеостазиса, т. е. иметь возможность возвращаться в состояние устойчивого равновесия, будучи выведенной из него внешним воздействием. Подобное устойчивое гомеостатическое поведение системы обусловлено наличием в звене управления отрицательной обратной связи. Рассмотренный принцип является ведущим в поведении, например, систем регулирования (в том числе и автоматического). Системы, реализующие в своем поведении в качестве основного принцип гомеостазиса, называют гомеостатическими, или *h*-системами. Управление в этих системах сводится к функции регулирования.

Принцип выбора решений. Сложные системы способны организовывать свое поведение на основе рационального выбора альтернатив из некоторого неединичного их множества (на основе выбора решений). Ведущим этот принцип становится для систем, реализующих так называемое индуктивное поведение, т. е. поведение, основанное на непосредственном опыте, наблюдении ситуации и принятии решения в зависимости от сложившейся ситуации без предвидения дальнейшего развития событий. Не обладая достаточным объемом памяти, такие системы не могут предвидеть будущие ситуации и, следовательно, не могут предварительно, заблаговременно принимать решения с учетом прогноза ситуаций. Такие системы называют решающими (без предвидения), или *s*-системами. Теория статистических ре-

шений в основном рассматривает именно эти решающие системы.

Принцип перспективной активности. Система может организовывать свое поведение, основываясь на предшествующем опыте в предположении, что будущие ситуации не могут существенно отличаться от прошлых. Обладающая достаточно емкой памятью и прогнозируя дальнейший ход развития ситуаций на основе ретроспективного анализа (т. е. анализируя развитие процесса в прошлом), система способна предвидеть будущее и планировать свое дальнейшее поведение (заблаговременно принимать решения, регламентирующие ее поведение в будущем). Системы, для которых принцип перспективной активности является ведущим в организации их поведения, называют предвидящими, или *p*-системами.

Принцип рефлексии. Система может организовывать свое поведение с учетом возможного мысленного представления о ее действиях распорядителя другой системы, с которой первая находится в определенных отношениях. Под рефлексией понимают отражение мыслительного процесса другого лица. Анализируя ход мысли, например, конкурента (в условиях конфликтной ситуации), ЛПР может продемонстрировать ему действия (быть может ложные намерения), стимулирующие принять решение, выгодное для ЛПР. Если конкурент такое решение принимает, то это значит, что ЛПР осуществляет рефлексивное принуждающее управление конкурентом, навязывая ему свою волю путем передачи стимулов для принятия решений. В этом случае ЛПР находится в первом ранге рефлексии, а конкурент — в нулевом (он не анализирует мыслительный процесс лица, принимающего решения). Однако и конкурент может находиться в первом ранге рефлексии (анализировать ход мысли ЛПР), и тогда рефлексивное управление затрудняется. Но ЛПР может подняться до второго ранга рефлексии (анализировать ход мысли конкурента о процессе мышления ЛПР). Оказавшись рангом рефлексии выше конкурента, ЛПР вновь становится способным рефлексивно управлять им.

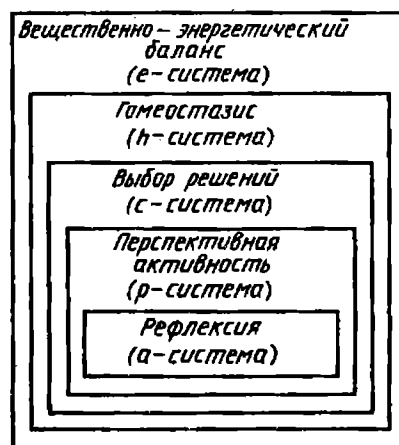


Рис. 1. Принципы усложняющегося поведения систем

Системы, организующие свое поведение на основе принципа рефлексии, должны быть наделены интеллектом (содержать в звене управления творчески мыслящих людей). Эти системы относятся к классу весьма сложных. Такие системы называют проницательными (рефлексивными), или *a*-системами.

Принципы поведения указаны в порядке усложнения систем. Система, для которой определенный принцип является ведущим, реализует в своем поведении все предшествующие принципы, но неспособна организовать свое поведение на основе последующих принципов. Так, например, предвидящие системы, реализуя первый, второй и третий принципы, а также принцип перспективной активности в качестве ведущего, неспособны реализовать принцип рефлексии.

На рис. 1 приведены принципы усложняющегося поведения систем в виде вложенных прямоугольников. Символически это значит, что если поставить точку внутри прямоугольника с надписью рефлексии, то эту точку будут содержать внутри себя все остальные прямоугольники (все принципы поведения реализуют системы, способные к рефлексивному управлению).

При анализе рационального поведения системы следует прежде всего установить ведущий принцип, положенный в основу поведения системы, т. е. установить тип системы. Затем необходимо выявить концепцию выработки решений, лежащую в основе органи-

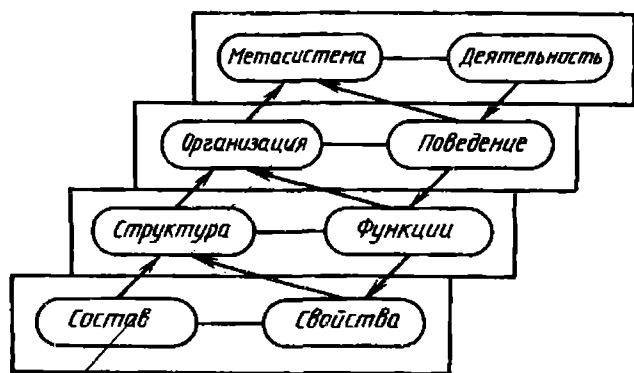


Рис. 2. Схема методологических уровней анализа систем

зации рационального поведения системы, т. е. в основе управления системой.

Опыт исследования эффективности технических систем показывает целесообразность введения четырех качественно различных методологических уровней анализа систем. На рис. 2 изображена схема методологических уровней анализа систем:

I уровень — «состав-свойства» (или элементарный уровень);

II уровень — «структура-функции» (или агрегативный уровень);

III уровень — «организация-поведение» (или системный уровень);

IV уровень — «метасистема-деятельность».

Объектом исследований в теории эффективности является операция, которая проводится в рамках S_0 -системы. Первый методологический уровень охватывает приемы и методы исследований элементов этой системы, их состав и свойства. Методология исследований на этом уровне основательно развита в различных естественных и технических науках. Большая роль в них отводится экспериментальным исследованиям. На этом уровне анализируются различные свойства элементов системы, выявляются их конструктивные, эксплуатационные, экономические характеристики и т. п. Теория эффективности на данном уровне исследует в основном качества элементов системы как совокупность их полезных свойств. На этом уровне понятие эффективности системы еще не вводится, поскольку система в целом и целенаправленный процесс ее функционирования не рассматриваются.

Объектом исследований на уровне «структура-функция» являются операции, проводимые в рамках ограниченных по своим масштабам и разнообразию функций S_0 -систем. Эффективность систем, изучаемых на данном уровне, может быть достаточно адекватно отражена, как правило, скалярным (обобщенным) показателем в метрической шкале. Обычно цель операции, проводимой в рамках этих систем, достаточно четко выражена и однозначно определена. Системы данного уровня имеют относительно простую структуру, связи между элементами стабильны, их число и разнообразие невелико, а внешняя среда имеет сравнительно постоянный характер. Как правило, в организации поведения этих систем ведущим является принцип перспективной активности, а сами они рассматриваются как подсистемы некоторой более сложной системы.

На методологическом уровне «организация-поведение» исследуются крупномасштабные операции (S_0 -системы большого масштаба). Обычно эффективность таких операций проявляется как их многомерное свойство, которое далеко не всегда удается адекватно отобразить скалярным показателем эффективности. Чаще приходится использовать векторный показатель. Системы данного уровня будем называть организациями. Организация может включать в свой состав несколько ТС, взаимодействия между которыми могут проявляться в самых различных формах.

В свою очередь, ТС, представляющие собой большие человеко-машинные системы со сложной иерархической структурой, включают в свой состав подсистемы (агрегаты), сложность которых не превышает сложности систем второго уровня исследования. Связи между подсистемами нестабильны (лабильны). Их интенсивность может изменяться во времени в зависимости от складывающихся ситуаций. Внешняя среда, как правило, имеет изменчивый динамический характер.

Системы данного уровня способны к самоорганизации (обладают L -качеством), и ведущим принципом их поведения часто является принцип рефлекс-

сии (α -система). Основа деятельности таких систем строится обычно на концепции адаптивизации.

Наличие в системах управления ТС людей, наделенных правом принимать решения и имеющих различные цели, вносит существенную неопределенность поведенческого характера и в значительной мере затрудняет формализованное описание S_0 -системы третьего уровня (S_0^{III}). Такая система может быть формализована лишь при условии, что достаточно четко определены гипотезы поведения субъектов системы, отражающие преследуемые ими цели. Очевидно, эффективность поведения организации будет определяться наряду с другими факторами стратегиями ее субъектов, их возможными рефлексиями, т. е. пониманием того, как представляют себе цели данного субъекта другие субъекты системы.

Сложность, а порой и невозможность верификации поведения систем данного уровня повышает роль теоретического обоснования выдвигаемых гипотез поведения ТС, что определяет необходимость исследований системы на более высоком методологическом уровне.

На уровне «метасистема-деятельность» исследуются глобальные системы (метасистемы), включающие в свой состав организации вместе с их внешней средой. Сложность подобных систем не допускает их формального представления. Анализ метасистемы и ее деятельности возможен лишь на вербальном (описательном) уровне. Деятельность метасистемы складывается из различных линий поведения организаций ее составляющих и преследует определенные глобальные цели. В рамках метасистемы можно рассматривать социальный и политический строй государства, систему его хозяйства в целом или отдельные отрасли экономики и т. д. В качестве метасистемы могут рассматриваться также экономические, военно-политические и другие крупные союзы и объединения государственного и международного уровня.

Деятельность метасистемы обусловлена поведением организаций,

объединенных метасистемой, ее организационными формами. В свою очередь, характер деятельности метасистемы, ее цели определяют поведение организаций, их взаимные отношения, а также структурный облик метасистемы.

Системный подход, как отмечалось, является основной методологической установкой исследования систем третьего уровня. Системный анализ есть способ реализации системного подхода к исследованию эффективности ТС, в основе которого лежит концепция метасистемы. Сущность этой концепции сводится к следующему. Исследуемая ТС третьего уровня рассматривается не изолированно, а как составная неотъемлемая часть метасистемы. С позиции метасистемы определяется полезность исследуемой системы третьего уровня, выражаемая ее эффективностью, которая связана не только со свойствами этой системы, но и со свойствами метасистемы. Таким образом, системный подход объединяет исследования на третьем и четвертом методологических уровнях. Этот подход называют системным не только потому, что исследуемую систему представляют в виде некоторой целостности, состоящей из элементов и наделенной определенной структурой, но главным образом потому, что эту систему рассматривают с точки зрения метасистемы, ее целей, на достижение которых направлена деятельность метасистемы.

2. ПРИНЦИПЫ СИСТЕМНЫХ ИССЛЕДОВАНИЙ ЭФФЕКТИВНОСТИ

Принципиальную неформализуемость метасистем можно объяснить, исходя, в первую очередь, из теоремы Геделя о неполноте формальных систем. Согласно этой теореме в рамках некоторой формальной системы невозможно вывести все истинные утверждения, относящиеся к объектам, описываемым средствами этой формальной системы. Необходима другая, более широкая система, в пределах которой формируются принципиальные положения для систем более низкого уровня.

ня, т. е. необходимо внешнее дополнение для того, чтобы справиться с геделевской трудностью. *Принцип внешнего дополнения* [7] является фундаментальной идеей теории систем. Для теории эффективности этот принцип имеет основополагающее значение.

Роль внешнего дополнения не сводится лишь к преодолению геделевской трудности, а позволяет в определенной мере ограничить изучаемую на третьем уровне систему; вычленив ее из метасистемы как некоторую целостность; выдвинуть гипотезы поведения субъектов системы и тем самым перейти к формализованному описанию систем на уровне «организация-поведение». Внешнее дополнение согласует цели организаций и их поведение с целями деятельности метасистемы. Внешне цели деятельности метасистемы и поведения входящих в нее организаций могут быть весьма различными.

Системный анализ, так же как и теория исследования операций, относится к нормативным теориям. Нормативная теория основывается на научных положениях, формируемых без специального обоснования статистикой, экспериментом. По уровню строгости формулировок различают три группы таких положений: аксиомы, постулаты, гипотезы. Системный анализ эффективности ТС в качестве исходных утверждений использует гипотезы ввиду сложности объекта исследования и большой степени неопределенности факторов, определяющих его эффективность.

Выдвижение гипотез осуществляется при исследовании деятельности метасистемы. Только на уровне метасистема-деятельность можно теоретически осмыслить поведение систем третьего уровня. Ввиду отсутствия внешнего дополнения метасистема неформализуема. Следовательно, вырабатываемое в рамках метасистемы внешнее дополнение, включающее различного рода гипотезы поведения систем, входящих в метасистему, формируется на вербальном уровне (в неформализованном виде). Воспринять в таком виде внешнее дополнение может лишь человек. Вот почему участие

человека принципиально необходимо в звене управления системы третьего уровня. Лицо, принимающее решение, формализуя внешнее дополнение, исходит из задач системы и использует при выработке решений определенные принципы. Внешнее дополнение позволяет ввести объективные критерии эффективности для систем третьего уровня, исходя из полезности этих систем в деятельности метасистемы. С изложенных позиций понятны роль и место теории принятия решений как научного направления в теории систем.

Решения принимаются и в процессе функционирования систем второго уровня. Однако внешнее дополнение для этих систем может быть сформировано в рамках исследований на уровне организация-поведение в таком формализованном виде, который обеспечивает выбор решения в системах второго уровня на основе четких формальных критериев. В этих условиях участие человека в звене управления таких систем не обязательно. В принципе эти системы могут быть автоматическими. В рамках исследований второго уровня формируется внешнее дополнение для исследования на первом уровне.

Таким образом, внешнее дополнение призвано исключить субъективизм исследователя, вызванный произволом в выборе критерия эффективности. Ошибка в выборе критерия может свести на нет все усилия исследователя при анализе эффективности систем.

Внешнее дополнение является тем логическим замыканием, которое в совокупности со свойствами (качеством) исследуемой системы, условиями и способами ее использования составляет необходимые условия для определения ее эффективности.

Важным принципом системного анализа эффективности в технике является *принцип декомпозиции систем*. В схеме рассматриваемых уровней исследований его главная роль видна в том, что сложную систему третьего уровня можно расчленить на ряд менее сложных подсистем, исследования которых могут быть проведены на втором уровне. Однако расчленение систем на подсистемы и общей задачи

на подзадачи требует в последующем согласования действий подсистем и решаемых ими задач с целью восстановления эмерджентных свойств системы, утрачиваемых при ее декомпозиции. Принцип декомпозиции выступает как средство, позволяющее снижать уровень сложности исследований. Обычно легче исследовать несколько систем второго уровня, чем одну систему третьего уровня.

Системные исследования эффективности в зависимости от степени обобщения учитываемых факторов можно разделить на *обобщенные* и *детальные*. В рамках обобщенных исследований в зависимости от целей анализа эффективности выделяют концептуальный и операционный уровни.

Целью *концептуального* исследования является установление общих тенденций развития изучаемого процесса, форм и способов его организации, выработка концепций по всем основным вопросам организации и проведения крупномасштабных операций, разработка систем целей и задач, принципов применения больших технических систем. Это исследование проводится с позиций метасистемы и организуется в основном на четвертом, а также на третьем (в общей части) методологических уровнях. Степень обобщения факторов здесь наиболее высокая.

Концептуальные исследования, как правило, проводятся на ранних этапах принятия решений для определения областей возможных стратегий на множестве концепций, устанавливаемых неформальным путем. Это обеспечивает содержательность и рациональную взаимосвязь всех остальных этапов исследования эффективности.

Операционное исследование имеет целью более подробное изучение направлений и вариантов действий в рамках концепций, рекомендованных к дальнейшему анализу концептуальным исследованием. Степень обобщения факторов здесь ниже, чем при концептуальных исследованиях, что позволяет определить функциональные структуры операций, технические средства, которые должны приводить к достижению поставленных задач,

а также перечни задач конкретным подсистемам (средствам), показатели и критерии их выполнения, сами подсистемы и связи между ними.

Операциональные исследования организуются на втором и частично на третьем методологических уровнях с общей задачей выработки практических рекомендаций по выбору стратегий систем третьего уровня и тактик действий систем второго уровня.

Детальные исследования направлены на анализ качества подсистем (агрегатов, узлов, элементов), входящих в состав систем, рассматриваемых на втором и третьем методологических уровнях. Эти исследования проводятся, как правило, на первом методологическом уровне; для них характерна низкая степень обобщения факторов.

Детальные исследования — последний уровень декомпозиции общей задачи исследования эффективности — обеспечивают решение всех вопросов, касающихся облика конкретных средств.

Концептуальные, операциональные и детальные исследования, взаимно дополняя друг друга, образуют единое целое.

Глубокое понимание методологических уровней исследования эффективности в технике, овладение методологическими принципами исследований на каждом уровне позволяют в значительной мере расширить роль метода моделирования и повысить эффективность прикладных научно-технических исследований в целом.

3. ОБЩАЯ СХЕМА ИССЛЕДОВАНИЙ ЭФФЕКТИВНОСТИ ОПЕРАЦИЙ

Формы исследований эффективности в технике весьма разнообразны. По-видимому, не может существовать универсальной рецептурной схемы проведения исследования эффективности операций с использованием технических систем, относящихся даже к одной отрасли техники. Однако может оказаться полезным некоторый перечень

обобщенных положений, придерживаясь которого исследователь организует свою деятельность при решении конкретной научно-технической задачи, разделяя ее на этапы. Ниже приведен перечень таких положений, составляющий обобщенную схему исследования эффективности операций.

I. Проблемный анализ

1. Определение существования проблемы (установление, есть ли проблема в действительности или она является мнимой), установление причин ее возникновения и взаимосвязи с другими проблемами; уяснение и формулировка проблемы, определение ее актуальности, разрешимости и срочности решения.

2. Выявление, анализ и описание ситуации, т. е. комплекса условий, в которых возникла или может возникнуть проблема.

3. Формирование и анализ проблемной ситуации с использованием эвристических и формальных методов; определение степени полноты и достоверности информации о проблемной ситуации и ее достаточности для оценки действительной необходимости и возможности решения проблемы с учетом сложившихся условий, а также для формирования целей деятельности, решающих проблему; выработка при необходимости комплекса мер для получения недостающей информации и дополнительного определения проблемной ситуации.

4. Формирование и анализ альтернативных целей $\alpha_0 = \{A_0^1, A_0^2, \dots, A_0^m\}$, достижение которых решает проблему, и выбор одной из них в качестве цели операции.

5. Анализ путей достижения цели и определение существенных ограничений (экономических, технических и др.), влияющих на выбор средств и способов достижения цели; декомпозиция сложной цели, расчленение ее на подцели и задачи.

6. Выбор средств достижения цели, обоснование уровня их качества в ряду усложняющихся уровней качества $\pi = \{R, I, C, A, L\}$; оценка наличных или потребных ресурсов.

7. Общая постановка задачи исследования эффективности операции.

II. Концептуальные исследования

1. Описание метасистемы, анализ ее деятельности.

2. Вычленение S_0 -системы, в рамках которой проводится (реализуется) операция.

3. Оценка информационной достаточности.

4. Обоснование методологического уровня исследования эффективности операции $v = \{I, II, III, IV\}$.

5. Установление типа операции (S_0 -системы), исходя из уровней усложняющегося поведения $\mu = \{e, h, c, p, a\}$.

6. Обоснование концепции рационального поведения $\gamma = \{П, О, А\}$.

7. Выдвижение гипотез поведения субъектов системы.

8. Обоснование состава и содержания внешнего дополнения, формирование требуемого результата операции (целеполагание), выбор пространства стратегий субъектов системы, установление основных ограничений (дисциплинирующих условий), вытекающих из деятельности метасистемы.

9. Выбор обобщенных показателей эффективности, описание результата операции (полезных эффектов и затрат), обоснование функции соответствия реального результата требуемому, выбор шкал показателей.

10. Обоснование принципа выработки концептуального решения и введение на этой основе критерия эффективности.

11. Концептуальное моделирование, кибернетическое описание операции (S_0 -системы), факторизация задачи (выявление эффектов факторов и их взаимодействий), установление направлений операционных исследований, уточнение и конкретизация задач исследования эффективности операции.

III. Операциональные исследования

1. Постановка задачи операционального исследования.

2. Оценка информационной достаточности.

3. Формирование (уточнение) множества стратегий по результатам концептуальных исследований.

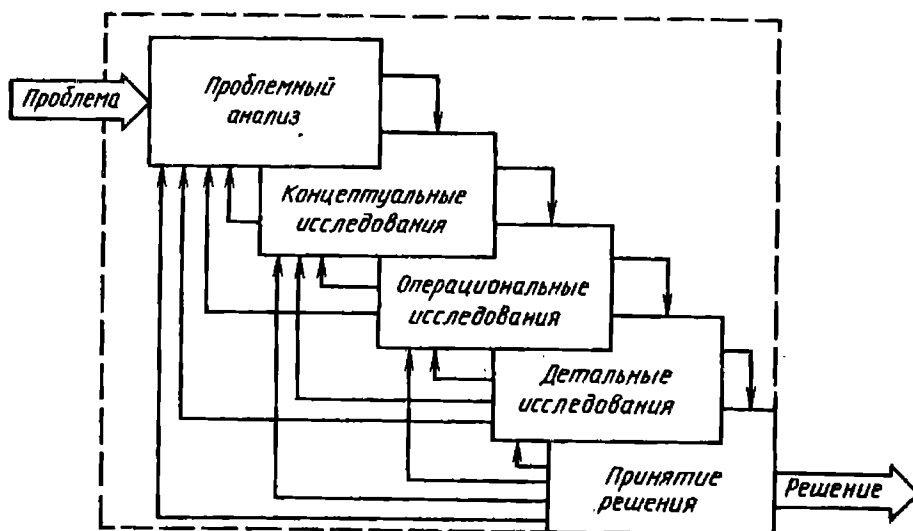


Рис. 3. Основные этапы исследования эффективности операций

4. Уточнение характеристик активных средств, описание их свойств, оценка наличных или потребных ресурсов.

5. Выбор способа и средств операционального моделирования.

6. Формирование операциональной модели, установление перечня исходных данных, введение операторов перехода и выхода, уточнение ограничивающих условий, проверка операциональной модели.

7. Уточнение показателей эффективности операции.

8. Обоснование принципа выработки операциональных решений и формирование критерия эффективности.

9. Операциональное моделирование, выработка рекомендаций для принятия решения в различных вариантах задачи исследования эффективности операции.

IV. Детальные исследования

1. Постановка задачи детального исследования.

2. Уточнение внешнего дополнения по результатам операционального исследования (требования к качеству элементов, требования к управлению качеством и т. д.).

3. Согласование и выбор показателя и критерия оценки качества элементов технических систем.

4. Определение характеристик качества по результатам эксперимента.

5. Решение задачи удовлетворения качеству или управления качеством элементов технической системы.

V. Принятие решения

1. Анализ результатов проведенных исследований.

2. Планирование контрольных мероприятий по внедрению принимаемого решения.

3. Утверждение принятого решения.

На рис. 3 приведена схема основных этапов исследования эффективности операций и показана связь между ними. Проблемный анализ непосредственно не входит в исследование эффективности операции. Он предваряет это исследование, являясь необходимым его начальным этапом. Этап принятия решения носит больше организационный характер, чем исследовательский, являясь завершающей стадией всей работы. Порядок пунктов внутри каждого этапа в зависимости от конкретного исследования может быть изменен, а часть пунктов — опущена. Кроме того, внутри этапов, а также между этапами возможны циклы, т. е. возвраты после выполнения определенных пунктов (этапов) к предшествующим пунктам в зависимости от получаемых промежуточных результатов.

С формальной точки зрения на этапе проблемного анализа реализуется отображение

$$a_0 \times \pi \rightarrow M_1, \quad (3.1)$$

где M_1 — множество метасистемных описаний определенных технических отраслей.

Отображение (3.1) каждой паре цель — средство ее достижения ставит

в соответствие конкретное метасистемное описание технической области, совокупность средств которой способна рационально решить выявленную проблему.

На этапе концептуальных исследований реализуется отображение

$$M_1 \times \mu \times v \times \gamma \rightarrow M_2, \quad (3.2)$$

где M_2 — множество вариантов внешнего дополнения с конкретным описанием S_0 -системы, в рамках которой реализуется исследуемая операция.

Операциональное исследование формально можно представить отображением

$$M_2 \times D \rightarrow U^*, \quad (3.3)$$

где D — множество вариантов исходных данных задачи выбора рационального решения; U^* — множество рациональных решений при различных вариантах исходных данных.

Этап принятия решения реализует отображение

$$U^* \times \sigma \rightarrow U^*, \quad (3.4)$$

где σ — множество возможных ситуаций, сложившихся к моменту принятия решения.

Последовательная реализация отображений (3.1)—(3.4) представляет собой формальную схему исследования эффективности операции. В перспективе подобный формальный подход может составить основу для автоматизации системных исследований эффективности в технике. Однако в настоящее время практическая реализуемость такого подхода существенно сдержана большой степенью неопределенности отображений (3.1)—(3.2). Тем не менее, по мере накопления системных исследований эффективности в технике, систематизации различных подходов к метасистемному описанию предметной области степень размытости (нечеткости) отображений (3.1)—(3.2) может быть в значительной мере снижена. Это откроет реальный путь к решению указанной проблемы на базе перспективной электронно-вычислительной техники и разработки соответствующего математического обеспечения в виде крупных проблемно-ориентированных имитационных систем (см. гл. 8).

Глава 4. Общее описание проблемы принятия решений при исследовании эффективности технических систем

1 ПРОБЛЕМА ПРИНЯТИЯ РЕШЕНИЙ

Проблема принятия решения связана с выбором направления (варианта) действий для достижения цели операции. Для этого необходимо решить ряд типовых задач, подготавливающих основное решение. Эти задачи и принимаемые по ним частные решения носят вспомогательный характер.

Проблема принятия решений при исследовании эффективности технических систем рассматривается в данной главе как проблема информационного взаимодействия типовых задач и формирования на их основе элементов модели проблемной ситуации с целью выработки в конечном итоге научно обоснованного решения.

В широком смысле решение есть процесс выбора одного (рационального) варианта действий или некоторого их подмножества из множества возможных. В узком смысле решение есть результат конкретного выбора варианта действий. Такой выбор осуществляет ЛПР, которое наделено определенными правами и полномочиями и несет всю полноту ответственности за последствия принимаемых решений.

Для того чтобы решить проблему, стоящую перед ЛПР, необходимо представить ее в виде одной или нескольких задач. Переходным этапом от проблемы к постановке формальных задач является проблемная ситуация, в ходе которой вербальная цель разбивается на подцели (задачи) и устанавливаются общие ограничения.

Пример 1. Анализируя проблему повышения рентабельности предпри-

ятия, ЛПР декомпозирует ее, например, следующим образом: «Повысить качество продукции. Снизить себестоимость продукции. При этом общие затраты на эти мероприятия не должны превосходить имеющихся средств».

Результатом анализа проблемной ситуации является формальная постановка задачи, которая указывает, каких результатов, в каких условиях и к какому сроку достичь.

Пример 2. «К четвертому кварталу следующего года повысить покупательную способность товаров не менее чем на 20 %, снизить их себестоимость не менее чем на 5 %; при этом суммарные затраты не должны превосходить 20 тыс. руб.».

В этом случае показателем качества продукции является уровень покупательной способности товаров.

Поставленная задача может быть решена различными способами.

Допустимые способы (в смысле наложенных ограничений) образуют множество стратегий (альтернатив).

Пример 3. Стратегиями в примере 2 могут являться: закупка в пределах выделенных ассигнований более качественного сырья, формирование премиального фонда за дополнительно реализованную продукцию, модернизация оборудования, повышение квалификации рабочих, сокращение управленческого аппарата, закупка лицензий и др.

Каждая из указанных стратегий является чистой стратегией в том смысле, что именно на ее реализацию расходуются все имеющиеся средства.

Если же указанные средства делятся в определенной пропорции между чистыми стратегиями, то такие сложные стратегии также входят в множество допустимых альтернатив.

Каждая из допустимых альтернатив обеспечивает, как правило, различный уровень решения поставленных задач. Например, имеются две альтернативы, приводящие к исходам операции со значениями показателей покупательной способности, себестоимости и затрат соответственно (25 %, 6 %, 9000 р.) и (29 %, 8 %, 15 000 р.). Указать, какая лучше из них, может лишь конкретное ЛПР применительно к понимаемой именно им данной задаче и именно в данный момент времени (сейчас и сегодня, а не завтра и потом). Это означает, что при выборе решения необходимо использовать помимо указанной еще и дополнительную информацию о предпочтениях ЛПР, его отношении к риску (из-за неопределенности условий проведения операции), об интересах других субъектов операции. Информация о предпочтениях (т. е. о представлениях ЛПР о «лучшем» и «худшем» на множествах целей, задач, условий, способов и исходах операций) может иметь качественный или количественный характер. Так, мнение ЛПР о том, что увеличение одних показателей вносит больший вклад в достижение цели операции по сравнению с изменением других, является качественной информацией об относительной важности показателей. Если же ЛПР может количественно оценить, насколько изменение одного показателя может быть компенсировано соответствующим изменением других, или же во сколько раз один показатель важнее другого, то такая информация о предпочтениях относится к количественной. Например, ЛПР может указать, что увеличение покупательной способности на 1 % эквивалентно снижению себестоимости продукции на 0,5 % в достаточно узкой области значений оцениваемых показателей.

В общем случае исход операции является неопределенным. Это вызвано неопределенностью условий проведения операции и действием факторов различной природы (см. пара-

граф 2 гл. 1). Получение значений показателей, характеризующих тот или иной исход операции, связано с решением задачи моделирования операции.

Проблема исследования эффективности операции с целью выработки решения включает три взаимосвязанных процесса (этапа): постановку задачи, получение результатов и анализ результатов.

Вопросы принятия решений, относящиеся к постановке задачи, связаны со структуризацией исходной информации о проблеме, анализом неопределенностей, формированием исходного множества стратегий, выбором показателя и критерия эффективности. Второй процесс связан с формированием модели операции (механизма ситуации) и получением оценок эффективности по результатам моделирования. Процесс анализа результатов предполагает решение задачи выбора на основе сформированного критерия эффективности или системы таких критериев.

Следует заметить, что постановка задачи, получение результатов и анализ результатов являются инвариантами для любого этапа и уровня исследования эффективности.

2. МОДЕЛЬ ПРОБЛЕМНОЙ СИТУАЦИИ

Для того чтобы охватить проблему выработки решения в целом, представить ее основные элементы, которые необходимо сформировать для получения окончательного решения о стратегии проведения операции, рассмотрим математическую модель проблемной ситуации. Эта модель отображает взаимосвязи основных элементов процесса выработки решения и последовательность формирования частных задач.

Обозначим: U — множество стратегий ЛПР; Λ — множество значений определенных и неопределенных факторов; G — множество исходов операции; Y — вектор характеристик (признаков) исхода $g \in G$, т. е. числовое выражение результата операции; H — модель, т. е. отображение, ставя-

шее в соответствие множествам стратегий U и факторов Λ множество результатов Y (G); W — показатель эффективности; Ψ — оператор соответствия «результат — показатель»; K — критерий эффективности; $\mathcal{P}$ — модель предпочтений ЛПР на элементах множества:

$D = \{U, \Lambda, G, Y, W, K\};$

θ — оставшая информация о проблемной ситуации.

Тогда модель проблемной ситуации представляется в виде системы:

$(U, \Lambda, H, G, Y, \Psi, W, K, \mathcal{P}, \theta).$ (4.1)

Наличие компоненты Λ как самостоятельного элемента в модели проблемной ситуации (4.1) предполагает, что множество значений неопределенных факторов при выработке решений будет либо обязательно установлено (задано извне), либо отыскание этих значений будет представлять самостоятельную задачу.

Проблема выбора показателя эффективности W связана с установлением вида функции соответствия результата операции Y (G) требуемому результату Y^{tr} .

Модель предпочтений $\mathcal{P}$ есть формализованное представление ЛПР о «лучшем» и «худшем» среди элементов некоторого множества. С помощью этой модели решаются важные частные задачи, связанные с формированием исходного множества альтернатив U , выделением существенных факторов Λ , определяющих условия проведения операции, построением моделей H и Ψ , выбором характеристик Y исходов $g \in G$, построением на их основе частных показателей эффективности, их агрегированием в некоторый обобщенный показатель, установлением критерия K и т. д.

Во многих практических случаях оказывается, что априорное задание одного из основных критериев эффективности (пригодности, оптимальности, адаптивности) приводит к выделению, как правило, некоторого множества «нехудших» альтернатив. В этом случае для однозначного выбора лучшей альтернативы необходимо формирование составного критерия — решающего правила, включающего как

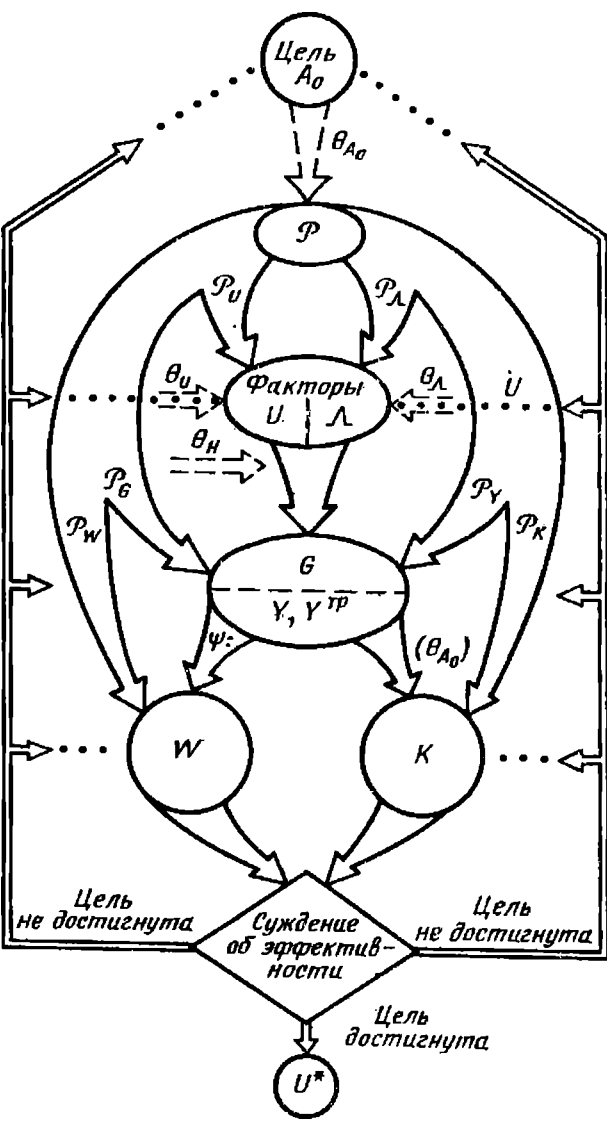


Рис. 1. Модель проблемной ситуации

формальные, так и неформальные предписания по вынесению суждения. Это решающее правило и задается элементами модели предпочтений $\mathcal{P}$ на множествах $G — \mathcal{P}_G, Y — \mathcal{P}_Y, W — \mathcal{P}_W$ и т. д.

Взаимосвязь указанных компонент модели проблемной ситуации приведена на рис. 1. Стрелками на рисунке показаны отношения между элементами проблемной ситуации. Сообразуясь с информацией θ_{A_0} о цели операции, ЛПР последовательно формирует множества U и Λ на основе информации θ_U и θ_Λ , опираясь на подмодели $\mathcal{P}_U$ и $\mathcal{P}_\Lambda$ модели предпочтений $\mathcal{P}$. Аналогично на основе подмоделей предпочтений $\mathcal{P}_G$ и $\mathcal{P}_Y$ и с учетом информации θ_H об имеющихся

средствах построения моделей H выбираются характеристики Y исхода G и устанавливается вид соответствия $H : U \times \Lambda \rightarrow Y(G)$, а также формируется величина требуемого результата Y^{TP} . Далее по информации Y , Y^{TP} с учетом предпочтений $\mathcal{P}_W$ о виде показателя эффективности устанавливается один из возможных видов метрики $\rho(Y, Y^{TP})$ и формируется модель Ψ «результат — показатель». Одновременно формируется критерий K по информации θ_{A_0} и подмодели предпочтения $\mathcal{P}_K$, которая может задавать критерий в форме решающего правила. На основе суждения о степени достижения цели операции либо осуществляется выбор лучшей альтернативы из множества $U^* \in U$, либо осуществляется возврат и коррекция элементов модели проблемной ситуации.

Задачи, соответствующие двум основным процессам принятия решений при исследовании эффективности, формируются на основе модели проблемной ситуации и имеют вид:

для процесса получения результатов

$$\Psi : \{Y | H : U \times \Lambda \xrightarrow{\theta} Y(G)\} \xrightarrow{\theta} W; \quad (4.2)$$

для процесса анализа результатов

$$\mathcal{P} \xrightarrow{\theta} K : U \xrightarrow{W} U^*. \quad (4.3)$$

В (4.2) запись $\{H : U \times \Lambda \xrightarrow{\theta} Y(G)\}$ означает множественность моделей, соответствующих системным уровням исследования. При этом, исходя из концепции метациели (внешнего дополнения), информация θ для низших уровней иерархии может быть получена либо с верхнего уровня, либо сформирована на основе информации от ЛПР рассматриваемого системного уровня.

В выражении (4.3) символом U^* обозначено подмножество «наилучших» с точки зрения ЛПР стратегий, из которых окончательно выбирают реализуемое решение $u^* \in U^*$. Символом A_0 на рис. 1 обозначена цель операции.

3. ЗАДАЧИ ПРИНЯТИЯ РЕШЕНИЙ

На основе модели проблемной ситуации (4.1) могут быть получены различные постановки частных задач принятия решений. Каждую такую задачу можно представить в виде логического высказывания типа: (Дано; требуется определить). В этом случае символ «;» разделяет две стороны проблемы — то, что известно, и то, что требуется найти.

Для этапа постановки задачи характерны следующие частные задачи принятия решений.

Задача структуризации исходной информации. При рассмотрении основных этапов был выделен самостоятельный пункт — проблемный анализ, который, являясь связующим информационным звеном между проблемой и моделью проблемной ситуации (4.1), содержательно раскрывается на этапе анализа проблемы. Поскольку все пункты проблемного анализа представляют структуризацию исходной информации о проблеме, то формально этот этап может быть сформулирован как следующая задача:

$$\begin{aligned} &\langle \theta; \theta_{A_0}, \theta_U, \theta_\Lambda, \theta_{\mathcal{P}} = \\ &= \{ \theta_{\mathcal{P}_G}, \theta_{\mathcal{P}_Y}, \theta_{\mathcal{P}_U}, \theta_{\mathcal{P}_\Lambda}, \theta_{\mathcal{P}_K} \} \rangle, \end{aligned} \quad (4.4)$$

где $\theta_{\mathcal{P}}$ — часть общей информации о проблеме, касающаяся предпочтений лица, принимающего решения:

на множестве возможных исходов G операции и различии их по предпочтительности относительно цели A_0 операции;

на множестве существенных характеристик Y исхода $g \in G$;

на множестве возможных способов (стратегий) U достижения цели операции;

на множестве Λ условий проведения операции;

на возможных концепциях (гипотезах) рационального поведения системы для определения в дальнейшем критерия эффективности.

Эта задача решается только эвристическими методами.

Задача анализа неопределенности. Классификация факторов по характеру неопределенности является одним из результатов решения задачи анализа неопределенности вида:

$$\langle \theta_{A_0}; \Lambda, \theta_{\Lambda} \rangle, \quad (4.5)$$

где θ_{A_0} — информация о цели операции; θ_{Λ} — информация о типе и характеристиках множества неопределенных факторов.

Решение этой задачи требует привлечения и анализа информации, получаемой на основе аналогий, исторического опыта, экспериментальных и статистических данных, результатов экспертиз. Особенностью данной задачи является применение для ее решения как формальных, так и неформальных (эвристических) методов. Тем не менее эта задача может успешно решаться лишь в том случае, если будут созданы мощные базы данных и базы знаний на основе современной информационно-вычислительной техники.

Это позволит оперативно получать, анализировать и обрабатывать информацию, касающуюся природы неопределенных факторов, диапазонов их изменения, априорного распределения вероятностей на них, психологических особенностей принятия решений другими субъектами исследуемой системы, типов взаимодействия между ними (нейтрализм, содействие, противодействие) и др.

К важнейшим задачам типа (4.5) относятся задачи математической статистики, касающиеся параметрического и непараметрического оценивания (см. т. 2), идентификации, кластеризации, прогнозирования [57]. Класс применяемых методов решения этих задач широк и достаточно хорошо разработан. Например, по существующим оценкам только методов прогноза насчитывается свыше 150. К ним относятся, прежде всего, морфологический, корреляционный, регрессионный, факторный и спектральный анализ, аппарат цепей Маркова (некоторые из этих методов приведены в т. 2), метод группового учета аргу-

ментов [25], методы распознавания образов и др.

Задача формирования исходного множества стратегий. Следуя модели проблемной ситуации, схема которой представлена на рис. 1, задачу формирования исходного множества стратегий наряду с задачей анализа неопределенности относят к одному из начальных этапов исследования эффективности операции, что определяет ее важное самостоятельное значение. Формальная постановка этой задачи выглядит следующим образом:

$$\langle \theta_{A_0}, \theta_U, \theta_{\Lambda}, \theta_{\mathcal{P}_U}, \Lambda, U \rangle. \quad (4.6)$$

Задача (4.6) не является тривиальной. Действительно, с одной стороны, множество альтернатив U , «заполняющее пустоту» между желаемым результатом (целью операции) и условиями ее проведения (Λ), должно быть по возможности более широким. Это обеспечит в дальнейшем свободу выбора решений ЛПР и сведет к минимуму возможность упустить «лучшее» решение $u^* \in U$.

Исходное множество стратегий U должно быть обозримым и, следовательно, достаточно узким, что позволит ЛПР в дальнейшем провести верификацию альтернатив на имеющихся в его распоряжении моделях при ограничениях на ресурсы (время, деньги и т. п.).

Многие компоненты задачи формирования исходного множества стратегий явно не заданы и эвристически формируются ЛПР.

Проблема удовлетворения противоречивых требований к множеству исходных стратегий U должна решаться по следующим основным направлениям, которые диктуются $\mathcal{P}_U$.

1. При решении вопроса о включении той или иной альтернативы u в множество U следует ориентироваться на цель операции, определяя, каков будет ответ на вопрос: обеспечивает ли данный вариант действий $u \in U$ достижение желаемого результата в условиях Λ хотя бы в принципе?

Уже этот этап позволяет ЛПР отсеять огромную часть потенциального множества альтернатив из-за их очевидной непригодности с точки зрения достижения цели операции. Такая «оче-

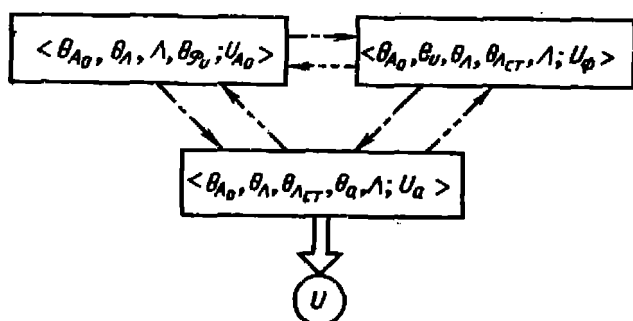


Рис. 2. Схема процесса формирования исходного множества стратегий

видность» весьма субъективна. Она отражает способность ЛПР увидеть среди множества возможных целей (если цель не задана экзогенно, извне) главную, а также предсказать возможные цели на будущее. Сформированное таким образом множество U_{A_0} можно назвать целевым.

2. Среди всех альтернатив множества U_{A_0} необходимо на основании информации θ_U выделить подмножество U_{Φ} физически реализуемых стратегий, отвечающих требованию: может ли быть данная альтернатива $u \in U_{A_0}$ реализована сейчас или в заданные на операцию сроки?

Решение этого вопроса связано с уровнем общетеоретических и конструкторских разработок, с прогнозом их развития. Кроме того, в множество U_{Φ} включаются только те альтернативы, которые удовлетворяют ограничениям по затратам, качественно оцениваемым ЛПР и экспертами. Причем эти затраты включают также средства, выделяемые на дополнительные исследования, обеспечивающие готовность альтернативы к заданному сроку начала операции.

3. Множество физически реализуемых целевых стратегий $U_{\Phi} \subseteq U_{A_0}$ должно быть проверено на недоминированность по полноте охвата факторов Λ — информация θ_{Λ} . Это означает, что ЛПР должно ответить на вопросы:

имеются ли среди множества U_{Φ} такие стратегии U_a , которые обеспечивают свободу выбора решений в ходе операции?

являются ли эти стратегии «гибкими» по отношению к изменяющимся неопределенностям природного и поведенческого характера?

Таким образом, для дальнейшего рассмотрения остаются лишь те стратегии $U_a \subseteq U_{\Phi}$, которые обеспечивают возможность эффективной корректировки первоначально намеченного плана операции при изменении условий ее проведения или инвариантные к условиям стратегии. Первое означает, что риск (потери), связанный с корректировкой плана в ходе операции, будет минимальным. Второе — выбор стратегий на основе прогноза поведения других субъектов операции обеспечит выгодное течение операции.

Такая трехэтапная процедура выделения исходного множества стратегий $U = U_a$ носит циклический характер, предполагающий возврат к «старому» на более высоком уровне в случае выявления противоречий на одном из приведенных этапов.

В целом процедуру формирования исходного множества стратегий можно представить схемой, изображенной на рис. 2, на которой стрелками с одной точкой показано направление «естественного» хода процесса, а стрелками с двумя точками — возврат в случае выявления противоречий. Символом $\theta_{A_{CT}}$ обозначена информация, касающаяся возможности включения альтернативы u в U к заданному сроку начала операции T при выделенных на ее разработку средствах C , а символом θ_a — информация о возможности корректировки первоначального плана операции u при изменении условий ее проведения.

Задача моделирования исходов операции. В рамках рассматриваемых моделей проблемной ситуации формальная запись задачи моделирования исходов операции, т. е. построения отображения H , выглядит следующим образом:

$$\langle \theta_{A_0}, U, \Lambda, \theta_H, \mathcal{P}_G, \mathcal{P}_Y; Y, \{H\} \rangle, \quad (4.7)$$

где $\{H\}$ — веер моделей.

Необходимость, в общем случае, рассмотрения не одной модели H ,

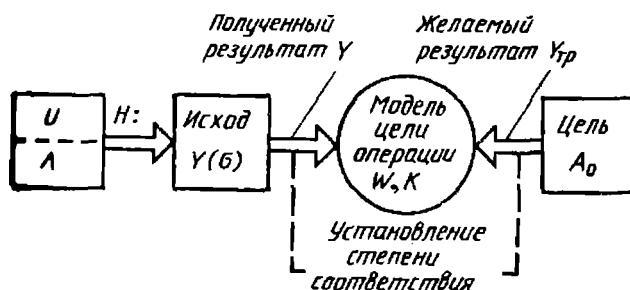


Рис. 3. Модель цели операции

а многоуровневой их иерархии $\{H\}$ определяется особенностью трехуровневой схемы исследования эффективности, при которой описание более высокого уровня зависит от обобщенных и факторизованных переменных низшего уровня. Такой подход при решении задачи моделирования исходов операции позволяет структуризовать процесс моделирования операции в целом, существенно облегчить оценивание ее эффективности, выявить наиболее значимые факторы, определяющие эффективность. Такая методология решения задачи (4.7) предъявляет вполне определенные требования к моделям, используемым на различных уровнях иерархии исследования.

Задача моделирования цели операции. Для того чтобы исследования были содержательными (т. е. были бы получены некоторые количественные оценки для выработки решения), прежде всего необходимо достаточно аккуратно и четко формализовать цель операции. Такая постановка задачи исследования эффективности операции вынуждает использовать количественные шкалы при формализации цели операции. Формализация цели операции должна базироваться на концепции внешнего дополнения, как необходимом условии получения обоснованного решения.

Цели и задачи операции гомоморфно отображаются в систему показателей их достижения. Вводится критерий эффективности. Система показателей эффективности и критерий образуют модель цели операции (рис. 3).

Задача моделирования цели операции формально ставится следующим образом:

$$\langle \theta_{A_0}, U, \Lambda, Y^{tr}, Y(G), \mathcal{P}_W, \mathcal{P}_K; W, K \rangle. \quad (4.8)$$

В некоторых случаях цель A_0 операции задается вышестоящим органом в виде параметров целеполагания Y^{tr} . Это обстоятельство отражено в (4.8) тем, что Y^{tr} стоит в левой части — «Дано». В общем случае параметры целеполагания являются результатом решения задачи:

$$\langle \theta_{A_0}, U, \Lambda, \mathcal{P}_Y, G; Y^{tr} \rangle.$$

Цель операции накладывает вполне определенный отпечаток на вид показателя эффективности и на сам критерий, точнее на принцип его выбора.

Числовая функция, определенная на множестве стратегий U , характеризующая степень достижения цели A_0 операции, может рассматриваться в качестве показателя эффективности лишь при соблюдении определенных требований.

А. Соответствие цели. ЛПР, зная значение показателя в той или иной ситуации, полностью представляет себе, в какой степени достигается цель.

Б. Содержательность и интерпретируемость. Сам показатель эффективности или его компоненты должны быть понятны ЛПР и иметь ясный физический смысл.

В. Измеримость. Показатели эффективности можно подвергать определенным математическим преобразованиям, допускаемым типом его шкалы.

Если эффективность операции не удается описать единственным показателем, то выбор числа частных показателей эффективности, образующих векторный показатель, представляет собой нетривиальную задачу. К векторному показателю эффективности предъявляются дополнительные требования.

Г. Минимальность числа используемых частных показателей. Человек может достаточно легко опериро-

вать с одним — тремя показателями без использования машинных методов обработки. В зависимости от степени подготовленности ЛПР верхняя граница размерности векторного показателя определяется величиной 7 ± 2 ; при этом с ростом числа показателей увеличивается время анализа получаемых результатов.

Д. Полнота. Векторный показатель эффективности должен содержать такое количество частных показателей, которое позволяло бы учитывать все определяющие признаки Y , адекватно отражающие результат операции.

Требования А—Д оказываются противоречивыми, что еще раз указывает на связь задачи принятия решения с проблемной ситуацией. Противоречивость проявляется в том, что невозможно удовлетворить всем требованиям сразу.

Например, ЛПР по значению показателя должен полностью представлять себе степень достижения цели (соответствие цели). Это приводит к необходимости удовлетворения требования полноты, которое противоречит требованию минимальности. Требование содержательности и интерпретируемости входит в противоречие с требованием измеримости, поскольку показатели эффективности могут не иметь адекватной физической шкалы (например, оценка эстетичности, эргономичности, новизны и т. п.) и их приходится описывать субъективными оценками.

В качестве показателей эффективности на концептуальном уровне исследований целесообразно выбирать агрегированные показатели, отражающие степень достижения целей операции.

На операциональном уровне исследований в качестве показателей эффективности целесообразно также использовать агрегированные показатели, которые должны характеризовать степень выполнения задач операции, решение которых приводит к достижению цели операции.

Требования к качеству решения этих задач могут быть получены на верхнем уровне — уровне концептуального исследования. Показателями качества на уровне детального исследования

могут являться как агрегированные, так и неагрегированные показатели качества элементов подсистем.

Выше были рассмотрены основные концепции принятия решений: пригодности, оптимизации и адаптивизации. На основе этих концепций формируют основные типы критериев эффективности, используемые при принятии решений.

Если операция проводится в условиях существенной неопределенности (целевой, природной и поведенческой), то в качестве критериев могут использоваться либо критерии пригодности (например, критерий гарантированного результата), либо оптимальности (например, критерий наибольшего гарантированного результата).

Однако использование этих подходов при выборе критерия в данной ситуации не всегда оправданно, так как в большинстве случаев гарантированные уровни являются недопустимо низкими и вследствие этого не несут никакой информации о рациональности действий.

Поэтому используют другой подход к формированию критериев, а именно — подход, основанный на концепции адаптивизации.

Этот подход предполагает прогнозирование возможных условий Λ и способов U на основе не только априорной (статической) информации, но также текущей динамической и прогнозной (виртуальной) информации. Используя такую информацию и применяя принцип оптимальности, можно получить не только целенаправленные, но и гибко ориентированные решения. Таким образом, переходим от статической модели принятия решения к динамической. Следовательно, суть этого подхода состоит не просто в выборе лучшего, как это имеет место в концепции оптимальности, а именно в движении к лучшему.

Математической формулировкой критерия адаптивности при выборе решений может, например, служить следующее выражение:

$$u^*(t, \tau) : \sup_{u(t) \in U(t, \tau)} W^t(u, t, \tau), \quad (4.9)$$

где u , U имеют тот же смысл, что и ранее; t — время; τ — упреждение прогноза.

Запись W^t означает, что цели операции, а следовательно, показатели и критерии могут меняться во времени, сам же показатель формируется как некоторое осреднение по факторам $\lambda \in \Lambda_j$.

Использование критерия (4.9) позволяет выбирать «нехудшую» стратегию $u^*(t, \tau)$ с точностью до ошибок прогнозирования на период прогноза τ . Эта стратегия обеспечивает наибольшее значение функции W^t на рефлексивно выбранном множестве допустимых стратегий $U(t, \tau)$, которое формируется к данному моменту времени t .

Таким образом, принцип принятия решений на основе концепции адаптивизации оказывается наиболее сложным по уровню выработки управляющих воздействий, но в то же время и наиболее совершенным.

Задача моделирования предпочтений. При решении всех изложенных выше задач перед ЛПР стоит проблема выбора объектов, подлежащих дальнейшему рассмотрению, из множества имеющихся. Осуществляя такой выбор, ЛПР руководствуется своей системой предпочтений.

Эта система предпочтений может быть выявлена и измерена в ходе контрольного предъявления ЛПР объектов из представленной совокупности. Выявленная и измеренная система предпочтений ЛПР называется моделью предпочтений.

Разработаны специальные методы моделирования предпочтений ЛПР, использующие формальные и неформальные процедуры.

Общая постановка задачи моделирования предпочтений формально записывается в следующем виде:

$$(D, \theta; \mathcal{P}_D), \quad (4.10)$$

где $D := \{U, \Lambda, Y, G, W, K\}$.

Задачу (4.10) надо понимать следующим образом. На основе имеющейся информации θ и результатов контрольного предъявления объектов, входящих в подмножество множества

D , выявить систему предпочтений $\mathcal{P}$ ЛПР.

Построение модели $\mathcal{P}$ может быть осуществлено с помощью получаемой от ЛПР специальной дополнительной информации, которая называется информацией $\Omega \subset \theta$ о предпочтениях. Типичными примерами такой информации являются независимость частных показателей по предпочтению, аддитивная независимость показателей, качественная информация об относительной важности, коэффициенты важности и др.

4. КЛАССИФИКАЦИЯ ЗАДАЧ ПРИНЯТИЯ РЕШЕНИЙ

В основу классификации задач принятия решений могут быть положены различные системы признаков. Однако сложившаяся практика исследования эффективности технических систем показывает, что наиболее общими и существенными признаками классификации являются: число лиц, принимающих решения; вид показателя эффективности; степень определенности информации о проблемной ситуации; зависимость элементов модели проблемной ситуации от времени.

По признаку числа ЛПР различают задачи индивидуального и группового принятия решений. Индивидуальные решения принимаются одним лицом, а групповые — коллективным органом (группой лиц). При групповом выборе решений определяющую роль играет задача согласования индивидуальных предпочтений членов группы. Главное здесь — объединить предпочтения отдельных лиц в единое мнение (групповое предпочтение).

В зависимости от вида используемого показателя эффективности задачи принятия решений подразделяют на задачи со скалярным и векторным показателем эффективности, которые часто еще называют соответственно скалярными и векторными задачами принятия решений.

По признаку степени определенности информации о проблемной ситуации различают задачи принятия решений в условиях определенности и неопределенности.

1. Методы решения основных классов задач

Факторы	Показатель эффективности	
	Скалярный	Векторный
Λ_F	Методы математического программирования	Методы принятия решений в условиях определенности
Λ_E	Методы стохастического программирования	Методы принятия решений в условиях стохастической неопределенности
Λ_J^N	Методы решения игр с «природой»	
Λ_J^B	Методы решения матричных игр ¹	Методы решения биматричных игр со строгим и нестрогим соперничеством ¹

¹ Классификация задач принятия решений в конфликтных ситуациях (Λ_J) проведена только по виду показателя эффективности и не учитывает структуру множеств стратегий игроков (конечное, бесконечное).

Примечание. Λ_F — множество определенных факторов; Λ_E — множество случайных факторов; Λ_J^N — множество факторов «природной» неопределенности; Λ_J^B — множество факторов «поведенческой» неопределенности.

Задачи принятия решений в условиях определенности (детерминированные задачи) характеризуются наличием полной и достоверной информации о проблемной ситуации, целях, ограничениях и последствиях принимаемых решений. В этих задачах относительно каждой стратегии $u \in U$ заранее, до проведения операции, известно, что она неизменно приводит к некоторому конкретному исходу (результату).

Задачи принятия решений в условиях неопределенности подразделяются в свою очередь на задачи в условиях *стохастической* и *нестохастической* неопределенности, а последние — на задачи принятия решений

в условиях *природной* и *поведенческой* неопределенности.

По способу описания неопределенностей нестохастического характера различают задачи принятия решений в *четкой* и *нечеткой* постановке.

Характерная особенность всех задач принятия решений в условиях неопределенности состоит в том, что исход операции зависит кроме стратегий оперирующей стороны и фиксированных факторов также от неопределенных факторов, не подвластных оперирующей стороне и не известных ей в момент принятия решения (или известных с недостаточной точностью). В результате влияния неопределенных факторов каждая стратегия $u \in U$

связанной с множеством возможных исходов операции, что существенно осложняет процесс выбора решения.

Признаку зависимости элементов от проблемной ситуации от времени различают *статические* и *динамические* задачи принятия решений. Статические задачи значительно больше статических, поскольку многие элементы динамических задач принятия решений зависят от изменений времени, описывающих поведение динамических объектов, участвующих в операции.

Классификация задач принятия

решений по перечисленным признакам приводит к различным комбинациям типов задач. Например, некоторая задача может быть классифицирована как статическая векторная задача группового принятия решений в условиях определенности. Возможны и другие комбинации.

Для решения различных типов задач разработаны определенные методы, основные из которых приведены в табл. 1. Задачи, не вошедшие в рубрикации таблицы, могут быть решены приведенными в ней методами с привлечением специальных приемов.

Глава 5. Особенности исследования эффективности технических систем на этапах жизненного цикла

1. ЖИЗНЕННЫЙ ЦИКЛ ТЕХНИЧЕСКОЙ СИСТЕМЫ

В процессе становления и развития техническая система проходит несколько этапов своего жизненного цикла. Пока нет единой классификации этих этапов для всего многообразия технических систем. Полагают, например, что жизненный цикл большой технической системы целесообразно разделить на последовательные этапы (рис. 1). Каждый из этапов может быть разделен на стадии или фазы [26]. Иногда выделяют этап совершенствования системы. Этапы во времени могут частично перекрывать друг друга. Неоднозначность выделения этапов, стадий или фаз жизненного цикла обусловлена многообразием самих систем, множественностью их целевого назначения, способов их производства и т. д. Часто этапы жизненного цикла системы объединяют в периоды. Различают обычно два периода: реализационный (или период создания системы) и эксплуатационный (или период полезной жизни системы).

Процесс становления и развития определенной технической системы является управляемым. Цель управления заключается в том, чтобы обеспечить наибольшую эффективность использования сил и средств при решении поставленной задачи на каждом этапе жизненного цикла системы.

Задачи, решаемые на каждом этапе, должны быть достаточно строго согласованы с глобальной целью, для достижения которой создается система. Основой согласования задач и целей каждого этапа жизненного цикла с общей целью являются иссле-

дования эффективности операции, направленные на достижение этой глобальной цели. Оценка эффективности операции служит также средством выработки и принятия решений на каждом этапе жизненного цикла, что в свою очередь является основой управления процессом становления и развития технической системы (ТС).

Исследования эффективности ТС на этапах жизненного цикла организуются в соответствии с общей схемой исследований (см. рис. 3, гл. 3). Этап НИР является начальным этапом жизненного цикла ТС. На этом этапе осуществляется проблемный анализ в форме поисковых НИР (при поиске возможных путей решения проблемы) и целевых НИР, в рамках которых могут быть организованы фундаментальные и прикладные исследования, направленные на обоснование рационального выбора средств достижения цели в операции. Прикладные исследования заканчиваются формированием требований к системе и разработкой технического задания. В системотехнике [48] этот этап называют *внешним проектированием системы (или системным проектированием)*.

Внешнее проектирование сложной технической системы представляет собой процесс метасистемного описания предметной области, формирования основной концепции системы, анализа условий обстановки и выработки внешнего дополнения для проведения исследований эффективности ТС, а также выработки требований к системе и разработки технического задания (ТЗ) на основе оценки эффективности ТС по обобщенным критериям эффективности.

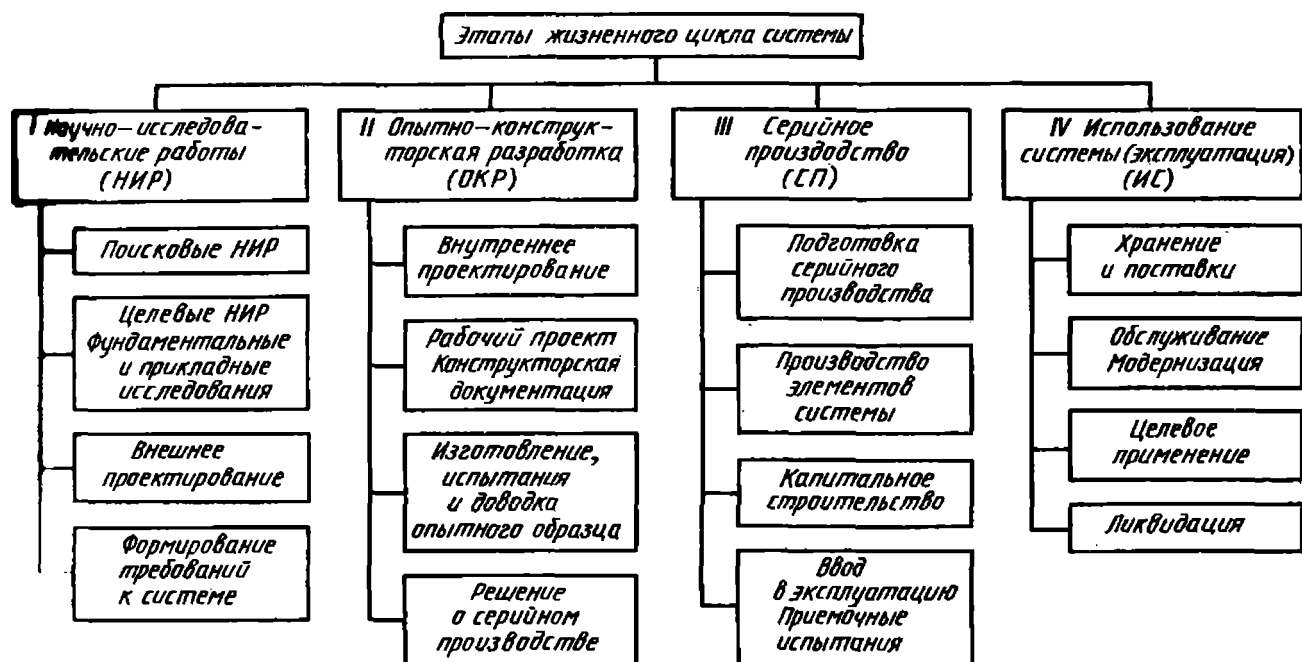


Рис. 1. Этапы жизненного цикла системы

Процесс внешнего проектирования имеет четыре стадии: определение цели проектирования; определение объекта проектирования; синтез математической модели объекта проектирования; формализация задачи проектирования. Стадии внешнего проектирования применены на рис. 2. На первой предварительной стадии анализируют прогно-

зируемые изменения внешней среды и оценивают рассогласования характеристик действующих систем относительно их требуемых значений, соответствующих допустимому уровню эффективности этих систем. Решение о разработке новой системы принимают при существенном снижении эффективности системы в прогнозируемых усло-

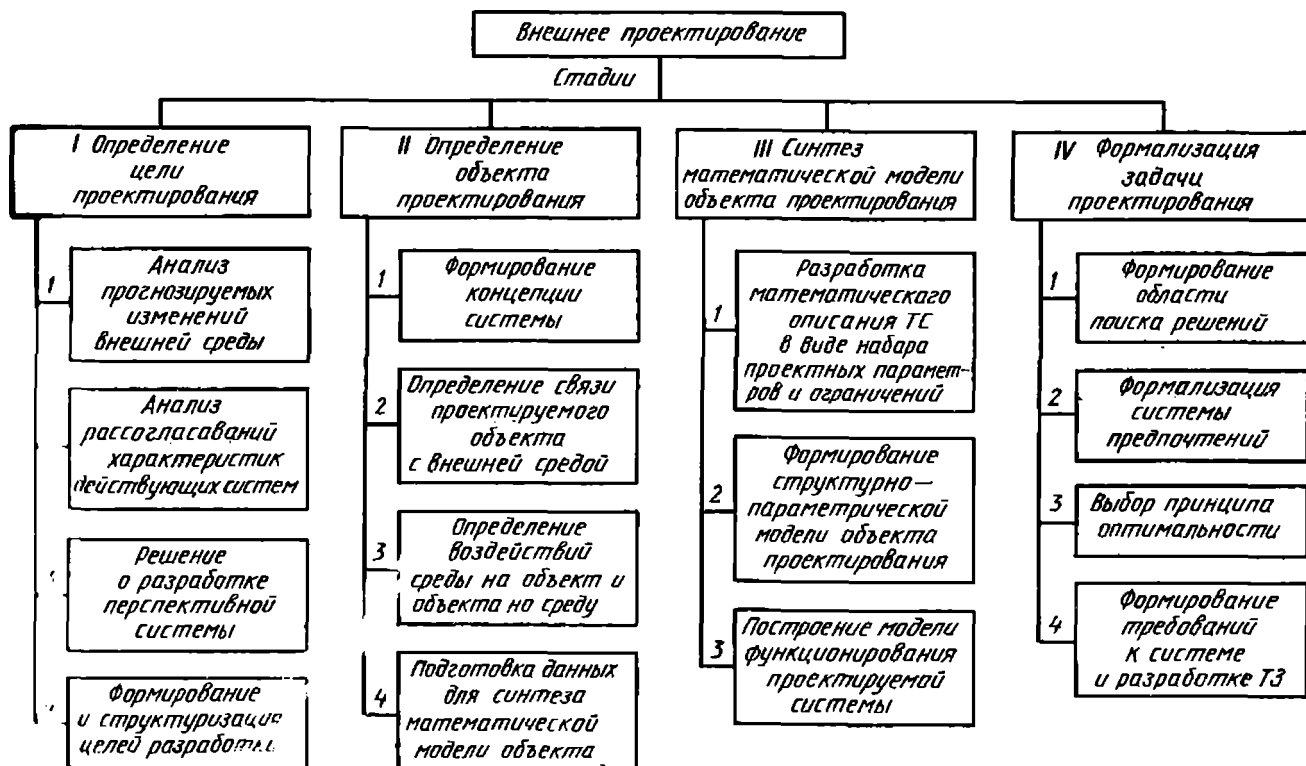


Рис. 2. Стадии внешнего проектирования

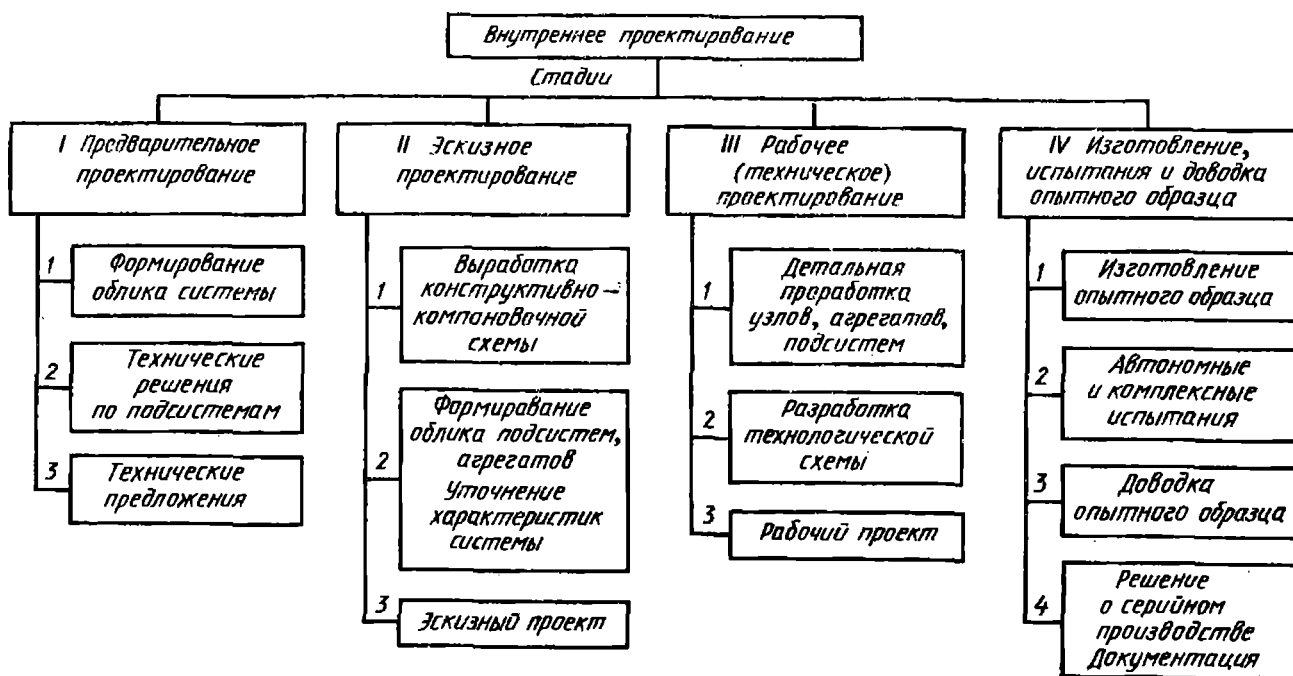


Рис. 3. Стадии внутреннего проектирования

виях обстановки. После всестороннего обоснования и принятия решения о разработке перспективной системы формируют цели разработки и проводят их структуризацию. Цели разработки тесно увязывают с той целью A_0 , для достижения которой создается система. Конкретное проектирование элементов, агрегатов и подсистем в системотехнике называют *внутренним проектированием*. Основными стадиями внутреннего проектирования являются предварительное, эскизное и рабочее проектирование. Стадии внутреннего проектирования приведены на рис. 3. Изготовление и доводка опытного образца также являются стадией внутреннего проектирования, которое заканчивается принятием решения о серийном производстве системы (или ее элементов).

2. СХЕМА ОБОБЩЕННОЙ ОПЕРАЦИИ

Методология системного анализа рекомендует рассматривать весь процесс создания и использования технической системы в рамках единого подхода. Следуя этим рекомендациям,

целесообразно ввести понятие обобщенной (укрупненной) S_{00} -системы, в пределах которой реализуется обобщенная операция (ОО), имеющая два последовательных периода: 1-й период — создание технической системы (реализационный период); 2-й период — использование созданной технической системы в качестве активного средства для достижения цели A_0 операции (период полезной жизни системы). В самом упрощенном виде задачу можно описать следующим образом. Для решения возникшей проблемы необходимо достичь цель A_0 . Активных средств достижения этой цели нет. Однако выделены (или могут быть выделены) материальные средства для разработки и производства технической системы, использование которой в операции в качестве активного средства дает возможность решать поставленную задачу, т. е. достичь цель. Требуется выбрать рациональный вариант технической системы, спроектировать и организовать процесс ее производства в соответствующие сроки так, чтобы потенциальная эффективность этой технической системы в операции, направленной на достижение цели, была по возможности наибольшей.

Таким образом, в обобщенной S_{00} -системе для достижения цели A_0 в качестве активных средств используют выделенные денежные ассигнования на создание и развитие перспективной технической системы, силы и средства проектно-конструкторских организаций и производственных объединений. Способы использования этих активных средств составляют множество стратегий первого периода обобщенной операции. Стратегиями второго периода операции являются способы использования активных средств, в состав которых входит созданная в первом периоде операции техническая система. Описанная методологическая схема, основой которой является обобщенная S_{00} -система, позволяет обоснованно подойти к согласованию целей и задач каждого этапа жизненного цикла сложной технической системы.

Цель A_0 обобщенной операции определяет назначение проектируемой системы и может быть реализована на этапе использования ТС. Цели предшествующих этапов жизненного цикла являются промежуточными целями обобщенной операции. Таким образом, этапы жизненного цикла системы являются этапами обобщенной операции, направленной на достижение цели A_0 , и формируются в результате декомпозиции последней. Способ декомпозиции обобщенной операции зависит от масштабов решаемой проблемы, сложности создаваемой технической системы, качества сил и средств, привлекаемых для ее создания, становления и развития, а также от других факторов. По-видимому, в этом причина неоднозначности деления жизненного цикла ТС на этапы, стадии и фазы.

На стадии поисковых НИР формируется множество альтернативных потенциальных целей $\alpha_0 = \{A_0^1, A_0^2, A_0^M\}$. Полагают, что достижение любой из этих целей решает возникшую проблему. Каждая альтернативная цель связана с определенной предметной областью. Поисковая НИР организуется в форме проблемного анализа, предваряющего концептуальные исследования эффективности, направленные на обоснование и выбор цели.

A_0 из множества α_0 , а также средств ее достижения.

Выбранная глобальная цель A_0 порождает множество локальных, частных целей и задач, которые играют роль средств достижения глобальной цели. Исследования целевых НИР направлены на формирование упорядоченности частных целей и задач, установление способов и средств их достижения. Определение множества частных целей и их упорядочение представляют собой процесс формирования программы достижения глобальной цели A_0 . Неопределенность в выборе способов и средств достижения частных целей и решения некоторых проблемных вопросов вынуждает иногда проводить фундаментальные исследования, результаты которых используют затем в прикладных исследованиях, формирующих теоретическую основу внешнего проектирования системы.

Процесс формирования множества промежуточных целей и задач ведут обычно последовательно, начиная с анализа факторов, непосредственно влияющих на достижение глобальной цели A_0 . При этом формируются промежуточные цели первого уровня. Затем выбирают цели второго уровня, достижение которых обуславливает выполнение целей первого уровня и так далее. На множестве промежуточных целей устанавливают отношение необходимости R (или предшествования), представляющее собой строгий частичный порядок. Так, если $A_i R A_j$, то считают, что цель A_i необходимо выполнять для достижения цели A_j (A_i предшествует A_j).

Отношение необходимости R на множестве промежуточных целей удобно представить в виде ориентированного графа. В частном случае, если нет ни одной цели, непосредственно предшествующей сразу двум целям и более, граф будет представлять собой дерево целей и задач. Если глобальная цель или какая-либо из промежуточных целей может быть достигнута не единственным способом, то на одном и том же множестве промежуточных целей можно задать различные отношения необходимости, соответствующие каждому варианту достижения цели.

Операции булевой алгебры позволяют формализовать процесс упорядочения целей и задач. В качестве примера рассмотрим следующую последовательную процедуру формирования упорядоченного множества промежуточных целей. Пусть тем или иным способом логического анализа установлено, что глобальную цель A_0 можно достичь двумя путями: выполнив либо цель A_1^1 , либо цель A_1^2 . Следовательно,

$$A_0 = A_1^1 \vee A_1^2,$$

где A_1^1 и A_1^2 — первая и вторая цели первого уровня. Цель кодируется бинарной переменной, принимающей лишь два значения: 1 (k -я цель A_i^k достигнута) и 0 (цель не достигнута). Для достижения цели A_1^1 необходимо (и достаточно) выполнить цели A_2^1 и A_2^2 , т. е.

$$A_1^1 = A_2^1 \wedge A_2^2.$$

В записи конъюнкции знак « $\wedge$ » можно опускать, т. е. $A_1^1 \wedge A_2^2 = A_1^1 A_2^2$. В дальнейшем будем пользоваться этим упрощением записи.

Кроме того, известно, что

$$A_1^2 = A_3^2 A_4^1.$$

Установлено, что

$$A_2^1 = A_3^1 A_3^2;$$

$$A_2^2 = A_3^2 A_3^3 A_3^4;$$

$$A_3^2 = A_3^2 A_3^5;$$

$$A_4^1 = A_3^4 A_3^6.$$

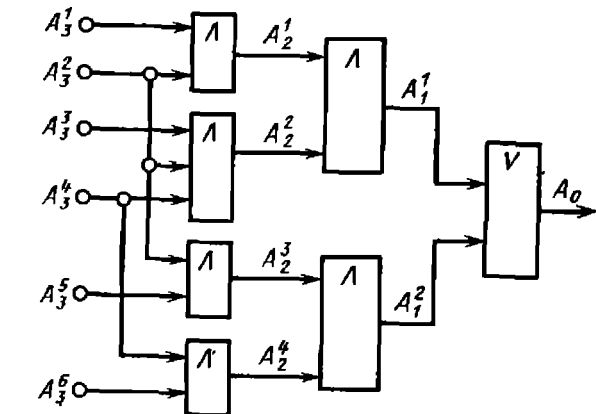


Рис. 4. Логическая схема упорядочения частных целей

Подставляя эти равенства в начальное выражение для A_0 , получим

$$A_0 = A_3^1 A_3^2 A_3^3 A_3^4 \vee A_3^2 A_3^4 A_3^5 A_3^6 = A_3^2 A_3^4 (A_3^1 A_3^3 \vee A_3^5 A_3^6).$$

Логическая схема упорядочения частных целей в рассматриваемом случае представлена на рис. 4. На схеме блок с индексом $\vee$ — есть дизъюнктор, а блоки с индексом $\wedge$ — конъюнкторы. Данную схему можно представить ориентированным графом без петель и циклов (рис. 5, а и б). Однако деревом этот граф не будет ввиду того, что A_3^2 непосредственно предшествует трем целям A_1^1 , A_2^2 и A_3^2 ; а цель A_4^1 — двум целям A_2^2 и A_3^4 . Если предположить, что A_3^2 непосредственно предшествует только A_1^1 и A_3^4 — только A_2^2 , то граф бинарного отношения необходимости на множестве целей примет форму дерева целей.

Отношение порядка на множестве промежуточных целей и задач структуризуется отнесением групп целей и задач к тому или иному этапу жизненного цикла системы. Формирование требований к системе и разработка технического задания являются целью внешнего проектирования ТС. Создание рабочего проекта и разработка конструкторской документации есть цель внутреннего проектирования, а ввод в эксплуатацию основных подсистем согласно выработанному плану является целью этапа серийного производства.

Определение цели проектирования является важнейшей стадией внешнего проектирования, за которой сле-

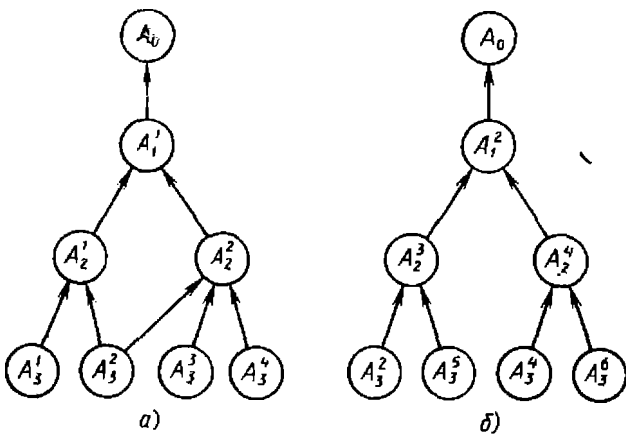


Рис. 5. Граф целей и задач

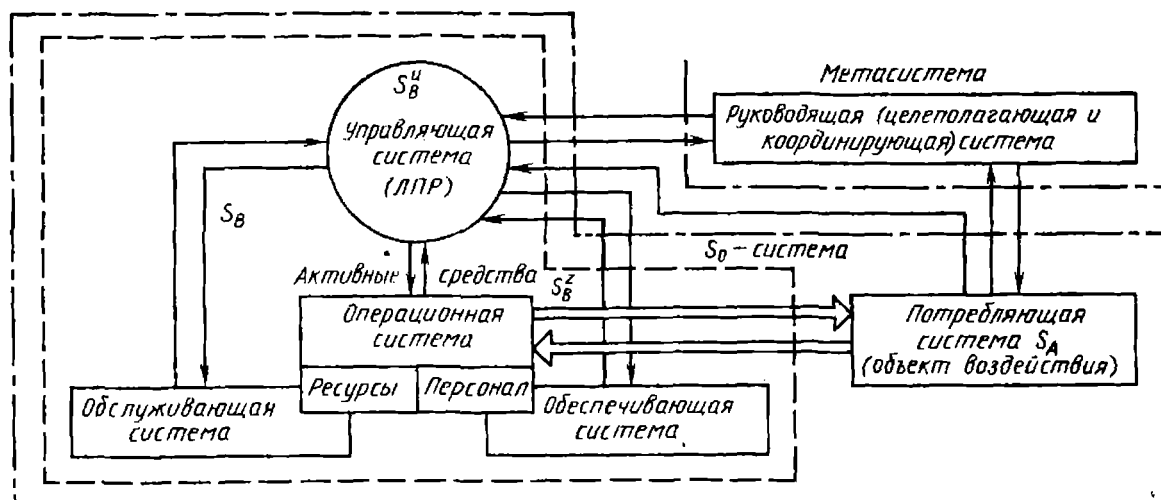


Рис. 6. Структурная схема S_0 -системы

дует стадия определения объекта проектирования. На этой стадии формируется концепция системы в целом, т. е. способ понимания и представления ее основных функций, вырабатывается общий замысел системы и пути его реализации. Кроме того, определяются связи проектируемой системы с внешней средой, устанавливается интенсивность этих связей и их направленность, а также изучаются способы взаимодействий внешней среды с проектируемой системой.

В соответствии с принципом информационной достаточности модель системы может быть построена, если объем сведений (знаний) о системе не ниже некоторого предельно допустимого уровня. Обычно объект проектирования определяют с целью получения сведений об этом объекте в объеме, превышающем минимальный уровень достаточности. Сбор, структуризация и уточнение сведений (знаний) об объекте проектирования называют *подготовкой данных для синтеза математической модели объекта*.

Формирование математической модели проектируемой системы обычно начинают с разработки математического описания технической системы в виде набора проектных параметров и ограничений. На основании такого описания строится структурно-параметрическая модель объекта проектирования, которая в общих чертах отражает облик проектируемой системы.

Структурно-параметрическая модель является основой построения модели функционирования S_0 -системы, в рамках которой проводится операция, направленная на достижение конечной цели A_0 . Вычленение S_0 -системы и построение ее функциональной модели являются определяющей стадией внешнего (системного) проектирования технической системы. На рис. 6 представлена структурная схема S_0 -системы. На схеме показаны основные компоненты S_0 -системы: управляющая система S_B^u ; активные средства S_B^z и потребляющая система S_A (или объект воздействия). На схеме показана, кроме того, руководящая система, которая входит в состав метасистемы и по отношению к S_0 -системе осуществляет функции целеполагания и координации. Управляющая система S_B^u содержит управляющий орган (лицо, принимающее решения) и средства управления (включая каналы связи).

Активные средства S_B^z обычно представляют собой некоторую операционную систему, перерабатывающую ресурсы с целью получения некоторого полезного эффекта при взаимодействии с потребляющей системой. Операционную систему может обслуживать персонал. Кроме того, в состав активных средств могут быть включены обслуживающая и обеспечивающая системы. Первая предназначена для под-

держания заданного режима работоспособности операционной системы; вторая предназначена для реализации условий, необходимых для нормального использования операционной системы. Управляющая система S_B^u и активные средства S_B^z (управляемая система) в совокупности образуют технический комплекс S_B . Потребляющая система S_A также может представлять собой технический комплекс со своим органом управления.

Объектом проектирования может быть технический комплекс S_B или активные средства S_B^z . Кроме того, часто в качестве объекта проектирования выступает либо операционная, либо обслуживающая, либо обеспечивающая системы.

В зависимости от выбора объекта проектирования устанавливается уровень его качества $\pi = \{R, I, C, A, L\}$. Как правило, A — качество вводится в случае, если объектом проектирования является полный комплект активных средств в операции, либо весь технический комплекс. Если объектом проектирования является подсистема, входящая в S_B^z (операционная система, обслуживающая или обеспечивающая система), то обычно ограничиваются уровнями качества R, I или C . Способность технической системы достичь цель операции (A — качество) выражается потенциальной эффективностью операции и характеризует качество активных средств как единой системы. Однако, если система является лишь частью активных средств, то ее качество не может быть охарактеризовано способностью достичь цель в операции. В этом случае ее качество оценивают либо устойчивостью, либо помехоустойчивостью, либо управляемостью, но иногда удается провести декомпозицию операции так, что исследуемая система является полномасштабным активным средством в новой операции. В этом случае можно ввести A — качество как характеристику этой системы, выражающую потенциальную эффективность этой новой операции.

Построение модели S_0 -системы является основой формализации задачи

проектирования. Эта стадия внешнего проектирования системы включает формирование множества допустимых решений, формализацию системы предпочтений ЛПР и установление критерия эффективности на основе введенного показателя эффективности операции. Такие исследования ведут на концептуальном уровне. При этом большое внимание должно быть уделено анализу априорной информации, способам ее получения и обработки. Внешнее проектирование заканчивается формированием рациональных требований к системе и разработкой технического задания. Выбор рациональных требований к системе осуществляется на основе оценивания эффективности различных вариантов требований. При оценивании эффективности вариантов на стадии внешнего проектирования необходимо по возможности проводить технико-экономическое обоснование каждого варианта.

Рациональные требования к системе на стадии внешнего проектирования приходится выбирать в условиях большой степени неопределенности ввиду того, что период использования проектируемой системы по назначению отдален по времени иногда на довольно большой срок; система предпочтений заказчика на стадии внешнего проектирования еще недостаточно четко выявлена; кроме того, часто слабо структурировано множество целей и задач, а множества допустимых стратегий на каждом этапе жизненного цикла достаточно велики и разнообразны.

По мере реализации (протекания) обобщенной операции степень неопределенности несколько уменьшается, так как часть неопределенных факторов становится определенной. Например, на завершающей фазе внешнего проектирования становится известным перечень требований к системе, так как уже выбран вариант системы из множества допустимых вариантов.

Обозначим: U^H — множество допустимых стратегий на этапе НИР; соответственно U^O , $U^П$, $U^И$ — множества стратегий на этапах ОКР, производства и использования системы. Тогда множество допустимых стратегий в обобщенной операции можно пред-

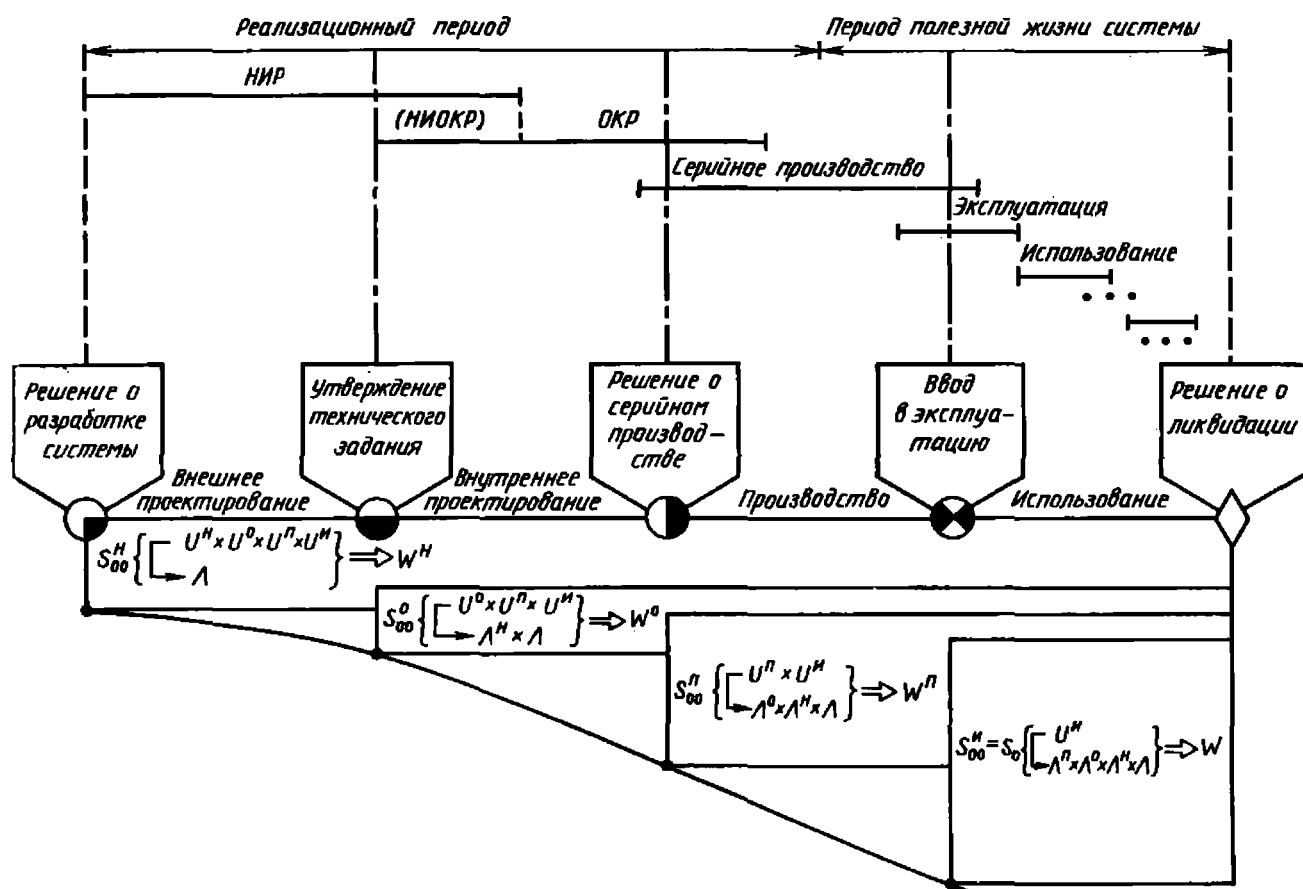


Рис. 7. Схема обобщенной операции

ставить в виде прямого произведения множеств стратегий на всех этапах жизненного цикла, т. е. $U = U^H \times U^O \times U^П \times U^И$. При завершении того или иного этапа число сомножителей в прямом произведении сокращается. Управляемые на предыдущем этапе факторы становятся на следующем этапе определенными факторами, формирующими условия нового этапа обобщенной операции. На рис. 7 приведена схема обобщенной операции, на которой показано постепенное сужение множества допустимых стратегий при переходе от одного этапа жизненного цикла к другому. Множество определенных факторов Λ расширяется вследствие поступления информации о решениях, принимаемых на каждом этапе жизненного цикла системы.

После каждого этапа жизненного цикла постепенно сокращается и объем системы S_{00} , реализующей обобщенную операцию. Системы S^H_{00} , S^O_{00} , $S^П_{00}$, реализующие соответствующие

этапы обобщенной операции, на рис. 7 изображены частично пересекающимися прямоугольниками. Высота такого прямоугольника условно показывает глубину проработки вопросов на соответствующем этапе, а сторона основания прямоугольника — широту охвата вопросов. Кривая в нижней части схемы показывает, как увеличивается глубина проработки и сокращается широта охвата вопросов с течением времени.

3. ОЦЕНИВАНИЕ ЭФФЕКТИВНОСТИ ТЕХНИЧЕСКОЙ СИСТЕМЫ В ПЕРИОД ЕЕ СОЗДАНИЯ

Система обеспечения информацией должна быть спланирована на стадии внешнего проектирования. При планировании этой системы должны быть указаны способы получения информации по ее видам на всех стадиях НИР и ОКР, порядок взаимодействий

с заказчиком с целью получения от него требуемой информации в установленные сроки. В состав информационной системы необходимо включить единый испытательный комплекс. Назначение последнего — организация системы испытаний всех видов. Единый испытательный комплекс (ЕИК) должен систематизировать все испытания, проводимые с системой и ее элементами, объединять их в общую систему планирования испытаний (экспериментов) и обработки результатов испытаний. ЕИК строится как общая информационная система испытаний, объединяющая в единую схему поисковые экспериментальные исследования на стадии поисковых НИР; лабораторные эксперименты на стадии целевых фундаментальных и прикладных исследований; экспериментальную проверку гипотез на стадии внешнего и внутреннего проектирования; автономные и комплексные заводские испытания элементов, агрегатов и подсистем; контрольные и приемочные испытания опытного образца системы; эксплуатационные испытания системы; периодические испытания технической системы и ее элементов в период эксплуатации.

Жизненный цикл системы отражается поэтапным развитием обобщенной операции как процесс постепенного накопления информации о свойствах создаваемой технической системы, условиях обстановки и способах ее использования в качестве активного средства в операции. По мере накопления информации о системе уточняется оценка эффективности операции, что приводит к более полному обоснованию принимаемых решений на каждом этапе жизненного цикла системы.

На стадии внешнего проектирования выбор рациональных требований к системе осуществляется на основе достаточно общих критериев эффективности, не приводящих к слишком сильному усечению (сужению) множества допустимых стратегий. На этой стадии использование сильного принципа оптимизации, приводящего к выбору небольшого числа оптимальных стратегий (или всего одной стратегии), не всегда оправданно потому, что на

ранних этапах жизненного цикла проследить ход развития операции во всех деталях нельзя.

На стадиях внешнего и внутреннего проектирования целесообразно организовать последовательный обмен информацией между представителями этих стадий проектирования. Обмен информацией может происходить по определенной схеме.

1. Лицо, принимающее решения на стадии внешнего проектирования ($ЛПР_H$), осуществляет выбор определенного подмножества U_H^* , т. е. выбирает несколько вариантов требований к системе, пользуясь в основном критерием пригодности. Эти варианты вместе с частными показателями эффективности $ЛПР_H$ сообщает лицу, принимающему решения на стадии внутреннего проектирования ($ЛПР_O$). Вектор частных показателей отражает систему предпочтений заказчика.

2. Лицо, принимающее решения на стадии внутреннего проектирования ($ЛПР_O$), формирует ряд альтернативных вариантов U_O^* будущей системы в соответствии с вариантами требования $ЛПР_H$. Каждый вариант рассматривается на стадии предварительного проектирования. На этой стадии формируется облик системы и вырабатываются технические предложения, исходя из оценки эффективности каждого варианта системы.

3. Информация о проведенных исследованиях эффективности каждого варианта облика системы передается $ЛПР_H$ для более полного оценивания эффективности системы в рамках уточненной математической модели функционирования проектируемой системы.

4. После нескольких циклов обмена информацией между $ЛПР_H$ и $ЛПР_O$ выбранный вариант $u^* \in U_H^* U_O^*$ всесторонне анализируется с целью удовлетворения требований заказчика. Оценка потенциальной эффективности проектируемой системы служит основой согласования требований заказчика с предлагаемыми характеристиками системы.

В процессе согласования варианта системы по мере обмена информацией между $ЛПР_H$ и $ЛПР_O$ может меняться

(усиливаться) принцип выбора и соответственно критерий эффективности.

Выбранные рациональные варианты технической системы оценивают затем на этапе серийного производства. Выбирают способ производства системы, при котором будут выдержаны требуемые характеристики. Соответствие полученных характеристик требуемым и

допустимость некоторых отклонений устанавливают оцениванием эффективности системы, проводимым на основе результатов приемочных испытаний. При этом критерий эффективности выбирают обычно таким же, каким пользуются при выборе рациональных стратегий на этапе использования технической системы.

Часть II

МОДЕЛИРОВАНИЕ И ОЦЕНИВАНИЕ ЭФФЕКТИВНОСТИ ТЕХНИЧЕСКИХ СИСТЕМ

Глава 6. Оценивание эффективности операций

1. СОДЕРЖАНИЕ ЗАДАЧИ ОЦЕНИВАНИЯ ЭФФЕКТИВНОСТИ

Процесс исследования эффективности операции независимо от того, на каком системном или методологическом уровне она рассматривается (см. гл. 3), может быть представлен в виде трех этапов: постановка задачи, получение результатов и анализ результатов (см. гл. 4). Рассмотрим этап получения результатов, на котором решается задача оценивания (измерения) эффективности операции. Поскольку результаты решения именно этой задачи являются искомыми на данном этапе, то его часто называют *этапом оценивания эффективности операции (технической системы)*.

На основе анализа результатов оценивания вырабатывается решение. Этот анализ осуществляется на третьем заключительном этапе процесса исследования эффективности операции и в зависимости от цели исследования направлен на решение задачи выбора стратегий из множества U допустимых стратегий или задачи оценки эффективности (см. параграф 1 гл. 1). Поскольку вторую задачу можно трактовать как частный случай задачи выбора, когда множество U выбора содержит только один элемент, то в дальнейшем задача оценивания эффективности операции рассматривается

в наиболее общей и характерной для практики постановке, ориентированной на решение задачи выбора стратегий.

Таким образом, оценивать эффективность операции необходимо для решения задачи выбора, так как только на основе вполне определенных результатов оценивания этот выбор можно осуществить, т. е. выработать решение. Степень обоснованности решения и его эффективность в значительной мере зависят от точности и достоверности результатов оценивания, их соответствия поставленной цели исследований и условиям реальной проблемной ситуации. Однако задача оценивания эффективности существует и имеет практический смысл постольку, поскольку существует проблема выбора. Это определяет ее роль и место в общей задаче исследования эффективности операции, главная цель которого — выработка научно обоснованного решения.

Основой для введения количественных оценок эффективности операции служит понятие показателя эффективности W как меры степени соответствия реального (прогнозируемого или полученного) результата Y операции требуемому $Y^{тр}$ или степени достижения цели A_0 операции, отображаемой требуемым результатом $Y^{тр}$.

Суть задачи оценивания эффективности состоит в измерении (количественном оценивании) уровня эффективности операции, т. е. в вычислении значения в общем случае векторного показателя эффективности $W = (W_1, W_2, \dots, W_m)$, где $W_i, i = 1, m$ — частные показатели эффективности операции.

Результатом решения этой задачи являются векторные числовые оценки $W(u)$, полученные для всех стратегий u из рассматриваемого множества U допустимых стратегий. Каждая такая оценка $W(u)$ характеризует уровень эффективности (степень достижения цели) операции при данной стратегии $u \in U$.

Содержание задачи оценивания эффективности операции определяют следующие основные процедуры:

построение математической модели операции и ее машинная реализация;

оценка качества модели и планирование экспериментов с ней;

вычисление значений $W(u)$ показателя эффективности W для всех стратегий $u \in U$ с использованием модели.

Если рассматривать оценивание эффективности операции как процесс выполнения указанных процедур, то видно, что в основе его лежит *математическое моделирование операции*.

В общем виде задачу оценивания эффективности операции можно представить формальной записью следующего вида:

$$W(u) = M[\rho(Y(u), Y^{TP})]; \quad (6.1)$$

$$\Psi: \{Y | H: U \times \Lambda \xrightarrow{\theta} Y(G)\} \xrightarrow{\theta} W, \quad (6.2)$$

где Λ — множество определенных и неопределенных факторов, формирующих обстановку операции; Y^{TP} — требуемый результат операции; Y — результат операции (вектор характеристик исхода $g \in G$ операции, важных для достижения цели A_0 операции, отображаемой требуемым результатом Y^{TP}); M — знак математического ожидания; ρ — функция соответствия; H — модель результата операции, позволяющая вычислить значения $Y(u)$ результата Y операции для каж-

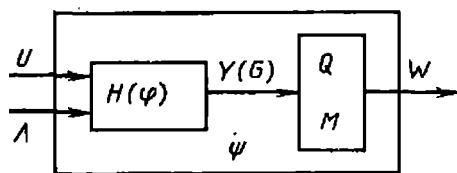


Рис. 1. Структурная схема модели операции

дой стратегии $u \in U$; θ — информация о проблемной ситуации.

Отображение Ψ в (6.2) является отображением множества U допустимых стратегий во множество значений показателя эффективности W с учетом (6.1) и обычно задается в форме определенной математической модели операции.

Приведенная формальная запись задачи оценивания эффективности операции задает в наиболее общем виде (6.2) модель операции с оператором выхода W в форме (6.1). Никаких ограничений на характер компонент модель (6.2) не накладывает и поэтому может использоваться как общая исходная основа для моделирования операций произвольной природы, назначения и сложности.

Структурная схема модели (6.2) операции представлена на рис. 1. На схеме приведены основные операторы, с использованием которых множество $U \times \Lambda$ входных воздействий отображается во множество значений показателя эффективности W .

Как следует из схемы, модель операции $\Psi: U \times \Lambda \rightarrow W$ представляет собой составной оператор $\Psi = H \circ Q \circ M$ оценивания эффективности операции, который является суперпозицией операторов H, Q и M . Модель $H: U \times \Lambda \rightarrow Y(G)$ в (6.2) представляет собой одну из форм оператора H моделирования результата (исхода) операции, которая может задаваться в функциональном или алгоритмическом виде и иметь различную математическую структуру. Операторы Q и M отображают множество Y результатов операции во множество значений показателя эффективности W . При этом множество Y с использованием оператора соответствия $Q: Y \times Y^{TP} \rightarrow \rho$ отображается во множество значений функции соответствия ρ ,

а оператор усреднения $M: \rho \rightarrow W$ переводит множество значений функции соответствия ρ во множество значений показателя эффективности W , т. е. $Q \circ M: Y \rightarrow W$

Суперпозиция операторов $Q \circ M$ определяет соответствие принятого показателя эффективности W цели A_0 операции, отображаемой требуемым результатом $Y^{\text{ТР}}$. Конкретный вид оператора $Q \circ M$ зависит от постановки задачи исследования эффективности операции, характера множества Λ факторов и связей между ними, учитываемых при оценивании эффективности.

Как уже отмечалось, операция формируется в рамках S_0 -системы, реализующей эту операцию, и поэтому в общем случае может рассматриваться как процесс функционирования S_0 -системы, т. е. как процесс изменения ее состояний во множестве Z всех возможных состояний S_0 -системы. Тогда операторы Φ и H (см. рис. 1), которые называют операторами перехода и выхода соответственно, могут быть символически записаны в виде соответствий:

$$\Phi: Z \times U \times T \times \Lambda \rightarrow Z;$$

$$H: Z \times U \times T \times \Lambda \rightarrow Y(G).$$

В общем случае, когда конкретные свойства моделируемой операции не заданы, к операторам Φ и H можно предъявить только одно требование: они должны давать возможность определять траектории Z_T и Y_T целиком, а также любые фрагменты этих траекторий. Каждая точка траектории Z_T состояний S_0 -системы и траектории Y_T результата операции характеризует для некоторого момента времени $t \in T$ состояние Z_t системы и значение Y_t результата Y операции. Конкретный вид обеих траекторий определяется множеством значений входных воздействий, реализованных за период функционирования S_0 -системы (все или некоторые составляющие U и Λ входа могут зависеть от времени $t \in T$), начальным состоянием Z_0 системы и операторами Φ и H .

Свойства операторов Φ и H представляют исследователю определенную свободу действия: информацию о входных воздействиях $U \times \Lambda$ он может

использовать для прогнозирования как результата Y операции, так и состояния S_0 -системы (или, например, технической системы, используемой в данной операции), причем (и это очень важно) в любой момент $t \in T$ можно переключиться с решения первой задачи на решение второй или решать их одновременно. Для многих реальных операций с использованием технических систем цель формируется в виде достижения этой системой некоторого требуемого состояния или одного из подмножества требуемых состояний из множества возможных.

Таким образом, структурная схема модели (6.2) операции (см. рис. 1) и ее описание дают общее представление о характере и последовательности действий при построении математической модели операции.

Исследователю, задавшемуся целью построить математическую модель конкретной операции, следует взять модель (6.2) за основу и, руководствуясь описанием структурной схемы (см. рис. 1), конкретизировать ее компоненты в соответствии с поставленной задачей исследования, используя всю имеющуюся в его распоряжении информацию о данной проблемной ситуации.

Однако это лишь общая стратегия действий. В действительности процесс перехода от модели (6.2) к модели конкретной операции, т. е. собственно процесс построения модели операции является очень сложным и трудоемким, особенно, когда речь идет об операциях, связанных с проектированием, созданием и применением больших технических систем.

Исходной информацией для построения математической модели операции служат качественные и количественные данные об элементах модели проблемной ситуации (4.1), сформированные на этапе постановки задачи в результате переработки и преобразования всей имевшейся информации о решаемой проблеме (см. гл. 4). Степень полноты и достоверности этих данных существенно влияет на процесс формализации, в ходе которого осуществляется переход от содержательного описания операции к ее математической модели.

Главное требование к модели — ее адекватность исследуемой операции, иначе невозможно получить положительные результаты моделирования, т. е. оценивание эффективности операции на неадекватной модели вообще теряет смысл. Модель считается *адекватной*, если она с достаточной степенью приближения на уровне понимания моделируемой операции ЛПР и исследователем отражает процесс функционирования S_0 -системы во внешней среде.

В процессе формализации исследователь должен определить, достаточно ли исходных данных об элементах проблемной ситуации (4.1) для построения адекватной модели. Если имеющихся данных недостаточно, то исследователь должен разработать процедуры получения дополнительной информации (см. гл. 11) и прежде всего о тех элементах проблемной ситуации (4.1), которые непосредственно фигурируют в (6.1) и (6.2). Для получения любой информации требуются затраты времени и средств. Поэтому необходимость в дополнительной информации должна каждый раз тщательно обосновываться.

При получении новой информации о моделируемой операции из внешних источников или из результатов моделирования ее модель пересматривается и уточняется с учетом новой информации, т. е. процесс моделирования, включая разработку и машинную реализацию модели, является *итерационным*. Этот итерационный процесс продолжается до тех пор, пока не будет построена такая модель, которую по мнению ЛПР и исследователя можно считать адекватной реальной операции в рамках поставленной задачи исследования эффективности операции (технической системы).

При оценивании эффективности крупномасштабных операций и сложных технических систем их модели рационально строить по блочному принципу. Каждый из этих блоков достаточно автономен, что выражается минимальным числом связей между ними. Поведение этих блоков должно быть хорошо изучено и для каждого из них построена математическая модель, которая в

свою очередь может содержать подблоки.

В прикладных задачах исследования эффективности операций в технике изучаются основные закономерности, связывающие эффективность операций с качеством и условиями функционирования технических систем, способами и условиями их применения в операциях. Поэтому одним из руководящих принципов построения математических моделей таких операций является адекватное отображение в моделях всей совокупности соотношений, описывающих уровень качества и структуру системы, ее поведение, взаимодействие с другими объектами (системами) и со средой и позволяющих изучать ее в динамике функционирования и развития.

Построение математической модели связано с выбором типа модели (аналитическая, статистическая, имитационная) в зависимости от цели и уровня (системного, методологического) исследования эффективности, необходимой точности и достоверности результатов оценивания, отведенного времени на исследование и т. д.

Сформированная математическая модель операции воплощается затем в конкретную машинную модель, разработка которой предполагает алгоритмизацию модели, выбор языка моделирования, программирование модели и проверку достоверности программы. Перед проведением рабочих расчетов на ЭВМ по составленной и отлаженной программе нужно оценить качество модели и при необходимости принять меры по его повышению (см. гл. 9).

Выполнение рабочих расчетов на ЭВМ осуществляется в соответствии с планом проведения эксперимента, в котором должны быть указаны комбинации переменных и параметров, для которых должно проводиться моделирование операции. Планирование машинного эксперимента в конечном итоге направлено на получение максимального объема необходимой информации о моделируемой операции при фиксированных затратах машинных ресурсов. Планирование экспериментов имеет особое значение при

проведении рабочих расчетов с использованием статистических и особенно имитационных моделей (см. гл. 9).

Заключительная процедура этапа оценивания эффективности операции — вычисление с использованием машинной модели значений $W(u)$ показателя эффективности W для всех допустимых стратегий $u \in U$. Этим завершается этап получения результатов.

Рассмотренные положения задают общую схему оценивания эффективности операций (технических систем). В дальнейшем раскрываются наиболее важные составляющие этой схемы, необходимые для решения практических задач исследования эффективности операций с использованием технических систем.

2. КЛАССИФИКАЦИЯ МЕТОДОВ МОДЕЛИРОВАНИЯ

Моделирование является одной из основных методологических концепций, играющей ведущую роль в процессе познания и практического овладения реальной действительностью. При исследовании эффективности сложных технических систем моделирование часто является единственно возможным подходом к их изучению. Концепция моделирования предполагает формирование некоторого условного образа (аналога) реальной системы и изучение его свойств с целью получения информации о реальной системе [65]. Этот образ реальной системы называют *моделью*. Однако по сравнению с реальной системой (оригиналом) модель может иметь совершенно иную природу. Между реальной системой и ее моделью должно быть установлено определенное соответствие (аналогия). От полноты этого соответствия зависит адекватность ¹ модели.

В теории эффективности методы математического моделирования занимают ведущее место. Однако при моделировании операций в технике дело в значительной мере осложняется тем,

что наряду с чисто физическими процессами функционирования разнообразных* технических подсистем, агрегатов, устройств приходится моделировать поведение людей в различных формах их взаимодействия между собой, что вынуждает обращаться к неформальным методам интуитивного моделирования, экспертного оценивания, анализа рефлексий и т. д.

В научной литературе существует большое разнообразие подходов к классификации моделей и методов моделирования [60, 65].

Различают материальное (предметное) и идеальное моделирование (рис. 2). В первом случае в качестве модели предполагается использование некоторого материального предмета. По природе аналогии материальное моделирование делят на физическое (макетирование, обеспечивающее аналогичность физической природы оригинала и модели) и аналоговое (обеспечивающее сходство процессов, протекающих в оригинале и модели). Идеальное моделирование основывается на мысленной идеализированной аналогии реального объекта и его модели, а по способу отражения реального объекта (или по глубине формализации) делится на знаковое (семиотическое) ² и интуитивное моделирование.

По способу представления семиотических моделей различают математическое, логическое и графическое моделирование. Математическое моделирование играет определяющую роль среди других форм знакового моделирования. Однако довольно трудно четко отделить логическое и графическое моделирование от математического ввиду их тесного переплетения (например, в форме логико-математического моделирования и т. п.).

По вычислимости различных показателей, отношений и т. п. методы математического моделирования делят на аналитические и алгоритмические. Аналитическое моделирование предполагает использование математической модели реального объекта в форме алгебраических, дифференциальных, интегральных и других уравнений,

¹ Под *адекватностью* понимают степень соответствия модели тому реальному явлению, для описания которого строится модель.

² Семиотика — теория знаковых систем.

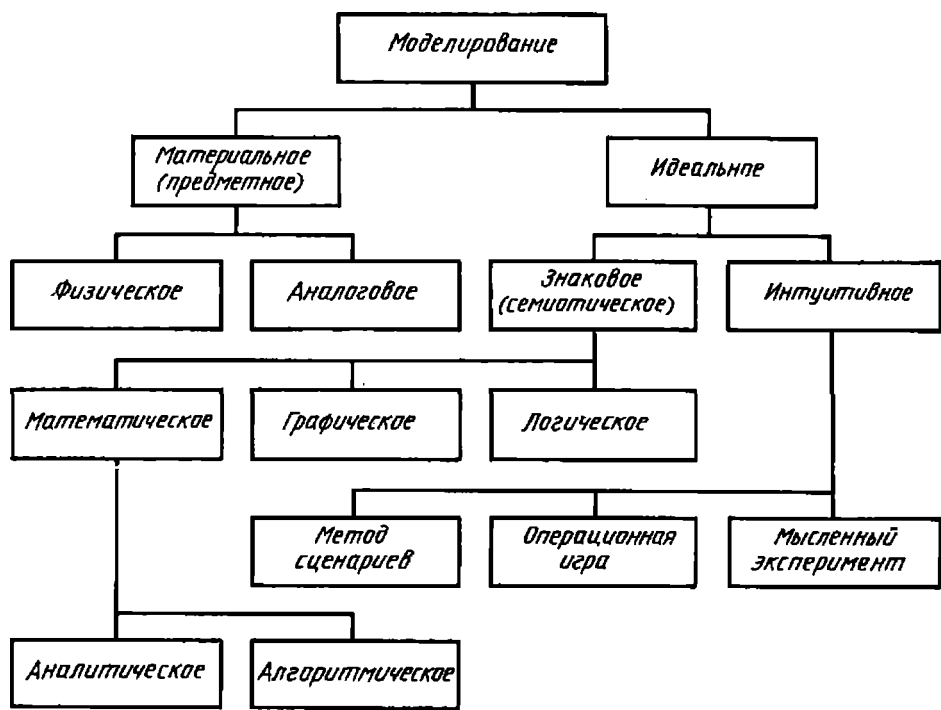


Рис. 2. Классификация методов моделирования

связывающих выходные переменные с входными, дополненных системой ограничений (в виде равенств или неравенств). При этом предполагается наличие однозначной вычислительной процедуры получения точного решения уравнений. При алгоритмическом подходе используемая математическая модель не допускает точного решения и вынуждает обращаться к различным рекуррентным методам, итеративным процедурам поиска приближенного решения. Такой подход при моделировании сложных систем является типичным.

Интуитивное моделирование проводится на вербальном (описательном) уровне. При этом методе не устанавливаются строгие количественные соотношения между моделируемыми явлениями, ограничиваясь лишь анализом качественных обобщенных понятий, отражающих общие тенденции развития явлений, направления изменения свойств изучаемых объектов и т. п. Такой подход осуществляется с целью выдвижения различного рода гипотез поведения субъектов сложных систем, формирования эвристик относительно взаимоотношений между активными элементами системы и их развития.

По способу формирования эвристик различают следующие формы интуитивного моделирования: метод сценариев, операционные игры, мысленный эксперимент и др. При исследовании эффективности технических систем мысленный эксперимент и операционные игры нашли широкое распространение. Метод сценариев менее распространен в практике. Однако в перспективе, по-видимому, он займет ведущее место среди форм интуитивного моделирования (при соответствующей методологической доработке). Интуитивное моделирование является основным методом моделирования деятельности метасистемы. Аналитическое моделирование не ограничивается исследованием достаточно простых систем. Концептуальные модели метасистемы часто представляют в аналитическом виде, учитывая лишь некоторые существенные связи и отношения между моделируемыми явлениями.

Основной формой системного анализа эффективности сложных технических систем является имитационное исследование, проводимое в рамках имитационных моделей. Обычно имитация реализуется на ЭВМ. При этом стремятся к наиболее полному учету всех существенных факторов. Собы-

тия при имитации разворачиваются во времени, как правило, в том порядке, в каком они следуют в реальной системе, но в измененной временной шкале. Действие случайных факторов учитывается с помощью специальных датчиков случайных чисел (имитаторов), настроенных на соответствующие вероятностные распределения. В определенном месте процесс имитации на ЭВМ может быть приостановлен для проведения, например, операционной игры, экспертного опроса или натурного эксперимента с использованием промежуточных данных, полученных при машинной имитации. Результаты игры, экспертизы или эксперимента могут быть использованы для продолжения имитации на ЭВМ.

Имитационное моделирование, являясь своеобразным экспериментированием с моделью реальной системы, открывает широкие возможности адекватного отображения исследуемых явлений. Оно синтезирует знаковое и интуитивное (иногда и материальное) моделирование. Таким образом, в имитационном моделировании используются практически все методы моделирования, отображенные на рис. 2.

Имитационное моделирование объединяет имитацию исследуемого явления и планирование имитационного эксперимента. Теория планирования эксперимента позволяет организовать имитационный эксперимент рациональным применительно к целенаправленному получению информации образом.

Имитационное моделирование обычно противопоставляют методам математического программирования. При исследовании операций типичным является следующий подход. Путем различного рода допущений реальную задачу идентифицируют с одной из стандартного класса задач математического программирования (линейного программирования, динамического программирования и т. п.). Для каждого класса задач разработаны алгоритмы решения, воспользовавшись которыми исследователь (при интересующих его исходных данных) может получить решение своей задачи. Этот путь решения оптимизационных задач хорошо зарекомендовал себя в прак-

тике исследования систем на уровне «структура—функция». Однако такой подход не всегда приводит к успеху при исследовании сложных систем уровня «организация—поведение». Упрощенное описание сложной системы, выполненное с целью сведения ее к одному из классов моделей математического программирования, часто приводит к утрате адекватности модели и, как следствие этого, к обесцениванию получаемых результатов.

При имитационном моделировании сложных систем во главу угла ставится требование адекватности модели. При имитационном подходе стремятся получить «портретную» модель сложной системы. Однако при высокой сложности реальной системы ее адекватная имитационная модель также становится в значительной мере сложной. Одна реализация имитируемого явления (прогон модели) может быть весьма продолжительной, а по одной реализации сделать какие-либо содержательные выводы о свойствах системы, как правило, не удастся. Проведение статистических имитаций в рамках модели при различных исходных данных может оказаться неприемлемым по времени. В этих условиях эффективными являются методы планирования эксперимента.

Имитационный эксперимент можно планировать либо с целью получения оптимальных значений факторов (задача оптимизации), либо с целью получения аналитической зависимости выходной переменной (отклика) от контролируемых переменных (факторов) в достаточно широкой области изменения последних (задача аппроксимации). Важным достоинством имитационного моделирования является возможность фиксации промежуточных значений различных показателей в процессе имитации подобно тому, как это имеет место при съеме данных в процессе функционирования реальной системы.

3. ПРИНЦИПЫ МОДЕЛИРОВАНИЯ ОПЕРАЦИЙ В ТЕХНИКЕ

Моделирование операции в технике — сложный, трудоемкий процесс, имеющий ярко выраженный творче-

ский характер. Поэтому дать исследователю пригодные во всех случаях готовые рекомендации по моделированию невозможно. Однако существуют определенные выработанные практикой принципы моделирования, придерживаясь которых осуществляют построение соответствующих моделей.

Принцип информационной достаточности. При полном отсутствии информации об исследуемой системе построение ее модели невозможно. При наличии полной информации о системе ее моделирование лишено смысла. Существует некоторый критический уровень априорных сведений о системе (уровень информационной достаточности), при достижении которого в принципе можно получить адекватную модель системы. Этот уровень определяется наличием информации об элементах системы и о существенных связях между ними, которые формируют эмерджентные свойства системы. При отсутствии этих сведений моделирование не приведет к получению новой информации о системе, т. е. окажется бесплодным. В понятие информационной достаточности входят также сведения о полезности системы в деятельности метасистемы (внешнее дополнение). Без этих сведений моделирование лишено целенаправленности.

Принцип параметризации. Некоторые относительно изолированные подсистемы характеризуются определенным параметром (возможно векторным). Значения этого параметра могут зависеть от ситуаций, складывающихся в исследуемой системе. Такие подсистемы в модели можно заменять соответствующим параметром, а не описывать процесс их функционирования. Зависимость параметра от ситуаций в общей модели обычно задают в виде таблицы, графика или аналитического выражения (формулы). Принцип параметризации позволяет иногда в значительной мере сократить объем и продолжительность моделирования исследуемой системы. Однако надо иметь в виду, что параметризация снижает адекватность модели и должна употребляться с известной осторожностью.

Принцип агрегирования В большинстве случаев сложную систему можно структурно представить состоящей из агрегатов (подсистем), для адекватного математического описания которых оказываются пригодными некоторые стандартные математические схемы. Достаточно универсальной математической схемой для моделирования агрегатов сложных технических систем явилась общая динамическая система (ОДС), реализующая процесс с дискретным вмешательством случая [9]. В качестве частных случаев ОДС охватывает модели, описываемые системой дифференциальных уравнений, конечные и вероятностные автоматы, системы массового обслуживания и т. п. В единую имитационную модель агрегаты объединяются с помощью оператора сопряжения. Принцип агрегирования позволяет в рамках единой имитационной системы достаточно гибко перестраивать имитационную модель в зависимости от задач исследования.

Принцип осуществимости. Создаваемая модель должна обеспечивать достижение поставленной цели исследования с вероятностью, существенно отличающейся от нуля, и за конечное время. Обычно задают некоторое пороговое значение P_0 вероятности достижения цели моделирования $P_{ц}(\tau)$, а также приемлемую границу τ_0 времени достижения этой цели τ . Считают модель осуществимой, если одновременно выполнены неравенства $P_{ц}(\tau) \geq P_0$ и $\tau \leq \tau_0$. При нарушении хотя бы одного из этих неравенств модель неосуществима.

Принцип рационального использования факторного пространства является основным в статистической теории планирования экспериментов [45]. Имитационное моделирование предполагает планирование активного машинного эксперимента. В пространстве управляемых переменных (в факторном пространстве) точки, отображающие условия экспериментов, следует выбирать в соответствии с определенным оптимальным планом эксперимента. При фиксированном числе опытов (реализаций) оптимальный план приводит к наибольшему объему информации, полученной из эксперимента.

В зависимости от цели исследований, априорной информации и других обстоятельств используют планы ортогональные, ротатабельные, D -оптимальные, Q -оптимальные и т. д. [66]. Принцип рационального использования факторного пространства позволяет в значительной мере экономить время расчетов на моделях.

Принцип множественности моделей. Любая модель реальной системы учитывает не все факторы, действующие в системе. Модель, как средство отражения реального явления, как бы усиливает действия одних факторов, ослабляя действия других. При использовании любой конкретной модели познаются лишь некоторые стороны реальности. Для более полного отражения действительности необходим ряд моделей («веер» моделей), позволяющий с разных сторон и с разной степенью детальности (в зависимости от цели исследования) отражать изучаемое явление. Принцип множественности моделей может быть рационально реализован в имитационной системе, которая позволяет достаточно гибко подходить к формированию моделей операций, выполняемых исследуемой системой в различных ситуациях.

4. КЛАССИФИКАЦИЯ МОДЕЛЕЙ ОЦЕНИВАНИЯ ЭФФЕКТИВНОСТИ

Приведенная в параграфе 2 классификация методов моделирования далеко не исчерпывает все возможные классификационные признаки и, следовательно, не может претендовать на полноту и завершенность. Однако в практике исследований эффективности операций такая классификация является достаточной и может служить основой для классификации моделей операций. Классификационных признаков для моделей гораздо больше, чем для методов моделирования. Рассмотрим частичную схему классификации моделей операций, которые используют при оценивании эффективности операций в технике.

На первое место среди классификационных признаков моделей можно поставить метод (способ) моделирова-

ния. По этому признаку почти каждому методу (см. рис. 2) соответствует определенный класс моделей, имеющий то же название, что и метод моделирования. Так, знаковому (семиотическому) моделированию соответствуют знаковые (семиотические) модели; математическому моделированию — математические модели, которые могут быть аналитическими или алгоритмическими. В основе деления семиотических моделей лежит система знаков, в которой записываются операторы Φ , H , Q и M . Модель называют математической, если указанные операторы записаны в математической символике. Если же для формирования операторов употреблена логическая символика, то модель будет логической. Такая модель может быть материализована в виде программы логических действий для ЭВМ или в виде блок-схемы последовательности логических операций. Знаковую модель называют графической, если операторы Φ и H , Q и M заданы графически (например, в форме номограмм, графиков).

На практике часто встречаются всевозможные комбинации различных знаковых (семиотических) моделей в виде, например, графоаналитических или логико-математических моделей. Программы для ЭВМ реализуют, как правило, логико-математическую модель.

Важным классификационным признаком моделей является мощность множества состояний моделируемой системы. По этому признаку модели делят на статические и динамические. Модель называется статической, если множество возможных состояний моделируемой системы является единственным, т. е. содержит всего один элемент. Система не меняет своего состояния в течение времени и, следовательно, оператор перехода Φ при формировании модели не рассматривается. Если множество состояний системы неединично и состояния могут меняться во времени, то соответствующую модель называют динамической. Динамические модели в зависимости от мощности множества состояний Z делят на модели с дискретными состояниями (множество состояний конечно

или счетно, т. е. не более, чем счетно) и модели с непрерывным множеством состояний (множество состояний ставится во взаимно-однозначное соответствие множеству точек числового интервала, имеющего мощность континуума). Не исключаются смешанные случаи дискретно-непрерывных множеств состояний.

Процесс смены состояний называют движением системы. Смена состояний системы может происходить в фиксированные моменты времени, множество которых дискретно и заранее определено. Такой процесс смены состояний называют последовательностью. Последовательности задают обычно разностными уравнениями. Если множество моментов смены состояний системы непрерывно в пределах некоторого временного интервала, то процесс движения такой системы описывают обычно дифференциальными уравнениями. При этом различают детерминированные модели (описывающие лишь причинно-следственные связи между различными факторами) и вероятностные модели. Детерминированные модели описывают процесс смены состояний в виде зависимости вектора состояний Z от времени и других факторов. Вероятностные модели отражают процесс движения системы в виде зависимости условной вероятности перехода в то или иное состояние системы от предшествующих состояний и других факторов. Вероятностная модель может быть записана не только для условной вероятности состояний, но и для других характеристик распределения состояний (например, вероятностные модели с непрерывным временем обычно записывают в виде уравнений относительно интенсивностей переходов в различные состояния). Вероятностные модели классифицируют обычно как случайные процессы, разделяя их на процессы с дискретным временем (случайные последовательности) и с непрерывным временем.

Понятие динамической модели (динамической системы) имеет еще одну содержательную интерпретацию. По характеру предсказаний различают стохастические и динамические закономерности. Закономерности дина-

мического типа позволяют точно, однозначно предсказать движение системы, имея информацию о начальном ее состоянии. Закономерности стохастические дают возможность предсказать дальнейшее движение системы лишь с некоторой вероятностью. В связи с этим по характеру предсказаний модели делят на динамические и стохастические. Первые употребляются в практике для описания реальных явлений с позиции жесткого детерминизма, а вторые — для описания явлений с учетом воздействия случайных факторов. Таким образом, динамическая модель отражает динамические закономерности движения системы, т. е. описывает процессы изменения состояния системы под воздействием только причинно-следственных (казуальных) связей.

В классе стохастических моделей наряду с вероятностными выделяют также статистические модели. Статистическая модель формируется по схеме эксперимента со случайным исходом. Операторы Φ и H включают в свой состав случайные факторы Λ_E . При этом случайное воздействие может подаваться на вход модели либо формироваться с помощью датчиков случайных переменных (имитаторов) в рамках самой модели. Оператор H на выходе статистической модели формирует случайную переменную Y . В статистических моделях оценивания эффективности оператор $\Psi = H \circ Q \circ M$ содержит алгоритм статистической обработки. Обычно по виду такого алгоритма статистические модели делят на модели регрессионного, дисперсионного или корреляционного анализа, параметрического или непараметрического оценивания и т. д. Кроме того, оператор Ψ может содержать нечетко заданные переменные (соотношения), которые могут быть определены своими функциями принадлежности. Соответствующие модели оценивания эффективности называют моделями с нечеткими переменными (отношениями) или моделями нечетких систем.

По виду оператора соответствия Q и, следовательно, по типу формируемого показателя эффективности W модели оценивания эффективности на-

зывают моделями среднего результата, гарантированного результата, вероятностной гарантии и т. д.

К специфическим особенностям моделей оценивания эффективности операций в технике относится форма представления результата операции Y и выбор оператора выхода H . Выходной оператор формирует результат операции. От выбора способа описания результата операции зависит качество оценивания эффективности операции. Переменная Y , описывающая результат, должна быть логически сопоставима с переменной Y^{TP} , определяющей требуемый результат (например, они должны быть соизмеримы, сравнимы и т. п.). В формализованном виде результат обычно задается в скалярной или векторной форме в виде числовых или нечисловых переменных, функций, функционалов и т. п. В содержательном плане результат отражает направленность исследования, определяется его целью и зависит от предметной области, в которой проводится оценивание эффективности. Содержание предметной области и вводимого результата операции часто дают основу для содержательной классификации моделей оценивания эффективности. С этой точки зрения различают следующие установившиеся виды моделей: модели технико-экономического

анализа, экономико-математические модели, строгого и нестрогого конфликта и т. п. По этому же принципу выделяют модели систем массового обслуживания, модели систем управления запасами и т. д.

Деление моделей на аналитические и имитационные является в известной степени искусственным и обусловлено способом представления оператора Ψ . Аналитическая модель оценивания эффективности обычно представляет собой уравнение, связывающее показатель эффективности W с различными существенными факторами. При этом полагают, что это уравнение, записанное в форме аналитического выражения, допускает точное решение (без использования приближенных численных методов). Отличительной особенностью имитационной модели является обычно сложность представления операторов Φ , H , Q и M . Операторы Φ и H реализуются, как правило, в алгоритмической форме на ЭВМ. При этом сам процесс моделирования может быть назван машинным экспериментом.

Имитационная модель оценивания эффективности позволяет достаточно полно описать результат операции и, следовательно, дает возможность всесторонне подойти к анализу эффективности операций.

Глава 7. Математическое моделирование операций

1. ОБЩАЯ СХЕМА ОПЕРАЦИИ

Моделирование операции является средством получения информации, необходимой для выработки решения относительно организации рационального поведения системы для достижения цели операции. Модель должна адекватно отображать исследуемую операцию в информационном аспекте, т. е. с точки зрения информационных связей между элементами системы, отражающими потоки информации внутри системы между ее элементами, а также между системой и внешней средой. Целесообразно поэтому в качестве модели операции использовать так называемую кибернетическую систему.

Под *кибернетической системой* (КС) понимают совокупность связанных между собой объектов любой природы (элементов системы), способных воспринимать, хранить, перерабатывать информацию, а также обмениваться информацией между собой и с внешней средой [43]. Материально-энергетические потоки в такой системе рассматриваются лишь в качестве носителей информации. Кибернетическая система является наиболее общей абстрактной моделью S_0 -системы, реализующей операцию. Поэтому общая схема операции может быть описана как кибернетическая система, а изучение последней направлено на выявление общих закономерностей, присущих различным конкретным операциям.

Кибернетическая система допускает следующее формальное описание. Каждый элемент ее может менять свои состояния с течением времени. Процесс функционирования элемента мо-

жет быть детерминированным или стохастическим. Связи между элементами являются носителями информационных потоков. Система связей задается оператором сопряжения R , который определяет структуру кибернетической системы. В общем случае структура системы может измениться в процессе функционирования системы под воздействием внешних и внутренних факторов.

Полное описание кибернетической системы как обобщенной модели операции включает: 1) семейство функций, определяющих изменения состояний всех элементов КС; 2) семейство функций, определяющих все входные сигналы элементов (от внешней среды и от других элементов в соответствии со схемой сопряжения); 3) семейство функций, описывающих изменения в структуре системы; 4) начальные состояния всех элементов КС; 5) начальную структуру системы.

Если хотя бы часть из перечисленных функций, входящих в описание КС, случайна, то кибернетическую систему называют вероятностной или стохастической. В противном случае система детерминирована. Очевидно, детерминированную систему можно рассматривать как частный вырожденный случай вероятностной системы.

Кибернетическая система, являясь моделью S_0 -системы, вводится для информационного (кибернетического) описания операции. Такое описание осуществляется с точки зрения только одного субъекта системы, а именно, с точки зрения ЛПР, его интересов и представлений об обстановке операции и целях других субъектов системы (их интересов, отражаемых соответствующими гипотезами поведе-

ния). В этом смысле информационное описание S_0 -системы является субъективным.

Кибернетическая система, включающая в свой состав несколько субъектов системы с несовпадающими интересами, называется *поведенчески неоднозначной системой*¹. Неоднозначность результата операции, реализуемой такой системой, обусловлена активными действиями наряду с ЛПР других субъектов системы, их стремлением к достижению своих (отличных от ЛПР) целей. Их интересы, стремления могут быть неполностью известны ЛПР. Поэтому выдвигаемые ЛПР гипотезы их поведения могут иметь разную степень достоверности.

Кибернетическая система управляема. Она включает в свой состав управляющие и управляемые подсистемы (агрегаты, элементы). Управляющая подсистема (система управления) содержит орган управления (распорядительный центр, распорядитель) с техническими средствами управления и систему связи для передачи команд управления исполнительным органам (управляемой подсистеме) и приема контрольной информации. S_0 -система может иметь несколько различных распорядительных центров с подчиненными им управляемыми подсистемами. В один из этих распорядительных центров (РЦ) входит ЛПР. Этот РЦ называют *основным*.

2. МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ ОПЕРАЦИИ

Состояние S_0 -системы определяется состояниями всех входящих в нее элементов. Функционирование системы есть процесс изменения ее состояний.

Пусть z_t — состояние S_0 -системы в момент времени t . При этом $z_t \in Z$, где Z — множество всех возможных состояний системы (пространство состояний). В общем случае z_t есть n -мерный вектор $z_t = \|z_{1t}, z_{2t}, \dots, z_{nt}\|$. Обозначим z^t траекторию процесса смены состояний до момента t , т. е.

при дискретном времени ($t = 0, 1, 2, \dots$)

$$z^t = \{z_0, z_1, z_2, \dots, z_t\}. \quad (7.1)$$

Состояние z_t S_0 -системы к моменту t в общем случае зависит от траектории процесса до момента t , т. е. от истории процесса:

$$z_t = \{z^l, l \leq t\}.$$

Кроме того, состояние z_t зависит от всей совокупности управляющих воздействий на управляемые подсистемы, агрегаты, элементы S_0 -системы со стороны ЛПР и других субъектов системы к моменту времени t , а также от воздействий внешних возмущающих факторов, часть из которых Λ_E имеет случайный характер, а другая часть может быть отнесена к природным и поведенческим неопределенностям $\Lambda_J = \Lambda_J^N \cup \Lambda_J^B$.

Таким образом, вход X S_0 -системы объединяет несколько составляющих:

- 1) стратегии ЛПР $u^t \in U$, где U — множество его допустимых стратегий;
- 2) стратегии других субъектов системы $v^t \in V$, где V — множество их допустимых стратегий ($V = \Lambda_J^B$);
- 3) случайные возмущения Λ_E ;
- 4) возмущающие воздействия нестохастического характера Λ_J^N , имеющие природную неопределенность.

Все перечисленные составляющие входа (или часть из них) могут зависеть от времени $t \in T$, где T — множество моментов времени. В этом случае $x(t) \in X$, где X — множество всевозможных вектор-функций входа (функциональное пространство входных переменных).

Теперь имеем

$$z_t = \varphi_z(z^{t-1}, u^{t-1}, v^{t-1}, \lambda_E, \lambda_J^N), \quad (7.2)$$

где z^{t-1} — траектория процесса смены состояний до момента t , ($z^{t-1} = \{z_0, z_1, z_2, \dots, z_{t-1}\}$); u^{t-1}, v^{t-1} — совокупность управлений (стратегии) ЛПР и других субъектов системы до момента t , ($u^{t-1} = \{u_0, u_1, \dots, u_{t-1}\}$, $v^{t-1} =$

¹ Поведенчески неоднозначную систему называют еще нерефлексной [43].

$= \{v_0, v_1, v_2, \dots, v_{t-1}\}$; λ_E, λ_J^N — значения возмущающих факторов.

Символическая запись оператора (7.2) удобна в виде соответствия

$$\Phi: Z \times U \times V \times T \times \Lambda \rightarrow Z, \quad (7.3)$$

где $\Lambda = \Lambda_F \cup \Lambda_E \cup \Lambda_J^N$.

Процесс z^t с учетом случайных возмущений Λ_E может быть случайным управляемым процессом $\hat{z}^t$. Для его описания вводится измеримое пространство $\langle Z, \Sigma_Z \rangle$, где Σ_Z — σ -алгебра на множестве Z . На этом пространстве задается система условных вероятностных мер:

$$\begin{aligned} \{\mu_t(\sum z/z^{t-1}, u^{t-1}, v^{t-1}) = \\ = P(\hat{z}_t \in \sum z/z^{t-1}, u^{t-1}, v^{t-1}); \\ v^t \in T\}. \end{aligned} \quad (7.4)$$

Эта система мер определяет случайный управляемый процесс с дискретным временем (случайную управляемую последовательность).

Процессы с непрерывным временем обычно описывают дифференциальными уравнениями

$$\dot{z} = \Phi_z(z, u, v, \lambda_E, \lambda_J^N). \quad (7.5)$$

Управляемый дискретный процесс можно представить как некоторую многошаговую (поэтапную) процедуру принятия решений ЛПР (выбор стратегий u_t ; $t = 0, 1, \dots$) и другими субъектами системы. В частном случае эта процедура может быть одношаговой, т. е. выбор стратегии может быть однократным. При поэтапном выборе стратегий на очередном шаге ЛПР выбирает стратегию, придерживаясь определенного правила рационального поведения, в зависимости от состояний системы и управлений на предыдущих шагах процесса, т. е.

$$u_t = \Phi_V(z^t, u^{t-1}, v^{t-1}). \quad (7.6)$$

Иногда приходится использовать рандомизированные стратегии. В этом случае вводится некоторый искусственный механизм случайного выбора управлений (чистых стратегий). Рандомизированную стратегию определяют следующим образом. Пусть $\langle U, \Sigma_U \rangle$ —

измеримое пространство с σ -алгеброй Σ_U . Условная вероятностная мера

$$\begin{aligned} v_t(\sum u/z^t, u^{t-1}, v^{t-1}) = \\ = P(u_t \in \sum u/z^t, u^{t-1}, v^{t-1}). \end{aligned} \quad (7.7)$$

Совокупность условных мер v_t ($t = 1, 2, \dots$) называют *рандомизированной стратегией*. Стратегии других субъектов системы (распорядительных центров) также могут быть рандомизированными. С точки зрения ЛПР выбор стратегий другими субъектами системы может представлять собой неопределенность нестохастического характера (поведенческая неопределенность). Однако в этих случаях иногда не вполне ясные представления о действиях других субъектов ЛПР может выразить с помощью функций принадлежности (в смысле Л. А. Заде) выбираемых стратегий к тому или иному подмножеству допустимых стратегий U [24].

Взаимоотношения между субъектами системы (включая ЛПР) определяют первичную структуру S_0 -системы. Субъекты между собой могут находиться в отношении строгого подчинения, образуя иерархическую централизованную структуру, нестрогого подчинения, отражаемого ромбовидной или матричной структурой, а также могут, не находясь в прямом подчинении, стимулировать желательные действия для себя со стороны другого субъекта путем поощрительных мероприятий, рефлексивного принуждения и т. д. Кроме того, субъекты системы могут между собой конфликтовать. Конфликт может принять различные формы от нестрогого соперничества до антагонизма. Рассмотрим иерархическую структуру отношений между субъектами системы, в которой на высшем иерархическом уровне находится распорядитель надсистемы, осуществляющий руководство S_0 -системой. Набор его допустимых стратегий (или политик) обозначим V^0 . ЛПР находится на первом иерархическом уровне и может выбирать стратегии u из множества допустимых стратегий U . На этом же уровне могут находиться и другие субъекты S_0 -системы, наборы допустимых стра-

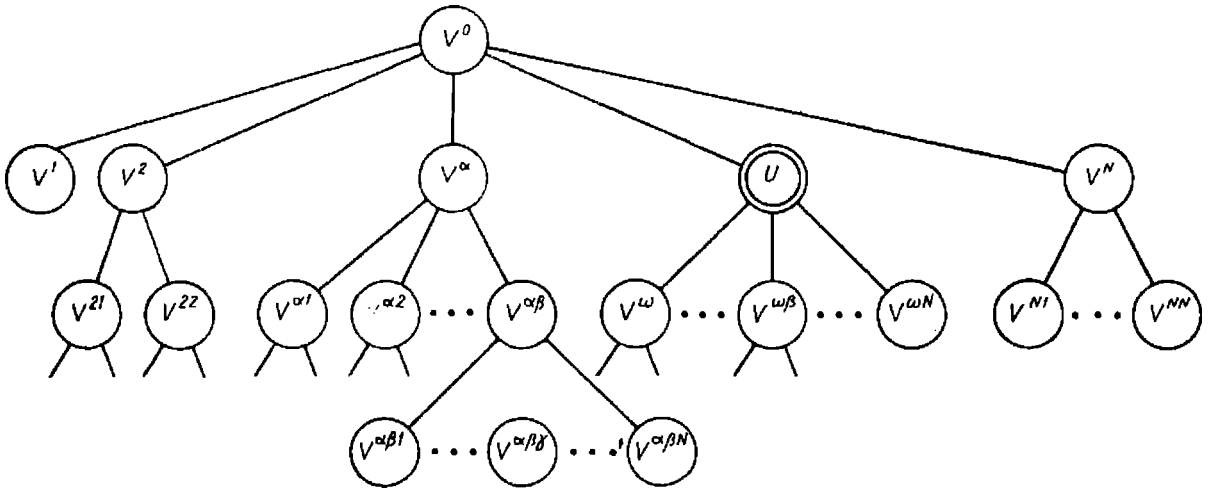


Рис. 1. Структура отношений субъектов системы

тегий которых обозначим V^α ($\alpha = 1, 2, \dots, N$). Подчиненные распорядителю α субъекты второго уровня иерархии имеют наборы допустимых стратегий $V^{\alpha\beta}$ ($\beta = 1, 2, \dots, N$) и т. д. Этими же индексами обозначим и подсистемы S_0 -системы, которыми управляют соответствующие распорядители. Так, $S_2^{\alpha\beta}$ означает подсистему, которой управляет распорядитель с набором допустимых стратегий $V^{\alpha\beta}$. На рис. 1 показана иерархическая структура отношений между субъектами системы. Внутри кружков показаны соответствующие наборы допустимых стратегий субъектов (распорядителей разных уровней иерархии). Таким образом, класс множеств допустимых стратегий V субъектов системы (исключая ЛПР) представляет собой совокупность множеств $V^\alpha, V^{\alpha\beta}, V^{\alpha\beta\gamma}, \dots$, ($\forall \alpha \in I_\alpha; \forall \beta \in I_\beta; \forall \gamma \in I_\gamma$), где I_α — множество натуральных чисел $\{1, 2, \dots, N\}$; $I_\beta = \{1, 2, \dots, N\}$; $I_\gamma = \{1, 2, \dots, N\}$ и т. д.

Высшему уровню иерархии присваивается индекс 0. I_α — множество индексов первого уровня. Индексы второго уровня есть пары из декартового (прямого) произведения множеств $I_\alpha \times I_\beta$. Индексы третьего уровня есть упорядоченные тройки из декартового произведения множеств $I_\alpha \times I_\beta \times I_\gamma$ и т. д.

Выход S_0 -системы отражает достигнутый к моменту времени t результат операции y^t . В общем случае y^t есть

управляемый случайный процесс $\{y_0, y_1, \dots, y_t\}$ с дискретным или непрерывным временем. Часто результат операции y^t вводят как функционал от выходного процесса. Выход S_0 -системы к моменту t зависит от процесса z^t смены состояний системы до момента t , от стратегий u^{t-1} ЛПР и других субъектов системы V^{t-1} , а также от случайных возмущений Λ_E и возмущений нестохастического характера Λ_J^N . Оператор выхода в общем виде можно записать следующим образом:

$$y^t = h(z^t, u^{t-1}, v^{t-1}, \lambda_E, \lambda_J^N). \quad (7.8)$$

Пусть Y — пространство выходной переменной. Оператор выхода удобно записать в символической форме в виде соответствия

$$H : Z \times U \times V \times T \times \Lambda \rightarrow Y \quad (7.9)$$

В сокращенной записи результат операции (реальный, фактический результат) обозначим $y(u, v)$, т. е. в виде зависимости y от стратегии ЛПР и других распорядителей системы (или $y(u)$, если последние отсутствуют).

При наличии случайных факторов Λ_E выход y^t может быть случайной переменной (случайным событием, случайной величиной или случайным процессом). Распределение выходной случайной переменной задается следующим образом. Вводится измеримое

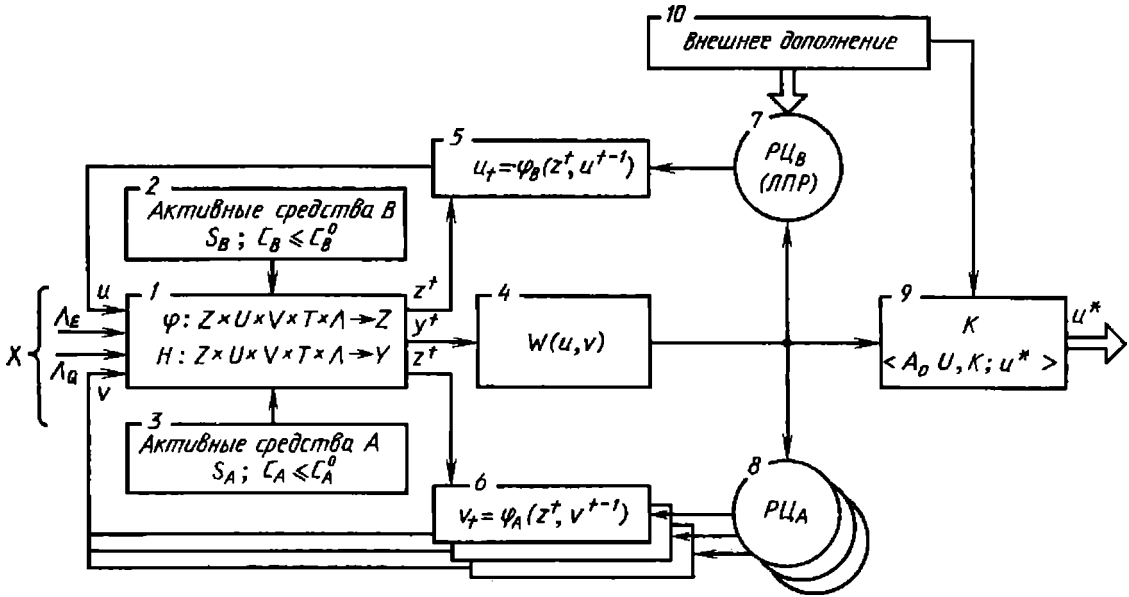


Рис. 2. Общая схема операции

пространство $\langle Y, \Sigma_Y \rangle$, где Σ_Y — σ -алгебра на Y . Условная вероятностная мера на этом пространстве определяется соотношением

$$\begin{aligned} \kappa_t(\Sigma_Y/z^t, u^{t-1}, v^{t-1}) &= \\ &= P(y_t \in \Sigma_Y/z^t, u^{t-1}, v^{t-1}). \end{aligned} \tag{7.10}$$

Совокупность условных мер κ_t ($t = 0, 1, 2, \dots$) образует выходной случайный управляемый процесс. Запись $y(u, v)$ означает в этом случае, что распределение случайной переменной y зависит от стратегий u и v .

На рис. 2 показана общая схема операции. Блок 1 отображает операторы перехода φ (7.3) и выхода H (7.9). Активные средства ЛПР отражены на схеме блоком 2. В формализованном виде активные средства представлены техническими характеристиками системы, которой управляет непосредственно ЛПР, и совокупностью ресурсов различного вида, используемых ЛПР в операции. Обычно на каждый вид ресурсов накладываются ограничения в виде неравенств

$$C_i^u \leq C_{i0}^u \quad (i = 1, 2, \dots, m),$$

где i — номер вида ресурсов; C_{i0}^u — запас ресурсов i -го вида.

Активные средства других субъектов системы отображены совокуп-

ностью блоков 3. Для них также могут существовать покомпонентные ограничения на используемые в операции ресурсы $C_i^{\alpha} \leq C_{i0}^{\alpha}$; $C_i^{\alpha\beta} \leq C_{i0}^{\alpha\beta}$ и т. д.

В блоке 4 формируется показатель эффективности операции. С этой целью на блок 4 из блока 1 подается значение результата операции $y(u, v)$, которое сопоставляется с требуемым результатом с помощью функции соответствия $\rho(y(u, v), y^{\text{TP}})$. Среднее значение (математическое ожидание) функции соответствия обычно принимается в качестве показателя эффективности $W(u, v)$ операции.

Стратегии ЛПР и других субъектов системы вырабатываются соответственно в блоке 5 и в группе блоков 6. При многошаговом процессе принятия решений из блока 1 в эти блоки подаются значения z^t процесса к моменту t . Блок 7 отображает основной распорядительный центр (ЛПР). Остальные распорядительные центры отражены серией блоков 8. Блок 9 отображает критерий эффективности K . В этом блоке формируется задача принятия решения $\langle A_0, U, K; u^* \rangle$, которая формулируется следующим образом: выбрать рациональную стратегию u^* из множества допустимых стратегий U , используя критерий K при сформулированной цели опера-

ции A_0 и предполагаемых действиях других субъектов системы, отражаемых их стратегиями $v \in V$.

Блок 10 формирует внешнее дополнение для задачи принятия решения. Этот блок не входит непосредственно в S_0 -систему. Однако в общей схеме операции внешнее дополнение необходимо для формирования цели операции в виде требуемого результата, установления множеств допустимых стратегий U , покомпонентных ограничений на ресурсы для ЛПР и других субъектов системы, выдвижения гипотез поведения последних, а также для обоснованного выбора критерия эффективности операции.

В общем случае можно допустить, что каждый субъект системы пользуется своим показателем эффективности, отражающим его интересы в операции. Обозначим показатели эффективности субъектов системы теми же индексами, что и множества их допустимых стратегий. Например, $W^{\alpha\beta}(u, v)$ есть показатель эффективности распорядителя, имеющего набор допустимых стратегий $V^{\alpha\beta}$. Показатель эффективности для ЛПР в этой общей схеме обозначим $W^{\omega}(u, v)$. Набор частных показателей эффективности операции задают обычно в виде вектора

$$(W^{\omega}(u, v), W^{\alpha}(u, v), W^{\alpha\beta}(u, v), \dots, W^{\alpha\beta\gamma}(u, v))^T \quad (7.11)$$

В наиболее сложных случаях компоненты вектора (7.11) сами могут быть векторами, т. е. частные показатели могут быть векторами, и тогда (7.11) записывают в виде матрицы.

Критерий эффективности K , которым пользуется ЛПР, формируется на основе определенных принципов принятия решений с учетом стремлений других субъектов системы к достижению своих целей в операции. Эти принципы принятия решений при использовании векторного показателя эффективности рассматриваются в части III.

3. ЧАСТНЫЕ ОПИСАНИЯ ОПЕРАЦИЙ

Рассмотренная общая схема операции позволяет подойти к исследованию большого класса конкретных операций как к частным случаям этой общей схемы.

1. Пусть S_0 -система имеет только один распорядительный центр — ЛПР. В этом случае блоки 3, 6, 8 в общей схеме операции (см. рис. 2) отсутствуют. Факторы поведенческой неопределенности в такой системе не имеют места. Допустим, кроме того, что операция проводится за один шаг, а случайные факторы Λ_F и факторы природной неопределенности не учитываются. При этом вход X S_0 -системы определяется стратегией $u \in U$ ЛПР. Функция соответствия $\rho(y(u), y^{\text{ТР}})$ совпадает с показателем эффективности $W(u)$ и, следовательно, существует функциональное соответствие между множеством допустимых стратегий U и множеством значений W показателя эффективности $W(u)$, т. е.

$$H: U \times \Lambda_F \rightarrow W. \quad (7.12)$$

Множество допустимых стратегий $U = \{u : \eta(u) \leq C\}$ объединяет стратегии, удовлетворяющие ограничениям $\eta(u) \leq C$, где $\eta(u)$ и C — векторы одинаковой размерности. Схема рассматриваемой операции изображена на рис. 3.

В этих условиях обычно ставится следующая задача: найти $u^* \in U$, при которой показатель эффективности достигает экстремального значения, т. е.

$$u^* : \text{extr}_{u \in U} W(u). \quad (7.13)$$

Методы решения этой задачи изложены в обширной литературе по исследованию операций. При этом достаточно глубоко изучены частные постановки задач математического программирования и найдены эффективные алгоритмы их решения. Задачи математического программирования классифицируются в зависимости от вида функции $W(u)$, ограничивающих неравенств $\eta(u) \leq C$ и ряда

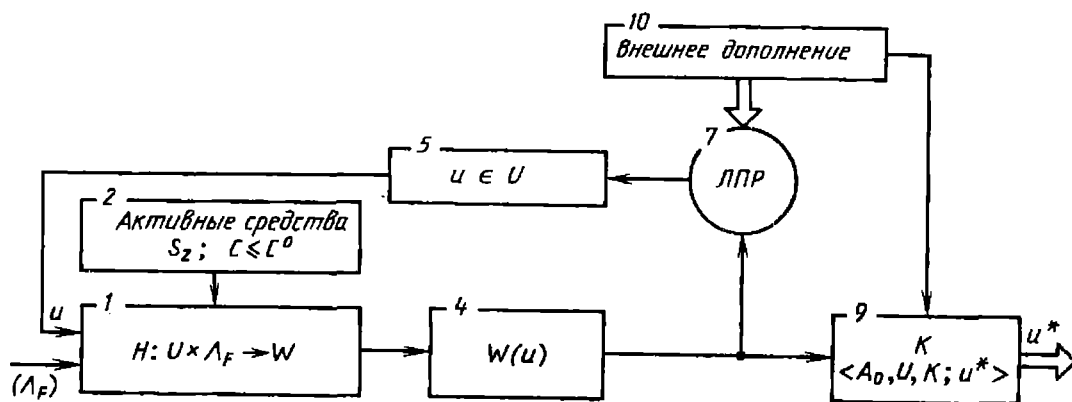


Рис. 3. Частная схема одноэтапной операции

других признаков. Так, если $W(u)$ есть линейная функция от аргумента $u = \{u_1, u_2, \dots, u_n\}$ и $\eta(u) \leq C$ — система линейных неравенств, то (7.13) есть задача линейного программирования. При несоблюдении хотя бы одного из этих условий задача (7.13) становится задачей нелинейного программирования. Последняя имеет широкое множество частных постановок задач. Так, при линейной системе ограничивающих неравенств и квадратичной зависимости $W(u)$ ставится задача квадратичного программирования. Если множество U и функция $W(u)$ выпуклы, то формулируют задачу выпуклого программирования. При целочисленных значениях $u \in U$ ставится задача целочисленного программирования. Ряд методов решения задач математического программирования изложен в т. 2 справочника.

2. Предположим, что в условиях рассмотренной задачи необходимо учесть действие случайных факторов Λ_E . Очевидно, при этом выход $\vartheta(u)$ будет случайной переменной и, следовательно, функция соответствия $\rho(\vartheta(u), u^T P)$ вводится как числовая случайная переменная (случайная величина). Показатель эффективности $W(u)$ представляет собой математическое ожидание функции соответствия. Система ограничивающих неравенств $\eta_i(u, \lambda_E) \leq C_i$ ($i = 1, 2, \dots, m$) также может иметь стохастический характер. Поэтому ее заменяют обычно либо системой

$$M[\eta_i(u, \lambda)] \leq C_i, \quad (7.14)$$

либо

$$P(\eta_i(u, \lambda) \leq C_i) \geq \alpha_i, \quad (7.15)$$

где α_i — заданные уровни вероятности того, что i -е неравенство $\eta_i(u, \lambda) \leq C_i$ будет выполнено. Оператор выхода H здесь примет вид:

$$H: U \times \Lambda \rightarrow R_e^1,$$

где R_e^1 — множество действительных чисел, на котором определяется функция соответствия.

В этих условиях может быть поставлена следующая задача стохастического программирования: найти $u^* \in U$, при которых показатель эффективности достигает экстремального (например, максимального) значения, т. е. $u^*: \max_{u \in U} W(u)$ при вероят-

ностных ограничениях (7.14) или (7.15).

Разнообразие различных частных задач стохастического программирования весьма обширно. Эти задачи классифицируются в зависимости от вида функции соответствия и ограничивающих неравенств, от типа распределений случайных переменных, учитываемых моделью операции, и ряда других признаков. В т. 2 справочника подробно рассмотрены методы решения задач стохастического программирования.

3. В условиях рассмотренных двух частных случаев теперь проводится многошаговая операция. S_0 -система на каждом шаге процесса ($t = 0, 1, 2, \dots$)

может менять свои состояния в соответствии с оператором перехода:

$$\Phi : Z \times U \times T \rightarrow Z \quad (7.16)$$

(процессы с полной информацией, в которых действием возмущающих факторов можно пренебречь) или

$$\Phi : Z \times U \times T \times \Lambda \rightarrow Z \quad (7.17)$$

(процессы с неполной информацией, учитывающие действия возмущающих факторов Λ).

Результат многошаговой операции обычно вводится как функционал выходного процесса к некоторому моменту времени (конец операции). Оператор выхода H удобно представить в виде

$$H : Z \times U \times T \times \Lambda \rightarrow R_e^1, \quad (7.18)$$

где R_e^1 — пространство значений функции соответствия $\rho(y(u), y^{TP})$.

Если действия возмущающих факторов Λ не учитываются, то R_e^1 отождествляется с пространством значений показателя эффективности W , т. е.

$$H : Z \times U \times T \rightarrow W. \quad (7.19)$$

Учет возмущений случайного характера приводит к необходимости введения вероятностной меры на пространстве R_e^1 . Тогда в качестве показателя эффективности принимают математическое ожидание функции соответствия, которая в этих условиях является случайной величиной.

Случайный характер функции соответствия обусловлен тем, что процесс смены состояний z^t системы является случайным управляемым процессом, и, следовательно, результат $y(u)$, вводимый как функционал от этого процесса, есть случайная величина, а функция $\rho(y(u), y^{TP})$ как функция случайного аргумента будет также случайной величиной.

В детерминированном случае (процессы с полной информацией) ЛПР, выбирая на каждом шаге многоэтапного процесса управления операцией стратегию $u_t \in U$, стремится получить экстремум показателя эффективности и не нарушить при этом систему различного рода ограничений. Такая задача является основной в сов-

ременной теории оптимального управления. Эффективные методы решения этой задачи в зависимости от конкретных условий ее постановки основаны на принципе оптимальности Беллмана (динамическое программирование) или на принципе максимума Понтрягина. При некоторых достаточно сильных ограничениях задача может быть поставлена в классическом варианте как задача вариационного исчисления. Принцип оптимальности и некоторые методы решения подобных задач изложены в т. 2 справочника.

При учете случайных возмущений (процессы с неполной информацией) приходится использовать вероятностное описание процесса. ЛПР, выбирая на каждом шаге определенную стратегию u_t из множества допустимых, изменяет лишь распределение вероятностей перехода системы из одного состояния в другое. Однако и в этом случае выбор стратегий подчинен стремлению ЛПР достичь экстремального значения показателя эффективности операции. Подобная задача также рассматривается в современной теории оптимального управления. Однако решение ее в общем виде не получено. Тем не менее достаточно полно разработаны различные методы управления случайными процессами при некоторых дополнительных предположениях. В частности, при описании функционирования многих технических систем используются марковские управляемые случайные процессы. В марковском случайном процессе распределение вероятностей переходов в различные состояния к моменту времени t зависит лишь от состояния z_{t-1} , в котором находилась система в момент времени $t-1$, но не зависит от траектории процесса до момента $t-1$. Таким образом, для марковского процесса имеет место следующее равенство условных вероятностей перехода в состояние z_t :

$$P(z_t/z^{t-1}) = P(z_t/z_{t-1}),$$

где $P(z_t/z^{t-1})$ — условная вероятность перехода системы в состояние z_t к моменту времени t , вычисленная в предположении, что до момента t система прошла состояния $z_0, z_1, \dots, z_{t-1}$.

$P(z_t/z_{t-1})$ — условная вероятность перехода системы в состояние z_t к моменту t , вычисленная в предположении, что в момент $t-1$ система находилась в состоянии z_{t-1} (здесь предистория процесса z^{t-2} не учитывается). Марковский случайный процесс называют *управляемым*, если условное распределение вероятностей состояний к моменту t формируется на основе выбранной ЛПР стратегии u_{t-1} из множества допустимых на предыдущем шаге процесса.

Системы (7.4) условных вероятностных мер для марковского управляемого случайного процесса можно записать в виде

$$\begin{aligned} \{\mu_t(\Sigma_z/z_{t-1}, u_{t-1}) = \\ = P(z_t \in \Sigma_z/z_{t-1}, u_{t-1}); \forall t \in T\}. \end{aligned} \quad (7.20)$$

В технических приложениях довольно широкое распространение получили полумарковские управляемые случайные процессы, которые вводятся следующим образом. Пусть множество состояний Z S_0 -системы конечно. Система может менять скачкообразно свои состояния в случайные моменты времени $t_0, t_1, t_2, \dots, t_n, \dots$ (t_0 — начальный момент; $t_0 < t_1 < t_2 < \dots < t_n < \dots$). Тогда это будет процесс с непрерывным временем. Предположим, что временные промежутки $\tau_i = t_i - t_{i-1}$ ($i = 1, 2, 3, \dots$) есть независимые по совокупности случайные величины, причем их распределения зависят от состояния системы в момент t_{i-1} и от состояния в момент t_i . При этом предположении и справедливости условия (7.20) процесс называют *полумарковским*. Из общего определения полумарковского процесса как частный случай получают марковский процесс с дискретным временем (все τ_i неслучайны, т. е. образуют регулярную последовательность чисел) и марковский процесс с непрерывным временем (все τ_i имеют экспоненциальные распределения).

Элементы теории марковских процессов изложены в т. 2 справочника.

4. Рассмотрим случай, когда в S_0 -системе активно действуют два распорядителя. Один из них является ЛПР.

С точки зрения его интересов вводится цель A_0 операции, проводится описание S_0 -системы и устанавливается множество допустимых стратегий U . Действия другого субъекта системы могут оказать существенное влияние на результат операции. До проведения операции эти действия для ЛПР представляют собой поведенческую неопределенность. Используя внешнее дополнение, ЛПР выдвигает гипотезы поведения партнера в операции (цели, отражающие его интересы, возможные реакции на действия ЛПР, множество допустимых стратегий V , принципы, лежащие в основе его рационального поведения, а также показатель эффективности, на основе которого он оценивает эффективность операции). На схеме операции (см. рис. 3) сохраняются почти все блоки. Однако в каждой группе блоков 3, 6 и 8 остается лишь по одному блоку.

Предположим, что с точки зрения ЛПР эффективность операции оценивается показателем $W^\omega(u, v)$, тогда как по мнению ЛПР другой субъект системы свои интересы в операции отражает показателем эффективности $W^\alpha(u, v)$. Можно, например, предположить, что критерием эффективности для ЛПР является правило, предписывающее выбирать стратегию $u^* \in U$ из условия максимизации $W^\omega(u, v)$, а критерий эффективности для партнера рекомендует ему выбирать стратегию $v \in V$ из условия минимизации $W^\alpha(u, v)$. Однако наличие взаимодействий партнеров в операции, проявляющееся в зависимости показателя одного из партнеров от стратегий другого, требует введения сложного критерия эффективности, учитывающего эти взаимодействия.

Допустим, в частности, что показатели эффективности обоих распорядителей совпадают:

$$W^\omega(u, v) = W^\alpha(u, v) = W(u, v). \quad (7.21)$$

Однако их интересы в операции прямо противоположны (ЛПР стремится к достижению максимального значения $W(u, v)$, а его партнер — к минимальному значению $W(u, v)$). Подобный

тип взаимодействия партнеров называют *конфликтом* (противодействием) в антагонистической форме. Именно такую ситуацию исследует теория антагонистических игр [16]. Другой частный случай взаимодействия появляется при выполнении условия (7.21) и при полностью совпадающих интересах партнеров (либо к достижению максимума $W(u, v)$, либо к достижению его минимума). Эта форма содействий партнеров называется *единством* и сводится к системе с одним распорядителем.

Отсутствие взаимодействий между субъектами системы приводит к ситуации нейтралитета. В этом случае показатели эффективности $W^w(u)$ и $W^a(u)$ зависят только от стратегии соответствующего распорядителя, и операция рассматривается в рамках системы с одним распорядителем—ЛПР.

Конфликт между субъектами может возникнуть не только в своей крайней антагонистической форме. Конфликтная ситуация возникает всякий раз, когда действия одного из партнеров противоречат стремлению к достижению своей цели другим партнером. Если партнеры в операции свои интересы оценивают разными показателями эффективности (условие (7.21) не выполняется), а любая стратегия одного из партнеров прямо противоречит стремлению, например, к максимизации показателя эффективности другого партнера, то такую ситуацию называют *строгим конфликтом* в отличие от *нестрогого конфликта*, при котором не все стратегии из множества допустимых прямо противоречивы стремлениям другого партнера.

Рациональное поведение партнеров в условиях конфликта обычно основывается на принципе равновесия по Нэшу [16]. При стремлении партнеров к максимизации своих показателей эффективности $W^w(u, v)$ и $W^a(u, v)$ равновесными в смысле Нэша стратегиями u^* и v^* конкурирующих партнеров будут стратегии, выбираемые из условия:

$$\begin{aligned} W^w(u^*, v^*) &\geq W^w(u, v^*); \\ W^a(u^*, v^*) &\geq W^a(u^*, v). \end{aligned} \quad (7.22)$$

В общем случае, ни один из партнеров не заинтересован в нарушении равновесного состояния.

Содействие партнеров в операции может проявляться не только в их единстве (в крайней, наиболее сильной форме), но также в форме строгого и нестрогого содружества или в форме коалиции. Эти формы содействия возникают в ситуациях, при которых стратегии одного из партнеров способствуют достижению своей цели в операции другим партнером. Рациональное поведение партнеров и в этих ситуациях организуется на основе определенных признаков равновесия, нарушение которых хотя бы одним из партнеров приводит к снижению эффективности операции.

Рассмотренные частные случаи показывают, что в не слишком сложных ситуациях (типичных для систем II методологического уровня, «структура-функция») исследование эффективности операции удается провести с использованием аналитических моделей S_0 -систем. Методическая схема исследования операций основывается на возможности идентификации реальной задачи с классом стандартных задач, например, математического программирования. Если реальную задачу удастся адекватно описать и поставить как задачу математического программирования, то дальнейший ход ее решения сводится к подбору готового разработанного алгоритма поиска решения при конкретных исходных данных. Однако при исследовании эффективности современных крупномасштабных операций с использованием больших технических систем (систем III методологического уровня) подобный подход может привести к существенной утрате адекватности модели, к недопустимому упрощению задачи и вследствие этого к выхолащиванию содержательной стороны получаемого математического решения.

Общая схема операции (см. рис. 2) может быть реализована в полном объеме в рамках заранее разработанной проблемно-ориентированной имитационной системы. Однако разработка имитационных систем требует больших затрат труда, времени и других материальных средств. Поэтому перед

началом исследований необходимо убедиться, что более простой подход к исследованию эффективности операции, основанный на идентификации решаемой задачи с задачей, например, математического или стохастического программирования, неприемлем.

4. МАТЕМАТИЧЕСКИЕ МОДЕЛИ ТЕХНИЧЕСКИХ СИСТЕМ

Реальная система состоит из элементов, которые под воздействием входных сигналов и в зависимости от своего состояния могут переходить в новое состояние и формировать выходной сигнал.

Этот выходной сигнал может быть воспринят другим элементом системы как входной, а в качестве реакции на него элемент перейдет в новое состояние и сформирует выходной сигнал. Образуется система связи элементов между собой, а также с внешней средой.

Процесс функционирования одного элемента сложной системы математически можно описать как *общую динамическую систему* (ОДС) [9]. Термин «общая» означает, что в рамках ОДС может учитываться не только причинный, но и стохастический характер движения системы (под движением ТС в широком смысле понимают процесс смены ее состояния).

Пусть z^t — вектор, характеризующий состояние ОДС в момент времени t ; $z_1^t, z_2^t, \dots, z_n^t$ — компоненты вектора z^t ; Z — множество различных возможных состояний системы, $z^t \in Z \forall t \in T$; x^t — входной сигнал в момент времени t , представляющий собой вектор с координатами $x_1^t, x_2^t, \dots, x_k^t$; y^t — вектор выходных сигналов в момент t ; $y^t = (y_1^t, y_2^t, \dots, y_k^t)^T$; $y \in Y \forall t \in T$.

Общей динамической системой (ОДС) называют систему, находящуюся в каждый момент времени $t \in T$ в одном из возможных состояний $z^t \in Z$ и способную переходить во времени из одного состояния в другое случайным образом или (и) под действием внутренних и внешних причин, совершая при этом движение $z(t)$.

Пусть Z — замкнутая ограниченная область в n -мерном пространстве. Из начального состояния z_0 , в котором система находится в начальный момент времени $t_0 \in T$, начинается движение $z(t)$ под действием внутренних причин. Это движение может задаваться любым аналитическим или алгоритмическим оператором. Движение продолжается либо до поступления в систему входного сигнала x , либо до выхода точки z на границу области Z .

В обоих этих случаях система скачком переходит в новое состояние. Величина и направление скачка в пространстве состояний могут быть случайными. Точка, отображающая состояние системы после скачка, находится в области Z .

Пусть в момент $t' > t_0$, когда система находилась в состоянии z' , поступил входной сигнал x' . Под его воздействием система совершит скачок, который определяется случайным вектором:

$$h' = z'' - z', \quad (7.23)$$

где z'' — вектор состояния системы после скачка, $z'' \in Z$. Распределение случайного вектора h' зависит от z' и t' . Из новой точки z'' (состояние системы после скачка) система движется под действием внутренних причин. Если в момент времени $t'' > t'$ поступит входной сигнал x'' , то система совершит скачок h'' в новое состояние, из которого продолжит движение под воздействием внутренних причин. При этом распределение вектора h'' будет зависеть в общем случае от состояния системы перед приходом сигнала x'' и от времени t'' .

Пусть в момент времени t^* система под действием внутренних причин приходит в точку z^* на границе области Z . Из точки z^* система совершает скачок внутрь области Z , определяемый случайным вектором $h^* = \bar{z} - z^*$, где $\bar{z}$ — состояние после скачка.

Каждый раз, когда система совершает скачок под воздействием входного сигнала или при выходе на границу области Z , на ее выходе появляется сигнал

$$y_t = y(t, z, x), \quad (7.24)$$

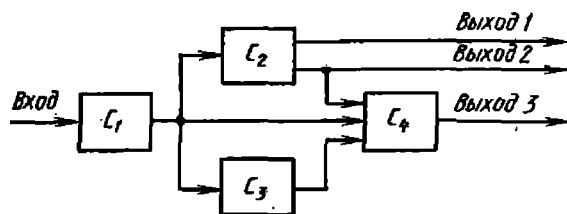


Рис. 4. Пример структурной схемы системы

где $\bar{t}$ — момент скачка; $\bar{z}$ — состояние, из которого совершается скачок; $\bar{x}$ — входной сигнал, под действием которого происходит скачок (при выходе на границу области Z сигнал $\bar{x}$ в (7.24) отсутствует).

Таким образом, ОДС задается следующими группами соотношений:

уравнениями границы области возможных состояний Z ;

уравнениями движения точки $z(t)$ внутри области Z ;

соотношениями для формирования распределения вероятностей скачков состояний при выходе на границу области Z ;

соотношениями для формирования распределения вероятностей скачков состояний при поступлении входных сигналов;

соотношениями для расчета выходных сигналов.

Моменты поступления входных сигналов ОДС могут быть случайными. Точка выхода на границу области состояний Z (точка z^*) определяется совместным решением уравнений движения системы под действием внутренних причин с уравнениями границы области.

ОДС может служить моделью широкого класса различных реальных систем. Однако в большинстве случаев на практике удобно использовать более простые модели по сравнению с ОДС. Такие модели могут быть описаны как частные случаи ОДС, получаемые путем введения некоторых упрощающих допущений.

Согласно модульному принципу моделирования математическая модель системы представима в виде совокупности модулей, описанных как ОДС или один из ее частных случаев, и оператора сопряжения. Оператор сопряжения описывает связи отдельных модулей (ОДС) между собой и с внешней средой. В общем виде оператор сопряжения R представляет собой ориентированный граф, вершины которого соответствуют отдельным модулям (ОДС), а выходящие из них дуги — связям. Наиболее удобна для практических целей табличная форма записи оператора сопряжения [9].

Пусть процессы функционирования модулей C_k и C_j (рис. 4) описаны как процессы в ОДС. Входной сигнал модуля C_j представляет собой m_j -мерный вектор:

$$x^t = (x_1^t, x_2^t, \dots, x_{i'}^t, \dots, x_{m_j}^t)^T \quad (7.25)$$

Выходной сигнал модуля C_k описывается n_k -мерным вектором:

$$y = \{y_1^t, y_2^t, \dots, y_{l'}^t, \dots, y_{n_k}^t\}. \quad (7.26)$$

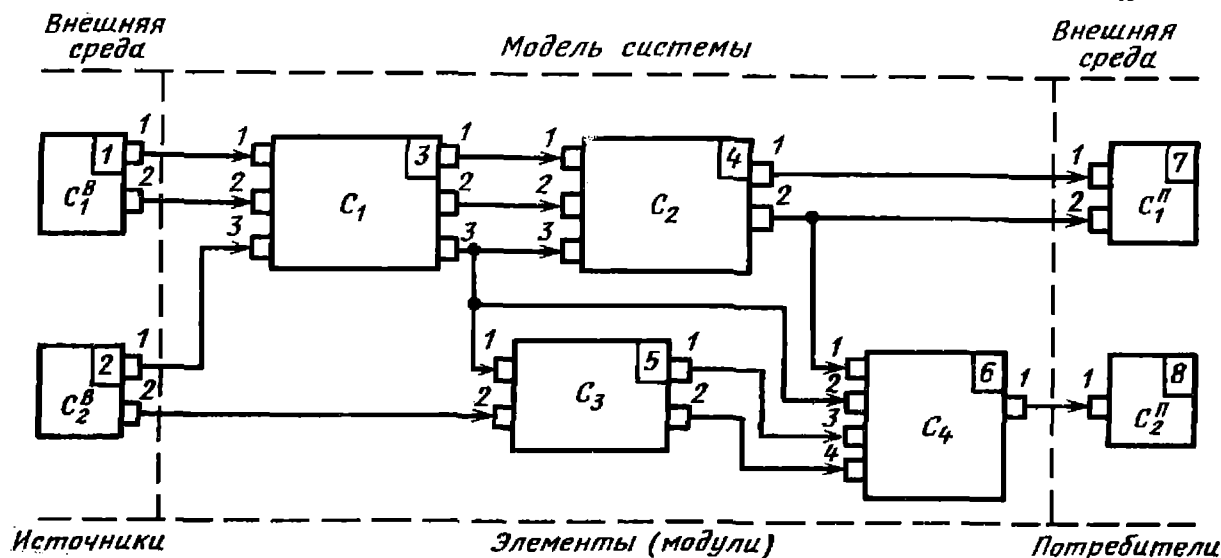


Рис. 5. Схема сопряжения модели

Если обозначить $Y_l^{(k)}$, $X_i^{(j)}$ — соответственно l -й выходной контакт модуля C_k и i -й входной контакт модуля C_j , то запись

$Y_l^{(k)} = R [X_i^{(j)}], \tag{7.27}$

где $R = \|r_{ij}\|$, означает, что l -й выходной контакт модуля C_k связан (передает сигнал) с i -м входным контактом модуля C_j (рис. 5). Теперь оператор сопряжения R можно задать в виде матрицы:

	1	2		M
C_1^B	r_{11}	r_{12}		r_{1M}
C_2^B	r_{21}	r_{22}		r_{2M}
$C_{\eta_1}^B$	$r_{\eta_1 1}$	$r_{\eta_1 2}$		
C_1	$r_{\eta_1+1, 1}$	$r_{\eta_1+1, 2}$		
C_2	$r_{\eta_1+2, 1}$	$r_{\eta_1+2, 2}$		
C_{η_2}	$r_{\eta_1+\eta_2, 1}$	$r_{\eta_1+\eta_2, 2}$		
$C_1^П$				
$C_2^П$				
$C_{\eta_3}^П$	$r_{\eta_1+\eta_2+\eta_3, 1}$	$r_{\eta_1+\eta_2+\eta_3, 2}$		$r_{\eta_1+\eta_2+\eta_3, M}$

(7.28)

В (7.28) $C_1^B, C_2^B, \dots, C_{\eta_1}^B$ — источники сигналов от внешней среды; $C_1, C_2, \dots, C_{\eta_2}$ —элементы системы (модули); $C_1^П, C_2^П, \dots, C_{\eta_3}^П$ — потребители сигналов (внешняя среда). Строки матрицы (7.28) соответствуют элементам системы, источникам и потребителям сигналов. Столбцы матрицы оператора сопряжения соответствуют номерам контактов. Пересечение строки и столбца определяет соответствующий

входной контакт. На этом пересечении ставится пара чисел:

$r_{ij} = \langle k, l \rangle, \tag{7.29}$

где k — номер модуля (источника), а l — номер его выходного контакта, с которым соединяется j -й входной контакт i -го модуля (потребителя). Поскольку связи r_{ij} для

$\forall i \in [1, \eta_1] \text{ и } \forall k \in [1, M] \tag{7.30}$

отсутствуют (источники сигналов от внешней среды $C_1^B, C_2^B, \dots, C_{\eta_1}^B$ не имеют входов), то верхняя часть табл. 1 (первые η_1 строк) обычно не изображается.

Схема сопряжения элементов дана на рис. 5. Система (модель) имеет четыре элемента (модуля) C_1-C_4 , два источника сигналов от внешней среды C_1^B, C_2^B и два потребителя сигналов $C_1^П, C_2^П$. Отдельные связи между модулями (источниками, потребителями) называются каналами передачи сигналов. Каждый входной контакт соединяется только с одним каналом. Каждый выходной контакт может быть соединен с любым числом каналов. Оператор сопряжения для такой схемы задается табл. 1.

На рис. 5 представлена одноуровневая схема сопряжения. Однако модель сложной системы может состоять из моделей подсистем, внутри каждой из которых задан свой оператор сопряжения.

Между моделями подсистем связи задаются оператором сопряжения более высокого уровня, чем операторы сопряжения элементов подсистем. Так формируется многоуровневая схема сопряжения. Описанная форма сопряжения (контакт с контактом) называется *канонической*.

Если оператор сопряжения не меняется во времени, то система (модель)

1. Оператор сопряжения элементов

Модули		Номера контактов			
Имя	Номер	1	2	3	4
C_1^B	1	—	—	—	—
C_2^B	2	—	—	—	—
C_1	3	1,1	1,2	2,1	—
C_2	4	3,1	3,2	3,3	—
C_3	5	3,1	2,2	—	—
C_4	6	4,2	3,3	5,1	5,2
$C_1^П$	7	4,1	4,2	—	—
$C_2^П$	8	6,1	—	—	—

обладает постоянной структурой. Для моделей ТС с переменной структурой

$$Y_i^{(k)} = R[x_i^{(j)}, t]. \quad (7.31)$$

Часто для моделей сложных систем с переменной структурой оператор сопряжения зависит от состояния системы и от входного сигнала (под состоянием сложной системы понимают совокупность состояний всех ее элементов).

В отдельных задачах оценивания эффективности оператор сопряжения может быть поставлен в зависимость от некоторого параметра $\gamma \in \Gamma$. В задачах синтеза оптимальной структуры γ — конструктивный параметр, подбираемый таким образом, чтобы оптимизировать поведение системы.

Если γ — случайная переменная, имеющая функцию распределения на множестве Γ , то система называется системой со *стохастической структурой*. В более сложных случаях параметр γ может быть неопределенным фактором нестохастического характера. Так, при исследовании конфликтных ситуаций оперирующая сторона не может полностью знать структуру системы конкурента. Информация о параметре γ может сводиться лишь к описанию принадлежности γ какому-либо подмножеству значений Γ .

Таким образом, в самом общем случае динамику связей модели сложной системы отражает оператор сопряжения

$$Y_i^{(k)} = R[x_i^{(j)}, z(t), \gamma(t), t]. \quad (7.32)$$

Параметр γ зависит от времени в задачах оценивания эффективности, подразумевающих оптимальное управление структурой.

Преимущества модульного принципа при построении математических моделей систем в виде ОДС проявляются в следующем [9].

1. Вектор состояний ОДС не имеет ограничений размерности, что обеспечивает удовлетворение различных неоднородных требований при формализации элементов и подсистем исследуемой системы.

2. Задающие поведение ОДС функции переходов и выходов опреде-

ляются в известной степени произвольным образом, что существенно облегчает аналитическое или алгоритмическое описание различных процессов и систем.

3. Описания функций переходов, функций выходов и операторов сопряжения ОДС в виде таблиц после соответствующей формально-логической обработки легко переводятся в программу для ЭВМ.

4. Применяемый в качестве методологической основы модульной концепции подход, реализующий схему «вход—выход», приводит к разграничению внутренних и внешних воздействий и взаимосвязей, что способствует более глубокому пониманию их сущности в процессе оценивания эффективности и выработки решений.

5. Связь между модулями (ОДС) в модели устанавливается путем обмена сигналами. Это позволяет определять структуру и класс сигналов таким образом, чтобы не возникали затруднения при необходимости расширения состава модулей и их изменения.

6. Модульная концепция на базе ОДС создает предпосылки для вовлечения в состав средств машинного моделирования мультипроцессорных систем. В результате параллельной обработки информации в ходе эксперимента с моделью сложной системы может быть получена существенная экономия машинного времени.

Все эти обстоятельства имеют решающее значение при выборе ОДС как типового (универсального) элемента математических моделей систем.

Частными случаями ОДС являются математические модели, у которых переход исследуемой системы из одного состояния в другое (скачок) отсутствует или происходит случайным образом. Ниже перечислены наиболее распространенные частные случаи ОДС.

1. *Динамическая система* (движение системы в пространстве состояний происходит только под действием внутренних причин и определенных входных сигналов, случайные скачки состояний не допускаются). Различают динамические непрерывные системы (ДНС), у которых состояния изменяются непрерывным образом, и ди-

намические дискретные системы, у которых сигналы и состояния рассматриваются только в дискретные моменты времени.

2. *Конечный автомат* (система при каждом поступлении входного сигнала в дискретные моменты времени мгновенно совершает скачок в новое случайное состояние; при этом множество состояний и множества входных и выходных сигналов конечны). Частными случаями конечных автоматов являются автоматы Мили, автоматы Мура, конечные преобразователи, функциональные элементы, генераторы (см. т. 2 справочника).

3. *Вероятностный автомат* (расширение конечного автомата на случай, когда функция переходов (7.3) задается стохастической матрицей $P_y = [P_{ij}(x)]$). Иногда упоминается как асинхронный автомат, универсальный стохастический инициальный автомат [27, 38]. Вероятностные автоматы используются при оценивании технических систем с ненадежными элементами, систем с изменяющейся во времени (случайным образом) структурой, а также систем, структура и состояние которых в некоторые моменты времени точно не известны и могут быть описаны лишь в терминах теории вероятностей: сети связи с ненадежными каналами, технологические и производственные системы при наличии случайных отказов оборудования или случайного количества работающих и т. д.

4. *Система массового обслуживания* (СМО) [на входы системы через дискретные, возможно, случайные интервалы времени поступают сигналы (заявки на обслуживание)]. Порядок и продолжительность прохождения сигнала через систему (продолжительность обслуживания) определяются дисциплиной обслуживания. При поступлении очередного сигнала (приход очередной заявки) или по окончании обслуживания какой-либо заявки система скачком переходит в новое состояние; выходные сигналы отсутствуют [13].

В рамках ОДС можно моделировать практически любые СМО: многоканальные (число входов $n > 1$), с различ-

ными дисциплинами обслуживания, с приоритетами заявок и т. п. (см. т. 2). Полное аналитическое описание допускают лишь простейшие СМО.

СМО используются в качестве моделей сетей связи, сетей технического обслуживания, технологических линий и производств, транспортных сетей и систем, систем управления запасами (склад, порт), систем телеобработки данных, вычислительных систем и сетей. В большинстве прак-

тически важных случаев используется алгоритмическое описание СМО.

5. *Кусочно-линейный агрегат* (КЛА). Для его описания в рамках ОДС принимается единственное допущение: движение под действием внутренних причин описывается линейными дифференциальными уравнениями [10]. С помощью КЛА моделируются процессы в сложных системах массового обслуживания, в различных линейных системах управления и регулирования.

Глава 8. Имитационные модели исследования эффективности

1. КЛАССИФИКАЦИЯ ИМИТАЦИОННЫХ МОДЕЛЕЙ

Одной из основных форм системного анализа эффективности операций (технических систем) является имитационное исследование, проводимое на имитационных моделях. Наибольшее распространение в практике исследования эффективности имеют цифровые имитационные модели, т. е. модели, реализуемые на универсальных цифровых ЭВМ или специальных цифровых вычислителях. Аналоговые имитационные модели, реализуемые с помощью аналоговых вычислительных машин (АВМ), и гибридные имитационные модели, создаваемые на специальных моделирующих комплексах, содержащих как аналоговые, так и цифровые блоки, не нашли достаточного применения при решении задач исследования эффективности. Рассмотрим цифровые имитационные модели, при этом термин «цифровые» опустим. Поскольку речь идет о машинных имитационных моделях, то методы их построения и использования рассмотрим, в первую очередь, в аспекте программной реализации этих моделей на ЭВМ.

Основным классификационным признаком имитационных моделей является способ отображения в них динамических процессов, протекающих в моделируемой системе. Принято определять динамическое поведение моделируемой системы в терминах понятий, называемых работами (или активностями), событиями и процессами.

Работы (активности). Независимо от содержания формализуемой в модели деятельности системы активность пред-

ставляется единым динамическим объектом, указывающим на совершение некоторой единицы работы. Хотя любая активность, в принципе, может быть представлена совокупностью более мелких функций, выполняемых в течение некоторого времени, в рамках конкретной модели она рассматривается как единый дискретный шаг. Например, при моделировании процесса жизненного цикла изделия на одном из верхних уровней моделирования этого процесса в качестве активностей можно представить этапы НИОКР, подготовки производства, выпуск продукции на серийном заводе и т. д. Каждая из активностей характеризуется временем ее выполнения и потребляемыми ресурсами. Очевидно, что при более детальном рассмотрении этого процесса каждый из этапов можно представить и в виде совокупности других более мелких активностей (работ).

События. События представляют собой мгновенное изменение некоторого объекта системы. Окончание работы (активности) в системе является событием. Это событие, в свою очередь, может послужить инициатором другой активности в системе. Событием может быть и изменение состояния объекта системы, например, занятие или освобождение станка, обслуживающего поток заготовок на производстве. В общем случае события можно разделить на две категории: события следования, которые управляют инициализацией работ внутри рассматриваемого процесса, и события изменения состояний, которые управляют выполнением работ, относящихся, в общем случае, к независимым процессам.

Процессы. Под процессом в имитационной модели понимают логически связанный набор активностей (работ). Некоторые процессы, в свою очередь, могут рассматриваться как работы (активности) в процессе более высокого уровня. Различие между активностями и процессами определяется в конечном счете степенью детализации моделируемого объекта. Операция, определенная в рамках модели одного уровня представления объекта как активность, может быть рассмотрена в виде процесса, включающего несколько активностей, в более детальной модели объекта. Например, рассмотрим модель функционирования некоторой вычислительной системы. В качестве элементарного процесса можно рассмотреть процесс ввода-вывода информации с магнитного диска. Этот процесс представляется состоящим из нескольких работ (активностей): установки головки записи-чтения, задержки на вращение диска и передачи данных. Если рассмотреть модель функционирования вычислительной системы применительно к выполнению программы, то в качестве элементарного процесса можно представить последовательность машинных команд программы, а составляющими ее активностями будут активность вычислений в оперативной памяти и активность обмена информацией программы с диском. Выполнение программы можно рассматривать как активность, входящую в процесс выполнения задания в рамках мультипрограммной вычислительной системы.

Процессы, включающие одни и те же типы работ и событий, относят к одному классу. Моделируемую систему в общем случае можно представить несколькими классами процессов. Каждый класс состоит из нескольких реализаций процессов, находящихся в любой момент времени в различных фазах выполнения. Реализация процесса однозначно идентифицируется среди процессов того же класса характеристиками его параметров.

Таким образом, чтобы описать моделируемую систему в терминах процессов, достаточно описать каждый класс процессов, из которых состоит система, и инициировать заданное

число реализаций каждого класса с соответствующими значениями их параметров. Описание каждого класса процессов состоит из указаний в определенной последовательности составляющих эти процессы условий, определяющих порядок изменения заданной очередности выполнения, а также воздействий, оказываемых процессами на характеристики и состояния элементов системы.

Важным классификационным признаком имитационных моделей является также способ регламентации событий и процессов в модели.

Так как в большинстве имитационных моделей имитируется поведение исследуемой системы на некотором интервале времени, принципиальным отличием имитационных моделей является порядок изменения временной координаты и способ согласования различных событий в системе. Эти два аспекта имитационных моделей известны как механизм системного времени.

В имитационных моделях оператор H формирования исхода задается алгоритмически. При этом действия, выполняемые на каком-либо этапе алгоритма (выполняемые каким-либо блоком модели), зависят от действий на других этапах (от действий и состояний других блоков модели). Эти действия должны быть скоординированы во времени (синхронизированы), иначе могут быть нарушены причинно-следственные связи в модели, что может привести к ошибочным результатам исследования.

Поскольку в реальной операции некоторые события могут происходить одновременно, а параллельные процессы могут взаимодействовать друг с другом, в имитационной модели имеется специальный алгоритм управления моделированием (механизм системного времени), обеспечивающий синхронизацию событий и процессов.

Синхронизация в имитационной модели означает такую организацию искусственного времени, которая обеспечивает появление событий (окончание этапов работы моделирующего алгоритма) в надлежащем порядке и с необходимыми интервалами между ними.

Существуют два основных механизма системного времени — задание времени с помощью постоянных и с помощью переменных интервалов времени (шагов) [60]. Второй из этих механизмов известен так же, как моделирование по особым состояниям.

По методу постоянного шага отсчет системного времени ведется через фиксированные, заранее выбранные исследователем интервалы времени. События в модели считаются наступившими в момент окончания этого интервала.

При моделировании по особым состояниям системное время каждый раз изменяется на величину, строго соответствующую интервалу времени до момента наступления очередного события. В этом случае события обрабатываются поочередно, в порядке их наступления, и регистрируются как одновременные лишь в том случае, если они одновременны в действительности.

Метод моделирования по особым состояниям сложнее в реализации, чем моделирование с постоянным шагом, так как для него требуется введение специальной процедуры (календарь событий) для определения моментов времени появления последующих событий и ранжирования их в порядке возрастания.

Выбор метода организации системного времени зависит от назначения модели, ее сложности, характера исследуемых процессов, имеющихся машинных ресурсов, требуемой точности результатов и т. д.

Метод фиксированных шагов предпочтительнее, если: события появляются регулярно, их распределение во времени достаточно равномерно; число событий велико, а продолжительность отдельного события мала; невозможно заранее установить, какие из событий являются существенными [73].

Метод моделирования по особым состояниям целесообразно использовать, если события распределяются во времени неравномерно или их продолжительность велика. Этот метод позволяет экономить машинное время при моделировании статических систем или систем периодического действия, в которых события могут дли-

тельное время не наступать. При построении моделей, используемых при исследовании эффективности операций, чаще применяют метод моделирования по особым состояниям.

В зависимости от способа взаимодействия пользователя с моделью в ходе исследования имитационные модели делят на автоматические и диалоговые. *Автоматическими* называются имитационные модели, взаимодействие пользователя с которыми сводится только к вводу исходной информации и управлению началом и окончанием работы моделей.

Имитационные модели и системы, позволяющие пользователю активно управлять ходом моделирования, называются *диалоговыми*.

2. МОДЕЛИРОВАНИЕ СЛУЧАЙНЫХ ФАКТОРОВ

При имитационном исследовании эффективности операций и технических систем приходится моделировать случайные факторы, представляемые в моделях как случайные события, величины, векторы, функции (процессы).

В основе всех методов и приемов моделирования случайных факторов лежит использование случайных чисел, имеющих равномерное распределение на интервале $[0, 1]$.

Случайные или «истинно» случайные числа формируются с помощью аналого-цифровых преобразователей на основе сигналов физических генераторов, использующих естественные первичные источники случайных шумов (радиоактивный распад, шумы электронных и полупроводниковых устройств и т. д.) [71].

Случайные числа, генерируемые аппаратно или программно на ЭВМ, называются *псевдослучайными*, так как хотя они и вырабатываются с помощью детерминированных рекурсивных формул, их статистические свойства совпадают со статистическими свойствами чисел, генерированных идеальными механизмами, выбирающими числа из интервала $[0, 1]$ независимо и с равной вероятностью.

В состав математического обеспечения современных ЭВМ серий ЕС,

СМ и «Электроника» входят специальные программы генерации (датчики) псевдослучайных чисел.

Моделирование случайных событий. Для моделирования отдельно взятого случайного события с вероятностью P достаточно одного равномерно распределенного на интервале $[0, 1]$ числа r . При попадании r в интервал $[0, P]$ считают, что событие A наступило, в противном случае — не наступило, т. е.

$$A, \text{ если } r \in [0, P]; \quad (8.1)$$

$$\bar{A}, \text{ если } r \in [P, 1].$$

Для моделирования полной группы несовместных событий $A_0 = \{A_1, A_2, \dots, A_N\}$ с вероятностями соответственно $P_1, P_2, \dots, P_N$ также достаточно одного значения r ; при этом

$$A = A_n \left| \sum_{i=0}^{n-1} P_i \leq r < \sum_{i=0}^n P_i; P_0 = 0, n = \overline{1, N}. \quad (8.2)$$

Если группа событий A_0 не полна, то вводят фиктивное событие A_{N+1} с вероятностью P_{N+1} , такой, что

$$\sum_{i=1}^{N+1} P_i = 1, \quad (8.3)$$

т. е. дополняют группу A_0 до полной, после чего генерируют число r и поступают согласно (8.2). При $A = A_{N+1}$ считают, что ни одно событие из исходной группы A_0 не наступило.

Моделирование непрерывных случайных величин. Основными методами моделирования непрерывных случайных величин с заданным законом распределения являются методы: нелинейных преобразований, исключения (метод Неймана) и композиций [10].

Из методов нелинейных преобразований наиболее широко используют метод обратной функции, основанный на следующем положении: случайная непрерывная величина y с функцией распределения $F(y)$ связана со случайной величиной r , равномерно распределенной на интервале $[0, 1]$, соотношением

$$y = F^{-1}(r), \quad (8.4)$$

где F^{-1} — обратная функция относительно F .

Например, случайную величину y с экспоненциальным законом распределения

$$f(y) = \frac{1}{m_y} \exp\left(-\frac{y}{m_y}\right)$$

можно моделировать, используя обратное преобразование, следующим образом:

$$y = -m_y \ln(1-r) \approx -m_y \ln(r).$$

Однако обратные преобразования существуют только для небольшого числа законов распределения со сравнительно простыми, выражаемыми в конечных квадратурах, функциями распределения; поэтому применение метода обратной функции ограничено.

Метод исключения основан на следующем положении: реализация (конкретное значение в эксперименте) y случайной величины y , имеющей плотность вероятности $f(y)$, является реализацией случайной величины x , имеющей плотность вероятности $g(x)$ в том случае, если отношение этих плотностей вероятностей ограничено сверху:

$$g(x)/f(x) \leq M < \infty \quad (8.5)$$

и

$$M \cdot r \leq g(y)/f(y), \quad (8.6)$$

где r — реализация случайной величины, имеющей равномерное распределение на интервале $[0, 1]$.

Метод исключения состоит в последовательной генерации пар случайных чисел $\{y_i, r_i\}$ и проверке условия (8.6). Если условие выполняется, случайное число y_i принимается как реализация моделируемой случайной величины x ; если не выполняется, то генерируется следующая пара чисел $\{y_{i+1}, r_{i+1}\}$, и процесс повторяется.

Среднее число повторений, необходимых для генерации одного значения случайной величины x , равно $1/M$. Случай, когда $f(y) = \text{const}$ (равномерное распределение), графически иллюстрируется на рис. 1. Точка $\{y, M \cdot r\}$ равномерно распределена в прямоугольнике A ; y исключается, если точка $\{y, M \cdot r\}$ расположена выше

кривой $g(y)$ и принимается за моделируемую реализацию x в противном случае (на рис. 1 в качестве $f(y)$ показано S_B — распределение Джонсона на интервале $[0, 1]$).

Метод исключения удобен для моделирования случайных величин, кривые плотности вероятности которых не имеют пиков, т. е. M невелико и для получения приемлемого значения y требуется небольшое машинное время.

Метод композиции основан на теоремах теории вероятностей, доказывающих представимость одной случайной величины композициями (как правило, линейными) двух или большего числа других случайных величин, имеющих легко реализуемые законы распределения. Так, согласно центральной предельной теореме (см. т. 2) распределение случайной величины

$$v = \frac{\sum_{i=1}^k r_i - \frac{k}{2}}{\sqrt{\frac{k}{2}}}, \quad (8.7)$$

где r_i — равномерно распределенные на интервале $[0, 1]$ случайные числа, с ростом k неограниченно приближается к нормальному распределению с параметрами $m_x = 0$ и $\sigma_x^2 = 1$ (выражение (8.7) для $k = 12$ широко используется для моделирования нормального закона). Известным примером применения метода композиции является также представление случайной величины, имеющей χ^2 -распределение с n степенями свободы, как суммы квадратов n независимых нормально распределенных случайных величин с параметрами $m_x = 0$, $\sigma_x^2 = 1$.

Выбор конкретного способа моделирования непрерывных случайных величин зависит от закона распределения моделируемой величины. При этом один и тот же способ моделирования может быть эффективен при одних значениях параметров распределения и неэффективен при других. Основные статистические распределения непрерывных случайных величин и алгоритмы их имитации приведены в табл. 1

Моделирование дискретных случайных величин. Основными методами

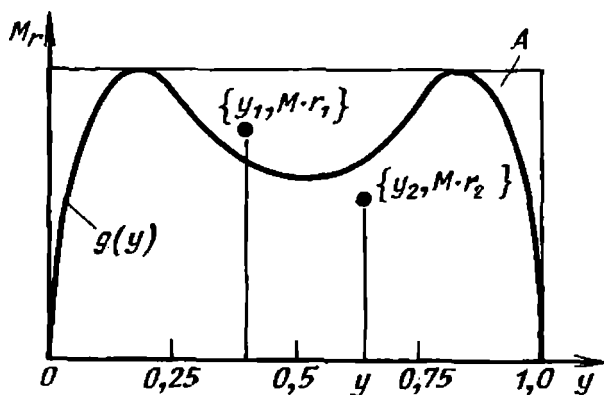


Рис. 1. Метод исключения

моделирования дискретных случайных величин являются метод последовательных сравнений и метод интерпретации.

В методе последовательных сравнений случайное число, равномерно распределенное на $[0, 1]$, последовательно сравнивают с вероятностью появления наименьшего из возможных значений моделируемой случайной величины, с суммой вероятностей появления наименьшего и следующего за ним значений, с суммой вероятностей наименьшего и двух следующих за ним значений и т. д. При первом выполнении условия

$$r \leq \sum_{i=1}^n P_i \quad (8.8)$$

процесс прекращается, и соответствующая случайная дискретная величина считается принявшей значение n . Процесс можно ускорить, применяя методы оптимизации перебора: дихотомии (деления шага пополам), перебора с предварительным ранжированием величин P_i по убыванию и т. д.

Величины P_i рассчитывают по функциям вероятности, соответствующим моделируемому закону.

Метод интерпретации основан на использовании физических явлений, описываемых моделируемыми законами распределения. Так, биномиальное распределение описывает число успехов в n независимых испытаниях с вероятностью успеха в каждом испытании p и вероятностью неудачи $q = 1 - p$. При моделировании этого распределения методом интерпретации выбирают n независимых случайных чисел, рав-

1. Основные непрерывные распределения случайных величин

Распределение	Плотность вероятности	Функция распределения	Параметры	Алгоритм датчика
Равномерное	$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b; \\ 0, & x < a, x > b \end{cases}$	$F(x) = \begin{cases} 0, & x < a; \\ \frac{x-a}{b-a}, & a \leq x \leq b; \\ 1, & x > b \end{cases}$	$m_x = \frac{a+b}{2};$ $D_x = \frac{(b-a)^2}{12}$	—
Показательное	$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0 \end{cases}$	$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0 \end{cases}$	$\lambda > 0;$ $m_x = \frac{1}{\lambda};$ $D_x = \frac{1}{\lambda^2}$	$x = -\frac{1}{\lambda} \ln r$
Нормальное	$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}}$	$F(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-m_x)^2}{2\sigma_x^2}} dx$	$-\infty < m_x < \infty;$ $\sigma_x > 0;$ $D_x = \sigma_x^2$	$R_N = m_x + \sigma_x \left[\sum_{i=1}^{12} r_i - 6 \right]$
Гамма-распределение	$f(x) = \begin{cases} \frac{\sigma^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{-\sigma x}, & x \geq 0; \\ 0, & x < 0 \end{cases}$	$F(x) = \begin{cases} \frac{\sigma^\lambda}{\Gamma(\lambda)} \int_0^x x^{\lambda-1} e^{-\sigma x} dx, & x \geq 0; \\ 0, & x < 0 \end{cases}$	$-\infty < \lambda < \infty;$ $\sigma > 0;$ $m_x = \frac{\lambda}{\sigma};$ $D_x = \frac{\lambda}{\sigma^2}$	Для λ — целых, $\lambda > 0$ $x = -\frac{1}{\sigma} \sum_{i=1}^{\lambda} \ln r_i$

Продолжение табл. 1

Распределение	Плотность вероятности	Функция распределения	Параметры	Алгоритм датчика
Бета-распределение	$f(x) = \begin{cases} \frac{1}{B(\lambda, \eta)} x^{\lambda-1} (1-x)^{\eta-1}, & 0 \leq x < 1; \\ 0, & x < 0, \quad x > 1 \end{cases}$	$F(x) = \begin{cases} 0, & x < 0; \\ \frac{1}{B(\lambda, \eta)} \int_0^x x^{\lambda-1} (1-x)^{\eta-1} \times \\ \times dx, & 0 \leq x < 1; \\ 1, & x \geq 1 \end{cases}$	$\begin{aligned} \lambda &\geq 0; \eta \geq 0; \\ m_x &= \frac{\lambda}{\eta + \lambda}; \\ D_x &= \frac{\eta \lambda}{(\eta + \lambda)^2 \times (\eta + \lambda + 1)} \end{aligned}$	Для λ, η — целых, $\lambda, \eta > 0$ $x = \frac{\sum_{i=1}^{2\lambda} R_{N_i}^2}{\sum_{i=1}^{2\lambda} R_{N_i}^2 \sum_{i=2\lambda+1}^{2\lambda+2\eta} R_{N_i}^2}$
Релея	$f(x) = \begin{cases} \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, & x > 0; \\ 0, & x \leq 0 \end{cases}$	$F(x) = \begin{cases} 1 - e^{-\frac{x^2}{2\sigma^2}}, & x > 0; \\ 0, & x \leq 0 \end{cases}$	$\begin{aligned} \sigma &> 0; \\ m_x &= \sigma \sqrt{\frac{\pi}{2}}; \\ D_x &= \sigma^2 \left(2 - \frac{\pi}{2} \right) \end{aligned}$	$x = \sqrt{-2\sigma^2 \ln r}$
Лапласа	$f(x) = \frac{\lambda}{2} e^{-\lambda(x-\mu)}$	$F(x) = \begin{cases} \frac{1}{2} e^{-\lambda(x-\mu)}, & -\infty < x \leq \mu; \\ 1 - \frac{1}{2} e^{-\lambda(x-\mu)}, & \mu < x < \infty \end{cases}$	$\begin{aligned} \lambda &> 0; \\ -\infty &< \mu < \infty; \\ m_x &= \mu; \\ D_x &= \frac{2}{\lambda^2} \end{aligned}$	$x = \mu - \frac{1}{\lambda} \ln \left(\frac{r_1}{r_2} \right)$

Распределение	Плотность вероятности	Функция распределения	Параметры	Алгоритм датчика
Эрланга	$f(x) = \begin{cases} \frac{(\sigma \cdot x)^{m-1} e^{-\sigma x}}{\frac{1}{\sigma} (m-1)!}, & x > 0; \\ 0, & x \leq 0 \end{cases}$	$F(x) = \begin{cases} 1 - e^{-\sigma x} \left[\sum_{k=0}^{m-1} (\sigma \cdot x)^k / k! \right], & x > 0; \\ 0, & x \leq 0 \end{cases}$	$\begin{aligned} \sigma &> 0; \\ m &> 0 - \text{целое;} \\ m_x &= \frac{m}{\sigma}; \\ D_x &= \frac{m}{\sigma^2} \end{aligned}$	$x = \frac{1}{\sigma} \ln \left[\sum_{i=1}^m r_i \right]$
Вейбулла	$f(x) = \begin{cases} \frac{\lambda}{\sigma} \left(\frac{x}{\sigma} \right)^{\lambda-1} e^{-\left(\frac{x}{\sigma}\right)^\lambda}, & x \geq 0; \\ 0, & x < 0 \end{cases}$	$F(x) = \begin{cases} 1 - e^{-\left(\frac{x}{\sigma}\right)^\lambda}, & x \geq 0; \\ 0, & x < 0 \end{cases}$	$\begin{aligned} \sigma &> 0; \\ \lambda &> 0, \\ m_x &= \sigma \cdot \Gamma \times \\ &\times \left(\frac{1}{\lambda} + 1 \right); \\ D_x &= \sigma^2 \times \\ &\times \left\{ \Gamma \left(\frac{2}{\lambda} + 1 \right) - \right. \\ &\left. - \left[\Gamma \left(\frac{1}{\lambda} + 1 \right) \right]^2 \right\} \end{aligned}$	$x = \sigma \left[-\ln r \right]^{\frac{1}{\lambda}}$

Примечание. $\Gamma(\lambda)$ — гамма-функция, $\Gamma(\lambda) = \int_0^\infty y^{\lambda-1} e^{-y} dy$; $B(\lambda, \eta)$ — бета-функция, $B(\lambda, \eta) = \frac{\Gamma(\lambda) \cdot \Gamma(\eta)}{\Gamma(\lambda + \eta)}$.

2. Основные дискретные распределения случайных величин

Распределение	Функция распределения	Функция вероятности	Параметры	Алгоритм датчика
Равновероятное	$F(x) = \begin{cases} 0, & x \leq x_1; \\ \sum_{j: x_j \leq x} x_j, & x_1 < x \leq x_n; \\ 1, & x > x_n \end{cases}$	$P(x_j) = \frac{1}{n},$ $j = 1, 2, \dots, n$	$m_x = \frac{1}{n} \sum_{j=1}^n x_j;$ $D_x =$ $= \frac{1}{n} \sum_{j=1}^n (x_j - m_x)^2$	—
Биномиальное	$F(x) = \sum_{k < x} C_n^k p^k (1-p)^{n-k}$	$P_n(x = k) =$ $= C_n^k p^k (1-p)^{n-k}$	$p > 0;$ $n > 0 \text{ — целое};$ $m_x = np;$ $D_x = np(1-p)$	$x = \sum_{i=1}^n a_i;$ $a_i = \begin{cases} 1, & r_i < p; \\ 0, & r_i \geq p \end{cases}$

Распределение	Функции распределения	Функции вероятности	Параметры	Алгоритм датчика
Геометрическое	$F(x) = 1 - (1 - p)^k$	$P(x = k) = p(1 - p)^{k-1}$	$p > 0;$ $m_x = \frac{1}{p};$ $D_x = \frac{1 - p}{p^2}$	$x = \left\lceil \frac{\log r}{\log(1 - p)} \right\rceil + 1$
Гипергеометрическое	$F(x) = \sum_{k < x} \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$	$P(x = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$	$n, N, M > 0;$ N, M, n — целые; $m_x = \frac{nM}{N};$ $D = \frac{nM(N-M)(N-n)}{N^2(N-1)}$	$x = \sum_{i=1}^n a_i;$ $a_i = \begin{cases} 1, & r_i < p_i; \\ 0, & r_i \geq p_i; \end{cases}$ $p_{i+1} = \frac{N_i p_i - a_i}{N - i};$ $p_1 = \frac{x}{N};$ $N_1 = N; N_i = N - i + 1$

номерно распределенных на $[0, 1]$, и подсчитывают количество тех из них, которые меньше p . Это число и является моделируемой случайной величиной.

Кроме указанных методов моделирования дискретных величин используют и другие, основанные на специальных (в частности, асимптотических) свойствах моделируемых распределений и на связи между различными распределениями [68].

Основные дискретные распределения и алгоритмы их имитации приведены в табл. 2.

3. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ИМИТАЦИОННЫХ МОДЕЛЕЙ

ТРЕБОВАНИЯ К ПРОГРАММНОМУ ОБЕСПЕЧЕНИЮ ИМИТАЦИОННЫХ ИССЛЕДОВАНИЙ

Инструментальную основу имитационных исследований составляет совокупность программ для ЭВМ, в виде которых реализованы как сами модели, отражающие процесс функционирования моделируемого объекта, так и соответствующие методы организации и проведения исследования на ЭВМ.

Разнообразие задач, возникающих на различных этапах имитационных исследований, и их сложность приводят к большому набору возможных методов реализации программного обеспечения решения этих задач. Кроме того, требования, предъявляемые к программному обеспечению, часто бывают противоречивы, что также приводит к увеличению числа методов реализации программ моделирования. Условно эти требования можно разделить на следующие основные группы: функциональные; технологические; эксплуатационные; экономические.

Группа функциональных показателей характеризует свойства программного обеспечения выполнять основные функции, возникающие при решении поставленных задач. Так, для программного обеспечения имитационных исследований эффективности технических систем к числу показателей этой группы следует отнести: степень кон-

цептуальности; удобство средств сбора, анализа и вывода данных моделирования; наличие средств имитации случайных процессов; наличие средств установки начального состояния модели; универсальность в применении.

Показатели технологичности программного обеспечения характеризуют его свойства, обуславливающие оптимальное распределение затрат ресурсов и времени на разработку, ввод в эксплуатацию, поддержание работоспособности и сопровождение программного обеспечения.

При анализе программного обеспечения имитационных исследований технологичность этих средств характеризуют показателями эффективности разных этапов процесса создания компонент программного обеспечения (программирования, отладки, комплексирования программ моделей сложной структуры и т. д.), модифицируемостью программ. К числу показателей технологичности относят также такие показатели, как структурируемость программного обеспечения и его стандартизацию.

Эксплуатационные показатели характеризуют приспособленность программного обеспечения к эксплуатации, которая заключается в простоте подготовки к работе, сопровождении программного обеспечения, удобстве решения задачи моделирования на ЭВМ. К этой группе показателей можно отнести следующие: эффективность машинных программ моделей; степень автоматизации процесса моделирования; совместимость с другими видами математического и программного обеспечения ЭВМ; мобильность прикладного программного обеспечения.

Экономические показатели характеризуют затраты на разработку и эксплуатацию программного обеспечения. К их числу относят: стоимость проектирования программ, их кодирования и отладки, документирования и стоимость эксплуатации с учетом модификаций.

Рассмотрим более подробно смысл основных из перечисленных выше показателей качества программного обеспечения решения задач исследования эффективности технических систем методами имитационного моделирования.

Функциональные показатели. Степень концептуальности программного обеспечения отражает удобство представления существенных черт конкретных задач моделирования в терминах соответствующих программных средств имитационного исследования, например, исходных данных и результатов моделирования, соответствия структуры модели структуре объекта исследования. Концептуальность программного обеспечения облегчает как процесс проведения пользователем модели имитационных исследований, так и процесс построения соответствующей имитационной модели. Требование концептуальности относится скорее к этапу разработки модели, чем к ее программной реализации. Однако возможность программирования модели в более общих и содержательных терминах, достигаемого при хорошей концептуальности программных средств реализации модели, облегчает как сам процесс программирования, так и обеспечивает получение более обозримой, а следовательно, и более простой в отладке программы.

Важным свойством программного обеспечения имитационных исследований является удобство описания структуры исследуемого объекта. Применительно к имитационным моделям вводят понятия их статической и динамической структур.

Статическая структура имитационной модели представляет собой не зависящую от времени конструкцию, в рамках которой определены состояния моделируемого объекта, т. е. возможные формы его существования. Поскольку при имитационном исследовании используются численные модели, то и состояние моделируемого объекта отражается в модели различными наборами чисел. Следовательно, программная реализация модели должна включать удобные средства определения различных массивов переменных, состав и числовые значения которых могут изменяться в процессе моделирования. Возможности программного обеспечения по изменению состава модели (генерирование новых элементов, а следовательно, изменение их числа), удобство описания характеристик элементов моделируемого

объекта и их связей друг с другом характеризуют качество рассматриваемого программного обеспечения с точки зрения отражения статической структуры модели.

Описание динамической структуры модели заключается в представлении динамических процессов, протекающих в моделируемом объекте. Поскольку имитационная модель отражает поведение объекта во времени, одним из важнейших аспектов реализации программы любой имитационной модели является решение задачи представления и управления течением времени в программе модели. Кроме того, важной задачей, тесно связанной с первой, является обеспечение согласованности функционирования различных элементов объекта моделирования во времени.

Особенностью отражения последовательности течения времени при машинной имитации является тот факт, что функционирование модели на ЭВМ протекает в искусственном (системном) времени. Хотя компоненты реального моделируемого объекта могут функционировать одновременно, компоненты машинной модели работают строго последовательно. Поэтому при программной реализации возникает еще и задача построения механизма задания времени для синхронизации действий компонент объекта на любом рассматриваемом интервале времени. В имитационных моделях принято два основных метода задания системного времени: с помощью фиксированных и переменных интервалов времени.

К числу важных функциональных характеристик программного обеспечения имитационных исследований следует отнести также возможности программного обеспечения по сбору, анализу и выводу данных моделирования. Исследование эффективности технических систем можно проводить либо на основе прослеживания динамики их функционирования по графикам траекторий имитируемых значений переменных моделей, либо по агрегированным характеристикам поведения системы. В последнем случае для построения этих характеристик используют методы статистического анализа результатов численного экс-

перимента. Качество программного обеспечения имитационных исследований в этом аспекте определяется насыщенностью соответствующего программного обеспечения средствами автоматического сбора, анализа и вывода данных. Качество аппарата сбора и вывода данных моделирования тем выше, чем меньше усилий требуется для программирования соответствующих заданий ЭВМ.

К числу стандартных данных численного эксперимента, которые, как правило, используются экспериментатором при исследовании эффективности технических систем по данным имитационного эксперимента, относят: число наблюдений, максимумы и минимумы для всех переменных; суммы квадратов для переменных, не зависящих от времени; взвешенные суммы и суммы квадратов переменных, зависящих от времени; гistogramмы распределения значений для переменных, не зависящих от времени; гistogramмы распределения кодов состояний для переменных, зависящих от времени; значения временных рядов через заданные интервалы времени.

Минимальный набор статистических характеристик, подпрограммы вычисления которых должны входить в состав программного обеспечения имитационных исследований, включает: и средние значения, дисперсии и средние квадратические отклонения, частотные распределения.

Одной из важнейших функциональных характеристик программного обеспечения является наличие в нем средств имитации случайных процессов. Аппарат генерирования случайных величин включает датчики случайных (псевдослучайных) чисел и стандартные программы преобразования этих чисел в случайные величины с различными статистическими распределениями. Методы моделирования случайных факторов, на основе которых работают программы генерации (датчики) псевдослучайных чисел, были описаны на с. 102.

Программное обеспечение имитационных исследований должно обеспечивать построение статистически приемлемых последовательностей

псевдослучайных чисел. В [46] сформулированы требования, которым должен удовлетворять идеальный генератор. Считается, что полученные с его помощью последовательности должны состоять из (1) равномерно распределенных, (2) статистически независимых, (3) воспроизводимых, (4) неповторяющихся чисел, (5) генератор должен работать быстро и (6) занимать минимальный объем машинной памяти.

Поскольку реализуемые на ЭВМ методы построения псевдослучайных чисел детерминированы, последовательности псевдослучайных чисел полностью воспроизводимы при задании одинаковых начальных значений. Возможность управления работой стохастических процедур модели имеет большое значение при проведении статистических испытаний модели. При вычислении стохастических показателей эффективности технических систем иногда возникает необходимость воспроизводить одни и те же последовательности псевдослучайных чисел или использовать в модели несколько независимых датчиков. На возможности управления генерированием псевдослучайных чисел основаны так называемые процедуры уменьшения дисперсии имитационного эксперимента. Использование специальных процедур воспроизведения имитации позволяет уменьшить разброс оценок характеристик системы, вычисленных по результатам статистической имитации или, что равнозначно, сократить объем имитационного эксперимента для получения заданной точности. Поэтому свойство управляемости работы стохастических процедур является важным требованием к соответствующему программному обеспечению.

Технологические показатели. Среди показателей качества программного обеспечения этой группы существенную роль играют показатели, характеризующие средства отладки программного обеспечения. Как правило, программы моделирования обладают сложной динамикой работы и крайне сложны и неудобны для отладки (передача управления в имитационных программах непрогнозируема). К числу средств отладки, которые должны входить в состав программного обеспечения ими-

тационного моделирования, следует отнести: привязку ошибок, выявленных в процессе компиляции и счета, к тексту исходной программы; подробный вывод состояния программы в момент обнаружения ошибки (подразумевается перечень входных точек и соответствующих параметров всех функций и подпрограмм, вызванных в этот момент); доступ к информации управляющей программы (это позволяет проследивать ход всех или части событий и использовать управляющую информацию для диагностики ошибок).

Этими же средствами можно пользоваться и для контроля прохождения программ во время их эксплуатации. Анализ очередности событий при имитации, происходящих при различных входных параметрах, является одной из форм имитационного исследования динамики функционирования сложных технических систем. Поэтому средства трассировки работы имитационной программы являются необходимым элементом программного обеспечения имитационных исследований.

Эффективность создания имитационных программ во многом определяется также степенью стандартизации применяемых программных средств. При разработке имитационных программ сложной структуры существенную роль играют такие факторы, как модульность программ, наличие средств комплексирования.

Эксплуатационные показатели. Для эксплуатации имитационных программ и, в первую очередь, программ много-разового использования наиболее существенную роль играет показатель эффективности выполнения программы. Программа тем более эффективна, чем она более быстродействующая и чем меньше занимает объем оперативной памяти ЭВМ.

При проведении имитационных исследований важным фактором становится степень автоматизации выполнения основных операций машинных экспериментов. Одним из путей решения этой задачи может быть использование средств системного программного обеспечения ЭВМ по выполнению стандартных операций обработки информации. К их числу обычно относят: банки данных, генераторы программ

отчетов, генераторы программ ввода данных и т. п. В связи с этим важным эксплуатационным свойством программных средств имитации становится их совместимость с системным программным обеспечением ЭВМ. Для достижения этого свойства программное обеспечение имитации должно быть открытым (допускать подключение других программных средств) и удовлетворять стандартам на связи между программами (по управлению, данным и т. п.).

В процессе эксплуатации программного обеспечения имитационных исследований может возникнуть необходимость переноса комплекса программ с одной ЭВМ на другую. Возможность такого переноса без значительных программных переделок и доработок принято называть *мобильностью программного обеспечения*.

Выбор рассмотренных факторов, определяющих основные требования, предъявляемые к программному обеспечению имитационных исследований, является неоднозначным. Анализ программных средств проводят и с несколько других точек зрения [73]. В конечном счете доминирующее влияние той или иной группы факторов определяет специфика решения конкретной задачи. Кроме того, перечисленные выше требования к программному обеспечению не являются независимыми друг от друга.

ВИДЫ ПРОГРАММНОЙ РЕАЛИЗАЦИИ ИМИТАЦИОННЫХ МОДЕЛЕЙ

Принято выделять три вида программного обеспечения, наиболее часто применяемого для задач моделирования: машинно-ориентированные и алгоритмические языки программирования; специализированные языки моделирования (в том числе и расширенные системы программирования для моделирования); проблемно-ориентированные пакеты программ моделирования.

Выделенные виды программных средств в определенной степени характеризуют уровни и этапы развития программного обеспечения имитационных исследований. Вопросы класси-

фикации средств программирования имитационных моделей, сравнительная характеристика языков моделирования, основанная на опыте их применения, подробно рассмотрены в [46, 60, 73]. Дадим краткую характеристику каждого из указанных видов программного обеспечения.

Машинно-ориентированные и универсальные языки программирования. Программное обеспечение, реализованное на машинно-ориентированных языках (автокоде, Ассемблере), было разработано на первом этапе использования ЭВМ для решения задач моделирования (при отсутствии более удобных средств программирования). Машинно-ориентированные языки используют в тех случаях, когда главным требованием к программной реализации модели является эффективность ее функционирования (минимальные время решения задачи на ЭВМ и объем оперативной памяти, занимаемой программой). В последнем случае считается, что потери, возникающие из-за неэффективности программного обеспечения моделирования, будут более значительны, чем затраты труда и времени на его разработку.

В большинстве практических ситуаций использование машинно-ориентированных языков для разработки прикладного обеспечения считается неразумным [73]. Наиболее рациональным следует признать использование машинно-ориентированных языков в качестве инструмента лишь программистов-профессионалов для построения сложных и ответственных системных программ или для реализации тех фрагментов имитационных моделей или систем, которые в наибольшей степени требуют оптимизации. В этом случае возможен подход, когда первоначальная версия программы пишется на языках высокого уровня (универсальных или специализированных), затем определяются оптимизируемые части программ, которые и переписываются уже на машинно-ориентированном языке. Большинство языков высокого уровня позволяет пользователю писать подпрограммы на языке Ассемблера для использования возможностей, не предусмотренных или

трудно программируемых в данном языке.

Значительное число задач моделирования все еще реализуется на универсальных алгоритмических языках программирования (Фортран, ПЛ/1 и др.), что характерно для текущего этапа программирования прикладных задач. Алгоритмические языки являются машинно-независимыми, что делает написанные на них программы более мобильными, чем при использовании машинно-ориентированных языков. Как правило, программы, реализованные на Фортране, ПЛ/1 и других языках высокого уровня, воспринимаются легче, чем программы на Ассемблере или автокоде. Процесс разработки программы становится более эффективным, так как, во-первых, операторы языка высокого уровня являются более мощными, чем команды Ассемблера, и, во-вторых, эти языки, как правило, имеют более развитые средства отладки и документирования программ. При наличии оптимизирующих компиляторов основных языков высокого уровня вопрос о более высокой эффективности программы, написанной на Ассемблере, снимается.

Машинно-ориентированные и универсальные языки программирования высокого уровня, как средство построения имитационных моделей, позволяют обеспечить, в первую очередь, высокие эксплуатационные показатели качества создаваемого программного обеспечения имитационных исследований. Однако у них следует отметить низкие функциональные, технологические и, как следствие, экономические показатели. В частности, общим недостатком машинно-ориентированных и универсальных языков программирования при использовании их в качестве средств разработки имитационных моделей является то обстоятельство, что в пределах этих средств программирования нет аппарата для организации автоматического сбора, анализа и вывода результатов моделирования, организации имитации (управления «модельным временем») и машинного эксперимента в целом. В связи с этим при разработке программ имитационных моделей помимо разработки собственно модели процесса дополни-

тельно пользователь должен решать и большое число трудоемких обеспечивающих задач.

Машинно-ориентированные и универсальные языки высокого уровня целесообразно использовать лишь для программирования небольших задач моделирования, когда применение других средств по какой-либо причине невозможно или нецелесообразно (отсутствие более эффективных средств в составе математического обеспечения используемой ЭВМ, требования по стыковке с компонентами других программных систем, в составе которых будет использоваться разрабатываемая модель). Указанные средства могут использоваться и для программирования обеспечивающих и расчетных функций в рамках больших имитационных систем.

Специализированные языки моделирования. Основное отличие большинства языков моделирования от универсальных языков программирования состоит в обеспечении ими решения тех специфических задач, которые возникают при имитационном исследовании.

В зависимости от используемых при динамической композиции моделируемых систем понятий принята и современная классификация языков моделирования: классы языков, ориентированные, соответственно, на работы, события или процессы. Часто выделяют еще один класс языков моделирования: языки, ориентированные на транзакты. В этих языках моделирование системы проводится путем движения транзактов (сообщений) через специализированные блоки действий модели. Специализированные блоки соединяются в структуры, отражающие логику действий и различные потоки моделируемой системы. Моделирующая программа создает (генерирует) транзакты, которые затем во времени проходят через блоки системы, воспроизводя логику функционирования моделируемой системы. Транзактные языки можно отнести к языкам процессов [46].

Любой объект, описанный в терминах одной из рассмотренных схем, можно описать и на основе двух других. Речь идет лишь об удобстве и эффективности представления моделируемой

системы. Сформулируем некоторые рекомендации по применению того или иного из рассмотренных способов представления динамики моделируемой системы [46].

Язык моделирования, основанный на механизме расписания событий, обеспечивает экспериментатору строгий контроль прохождения программы.

Язык работ удобен для описания систем с большим числом разнообразных ресурсов, так как он позволяет записывать условия их доступности в компактной форме.

Объединение нескольких событий в одну программу, допустимое языком процессов, облегчает описание динамики системы. При этом, с одной стороны, сокращается число необходимых «вспомогательных» операторов, а с другой — проясняется общая динамическая структура системы, поскольку логика каждого процесса сосредотачивается в одной программе.

Расширенные системы программирования для моделирования. Весьма перспективным направлением в развитии программных средств построения имитационных моделей является создание систем программирования для имитационного моделирования по принципу расширения некоторой базовой системы. В качестве базового языка таких систем применяют один из развитых универсальных алгоритмических языков программирования, а в качестве надстройки — используют соответствующие операторы одного из языков моделирования. Такой подход позволяет использовать наиболее сильные стороны каждого из рассмотренных средств разработки имитационных моделей. Как правило, в состав таких систем включают и общеалгоритмические средства (средства динамического управления распределением памяти, средства работы с различными типами данных, различные операторы ввода-вывода, средства доступа к адресам размещения данных и т. п.), и специфические средства языков моделирования, позволяющие эффективно описывать и управлять элементами моделируемого объекта (средства описания классов элементов, управление модельным временем и

синхронизация взаимодействия элементов). Примерами реализации рассмотренных средств программирования имитационных моделей являются: Дискретно Ориентированная Система Имитационного Моделирования (ДИОСИМ), разработанная на базе языков ПЛ/1, НЕДИС и СИМУЛА-67; система моделирования дискретных процессов (ДИСМ); ПЛ/1 Имитационная Система (ПЛИС), созданная на базе языков ПЛ/1 и GPSS, и т. п.

Проблемно-ориентированные пакеты программ моделирования. Пакеты программ моделирования разрабатывают, как правило, для некоторого формального класса задач со сравнительно узкой областью применения. Существует несколько типов пакетов программ моделирования. Наряду с собственно программами имитации в состав пакета входят и сервисные программы, облегчающие работу пользователя. К числу сервисных относятся программы ввода исходных данных о моделируемом объекте, программы преобразования описания объекта моделирования к стандартной формальной схеме, принятой в рамках данного пакета, программы обработки данных имитационного эксперимента. В частности, в сервисные программы могут быть включены программы-генераторы, которые по описанию задачи на формализованном входном языке генерируют программы имитации. Общим для пакетов программ моделирования является то, что пользователь для решения задачи использует готовое программное обеспечение, настраивая его на решение своей задачи.

При условии, что задача исследования хорошо формализуется в рамках математического аппарата пакета, такое моделирование может существенно сократить время на подготовку задачи к счету на ЭВМ. Недостатком этого подхода является, как правило, сравнительно узкая область применения каждого проблемно-ориентированного пакета программ. Чтобы расширить область применения пакета, часть его программного обеспечения делают открытым, т. е. допускается возможность включения в его состав и программ пользователей, реализующих

нестандартные операции по моделированию.

Интересным примером построения проблемно-ориентированных пакетов программ моделирования является пакет, получивший название универсальной автоматизированной имитационной модели (УАИМ) [9]. Основная концепция УАИМ предполагает переложение на ЭВМ большей части работ, выполняемых при подготовке к моделированию и при проведении исследований на модели. Автоматизация основных процессов имитационного исследования сложных технических систем (описание объекта моделирования, построение моделирующего алгоритма, его программирование, организация вычислений на ЭВМ) осуществляется на базе заранее разработанных программ, объединенных в единый программный комплекс.

Средства УАИМ позволяют настраиваться на моделирование любого объекта, формализуемого в рамках используемого математического аппарата. «Универсальность» УАИМ определяется возможностью моделировать в ее рамках многоуровневые сложные системы достаточно произвольной структуры с элементами, которые являются динамическими системами в широком смысле слова (удовлетворяющих весьма общим предположениям относительно характера функционирования) [9, 10]. Программы имитации конкретного пакета (конкретной АИМ) создаются только для некоторой стандартной формы описания моделируемого объекта и реализуются в виде стандартного программного модуля.

Структурный аспект взаимодействия элементов моделируемой системы в процессе имитации описывается моделью сопряжения элементов сетью идеальных каналов связи. Реализация УАИМ осуществляется таким образом, что структурные и функциональные характеристики моделируемых объектов входят не в описание ее функций и структуры, а являются исходными данными для моделирования. Это обеспечивает гибкость пакета и возможность исследования в его рамках достаточно больших систем (описание характеристик элементов моделируе-

мой системы хранится во внешней памяти ЭВМ).

Кратко суть механизма имитации средствами рассматриваемого пакета состоит в следующем. Имитация процесса функционирования сложной системы представляется совокупностью следующих операций: имитацией функционирования элемента системы, имитацией взаимодействия между элементами и управлением очередностью системных событий. Имитация функционирования элемента системы осуществляется путем настройки стандартного модуля модели (представляющего собой, например, общую динамическую систему с дискретным вмешательством случая) на исходный элемент системы и реализации на ЭВМ полученной модели элемента. Настройка модуля на данный элемент заключается в поиске в архиве во внешней памяти ЭВМ строки, соответствующей данному элементу, и считывании параметров элемента в оперативную память. Реализация на ЭВМ полученной модели состоит в определении момента выхода на границу пространства состояний динамической системы, формализующей элемент, реализации скачка состояния в момент выхода на границу или в момент поступления сигнала.

Имитация взаимодействия между элементами состоит в определении адресов передачи и содержания выходного сигнала имитируемого элемента и передачи его другим элементам системы в соответствии с указанными адресами. Управление очередностью модельных событий осуществляется по принципу планирования событий с использованием механизма переменного шага, т. е. модуль настраивается на тот элемент системы, очередной опорный момент функционирования которого наиболее близок к рассматриваемому текущему моменту модельного времени.

Подробнее с эксплуатируемыми в настоящее время программами имитации сложных систем, построенными с учетом описанных принципов, можно познакомиться в [9].

4. ПРОБЛЕМНО-ОРИЕНТИРОВАННЫЕ ИМИТАЦИОННЫЕ СИСТЕМЫ ИССЛЕДОВАНИЯ ЭФФЕКТИВНОСТИ

СОСТАВ И СТРУКТУРА ИМИТАЦИОННОЙ СИСТЕМЫ

Развитие методики и средств (технических и программных) имитационного исследования эффективности сложных технических объектов привело к созданию специальных моделирующих систем, получивших название имитационных систем или проблемно-ориентированных имитационных систем, что подчеркивает их ориентацию на обеспечение исследования эффективности объектов определенного типа или класса. Проблемно-ориентированные имитационные системы являются средством проведения имитационных исследований эффективности крупномасштабных операций в технике. В рамках имитационной системы реализуется общая схема операции, описываемая как кибернетическая система. Агрегативный принцип построения имитационной системы, формализация элементов S_0 -системы или определенного фрагмента операции моделью общей динамической системы позволяют исследователю построить кибернетическую систему достаточно высокой сложности, обеспечивающую ее требуемую адекватность моделируемому процессу. Однако построение адекватных моделей, базирующихся на надежной и доброкачественной информации, является необходимым, но не достаточным условием успешного проведения имитационных исследований. Важно также предоставить исследователю широкий набор удобных средств построения моделей, постановки и реализации машинных экспериментов, расчета показателей эффективности и принятия решений. Таким образом, под проблемно-ориентированной имитационной системой в общем случае понимается совокупность моделей, имитирующих протекание изучаемого процесса (или класса процессов), объединенная со специальной системой вспомогательных

программ и информационной базой, позволяющих достаточно просто и оперативно реализовать варианты расчеты [43].

На рис. 2 приведена примерная структура проблемно-ориентированной имитационной системы, отражающая принципиальные стороны ее функционирования. Как правило, проблемно-ориентированная имитационная система предназначена для работы различных категорий пользователей (как по уровню подготовки для работы с ЭВМ, так и по знаниям методов и математического аппарата моделирования). Поэтому имитационные системы должны включать в себя развитый блок обеспечения взаимодействия с пользователем, предоставляющий ему удобный и простой контакт с системой. Очевидно, что пользователь должен получить и исчерпывающую информацию о возможностях системы, ее конструкции, правилах работы и т. п. Обеспечению этих функций служит блок информационно-справочного обеспечения исследователя. Центральной частью имитационной системы является управляющий блок, предназначенный координировать всю работу системы, обеспечивая правильную последовательность этапов имитационного исследования в рамках конкретных задач. Управляющий блок непосредственно связан с блоком обеспечения взаимодействия с пользователем, выполняющим функции программной поддержки человеко-машинного режима функционирования имитационной системы. Основу имитационной системы составляют собственно модели и информация об объектах исследования, объединенные, соответственно, в Банках моделей и модулей системы, а также в Банке данных (информационном блоке системы). Управляющий блок имитационной системы координирует работу с моделями и данными через соответствующие блоки обслуживания Банков моделей, модулей и данных. Имитационная система должна обеспечивать решение достаточно широкого класса задач в своей проблемной области. Банк модулей формируется из стандартных модулей, имитирующих функционирование элементов исследуемых тех-

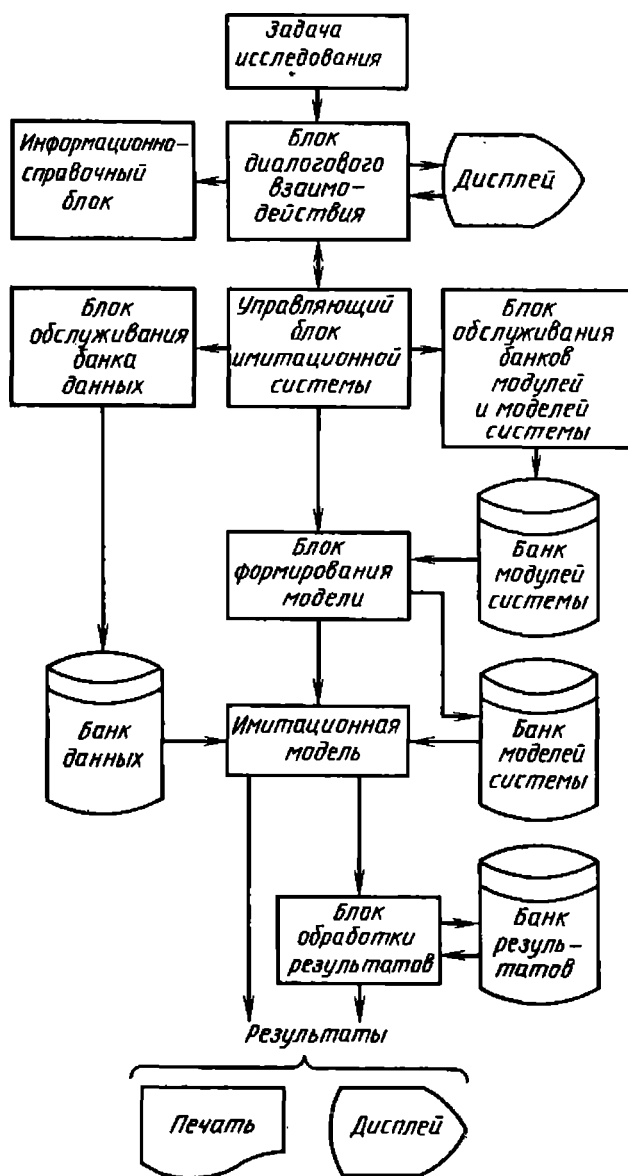


Рис. 2. Структура проблемно-ориентированной имитационной системы

нических систем. Для построения конкретной имитационной модели служит блок формирования модели.

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ИМИТАЦИОННОЙ СИСТЕМЫ

Реализация имитационной системы на ЭВМ связана с разработкой большой программной системы (БПС), процесс создания которой качественно отличается от создания автономных программ решения небольших задач.

При разработке БПС необходимо: 1) на первом этапе разработки проблемно-ориентированной системы (ПО-системы) провести архитектурное ее проектирование и представить систему как комплекс двух подсистем: функ-

циональной (проблемно-ориентированной) и управляющей, инвариантной относительно классов решаемых задач;

2) установить основополагающие приемы управления данными в ПО-системе на основе интегрированной базы данных;

3) разрабатывать программное обеспечение ПО-системы в рамках единого технологического процесса;

4) формализовать совокупность знаний и представлений о классе решаемых задач в виде модели предметной области (МПО) и разработать приемы и программные средства построения функциональных программ системы на базе МПО;

5) решить вопросы организации исследований в рамках ПО-имитационной системы, т. е. вопросы планирования машинных экспериментов.

При построении программного обеспечения имитационной системы:

1) определяют требования к создаваемой системе, общую архитектуру ПО-имитационной системы (этап архитектурного проектирования), структуру основных блоков системы на основе модульного анализа алгоритмов (этап технического проектирования);

2) выбирают и (или) создают технологические и инструментальные средства ПО-имитационной системы;

3) с использованием технологических и инструментальных средств системы реализуют построение функциональных компонент (этот этап включает кодирование стандартных модулей системы и комплексную отладку моделей);

4) документируют систему и осуществляют пробную эксплуатацию.

Определенные этапы можно выполнять параллельно, в частности, документирование системы ведется, как правило, на всех этапах.

С учетом особенностей исследования эффективности технических систем и основных положений методологии имитационного исследования к числу основных требований, предъявляемых к ПО-имитационной системе, целесообразно отнести следующие:

открытость системы, обеспечивающую возможность ее развития для решения новых задач;

адаптивность системы к задачам исследования и категории ее пользователей;

согласование моделей системы по информации, показателям и критериям эффективности;

соответствие структуры системы моделей уровням обобщения информации и уровням решаемых задач;

обеспечение заданной точности моделирования и практической реализуемости исследований с учетом приемлемого времени решения на ЭВМ;

обеспечение удобных режимов работы в системе, основанных на интерактивном характере решения задач.

Для выполнения сформулированных требований в основу построения программного обеспечения имитационной системы могут быть положены следующие основные принципы:

многомодельность системы, означающая наличие в имитационной системе множества моделей, а в рамках моделей — множества их программных реализаций;

многовариантность компонент системы, понимаемая как многократность модификаций существующих программ, отражающая вновь принимаемые решения о методах расчета или организации программы;

модульность программного обеспечения системы;

иерархичность построения системы; единство математических моделей формализации исследуемых процессов и типизация на этой основе функциональных структур моделей;

единство информационного представления данных о характеристиках исследуемых технических систем в моделях;

параметризация ряда характеристик технических систем как средство решения задачи компоновки модели сложного объекта на основе информации, получаемой на более подробных моделях его подсистем.

При проектировании программ моделей имитационной системы целесообразно выделять два уровня: архитектурное и техническое. Архитектурное проектирование подразумевает выделение основных блоков программ моделей решения задачи и определение их взаимоподчиненности и связей. Под

блоком программы понимается совокупность модулей одного функционального назначения. Техническое проектирование заключается в структуризации каждого блока. Структуризация включает в себя выделение основных модулей блока, определение их взаимосвязи по управлению и данным, задание логики функционирования. При этом целесообразно стремиться уже на первых шагах этапа технического проектирования определить, по-возможности, окончательную структуру блока, которая в наибольшей степени отражала бы существо моделируемого программой блока процесса. Для любого функционального модуля в составе блока может быть разработано несколько программных реализаций, различающихся методом решения, полнотой и глубиной формализации рассматриваемого процесса (принцип многовариантности компонент системы). Переход к другому варианту блока (модели) будет осуществляться путем замены одного или нескольких модулей, реализующих более полное для конкретного варианта блока (модели) или, наоборот, более упрощенное описание моделируемого процесса.

Рассмотренный подход к представлению и формированию компонент прикладного программного обеспечения имитационной системы обеспечивает довольно ясную методику как построения программ модулей системы, так и последующего ее развития для решения новых задач на основе новых вариантов моделей (требование открытости системы и реализация принципа ее многомодельности). Такая методика представления компонент системы отражает ту особенность имитационного исследования, которая заключается в возможности корректировки постановок задач в процессе исследований, что приводит к необходимости постоянного видоизменения функционального программного обеспечения, т. е. порождает эволюционный характер программного обеспечения имитационной системы.

Существенное достоинство принятого подхода — создание на ранней стадии разработки системы некоторого работоспособного варианта и увеличе-

ние полноты этого варианта с каждым последующим шагом развития системы. Возможность эксплуатации еще не полностью готовой системы позволяет совместить стадии ее освоения и разработки так, что срок включения системы в эксплуатацию (практического использования системы) может быть существенно сокращен.

Выше рассмотрены, в основном, методические аспекты технологии построения проблемно-ориентированной имитационной системы. Однако технология разработки сложных программных систем имеет еще один аспект — инструментальный. С учетом таких характеристик имитационной системы, как ее открытость и определенная универсальность в рамках решаемого класса задач, инструментальным аспектам технологии построения и функционирования проблемно-ориентированной имитационной системы должна отводиться особенно важная роль, поскольку качество и полнота инструментальных средств во многом определяют эксплуатационные характеристики системы.

Инструментальные средства имитационной системы, реализованные в виде программ, организационно объединены в рамках специального системного программного обеспечения имитационной системы (ССПО ИС). Основная цель создания ССПО ИС состоит в предоставлении пользователю системы удобных средств работы для исследования и решения конкретной проблемы. Можно выделить следующие программные средства ССПО ИС:

- обеспечивающие исследователя информацией о возможностях системы и методиках работы в ее информационной и программной средах;

- позволяющие ускорить и по возможности автоматизировать процесс разработки конкретного варианта модели;

- упрощающие процесс подготовки необходимой исходной информации для решения задачи;

- обеспечивающие рациональные режимы проведения исследований на базе моделей имитационной системы.

Информационно-справочное ССПО. С целью повышения уровня автоматизации основных этапов работы в рам-

ках системы и эффективности ее использования необходимо обеспечить пользователя системы информацией: 1) о возможностях имитационной системы для исследования эффективности конкретной операции (технической системы); 2) о методике решения поставленной задачи. Наряду с предоставлением пользователю информации в виде обычных справок современные программные системы должны обеспечивать и режимы автоматизированного обучения с оперативной выдачей пользователю (обучаемому) оценок и коррекции его ответов на поставленные вопросы. В качестве обучающей тематической среды используют тестовые примеры решения типовых задач исследования эффективности основных для данной имитационной системы объектов моделирования на всех или наиболее сложных для освоения этапах исследования эффективности операций. Режимы оперативного взаимодействия пользователя с имитационной системой поддерживаются соответствующими программными средствами.

ССПО автоматизации разработки и конструирования программ. К средствам автоматизации разработки конкретного варианта модели относятся средства кодирования компонент моделей и средства их комплексирования. Средства кодирования объединены в базовую подсистему программирования, основу которой могут составить как универсальные, так и специализированные языки программирования.

Базовую подсистему программирования используют для формирования компонент Банка модулей системы на этапе создания проблемно-ориентированной имитационной системы, а также на этапах эксплуатации системы при решении конкретных задач при отсутствии необходимых компонент программного обеспечения.

Одной из важнейших задач для многоэлементной программной системы является задача автоматизации комплексирования ее составляющих в виде законченных и единых программных единиц (задач или заданий для ЭВМ). При многоуровневом построении программной имитационной системы целе-

сообразно эту задачу решать на нескольких уровнях. В частности, можно выделить несколько уровней комплексирования в имитационной системе и несколько соответствующих им классов программных компонент системы.

Модули-макроопределения. Это низший уровень, реализующий сборку программ моделей средствами базовой подсистемы программирования. Например, возможность ввода в язык программирования новых дополнительных классов объектов, в частности, динамических [новых видов транзактов со специальными подпрограммами обработки типовых (для конкретно решаемой задачи) ситуаций моделирования] позволяет повысить проблемную ориентацию подсистемы программирования и тем самым ускорить процесс «допрограммирования» компонент имитационной системы.

Программные модули, реализующие типовые математические схемы формализации моделируемых элементов системы (например, кусочно-линейный агрегат, вероятностный автомат и т. п.) На этом уровне комплексирование может быть реализовано в рамках стандартных схем сопряжения.

Подмодели. Подразумевается формирование модели сложного технического объекта при условии разбиения процесса его функционирования на ряд подпроцессов, описываемых самостоятельно. При этом подпроцессы могут развиваться во времени последовательно либо параллельно. При последовательном соединении подпроцессов связь подмоделей осуществляется по входам и выходам в соответствии с заданным (сформированным) сценарием имитации. При параллельном соединении подпроцессов, как правило, внутри подпроцессов должны быть предусмотрены и точки их активации и реактивации, позволяющие переключаться с одного подпроцесса на другой. Соответственно и программа управления имитацией будет несколько сложнее, чем для последовательного соединения подпроцессов.

Модели и расчетные модули. Имитационное моделирование кроме непосредственной имитации включает и операции как предшествующие ими-

тации (подготовка данных), так и последующие (сбор статистических данных, их обработка, расчет показателей эффективности и т. д.). В целях автоматизации процесса исследования на ЭВМ соответствующие обеспечивающие, имитационные и расчетные программы должны быть скомплексированы и состыкованы.

Поставленная задача решается программами управления имитационным экспериментом. В качестве базового программного обеспечения для этих целей могут быть использованы программы-мониторы, входящие в состав специального математического обеспечения ЭВМ, базирующиеся на представлении знаний о структурах задач в виде семантических сетей.

ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ИМИТАЦИОННОЙ СИСТЕМЫ

Проблема оптимальной организации данных является одной из важнейших при создании сложных моделирующих комплексов. Решение задачи повышения эксплуатационных характеристик таких комплексов (адаптивность, оперативность и т. п.) находится в прямой зависимости от организации процесса обработки информации. Причем, если на время непосредственного решения задач на ЭВМ в рамках проблемно-ориентированной имитационной системы наибольшее влияние оказывает качество алгоритмов переработки информации, то для повышения таких свойств системы, как ее адаптивность к задачам исследования, обеспечение многовариантности расчетов, уменьшение времени на подготовку задачи моделирования, большая роль принадлежит именно вопросам организации информационного обеспечения имитационной системы. Кроме того, в целях обеспечения открытости и расширяемости функций системы, преемственности в разработке ее компонент требуется информация не только о текущей структуре данных о моделируемых объектах системы, но и сведения о тенденциях развития как самой моделируемой системы, так и системы средств ее моделирования (самого комплекса моделей). Все это вызывает

необходимость в такой организации данных, которая позволяла бы осуществлять первоначальное формирование, модификацию состава и структуры данных, их выборку, не изменяя программного обеспечения имитационной системы.

Теоретической и практической основой решения этой задачи является методология построения *БАЗ ДАННЫХ*, позволяющая осуществлять описание данных на нескольких уровнях и тем самым обеспечивающая независимость программных компонент системы от конкретных структур и состава обрабатываемых данных. Неотъемлемой частью системного программного обеспечения становится система управления базами данных (СУБД), которую в совокупности с базами данных принято называть *БАНК ДАННЫХ*.

Исследование эффективности операции представляет собой процесс обработки информации о моделях исследуемого технического объекта. До проведения имитационного эксперимента исходная информация о моделях объекта исследования содержит сведения о его состоянии и характеристиках (в основном, сведения о состоянии и характеристиках его элементов) и, как правило, не содержит сведения о его эмерджентных (системных) свойствах (или такие сведения нуждаются в уточнении). В частности, речь идет о значениях показателей, характеризующих эффективность операции $W(u)$. В результате проведения имитационного эксперимента формируется дополнительная информация о моделях объекта исследования, содержащая как уточненные сведения о состоянии его подсистем Z , так и необходимые сведения о системных свойствах объекта (значения частных $W_i(u)$ и обобщенных $W(u)$ показателей эффективности операции).

Таким образом, подсистема информационного обеспечения предназначена для организации хранения и оперативной выдачи исходной и дополнительной информации в виде сведений о моделях исследуемого объекта.

В состав подсистемы информационного обеспечения входят: интегрированные базы данных (БД) моделей

объекта исследования; система управления базами данных (СУБД), включающая комплексы: программ ввода данных в базы данных, программ формирования запросов моделей к базам данных системы, программ информационной связи между моделями объекта исследования и программ поддержания и ведения баз данных.

Процесс проектирования подсистемы информационного обеспечения имитационной системы состоит из следующих этапов:

I — информационное обследование объекта исследования;

II — проектирование БД имитационных моделей объекта исследования;

III — выбор базовой СУБД с последующей реализацией БД подсистемы информационного обеспечения и реализацией комплексов программ обслуживания баз данных.

Информационное обследование моделируемой технической системы предполагает изучение и определение перечня учитываемых при моделировании характеристик системы, их структуры, взаимосвязи различных подсистем по информации.

Основой проектирования БД является разработка логической структуры базы данных (ЛС БД). Для описания логических структур БД используют различные методы: иерархический, сетевой, реляционный [37]. При разработке логических структур БД имитационных моделей в рамках проблемно-ориентированной имитационной системы исследования эффективности целесообразно использовать реляционный метод.

ЛС БД реляционного типа имеют следующие преимущества перед ЛС БД иерархического и сетевого типов [37]:

а) гибкость; операции добавления и удаления атрибутов требуют минимальных затрат;

б) точность; отношения, представленные в виде реляционных таблиц (РТ), по своей природе обладают более точным смыслом и поддаются математически точным методам манипулирования;

в) простота и надежность защиты данных; для каждой РТ можно задать правомерность доступа;

г) простота ведения; физическое размещение РТ оказывается намного проще, чем размещение иерархических и сетевых структур;

д) независимость данных проще обеспечить в ЛС БД реляционного типа;

е) ясность; ЛС БД иерархического или сетевого типа при изображении связи с помощью стрелок выглядит нагляднее, пока число стрелок невелико; рост БД приводит к такому увеличению числа связей, что их невозможно достаточно четко отразить на схеме.

Атрибуты исходной ЛС БД имитационных моделей описывают:

состав и подчиненность друг другу подсистем объекта моделирования;

характеристики (неизменные свойства) подсистем;

состояния (изменяемые свойства) подсистем;

взаимодействие подсистем в различных состояниях;

системные свойства объекта исследования.

ЛС БД имитационной модели рекомендуется разрабатывать на основе анализа выбранного математического метода имитационного моделирования. В этом случае ЛС БД описывает не возможные реализации (т. е. конкретные имитационные модели) математического метода, а сам этот метод. Такой подход обеспечивает выполнение одного из важнейших принципов информационного обеспечения имитационной системы — независимости данных в БД. Так, если для построения имитационных моделей используют агрегативный подход, согласно которому исследуемый объект представляется в виде системы связанных между собой кусочно-линейных агрегатов (КЛА), то атрибуты исходной ЛС БД имитационных моделей будут описывать: входы, выходы, параметры, состояния КЛА; сведения о программах обработки динамики функционирования КЛА; связи КЛА между собой, по входам и выходам; очередность запуска КЛА.

При выборе базовой СУБД в проблемно-ориентированной имитационной системе исследования эффективности выдвигаются следующие требования к ее характеристикам:

возможность реализации выбранного метода организации ЛС БД (в частности, реляционного метода);

надежность в эксплуатации;

возможность описания данных и манипулирования данными на логическом уровне;

наличие развитых средств защиты данных;

возможность работы в БД из прикладных программ имитационной системы (открытый тип СУБД).

Реализация БД осуществляется в рамках СУБД и заключается в определении физической структуры БД (т. е. в определении способов физического хранения данных и доступа к ним во внешней памяти ЭВМ — дисках, лентах и т. д.). ЛС БД описывается на языке описания данных СУБД; полученное описание обрабатывается специальным транслятором СУБД. При трансляции описания ЛС БД происходит автоматическое определение физической структуры БД.

ДИАЛОГОВОЕ ОБЕСПЕЧЕНИЕ ИМИТАЦИОННОЙ СИСТЕМЫ

Включение в программное обеспечение имитационной системы средств диалога обеспечивает более высокий технологический уровень имитационного эксперимента, позволяющий исследователю восполнять недостаток существующих методов имитации своим опытом и интуицией.

Так, исследователь получает возможность:

оперативно вмешиваться в процессы имитации, рассчитывать показатели эффективности и анализировать промежуточные данные сразу же после их получения;

формировать условия для следующего расчетного варианта после анализа данных, получаемых в предыдущих расчетах;

осуществлять выбор метода расчета показателей эффективности в процессе решения задачи;

варьировать в процессе моделирования временем решения задачи и точностью полученных результатов;

взаимодействовать не только с имитационной моделью, но и с базой данных имитационной системы;

оперативно вносить изменения в модель исследуемого объекта в процессе решения задачи в связи с уточнением ее постановки;

использовать эвристические методы решения задачи;

гибко управлять различными взаимосвязями между отдельными расчетными программами.

При организации диалога важное значение имеет установление приоритета партнеров, что определяется самой природой процесса диалогового взаимодействия, когда инициатива в ведении диалога в силу различных причин переходит от человека к ЭВМ и наоборот.

В практике нашли распространение следующие типовые режимы:

диалог, инициируемый ЭВМ («ведущий — ЭВМ»), когда управление диалогом осуществляет только администратор системы (лицо, сопровождающее программное и информационное обеспечение имитационной системы);

диалог равных партнеров, когда управление осуществляется совместно пользователями и ЭВМ;

диалог, инициируемый человеком («ведущий — человек»), когда управление диалогом осуществляют пользователи.

При исследовании эффективности операций средствами проблемно-ориентированной имитационной системы используют различные режимы диалога. Так, при постановке задачи на моделирование, формировании и анализе концептуальной модели диалог позволяет исследователю более эффективно усваивать условия задачи, оперативно получать информацию о состоянии системы и отдельных составляющих ее элементов, а также об используемых моделях и наборе работоспособных методов решения задач. При построении имитационной модели по запросу может быть выдана краткая характеристика модулей пакета прикладных программ и необходимая дополнительная информация по методике построения имитационной модели. На этих этапах решения задачи тип используемого режима диалога определяется, в основном, уровнем подготовки исследователя. Для неподготовленных поль-

зователей наиболее удобным является режим «ведущий — ЭВМ».

При планировании имитационного эксперимента исследователь может применять также любой из трех диалоговых режимов в зависимости от своего уровня знания системы и структуры диалога. При более высоком уровне подготовки пользователей предпочтение отдается режиму «ведущий—человек».

При проведении имитационного эксперимента исследователь, как правило, получает стандартно-контрольную информацию, поэтому диалог чаще организуется в режиме «ведущий—ЭВМ». Однако и на этом этапе можно использовать режим «ведущий—человек».

На этапе обработки и анализа результатов пользователи оценивают полученные результаты моделирования, определяют необходимость дополнительных запросов, анализируют ответы на них, формируют новые варианты задач и делают предварительные выводы по полученным результатам моделирования. При отрицательных результатах моделирования могут меняться исходные данные и проводиться повторное моделирование до тех пор, пока не будет получен положительный результат, после чего формируются окончательные выводы и документируются полученные результаты. На этом этапе большое распространение получил режим диалога «равных партнеров».

Подробнее вопросы организации диалога при имитационном исследовании эффективности технических систем изложены в [42].

ТЕХНОЛОГИЧЕСКАЯ СХЕМА РЕШЕНИЯ ЗАДАЧ СРЕДСТВАМИ ИМИТАЦИОННОЙ СИСТЕМЫ

Для успешного решения задач средствами имитационной системы исследователь должен обладать навыками и методикой проведения имитационных исследований.

На рис. 3 представлена технологическая схема проведения исследований на основе средств проблемно-ориентированной имитационной системы. На схеме отражены основные ком-

поненты программного и информационного обеспечения системы, показаны их взаимосвязь и последовательность применения в процессе решения задач исследования эффективности.

Последовательность решения задачи исследования эффективности технической системы с использованием средств рассмотренного моделирующего комплекса можно представить следующим образом.

В соответствии с изложенной концепцией многоуровневых исследований эффективности операции на первых этапах проводится неформальный анализ исследуемой проблемы и решается задача приближенной оценки эффективности изучаемой системы на основе ее концептуальной модели. Используя результаты исследований первых этапов, исследователь на этапе операционных исследований приступает к непосредственному решению задачи в среде имитационной системы.

При этом типовую последовательность его действий можно представить в виде подэтапов.

1. Ознакомление с возможностями системы и обучение приемам работы в ее среде. На этом этапе в качестве средств программного обеспечения можно использовать пакеты прикладных программ (ППП) информационно-справочных систем и обучающих систем.

2. Исследование вопроса о программной достаточности решения своей задачи; осуществляется поиск необходимых программных средств в Банке моделей и Банке модулей имитационной системы; относительно программной достаточности обеспечения решения задачи возможны три ситуации:

а) все программное обеспечение для решения задачи есть (имеется необходимая модель в Банке моделей системы), т. е. поставленная задача подходит под типовую в рамках используемой версии имитационной системы;

б) все компоненты программного обеспечения есть (Банк модулей сформирован полностью), но не построена требуемая модель исследуемого объекта или требуемый план проведения машинного эксперимента (нет модели в Банке моделей системы);

в) нет отдельных компонент в Банке модулей имитационной системы, необ-

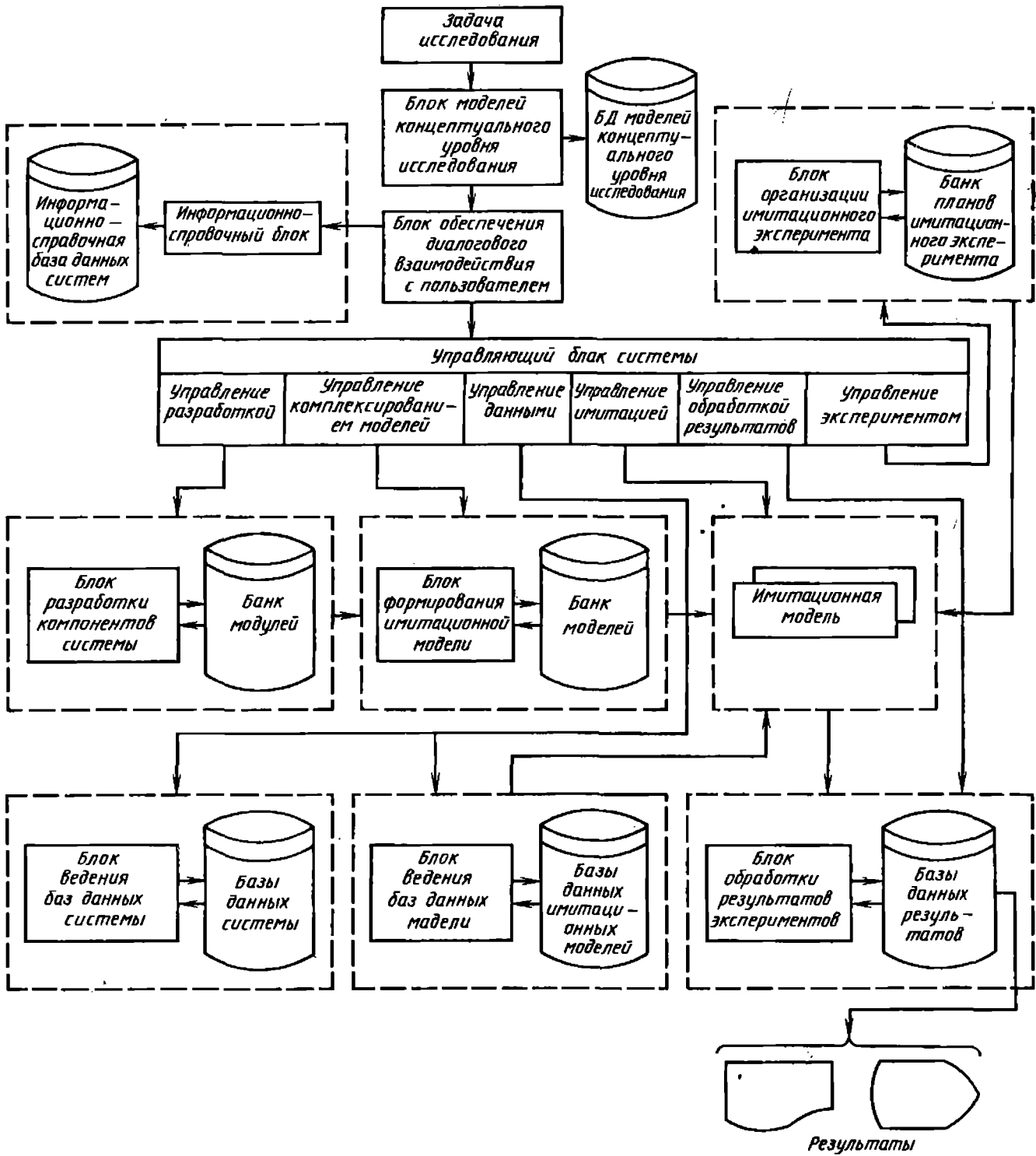


Рис. 3. Технологическая схема решения задач в рамках имитационной системы

ходимых для построения и исследования модели; последние две ситуации возникают с учетом требования открытости имитационной системы, поддерживаемого принципом ее многомодельности; во втором случае возникает задача комплексирования соответствующих модулей в рамках модели исследования; в третьем случае требуется разработка дополнительных модулей и их последующее комплексирование.

3. Решение вопроса информационной достаточности проведения исследований. Определяется наличие требуемых данных для задачи. С использованием СУБД имитационной системы на основе единой информационной базы данных системы формируется база данных решаемой задачи.

4. Выбор плана проведения исследований на ЭВМ. Сценарий исследования может быть готов и находиться в Банке

планов эксперимента (в виде управляющих операторов используемой программы системы-монитора), либо конкретный сценарий формируется (разрабатывается и кодируется на входном специализированном языке оператора-исследователя) и записывается в Банк планов.

5. По сформированному плану проводятся расчеты на ЭВМ, состоящие в имитации функционирования исследуемой системы, сборе статистических данных.

6. На основе полученных на предыдущем этапе результатов проводится их обработка и рассчитываются значения выбранных показателей эффективности. В соответствии с общей схемой исследования эффективности операции осуществляется переход к этапу принятия решений, после которого возможно обращение вновь к средствам имитационной системы, либо осуществляется возврат к одному из предыдущих этапов для проведения уточняющих исследований.

Глава 9. Оценка качества моделей и планирование экспериментов

1. ОЦЕНКА КАЧЕСТВА МОДЕЛЕЙ

Оценка качества является завершающим этапом работ по созданию модели операции (технической системы). Цель оценки качества модели — не столько получение конкретной числовой оценки значения некоторого показателя качества, сколько выявление того, где и как проявляются недостатки разработанной модели. Результаты оценки качества модели должны использоваться для определения направлений и способов ее совершенствования, планирования испытаний модели и для решения вопросов о возможности ее эксплуатации.

Модель системы состоит из математического описания исследуемой системы, алгоритма моделирования, программы для ЭВМ и рабочей документации. Поэтому при оценке качества модели учитывается качество каждой из перечисленных компонент.

Качество моделей оценивается по трем группам свойств: целевым, эксплуатационным и модификационным.

Целевыми называются свойства модели, характеризующие степень ее соответствия целям и задачам конкретного исследования. К целевым свойствам моделей относятся адекватность, устойчивость, точность и результативность.

Эксплуатационными называются свойства моделей систем, определяющие удобство эксплуатации модели пользователем. К ним относятся: производительность, надежность, защищенность, эргономичность, завершенность. Эксплуатационные свойства делятся на относительные (т. е. зависящие от конкретной задачи) — производительность, надежность и абсолют-

ные (не зависящие от конкретной задачи) — защищенность, эргономичность, завершенность.

Модификационными называются свойства моделей, определяющие удобство внесения изменений в модель (в случае необходимости, например, устранения ошибок в программе), простоту модификации модели при развитии и совершенствовании исследуемой системы или при переходе на другой уровень исследования (например, при повышении степени детализации математического описания).

К модификационным свойствам относятся понятность, структуризованность, расширяемость, доступность.

Модификационные свойства не связаны с конкретной задачей (они характеризуют возможности применения модели для решения множества задач определенного класса).

Перечисленные группы свойств определяют в целом качество модели. Эти свойства оценивают иногда с помощью специальных методик [69].

Среди свойств, определяющих качество моделей, особое место занимает точность результатов моделирования. В частности, важной характеристикой имитационной модели, включающей в свой состав случайные факторы, является дисперсия σ_y^2 переменной y , наблюдаемой в процессе машинного эксперимента (дисперсия воспроизводимости эксперимента). Точность оценки, например, среднего результата можно повысить увеличением объема испытаний (увеличением числа прогонов имитационной модели) при фиксированной дисперсии воспроизводимости σ_y^2 . Этот путь, однако, может привести к неприемлемо большим затратам машинного времени.

Другим путем повышения точности моделирования является понижение дисперсии воспроизводимости. В практике имитационного моделирования используют оба указанных приема. Обычно стремятся уменьшить дисперсию воспроизводимости, применяя специальные методы. При фиксированном значении этой дисперсии и с учетом заданных требований к точности моделирования определяют необходимый объем испытаний.

Если меры, принятые для понижения дисперсии воспроизводимости, приводят к тому, что необходимый объем испытаний не выходит за пределы допустимого, то качество имитационной модели (в рассмотренном смысле) считают приемлемым.

Некоторые наиболее употребительные в практике имитационных исследований методы понижения дисперсии (МПД) и установления необходимого объема испытаний рассмотрены на с. 130. Совокупность этих методов обычно относят к так называемому тактическому планированию экспериментов [73].

Цель методов стратегического планирования имитационных экспериментов — получение максимального объема информации об изучаемом явлении путем рационального использования факторного пространства при ограниченном числе испытаний. Один из распространенных в практике методов планирования факторных экспериментов (метод Бокса — Уилсона) рассмотрен на с. 136.

При имитационных исследованиях эффективности технических систем важное место занимают вопросы выбора переменных, наблюдаемых в ходе машинного эксперимента (параметров наблюдаемого процесса, характеристик изучаемой операции и т. д.). Наблюдаемые переменные в имитационной модели должны адекватно отражать исходы операции [конечные и (или) промежуточные]. Как правило, имитационная модель операции представляет собой одну из форм оператора H формирования исхода:

$$H: U \times \Lambda \rightarrow Y,$$

которая при моделировании больших систем имеет сложную разветвленную

математическую структуру и обычно задается в алгоритмическом виде.

Множество исходов Y операции с использованием операторов Q и M отображается во множество значений показателя эффективности W ; при этом оператор соответствия Q (выражаемый функцией соответствия) отображает множество Y во множество значений функции соответствия ρ :

$$Q: Y \times Y^{\text{тр}} \rightarrow \rho,$$

а оператор усреднения M переводит множество значений функции соответствия ρ во множество значений показателя эффективности W :

$$M: \rho \rightarrow W.$$

Оценка качества имитационной модели операции зависит не только от дисперсии воспроизводимости σ_y^2 , но и от вида оператора $Q \circ M$, представляющего собой суперпозицию соответствующих операторов, т. е. $Q \circ M: Y \rightarrow W$.

Конкретный вид этого оператора определяется постановкой задачи исследования эффективности операции и является основой для формирования требований к качеству имитационной модели, в частности к ее точности.

Качество оценивания эффективности операции, таким образом, определяется: видом оператора $Q \circ M$, определяющего соответствие принятого показателя эффективности цели операции; видом оператора формирования исхода H , определяющим методические погрешности математического описания моделируемой операции; планом эксперимента, определяющим степень рациональности использования факторного пространства (полноту информации об исходах операции, извлекаемой в ходе машинного эксперимента); дисперсией воспроизводимости и объемом испытаний. Последние три фактора характеризуют точность статистической оценки параметров наблюдаемых переменных.

Справедливо следующее выражение:

$$\mu_s = I_W \cdot f \{ \Delta Y_m, \Delta Y_n [\sigma_y^2, P_s(N_o)], \Xi \},$$

где μ_s — некоторая мера качества оценивания эффективности операции;

I_W — булева переменная, определяемая следующим образом:

$$I_W = \begin{cases} 1, & \text{соответствует } A_0, \\ 0, & \text{не соответствует } A_0; \end{cases}$$

A_0 — цель операции; ΔY_M — методическая погрешность математического описания моделируемой операции; ΔY_{Π} — погрешность измерения (вычисления) характеристик наблюдаемых переменных, зависящая от дисперсии воспроизводимости σ_y^2 , плана машинного эксперимента $P_{\Sigma}(N_0)$ и объема испытаний N_0 ; f — невозрастающая функция соответствующих аргументов; Ξ — прочие факторы, влияющие на качество оценивания эффективности.

Величины ΔY_M и ΔY_{Π} определяют точность имитационной модели операции.

Методическая погрешность ΔY_M в первую очередь зависит от объема и степени достоверности доступной исследователю информации о моделируемой операции и условиях ее проведения и, как правило, не поддается оценке формальными методами (за исключением случаев, когда результаты моделирования могут быть сопоставлены с результатами реальной операции).

Погрешность измерения ΔY_{Π} для имитационных моделей, содержащих случайные факторы, может быть оценена статистически; при этом, для прочих равных условий, чем больше объем испытаний, тем меньше величина ΔY_{Π} . В действительности, однако, объем испытаний ограничен имеющимися в распоряжении исследователя ресурсами (денежными, временными, машинными и т. д.).

Методы понижения дисперсии воспроизводимости и планирования эксперимента позволяют существенно уменьшить погрешность ΔY_{Π} , а следовательно, повысить качество имитационной модели и качество оценивания эффективности операции в целом.

2. МЕТОДЫ ПОВЫШЕНИЯ КАЧЕСТВА ОЦЕНОК ПОКАЗАТЕЛЕЙ ЭФФЕКТИВНОСТИ

При имитационных исследованиях часто возникает необходимость оценки средних значений $\bar{y}$ выходных харак-

теристик $y \in Y$ имитационной модели (средних значений результата операции).

Средние значения $\bar{y}$ могут представлять самостоятельный интерес, например, при изучении поведения моделируемой системы в различных условиях, при оценке влияния тех или иных факторов на исход операции и т. д. Значения $\bar{y}$ можно также использовать для вычисления значения показателя эффективности исследуемой операции.

Во всех подобных случаях повышение точности оценивания (вычисления) средних значений $\bar{y}$ ведет к повышению качества имитационного исследования в широком смысле. Однако особый интерес представляют случаи, когда $\bar{y}$ применяют для вычисления оценок показателя эффективности W , так как именно эти оценки являются основой для принятия решений в области создания и совершенствования систем и управления операциями.

С этой точки зрения методы повышения качества (в частном случае — точности) вычисления средних $\bar{y}$ являются одновременно и методами повышения качества оценивания показателей эффективности. Ниже рассмотрены наиболее употребительные из этих методов.

В качестве оценок средних обычно используют выборочные средние, которые определяются по конечной выборке значений $\bar{y}$ из генеральной совокупности Y . В силу случайности выборки (см. т. 2) выборочные средние не равны в точности истинным средним (математическим ожиданиям). Однако чем больше объем выборки (т. е. чем больше число прогонов имитационной модели), тем выше вероятность того, что выборочные средние будут близки к истинным средним.

Объем выборки, необходимый для вычисления выборочного среднего с заданной точностью, зависит от вида распределения случайной величины y , а также от того, коррелированы или не коррелированы между собой случайные элементы выборки.

При проведении экспериментов со статическими имитационными моделями элементы выборки не зависят от времени (обычно относятся к одному

и тому же моменту времени) и оказываются некоррелированными, если для моделирования случайных факторов в рассматриваемой операции используются генераторы случайных чисел, не связанные между собой и выдающие некоррелированные последовательности чисел.

При проведении экспериментов с динамическими имитационными моделями в силу внутренних свойств исследуемой системы, коррелированности воздействий, представляющих собой случайные функции времени, и других причин выходные характеристики модели (элементы выборки), представляющие собой, как правило, значения одной и той же (в общем случае, векторной) случайной функции времени, взятые для разных моментов времени, оказываются коррелированными.

Если случайные значения выходных характеристик имитационной модели некоррелированы и распределены одинаково (последнее допущение справедливо в тех случаях, когда законы распределения учитываемых в исследуемой операции случайных факторов не изменяются от прогона к прогону), то в силу центральной предельной теоремы (см. т. 2) величину $\bar{y}$ (выборочное среднее) можно считать нормально распределенной. В этом случае число прогонов N имитационной модели, необходимое для того, чтобы истинное среднее y^0 (скаляр) с вероятностью $(1 - \alpha)$ лежало в интервале $\bar{y} \pm b$, определяется следующим образом:

$$N = \frac{Z_{\alpha/2}^2 \cdot \sigma_y^2}{b^2}, \quad (9.1)$$

где $Z_{\alpha/2}$ — квантиль порядка $\alpha/2$ стандартного нормального распределения; σ_y^2 — дисперсия случайной величины y ; b — доверительный интервал.

Если значение дисперсии σ_y^2 до начала имитационного эксперимента неизвестно, целесообразно выполнить пробную серию из L прогонов и вычислить на ее основе выборочную дисперсию

$$\bar{\sigma}_y^2 = \frac{1}{L-1} \sum_{i=1}^L (y_i - \bar{y}^L)^2, \quad (9.2)$$

где y_i — значение выходной характеристики модели в результате i -го прогона ($i = 1, 2, \dots, L$); $\bar{y}^L$ — выборочное среднее, вычисленное по результатам L пробных прогонов.

Оценку $\bar{\sigma}_y^2$ подставляют в (9.1) вместо σ_y^2 и получают предварительную оценку необходимого числа прогонов N^* . После этого выполняют оставшиеся $N^* - L$ прогонов, периодически уточняя оценку (9.2) и необходимое число прогонов. Описанный способ определения необходимого объема выборки при неизвестной до начала эксперимента дисперсии называется последовательным.

Если элемент выборки — векторная величина, т. е.

$$y_i = |y_{i1}, y_{i2}, \dots, y_{iq}|^T,$$

$$(i = 1, 2, \dots),$$

оценку необходимого числа прогонов выполняют отдельно для каждой из компонент вектора y . Наибольшее из полученных значений N_j , ($j = 1, 2, \dots, q$) принимают в качестве окончательного числа прогонов.

Если элементы выборки коррелированы между собой, то применение описанного способа определения числа прогонов модели N , основанного на независимости результатов эксперимента, занижает оценку дисперсии и приводит к использованию меньших значений N , чем необходимо для обеспечения заданной точности.

В этом случае целесообразно использовать один из следующих методов определения оценок математического ожидания [46].

1. Выборку разбивают на интервалы, превышающие по длине интервалы «сильной» зависимости данных; оценку дисперсии проводят для каждого такого интервала отдельно, а полученные значения $\sigma_{y_k}^2$, ($k = 1, 2, \dots, K$), где K — число интервалов, осредняются; недостаток метода — невозможность до начала эксперимента достаточно точно определить число и длину интервалов.

2. Выборку формируют как запись достаточно большого числа реализаций процесса функционирования ТС; среднее и дисперсию вычисляют по сово-

купности данных в каждый момент времени, а не по совокупности данных вдоль одной траектории; полученные оценки затем осредняют по времени; для определения требуемого числа «реализаций» используют формулы (9.1) и (9.2); метод обеспечивает высокую точность оценивания, но его реализация требует больших затрат машинного времени.

3. Необходимое число прогонов имитационной модели (количество элементов выборки) определяют по приближенной формуле, учитывающей автокорреляцию данных [73]:

$$N \approx \frac{4Z_{\alpha/2}^2 \bar{\sigma}_y^2 \left[1 + 2 \sum_{k=1}^s \left(1 - \frac{k}{s+1} \right) r^*(k) \right]}{b^2 \bar{y}^*}, \quad (9.3)$$

где $r^*(k)$ — определяемые по результатам дополнительных (пробных) экспериментов выборочные коэффициенты автокорреляции с шагом k ;

$$r^*(k) = \frac{1}{\bar{\sigma}_y^2 (n-1)} \times \sum_{j=1}^{n-k} (y_j^* - \bar{y}^*) (y_{j+k}^* - \bar{y}^*); \quad (9.4)$$

y_j^* , ($j = 1, 2, \dots, n$) — выборочные значения выходной характеристики модели в дополнительной серии из n прогонов; $s \leq n-1$ — глубина учета автокорреляции.

Для большинства приложений достаточно принимать $s \approx 0,1n$. Используемые в (9.3) и (9.4) оценки выборочного среднего $\bar{y}^*$ и выборочной дисперсии $\bar{\sigma}_y^2$ должны определяться по результатам отдельных дополнительных экспериментов.

В случае векторных величин y третий метод применяют отдельно для каждой из компонент вектора y .

Наиболее точные оценки могут быть получены с помощью методов спектрального анализа [5].

Рассмотренные методы определения числа прогонов имитационной модели

(количества элементов выборки) относятся к группе методов простой случайной выборки (ПСВ). Основным недостатком ПСВ — медленная сходимость выборочных средних к истинным средним с ростом объема выборки N , пропорциональная $\sqrt{N}$.

Медленность сходимости при высоких требованиях к точности оценивания средних приводит к необходимости применения методов уменьшения ошибок (повышения качества оценивания), не требующих увеличения объема выборки. Такие методы, получившие название *методов понижения дисперсии* (МПД), учитывают информацию о статистических свойствах моделируемых воздействий на ТС, об особенностях структуры и динамики исследуемой ТС, о распределениях выходных характеристик модели.

Все методы, использующие при вычислении выборочных средних дополнительную информацию, делят на три группы: методы, применяемые после того, как выборка некоторого объема уже сформирована (пассивные методы оценивания); методы, предусматривающие формирование выборки специальным образом (активные методы оценивания); методы, в которых для получения оценок средних используются значения и реализации некоторых вспомогательных величин и процессов, косвенно связанных с выходными характеристиками модели ТС (косвенные методы оценивания).

ПАССИВНЫЕ МЕТОДЫ ПОВЫШЕНИЯ КАЧЕСТВА ОЦЕНИВАНИЯ ПОКАЗАТЕЛЯ ЭФФЕКТИВНОСТИ

Пусть в имитационной модели операции нужно учесть m случайных факторов $\lambda_1, \lambda_2, \dots, \lambda_m$. Поскольку для моделирования каждого из значений факторов λ_i , ($i = 1, 2, \dots, m$) используют одно или несколько значений случайных чисел, имеющих равномерное распределение на интервале $[0,1]$, то справедлива запись:

$$B: R \rightarrow \Lambda, \\ \text{где } B — \text{оператор преобразования; } \Lambda = \bigcup_{j=1}^m \Lambda_j — \text{множество значений слу-}$$

чайных факторов, учитываемых при моделировании; $R = \bigcup_{k=1}^l R_k$, $l \geq m$ — множество значений случайных величин, имеющих равномерное распределение на интервале $[0,1]$; Λ_j , ($j = 1, 2, \dots, m$) — множество значений j -го случайного фактора; R_k , ($k = 1, 2, \dots, l$) — множество значений k -й случайной величины.

Множество R представляет собой l -мерное пространство векторов r со случайными координатами и называется выборочным пространством.

Поскольку имитационная модель $H (U \times \Lambda \rightarrow Y)$ при каждой конкретной стратегии управления $u \in U$ каждому вектору $|\lambda_1^*, \lambda_2^*, \dots, \lambda_m^*|^T$ значений случайных факторов ставит в соответствие единственное значение выходной характеристики y^* , то можно утверждать, что каждой точке r^* выборочного пространства R также соответствует единственное значение $y^* = |y_1^*, y_2^*, \dots, y_q^*|^T$

Для рассматриваемых ниже методов понижения дисперсии требуется, чтобы корреляция между каждым случайным вектором r , моделируемым в эксперименте, и выходными характеристиками модели была положительной и, по возможности, большой, т. е. большим значениям переменных r_j , ($j = 1, 2, \dots, l$) должны соответствовать большие значения характеристик y .

Это условие в имитационных моделях легко реализуемо, если учесть, что законы распределения случайных величин x_1 и $x_2 = 1 - x_1$ тождественны, где x_1 — случайное число, имеющее равномерное распределение на интервале $[0,1]$.

Наличие свободы выбора формы выборочного пространства является одним из основных отличий машинного эксперимента от натурального (эксперимента с реальной системой).

Метод регрессионной выборки. Введем случайную величину z , заданную на R следующим образом:

$$z(c) = y - x(c) + M[x(c)], \quad (9.5)$$

где вспомогательная переменная $x(c)$, определенная на R , является случайной величиной, зависящей от L -мер-

ного вектора параметров $c = |c_1, c_2, \dots, c_L|^T$.

Математическое ожидание $M[x(c)]$ случайной величины $x(c)$ должно быть известно. Из (9.5) следует, что $M[z(c)] = \bar{y}$, а дисперсия $z(c)$

$$\sigma_z^2 = \sigma_y^2 + \sigma_x^2 - 2Sp\{K_{xy}[x, y]\},$$

где $Sp\{K_{xy}[x, y]\}$ — след корреляционной матрицы между переменными $x(c)$ и y . Если вспомогательную переменную $x(c)$ ввести так, чтобы выполнялось условие $\sigma_z^2 < \sigma_y^2$, то для оценки величины $\bar{y}$ целесообразно использовать $M[z(c)]$.

Для того чтобы дисперсия σ_z^2 , определяющая точность вычисления $\bar{y}$, была меньше σ_y^2 , необходимо, чтобы $Sp\{K_{xy}[x, y]\}$ был положительным числом, желательно большим, т. е. корреляция между y и $x(c)$ должна быть близка к +1.

Минимум дисперсии σ_z^2 обеспечивается в том случае, если x является регрессией на R (отсюда и название метода), т. е. если вектор c минимизирует сумму квадратов разностей величин y_j и x_j , ($j = 1, 2, \dots, N$), где y_j и x_j — элементы соответствующих выборок, N — объем каждой из выборок. Компоненты вектора параметров c находятся известными способами простой (если y — скаляр) или множественной (y — вектор) регрессии (см. т. 2). Если вектор параметров c и

$$M[z] = \frac{1}{N} \sum_{j=1}^N \{y_j - x_j + M[x]\} \quad (9.6)$$

вычисляют с помощью одной и той же выборки, то $M[z]$ является смещенной оценкой для $\bar{y}$ из-за наличия корреляции. Смещение можно практически исключить, если разделить выборку на несколько групп (в простейшем случае — на две). Для каждой i -й группы отдельно вычисляют вектор параметров c_i и $M_i[z]$ с использованием среднего значения:

$$\bar{c} = \frac{1}{D-1} \sum_{j=1}^{D-1} c_j | j \neq i, \quad (9.7)$$

где D — число групп.

Внутри каждой группы величины x и s независимы, а $M_i[z]$ — несмещенные оценки. Искомое выборочное среднее определяется осреднением полученных частных оценок:

$$M[z] = \frac{1}{D} \sum_{i=1}^D M_i[z]. \quad (9.8)$$

Дисперсия величины $M[z]$, определенная таким способом, будет несколько выше истинного значения вследствие остаточной коррелированности между группами данных, но ниже, чем σ_y^2 .

Метод компенсации [73]. Сущность метода компенсации заключается в вычислении $\bar{y}$ как среднего двух оценок $\bar{y}_1$ и $\bar{y}_2$, имеющих сильную отрицательную корреляцию. Дисперсия оценки $\bar{y} = (\bar{y}_1 + \bar{y}_2)/2$ имеет вид

$$\sigma_{\bar{y}}^2 = \frac{1}{4} (\sigma_{\bar{y}_1}^2 + \sigma_{\bar{y}_2}^2) + \frac{1}{2} Sp \{K_{\bar{y}_1, \bar{y}_2} [\bar{y}_1, \bar{y}_2]\}. \quad (9.9)$$

Вследствие отрицательной корреляции эта величина существенно меньше, чем при использовании независимых оценок $\bar{y}_1$ и $\bar{y}_2$.

Оценки $\bar{y}_1$ и $\bar{y}_2$ получают по отдельным выборочным подпространствам $R' = \{R'_1, R'_2, \dots, R'_l\}$ и $R'' = \{R''_1, R''_2, \dots, R''_l\}$, которые в простейшем случае связаны между собой следующим образом:

$$r''_k = \bar{1} - r'_k | r'_k \in R', r''_k \in R'' \quad (k = 1, 2, \dots), \quad (9.10)$$

где $\bar{1}$ — l -мерный единичный вектор.

Возможны и другие способы формирования пространств R' и R'' . Единственным требованием к ним является обеспечение отрицательной (по возможности, сильной) корреляции между оценками $\bar{y}_1$ и $\bar{y}_2$. В большинстве практически важных случаев дисперсия оценки при использовании метода компенсации уменьшается в 2—10 раз, что делает его одним из наиболее эффективных МПД [73].

Широко применяется также метод испытаний с дополняющей переменной

[46], вводимой на R таким образом, чтобы корреляция между ней и оцениваемой переменной была близка к — 1.

Преимущества пассивных методов понижения дисперсии по сравнению с активными методами — возможность их применения уже после того, как выборка значений сформирована.

Эти методы осуществляются в два этапа: сначала формируют выборочное пространство R и выборку, а затем вычисляют оценки выборочного среднего. Эти этапы формально независимы, за исключением того, что при формировании пространства R необходимо предусмотреть включение в него значений, позволяющих вычислять вспомогательные переменные или формировать взаимокompенсирующие выборочные подпространства R' и R'' .

Важным достоинством пассивных методов оценивания является возможность их применения без существенных изменений для любых статистических или имитационных моделей.

АКТИВНЫЕ МЕТОДЫ ПОВЫШЕНИЯ КАЧЕСТВА ОЦЕНИВАНИЯ ПОКАЗАТЕЛЯ ЭФФЕКТИВНОСТИ

Активные методы оценивания среднего основаны на извлечении выборки специальным образом с использованием информации о структуре пространства значений исходов операции y . Для большинства активных методов требуется разделение выборки на части, называемые *слоями*. При этом необходимо, чтобы значения элементов выборки как можно меньше различались внутри одного и того же слоя и как можно больше — между различными слоями.

С помощью случайного отбора элементов внутри слоев для каждого из них можно оценить $\bar{y}$ с точностью, зависящей от разброса значений y внутри слоя. Полученные оценки затем используют для вычисления математического ожидания по выборке в целом:

$$\bar{y} = \frac{1}{N} \sum_{i=1}^K N_i \bar{y}_i, \quad (9.11)$$

где N и N_i — соответственно объем выборки и ее i -го слоя; K — число

слоев; $\bar{y}_i$ — оценка $\bar{y}$, вычисленная для i -го слоя.

Если предположить, что оценки $\bar{y}_i$, ($i = 1, 2, \dots, K$) независимы, то из (9.11) получим

$$\sigma_y^2 = \frac{1}{N^2} \sum_{i=1}^K N_i^2 \sigma_{y_i}^2, \quad (9.12)$$

где $\sigma_{y_i}^2$ — дисперсия y , вычисленная для i -го слоя.

При удачном выборе слоев, когда дисперсии $\sigma_{y_i}^2$ малы, оценка (9.11) также будет иметь малую дисперсию и, следовательно, будет предпочтительнее оценки, полученной методами простой случайной выборки.

В большинстве практических задач, в особенности при векторных величинах $y = |y_1, y_2, \dots, y_q|^T$, возможности уменьшения разброса значений y внутри слоев ограничены. Это объясняется тем, что для одной и той же системы слоев зачастую невозможно минимизировать дисперсии всех или хотя бы нескольких величин $\bar{y}_i$, ($i = 1, 2, \dots, K$). Кроме того, слои должны быть четко отделены один от другого, что не всегда выполнимо.

Часто можно использовать естественное разбиение выборки на слои, например экономические данные можно разделить по районам, предприятиям, видам продукции и т. д.

Методы оценивания среднего с помощью расслоенной выборки отличаются способом взвешивания послойных оценок. Наиболее широко используют метод пропорционального отбора и метод оптимального отбора.

В методе пропорционального отбора объемы послойных выборок определяются пропорционально объемам соответствующих слоев:

$$n_i = nN_i/N, \quad (9.13)$$

где n_i — объем i -й послойной выборки, т. е. число элементов, случайным образом взятых из i -го слоя выборки для вычисления соответствующего сред-

него; n — суммарный объем послойных выборок;

$$n = \sum_{i=1}^K n_i. \quad (9.14)$$

В методе оптимального отбора объемы послойных выборок полагают пропорциональными среднеквадратическим отклонениям по соответствующим слоям:

$$n_i = n \frac{N_i \sigma_{y_i}}{\sum_{j=1}^K N_j \sigma_{y_j}}. \quad (9.15)$$

Состав оптимальной выборки зависит от $\sigma_{y_1}, \sigma_{y_2}, \dots, \sigma_{y_K}$ — среднеквадратических отклонений по слоям; поэтому он обычно неодинаков для различных компонент вектора y . Метод оптимального отбора целесообразно применять в тех случаях, когда важно повысить точность оценивания какой-либо одной выходной характеристики модели.

Недостаток метода — необходимость знания величин $\sigma_{y_1}, \sigma_{y_2}, \dots, \sigma_{y_K}$. На практике эти величины зачастую априорно неизвестны и их необходимо определять по дополнительным выборкам.

Получаемые оценки σ_{y_i} оказываются приближенными, и точность метода вследствие этого снижается. Для получения достаточно надежных оценок σ_{y_i} необходимы выборки, содержащие не менее 30 элементов в каждом слое. Можно также использовать априорную информацию, экспертное оценивание и т. д.

КОСВЕННЫЕ МЕТОДЫ ПОВЫШЕНИЯ КАЧЕСТВА ОЦЕНИВАНИЯ ПОКАЗАТЕЛЯ ЭФФЕКТИВНОСТИ

Зачастую одни выходные характеристики модели легче получить (вычислить), чем другие. Если существует взаимосвязь (в частном случае — функциональная зависимость) между несколькими выходными характеристиками, то может оказаться целесообразным оценивать по результатам имитационного эксперимента те из

них, которые легче моделируются, а полученную информацию использовать для оценки остальных.

В отдельных случаях возможна замена исследуемого процесса некоторым другим, имеющим то же математическое ожидание, но меньший разброс значений.

Методы оценивания средних значений выходных характеристик по результатам имитационного эксперимента, основанные на использовании указанных или аналогичных им возможностей, получили название *косвенных*. Эти методы связаны с изучением процессов и их характеристик, лишь косвенно связанных с исследуемыми.

Метод сопутствующих переменных. Пусть необходимо оценить среднее значение скалярной выходной характеристики y по выборке, разделенной на K слоев. Допустим также, что существует возможность вычисления вместо y некоторой выходной характеристики x , называемой сопутствующей переменной.

В данном методе предполагается, что регрессия y на x линейна, так что в каждом i -м слое, ($i = 1, 2, \dots, K$) распределение y при заданном x^* имеет среднее

$$\bar{y}_i = a_0 + a_1 x^* \quad (9.16)$$

и среднее квадратическое отклонение

$$\sigma_{y_i} = \bar{\sigma}_{y_i} \sqrt{1 - k_{xy}^2}, \quad (9.17)$$

где a_0 и a_1 — известные постоянные величины; k_{xy} — коэффициент корреляции между x и y ; $\bar{\sigma}_{y_i}$ — среднее квадратическое отклонение в i -м слое, вычисленное без учета корреляции.

Пусть из каждого i -го слоя извлекается выборка объема n_i с элементами x_{ij} ($i = 1, 2, \dots, K$; $j = 1, 2, \dots, n_i$). Тогда сумма значений y по всем слоям

$$y_\Sigma = Na_0 + a_1 \sum_{i=1}^K \frac{N_i}{n_i} \sum_{j=1}^{n_i} x_{ij}, \quad (9.18)$$

где $N = \sum_{i=1}^K N_i$ — суммарный объем слоев; N_i — объем i -го слоя.

Дисперсия величины y_Σ :

$$\sigma_{y_\Sigma}^2 = a_1^2 \sum_{i=1}^K N_i^2 \left(\frac{1}{n_i} - \frac{1}{N_i} \right) \sigma_{x_i}^2, \quad (9.19)$$

где $\sigma_{x_i}^2$ — дисперсия x в i -м слое.

Если бы в каждом случае вместо x измерялось (моделировалось) значение y , то

$$y_\Sigma = \sum_{i=1}^K N_i \bar{y}_i \quad (9.20)$$

и дисперсия этой оценки была бы

$$\bar{\sigma}_{y_\Sigma}^2 = \sum_{i=1}^K N_i^2 \left(\frac{1}{n_i} - \frac{1}{N_i} \right) \sigma_{y_i}^2. \quad (9.21)$$

Поскольку справедливо соотношение

$$\sigma_{x_i}^2 = \frac{k_{xy}^2 \sigma_{y_i}^2}{a_1^2},$$

то (9.19) можно записать в виде:

$$\sigma_{y_\Sigma}^2 = k_{xy}^2 \sum_{i=1}^K N_i^2 \left(\frac{1}{n_i} - \frac{1}{N_i} \right) \sigma_{y_i}^2. \quad (9.22)$$

Так как $k_{xy}^2 \leq 1$, то из сопоставления (9.21) и (9.22) следует, что $\sigma_{y_\Sigma}^2 \leq \bar{\sigma}_{y_\Sigma}^2$, т. е. при выполнении условия (9.16) использование сопутствующей переменной x никогда не приведет к менее точной оценке суммы значений y в совокупности, чем непосредственные оценки $\bar{y}$, полученные по выборке той же структуры.

Все оценки в методе сопутствующих переменных существенно зависят от того, выполняются или нет сделанные допущения о коэффициентах регрессии, средних по выборочному пространству и т. д.

Если допущения не выполняются, оценки могут оказаться сильно смещенными [73].

Метод значимой выборки [46] основан на замене исследуемого процесса

другим, имеющим то же математическое ожидание, но меньшую дисперсию.

Пусть на выборочном пространстве определена вспомогательная переменная

$$z = \frac{y f(r)}{f^*(r)}, \quad (9.23)$$

где y — случайная величина, среднее которой ищется; $f(r)$ — плотность вероятности случайного вектора r (элемента выборочного пространства); $f^*(r)$ — вспомогательная функция плотности вероятности, отличная от нуля для всех $r \in R$.

Так как $y = y(r)$, то из (9.23) следует, что математическое ожидание величины y с плотностью вероятности $f(r)$ равно математическому ожиданию величины z с плотностью вероятности $f^*(r)$.

Дисперсия величины z :

$$\sigma_z^2 = \int_R \frac{y^2 f^2(r)}{[f^*(r)]^2} f^*(r) dr - \bar{z}^2, \quad (9.24)$$

где $\bar{z}$ — среднее значение величины z .

Если все значения y положительны, то $\sigma_z^2 = 0$ при

$$f^*(r) = \frac{1}{\bar{y}} \cdot y \cdot f(r). \quad (9.25)$$

В действительности $\bar{y}$ и $f(r)$ неизвестны (иначе эксперимент с моделью был бы не нужен), и дисперсия σ_z^2 практически не может быть сведена к нулю. Целесообразно эмпирически подобрать вид функции $f^*(r)$, определив ее параметры на основе априорной информации об исследуемой ТС или с помощью дополнительной выборки.

Если компоненты вектора r независимы (некоррелированы), что практически всегда выполнимо в имитационной модели, то справедливо

$$f^*(r) = \prod_{i=1}^l f_i^*(r_i).$$

Опытным путем установлено [46], что целесообразно выбирать

$$f_i^*(r_i) = \frac{a_i^{r_i} \ln(a_i)}{a_i - 1}, \quad (9.26)$$

где $a_1, a_2, \dots, a_l$ — постоянные величины.

Полагают $a_1 = a_2 = \dots = a_l = c$, и величину c находят из условия $\frac{d}{dc}(\sigma_z^2) = 0$, подставляя в (9.24) значение $\bar{z}$, полученное с помощью дополнительной выборки.

3. ПЛАНИРОВАНИЕ ИМИТАЦИОННЫХ ЭКСПЕРИМЕНТОВ

Целью имитационных экспериментов является возможно более глубокое изучение моделируемых систем при ограниченных затратах. С этой целью необходимо планировать и проектировать не только модель системы, но и процесс проведения экспериментов с ней.

В практике имитационных исследований наиболее распространены следующие типы экспериментов:

сравнение средних и дисперсий результатов операций для различных альтернатив;

определение значимости влияния тех или иных факторов и необходимости их учета при исследовании конкретной системы;

отыскание оптимальных альтернатив (в частности, стратегий управления) на некотором множестве возможных значений.

Математические методы отыскания целесообразных планов проведения экспериментов указанных типов и проведения расчетов с целью вычисления искомых оценок получили название *методов планирования эксперимента*.

Формализованную общую схему эксперимента можно представить следующим образом.

Пусть некоторая система имеет k контролируемых входов $\bar{X} = \|x_1, x_2, \dots, x_k\|^T$, называемых *факторами*.

Вектор $\bar{X}$ изображается точкой в пространстве переменных x_j , $j = 1, 2, \dots, k$ (в факторном пространстве). Эксперимент называют *активным*, если исследователь по своему усмотрению может изменить значения (уровни) факторов. В противном случае (иссле-

дователь контролирует лишь уровни факторов, но не управляет ими) эксперимент называют *пассивным*.

Обозначим: i — номер опыта ($i = 1, 2, \dots, n$); $\bar{x}_i$ — комплекс условий i -го опыта ($\bar{x}_i = \|x_{i1}, x_{i2}, \dots, x_{ik}\|$); x_{ij} — уровень j -го фактора в i -м опыте.

Система имеет также ν входов $\omega_1, \omega_2, \dots, \omega_\nu$, которые в имитационной модели полагают случайными и задают в процессе имитации с помощью датчиков случайных чисел (имитаторов). Управлять этими входами, в силу их случайности, исследователь не может.

Система характеризуется $(\nu + 1)$ выходными параметрами $y, v_1, v_2, \dots, v_\nu$, причем y — основной параметр (в частности, показатель эффективности).

Основной выходной параметр y (наблюдаемая переменная) является скалярной величиной; при этом предполагается, что имеет место следующая модель наблюдения:

$$y = \eta(\bar{x}) + \varepsilon(\bar{x}), \quad (9.27)$$

где $\eta(\bar{x})$ — функция отклика (неслучайная функция факторов); $\varepsilon(\bar{x})$ — ошибка опыта (случайная величина, распределение которой определяется значением вектора $\bar{x}$, $M[\varepsilon(\bar{x})] = 0$).

Очевидно, y является случайной переменной, так как зависит от случайной величины $\varepsilon(\bar{x})$.

Остальные ν выходных параметров в общем случае также зависят от входов и в процессе эксперимента должны находиться в некоторых пределах, задаваемых, например, системой неравенств:

$$\psi_\alpha[v(x)] \leq 0, \alpha = 1, 2, \dots, a. \quad (9.28)$$

Система неравенств (9.28) выделяет в факторном пространстве область эксперимента G (область допустимых значений факторов). Функция отклика $\eta(\bar{x})$ математически отражает механизм изучаемого явления (поведение исследуемой системы).

Общая задача эксперимента с моделью системы — получение информации о функции отклика. Наиболее важными являются экспериментальные исследования с помощью моделей, на-

правленные на отыскание некоторых значений $x_1^*, x_2^*, \dots, x_k^*$ управляемых факторов, для которых основной выходной параметр достигает своего экстремального значения (экстремальный эксперимент).

В процессе экстремального эксперимента при фиксированных значениях управляемых факторов наблюдают значение переменной y . По значению (или по нескольким значениям) этой переменной, полученному из опыта, можно лишь приближенно оценить функцию отклика. Значение этой функции в фиксированной точке $x_1, x_2, \dots, x_k$ факторного пространства равно условному математическому ожиданию наблюдаемой переменной y в этой точке, т. е.

$$\eta(x) = M[y(x)], \quad (9.29)$$

так как по предположению $M[\varepsilon(\bar{x})] = 0$.

В статистике функция, связывающая условное математическое ожидание случайной переменной с контролируемыми переменными, называется *регрессией* (см. т. 2).

Если вид функции отклика $\eta(x)$ неизвестен, то полезным оказывается ее представление в виде полинома. В этом случае функция отклика записывается в виде

$$\begin{aligned} \eta(x) = & \beta_0 + \sum_{j=1}^k \beta_j x_j + \sum_{j < l} \beta_{jl} x_j x_l + \\ & + \sum_{j=1}^k \beta_{jj} x_j^2 + \end{aligned} \quad (9.30)$$

где $\beta_0, \beta_j, \beta_{jl}, \beta_{jj}$ — неизвестные теоретические коэффициенты регрессии.

Функция отклика считается полностью известной, если определены все коэффициенты β , входящие в формулу (9.30).

Вследствие влияния на результаты эксперимента случайных факторов $\omega_1, \omega_2, \dots, \omega_\nu$ истинные значения коэффициентов могут быть оценены лишь

приблизительно, т. е. может быть построена лишь функция

$$\hat{y} = b_0 + \sum_{j=1}^k b_j x_j + \sum_{j < l} b_{jl} x_j x_l + \sum_{j=1}^k b_{jj} x_j^2 + \quad (9.31)$$

где $\hat{y}$ — оценка значения функции отклика, предсказанная по уравнению регрессии (9.31); b_0, b_j, b_{jl}, b_{jj} — статистические оценки коэффициентов β , которые получают с помощью метода наименьших квадратов.

Этот метод дает эффективные оценки коэффициентов регрессии в случае, если ошибка опыта $\varepsilon(\bar{x})$, а следовательно, и сама переменная y имеют нормальное распределение. Дисперсия наблюдаемой переменной y , очевидно, равна дисперсии ошибки опыта, т. е.

$$\sigma_y^2 = \sigma^2 \{\varepsilon\};$$

σ_y^2 называют дисперсией воспроизводимости эксперимента. Она характеризует качество эксперимента (эксперимент называют идеальным при $\sigma_y^2 = 0$).

Точность оценки математического ожидания выходного параметра повышается с увеличением числа опытов, по которым оцениваются коэффициенты регрессии. Разность между числом опытов, по которым оценивают коэффициенты регрессии, и числом оцениваемых коэффициентов m называют числом степеней свободы h , т. е.

$$h = n - m. \quad (9.32)$$

Эксперимент называют *ненасыщенным* при $h > 0$, *насыщенным* при $h = 0$ и *сверхнасыщенным* при $h < 0$. Метод наименьших квадратов в регрессионном анализе используют при $h \geq 0$. Большое влияние на точность оценки коэффициентов оказывает расположение точек, в которых проводятся опыты, в факторном пространстве.

Пусть для оценки функции отклика предполагается провести n измерений выходного параметра y . В этом случае набор n координат точек в факторном пространстве $(x_{11}, x_{12}, \dots, x_{1k}), (x_{21}, x_{22}, \dots, x_{2k}), \dots, (x_{n1}, x_{n2}, \dots, x_{nk})$,

в которых будут проведены измерения (опыты) на множестве допустимых значений G , называется *планом эксперимента*. В общем случае координаты некоторых точек могут совпадать, т. е. в одной и той же точке может быть поставлено несколько опытов.

Основная задача планирования экстремального эксперимента — из всех допустимых выбрать такой план, который позволил бы оценивать функцию отклика с наибольшей точностью при фиксированном числе опытов n или, наоборот, выбрать такой допустимый план, при котором для оценки функции отклика с заданной точностью требуется минимальное число опытов.

Экстремальный эксперимент состоит из нескольких этапов. На каждом этапе с помощью модели экспериментально исследуют ограниченный участок (подобласть) факторного пространства, в отдельных точках которого по определенному плану ставят опыты. После каждого этапа полученные данные обрабатывают с целью получения информации о положении подобласти G^* , содержащей экстремальную точку $(x_1^*, x_2^*, \dots, x_k^*)$.

Околоэкстремальная область (или область экстремума) задается условием

$$\begin{aligned} &g(x_1^*, x_2^*, \dots, x_k^*) - \\ &- g(x_1, x_2, \dots, x_k) \leq \\ &\leq \delta; \quad \forall (x_1, x_2, \dots, x_k) \in G^*. \end{aligned} \quad (9.33)$$

Выход в околоэкстремальную область обычно осуществляют методом крутого восхождения. С этой целью в начальной подобласти факторного пространства, выбираемой на основе априорной информации, ставят серию опытов по определенному плану для оценки градиентного направления (градиента функции отклика). Далее делают несколько шагов в этом направлении до перевала (гребня) функции отклика, т. е. до точки факторного пространства, в которой знак приращения функции отклика меняется на очередном шаге. В области перевала вновь ставят серию опытов, по результатам которых оценивают направление градиента, и продолжают движение

в новом направлении. Признаком выхода в околэкстремальную область является незначительное отличие от нуля всех линейных коэффициентов регрессии b_j ($j = 1, 2, \dots, k$). В околэкстремальной области уточняют положение экстремума функции отклика, используя более сложные планы, чем планы для оценки градиента.

Начальные серии опытов планируют таким образом, чтобы можно было оценить лишь линейные коэффициенты регрессии, характеризующие направление градиента функции отклика. Поэтому на начальных этапах используют простейшую математическую модель эксперимента — линейную (полином первого порядка).

Линейная модель, используемая для аппроксимации соответствующего участка поверхности отклика, считается адекватной, если проводимая на ее основе оценка градиентного направления статистически удовлетворяет заданной точности. Получить пригодную для использования информацию о направлении градиента можно только с помощью адекватной модели, поэтому проверка адекватности — неотъемлемая часть каждого из этапов планирования эксперимента.

Линейная модель поверхности отклика от k переменных содержит $(k + 1)$ коэффициентов:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k. \quad (9.34)$$

О достижении области экстремума G^* свидетельствуют малые значения всех коэффициентов линейной регрессии. Если экстремум лежит на границе G , то значения коэффициентов регрессии могут быть любыми, однако в этом случае сам факт выхода на границу области G свидетельствует о достижении экстремума.

Эксперимент оканчивают при достижении области G^* , так как в этом случае экстремальные условия найдены с требуемой точностью.

Ниже приведены методы планирования серий экспериментов с целью предсказания направления градиента.

Рассмотрим схему полного факторного эксперимента (ПФЭ).

Область проведения эксперимента устанавливают на основе априорной информации об исследуемой системе. На основе этой информации выбирают и подобласть экспериментальной области, в которой ставят первую серию опытов.

Ниже рассматриваются экстремальные эксперименты, в которых каждый фактор варьируют на двух уровнях. В этом случае число точек, в которых проводят опыты, равно 2^k , где k — число факторов. Каждый из k факторов может быть установлен либо на нижнем, либо на верхнем уровне, которые расположены симметрично относительно некоторого нулевого уровня каждого фактора.

Интервалом варьирования фактора называется некоторое число J , прибавление которого к нулевому (основному) уровню дает верхний уровень, а вычитание — нижний. Нулевые уровни каждого фактора представляют собой точку в пространстве G , называемую *центром эксперимента (центром плана)*.

Для упрощения записи используют следующий код для обозначения уровней фактора: $x_j = +1$ (или «+») — верхний уровень j -го фактора; $x_j = -1$ (или «-») — его нижний уровень.

Переход к кодовым значениям уровней факторов задается преобразованием:

$$x_j = \frac{\bar{x}_j - \bar{x}_{j0}}{J_j} \quad (j = 1, 2, \dots, k), \quad (9.35)$$

где x_j — кодированное значение фактора; $\bar{x}_j$ — натуральное значение фактора; $\bar{x}_{j0}$ — натуральное значение нулевого уровня; J_j — интервал варьирования.

Эксперимент, в котором реализуются все возможные сочетания уровней факторов, называется *полным факторным экспериментом* (ПФЭ). Для двух уровней он называется ПФЭ типа 2^k , для L уровней — типа L^k .

План ПФЭ оформляют в виде матрицы планирования (табл. 1). Геометрическая интерпретация ПФЭ типа 2^2 показана на рис. 1.

1. Матрица ПФЭ типа 2²

Номер опыта	x_1	x_2	y
1	—	—	y_1
2	+	—	y_2
3	—	+	y_3
4	+	+	y_4

2. Матрица ПФЭ типа 2³

Номер опыта	x_1	x_2	x_3	y
1	—	—	+	y_1
2	+	—	—	y_2
3	—	+	—	y_3
4	+	+	+	y_4
5	—	—	—	y_5
6	+	—	+	y_6
7	—	+	+	y_7
8	+	+	—	y_8

Матрица планирования 2³ строится следующим образом. В столбцах первых двух факторов записывают матрицу типа 2² (в качестве начальной всегда выступает матрица типа 2²). В столбце 3-го фактора для первых 2² опытов записывают произведения столбцов x_1 и x_2 матрицы. Для последующих 2² опытов в столбцах первых двух факторов матрицу планирования повторяют, а в столбце 3-го фактора записывают произведения столбцов матрицы 2² с обратными знаками.

Матрица ПФЭ типа 2³ представлена в табл. 2.

Матрицы ПФЭ, построенные таким способом, обладают следующими свойствами:

а) симметричности; сумма всех элементов столбца любого фактора равна нулю:

$$\sum_{i=1}^n x_{ij} = 0; \quad j = 1, 2, \dots, k,$$

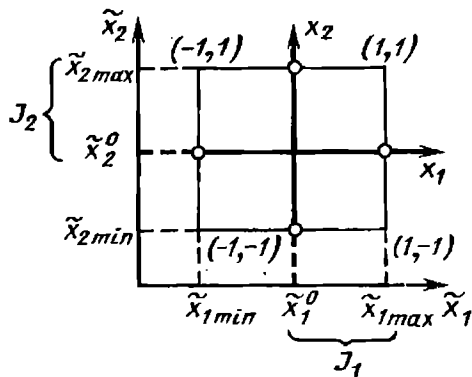


Рис. 1. Геометрическая интерпретация ПФЭ типа 2²

где i — номер опыта, j — номер фактора;

б) нормированности; сумма квадратов элементов любого столбца равна числу опытов:

$$\sum_{i=1}^n x_{ij}^2 = n; \quad j = 0, 1, 2, \dots, k;$$

в) ортогональности; сумма почленных произведений любых двух столбцов равна нулю:

$$\sum_{i=1}^n x_{ij}x_{il} = 0; \quad j, l = 0, 1, 2,$$

$$\dots, k, \quad j \neq l.$$

Матрицы ПФЭ типа 2^k обладают и свойством ротатабельности — независимостью дисперсии ошибки предсказания значений функции отклика от

направления градиента. Последнее свойство особенно важно, так как до начала эксперимента исследователь обычно не знает направление градиента и стремится поэтому принять план, точность которого одинакова во всех направлениях.

Важным этапом планирования эксперимента является выбор интервалов варьирования факторов. С одной стороны, интервал не должен быть настолько мал, чтобы уровни факторов были неразличимы на фоне случайных воздействий $\omega_1, \omega_2, \dots, \omega_r$, а с другой стороны — не должен быть настолько велик, чтобы нижний и верхний уровни располагались по разные стороны экстремума (в этом случае точность предсказания направления градиента может оказаться очень низкой) или выходили за пределы допустимой области G .

ПФЭ позволяет количественно оценить все линейные эффекты факторов и эффекты взаимодействия факторов. Взаимодействие факторов возникает в том случае, если эффект одного фактора зависит от уровня, на котором находится другой фактор.

Для расчета коэффициентов регрессии по результатам эксперимента строят специальную матрицу. За основу берут матрицу планирования ПФЭ (эта матрица имеет k столбцов и 2^k строк). Затем слева к ней добавляют столбец «фиктивной» переменной x_0 , состоящей только из «+1». Справа к начальной матрице дописывают столбцы, представляющие собой все возможные произведения факторов опытов при заданном уровне факторов.

В табл. 3 представлена матрица для расчета коэффициентов регрессии ПФЭ типа 2^3 с учетом взаимодействий факторов.

Столбцы произведений факторов получены почленным перемножением столбцов соответствующих факторов.

При обработке данных по методу наименьших квадратов ортогональность матрицы ПФЭ позволяет получить независимые оценки коэффициентов уравнения регрессии для неполного полинома (т. е. полинома, не имею-

3. Матрица для расчета коэффициентов регрессии ПФЭ типа 2^3 с учетом взаимодействий факторов

Номер опыта	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	y
1	+	—	—	+	+	—	—	+	y_1
2	+	—	+	—	—	+	—	+	y_2
3	+	+	—	—	—	—	+	+	y_3
4	+	+	+	+	+	+	+	+	y_4
5	+	—	—	—	+	+	+	—	y_5
6	+	—	+	+	—	—	+	—	y_6
7	+	+	—	+	—	+	—	—	y_7
8	+	+	+	—	+	—	—	—	y_8

щего членов, содержащих степени независимых переменных) вида

$$\begin{aligned} y = & b_0 + \sum_{j=1}^k b_j x_j + \sum_{j<l} b_{jl} x_j x_l + \\ & + \sum_{j<l<n} b_{jln} x_j x_l x_n + \dots + \\ & + b_{1,2\dots k} x_1 x_2 \dots x_k \end{aligned} \tag{9.36}$$

$(j, l, n = 1, 2, \dots, k).$

Независимость оценок означает, что оценка любого коэффициента не зависит от того, какие значения принимают оценки других коэффициентов. Выра-

жения для коэффициентов полинома (9.36) имеют следующий вид:

$$\left. \begin{aligned} b_0 &= \frac{\sum_{i=1}^n x_0 y_i}{n} = \frac{\sum_{i=1}^n y_i}{n}; \\ b_j &= \frac{\sum_{i=1}^n x_{ij} y_i}{n}; \\ b_{jl} &= \frac{\sum_{i=1}^n x_{ij} x_{il} y_i}{n}; \\ b_{jlr} &= \frac{\sum_{i=1}^n x_{ij} x_{il} x_{ir} y_i}{n}. \end{aligned} \right\} \quad (9.37)$$

Для определения какого-либо коэффициента необходимо алгебраическую сумму произведений функции отклика на знак соответствующего столбца разделить на число опытов n .

Независимость оценок коэффициентов обеспечивается только специальным выбором плана эксперимента. Если план составлен произвольно, то оценки оказываются коррелированными, что затрудняет их физическую интерпретацию и не позволяет придать четкий физический смысл уравнению регрессии.

Возможна следующая физическая интерпретация коэффициентов регрессии (применительно к неполному полиному): β_0 есть значение y в центре эксперимента; β_j численно равен приращению y при переходе j -го уровня с нулевого фактора на верхний (вклад j -го фактора); β_{jl} численно равен нелинейной части приращения y при одновременном переходе факторов x_j и x_l с нулевого уровня на верхний. Таким образом, если факторы x_j и x_l одновременно перевести с нулевого уровня на верхний, то приращение y будет складываться из линейной части ($\beta_j + \beta_l$), равной сумме коэффициентов при линейных членах, и нелинейной части, равной коэффициенту β_{jl} .

Эти утверждения справедливы лишь для модели вида (9.36), включающей только линейные эффекты и эффекты

взаимодействия. Между тем существенными могут оказаться коэффициенты при квадратах факторов, кубах и т. д. В этом случае уже нельзя раздельно вычислить оценки всех теоретических коэффициентов регрессии на основе ПФЭ типа 2^k .

Таким образом, из ПФЭ нельзя извлечь информацию о квадратичных членах и членах более высокого порядка. Для таких моделей следует выбирать планы с числом уровней более двух.

На начальных этапах экстремального эксперимента зачастую достаточно иметь лишь приблизительную информацию о направлении градиента, поэтому нет необходимости ставить все опыты в виде ПФЭ.

Эффективным приемом сокращения числа опытов является *дробный факторный эксперимент*.

При оценке направления градиента иногда можно ограничиться линейным приближением поверхности отклика:

$$y = b_0 + \sum_{i=1}^k b_i x_i. \quad (9.38)$$

Для оценки коэффициентов (9.38) необходимо провести не менее $(k+1)$ опытов, в то время как ПФЭ типа 2^k требует проведения 2^k опытов.

Однако сокращение числа опытов приводит к потере информации о поверхности отклика и допустимо лишь в том случае, если эта потеря несущественна. Планировать сокращенный эксперимент следует так, чтобы с максимальной точностью оценить коэффициенты b_j при линейных членах.

Если, например, в ПФЭ типа 2^3 (см. табл. 3) осуществить лишь первые четыре опыта, то коэффициенты (9.36) не могут быть оценены раздельно, так как все столбцы «усеченной» матрицы повторяются. Система смешивания оценок запишется в этом случае так:

$$\begin{aligned} b_0 &\rightarrow \beta_0 + \beta_{123} \text{ (так как } x_0 = x_1 x_2 x_3); \\ b_1 &\rightarrow \beta_1 + \beta_{23} \text{ (» » } x_1 = x_2 x_3); \\ b_2 &\rightarrow \beta_2 + \beta_{13} \text{ (» » } x_2 = x_1 x_3); \\ b_3 &\rightarrow \beta_3 + \beta_{12} \text{ (» » } x_3 = x_1 x_2). \end{aligned}$$

Если функция отклика линейна, то все эффекты взаимодействия отсутствуют, т. е.

$$\beta_{123} = \beta_{23} = \beta_{13} = \beta_{12} = 0.$$

Только в этом случае оценки b_j ($j = 0, 1, 2, 3$) будут оценками линейного уравнения теоретической регрессии:

$$b_0 \rightarrow \beta_0, b_1 \rightarrow \beta_1, b_2 \rightarrow \beta_2, b_3 \rightarrow \beta_3.$$

Таким образом удастся получить отдельные оценки коэффициентов при линейных членах, проводя число опытов в 2 раза меньше, чем в ПФЭ.

В этом случае «усеченная» матрица обладает всеми перечисленными выше свойствами. Тот же результат можно получить, если использовать только вторую половину матрицы.

Если из табл. 3 произвольно выбрать четыре опыта, например, 1, 4, 5 и 8, то система смешивания будет следующей:

$$b_0 \rightarrow \beta_0 + \beta_{12}; b_1 \rightarrow \beta_1 + \beta_2;$$

$$b_2 = \beta_2 + \beta_1; b_3 = \beta_3 + \beta_{12}.$$

т. е. отдельные оценки коэффициентов регрессии при произвольном выборе «усеченного» плана эксперимента нельзя получить даже для линейной модели регрессии.

Если в ПФЭ осуществляют только часть опытов, то матрица планирования такого сокращенного эксперимента носит название *дробной реплики от ПФЭ*. Реплика, включающая только половину опытов ПФЭ, называется *полуреplikой* ($1/2$ — репликой), включающая четвертую часть опытов — $1/4$ — *репликой* и т. д. Обозначаются они, соответственно, 2^{k-1} , 2^{k-2} и т. д.

Система смешивания оценок полностью определяется уравнениями столбцов. Уравнения совпадающих столбцов, с помощью которых задается система смешивания оценок, называются *генерирующими соотношениями* (например, $x_3 = x_1x_2x_3$). При выборе полуреплик используют априорную информацию об эффектах взаимодействия факторов. При этом систему смешивания оценок стремятся подобрать такой, чтобы линейные эффекты факторов смешивались только с эффектами взаи-

модействия (а не между собой), величины которых близки к нулю.

Полуреплика задается *определяющим контрастом* — генерирующим соотношением, в левой части которого стоит фиктивная переменная x_0 . Например, полуреплику 2^{3-1} , включающую в себя опыты с первого по четвертый, можно задать следующим определяющим контрастом:

$$x_0 = x_1x_2x_3,$$

из которого можно получить все генерирующие соотношения, последовательно умножая определяющий контраст на каждый столбец матрицы (при соблюдении следующих правил: $x_0x_j = x_j$; $x_jx_j = 1$),

$$x_1 = x_2x_3; x_2 = x_1x_3; x_3 = x_1x_2.$$

Реплики большей дробности задаются *обобщающим определяющим контрастом*. Например, для реплики 2^{5-2}

$$x_0 = x_1x_2x_3x_4x_5 = x_1x_2x_3 = x_4x_5.$$

С увеличением дробности реплик растет сложность смешивания оценок.

Проверка однородности дисперсии воспроизводимости. Дисперсия воспроизводимости $\sigma^2\{y\}$, характеризующая ошибку отдельного опыта, служит основой статистического анализа результатов эксперимента.

В некоторых случаях эта дисперсия исследователю известна, однако чаще всего ее необходимо оценить. Оценку выполняют по результатам серии параллельных (проводимых в одной и той же точке) опытов. Если не известно, однородна ли дисперсия воспроизводимости во всей области эксперимента, то проводят несколько серий опытов в различных точках.

Оценка дисперсии воспроизводимости в i -й точке факторного пространства вычисляется по формуле

$$\sigma_i^2\{y\} = \frac{1}{r_i - 1} \sum_{l=1}^{r_i} (y_{il} - \bar{y}_i)^2, \quad (9.39)$$

где r_i — число параллельных опытов в i -й точке области эксперимента; l — номер параллельного опыта ($l = 1, 2, \dots, r_i$); y_{il} — результат l -го опыта

в i -й точке области эксперимента; $\bar{y}_i$ — среднее значение функции отклика по всем параллельным опытам в i -й точке;

$$\bar{y}_i = \frac{1}{r_i} \sum_{l=1}^{r_i} y_{il}. \quad (9.40)$$

Дисперсию воспроизводимости по всей области эксперимента находят осреднением оценок по всем n точкам, в которых проводились серии параллельных опытов:

$$\sigma_y^2 = \frac{\sum_{i=1}^n h_i \sigma_i^2 \{y\}}{\sum_{i=1}^n h_i}, \quad (9.41)$$

где $h_i = r_i - 1$ — число степеней свободы в i -й точке.

Если число опытов в каждой точке одинаково и равно r , то с учетом (9.39)

$$\sigma^2 \{y\} = \frac{1}{n(r-1)} \sum_{i=1}^n \sum_{l=1}^r (y_{il} - \bar{y}_i)^2. \quad (9.42)$$

Однородность дисперсии воспроизводимости является необходимым условием для проведения регрессионного анализа, поскольку неоднородность дисперсии ведет к большим ошибкам вычисления коэффициентов регрессии.

Для проверки однородности дисперсии воспроизводимости в общем случае используется критерий Бартлетта:

$$\kappa = \frac{1}{c} \left[h \lg \sigma^2 \{y\} - \sum_{i=1}^n h_i \lg \sigma_i^2 \{y\} \right], \quad (9.43)$$

где

$$c = 0,4343 \left[1 + \frac{1}{3(N-1)} \times \times \left\{ \sum_{i=1}^n \frac{1}{h_i} - \frac{1}{h} \right\} \right]; \quad h = \sum_{i=1}^n h_i;$$

N — общее число опытов; κ приближенно подчиняется χ^2 — распределению с $(n-1)$ степенями свободы (см. т. 2).

Если

$$\kappa < \chi_{\text{табл}}^2$$

для заданного уровня значимости (обычно 5 %), то считают, что дисперсия воспроизводимости однородна.

Если число опытов в каждой точке одинаково и равно n , то используют более простой критерий Кохрена:

$$G = \frac{\sigma_{\max}^2 \{y\}}{\sum_{i=1}^n \sigma_i^2 \{y\}}. \quad (9.44)$$

Если $G < G_{\text{табл}}$, то дисперсия для заданного уровня значимости считается однородной. В формуле (9.44) $\sigma_{\max}^2 \{y\}$ — максимальная из дисперсий по всем точкам. Возможны и другие критерии однородности дисперсии.

Если в каждой точке плана проводится не один, а серия опытов, то в матрице ПФЭ или в матрице дробной реплики вместо значения y_i параметра оптимизации в i -й точке для определения коэффициентов регрессии в формуле (9.37) следует подставлять среднее значение параметра оптимизации в этой точке:

$$\bar{y}_i = \frac{\sum_{l=1}^{r_i} y_{il}}{r_i},$$

при этом в формуле (9.37) n — число точек, в которых проводились параллельные опыты.

Если установлено, что дисперсия воспроизводимости неоднородна, то следует опытным путем подобрать некоторое преобразование Φ параметра оптимизации:

$$y_\Phi = \Phi(y),$$

минимизирующее неоднородность. Часто оказывается подходящим логарифмическое или экспоненциальное преобразование.

Проверка адекватности уравнения регрессии основана на сопоставлении рассеяния значений параметра оптимизации y , полученных в результате моделирования, относительно линии регрессии с рассеянием этих значений относительно своих математических ожиданий, характеризуемым дисперсией воспроизводимости.

Для определения рассеяния y относительно линии регрессии вычисляют остаточную сумму квадратов отклонений экспериментальных значений от значений, полученных по уравнению регрессии:

$$S_R = \sum_{i=1}^n \sum_{l=1}^{r_i} (y_{il} - \hat{y}_i)^2, \quad (9.45)$$

где $\hat{y}_i$ — значение параметра оптимизации в i -й точке области эксперимента, предсказанное по уравнению регрессии. Остаточную сумму S_R , отнесенную к одной степени свободы, называют **дисперсией адекватности**:

$$\sigma_R^2 = \frac{S_R}{h}, \quad (9.46)$$

где $h = \sum_{i=1}^n r_i - m$ — число степеней свободы дисперсии адекватности; m — число оцениваемых коэффициентов регрессии.

Адекватность модели проверяют сравнением дисперсий адекватности и воспроизводимости по критерию Фишера с соответствующим числом степеней свободы (см. т. 2):

$$F = \frac{\sigma_R^{*2}}{\sigma^2\{y\}}. \quad (9.47)$$

При $F < F_{\text{табл}}$ гипотеза об адекватности модели не отвергается.

Значимость коэффициентов регрессии проверяют с целью исключения из уравнения регрессии факторов, слабо влияющих на параметр оптимизации y , и для установления факта попадания в область экстремума.

Обычно проверяют, значимо ли отличие от нуля того или иного коэффициента. Коэффициент b_j считается значимым, если он значимо отличается

от нуля в смысле, например, t — критерия Стьюдента (см. т. 2):

$$t = \frac{b_j}{\sigma^*\{b_j\}},$$

где $\sigma^*\{b_j\}$ — среднее квадратическое отклонение b_j .

При ортогональном планировании $\sigma\{b_j\}$ одинаковы для всех коэффициентов регрессии; поэтому

$$\sigma^2\{b_j\} = \frac{\sigma^2\{y\}}{\sum_{i=1}^n r_i}$$

и

$$t = \frac{b_j \sqrt{\sum_{i=1}^n r_i}}{\sigma\{y\}}. \quad (9.48)$$

Вычисленные по этой формуле значения t сравнивают с табличным значением $t_{\text{табл}}$ для заданного уровня значимости и числа степеней свободы

$$h = \sum_{i=1}^m h_i.$$

Если $t \geq t_{\text{табл}}$, то b_j значимо отличается от нуля.

Незначимые коэффициенты регрессии исключают и вновь проводят проверку адекватности модели; при этом следует иметь в виду следующие обстоятельства.

1. Незначимое отличие коэффициента регрессии от нуля еще не означает, что он равен нулю; причиной незначимости может быть также неудачный выбор интервала варьирования.

2. Если хотя бы один из коэффициентов при членах взаимодействия значимо отличается от нуля, то это свидетельствует о неадекватности линейной модели и о необходимости ее проверки и обеспечения адекватности.

3. Если все коэффициенты при линейных членах значимы, а при взаимодействиях — нет, то можно переходить к следующему этапу экстремального эксперимента — движению по направлению градиента (крутому восхождению). Если коэффициенты незначимы и при линейных членах,

то следует либо провести эксперимент с измененными (увеличенными или уменьшенными) интервалами варьирования факторов, либо, если и после этого коэффициенты будут незначимы, прекратить эксперимент, поскольку достигнута область экстремума.

4. Если при адекватной линейной модели только часть линейных коэффициентов значимо отличается от нуля, то следует повторить весь эксперимент при расширенных интервалах варьирования для тех факторов, при которых коэффициенты незначимы. Если и после этого те же коэффициенты оказались незначимыми, то соответствующие факторы можно исключить из дальнейших исследований.

5. При неадекватной линейной модели целесообразно заново провести весь эксперимент. При этом желательно центр эксперимента перенести в точку с лучшим из полученных значений параметра оптимизации, а интервалы варьирования уменьшить в обратной пропорции относительно соответствующих коэффициентов регрессии.

6. При неадекватной модели в некоторых случаях целесообразно начать движение по направлению градиента. Иногда это приводит к успеху быстрее, чем повторение экспериментов в прежних условиях.

Таким образом, общая процедура статистического анализа уравнения регрессии сводится: 1) к оценке дисперсии воспроизводимости в каждой строке плана эксперимента; 2) к проверке однородности дисперсий воспроизводимости, адекватности модели, значимости коэффициентов регрессии; 3) к исключению незначимых коэффициентов и проверке адекватности модели со значимыми коэффициентами; 4) к интерпретации уравнения регрессии в терминах моделируемой системы.

Рассмотрим этап крутого восхождения, т. е. движение по направлению градиента.

Градиентом функции отклика $\eta(x_1, x_2, \dots, x_k)$ в точке $x^0 = (x_1^0, x_2^0, \dots, x_k^0)$ называется вектор с координатами

$$\left(\frac{\partial \eta}{\partial x_1}, \frac{\partial \eta}{\partial x_2}, \dots, \frac{\partial \eta}{\partial x_k} \right)_{x=x^0},$$

т. е.

$$\text{grad } \eta(x^0) = \left\{ \frac{\partial \eta}{\partial x_1}, \frac{\partial \eta}{\partial x_2}, \dots, \frac{\partial \eta}{\partial x_k} \right\}_{x=x^0},$$

характеризующий направление максимального роста этой функции. Для того чтобы кратчайшим путем достичь области экстремума (вершины поверхности отклика), следует двигаться по направлению градиента.

Оценкой градиента является вектор линейных членов регрессии:

$$\text{grad}^* \eta(\bar{x}^0) = \{b_1, b_2, \dots, b_k\}. \quad (9.49)$$

Пошаговое движение по градиенту осуществляется следующим образом. Выбирают шаг α . Затем к координатам нулевой точки (центр эксперимента) добавляют значение шага α , измененное пропорционально соответствующим коэффициентам регрессии, в результате чего получают координаты новой точки, расположенной в направлении градиента функции отклика.

При переходе от кодированного факторного пространства к натуральному градиент меняет направление на противоположное, поэтому движение по градиенту следует осуществлять только в натуральном факторном пространстве.

Переход осуществляется по (9.37). Подставив найденные с помощью формулы (9.37) кодированные значения факторов в уравнение регрессии, получим значения коэффициентов, соответствующие натуральному факторному пространству.

Единых правил для выбора шага не существует. Целесообразно иметь такой шаг, чтобы первая точка крутого восхождения располагалась вне области эксперимента. Однако сделанный шаг не должен выводить точку за пределы области определения факторов. Обычно шаг подбирают для фактора, коэффициент при котором наибольший по абсолютной величине. Пусть таковым является фактор $\bar{x}_\mu$ и для него выбран шаг $\delta \bar{x}_\mu$.

Шаги по остальным факторам рассчитывают так, чтобы полученная точка располагалась в направлении гра-

диента, а фактор $\tilde{x}_\mu$ изменился на величину $\delta\tilde{x}_\mu$.

Пропорциональный пересчет шага для j -го фактора выполняют по формуле

$$\delta\tilde{x}_j = b_j J_j \frac{\delta\tilde{x}_\mu}{b_\mu J_\mu} \quad (j = 1, 2, \dots, k), \quad (9.50)$$

после чего определяются координаты новой точки $\{\tilde{x}_1^0 + \delta\tilde{x}_1, \tilde{x}_2^0 + \delta\tilde{x}_2, \dots, \tilde{x}_k^0 + \delta\tilde{x}_k\}$. Прибавляя к полученным координатам новой точки значения соответствующих шагов $\delta\tilde{x}_j$, получают условия эксперимента на втором шаге и т. д.

Обязательная реализация плана эксперимента на каждом шаге крутого восхождения не требуется. Целесообразно переходить от шага к шагу,

если значение параметра оптимизации y , рассчитываемое как средний результат серий опытов, в каждой новой точке больше, чем в предыдущей. Если на очередном шаге параметр оптимизации не изменился или уменьшился, целесообразно предыдущую точку принять за центр нового плана с тем, чтобы уточнить уравнение регрессии в данной точке, и, возможно, определить новое направление градиента. Если улучшить ранее достигнутый результат и после этого не удастся, можно прекратить эксперимент (поскольку область экстремума достигнута) или перейти к планам более высокого порядка, позволяющим учесть дополнительные нелинейные (например, квадратичные) члены регрессии и, следовательно, более точно описать поверхность отклика в области экстремума.

Часть III

МЕТОДЫ ВЫРАБОТКИ РЕШЕНИЙ

Глава 10. Процесс выработки решений при исследовании эффективности и качества технических систем

1. ВЕРТИКАЛЬНАЯ ДЕКОМПОЗИЦИЯ ПРОЦЕССА ВЫРАБОТКИ РЕШЕНИЙ ПРИ ИССЛЕДОВАНИИ ЭФФЕКТИВНОСТИ И КАЧЕСТВА

Неопределенность компонентов модели проблемной ситуации (см. гл. 4) приводит к необходимости рассматривать задачу принятия решений с позиций системного подхода, т. е. проводить исследования по последовательно убывающим уровням обобщения с учетом комплекса межуровневых и внутриуровневых взаимосвязей.

Например, с общесистемных позиций любая операция, реализуемая в рамках S_0 -системы, может быть декомпозирована как по элементам, так и по процессам (подпроцессам). Элементами операции являются активные ресурсы (см. рис. 1 гл. 4), условия и способы ее проведения. Процессы, сопровождающие операцию, могут быть условно разделены на целевые, которые непосредственно приводят

к целевому результату Y^r , и функциональные — обеспечивающие целевые процессы. К последним, как правило, относят: сбор и обработку информации, планирование, управление и контроль и т. д.

Выводы, вытекающие из исследования операции, и рекомендации для выработки обоснованного суждения с последующим принятием решения могут быть конструктивными (т. е. носить нетривиальный характер в отличие от суждений типа «чем больше, тем лучше») лишь в том случае, если они базируются на количественных или числовых оценках. Последнее обстоятельство требует привлечения метода математического моделирования в качестве основного средства исследования эффективности операции. Выявленные при экспериментировании с построенной моделью свойства и закономерности затем обобщают и переносят на реальный объект исследования. Формально процесс исследования может быть представлен следующей совокупностью отображений:

$$\left. \begin{aligned} \{X, R\} \times T &\rightarrow \{X, R\} \xrightarrow{\theta^{(k)}} \{X^r(k), R^r(k)\} \times T^r \rightarrow \\ &\rightarrow \{X^r(k), R^r(k)\} \Rightarrow \theta_z^r(k) \Rightarrow \theta_z(k) \Rightarrow \theta^{(k+1)} = \\ &= \theta^{(k)} \cup \theta_z(k), \quad k = 0, 1, 2, \dots, \end{aligned} \right\} \quad (10.1)$$

где $\{X, R\}$ — множество элементов X S_0 -системы и связей между ними; T — множество моментов времени; $\{X, R\} \times T \rightarrow \{X, R\}$ — процесс функционирования S_0 -системы; $\{X^r(k), R^r(k)\} \times T^r \rightarrow \{X^r(k), R^r(k)\}$ — гомоморфное отображение процесса $\{X, R\} \times T \rightarrow \{X, R\}$, т. е. модель операции; $\theta^{(k)}$ — информация об операции на k -м шаге исследования; $\theta_z^r(k)$ — новая информация, полученная по результатам моделирования на k -м шаге; $\theta_z(k)$ — новые знания на k -м шаге, полученные на основе информации $\theta_z^r(k)$.

Таким образом, (10.1) вскрывает сущность исследования как процесса получения новых знаний об объекте исследования (S_0 -системе) для принятия решения. Качество нового знания $\theta_z(k)$ (его достоверность и прагматика), приобретенного на k -м шаге исследования, зависит, во-первых, от логической организации процесса исследования, задаваемого структурой составного отображения (10.1); во-вторых, от адекватности отображения (модели операции)

$$\{X, R\} \times T \xrightarrow{\theta^{(k)}} \{X^r(k), R^r(k)\} \quad (10.2)$$

целям исследования¹; в-третьих, от способов получения, обработки и анализа результатов для формирования информации $\theta_z^r(k)$; в-четвертых, от интерпретации информации $\theta_z^r(k)$ в элементы новых знаний $\theta_z(k)$ об операции; в-пятых, от обобщения новых непротиворечивых сведений в новую информацию (знание) $\theta^{(k+1)}$.

Это превращает задачу принятия решения в многошаговый процесс, который может быть декомпозирован как по «вертикали», так и по «горизонтали». Именно такая декомпозиция

и определяет логическую организацию процесса исследования и принятия решения. Множество $\{k\}$ шагов исследования формально можно представить номерами элементов декартова произведения:

$$\mathcal{X}^B \times \mathcal{X}_k^{r^B} \Rightarrow \{k\}, \quad (10.3)$$

где $\mathcal{X}^B = \{k^B\}$ — множество уровней принятия решений «вертикальной» декомпозиции; $\mathcal{X}^r = \{k^r\}$ — множество шагов исследования «горизонтальной» развертки уровня $k^B \in \mathcal{X}^B$.

Хотя формально в (10.1) множество $\{k\}$ является бесконечным, тем не менее при решении всех практических задач оно имеет верхнюю грань, которая определяется обобщенным ресурсом на проведение исследования (отпущенное время, вычислительные ресурсы, уровень прикладных и общенаучных знаний).

Способ представления модели (10.2) зависит от цели исследования. Если целью исследования является лишь установление наиболее общих закономерностей для принятия решений относительно выработки долговременных концепций развития большой технической системы, то модель (10.2) должна строиться в наиболее общей форме, предполагающей высокую степень агрегирования исходной информации для выявления тренда. Построение чрезмерно детальной модели в этом случае зачастую не только невозможно, но и просто вредно, так как излишний учет факторов затрудняет выявление «целевой» тенденции среди возникающих побочно. Кроме того, в этом случае неоправдано растут затраты на построение модели и экспериментирование с ней, обработку и обобщение получаемой информации, что снижает мощность множества $\{k\}$ и, следовательно, снижает полноту нового знания $\theta^{(k+1)}$.

В основе декомпозиции процесса принятия решения при исследовании эффективности и качества технических систем лежит важнейший общесистемный принцип — принцип внешнего дополнения. Согласно этому принципу проверка истинности получаемых результатов (информации $\theta_z(k)$) возможна только в рамках метасистемы

¹ Поскольку при исследовании эффективности операция наиболее часто используют критерии пригодности или оптимальности, то оценивание эффективности проводится по конечному состоянию $\{X^r, R^r\}$. В случае использования критерия адаптивности правая часть (10.2) трансформируется в $\{X^r(k), R^r(k)\} \times T^r$.

относительно исследуемой технической системы.

В любом из этих случаев необходимо, прежде всего, сформировать комплекс условий функционирования рассматриваемой системы. Комплекс условий понимается как некоторый необходимый минимум дополнительной информации, позволяющей «замкнуть» рассматриваемую техническую систему (сформировать модель цели операции для нее, установить условия проведения операции) и свести рассматриваемую проблему исследования к одной из типовых задач принятия решения.

Таким образом, процесс исследования и принятия решения носит иерархический характер («вертикальная» декомпозиция), т. е. осуществляется последовательное привлечение информации от выше- и нижележащих системных уровней исследования с детализацией ее на рассматриваемом уровне («горизонтальная» декомпозиция).

В соответствии с этим «вертикальная» декомпозиция задает своеобразную иерархию уровней исследования, состоящую как минимум из двух соподчиненных этапов принятия решения — концептуального и операционального. Соподчиненность не означает строгую зависимость, а отражает диалектическое единство уровней принятия решений. Хотя концептуальный уровень и является ведущим, в то же время результаты операциональных исследований обуславливают соответствующую корректировку задачи принятия решения на концептуальном уровне.

При исследовании эффективности и качества больших технических систем возникает необходимость во введении и третьего уровня принятия решений, подчиненного операциональному уровню, на котором проводится детальный анализ качества элементов технической системы и вырабатываются комплексные решения по его удовлетворению.

Таким образом, выделяют следующие уровни принятия решения: $k^B = I$ — концептуальный (вышележащий); $k^B = II$ — операциональный

(заданный); $k^B = III$ — детальный (нижележащий).

Концептуальный уровень — это уровень принятия решений относительно целей функционирования S_0 -системы и внешних условий, т. е. относительно комплекса условий функционирования исследуемой технической системы. Он обеспечивает выявление рациональной логики процесса функционирования и его информационную достаточность, является тем связующим звеном, которое придает целенаправленный, логически увязанный характер всему процессу принятия решений. Принимаемые на этом уровне решения должны базироваться на наиболее надежной информации, носящей, как правило, качественный характер. Эта информация вскрывает общую структуру предпочтений $\mathcal{P}$ ЛПР, обеспечивающую содержательность и непротиворечивость принимаемых решений на операциональном и детальном уровнях.

На операциональном уровне проверяют степень выполнения частных задач (подцелей) операции и синтезируют функциональную структуру технической системы, участвующей в операции, если решается задача разработки (проектирования) системы. На этом же уровне принимают решения по формированию требований к подсистемам (агрегатам, узлам) системы.

Детальный уровень предполагает анализ качества элементов подсистем и при необходимости решение задачи инженерного синтеза.

В случае выявления противоречий и «узких» мест в ходе принятия решений по уровням в направлении «сверху вниз» происходит возврат на вышележащий уровень с целью коррекции ранее принятых решений.

Такая процедура предусматривает уточнение принимаемых частных решений и модели предпочтений $\mathcal{P}$ на основе выявления тех элементов модели проблемной ситуации, которые были определены не до конца. Подобная корректировка с движением «снизу вверх» представляет собой обратные связи в «вертикальной» декомпозиции. Именно они обеспечивают непротиворечивость получаемой новой ин-

формации для принятия решения. Полнота этой информации обеспечивается в результате пошагового расширения моделей принятия решения при «горизонтальной» декомпозиции процесса, описываемого (10.1).

Этот процесс обеспечивает «вложенность» модели предпочтений и «вложенность» эффективного множества альтернатив:

$$\mathcal{P}^0 \subseteq \mathcal{P}^1 \subseteq \mathcal{P}^2 \subseteq \dots \subseteq \mathcal{P}^*; \\ U \supseteq U_1 \supseteq U_2 \supseteq \dots \supseteq U^*. \quad (10.4)$$

При этом «вложенность» достигается не просто некоторым расширением априорной информации $\theta^{(k-1)}$ в результате различного рода непротиворечивых и содержательных сообщений с выше- и нижележащих уровней, но и получением качественно новой информации путем эвристического (интеллектуального) анализа имеющейся информации. Именно ЛПР обеспечивает концептуальную согласованность всей информации при выработке решения.

Приведенная информационно-логическая схема принятия решения («вертикальная» декомпозиция) носит универсальный характер в том смысле, что она может быть применена для анализа системы любой степени детализации (до тех пор, пока она остается системой в понимании исследователя). В этом случае сама исследуемая техническая система (подсистема, агрегат, узел) фиксируется на операциональном уровне; концептуальный уровень формируется для вскрытия целей ее функционирования на основе установления общих тенденций развития метасистемы (установления комплекса условий функционирования) и привлекается информация об элементах технической системы и их возможностях с нижележащего детального уровня.

2. ОСОБЕННОСТИ ВЫРАБОТКИ РЕШЕНИЙ ПРИ ПРОВЕДЕНИИ ИССЛЕДОВАНИЙ ПО УРОВНЯМ «ВЕРТИКАЛЬНОЙ» ДЕКОМПОЗИЦИИ

Цели и задачи каждого уровня «вертикальной» декомпозиции могут быть решены лишь в том случае, если

указано некоторое общее свойство, по которому объекты рассматриваемого уровня принятия решения (концепция проведения операции, стратегия применения технической системы, элемент системы) различаются по предпочтительности с точки зрения оперирующей стороны. Таким обобщенным динамическим свойством для концептуального уровня принятия решения является «потенциальная эффективность».

Термин «потенциальная эффективность» отражает то обстоятельство, что при принятии решений, направленных на оценку принципиальной приемлемости той или иной концепции проведения операции, на установление возможных стратегий ЛПР и других субъектов S_0 -системы или оценку влияния отдельных факторов, используется лишь наиболее общая агрегированная информация качественного характера.

Потенциальная эффективность операции отражает способность (A — качество) активных средств, используемых в операции, решить поставленную задачу. Оценка потенциальной эффективности в общем виде заключается в вынесении суждения о степени достижения цели операции при идеальном способе использования активных средств. При этом ЛПР сознательно абстрагируется от некоторых действительных ограничений, накладываемых связями между реально существующими элементами, системами или процессами, рассматривая их как некоторое объединение «идеальных» объектов (элементов, систем, процессов), которые могут образовывать необходимые операциональные структуры и при этом не терять своих свойств (качеств). Само образование этих структур можно интерпретировать как идеальное управление элементами, системами или процессами.

Таким образом, потенциальная эффективность характеризуется верхней границей значений показателя эффективности (предельной эффективностью). Поэтому часто оценивание потенциальной эффективности называют верхним оцениванием.

Множество «идеальных» объектов (элементов, систем, процессов), с по-

мощью которого можно сформировать любую виртуальную (мысленно допускаемую) операциональную структуру, приводящую к достижению цели A_0 операции хотя бы в принципе, называют структурно-операциональным базисом $\{X^1\}$ (т. е. $k^B = 1$).

После введения отношения R^1 на множестве элементов структурно-операционального базиса $\{X^1\}$ полученную систему $\{X^1, R^1\}$ (вариант операциональной структуры) рассматривают как идеальное представление потенциальных возможностей подсистем и элементов нижележащих уровней, а их объединение с конкретной гипотезой поведения определяет концепцию проведения операции:

$$\{X^1, R^1\} \times T^1 \Rightarrow \{X^1, R^1\}_G. \quad (10.5)$$

В отображении (10.5) вариант операциональной структуры $\{X^1, R^1\}$ и множество моментов операционального времени T^1 выбираются таким образом, чтобы обеспечить наибольшую «близость» (в смысле некоторой метрики, построенной на основе модели предпочтений $\mathcal{P}_G$) исхода $\{X^1, R^1\}_G$ операции к ее желаемой цели — исходу $\{X^1, R^1\}_{A_0}$. Целевое состояние $\{X^1, R^1\}_{A_0}$ на концептуальном уровне принятия решения задается достаточно нечетко. Например, оно может описываться некоторой областью желаемых состояний в пространстве $\{X^1, R^1\}$, если оперирующая сторона располагает необходимой информацией, чтобы описать границы этой области, или оно может задаваться как некоторая «идеальная» (принципиально недостижимая) точка, к которой целесообразно стремиться.

В первом случае пространство желаемых состояний может задаваться, например, с помощью аппарата теории нечетких множеств (в смысле Л. А. Заде), а для вынесения суждения о потенциальной эффективности целесообразно применять критерий пригодности по заданному уровню степени принадлежности результата к нечеткому желаемому состоянию или критерий оптимальности, максимизирующий эту степень принадлежности.

Во втором случае «идеальную» точку обычно устанавливают в «бесконечности», а для задания отношения предпочтения используют либо критерий оптимальности, либо критерий адаптивности.

Использование критериев адаптивности и оптимальности предполагает получение достаточно «сильной» информации, позволяющей «свернуть» частные характеристики y_i исхода $\{X^1, R^1\}_G$ в некоторую скалярную функцию (обобщенный показатель потенциальной эффективности). Если информации для получения обобщенного показателя недостаточно, то строится решающее правило (составной критерий) на основе качественной информации, и в этом случае отношение предпочтения может оказаться несвязным. В промежуточном случае может быть указано некоторое множество возможных «сверток» частных характеристик (показателей) исхода $\{X^1, R^1\}_G$ и (или) множество параметров для одного вида «свертки», что также приведет в общем случае к выделению некоторого недоминируемого подмножества множества исходов $\{X^1, R^1\}_G$.

На основе введенного понятия потенциальной эффективности различные концепции проведения операции упорядочиваются по предпочтительности на исходах $\{X^1, R^1\}_G$, после чего подмножество исходов, соответствующих «нехудшим» концепциям, рассматривают в качестве потенциально достижимых целей:

$$\{X^1, R^1\}_G = \{X^1, R^1\}_{A_0^{II}} \quad (10.6)$$

для операционального уровня, задавая для него внешнее дополнение. В (10.6) A_0^{II} — множество подцелей (задач), получаемое в результате декомпозиции цели A_0^I для операционального уровня.

Подмножество «нехудших» (недоминируемых¹) концепций и их вариантов, получаемых в результате «привязки» вариантов $\{X^1, R^1\}$ операциональной структуры к определен-

¹ Смысл понятия «недоминируемый» определяется на с. 184, поэтому здесь термины «нехудший» и «недоминируемый» являются синонимами.

ным моментам операционального времени T^I [см. (10.5)], должно быть достаточно широким. Это в дальнейшем обеспечит свободу выбора решений на нижележащих уровнях принятия решений.

Если в результате анализа на концептуальном уровне выделенной оказалась лишь единственная недоминируемая концепция (вариант) или ЛПР считает, что их множество узко, то необходимо расширить представительный набор альтернатив. Обычно это осуществляется одним из двух возможных способов: генерацией новых концепций проведения операции или назначением некоторых уступок по характеристикам Y^{TP} желаемого состояния (или по значениям компонент показателя эффективности W , соответствующим наилучшему исходу).

Учитывая, что все концепции проведения операции формируют с использованием одних и тех же элементов $\{X^I\}$ структурно-операционального базиса, логично ожидать, что предпочтения на альтернативных концепциях слабо изменятся в ходе конструктивного раскрытия элементов $\{X^I\}$ и R^I на операциональном, а затем и детальном уровнях.

Это означает, что если альтернатива была «сильно» доминируемой (значения частных характеристик y_i или показателя W для исхода $\{X^I, R^I\}_G$ при данной альтернативе существенно хуже аналогичных характеристик других альтернатив) на уровне концептуального исследования, то она будет доминироваться и на остальных нижележащих уровнях.

Если некоторые альтернативы относились к классам недоминируемых или «слабо» доминируемых в указанном смысле, то предпочтения между ними могут незначительно измениться, причем нарушение порядка следования альтернатив по предпочтительности зависит от допускаемых уступок Δy_i^{TP} или ΔW . При этом чем больше допускаемая уступка, тем в среднем меньше степень нарушения порядка следования альтернатив по предпочтительности. Нарушения в порядке предпочтительности альтернатив могут появиться также вследствие того,

что при формировании отношения R^I на элементах $\{X^I\}$ не всегда оказывается возможным заранее учесть эмерджентные свойства получаемых операциональных структур $\{X^I, R^I\}$.

Для каждой концепции проведения операции из выделенного их недоминируемого подмножества на операциональном уровне формируется общий для всех концепций структурно-функциональный базис $\{X^{II}\}$. На элементах базиса $\{X^{II}\}$ строятся функциональные структуры $\{X^{II}, R^{II}\}$ (в общем случае и подпроцессы $\{X^{II}, R^{II}\} \times T^{II} \Rightarrow \{X^{II}, R^{II}\}$).

После выделения базиса $\{X^{II}\}$ решается задача оценки эффективности процесса проведения операции:

$$\{(\{X^{II}, R^{II}\} \times T^{II}), R^I\} \times T^I \Rightarrow \Rightarrow \{X^I, R^I\}_{A_o^{II}} = \{X^I, R^I\}_G, \quad (10.7)$$

которая при заданных на концептуальном уровне отношении R^I и операциональном времени T^I декомпозируется на задачу оценки эффективности подпроцесса:

$$\begin{aligned} \{X^{II}, R^{II}\} \times T^I &\Rightarrow \\ \Rightarrow \{X^{II}, R^{II}\}_{A_o^{II}}|_{R^I, T^I} &= \{X^I\}_G \subset \\ \subset \{X^I, R^I\}_G|_{R^I}. & \quad (10.8) \end{aligned}$$

В зависимости от полноты исходной информации с нижележащего detailного уровня задача (10.8) может решаться по критериям пригодности, оптимальности или адаптивности для обеспечения достаточно представительного множества недоминируемых подпроцессов.

На детальном уровне исследования вновь формируется теперь уже структурно-элементный базис $\{X^{III}\}$, на основе которого синтезируются элементные структуры $\{X^{III}, R^{III}\}$ (или элементарные подпроцессы $\{X^{III}, R^{III}\} \times T^{III}$) и решается задача оценки качества элементов $\{X^{III}\}$:

$$\begin{aligned} \{(\{X^{III}, R^{III}\} \times T^{III}), R^{II}\} \times T^{II} &\Rightarrow \\ \Rightarrow \{X^{III}, R^{III}\}_{A_o^{III}} &= \{X^{II}, R^{II}\}_G. \end{aligned} \quad (10.9)$$

Эта задача при фиксированных R^{II} и T^{II} также декомпозируется на задачу оценки качества элементарного подпроцесса:

$$\begin{aligned} \{X^{III}, R^{III}\} \times T^{III} &\Rightarrow \\ &\Rightarrow \{X^{III}, R^{III}\}_{A_0^{III}}|_{R^{II}, T^{II}} = \\ &= \{X^{II}\}_G \subset \{X^{II}, R^{II}\}_G|_{R^{II}}. \quad (10.10) \end{aligned}$$

Концептуальная диаграмма принятия решения при исследовании эффективности и качества, соответствующая задачам (10.5)–(10.10), приведена на рис. 1.

На концептуальном уровне исследования можно оценить лишь потенциальную степень достижения цели $A_0 \equiv A_0^I$ операции (потенциальную эффективность) применительно к той или иной концепции ее проведения. На операциональном уровне может быть оценена степень выполнения частных задач операции, а на детальном уровне — степень удовлетворения качества элементов $\{X^{III}\}$ предъявляемым к ним требованиям.

Окончательное суждение о предпочтительности тех или иных концепций, подпроцессов или элементов S_0 -системы (в смысле достижения цели операции) выносится на основе результатов оценивания потенциальной эффективности, эффективности и качества.

Обозначим через W показатель достижения цели операции, через $W_{п.э}$ — показатель потенциальной эффективности концепции проведения операции [процесса, задаваемого (10.5)], а через $W_э$ и $W_к$ — показатели эффективности и качества соответственно процессов, задаваемых (10.8) и (10.10). Тогда W можно представить как некоторую композицию (символ композиции — $\otimes$).

$$\begin{aligned} W &= W_{п.э} \otimes W_э (W_{п.э}) \otimes \\ &\otimes W_к (W_{п.э}, W_э). \quad (10.11) \end{aligned}$$

В выражении (10.11) запись $W_э(W_{п.э})$ и $W_к(W_{п.э}, W_э)$ означает, что соот-

ветствующий показатель определяется концепцией проведения операции, а следовательно, и целевыми требованиями к процессам, подпроцессам и элементам S_0 -системы.

Выбор рациональной стратегии проведения операции (выработка решений) является основой управления. Цель управления — максимальное использование потенциальных возможностей, заложенных в систему, для решения поставленных задач. Следовательно, управление тем эффективнее, чем ближе значение показателя эффективности W операции к предельному его значению $W_{п.э}$, т. е. к потенциальной эффективности. Поэтому в качестве показателя эффективности управления, естественно использовать относительное значение показателя эффективности операции

$$W'_{упр} = \frac{W_{п.э} - W}{W_{п.э}}. \quad (10.12)$$

Эффективность управления тем выше, чем меньше величина $W'_{упр}$; поэтому используют следующее допустимое преобразование:

$$W_{упр} = 1 - W'_{упр}, \quad (10.13)$$

приводящее в соответствие увеличение показателя эффективности $W_{упр}$ управлению увеличению эффективности управления; $0 \leq W_{упр}, W'_{упр} \leq 1$. Понятие эффективности управления является достаточно широким. Действительно, анализ структуры показателей (10.11) и (10.12) показывает, что суждение об эффективности управления в целом можно сделать на основании суждений об эффективности процессов, протекающих в системе (показатель $W_э$), и качестве элементов системы (показатель $W_к$). Например, если концепция проведения операции определена и заданы подпроцессы, то можно предъявить требования к качеству элементов системы (если фиксированы связи R^{III} между ними) или установить необходимые связи на элементах системы (если элементная база X^{III} задана), исходя из требуемого уровня W^{TP} эффективности операции.

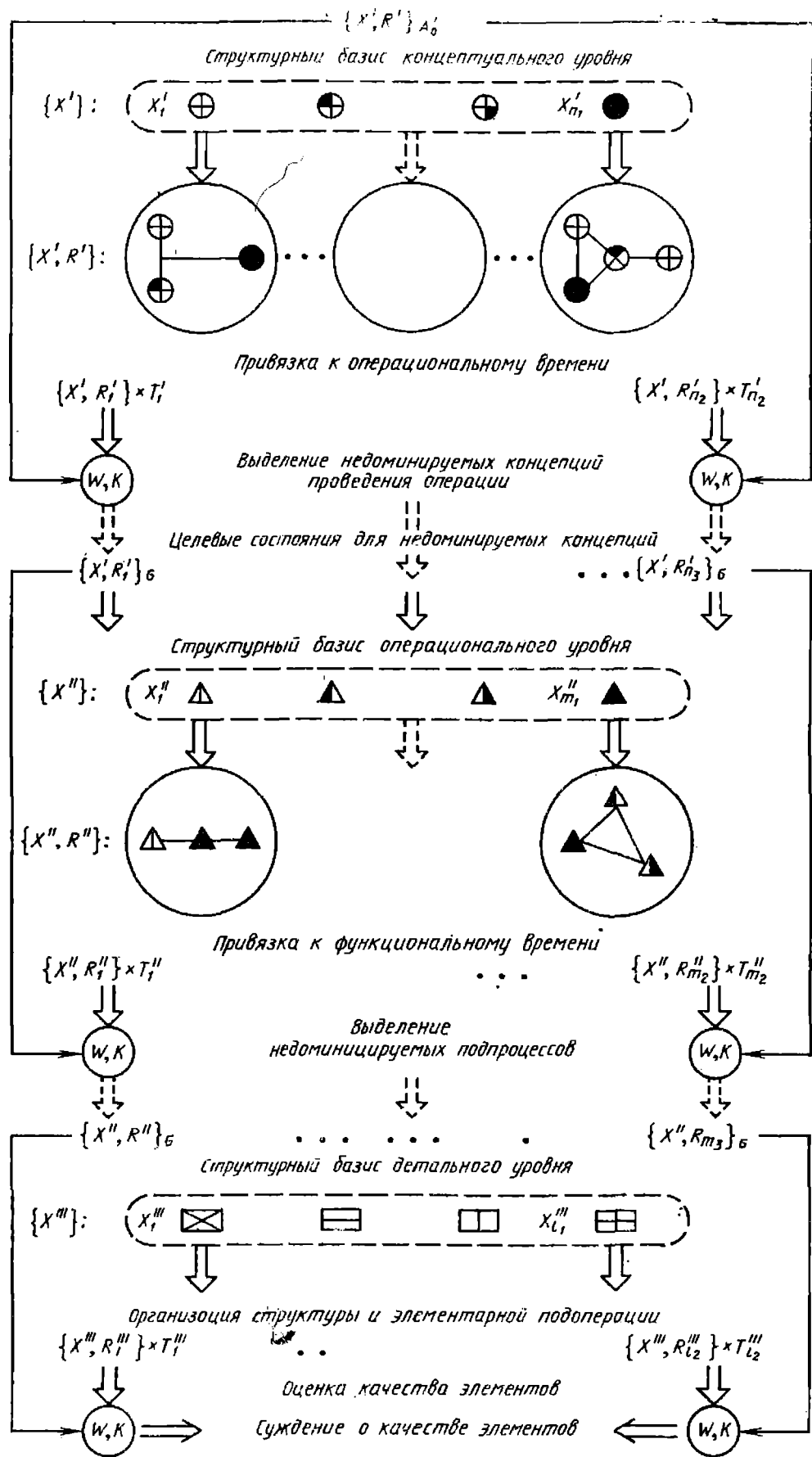


Рис. 1. Концептуальная диаграмма принятия решения при исследовании эффективности и качества

В этом случае могут ставиться следующие задачи инженерного синтеза:

$$X^{III}: \max_{X^{III} \in X(R^{III}, T^{III}, W_3)} W_K(X^{III})$$

или $X^{III}: W_K(X^{III}) \geq W_K^{TP}(W_3);$ (10.14)

$$R^{III}: \max_{R^{III} \in R(X^{III}, T^{III}, W_3)} W_K(R^{III})$$

или $R^{III}: W_K(R^{III}) \geq W_K^{TP}(W_3).$ (10.15)

Задача, определяемая (10.14), является задачей разработки элементной базы с характеристиками качества, обеспечивающими наибольший либо требуемый уровень показателя W_K при заданных элементной структуре R^{III} узлов и агрегатов и процессах, протекающих в системе. Задачу, определяемую (10.15), можно интерпретировать, например, как задачу повышения надежности узлов и агрегатов путем выбора способа связи R^{III} резервируемых элементов $\{X^{III}\}$. Аналогично может ставиться задача отыскания моментов времени из множества $\{T^{III}\}$, задающих последовательность элементарных подпроцессов, обеспечивающих требуемое качество работы узлов и агрегатов (например, задача отыскания моментов времени включения элементов, находящихся в холодном или горячем резерве, обеспечивающих требуемое качество переходных процессов в агрегате или узле).

Аналогично задачам, определяемым (10.14), (10.15), ставятся задачи оперативного синтеза, т. е. задачи принятия решения относительно облика подсистем (состав элементов $\{X^{II}\}$ узлов и агрегатов подсистем и связей R^{II} между ними), а также организации взаимодействия между подсистемами в моменты времени из множества $\{T^{II}\}$.

Качество принимаемых решений во многом определяется адекватностью вида композиции (свертки), задаваемой (10.11), цели и механизму операции. Например, если интерпретировать значения показателей, стоящих в правой части (10.11), как меру вклада в степень достижения цели операции,

то свертку показателей можно представить в мультипликативном виде:

$$W = W_{п.э} W_3 W_K. \quad (10.16)$$

Сомножители в (10.16) записаны без аргументов в скобках, что означает их принадлежность к конкретной концепции, а сами они представляют однородные числовые оценки из диапазона $[0, 1]$.

Другим возможным способом организации свертки является функция вида $W = \min \{W_{п.э}, W_3, W_K\},$ (10.17) которая соответствует вынесению суждения по «узкому месту».

Композиция, задаваемая (10.11), может учитывать и эмерджентные свойства системы, проводящей операцию, которые возникают при совместном функционировании подсистем или объединении элементов определенного качества в узлы, агрегаты и подсистемы. При этом (10.11) может быть представлена в виде

$$W = W_{п.э}^{\gamma_I} W_3^{\gamma_{II}} W_K^{\gamma_{III}}; \gamma_I, \gamma_{II}, \gamma_{III} \geq 0. \quad (10.18)$$

Коэффициенты γ_i в (10.18) назначают лица, имеющие опыт создания больших технических систем и хорошо представляющие физику процессов, протекающих в исследуемой системе. Коэффициенты $\gamma_I, \gamma_{II}, \gamma_{III}$ характеризуют предельное процентное увеличение показателя W при возрастании показателей $W_{п.э}, W_3, W_K$ на 1 % соответственно.

В случае выявления противоречий на том или ином уровне принятия решений осуществляется последовательный возврат к вышележащим уровням для коррекции целевых требований к данному уровню и расширения множества альтернатив.

3. ТРЕБОВАНИЯ К ПРОЦЕССУ ВЫРАБОТКИ РЕШЕНИЯ ПРИ ПРОВЕДЕНИИ ИССЛЕДОВАНИЙ ПО УРОВНЯМ «ВЕРТИКАЛЬНОЙ» ДЕКОМПОЗИЦИИ

Информационно-логическая схема принятия решения при исследовании потенциальной эффективности, эф-

фективности и качества, структура которой отображена концептуальной диаграммой на рис. 1, обеспечивает последовательное усложнение и расширение информации для принятия решения от уровня к уровню. Такой процесс принятия решения позволяет уже на ранних этапах исследования «отсечь» заведомо неприемлемые концепции проведения операции, выделить «узкие места», подлежащие дальнейшей проработке, и сосредоточить на них внимание, устранить возникающие противоречия. Этим достигается не только значительное сокращение затрат на исследование, но и обеспечивается оперативный анализ весьма широкого множества конкурирующих вариантов.

Основной акцент в этом процессе приходится на концептуальный уровень, придающий целостный, всесторонний, логически стройный характер всему процессу исследования и выработки решения. Результаты этого уровня во многом, если не во всем, предопределяют генерацию новых идей и научных проработок в смежных областях науки, техники и производства, затрагивая в то же время вопросы выбора направлений технической политики.

Особое место концептуального уровня в общей иерархии процесса принятия решения, его цели и задачи предъявляют вполне определенные требования к организации этого процесса, к принципам выработки решений и самим средствам концептуального исследования (моделям). Эти требования могут быть сформулированы в виде основополагающих принципов концептуального исследования; к ним относятся принципы: эвристической генерации; высокой разрешающей способности; экспресс-анализа; минимального базиса; вложенности; однозначной семантики; веера.

Принцип эвристической генерации требует широкого охвата всех элементов, составляющих содержание основных проблемных вопросов оценки эффективности и принятия решений при исследовании технических систем. Он обеспечивает полноту множества недоминируемых концепций проведения операции на основе не-

формального выявления целей, разделения их на классы «обязательных», «желательных», «возможных» с последующим формированием способов их достижения (хотя бы в принципе), что позволяет решить вопрос о множестве моделей цели операции (т. е. о совокупности показателей и критериев эффективности). При этом критерий эффективности, как составную часть модели цели операции, формируют на основе вскрытия системы предпочтения оперирующей стороны и эвристической рефлексивной генерации предпочтений других субъектов системы.

Согласно принципу высокой разрешающей способности необходимо выбирать такие средства получения и анализа результатов, которые обеспечивают действенность концептуального исследования. Так как целью концептуального исследования является установление наиболее общих закономерностей протекания операции, то используемый на данном уровне аппарат должен обеспечивать высокую разрешающую способность именно к выявлению главных тенденций и закономерностей.

Требование оперативности решения задачи анализа при концептуальном исследовании является одним из основных, так как главная задача концептуального уровня, вытекающая из его основной цели, заключается в выделении недоминируемого множества альтернатив из широкого набора концептуальных вариантов, мощность которого априори не известна. Более того, это множество может расширяться в результате эвристической генерации принципиально новых альтернатив при исследовании по направлению «снизу вверх» или при уточнении систем предпочтений субъектов операции.

Возврат к вышележащему уровню происходит лишь после того, как полностью исчерпаны потенциальные возможности реализации рассматриваемой концепции. Оценка потенциальных возможностей может быть оперативно проведена лишь на основе применения аппарата оптимизации. Поэтому сформулированное требование оперативного анализа альтерна-

тивных вариантов, предполагающего оптимизацию процесса протекания операции, выражается в принципе «экспресс-анализа».

Согласно принципу минимального базиса при формировании концепций проведения операции необходимо использовать по возможности наименьший инвариантный альтернативам набор $\{X^{k^B}\}$, $k^B = I, II, III$ элементов (подопераций, процессов, подпроцессов, подсистем, агрегатов, узлов). При этом каждый элемент, включаемый в набор, должен выводиться из общей согласованной системы допущений о роли, месте и возможностях этого элемента в операции, которая в дальнейшем может быть лишь развита и уточнена¹.

Следование этому принципу обеспечивает не только концептуальную и информационную согласованность выводов и рекомендаций по анализу различных вариантов, но и сравнимость их по предпочтительности на нижележащих уровнях без опасения «потерять» какой-либо недоминируемый вариант (см. рис. 1).

Концептуальная согласованность принимаемых решений, основанная на одном и том же минимальном наборе $\{X^{k^B}\}$, обеспечивается последовательным расширением получаемой новой информации:

$$\theta^{(k)} \subseteq \theta^{(k+1)} \subseteq \dots \subseteq \theta^{(k+l)} \quad (10.19)$$

в ходе не только «вертикальной», но и «горизонтальной» декомпозиции процесса принятия решений (см. (10.1)).

Это позволяет последовательно сужать множество недоминируемых вариантов [согласно (10.4)] на каждом уровне k^B , равном I, II, III исследования. Организация процесса исследования, задаваемого (10.1) в соответствии с (10.19) и (10.4), составляет существо принципа вложенности.

Ряд этапов концептуального исследования носит во многом неформальный характер (формирование цели

и задач операции, описание предпочтений субъектов системы, задание исходного множества альтернатив и др.). Требование системности исследования вызывает необходимость привлечения информации, получаемой от специалистов различного профиля, уровня руководства и компетентности. Это, как правило, приводит к неоднозначной интерпретации различных понятий и получаемых результатов.

Данное обстоятельство не только тормозит исследование, но зачастую вызывает исключение части существенной информации $\theta_z(k)$ (см. (10.1)) об эмерджентных свойствах процесса в силу того, что акцент, придаваемый «узким» специалистом той или иной информации, может не совпадать с ценностью этой информации относительно достижения цели операции.

Следовательно, организация процесса выработки решения должна обеспечивать свободное обсуждение всех возникающих вопросов, однозначную интерпретацию получаемых результатов всеми участниками исследования и организацию обратной связи («исполнители-руководитель»), позволяющую учесть особое мнение того или иного специалиста. Этим обеспечивается реализация принципа однозначной семантики.

Концептуальные исследования проводятся на ранних этапах принятия решений в условиях высокой степени неопределенности исходной информации относительно «механизма» протекания процесса.

Для обеспечения информационной достаточности концептуального исследования и наиболее полного адекватного описания рассматриваемого процесса естественным является последовательный «частный» анализ его отдельных свойств. Это приводит к необходимости рассматривать не одну большую, комплексную модель функционирования системы (построение которой зачастую просто невозможно из-за указанных выше причин), а представительный набор — «веер» моделей.

При этом в зависимости от задачи исследования и сделанных допущений из множества всех мыслимых характеристик $\{Y(G)\}$ исхода G выделяется их некоторое ограниченное подмно-

¹ Развитие и уточнение предполагает содержательное непротиворечивое раскрытие свойств элемента без изменения его основной функциональной концепции.

жество $C_{Y(G)} \subset \{Y(G)\}$. На выбор множества $C_{Y(G)}$ влияют возможности имеющегося математического аппарата описания механизма формирования результата $Y(G)$ и оптимизации его параметров (характеристик y_i , $i = \overline{1, m}$).

Так как получение результата $\{Y(G)\}$ не является самоцелью, а служит основанием для вычисления заданного показателя эффективности W операции, то каждому подмножеству $C_{Y(G)}$ в общем случае соответствует свой набор частных компонент векторного показателя эффективности, т. е. можно говорить о веере моделей «исход — показатель», задаваемых оператором Ψ . Таким образом, каждая модель веера предназначена для анализа лишь ограниченного подмножества характеристик процесса и вскрытия на этой основе некоторых из его общих закономерностей.

Пример 1. При исследовании системы, содержащей манипулятор, рассматривается элементарный процесс функционирования робота в виде операции по перемещению детали со стеллажа к приемному отверстию многоцелевого станка. Исход этой операции описывается следующими пространственно-временными характеристиками: y_1, y_2, y_3 — местоположение детали относительно приемного отверстия; $y_4 = t$ — время доставки.

Заданы требуемые значения y_i^{TP} , $i = \overline{1, 4}$. В зависимости от цели исследования процесса выполнения элементарной операции могут быть построены:

модель среднего результата (ошибки) в конце операции

$$W(u) = M[\rho(y(u))] = \\ = \Psi(u, \Lambda_F, \Lambda_E, y_4 = y_4^{TP}),$$

$$\text{где } \rho(y(u)) = \sqrt{\sum_{i=1}^3 (y_i(u) - y_i^{TP})^2};$$

динамические модели среднего результата операции

$$W_i(u, t) = M[y_i(u, t)] = \\ = \Psi_i(u, t, \Lambda_F, \Lambda_E), \quad i = \overline{1, 3};$$

динамическая модель качества программного управления

$$W_j(u, t) = \\ = M[(y_j(u, t) - y_j^{prog}(t))^2] = \\ = \Psi_j(u, t, \Lambda_F, \Lambda_E, y_j^{prog}(t)), \\ j = \overline{1, 3},$$

где $y_j^{prog}(t)$ — программная траектория j -й характеристики;

оптимизационные модели (ставятся в форме задачи Лагранжа, Майера и Больца, см. т. 2).

Таким образом, веер моделей в общем случае представляет различные варианты декомпозиции «портретной» модели

$$W = \Psi(u, \Lambda_F, \Lambda_E, \Lambda_J^N, \Lambda_J^B), \quad (10.20)$$

которая связывает значение показателя эффективности W со стратегией u ЛПР, а также со значениями определенных Λ_F , случайных Λ_E , природно неопределенных Λ_J^N факторов и факторов поведенческой неопределенности Λ^B (противодействия и содействия).

Вид портретной модели не может быть полностью определен на данном уровне исследования. Однако фиксация отдельных аргументов в (10.20) путем различных допущений о значениях аргументов, выбора принципов их осреднения, выдвижения гипотез о поведении других субъектов системы позволяет: замкнуть каждую из моделей веера; существенно понизить размерность модели; отсеять незначимые факторы среди аргументов (10.20); повысить разрешающую способность каждой модели веера по вскрытию закономерностей процесса путем осреднения неопределенных факторов, которые не вызывают выхода системы за границы области ее гомеостазиса и могут быть парированы управляющими воздействиями.

При формировании веера моделей, прежде всего, следует ориентироваться на цель исследования, а затем выбирать форму представления модели.

Если цель исследования состоит лишь в установлении закономерностей формирования конечного результата

безотносительно динамики развития процесса, то могут быть использованы статические модели, задаваемые, например, системами алгебраических или трансцендентных уравнений, некоторыми оптимизационными задачами, выражениями для финальных вероятностей и т. п.

Если при анализе процесса важным является его развитие, то используют динамические модели, которые также могут быть оптимизационными, а форма представления модели — как дискретная, так и непрерывная (разностные, дифференциальные, интегродифференциальные уравнения).

В зависимости от способа описания гипотез относительно поведения других субъектов S_0 -системы могут быть построены модели как со строгим (стратегические и нестратегические), так и нестрогим соперничеством, учитывающими или не учитывающими способности других субъектов к рефлексивному выбору ими стратегий своего поведения. В последнем случае модель строят как адаптивизационную.

Каждый из семи рассмотренных принципов отражает вполне определенный аспект концептуального уровня принятия решения относительно его синтаксиса, семантики, прагматики.

К синтаксическим относятся принципы эвристической генерации, минимального базиса и вложенности; к семантическим — принцип однозначной семантики и частично вложенности; к прагматическим — принципы высокой разрешающей способности, экспресс-анализа и веера.

Между указанными принципами существует вполне определенная диалектическая взаимосвязь. Например, принципы эвристической генерации и минимального базиса, а также принципы веера, однозначной семантики и высокой разрешающей способности являются взаимопротиворечивыми. Частично снимает это противоречие принцип вложенности.

На рис. 2 представлены принципы концептуального уровня принятия решения, которыми следует руководствоваться при организации процесса, задаваемого выражением (10.1), и дана их мнемоническая и формальная интерпретация. Прагматические прин-

ципы выделены на рисунке одной окружностью; синтаксические — двумя сплошными окружностями; семантические — сплошной и штриховой окружностями (с учетом этого принцип вложенности имеет семантическую и синтаксическую мнемонику).

Принципы концептуального уровня имеют общеметодологическое значение для всех уровней иерархии процесса принятия решения. Однако при выработке решения на других уровнях они трансформируются в принципы, отражающие целевые требования каждого из нижележащих уровней.

На операциональном уровне ($k^B = II$) синтаксические принципы концептуального уровня ($k^B = I$) трансформируются в синтаксический принцип операционально-структурного анализа, семантические — в принцип эмерджентной композиции, а прагматические — в принцип оптимальной информационной связности (рис. 3).

Согласно принципу операционально-структурного анализа на операциональном уровне необходимо формировать структурно-функциональный базис $\{X^{II}\}$, исходя из согласования (вложенности) этих элементов с элементами операционального базиса $\{X^I\}$ вследствие декомпозиции (анализа) его на минимальный набор составных частей, выделяемых, в основном, эвристическими методами.

В результате такой декомпозиции существенно важно восстановить впоследствии [при решении задачи (10.7)] эмерджентные свойства структурно-операционного базиса $\{X^I\}$ и операциональных структур $\{X^I, R^I\}$.

В соответствии с принципом эмерджентной композиции, базирующемся на концептуальных принципах вложенности и однозначной семантики, необходимо выбирать такие связи R^{II} и такие моменты функционального времени T^{II} , чтобы обеспечивалось восстановление (эквивалентность) эмерджентных свойств подпроцесса, описываемого выражением (10.8).

Конечной целью операционального уровня принятия решения является оценка эффективности операции, основанная на оценке степени достижения подцелей (задач) операции.

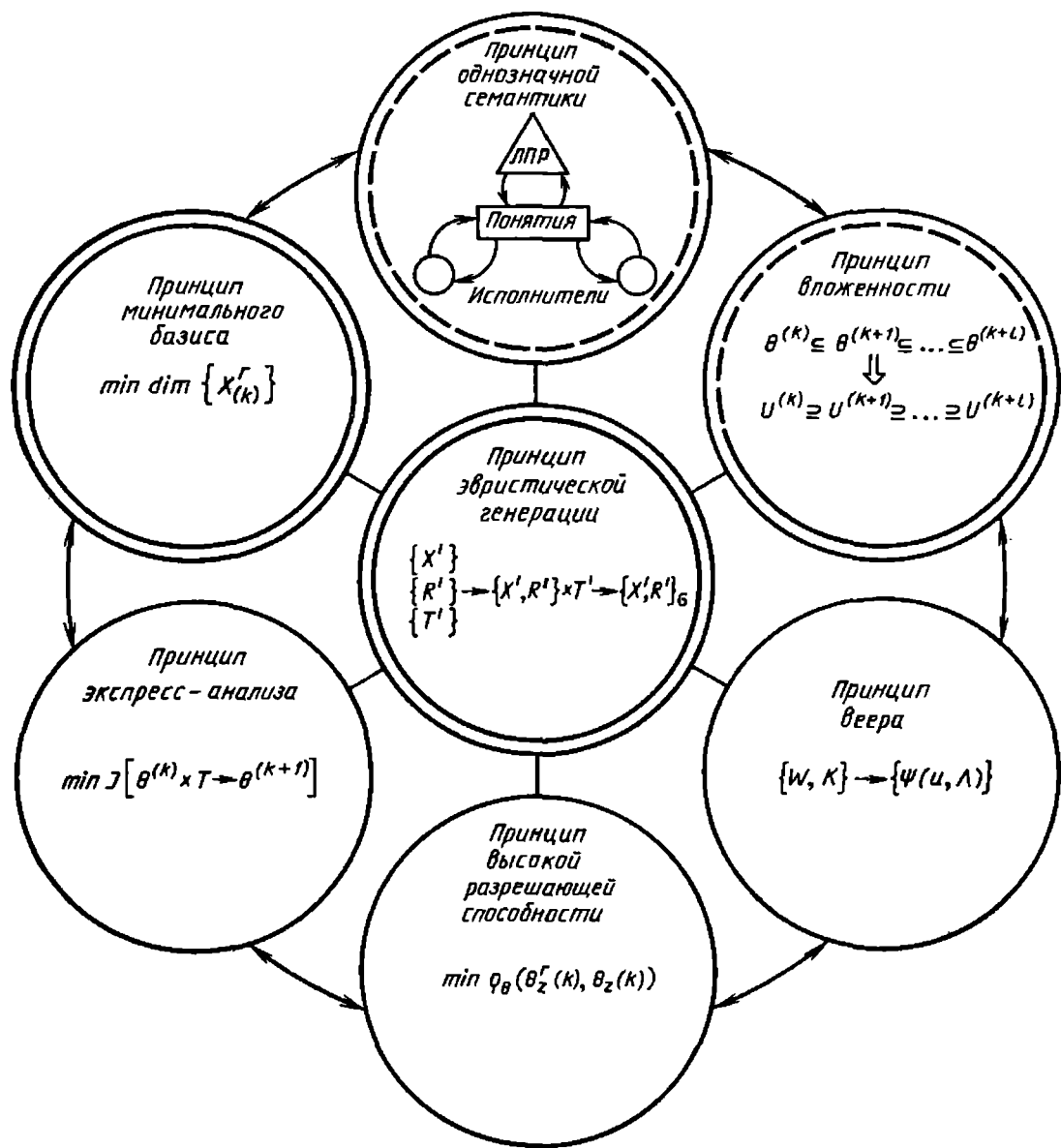


Рис. 2. Принципы концептуального уровня принятия решений

При этом адекватность получаемых выводов и рекомендаций будет определяться:

- правильностью отображения общих закономерностей процесса проведения операции в частные закономерности, свойственные отдельным процессам и подпроцессам решения задач операции;
- наиболее полным учетом различных свойств операции, вскрытых на основе использования веера моделей;
- выбором оптимальных управлений, обеспечивающих наиболее полное использование потенциальных возможностей системы, реализующей операцию;
- высокой точностью оценки эффективности операции.

Все эти требования, основанные на прагматических аспектах концептуального исследования, выражаются принципом оптимально-информационной связности решений операционного уровня.

Для детального уровня принятия решения ($k^B = III$) существенным является принцип функционально-структурного анализа качества. Согласно этому принципу необходимо проводить оценку соответствия качества имеющихся элементов (узлов, агрегатов) или предъявлять к ним обоснованные требования с тем, чтобы обеспечить выполнение заданных функций подсистемами, образованными из элементов $\{X^{III}\}$ посредством организации элементных структур

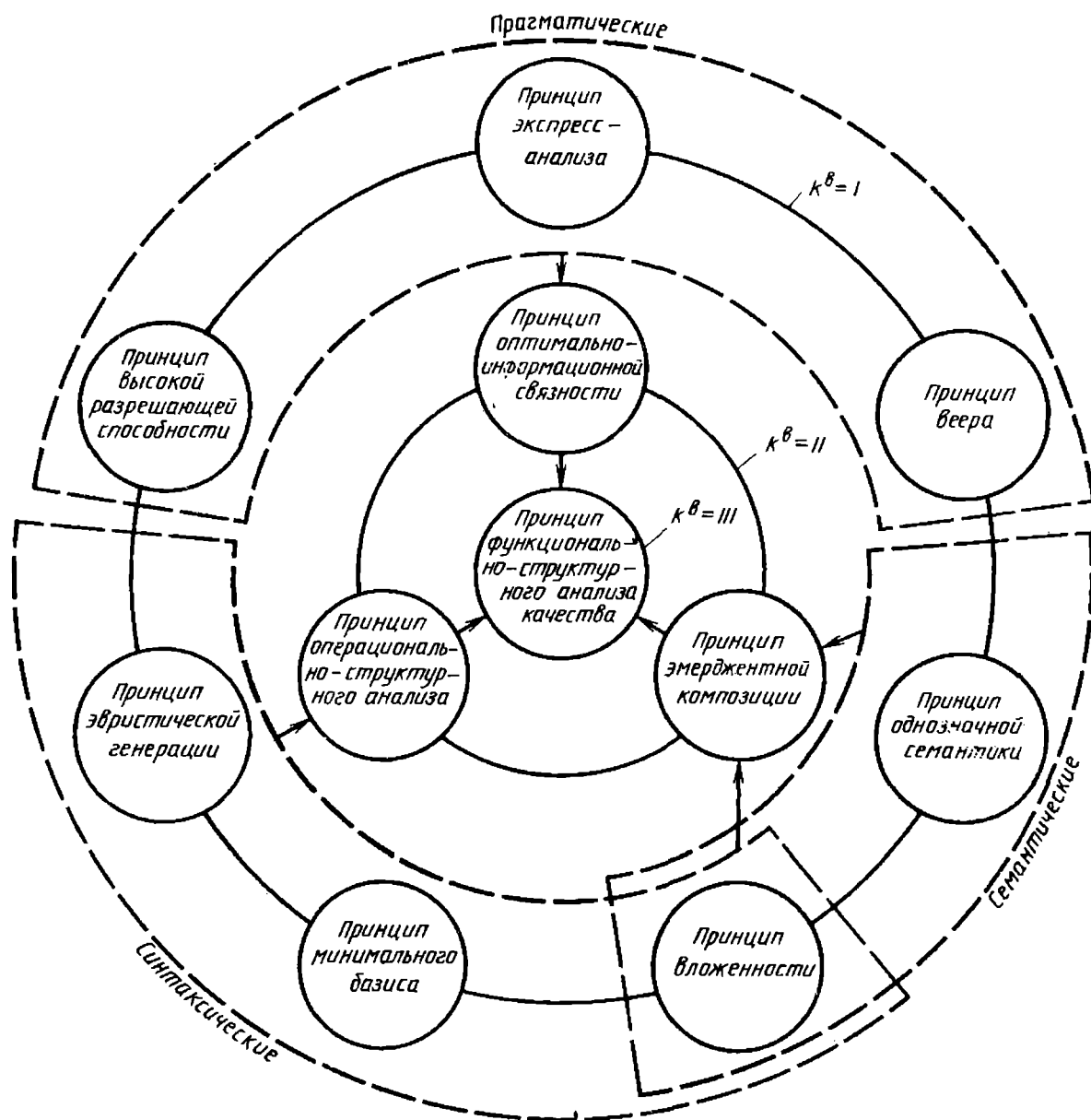


Рис. 3. Иерархия принципов выработки решений при исследовании эффективности и качества

$\{X^{III}, R^{III}\}$ или элементарных подпроцессов $\{X^{III}, R^{III}\} \times T^{III}$.

Соответствие качества оценивают на основе подпроцесса, задаваемого (10.9). Принцип функционально-структурного анализа качества индуцируется принципами операционального уровня.

4. ОБЩАЯ СХЕМА ПРОЦЕССА ПРИНЯТИЯ РЕШЕНИЙ НА ЗАДАННОМ УРОВНЕ ДЕКОМПОЗИЦИИ ИССЛЕДОВАНИЯ

Процесс принятия решения на каждом из рассмотренных уровней включает этапы, тесно увязанные с этапами

оценивания эффективности операции, которые содержат: постановку задачи, выбор показателя и критерия эффективности, оценивание эффективности по результатам математического моделирования, интерпретацию полученных результатов и выработку суждения об эффективности с целью принятия дальнейших решений.

Процесс выработки решения, представляющий логическую последовательность действий на заданном уровне k^B , равном I, II, III, связанных единой целью, включает этапы, представленные на схеме процесса выработки решений на уровне «вертикальной» декомпозиции (рис. 4).

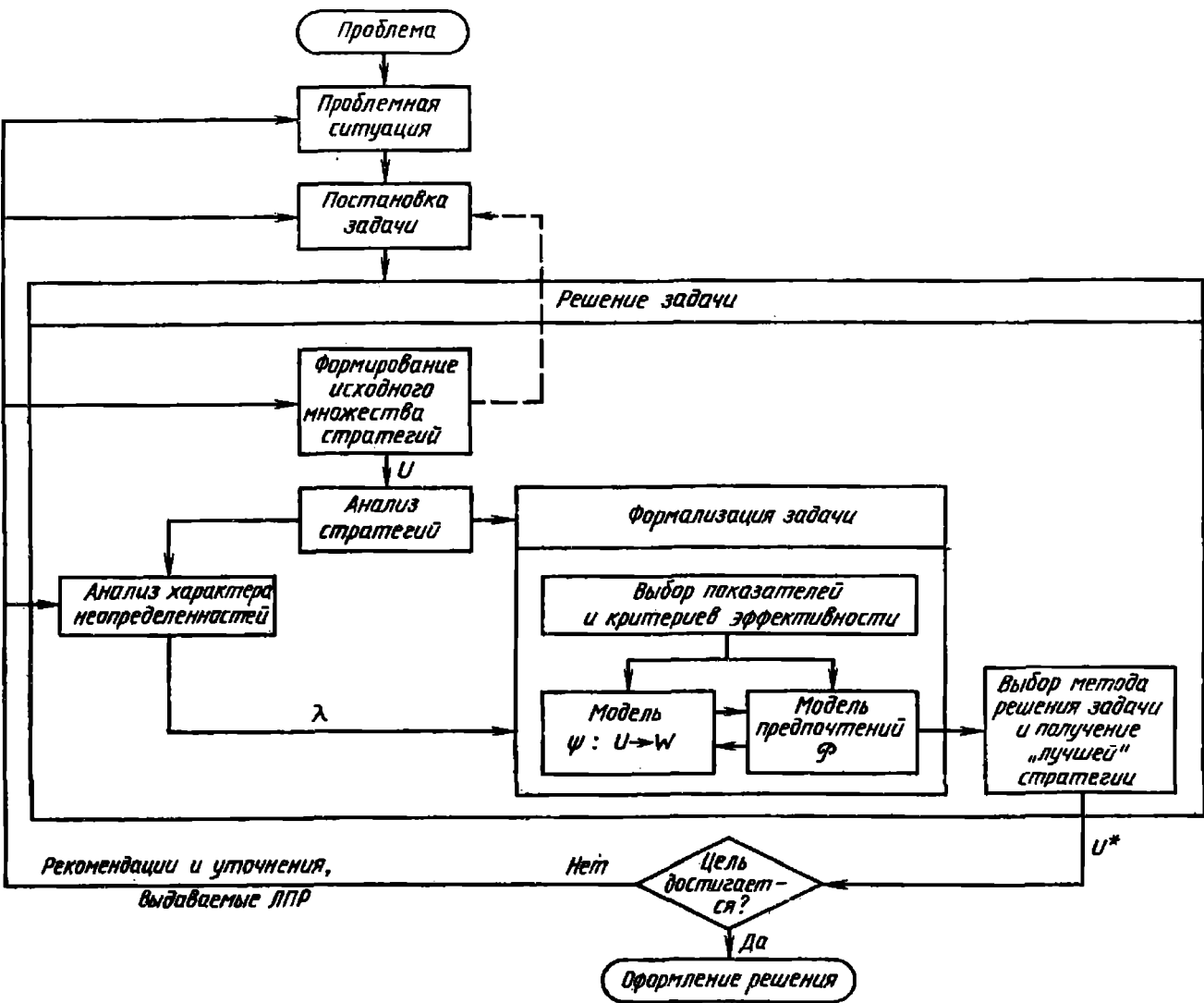


Рис. 4. Схема процесса принятия решения на уровне «вертикальной» декомпозиции

Выработка решения начинается с анализа проблемы и проблемной ситуации. На этих этапах осуществляется структуризация проблемы и устанавливаются цели, проводится их декомпозиция на подцели и задачи, формируются общие условия проведения операции. Этот этап раскрывается в задаче структуризации исходной информации, позволяющей затем гомоморфно отобразить результаты ее решения в исходные данные, используемые при формировании математической модели. Применительно, например, к концептуальному уровню задача может быть раскрыта следующим образом: определить проблемы, установить главное противоречие между устремлением ЛПР и его возможностями, наметить операцию, определить понятие желаемого исхода (результата), установить общие пред-

почтения на элементах проблемной ситуации. В результате такого анализа формулируется постановка задачи, т. е. цель конкретизируется по срокам ее достижения, устанавливаются целевые и другие ограничения, окончательно фиксируется комплекс условий проведения операции. Это позволяет перейти к решению задачи. Решение задачи представляет собой процесс, состоящий из последовательности задач принятия решений, задаваемых выражениями (4.5)–(4.8), (4.10). Особенностью процесса решения указанных задач является его итерационный характер по шагам $k^r \in \mathcal{K}^r$ (k^v) «горизонтальной» декомпозиции, обусловленной наличием обратных связей. Например, этап формирования исходного множества стратегий [за-

дача (4.6)] тесно увязан не только с последующими этапами общей задачи получения (4.2) и анализа (4.3) результатов, но и с первым этапом — постановкой задачи. Это означает, что если не нашлось ни одной стратегии, приводящей к достижению цели ($U = \emptyset$), то необходимо вернуться к этапу анализа проблемы и постановки задачи. Если множество $U \neq \emptyset$, то это не означает, что не требуется корректировка задач — они могут быть расширены или изменены. Кроме того, полученное множество стратегий U оказывается своеобразным внешним дополнением для решения последующих задач: ориентируясь на множество стратегий, определяют требования к моделям H и Ψ и способом учета в них факторов Λ выбирают приемы верификации модели и методы проверки ее адекватности, формируют модели анализа получаемых результатов.

Получение значений показателя эффективности $W(u)$ и построение модели предпочтений $\mathcal{P}$ позволяет выбрать наиболее эффективный метод решения задачи и получить множество $U^* \subseteq U$ «лучших» стратегий, из которого в дальнейшем ЛПР выделяет стратегию $u \in U^*$.

Если цель операции не достигнута, то осуществляется возврат к этапу анализа проблемной ситуации или к любому другому предшествующему этапу. Введение подобной обратной связи позволяет ЛПР не только рационально распределять имеющиеся ресурсы в рамках выбранной концепции проведения исследования, но и формировать новые концепции.

Таким образом, на первых итерациях проводят самый «грубый» учет неопределенности и отыскивают приемлемые стратегии. В этом случае не требуется нахождение сразу «наилучших» решений; поэтому в качестве критерия эффективности используют критерий пригодности. По мере получения дополнительной информации на следующих итерациях применяют более совершенные критерии (либо оптимальности, либо адаптивности).

По достижению желаемого результата процесс выработки решения на рассматриваемом системном уровне

заканчивают и осуществляют переход к следующему уровню «вертикальной» декомпозиции.

После того, как ЛПР удовлетворится результатами решения задач на всех уровнях, осуществляют юридическое оформление решения и его реализацию. Но процесс принятия решений на этом не заканчивается. Остаются еще вопросы, касающиеся проведения в жизнь принятого решения: решения на этапах научно-исследовательских работ и опытно-конструкторских разработок, испытаний на этапах производства образца и его эксплуатации.

На всех указанных этапах жизненного цикла осуществляются контроль работ и коррекция ранее принятых решений.

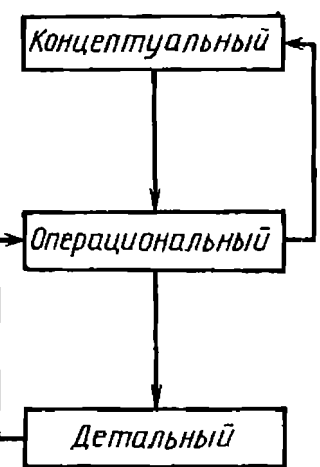
Основным средством исследования эффективности и качества технических систем является математическое моделирование. Поэтому выбор формы представления моделей H и Ψ на каждом уровне процесса принятия решения имеет первостепенное значение. При выборе моделей H и Ψ следует руководствоваться требованиями к процессу принятия решения, сформулированными в виде принципов.

Сравнение достоинств и недостатков моделей Ψ и $\{H\}$, построенных с использованием аналитических и статистических методов моделирования [6, 19, 53], позволяет указать наиболее предпочтительные из них для каждого уровня исследования.

Так, при исследовании основных концепций проведения операции, определении предварительного перечня подцелей операции и задач подсистем сложной технической системы, формировании ее «концептуального» облика необходимо многократно и оперативно решать задачу синтеза через анализ с целью отбраковки заведомо «худших» альтернатив. В этом случае не имеет смысла создавать весьма сложные стохастические модели¹, а модели Ψ и H целесообразно представлять в виде обобщенных аналитических моделей и использовать

¹ Пока не существует эффективного математического аппарата оптимизации сложных стохастических процессов.

1. Иерархия уровней принятия решений

Уровень принятия решения	Объект исследования	Цель исследования	Модель	Показатели и критерии эффективности
	Система	Анализ концепций проведения операции. Определение перечня подцелей и задач, подсистем, условий их функционирования. Формирование «облика» системы	Аналитическая	Степень достижения цели операции. Критерий пригодности. Критерий адаптивности
	Подсистема	Анализ способов выполнения задач подсистемами. Определение обобщенного облика подсистем и средств, общие требования к качеству их элементов	Имитационная	Степень выполнения задач подсистемами. Критерий пригодности. Критерий оптимальности
	Элемент	Детальный анализ качества элементов	Статистическая	Показатели качества элементов (технические характеристики). Критерий оптимальности

методы оптимизации. Такие модели относят к классу концептуальных.

На операциональном уровне исследований, когда определены цели, задачи и условия функционирования подсистемы, выявлена рациональная логика развития операции, можно учесть дополнительные факторы и построить более сложную модель для оценивания эффективности выполнения задач подсистемами сложной технической системы.

Результатом этого этапа являются обобщенный облик подсистем и средств достижения цели, формулировка общих требований к качеству их элементов. Используемые на операциональ-

ном уровне модели, как правило, реализуются в виде сложных имитационных систем.

Уровень детального исследования предполагает создание подробных математических, физических и натурных моделей элементов подсистем для анализа их качества. Так как на этом этапе оперируют, как правило, фактическим материалом, используя методы планирования эксперимента, математической статистики и другие, те модели этого уровня являются в основном статистическими.

В табл. 1 приведены особенности каждого уровня общей иерархии принятия решений и их взаимосвязь.

Такая трехуровневая декомпозиция общей задачи принятия решений позволяет установить жизнеспособность выдвинутой концепции проведения операции, порождает единый (системный) взгляд на операцию и процесс принятия решений как с точки зрения

ее цели, так и с учетом возможности других подсистем и средств. Она позволяет проводить оценку принимаемых решений на нижележащих уровнях с использованием известных методов исследования операций (условия и цели точно определены).

Глава 11. Выявление и измерение предпочтений

Исследование эффективности операций в технике направлено в конечном итоге на выработку научно обоснованных организационно-технических решений, связанных с созданием и применением технических систем. Решение, независимо от его конкретного содержания и характера, — это результат сознательного выбора одной наилучшей стратегии u^* из множества U допустимых стратегий, сформированного в рамках рассматриваемой проблемной ситуации. Поэтому задача выбора наилучшей стратегии $u^* \in U$ составляет ядро исследования эффективности операций и решается на основе сопоставления (сравнения) всех допустимых стратегий $u \in U$ по их эффективности.

Ответственность за принятое решение несет ЛПР, и ему принадлежит право окончательного выбора стратегии. Поэтому вполне закономерно, что ЛПР выберет такую стратегию $u \in U$, которая является наилучшей в соответствии с его собственными представлениями о достоинствах и недостатках сравниваемых стратегий с точки зрения достижения стоящей перед ним цели. Другими словами, каждое ЛПР имеет свою систему предпочтений, из которой оно исходит при выборе наилучшей, т. е. наиболее предпочтительной по эффективности стратегии в конкретной проблемной ситуации.

Исследование эффективности операции, цель которого помочь ЛПР принять правильное решение, может привести к практически полезным результатам лишь в том случае, если математические методы и модели оценивания (измерения) эффективности операции и выбора наилучшей стра-

тегии (или нескольких лучших стратегий) учитывают и адекватно описывают систему предпочтений ЛПР в данной проблемной ситуации. Пренебрежение этим требованием при формализации задачи исследования эффективности операции неизбежно приводит к тому, что полученные результаты (рекомендации) не согласуются с предпочтениями ЛПР и поэтому им не воспринимаются или даже полностью отвергаются.

Поэтому исследование эффективности операции требует решения ряда сложных проблем, связанных с выявлением и измерением предпочтений ЛПР и построением модели предпочтений $\mathcal{P}$, которая является важнейшим элементом модели проблемной ситуации (4.1).

1. ПРЕДПОЧТЕНИЯ И ИХ СВОЙСТВА

ПОНЯТИЕ ПРЕДПОЧТЕНИЯ

Под *предпочтением* понимают любую форму упорядочения элементов множества D , т. е. устранения неопределенности в выборе некоторого элемента $d^* \in D$ или некоторого подмножества $D^* \in D$. Поскольку термин «предпочтительный» означает «лучший среди других», то выбранный по этому признаку из множества D единственный элемент d^* называют наилучшим, а подмножество D^* элементов — множеством наилучших элементов в рассматриваемом множестве D .

При формировании модели проблемной ситуации и ее отдельных элементов, а также в процессе исследова-

ния эффективности операции ЛПР, экспертам и исследователям приходится решать большое число разнообразных по характеру и содержанию задач, которые сводятся к задаче выбора.

В самом общем виде задача выбора формулируется следующим образом: задано множество D элементов (объектов), из которого необходимо выделить один лучший в определенном смысле элемент d^* или подмножество D^* лучших элементов. Природа множества D может быть различной.

Выделение требуемых элементов из множества D осуществляется индивидом (ЛПР, экспертом, исследователем) на основе некоторого правила выбора (решающего правила) K . Тогда задачу выбора в зависимости от ее конкретной постановки формально можно представить в виде $\langle D, K; d^* \rangle$ или $\langle D, K; D^* \rangle$, где d^* и D^* — решения этих задач.

Правило выбора K в явном или неявном виде отражает концепцию рационального поведения γ индивида (см. гл. 3) и его систему предпочтений $\mathcal{P}$ в конкретной ситуации выбора, т. е. $\{\gamma, \mathcal{P}\} \rightarrow K$.

Система предпочтений $\mathcal{P}$ индивида лежит в основе правила выбора K и представляет собой совокупность его представлений о достоинствах и недостатках сравниваемых элементов $d \in D$ с точки зрения их соответствия той цели, которую он преследует при выборе. Предполагается, что каждый индивид имеет свою собственную систему предпочтений на множестве D , которая отражает объективную характеристику решения и психологию его мышления и которой он руководствуется при рациональных действиях. В общем случае предпочтения индивида непостоянны и могут изменяться даже в одной и той же ситуации выбора.

Система предпочтений индивида на множестве D элементов выбора обладает свойством *полноты*, если он умеет сравнить между собой любые два элемента $d_1, d_2 \in D$ и всегда вынести одно из трех альтернативных суждений: а) d_1 предпочтительнее d_2 ; б) d_1 и d_2 одинаково предпочтительны; в) d_2 предпочтительнее d_1 . Кроме того, пред-

почтения индивида на множестве D должны обладать свойством *направленности*. Это означает, что если, например, индивид при сравнении элементов d_1 и d_2 выносит суждение « d_1 предпочтительнее d_2 », а при сравнении элементов d_2 и d_3 — « d_2 предпочтительнее d_3 », то при сравнении элементов d_1 и d_3 его вывод должен быть однозначен: « d_1 предпочтительнее d_3 ».

Во многих практических задачах выбора сравнение элементов по предпочтительности осуществляется не непосредственно, а с помощью количественных или качественных показателей (характеристик) тех свойств элементов $d \in D$ или последствий их выбора, которые представляются индивиду важными с точки зрения поставленной цели выбора. Значения этих показателей могут быть определены путем объективных или субъективных измерений в тех или иных шкалах (см. параграф 3 гл. 1).

Решение задач выбора $\langle D, K; d^* \rangle$ и $\langle D, K; D^* \rangle$ на основе сопоставления векторных оценок элементов $d \in D$ (а именно этот случай типичен для практики) представляет значительные трудности для индивида, связанные прежде всего с мотивацией его системы предпочтений на множестве векторных оценок и формированием правила выбора K . В условиях неопределенности, когда нарушается однозначное соответствие между элементами $d \in D$ и их векторными оценками, эти трудности возрастают еще более.

Для их преодоления необходимо комплексное использование логического мышления и интуиции человека, математических методов и вычислительных средств, позволяющих выработать обоснованное решение в сложных ситуациях. В первую очередь это относится к главной задаче исследования эффективности операции — задаче $\langle U, K; u^* \rangle$ выбора наилучшей стратегии u^* из множества U допустимых стратегий по критерию эффективности K .

В процессе исследования эффективности операции процедуры выбора формируются ЛПР, экспертами и исследователями, т. е. творчески, неформальным образом, с применением формальных средств — математических методов и ЭВМ, которые значительно

усиливают интеллектуальные возможности человека и повышают обоснованность вырабатываемых решений, но вместе с тем требуют формализованного описания системы предпочтений ЛПР для явного задания критерия K эффективности.

Для построения модели предпочтений $\mathcal{P}$ ЛПР необходимо выявить и измерить его предпочтения на элементах проблемной ситуации (4.1).

Под выявлением предпочтений ЛПР понимают процесс получения информации о предпочтениях этого лица на множестве выбора и построения на ее основе модели, описывающей эти предпочтения. Существует две постановки задачи выявления предпочтений и соответственно два подхода к ее решению [41].

В первой постановке задача состоит в том, чтобы на основании информации о ранее принятых решениях в реальных ситуациях выбора построить модель предпочтений, соответствующую этим решениям. Такая постановка задачи апостериорного выявления предпочтений может иметь место в тех случаях, когда ситуация выбора может повторяться многократно в неизменных (или почти неизменных) условиях и есть возможность накопить достаточный объем информации о принятых решениях.

Во второй постановке задача состоит в том, чтобы построить модель предпочтений ЛПР до принятия им реального решения. Необходимую для этого информацию получают от него посредством специальной процедуры опроса. Модель предпочтений, построенная на основе этой информации, используется для решения задачи выбора. Такая постановка задачи априорного выявления предпочтений наиболее характерна для прикладных задач.

Выявление системы предпочтений в рамках векторных задач принятия решений основано на высказывании ЛПР некоторых суждений о важности частных показателей эффективности, о влиянии определенных изменений отдельных компонент векторных оценок или набора этих компонент на общую предпочтительность элементов множества выбора и др. Совокупность подобных утверждений представляет

собой информацию о системе предпочтений ЛПР. Закономерности, устанавливаемые в процессе анализа получаемой информации, принимаются в качестве допущений о системе предпочтений (например, допущения о разных видах независимости показателей эффективности).

При исследовании возможных способов выявления системы предпочтений ЛПР необходимо: а) четко формулировать допущения, лежащие в основе получения информации о предпочтениях, и проверять их справедливость; б) четко выделять виды запрашиваемой информации о предпочтениях и проверять возможность их получения; в) рассматривать возможные способы получения информации каждого вида и выбирать из них наиболее простые и удобные для ЛПР.

Измерение предпочтений есть отображение альтернативных вариантов решений на числовую ось. Это отображение осуществляется специальной функцией, которая описывает комплексную (интегральную) оценку положительных и отрицательных последствий решения и, следовательно, характеризует его эффективность. Для построения этой функции необходимо соотнести разнородные показатели степени достижения целей в общую оценку эффективности.

Получение функциональной зависимости эффективности вариантов решений (стратегий) от частных показателей эффективности в виде формулы возможно только в простейших случаях. Как правило, эта зависимость носит более сложный характер причинно-следственных связей и не описывается простыми формальными соотношениями.

В связи с этим существенная роль в установлении соотношений между частными показателями и общей оценкой предпочтений вариантов решений принадлежит ЛПР и экспертам.

В общем случае предпочтения экспертов могут не совпадать с предпочтениями ЛПР. Это помогает ему критически осмыслить различные точки зрения, еще раз взвесить возможные последствия решений, четко осознать свои предпочтения. Привлечение экспертов к формированию и выбору стра-

тегий и других объектов проблемной ситуации — это использование коллективного опыта и знаний, позволяющих повысить обоснованность принимаемых решений.

Выявление и измерение предпочтений ЛПР является наиболее сложным и ответственным этапом задачи исследования эффективности операции. От того, насколько построенная модель предпочтений $\mathcal{P}$ ЛПР соответствует его реальным предпочтениям, во многом зависит успех всего исследования и ценность полученных результатов.

Построение модели предпочтений $\mathcal{P}$ ЛПР, учет концепции его рационального поведения и принятие определенных допущений позволяют сформировать для конкретной задачи критерий эффективности K . Именно на этапе построения критерия эффективности проводится конкретизация понятия «предпочтение» и тем самым определяется упорядочение вариантов решений.

Если использование сформированного критерия эффективности K обеспечило выбор наилучшей стратегии u^* из множества U допустимых стратегий или сужение этого множества до требуемого подмножества U^* лучших стратегий, то задача выбора решена. Если этого не произошло, то необходима дополнительная информация о предпочтениях ЛПР.

Возможно, что дополнительную информацию о предпочтениях ЛПР в явной форме получить нельзя или ее получение нерационально вследствие ограниченности времени или больших затрат ресурсов. Тогда ЛПР проводит неформальный анализ выделенных стратегий и выбирает из них наилучшую в соответствии со своими предпочтениями, существующими в неявной форме. Если дополнительную информацию о предпочтениях ЛПР можно получить в явной форме, то осуществляется уточнение модели предпочтений $\mathcal{P}$ ЛПР и соответствующая коррекция критерия эффективности K . Это можно пояснить на результатах примера, рассмотренного в параграфе 1 гл. 4.

Пример 1. Две альтернативы характеризуются следующими значениями векторного показателя эффективности:

$W(u_1) = (25,6,19\ 000)$ и $W(u_2) = (22, 8, 15\ 000)$. Обе альтернативы удовлетворяют требованиям поставленной задачи. Однако ЛПР, располагая дополнительной информацией о проблемной ситуации, которая не была использована при решении задачи выбора, пришел к выводу о том, что стратегия u_1 приводит к оцененному результату уже в конце второго квартала, тогда как вторая стратегия u_2 — лишь к указанному при постановке задачи сроку (четвертый квартал). При этом ЛПР считает, что увеличение покупательной способности на 25 % в конце второго квартала предпочтительнее увеличения ее на 22 %, но в четвертом квартале, несмотря на увеличение себестоимости продукции и большие затраты на реализацию плана. В результате такого неформального анализа ЛПР предпочел выбрать стратегию u_1 .

В процессе исследования эффективности операции уточнение предпочтений ЛПР и корректировка модели его предпочтений могут осуществляться неоднократно.

СПОСОБЫ ВЫРАЖЕНИЯ ПРЕДПОЧТЕНИЙ

Введенное понятие системы предпочтений основывается на том, что ЛПР может выразить свое отношение к тому или иному элементу путем попарного их сравнения с указанием более предпочтительного объекта в каждой паре или указания о том, что объекты равноценны. Иногда допускается делать вывод по результатам контрольных предъявлений, что элементы несравнимы между собой. Это означает, что ЛПР не считает, что первый предпочтительнее второго, или что второй предпочтительнее первого, или что они равноценны. Подобная ситуация возникает тогда, когда элементы похожи друг на друга и при одном предъявлении ЛПР указывает на предпочтительность одного, при некотором другом предъявлении — второго, а в некоторых случаях указывает, что они равноценны. Этот способ выражения предпочтения самый простой, так как элементы сравниваются лишь в парах без учета отношения ЛПР к другим

элементам. Этот способ выражения предпочтений называется попарным сравнением и относится к широкому классу способов, определенному как элементарные суждения.

Результаты попарного сравнения часто представляют в виде матрицы попарных сравнений. В этом случае выбирается некоторая шкала (в общем случае множество символов), например, со значениями: 1 — отражает факт предпочтительности; 0,5 — факт равноценности; 0 — факт непредпочтительности; * — факт несравнимости элементов; на пересечении соответствующих строк и столбцов матрицы, составленной из элементов, ставится тот или иной символ (рис. 1, б). Кроме попарного сравнения для выражения предпочтений могут быть использованы другие формы элементарных суждений. Рассмотрим их в порядке возрастания сложности получения информации от ЛПР.

Сортировка. ЛПР предъявляется исходное множество элементов, которое оно должно разделить на некоторое количество классов.

Пример 2. Рассматривая проблему повышения рентабельности предприятия (см. пример 1 гл. 4), ЛПР может разделить общую цель на следующие классы подцелей, различающиеся очередностью их достижения. Оно может выделить следующие классы: «обязательные» цели (без их достижения невозможно решение проблемы и они должны решаться в первую очередь; к ним относятся указанные в примере задача повышения покупательной способности, задача снижения себестоимости, сокращение затрат на решение проблемы); «желательные» цели (например, дополнительное создание резервных фондов материального стимулирования и перспективного развития производственной базы); «возможные» цели (например, повышение качества продукции до уровня мировых стандартов, завоевание престижа предприятием) и др.

В этом случае ЛПР осуществляет не просто попарное сравнение элементов безотносительно ко всем остальным, а выделяет классы, как бы охватывая всю совокупность элементов. Поэтому сортировка требует от ЛПР большей

подготовленности, чем просто попарное сравнение.

Ранжирование. Это способ выражения предпочтения путем представления элементов в виде последовательности в соответствии с возрастанием или убыванием их предпочтительности. Различают строгое и нестрогое ранжирование. В первом случае не допускается указывать на равноценность некоторых рядом поставленных элементов. Ранжирование часто представляют как приписывание элементам оценок в ранговой шкале. При этом допускается как прямое ранжирование (когда более предпочтительному элементу приписывается меньшее значение оценки), так и обратное ранжирование (более предпочтительным является элемент с более высокой оценкой). В последнем случае часто рангом элемента считают номер его места, которое он занимает в ранжированном ряду при обратной нумерации мест (при этом полагается, что равноценные объекты находятся на одном и том же месте). Например, места в социалистическом соревновании имеют прямую ранжировку, а классификатор профессий (разрядная сетка) — обратную ранжировку.

Элементарные суждения в виде результатов попарного сравнения, сортировки и ранжирования выражаются всегда в качественных шкалах. Это определяет вид допустимых над ними математических преобразований, приводящих к осмысленным выводам (см. параграф 3 гл. 1).

Промежуточную шкалу между качественной и количественной имеют элементарные суждения, выраженные в балльных оценках. Каждому элементу из множества предъявленных ЛПР ставит в соответствие число (балл), характеризующее его субъективное мнение о предпочтительности объектов. Указанные числа (балльные оценки) выбираются из специальной балльной шкалы. Обычно значения (градации) балльной шкалы представляют собой ряд чисел, отстоящих друг от друга на одинаковом расстоянии. Чаще всего берут первые натуральные числа (1, 2, 3 ...), но иногда и действительные в том случае, когда шкала имеет большое число градаций

при сравнительно узком диапазоне изменения баллов (например, при диапазоне шкалы от 1 до 5 допускаются оценки типа 4,85). Оценивание в балльной шкале проводится тогда, когда предпочтительность элемента устанавливается по правилам, не допускающим двойного толкования.

Попарное выражение предпочтения как доли суммарной интенсивности. Этот способ сходен со способом попарного сравнения и отличается от него тем, что ЛПР должен указать, какой из двух элементов является более предпочтительным, а также как суммарная интенсивность предпочтения, приходящаяся на оба рассматриваемых элемента, распределяется между ними. Обычно мера суммарной интенсивности предпочтения, приходящаяся на сравниваемые элементы, принимается равной 1. Например, ЛПР может указать, что покупательная способность продукции составляет 0,75, а уровень себестоимости — 0,25 при сравнении их по вкладу в уровень рентабельности предприятия, т. е. целевой эффект (см. примеры 1 и 2 главы 4).

Попарное выражение предпочтения как доли относительной интенсивности. В отличие от предыдущего случая ЛПР должен указать — во сколько раз один элемент предпочтительнее другого. Так, в условиях рассматриваемого примера достижение цели по повышению покупательной способности в 3 раза предпочтительнее для ЛПР, чем достижение цели по снижению себестоимости ($0,75 : 0,25 = 3$).

Попарное выражение предпочтения как доли суммарной или относительной интенсивности требует использования шкал, близких к количественным.

При выражении предпочтений коэффициентами важности ЛПР просят указать степень влияния изменения значения частного показателя эффективности на результат операции. Степень этого влияния должна измеряться коэффициентами важности, которые удовлетворяют требованиям неотрицательности и нормировки (сумма их равна единице). Коэффициент важности тем больше, чем более предпочтительным для ЛПР является изменение показателя в соответствующем на-

правлении по вкладу в целевой эффект. Для одинаковых по предпочтительности частных показателей эффективности значения коэффициентов важности одинаковы. Получение такой информации от ЛПР представляет наибольшую трудность. В дальнейшем полученные коэффициенты важности используются при агрегировании частных показателей в функцию эффективности.

Рассмотренные способы выражения предпочтений ЛПР относятся к случаю сравнения исходов операции G или значений показателей W . При выборе лучшей стратегии проведения операции (т. е. установлении предпочтений на множестве U) необходимо дополнительно учитывать условия ее проведения, т. е. ориентироваться на более «правдоподобную» ситуацию. В этом случае необходимо выявить предпочтения ЛПР относительно условий Λ . Для этого применяют два основных способа.

При выражении предпочтения субъективными вероятностями от ЛПР необходимо получить оценки степени возможности реализации того или иного комплекса условий проведения операции. Эти оценки выражаются в виде неотрицательных чисел, в сумме равных 1, причем каждое из этих чисел отражает степень уверенности ЛПР в том, что при проведении операции может реализоваться ситуация, соответствующая определенному параметру $\lambda \in \Lambda$. Существуют условия [14] для получения логически непротиворечивых выводов с использованием субъективных вероятностей. При выполнении этих условий с субъективными вероятностями можно обращаться так, как если бы это были обычные «объективные» вероятности.

Способ выражения предпочтений лингвистическими переменными основан на той предпосылке, что элементами мышления человека являются не числа, а элементы некоторых нечетких множеств, или классов объектов, для которых переход от «принадлежности» к «непринадлежности» не скачкообразен, а непрерывен [24]. Такой подход используют при принятии решений, когда определяющими являются не количественные,

а качественные характеристики. Например, при формулировке «возможных» целей предприятия в примере 2 ставилась задача «завоевания престижа». В этом случае уровень престижа может измеряться лингвистической переменной со значениями «низкий», «средний», «высокий», «очень высокий». Далее вводят некоторую функцию принадлежности как способ формализации субъективного смысла этих качественных показателей. Основные понятия, определения и операции над нечеткими множествами даны в приложении к гл. 11.

ОТНОШЕНИЯ

Кроме рассмотренных выше элементарных суждений для математического описания предпочтений в моделях принятия решений используется универсальное их представление в виде отношений.

Отношение — это математическое понятие для обозначения подмножества прямого декартова произведения множеств.

Наиболее употребительными в практике принятия решений являются бинарные отношения, так как они хорошо связываются с традиционными способами выражения элементарных суждений.

Бинарным отношением R на множестве элементов D называется подмножество упорядоченных пар $\{d', d''\}$ множества $D \times D$ всех таких пар.

Символом $D \times D$ обозначают прямое декартово произведение. Элементами множества D могут быть, например, исходы операции (в этом случае $D = G$). Если декартово произведение состоит более чем из двух «сомножителей» ($D \times D \times D$, $D \times D \times D \times D$, ...), то его элементами являются упорядоченные тройки, четверки элементов и т. д. В этом случае принципиально можно рассматривать тернарные, тетрарные и другие отношения.

Бинарные отношения могут быть использованы для универсального описания связей между элементами различной природы: для описания связности электрических и информационных сетей, иерархических структур управления и т. п.

Бинарные отношения есть множества специального вида, поэтому их описание основывается на обычных способах задания множеств: перечислением элементов множества R , указанием общих свойств этих элементов, графом, матрицей смежности, подмножеством точек в декартовой системе координат.

Пример 3. Пусть R выражает мнение ЛПР о том, что частный показатель W_i не менее предпочтителен, чем показатель W_j , для векторного показателя $W = (W_1, W_2, W_3, W_4)$, где W_1 — покупательная способность; W_2 — себестоимость; W_3 — затраты; W_4 — время на реализацию стратегии (см. пример 1). ЛПР, например, считает, что увеличение покупательной способности не менее предпочтительнее снижения себестоимости и затрат, а снижение себестоимости, в свою очередь, не менее предпочтительнее снижения затрат (при условии, что затраты не превышают допустимой нормы); фактор времени при этом не менее важнее снижения себестоимости и затрат. В этом случае отношение R можно записать:

$$R = \{(W_i, W_j) \mid (W_1, W_1), (W_1, W_2), (W_1, W_3), (W_2, W_2), (W_2, W_3), (W_3, W_3), (W_4, W_2), (W_4, W_3), (W_4, W_4)\}$$

— как прямое перечисление показателей, связанных введенным отношением;

$R = \{(W_i, W_j) \mid W_i, W_j \in W; W_i \text{ не менее предпочтительнее (важнее), чем } W_j\}$ — как указание общих свойств элементов.

Это же отношение может быть представлено в виде графа, матрицы смежности и множеством точек в декартовой системе координат (соответственно случай $a \rightarrow b$ на рис. 1); направление стрелок на рис. 1, a соответствует направлениям рассматриваемого предпочтения между элементами, которые они связывают. Петли на графе обозначают тот факт, что элемент не менее предпочтителен самого себя.

С использованием указанных способов графического представления отношений весьма удобно анализировать

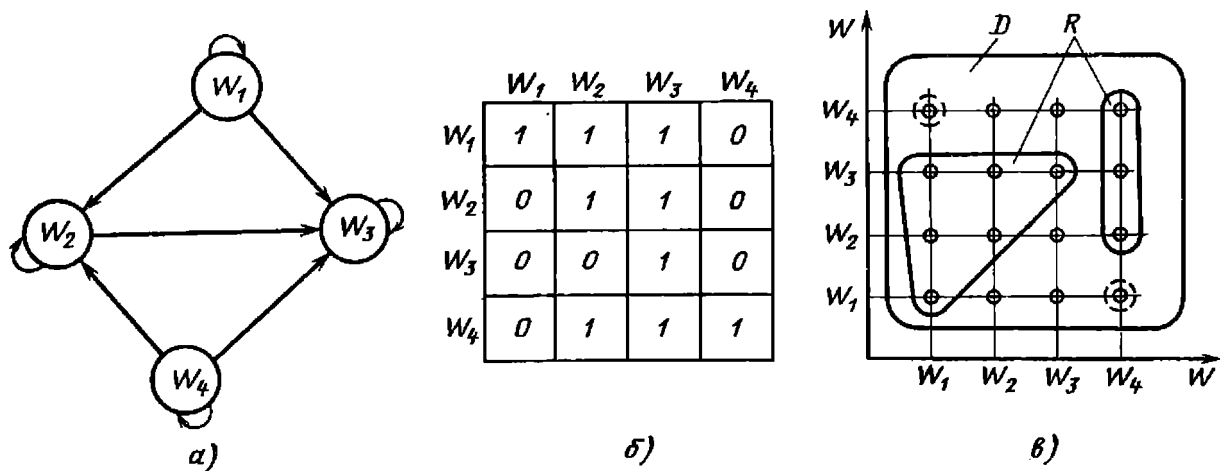


Рис. 1. Графическое представление бинарного отношения «не менее предпочтителен (важен)»

их свойства. Для формального описания свойств бинарных отношений обозначим: $(d', d'') \in R$ или $d' R d''$ — если элементы d', d'' связаны отношением R и $(d', d'') \notin R$ или $d' \not R d''$ — если элементы не связаны отношением R .

Свойства бинарных отношений. Если для любого элемента $d \in D$ выполняется условие $(d, d) \in R$, то отношение R *рефлексивно* (рис. 2, а).

Если для любой тройки элементов $d, d', d'' \in D$ удовлетворяется условие, что из $(d, d') \in R$ и $(d', d'') \in R$ следует, что $(d, d'') \in R$, то такое отношение называется *транзитивным* (рис. 2, е).

Отношение R называется *симметричным*, если из $(d, d') \in R$ всегда следует $(d', d) \in R$ (рис. 2, б).

Отношение R называется *связным* (линейным, полным), если для любых

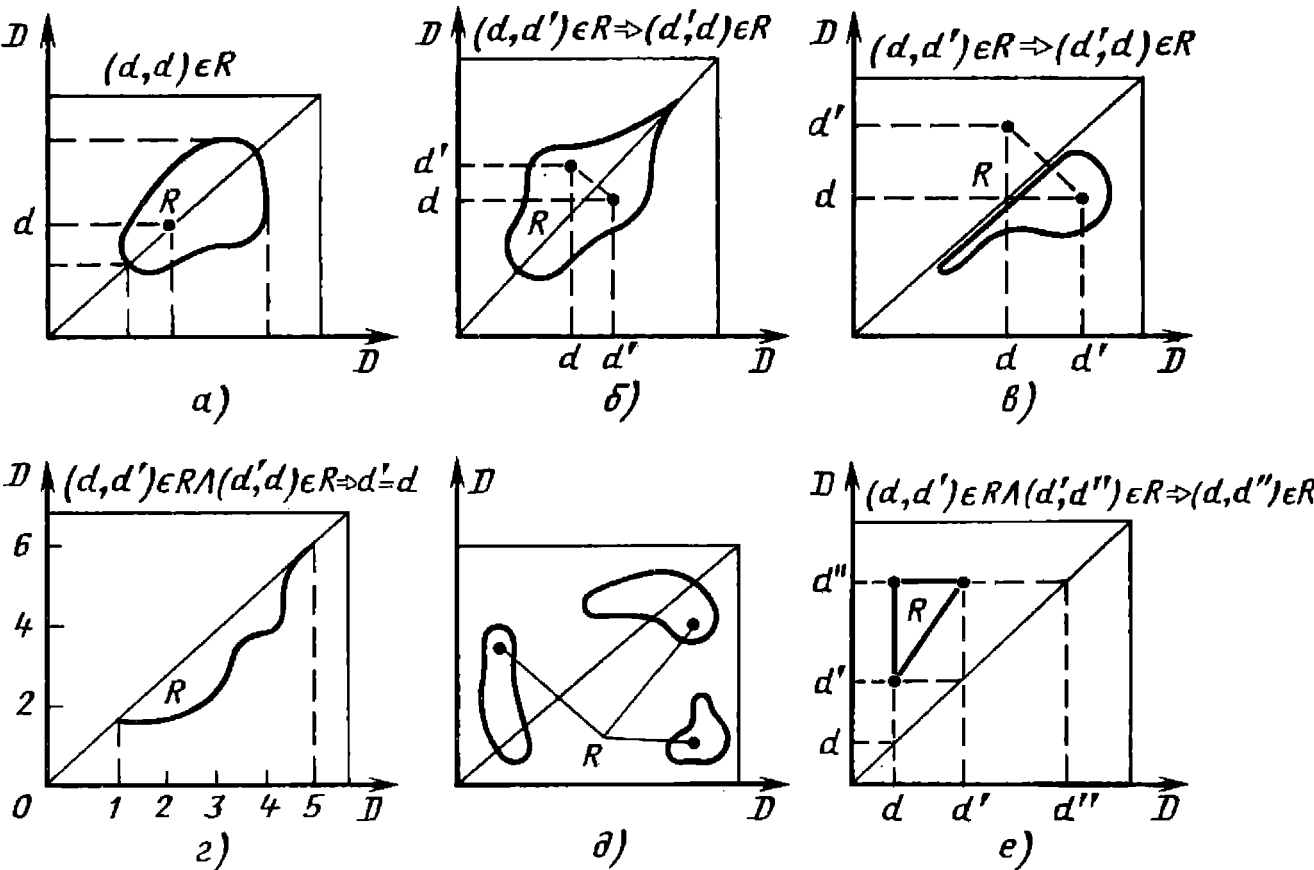


Рис. 2. Геометрическая интерпретация бинарного отношения

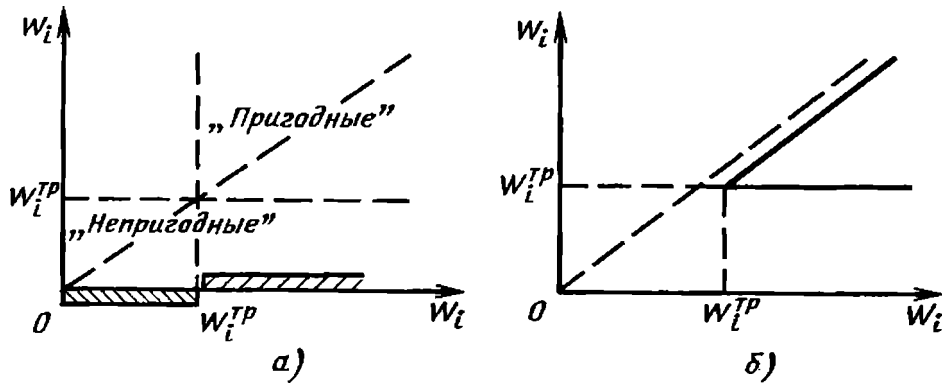


Рис. 3. Отношения эквивалентности и строгого частичного порядка

двух несовпадающих элементов $(d, d') \in D$ справедливо хотя бы одно из двух утверждений: либо $(d, d') \in R$, либо $(d', d) \in R$. Каждое из свойств бинарных отношений может иметь «антипода». Например, отношение может быть несвязным (рис. 2, д).

Если отношение R справедливо только для несовпадающих элементов из D , то оно называется *антирефлексивным*, т. е. из (d', d'') следует, что d' не есть d'' .

Если отношение R не является симметричным, то в зависимости от природы элементов $d \in D$ (числовые или нечисловые характеристики) вводятся свойства антисимметричности (для числовых) и асимметричности (для нечисловых).

Отношение R называется *антисимметричным*, если из $(d, d') \in R$ и $(d', d) \in R$ следует равенство $d = d'$ (рис. 2, г).

Отношение R *асимметрично*, если из $(d, d') \in R$ следует, что $(d', d) \notin R$ (рис. 2, в).

Дадим характеристику свойств бинарного отношения на примере рис. 1. Введенное отношение рефлексивно. Признаком этого является наличие петель на графе, единиц — на диагонали матрицы смежности. Отношение R несвязно (отсутствует стрелка между W_1 и W_4 на графе), транзитивно.

Кроме того, отношение R асимметрично (граф однонаправленный), область R в D на рис. 1, в геометрически несимметрична относительно диагонали.

Если бы отношение R рассматривалось только на элементах W_1, W_2, W_3 , то оно обладало бы свойствами

рефлексивности, транзитивности и связности.

В теории принятия решений особое место занимают отношения, обладающие специальным набором указанных свойств. Это отношения эквивалентности, строго частичного порядка, квазипорядка и порядка.

Эквивалентностью называется симметричное, рефлексивное, транзитивное отношение. Это отношение имеет большое значение при формализации процессов и явлений (рис. 3, а). В математике оно связано с понятием разбиения множеств на классы. Если отношение R есть эквивалентность на множестве D , то элементы d и d' относятся к одному классу D_j разбиений тогда и только тогда, когда $(d, d') \in R$. И наоборот, если дано разбиение D на классы $\{D_j\}$, то пара $(d, d') \in D_j$ эквивалентна.

Строгим частичным порядком называется антирефлексивное транзитивное отношение (рис. 3, б).

Квазипорядком называется рефлексивное и транзитивное отношение.

Порядком называется антисимметричное рефлексивное транзитивное отношение.

В примере 3 множество всех значений векторного показателя эффективности $W = (W_1, W_2, W_3, W_4)$ разбивается на два класса: класс пригодных значений, для которого $W_1 \geq 20\%$; $W_2 \geq 5\%$; $W_3 \leq 20\,000$ р. и $W_4 \leq t_0 + 12$ месяцев (t_0 — момент времени анализа альтернатив); класс непригодных значений, для которого не выполняется хотя бы одно из указанных условий.

В этом случае все альтернативы каждого класса эквивалентны, т. е. в первом классе находятся все одинаково пригодные, во втором — все одинаково непригодные. Отношение эквивалентности R задается указанием общего свойства: «быть пригодным». Использование дополнительной информации о предпочтении, например, « W_i не менее предпочтительнее (важнее) W_j » позволяет на выделенном классе пригодных альтернатив построить квазипорядок, который в данном случае оказывается несвязным (см. рис. 1). Если от ЛПР была бы получена информация типа: « W_i строго предпочтительнее W_j » и при этом она касалась бы любой пары альтернатив, то полученное на основе этой информации отношение было бы строгим частичным порядком, который обладает свойством антисимметричности (значения показателей — действительные числа). Для каждого из показателей W_i отношения эквивалентности и строгого частичного порядка показаны на рис. 3.

Каждое из рассмотренных элементарных суждений (как способ выражения предпочтений) может быть охарактеризовано с помощью свойств бинарных отношений.

Попарное сравнение в общем случае обладает только свойством рефлексивности. Поскольку сравнение элементов проводится только в парах без учета остальных элементов, свойство транзитивности выявленного отношения предпочтения, как правило, отсутствует, а так как допускается указывать на несравнимость элементов, то отсутствует и свойство связности. Например, при попарном сравнении ЛПР может указать, что увеличение покупательной способности W_1 предпочтительнее снижения себестоимости W_2 , а снижение себестоимости предпочтительнее снижения затрат W_3 (при попарном сравнении последних). Из этого еще нельзя заключить, что W_1 предпочтительнее W_3 , так как при их попарном сравнении ЛПР может даже указать, что снижение затрат предпочтительнее повышения покупательной способности. Такой случай называется *нетранзитивностью в суждениях ЛПР*. Это обстоятельство яв-

ляется одним из главных недостатков одного из самых простых способов выражения элементарных суждений.

Сортировка может задавать либо отношение эквивалентности, либо толерантности (рефлексивное, симметричное отношение) на предъявленном ЛПР множестве элементов. Так как среди предъявленных элементов ЛПР может уверенно отнести к тому или иному классу лишь элементы, субъективно «сильно» различающиеся между собой, а среди оставшихся есть «похожие», то транзитивность на границах между классами может нарушиться. В результате этого отношение становится только рефлексивным и симметричным, что является определенным недостатком сортировки. Такое отношение называется *толерантностью*.

Ранжирование задает отношение квазипорядка. Если ранжирование строгое, то выявленное отношение является строгим частичным порядком.

Способы балльного оценивания субъективных вероятностей и выражения предпочтения коэффициентами важности устанавливают отношение порядка на предъявленном множестве элементов.

Наиболее серьезными недостатками части элементарных суждений является отсутствие связности и транзитивности, что не позволяет осуществить однозначный выбор. В этом случае необходимо либо привлечь дополнительную информацию, раскрывающую неопределенность и неоднозначность суждений ЛПР, либо использовать ряд непротиворечивых гипотез для устранения указанных недостатков. В качестве определяющей гипотезы при выработке решения выдвигаются предположения о транзитивности и связности суждений ЛПР. Исходя из этой гипотезы, некоторое нетранзитивное отношение R можно аппроксимировать «ближайшим» к нему наименьшим транзитивным отношением $\hat{R}$, включающим в себя R . Такая операция называется *транзитивным замыканием отношения R* , которое строится следующим образом:

$$\hat{R} = R \cup R^2 \cup R^3 \cup \dots \cup \dots,$$

где $R^2 = R \otimes R$ — композиция отношения R ;

$$R^3 = R^2 \otimes R, \quad R^4 = R^3 \otimes R, \dots,$$

а композиция R^2 определяется по правилу перемножения матриц смежности отношений R с заменой арифметических операций операциями булевой алгебры (см. т. 2).

Пример 4. Построим транзитивное замыкание для примера 1, когда при попарном сравнении установлено, что W_1 не менее предпочтительнее W_2 , а W_2 не менее предпочтительнее W_3 . Указанное отношение несвязно и не-транзитивно. Матрица смежности этого отношения выглядит следующим образом:

$$R = \begin{matrix} & \begin{matrix} W_1 & W_2 & W_3 \end{matrix} \\ \begin{matrix} W_1 \\ W_2 \\ W_3 \end{matrix} & \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} \end{matrix}.$$

Композиции отношения имеют вид:

$$R^2 = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix}, \quad R^3 = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix}.$$

Так как R^3 совпадает с R^2 , то композиции более высокого порядка искать не нужно. В этом случае транзитивное замыкание $\check{R}$ совпадает с R^2 , т. е. появляется связь между элементами W_1 и W_3 .

Потеря транзитивности, как правило, возникает в том случае, когда ЛПР не может четко выразить суждение об отношении на множестве элементов, что, например, при сортировке приводит к толерантности. В этом случае иногда применяют аппарат задания нечетких отношений предпочтений с использованием лингвистической переменной.

НЕЧЕТКИЕ ОТНОШЕНИЯ

Переход от обычного (четкого) отношения к нечеткому в принципе тот же, что и переход от обычного множества к нечеткому. В практике чаще исполь-

зуют задание нечетких отношений на обычных множествах.

В этом случае нечетким отношением $\tilde{R}$ на обычном множестве D называется [32, 50] нечеткое подмножество прямого декартова произведения $D \times D$, характеризующееся функцией принадлежности $\mu_R: D \times D \rightarrow [0, 1]$. Значение $\mu_R(d, d')$ этой функции принимается как субъективная мера отношения $(d, d') \in R$. Обычное отношение есть частный случай нечеткого отношения с релейной функцией принадлежности:

$$\mu_R = \begin{cases} 1, & \text{если } (d, d') \in R; \\ 0, & \text{если } (d, d') \notin R. \end{cases}$$

Пример 5. Отношение $\tilde{R}$ для случая, изображенного на рис. 1, а, полученное в результате опроса экспертов, представляется в виде функции принадлежности:

$$\mu_R = \begin{vmatrix} 1 & 0,9 & 0,6 & 0,6 \\ 0,1 & 1 & 0,7 & 0,2 \\ 0,4 & 0,3 & 1 & 0,3 \\ 0,4 & 0,8 & 0,7 & 1 \end{vmatrix}.$$

Если ЛПР считает, что факт превосходства в важности имеет место при значениях функции принадлежности $\mu_R \geq 0,5$, то модель предпочтения можно представить «четкой» аппроксимацией, изображенной на рис. 1, б.

Приведенная матрица соответствует случаю, когда при попарном сравнении элементов накладывается дополнительное условие нормировки, из которого следует, что $\mu_R(d, d') = 1 - \mu_R(d', d)$. В общем случае элементы матрицы смежности μ_R могут быть свободны от этого ограничения.

По аналогии с обычным отношением нечеткое отношение можно представить ориентированным графом, каждой дуге которого приписано число μ_R , или же поверхностью в декартовой системе координат.

При работе с нечеткими отношениями, построенными на обычном множестве D , используют следующие основные операции [24, 32, 50].

Объединение нечетких отношений.

Заданы отношения $\tilde{R}_A$ и $\tilde{R}_B$ на множестве D . Их объединением называется отношение $\tilde{R}_C = \tilde{R}_A \cup \tilde{R}_B$, имеющее функцию принадлежности:

$$\mu_{\tilde{R}_C}(d, d') = \max \{ \mu_{R_A}(d, d'), \mu_{R_B}(d, d') \}.$$

Их пересечением называется отношение $\tilde{R}_F = \tilde{R}_A \cap \tilde{R}_B$, имеющее функцию принадлежности:

$$\mu_{R_F}(d, d') = \min \{ \mu_{R_A}(d, d'), \mu_{R_B}(d, d') \}.$$

Дополнением нечеткого отношения $\tilde{R}$, заданного на множестве D , является отношение $\tilde{R}'$, имеющее функцию принадлежности:

$$\mu_{R'}(d, d') = 1 - \mu_R(d, d').$$

Обратное к $\tilde{R}$ нечеткое отношение $\tilde{R}^{-1}$ определяется на множестве D следующим образом:

$$(d, d') \in \tilde{R}^{-1} \Leftrightarrow (d', d) \in R \text{ для любых } d, d' \in D, \text{ причем } \mu_{R^{-1}}(d, d') = \mu_R(d', d).$$

Как и для обычных отношений, важное место занимает композиция нечетких отношений $\tilde{R} \circ \tilde{R} = \tilde{R}_A \otimes \tilde{R}_B$, которая задается функцией принадлежности вида

$$\mu_{\circ}(d', d'') = \max_{d \in D} \min \{ \mu_{R_A}(d', d), \mu_{R_B}(d, d'') \},$$

т. е. она получается с помощью тех же операций, что и композиция обычных отношений.

Существуют и другие определения композиции нечетких отношений: минимаксное с функцией принадлежности

$$\mu_{\circ}(d', d'') = \min_{d \in D} \max \{ \mu_{R_A}(d', d), \mu_{R_B}(d, d'') \}$$

и максимумумпликативное с функцией принадлежности

$$\mu_{\circ}(d', d'') = \max_{d \in D} \{ \mu_{R_A}(d', d) \times \mu_{R_B}(d, d'') \}.$$

Однако в практике принятия решений чаще используется максиминное определение композиции.

Нечеткие отношения могут обладать следующими свойствами.

Рефлексивность. Нечеткое отношение $\tilde{R}$, заданное на множестве D , обладает свойством рефлексивности, если для любого $d \in D$ выполняется условие

$$\mu_R(d, d) = 1.$$

Например, отношение «примерно одинаковы по предпочтительности» рефлексивно.

Симметричность. Нечеткое бинарное отношение $\tilde{R}$, заданное на множестве D , называется симметричным, если для любых элементов $d, d' \in D$ выполняется условие

$$\mu_R(d, d') = \mu_R(d', d).$$

Например, нечеткое отношение «показатели близки по предпочтительности» симметрично.

Транзитивность. Нечеткое отношение $\tilde{R}$, заданное на множестве D , называется транзитивным, если для любых трех пар из произвольных элементов $d, d', d'' \in D$ выполняется условие

$$\mu_R(d, d'') \geq \max_{d'} \min \{ \mu_R(d, d'), \mu_R(d', d'') \}.$$

Например, транзитивным нечетким отношением является «один показатель много важнее другого».

Нечеткое отношение $\tilde{R}$ транзитивно, если $\tilde{R} \otimes \tilde{R} \subseteq \tilde{R}$, т. е. свойство транзитивности зависит от способа определения произведения нечетких отношений: максиминное, минимаксное, максимумумпликативное. В дальнейшем везде используется максиминная форма представления произведе-

ния, так как она более близко связана с обычной операцией произведения четких отношений.

Отношение называется *антирефлексивным*, если для любого элемента $d \in D$ выполняется условие

$$\mu_R(d, d) = 0.$$

Примером антирефлексивного отношения может служить высказывание «один показатель важнее другого». Дополнение рефлексивного отношения является антирефлексивным.

Нечеткое отношение $\tilde{R}$ является асимметричным, если для любых элементов $d, d' \in D$ из условия $\mu_R(d, d') > 0$ следует, что $\mu_R(d', d) = 0$.

Примером такого отношения является «один показатель важнее другого».

По аналогии с транзитивным замыканием обычного бинарного отношения вводится транзитивное замыкание нечеткого бинарного отношения.

Транзитивным замыканием нечеткого отношения $\tilde{R}$ называется нечеткое

отношение $\tilde{R}^{\circ} = \tilde{R} \cup \tilde{R}^2 \cup \tilde{R}^3 \cup \dots$, где $\tilde{R}^2 = \tilde{R} \otimes \tilde{R}$, а $\mu_{R^2}(d, d'') = \max_{d'} \min \{ \mu_R(d, d'), \mu_R(d', d'') \}$ для любых $d, d', d'' \in D$. Аналогичным образом определяется $\mu_{R^k}(d, d'')$.

Пример 6. Рассмотрим операцию транзитивного замыкания для введенного отношения $\tilde{R}$ на графе, изображенном на рис. 1, а:

$$\mu_R = \begin{vmatrix} 1 & 0,9 & 0,6 & 0,6 \\ 0,1 & 1 & 0,7 & 0,2 \\ 0,4 & 0,3 & 1 & 0,3 \\ 0,4 & 0,8 & 0,7 & 1 \end{vmatrix}.$$

Композиции отношения $\tilde{R}$ имеют вид:

$$\mu_{R^2} = \begin{vmatrix} 1 & 0,9 & 0,7 & 0,6 \\ 0,4 & 1 & 0,7 & 0,3 \\ 0,4 & 0,4 & 1 & 0,4 \\ 0,4 & 0,8 & 0,7 & 1 \end{vmatrix};$$

$$\mu_{R^3} = \begin{vmatrix} 1 & 0,9 & 0,7 & 0,6 \\ 0,4 & 1 & 0,7 & 0,4 \\ 0,4 & 0,4 & 1 & 0,4 \\ 0,4 & 0,8 & 0,7 & 1 \end{vmatrix};$$

$$\mu_{R^4} = \begin{vmatrix} 1 & 0,9 & 0,7 & 0,6 \\ 0,4 & 1 & 0,7 & 0,4 \\ 0,4 & 0,4 & 1 & 0,4 \\ 0,4 & 0,8 & 0,7 & 1 \end{vmatrix}.$$

Так как μ_{R^3} совпадает с μ_{R^4} , то функции принадлежности композиций более высокого порядка искать не нужно, а функция принадлежности транзитивного замыкания определяется объединением $\mu_R, \dots, \mu_{R^4}$:

$$\mu_{\tilde{R}^{\circ}} = \mu_R \cup \mu_{R^2} \cup \mu_{R^3} = \begin{vmatrix} 1 & 0,9 & 0,7 & 0,6 \\ 0,4 & 1 & 0,7 & 0,4 \\ 0,4 & 0,4 & 1 & 0,4 \\ 0,4 & 0,8 & 0,7 & 1 \end{vmatrix}.$$

Аналогом системы предпочтений для задач принятия решений с нечетко выраженным отношением предпочтения служит одно из трех возможных нечетких высказываний ЛПР при попарном сравнении двух произвольных элементов из предъявленного множества: а) первый элемент «не хуже» второго; б) оба предъявленных элемента «примерно одинаковы»; в) второй элемент «не хуже» первого; причем каждое из этих суждений может быть описано значениями функции принадлежности.

2. ОЦЕНИВАНИЕ ПРЕДПОЧТЕНИЙ

ОТНОШЕНИЯ ПРЕДПОЧТЕНИЯ И ИХ СВОЙСТВА

Бинарные отношения, являясь универсальным способом описания связей между элементами произвольной природы, широко используются в практике принятия решений. С их помощью формально задаются и описываются

свойства всех отношений предпочтения. Основными отношениями предпочтения являются следующие:

отношение строгого предпочтения $\succ$;
отношение безразличия $\sim$.

В этом случае запись $d' \succ d''$ означает, что элемент d' строго предпочтительнее элемента d'' , т. е. при предъявлении ЛПР только двух указанных элементов оно всегда будет явно предпочитать элемент d' . Запись $d' \sim d''$ означает, что элементы одинаковы по предпочтительности и если предъявление ограничить только этими двумя элементами, то ЛПР всегда безразлично, какой из них выбрать.

На основе отношений строгого предпочтения и безразличия вводят дополнительно: отношения нестрогого предпочтения $\succcurlyeq$, несравнимости $\nlessgtr$ и неразличимости $\#$, а также различные градации указанных отношений.

Отношение нестрогого предпочтения $d' \succcurlyeq d''$ означает, что элемент d' , по мнению ЛПР, не менее предпочтителен, чем d'' , т. е. при их предъявлении ЛПР указывает либо, что $d' \succ d''$, либо, что $d' \sim d''$. Формально отношение $\succcurlyeq$ есть объединение $\succ \cup \sim$.

Отношение несравнимости $d' \nlessgtr d''$ означает, что ЛПР не понятно, как выразить отношения между элементами d' , d'' , т. е. оно не может однозначно утверждать, что $d' \succ d''$ или $d'' \succ d'$, или $d' \sim d''$.

Отношение неразличимости $d' \# d''$ означает, что либо ЛПР не может сравнить элементы d' , d'' ($d' \nlessgtr d''$), либо считает их эквивалентными ($d' \sim d''$).

Формально отношения несравнимости и неразличимости можно представить так:

$$d' \nlessgtr d'' \Leftrightarrow d' \not\succcurlyeq d'' \wedge d'' \not\succcurlyeq d';$$

$$d' \# d'' \Leftrightarrow d' \nlessgtr d'' \vee d'' \sim d'.$$

По смыслу введенные отношения предпочтения обладают следующими свойствами бинарных отношений:

$\succ$ — антирефлексивно и асимметрично;

$\sim$ — рефлексивно и симметрично;

$\succcurlyeq$ — рефлексивно.

В зависимости от наличия дополнительных свойств у указанных отношений вводятся различные их градации,

представленные в табл. 1. Отношения 2—8 в табл. 1 охарактеризованы как транзитивные, в общем случае они таковыми могут и не быть (например, если эти отношения выявлены в результате попарного сравнения). Транзитивность теряется в том случае, если ЛПР в процессе контрольных предъявлений оценивает объекты по разным целевым признакам. Отношение нестрогого предпочтения $\succcurlyeq$ является следствием нестрогого ранжирования элементов. В результате этого все множество предъявления элементов разбивается на различающиеся по предпочтительности классы, внутри которых элементы одинаковы по предпочтительности. Если каждый класс содержит только один элемент и проведено нестрогое ранжирование, то такое отношение есть связный нестрогий порядок. Если при этом ранжирование строгое — связный строгий порядок (серия). Если в каждом классе более одного элемента, проведено строгое ранжирование между классами, но внутри класса элементы неразличимы (либо несравнимы, либо эквивалентны), то полученное отношение есть *квазисерия*. Любое частичное отношение отличается от связного тем, что классы элементов из множества предъявления нельзя полностью упорядочить по предпочтительности (это можно сделать лишь частично).

В практике выявления и оценивания предпочтений обычно стремятся добиться непротиворечивости суждений ЛПР, поэтому везде в дальнейшем всегда будем считать, что отношения строгого предпочтения $\succ$, безразличия $\sim$ и нестрогого предпочтения транзитивны, так что $\succ$ — строгий частичный порядок, $\sim$ — эквивалентность, а $\succcurlyeq$ — квазипорядок.

НЕЧЕТКИЕ ОТНОШЕНИЯ ПРЕДПОЧТЕНИЯ И ИХ СВОЙСТВА

С использованием нечетких бинарных отношений в отношении нестрогого предпочтения $R_{\succcurlyeq}$ на множестве D задаются и описываются свойства нечетких отношений строгого предпочтения $R_{\succ}$ и отношения безразличия $R_{\sim}$. При этом обычно считают, что нечет-

1. Модельные отношения предпочтений и их свойства

№ по пор.	Отношения предпочтения Бинарные отношения	Свойства бинарных отношений						
		Транзитив- ность	Рефлексив- ность	Антиреф- лексивность	Симметрич- ность	Асимметрич- ность	Нерефлексив- ность	Связность
1	$\approx, \#$ Толерантность		+		+			
2	$\sim$ Эквивалентность	+	+		+			
3	$\succcurlyeq$ Несвязный квазипорядок	+	+					
4	$\succcurlyeq$ Связный квазипорядок	+	+					+
5	$\succcurlyeq$ Связный нестрогий порядок	+		+		+ *1		+
6	$\succ$ Частичный строгий порядок	+		+		+ *1		
7	$\succ$ Связный строгий порядок	+		+		+ *1		+
8	$\succ$ Квазисерия *2	+				+	+	

*1 Данное свойство вытекает из предыдущих свойств.

*2 $d' \succ d'' \wedge d'' \succ d' \Rightarrow d' \# d''$.

кое отношение нестрогого предпочтения $R_{\succsim}$ обладает только свойством рефлексивности, т. е. $\mu_{R_{\succsim}}(d', d'') = 1$ для любых $d' \in D$ [32, 50], а для любой пары элементов $d', d'' \in D$ значение функции принадлежности $\mu_{R_{\succsim}}(d', d'')$ понимается как степень выполнения предпочтения «элемент d' «не хуже» элемента d'' ». Равенство $\mu_{R_{\succsim}}(d', d'') = 0$ означает, что либо $\mu_{R_{\succsim}}(d'', d') > 0$ («элемент d'' «не хуже» элемента d' »), либо — $\mu_{R_{\succsim}}(d'', d') = 0$ (элементы d' и d'' несравнимы между собой).

По заданному нечеткому отношению $R_{\succsim}$ нечеткие отношения $R_{\succ}$ и $R_{\sim}$ определяются следующим образом.

Нечеткое отношение строгого предпочтения имеет функцию принадлежности:

$$\mu_{R_{\succ}}(d', d'') = \begin{cases} \mu_{R_{\succsim}}(d', d'') - \mu_{R_{\succsim}}(d'', d'), & \text{если} \\ \mu_{R_{\succsim}}(d', d'') \geq \mu_{R_{\succsim}}(d'', d'), & \\ 0, & \text{если} \\ \mu_{R_{\succsim}}(d', d'') < \mu_{R_{\succsim}}(d'', d'). & \end{cases}$$

Нечеткое отношение безразличия имеет функцию принадлежности:

$$\begin{aligned} \mu_{R_{\sim}}(d', d'') &= \\ &= \max \{1 - \max \{\mu_{R_{\succ}}(d', d''), \\ &\mu_{R_{\succ}}(d'', d')\}, \min \{\mu_{R_{\succ}}(d', d''), \\ &\mu_{R_{\succ}}(d'', d')\}\}. \end{aligned}$$

Помимо отношения безразличия, которое транзитивно, симметрично и рефлексивно, вводится нечеткое отношение квазиэквивалентности $R_{\#}$ с функцией принадлежности $\mu_{R_{\#}}(d', d'') = \min \{\mu_{R_{\succ}}(d', d''), \mu_{R_{\succ}}(d'', d')\}$.

В отличие от отношения безразличия отношение $R_{\#}$ в общем случае не транзитивно. В противном случае $R_{\#} = R_{\sim}$.

Таким образом, рассмотренные отношения предпочтения обладают следующими свойствами нечетких бинарных отношений: $R_{\sim}$, $R_{\#}$ — рефлексивны и симметричны; $R_{\succ}$ — не рефлексивно и антисимметрично; во всех случаях, если исходное отношение $R_{\succsim}$ транзитивно, то тем же свойством обладают и нечеткие отношения $R_{\#}$ и $R_{\succ}$, а $R_{\sim}$ — всегда транзитивно.

Как и в случае обычных нечетких отношений, важным свойством является свойство связности, которое приводит к тому, что во множестве D все элементы сравнимы по предпочтительности. В случае нечеткого отношения однозначно можно определить лишь полное отсутствие связности: нечеткое отношение $\tilde{R}$ не связно тогда и только тогда, когда среди элементов множества D найдутся такие d' и d'' , для которых $\mu_R(d', d'') = \mu_R(d'', d') = 0$. Поэтому свойство связности нечеткого отношения понимается шире, чем в случае обычного отношения (функция принадлежности которого релейна $\mu_R = \{0; 1\}$), т. е. рассматривают различные степени связности отношения.

Отношение $\tilde{R}$ называется h -связным, если его функция принадлежности удовлетворяет условию

$$\begin{aligned} \max \{\mu_R(d', d''), \mu_R(d'', d')\} &> h, \\ \forall d', d'' \in D, h \in [0, 1]. \end{aligned}$$

Так, для примера 5 отношение предпочтения с функцией принадлежности μ_R является 0,55-связным, т. е. из каждых двух элементов, по крайней мере, один хуже другого со степенью, большей 0,55.

Особое место в задачах принятия решений с использованием нечетких отношений имеют слабо связные ($h = 0$) и сильно связные отношения. Для сильно связного отношения функция принадлежности удовлетворяет условию

$$\begin{aligned} \max \{\mu_R(d', d''), \mu_R(d'', d')\} &= 1, \\ \forall d', d'' \in D, \end{aligned}$$

что эквивалентно утверждению

$$\mu_R(d', d'') = 1 - \mu_R(d'', d').$$

Модельные отношения нечетких отношений предпочтения по аналогии с обычными отношениями имеют различные градации, которые можно сопоставить с приведенными в табл. 1. В этом случае обычное отношение несравнимости и неразличимости задается нечетким отношением «сходство безразличия — подобия», а отношение нестрогого предпочтения, моделируемое несвязным квазипорядком, — «предпорядком»; отношения строгого предпочтения — «нестрогим» и «строгим порядком». Дополнительно используют нечеткие отношения «различия» (антирефлексивное, транзитивное, симметричное) и «несходства» (антирефлексивное, симметричное).

Для того чтобы осуществить однозначный выбор среди множества допустимых альтернатив, выявленные предпочтения ЛПР должны в наибольшей степени соответствовать «идеальной» системе предпочтений. Система предпочтений в таком определении в общем случае должна отвечать следующему набору свойств: транзитивность, асимметричность (или в крайнем случае рефлексивность) и связность. Для обеспечения указанного набора свойств необходимо получить от ЛПР очень «сильную» и объемную информацию как относительно его предпочтений, так и относительно существа рассматриваемой проблемы. При анализе слабо структурированных проблем на начальных этапах исследования в условиях существенной неопределенности и дефицита времени удовлетворить таким требованиям практически невозможно. Это заставляет последовательно выявлять предпочтения ЛПР по мере уяснения им целей и задач, условий проведения операции и получения промежуточных результатов. Каждый отдельный из применяемых способов выявления предпочтений не обеспечивает требуемого набора свойств и лишь частично вскрывает систему предпочтений и, следовательно, лишь частично снимает неопределенность в

выборе. Степень снижения неопределенности зависит от того, какими возможностями обладают те или иные способы выражения предпочтений, т. е. на обеспечение каких именно требований к системе предпочтений они ориентированы в первую очередь. Выбирая тот или иной способ, необходимо руководствоваться, с одной стороны, степенью снижения неопределенности в выборе путем обеспечения конкретного свойства структуры предпочтений, а с другой стороны, — возможностью использовать данный способ на данном шаге принятия решений. Это приводит к противоречию между желанием использовать более простые способы и стремлением максимально вскрыть систему предпочтений ЛПР.

Для решения указанного противоречия рассмотренные способы выражения предпочтения применяют в комплексе, обеспечивая их «вложенность» по последовательному сужению неопределенности в выборе в направлении обеспечения: 1) транзитивности на классах элементов из предъявленной совокупности (способ сортировки с последующим ранжированием классов); 2) асимметричности или хотя бы рефлексивности внутри класса (парное сравнение, балльное оценивание) и получения подклассов «нехудших» элементов; 3) связности элементов внутри подклассов (парное сравнение элементов в долях суммарной или относительной интенсивности, получение коэффициентов важности); 4) проверки согласованности предпочтений (для устранения противоречивых суждений, возникших при декомпозиции задачи).

Моделируемые отношения в случае 1 представляются квазисерией или строгим порядком (частичным или связным), в случае 2 — частичным квазипорядком или эквивалентностью, в случае 3 — связным квазипорядком, т. е. при выборе конкретного способа выявления предпочтений необходимо ориентироваться на то, какие свойства системы предпочтений будут выявлены (см. табл. 1).

В задачах принятия решений отношения предпочтения задают либо аналитическими выражениями, либо логическими условиями, либо алгорит-

мами. Это и определяет конкретный способ выражения и измерения предпочтений.

ФУНКЦИЯ ВЫБОРА И ЯДРО ОТНОШЕНИЯ

Необходимость формализации предпочтений ЛПР для постановки и решения прикладных задач выбора наилучшей альтернативы требует измерения предпочтений для использования их во вложенной процедуре поэтапного выбора с последовательным отбрасыванием части стратегий и сведением к минимуму мощности окончательного подмножества альтернатив, которые в идеале должны оказаться эквивалентными. Это наиболее рациональный принцип выбора стратегий.

В зависимости от «близости» выявленного отношения предпочтения ЛПР к «идеальной» структуре системы предпочтений, приводящей к однозначному выбору, мощность выделяемого конечного подмножества $D^* \subset D$, служит мерой совершенства выявленной системы предпочтений.

Наиболее совершенным является отношение связного квазипорядка. Оно позволяет в максимальной степени сузить множество D^* , в которое будут включаться только так называемые наилучшие элементы. Элемент $d^* \in D$ называется *наилучшим во множестве D* , если он не менее предпочтителен любого другого элемента $d \in D$, т. е.

$$d^* \geq d, \quad \forall d \in D.$$

Наилучший элемент в этом случае «единственен» с точностью до эквивалентности: если d_1^*, d_2^* — два наилучших элемента, то $d_1^* \sim d_2^*$.

Отношение связного квазипорядка ставит в соответствие множеству D подмножество его наилучших объектов $D^* = \{d^*\}$. Это соответствие называется *функцией выбора*:

$$C_{\geq}: D \rightarrow D^*, \text{ где } D^* = \\ = \{d^* \mid d^* \geq d, \quad \forall d \in D\}.$$

Если множество D^* не пусто, то выбор осуществляется только среди этих наилучших элементов $d^* \in D^*$. В случае нарушения связности мно-

жество D^* оказывается пустым. Например, как для графа на рис. 1, а.

Для сужения исходного множества D это вынуждает оперировать лишь частью отношения нестрогого предпочтения ($\geq$) и выделять так называемые недоминируемые (максимальные) элементы. Элемент $d^m \in D$ называется *недоминируемым* по отношению строгого предпочтения ($\succ = \geq \setminus \sim$), если среди остальных элементов множества D не существует ни одного элемента $d \in D$, который был бы строго предпочтительнее элемента d^m . Поскольку отношение строгого предпочтения есть лишь строгий частичный порядок, то таких элементов d^m может быть несколько. Подмножество недоминируемых элементов $\{d^m\}$ называется *ядром отношения строгого предпочтения на множестве D* :

$$M_{\succ} = \{d^m \mid \exists d \in D : d \succ d^m; \\ \forall d, d^m \in D\}.$$

Все элементы d^m , входящие в ядро отношения $M_{\succ}$, несравнимы между собой, т. е., например, $d_1^m \not\geq d_2^m$, $d_2^m \not\geq d_1^m$.

Для графа, изображенного на рис. 1, а, $M_{\succ} = \{W_1, W_4\}$. Из определения наилучшего и недоминируемого элементов непосредственно следует, что наилучший объект всегда является и недоминируемым, а обратное утверждение неверно.

При поэтапном использовании информации от ЛПР исходное отношение строгого предпочтения $\succ^1$ может быть расширено до $\succ^2 \supset \succ^1$. В этом случае оказывается, что $M_{\succ^2} \subseteq M_{\succ^1}$, т. е. элемент, недоминируемый по отношению $\succ^2$, является недоминируемым и по отношению $\succ^1$. Это и объясняет смысл «вложенности» процедуры поэтапного выбора с последовательным отбрасыванием части альтернатив.

Если в процедуре поэтапного выбора используются нечеткие отношения¹,

¹ Для нечетких отношений однозначно определяется лишь полное отсутствие связности.

то на каждом шаге можно выделять лишь недоминируемое множество альтернатив, используя нечеткое отношение строгого предпочтения $R_{>}$. В этом случае при некотором фиксированном $d' \in D$ функцию принадлежности $\mu_{R_{>}}(d', d)$ можно рассматривать как функцию принадлежности нечеткого множества элементов d , которые строго доминируются элементом d' [80]. Тогда ядро нечеткого отношения строгого предпочтения $M_{R_{>}}$ представляет собой пересечение дополнений, введенных в D подмножеств с функциями принадлежности $M_{R_{>}}(d', d)$:

$$\mu_{M_{R_{>}}}(d) = \min_{d' \in D} [1 - \mu_{R_{>}}(d', d)],$$

$$\forall d \in D,$$

или

$$\mu_{M_{R_{>}}}(d) = 1 - \max_{d' \in D} \mu_{R_{>}}(d', d),$$

$$\forall d \in D;$$

$$\mu_{M_{R_{>}}}(d) = 1 - \max_{d' \in D} [\mu_{R_{>}}(d', d) - \mu_{R_{>}}(d, d')].$$

Последнее выражение представляет формальную запись способа обработки исходной нечеткой информации (заданной в форме нечеткого отношения нестрогого предпочтения $R_{>}$) для выделения в D ядра недоминируемых элементов $M_{R_{>}}$.

Значение $\mu_{M_{R_{>}}}(d)$ представляет собой степень, с которой любой элемент d не доминируется ни одним элементом $d' \in D$, т. е. если $\mu_{M_{R_{>}}}(d) = \gamma$, то элемент d может доминироваться другими элементами $d' \in D$, но со степенью не выше, чем $1 - \gamma$.

Так как значение $\mu_{M_{R_{>}}}(d)$ есть степень недоминируемости элемента d , то, естественно, при заданной нечеткой информации выбирать элементы, которые имеют большую степень принадлежности нечеткому множеству $M_{R_{>}}$.

Пример 7. Для нечеткого отношения, заданного в примере 5, ядро нечеткого отношения определяется следующими операциями:

$$\mu_{R_{>}}(W_i, W_j) = \begin{vmatrix} 1 & 0,9 & 0,6 & 0,6 \\ 0,1 & 1 & 0,7 & 0,2 \\ 0,4 & 0,3 & 1 & 0,3 \\ 0,4 & 0,8 & 0,7 & 1 \end{vmatrix};$$

$$\mu_{R_{>}}(W_i, W_j) = \begin{vmatrix} 0 & 0,8 & 0,2 & 0,2 \\ 0 & 0 & 0,3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0,6 & 0,4 & 0 \end{vmatrix};$$

$$\mu_{M_{R_{>}}}(W_i) = \| 1 \quad 0,2 \quad 0,6 \quad 0,8 \|.$$

Анализ ядра нечеткого отношения $M_{R_{>}}(W_i)$ показывает, что наибольшей степенью недоминируемости обладают показатели W_1 и W_4 , причем показатель W_1 четко не доминируется. Замена нечеткого отношения $>$, приводящего к ядру $M_{>} = \{W_1, W_4\}$, нечетким отношением $R_{>}$ с нечетким ядром $\mu_{M_{R_{>}}} = \| 1; 0,2; 0,6; 0,8 \|$ привела к сходным результатам (наибольшей степенью недоминируемости обладают показатели W_1 и W_4). Но, кроме того, более «тонкая» информация о предпочтениях, заданная «размытым» отношением $R_{>}$, позволяет утверждать, что более предпочтительным (по степени недоминируемости) является показатель W_1 . В этом случае показатель W_1 называется четко недоминируемым элементом, т. е.

$\mu_{M_{R_{>}}}(W_1) = 1$,

$$\mu_{M_{R_{>}}}(W_1) = 1,$$

Подобные четко недоминируемые элементы $d^m \in D$ имеют особое значение в задачах принятия решений, так как подмножество таких элементов (если их несколько) рассматривают в некотором смысле как подмножество наилучших элементов или как четкое решение нечетко поставленной задачи выбора.

При h -связном нечетком отношении нестрогого предпочтения $R_{\succ}$ любые два элемента $d_1^m, d_2^m \in D$, являющиеся четко недоминируемыми, эквивалентны со степенью, большей h . При слабой связности нечеткого отношения $R_{\succ}$ любые два четко недоминируемые элемента эквивалентны с положительной степенью. При сильно связном нечетком отношении любые элементы $d_{m1}, d_{m2}^m \in D$ определено эквивалентны (со степенью 1), т. е. выбор любого из них является обоснованным в данной задаче, а их множество есть множество наилучших элементов.

ФУНКЦИЯ ЭФФЕКТИВНОСТИ

Рассмотренные способы выражения предпочтений ЛПР предполагают предъявление ему некоторого ограниченного контрольного множества элементов из D с целью прямого их оценивания по предпочтительности. Из этого следует, что выявленное отношение предпочтения R (четкое или нечеткое) относится только к предъявленным элементам. Если нужно установить предпочтения ЛПР на других элементах, входящих во множество D , то необходимо сформировать новое контрольное предъявление и вновь провести оценивание предпочтений. Такая процедура должна повторяться до полного исчерпания множества D , что практически невозможно по двум причинам. Во-первых, множество D может иметь мощность континуума, а во-вторых, даже если множество D конечно, ЛПР не всегда может длительно консультироваться с исследователем.

Использование выборочного метода и выполнение ряда предпосылок позволяют преодолеть указанные трудности. Отношение ЛПР, выявленное по результатам контрольного предъявления множества элементов, аппроксимируют некоторой числовой функцией, которая каждому элементу $d \in D$ ставит в соответствие действительное число, величина которого согласуется с представлением ЛПР о степени предпочтительности данного элемента (по достижении цели операции, условиям ее проведения, затратам и

т. п., характеристикам каждого элемента $d \in D$). Такую функцию называют *функцией эффективности*¹. С помощью этой функции уже без непосредственного привлечения ЛПР может решаться задача выбора.

В качестве функции эффективности можно рассматривать любое соответствие (функцию, функционал, оператор), для которого выполняется соотношение:

$$d' R d'' \Leftrightarrow W_e(d') \geq W_e(d''), \quad (11.1)$$

где R — отношение нестрогого предпочтения на D (четкое или нечеткое); $W_e(d)$ — значение функции эффективности.

Если в (11.1) неравенство строгое, то $d' \succ d''$ (или $d' R_{\succ} d''$), если неравенство выполняется как равенство, то $d' \sim d''$ (или $d' R_{\sim} d''$). Более предпочтительному элементу в этом случае приписывается большее значение функции эффективности.

Численное представление функции эффективности, как следует из (11.1), возможно в том случае, если отношение: а) R — связно; б) R — транзитивно; в) R — асимметрично и при этом г) предпочтении на множестве D не изменяются скачком.

Требование а) указывает на обязательную сравнимость любых двух элементов $d', d'' \in D$. Предпочтения ЛПР должны быть согласованы (непротиворечивы), что означает транзитивность отношения. Улучшение отдельных характеристик элемента $d \in D$ должно приводить к различию в оценке его предпочтительности. Выполнение этого условия должно приводить к асимметрии отношения R и монотонности функции эффективности. Таким образом, из третьего требования вытекает, что для любого

¹ Часто эту функцию называют функцией ценности или функцией полезности. Эти названия укоренились с тех пор, когда ошибочно считали, что каждому элементу из предъявленной совокупности присуща некоторая объективная полезность (ценность), которую и отражают субъективные предпочтения ЛПР. Авторы считают термин «функция эффективности» более соответствующим смыслу задач принятия решений на основе анализа результатов оценивания эффективности операции.

элемента d достаточно малое изменение одной из его характеристик в смысле улучшения может быть компенсировано определенным изменением другой характеристики, что приводит к эквивалентности исходного и полученного таким образом нового элемента. Условие $\sim$ требует наличия именно такого свойства у отношения R .

Каждый элемент $d \in D$ описывается определенным набором характеристик $x = (x_1, x_2, x_3, \dots, x_{n_0})$. Тогда любой элемент $d \in D$ можно представить точкой в пространстве характеристик $X = X_1 \times X_2 \times \dots \times X_{n_0}$, где X_i — шкала i -й характеристики $i = \overline{1, n_0}$, а каждой точке поставить в соответствие значение функции эффективности. При выполнении условий а—г отношение нестрогого предпочтения R — квазипорядок, а это указывает на существование эквивалентных элементов, т. е. поверхностей равноценности для заданного уровня функции эффективности:

$$W_e(X) = \text{const.} \quad (11.2)$$

Построение функции эффективности представляет самостоятельную задачу. Не всегда эта задача имеет решение. Тогда применяют другие эвристические методы, позволяющие осуществить выбор.

Так как по (11.1) большие значения функции эффективности соответствуют более предпочтительным элементам, то задачу выбора наилучшего элемента $d^* \in D$ (или получения множества D^* таких элементов) можно рассматривать как обычную задачу оптимизации:

$$d^* : \max_{x \in X} W_e(d(x)). \quad (11.3)$$

Из (11.1) также следует, что функция эффективности полностью задает структуру предпочтений ЛПР (отношение R). Однако структуру предпочтений можно описать целым классом функций, удовлетворяющих (11.1), т. е. функция эффективности определяется с точностью до монотонного преобразования и, следовательно, имеет порядковую шкалу.

В качестве характеристик исходов G (т. е. $D = G$) используются частные показатели эффективности, надежности, качества, которые должны иметь ясный физический смысл. Именно это требование позволяет ЛПР уяснить задачу выбора и упорядочить исходы по предпочтительности. Однако для агрегированной характеристики исхода — функции эффективности это требование не обязательно, что существенно упрощает решение задачи (11.3).

Если решением задачи (11.3) является множество наилучших элементов D^* , то для того чтобы провести окончательный выбор, необходимо привлечь дополнительную информацию от ЛПР, расширяющую исходное отношение предпочтения R . В качестве такой информации выступают сведения о характеристиках исхода операции, которые по какой-либо причине не были учтены при начальной постановке задачи. В конечном итоге вложенность отношений $R_1 \subset R_2 \subset \dots \subset R_t$, а следовательно, и вложенность множеств $D_1^* \supset D_2^* \supset \dots \supset D_t^*$ приводят к тому, что на шаге t множество D_t^* оказывается достаточно узким и это позволяет осуществить однозначный выбор.

3. ВЫЯВЛЕНИЕ И ИЗМЕРЕНИЕ ПРЕДПОЧТЕНИЙ МЕТОДАМИ ЭКСПЕРТНОГО ОЦЕНИВАНИЯ

ОБЩАЯ ХАРАКТЕРИСТИКА И КЛАССИФИКАЦИЯ МЕТОДОВ ЭКСПЕРТНОГО ОЦЕНИВАНИЯ

Методы экспертного оценивания используют во всех случаях, когда оцениваемые объекты или их характеристики субъективны и поэтому получить требуемые данные о них путем объективных измерений невозможно (например, с помощью измерительных приборов, действие которых основано на использовании физических законов). Под методом экспертного оценивания обычно понимают комплекс логических и математических процедур, направленных на получение от специалистов-экспертов информации, ее ана-

лиз и обобщение с целью выбора рациональных решений [3, 6, 22, 35].

Идея экспертного оценивания состоит в том, что для получения необходимой новой информации из имеющейся исходной формируется группа экспертов, которые проводят интуитивно-логический анализ того или иного вопроса с целью вынесения качественного или количественного суждения по нему. Суждения экспертов определенным образом обрабатываются с использованием соответствующего математического аппарата, в результате чего получают так называемые *экспертные оценки*. Эти оценки, по-прежнему, субъективны, и степень доверия к ним зависит от знаний, опыта и психологии экспертов.

При построении элементов модели проблемной ситуации и установлении отношений между ними все многообразие задач, решаемых экспертами, сводится к формированию объектов и (или) оценке характеристик. Формирование объектов включает формулировку целей, условий, ограничений, стратегий, показателей эффективности, определение признаков и показателей для описания свойств объектов и их взаимосвязей и т. п. При оценивании характеристик эксперты проводят «измерение» достоверности гипотез и событий, важности целей и показателей эффективности, приоритетов, предпочтений и т. п.

Как следует из перечисленных выше задач, методы экспертного оценивания можно использовать в процессе исследования эффективности для получения различных видов информации, включая информацию о предпочтениях. Вместе с тем при практическом использовании этих методов необходимо иметь в виду, что экспертная оценка — это не решение, а лишь полезная информация, помогающая выработать обоснованное решение. Принимать решение на основе своих предпочтений может только ЛПР, и оно несет за них ответственность. Эксперты несут ответственность только за свои рекомендации. В общем случае предпочтения экспертов могут не совпадать с предпочтениями ЛПР. Это помогает ему критически осмыслить различные точки зрения, уточнить или изменить свои

предпочтения и тем самым уменьшить возможность принятия ошибочных решений. Поэтому при использовании экспертных методов для выявления и измерения предпочтений в роли эксперта должно выступать только лицо, принимающее решение (индивидуальное или групповое). Следует также обратить внимание на существенное различие между групповой экспертизой и коллективным принятием решения.

В зависимости от объема и качества исходной информации разработаны специальные методы как самой экспертизы, так и методы обработки экспертной информации. Если исходная информационная база недостаточна или задача является уникальной, то применяют методы **индивидуального экспертного оценивания**: «интервью», аналитических докладных записок, сценария.

Для ситуации с большим объемом информации обычно привлекают не одного, а нескольких специалистов и используют методы **группового экспертного оценивания**: анкетирования, «комиссий», коллективной генерации идей, комплексные.

Метод интервью заключается в опросе эксперта по заранее сформулированным вопросам, на которые эксперт дает ответы экспромтом.

Метод аналитических докладных записок предполагает длительную и тщательную самостоятельную работу эксперта над анализом тенденций развития, оценкой текущего состояния и путей развития объекта исследования.

Методом сценария устанавливается логическая последовательность событий развития будущего состояния объекта исследования, исходя из существующей ситуации.

Методы интервью и аналитических докладных записок используют в задачах формирования исходного множества стратегий, задачах анализа неопределенностей, а метод сценария — при определении цели и исходов операции, выбора показателей и критериев эффективности.

Основные преимущества этих методов состоят в их оперативности, возможности в полной мере использовать индивидуальные способности эксперта, отсутствии давления авторитетов, низ-

ких затрат на экспертизу. Главным недостатком этих методов является высокая степень субъективности получаемых оценок из-за ограниченности знаний одного специалиста.

Методы группового экспертного оценивания основаны на принципе выявления коллективного мнения. Основное преимущество этих методов над индивидуальными заключается в возможности разностороннего анализа проблемы. Кроме того, коллективная ответственность часто позволяет экспертам принимать более рискованные решения по поставленным перед ними вопросам, а оценки, полученные от группы, чаще содержат «искомую» оценку, чем при индивидуальной экспертизе. Это является весьма важным обстоятельством как в задаче анализа неопределенности, так и при формировании исходного множества стратегий. Недостатками этих методов является сложность самой процедуры получения информации, сложность получения группового мнения экспертов из индивидуальных мнений, возможность оказания давления авторитета в группе.

Указанные недостатки могут быть существенно ослаблены путем применения научной организации экспертизы.

При использовании метода анкетирования опрос экспертов осуществляется анонимно и проводится в несколько туров. Это позволяет каждому эксперту не идти на компромисс, до конца отстаивать свое собственное мнение или корректировать его без чьего-либо давления сверху.

Метод комиссий предполагает свободное обсуждение проблемы между экспертами. Основной недостаток — стремление каждого эксперта к компромиссу.

Метод коллективной генерации идей основан на выдвижении большого числа идей и предположений по решению поставленной проблемы. Процесс генерации идей и их анализ полностью разделены. Обязательно фиксируются все высказанные идеи, запрещается какая-либо оценка выдвигаемых идей. Этот метод используют, как правило, в «тупиковых» ситуациях.

Поскольку при экспертизе специалисты выносят суждения о предпочтительности того или иного способа, указывают возможные сроки наступления некоторых событий, либо формируют альтернативы и упорядочивают их в исходное множество, то в качестве основных элементов экспертного оценивания используются попарное сравнение, сортировка и ранжирование, балльное оценивание и их модификации. После получения информации от экспертов в необходимой форме проводится ее анализ и обработка с целью выявления согласованности мнений отдельных экспертов и формирование (если это необходимо) коллективного мнения. Эта задача решается с помощью специальных процедур, позволяющих выявлять и учитывать компетентность экспертов (метод выявления компетентности экспертов, метод Дельфи, процедура формирования группового мнения). Различные схемы сочетания ограниченного числа этих процедур позволяют решать широкий класс задач различного целевого назначения.

ОРГАНИЗАЦИЯ И ПРОВЕДЕНИЕ ЭКСПЕРТНОГО ОЦЕНИВАНИЯ

Под *организацией экспертного оценивания* понимается комплекс взаимосвязанных мероприятий, определяющих цель работы, условия и способы ее проведения, обеспечение процесса экспертного оценивания, права и обязанности привлекаемых лиц.

Задачу на проведение экспертного оценивания ставит ЛПР (заказчик). Задачу оформляют в виде руководящего документа — решения на проведение экспертного оценивания. Для подготовки решения и руководства всей дальнейшей работой назначают руководителя экспертизы, который определяет состав группы управления.

Условно можно выделить следующие основные этапы проведения экспертного оценивания:

- формирование цели и задач оценивания;

- формирование группы управления и оформление решения на проведение оценивания;

- выбор метода получения экспертной информации и способа ее обработки;

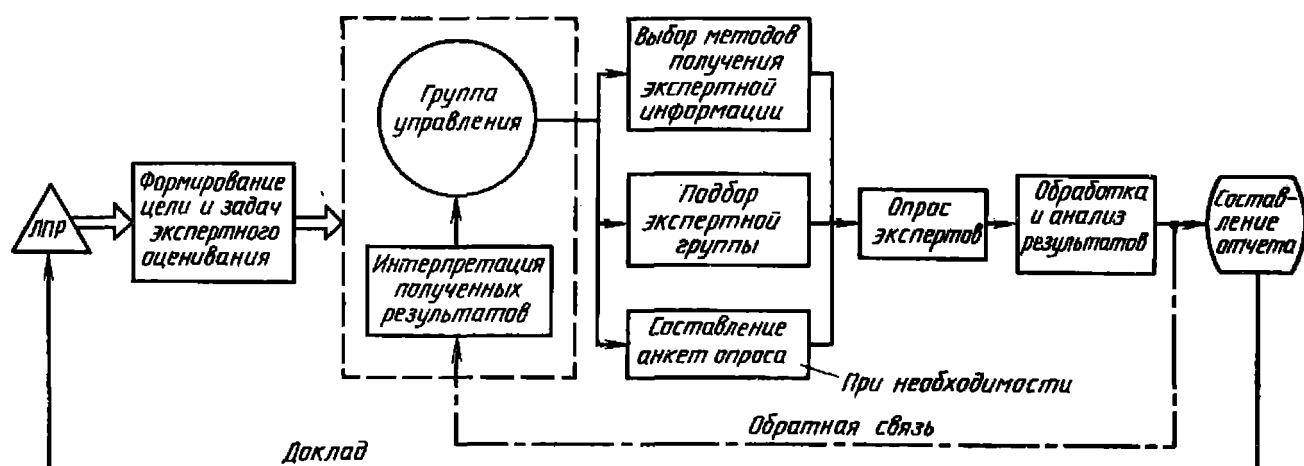


Рис. 4. Взаимосвязь этапов экспертного оценивания

подбор экспертной группы и формирование при необходимости анкет опроса;

опрос экспертов (экспертиза);

обработка и анализ результатов экспертизы;

интерпретация полученных результатов.

составление отчета.

Этап интерпретации полученных результатов необходим для организации обратной связи в процессе экспертного оценивания. Обратную связь с экспертами группа управления осуществляет либо методом Дельфи, либо методом совещаний с обсуждением результатов анонимных опросов. Взаимосвязь указанных этапов представлена на рис. 4.

Этап формирования цели и задач экспертного оценивания является основным, так как от него зависит надежность получаемого результата и его прагматическая ценность. Формирование цели и задач экспертного оценивания во многом определяется существом решаемой проблемы. Должны быть учтены: надежность и полнота имеющейся исходной информации, требуемая форма представления результатов (качественная или количественная), возможные области использования полученной информации, сроки ее представления, наличие имеющихся ресурсов и возможность привлечения специалистов других отраслей знания.

На группу управления возлагается не только вся организационно-плановая работа по обеспечению благопри-

ятных условий для эффективной творческой деятельности экспертов, но и аналитическая работа по подбору экспертной группы, определению методов получения и обработки информации, составлению анкет опросов, содержательной интерпретации получаемых результатов. Для решения этих задач необходимо включить в группу управления высококвалифицированных коммуникабельных специалистов как в области рассматриваемой проблемы, так и в других областях — математики, психологии, социологии.

Экспертную группу обычно подбирают в несколько этапов. Вначале устанавливают отрасли знаний, так или иначе связанные с рассматриваемой проблемой. Затем намечают список «потенциальных» экспертов, которые по своим профессиональным качествам компетентны в этих областях знаний. Такой предварительный отбор проводится на основе легко доступной информации о профессиональной подготовке кандидата — должность, ученое звание и степень, стаж практической деятельности, число публикаций, участие в других экспертизах. При этом желательно, чтобы кандидат в экспертную группу имел широкий общий кругозор и эрудицию. Сама группа не должна, по возможности, состоять из представителей одной отрасли или одной узкой специальности, чтобы исключить влияние ведомственных интересов на цели экспертизы и не сделать получаемые результаты тенденциозными. Если предполагается ис-

2. Оценки (в баллах) профессиональных качеств экспертов

Занимаемая должность	Специалист без степени	Кандидат наук	Доктор наук	Академик, член-корреспондент
Ведущий инженер (конструктор)	1	—	—	—
Младший научный сотрудник	1	1,5	—	—
Старший научный сотрудник	—	2,25	3,0	—
Заведующий лабораторией, сектором, руководителем группы	2,0	3,0	4,0	6,0
Заведующий отделом, заместитель заведующего отделом	2,5	3,75	5,0	7,5
Руководитель комплекса, отделения	3,0	4,5	6,0	9,0
Директор, заместитель директора	4,0	6,0	8,0	12,0

пользовать в экспертизе метод комиссий или другой метод открытых опросов, то состав группы экспертов должен быть однородным, чтобы исключить «давление авторитетом». Первоначально намеченный список кандидатов подвергается дальнейшему анализу для решения вопроса о численном составе экспертной группы, который зависит от ограничений финансового, временного и организационного характера. В окончательный список экспертной группы включают тех кандидатов, которые являются наиболее компетентными.

Компетентность эксперта есть степень его квалификации в определенной области знаний. Она определяется посредством анализа его профессиональной деятельности, широты кругозора по перспективам развития рассматриваемой проблемы.

В зависимости от профессиональной подготовки эксперта (должность, ученое звание, степень) ему приписывают определенный балл. В табл. 2 в качестве примера приведена оценка профессиональных качеств эксперта, работающего в научно-исследовательской организации [22]. Такая информация позволяет предварительно отобрать и ранжировать кандидатов по их компетентности.

В практике экспертного оценивания широко используют оценку компетентности эксперта методами самооценки и взаимной оценки. В первом случае

компетентность эксперта оценивают коэффициентом k ($0 \leq k \leq 1$), который определяют на основе суждений эксперта о степени своей информативности по решаемой проблеме (коэффициент $k_{\text{и}}$) и степени аргументации своих мнений (коэффициент $k_{\text{а}}$) [22]:

$$k = 0,5 (k_{\text{и}} + k_{\text{а}}).$$

Коэффициент информативности $k_{\text{и}}$ получают на основе самооценки эксперта по десятибалльной шкале, как

$$k_{\text{и}} = 0,1 X_{\text{и}},$$

где $X_{\text{и}}$ — балл, выставленный экспертом.

Значение коэффициента $k_{\text{а}}$ определяют следующим образом. Эксперту предъявляют табл. 3 без цифр, в которой он отмечает, какой источник им оценивается по соответствующим градациям: В, С, Н. Группа управления переводит отметки эксперта в шкалу эталонной табл. 3 и вычисляет коэффициент $k_{\text{а}}$ путем суммирования цифр, соответствующих позициям таблицы, отмеченным экспертом.

Другой метод заключается в вычислении коэффициентов компетентности на основе матриц, составленных по результатам взаимной оценки экспертов. Каждый участник должен оценить объем и качество знаний других кандидатов по вопросам анкеты. При составлении анкет по этому методу необходимо заранее выявлять возможные цели экспертов, противореча-

3. Оценки коэффициента аргументации эксперта

Источники аргументации	Степень влияния источника на Ваше мнение		
	В (высокая)	С (средняя)	Н (низкая)
Проведенный Вами теоретический анализ	0,3	0,2	0,1
Ваш производственный опыт	0,5	0,4	0,2
Обобщение работ отечественных авторов	0,05	0,05	0,05
Обобщение работ зарубежных авторов	0,05	0,05	0,05
Ваше личное знакомство с состоянием дел за рубежом	0,05	0,05	0,05
Ваша интуиция	0,05	0,05	0,05

щие целям экспертизы, т. е. исключать те причины, которые могут побудить эксперта сознательно искажать оценки знаний других кандидатов. Использовать этот метод можно, если кандидаты знают друг друга по совместной деятельности.

Одной из эффективных модификаций метода взаимной оценки, которая позволяет не только оценить компетентность уже отобранных кандидатов, но и одновременно выявить возможно полное множество специалистов по рассматриваемой проблеме и сформировать их список, является следующая процедура. Членам группы управления предлагается высказать свое мнение о включении кандидатов в группу экспертов. Названных новых лиц просят, в свою очередь, сделать то же самое. Проведя несколько туров такого опроса, составляют удовлетворительный по полноте список кандидатов. По результатам опроса составляется матрица смежности, элементами которой являются единицы или нули в зависимости от того, высказался ли кандидат с номером i в пользу включения в группу кандидата с номером j или нет соответственно (i — номер строки матрицы смежности, j — номер столбца). После определения матрицы $Z = \|z_{ij}\|$, $i, j = \overline{1, n}$ вычисляют коэффициенты компетентности канди-

датов γ_i по следующему алгоритму.

Алгоритм 1. 1. Задать критерий останова и требуемую точность ε вычисления γ_i , $i = \overline{1, n}$.

2. Положить $t = 0$ и все $\gamma_i^{(t)} = \frac{1}{n}$, $i = \overline{1, n}$.

3. Положить $t = t + 1$.

4. Вычислить

$$\gamma_j^{(t)} = \frac{\sum_{i=1}^n z_{ij} \gamma_i^{(t-1)}}{\sum_{i=1}^n \sum_{j=1}^n z_{ij} \gamma_i^{(t-1)}}, \quad j = \overline{1, n}. \quad (11.4)$$

5. Проверить условие $|\gamma_i^{(t)} - \gamma_i^{(t-1)}| \leq \varepsilon$, $i = \overline{1, n}$. Если условие выполняется, то перейти к п. 6, иначе — к п. 3.

6. Прекратить вычисления. Полученные $\gamma_i^{(t)}$ принять в качестве коэффициентов компетентности γ_i .

В этом алгоритме критерием останова является условие п. 5. Обычно величину ε выбирают на один-два порядка меньше величины $\frac{1}{n}$. Иногда в качестве критерия останова используют условие $t = t^{\text{пр}}$ (обычно $t^{\text{пр}}$ вы-

бирают в диапазоне 3—5, что обусловлено быстрой сходимостью процесса). Содержательный смысл алгоритма состоит в том, что на первой итерации ($t = 1$) подсчитывают отношение суммы «голосов», поданных за кандидата с номером j , к общему числу всех «голосов» (сумма единиц в матрице смежности). На последующих итерациях ($t > 1$) голоса «взвешиваются» коэффициентами компетентности кандидатов $\gamma_i^{(t-1)}$, вычисленными на предыдущей итерации.

Третий метод оценки компетентности кандидатов состоит в проверке достоверности их суждений по результатам контрольных экспертиз. Контрольная экспертиза предусматривает опрос экспертов по тем вопросам, для которых уже известны достоверные ответы, не известные заранее участникам. При этом необходимо, чтобы эти вопросы были связаны с рассматриваемой проблемой и информационно обеспечены настолько, чтобы эксперт мог дать по ним числовую оценку в заданной шкале. Коэффициент (достоверности) в этом случае определяется как отношение числа вопросов, на которые эксперт дал правильные ответы, к общему числу поставленных вопросов:

$$\gamma_i^{(д)} = \frac{N_{\Pi_i}}{N}, \quad i = \overline{1, n},$$

где N_{Π_i} — число правильных ответов i -го эксперта; N — общее число поставленных вопросов.

Этот показатель характеризует эксперта как «измерительный прибор».

Контрольная экспертиза для оценки компетентности экспертов может использоваться даже тогда, когда не известны достоверные ответы на поставленные вопросы. В этом случае используют подход, основанный на обработке нормированных балльных оценок. По этой методике сначала полагают, что все эксперты равнокомпетентны. Затем вычисляют средний нормированный балл каждого из оцениваемых элементов по индивидуальным оценкам экспертов и взвешенную сумму оценок каждого из экспертов. В качестве весовых коэффициентов используют полученные средние оцен-

ки элементов. Вычисленные оценки нормируются их суммой и используются в качестве уточненных значений коэффициентов компетентности. Смысл этой операции состоит в том, что для эксперта, оценки которого ближе к среднегрупповым, коэффициент его компетентности увеличивается. Такая процедура может повторяться неоднократно в ходе контрольных экспертиз.

Пример 8. Четыре эксперта, оценивая относительную важность покупательной способности, себестоимости и затрат в проблеме повышения рентабельности предприятия (см. пример 3) в десятибалльной шкале, дали следующие оценки:

Эксперт	Оцениваемый элемент		
	Покупательная способность W_1	Себестоимость W_2	Затраты W_3
1	8	4	2
2	10	2	1
3	4	8	2
4	5	4	3

Для всех экспертов коэффициент компетентности полагаем равным 0,25. Таблица соответствующих нормированных балльных оценок, определяемых как отношение баллов по каждому из оцениваемых элементов к их суммарным баллам, выставленным каждым из экспертов, имеет вид:

Эксперт	Элемент		
	W_1	W_2	W_3
1	0,57	0,29	0,14
2	0,77	0,15	0,08
3	0,29	0,58	0,13
4	0,42	0,33	0,25

При этом средние баллы, приходящиеся на оцениваемые элементы, со-

ставляют $\frac{1}{4} (0,57 + 0,77 + 0,29 + 0,42) = 0,51$ для W_1 , 0,34 для W_2 и 0,15 для W_3 . Взвешенные суммы относительных балльных оценок экспертов составляют для первого эксперта $0,57 \times 0,51 + 0,29 \times 0,34 + 0,14 \times 0,15 = 0,41$; для второго эксперта 0,45; для третьего 0,31; для четвертого 0,36. Сумма полученных взвешенных суммарных оценок определится величиной $0,41 + 0,45 + 0,31 + 0,36 = 1,53$, а уточненные коэффициенты γ_i компетентности экспертов будут равны: $0,41 : 1,53 = 0,26$ — для первого; 0,3 — для второго; 0,2 — для третьего и 0,24 — для четвертого.

Использование указанных методов отбора кандидатов в экспертную группу обеспечивает более высокую надежность экспертизы, чем волонтаристское назначение экспертов.

Описанные методы экспертного оценивания базируются на соответствующих процедурах опроса, которые различаются по форме общения с экспертом и способу постановки ему вопросов. К таким процедурам относятся: индивидуальные и групповые, очные и заочные, открытые и закрытые опросы. При выборе конкретной процедуры опроса необходимо учитывать не только реальные ограничения проведения экспертизы, но и достоинства и недостатки этих процедур. Дополнительно к ним можно указать, что очные опросы предпочтительнее заочных по информативности, так как позволяют исключить возможное неправильное истолкование экспертом вопросов анкеты. Более того, при заочном опросе эксперт вообще может не дать ответа на некоторые вопросы ввиду их непонимания, хотя при очном опросе это может быть оперативно устранено путем переформулировок и уточнений поставленных вопросов. С учетом затрат на получение информации и исключения психологического давления на эксперта со стороны группы управления (которое может исказить получаемую информацию) более предпочтителен заочный опрос.

Способ постановки вопросов эксперту регламентирует определенную сво-

боду его ответов. Если группа управления желает получить конкретный ответ по интересующей проблеме в том случае, когда нет уверенности, что сам эксперт захочет дать информацию, как правило, используют процедуру закрытого опроса. Эта процедура предусматривает постановку перед экспертом таких вопросов, в формулировке которых заведомо содержится перечень альтернативных ответов. Если вопрос предусматривает ответ в форме только «да» или «нет», то такой вопрос называется *чисто закрытым*. Если требуется указать один из более, чем двух предлагаемых вариантов ответа, то такой вопрос называется *выборным*.

Достоинство этой процедуры в том, что опрос проводится оперативно, эксперты хорошо понимают поставленную задачу и уверенно работают, имея перед собой набор альтернативных ответов. Однако «навязывание» экспертам ответов исключает возможность выражения ими своего мнения в том случае, если оно не совпадает с альтернативными ответами и снимает с него ответственность.

Процедуру закрытого опроса целесообразно использовать при анализе характера неопределенностей, выборе метода моделирования и определении типа модели, анализе полученных результатов моделирования, т. е. в тех случаях, когда набор альтернатив имеется или очевиден.

Процедура открытого опроса дает полную свободу ответов эксперта по рассматриваемой проблеме. Вопрос ставится в наиболее общей форме, отражая лишь существо задачи. Такая процедура применяется в ситуациях, требующих нетривиального решения (по своему целевому назначению эта процедура сходна с методом коллективной генерации идей). Недостатки этой процедуры заключаются в том, что для нее требуются неформальные методы обработки полученной информации в виду «свободной» интерпретации вопроса и ответа, а также высокая квалификация экспертов.

Процедуру открытого опроса, как правило, применяют в задачах формирования исходного множества стратегий, выбора показателей и критериев эффективности, прогнозирования по-

ведения других субъектов операции и в других случаях, требующих обнаружения с помощью экспертов неясных элементов проблемной ситуации.

Для устранения недостатков, присутствующих каждой из процедур, необходимо использовать эти процедуры комплексно. Например, при проведении закрытого опроса можно предоставить экспертам возможность вообще не выбирать одну из представленных альтернатив, или выдвигать собственный ответ по поставленным вопросам.

В основе большинства из описанных методов лежит такое средство сбора информации, как опросный лист — анкета. По целевому назначению анкеты можно разделить на анкеты-справки, содержащие данные о профессиональной подготовке эксперта, используемые на этапе формирования экспертной группы, и анкеты-опросники, содержащие логически увязанную систему вопросов по проблеме, интересующей группу управления. Как правило, первоначальный опрос проводится по анкетам, содержащим открытые вопросы, требующие, например, ответа в виде: «1. Каковы по Вашему мнению особенности развития предприятия в следующей пятилетке? 2. Каковы по Вашему мнению направления повышения рентабельности предприятия в этот период? 3. Перечислите доводы в пользу указанных Вами направлений» и т. п.

Обработка полученных результатов по этим анкетам приводит к множеству альтернатив (см. пример 3 гл. 4). Последующий опрос, как правило, проводят по анкетам с вопросами закрытого типа. При этом в анкете не должно содержаться вопросов, допускающих двойное толкование. Само построение вопросов должно быть таким, чтобы эксперт последовательно раскрывал существо проблемы, каждый раз опираясь на информацию, содержащуюся в предыдущих вопросах. Это означает, что ответы на первые в общей иерархии вопросы должны базироваться на самой надежной и доступной для эксперта информации и носить по возможности качественный характер. Эксперт должен указать, например, различие альтернатив по

предпочтительности: «Что является более предпочтительным с точки зрения рентабельности предприятия: покупательная способность, себестоимость продукции, затраты?»

Последующие вопросы анкеты должны составляться так, чтобы для ответа на них требовалась более «совершенная» информация, например, в форме диапазонов значений интересующих факторов. Например: «1. Каково Ваше мнение о минимальном значении уровней покупательной способности, себестоимости и затрат, обеспечивающих решение задачи повышения рентабельности предприятия в следующей пятилетке из числа представленных значений? 2. Каковы по Вашему мнению максимально достижимые значения уровней этих показателей из числа представленных?» (эти вопросы могут ставиться и как открытые).

Последними вопросами анкеты должны быть такие, для ответа на которые требуется информация в виде точечной оценки (числа). Если, например, цель экспертизы заключается в выявлении относительного вклада каждого из факторов в достижение цели операции, то последним в анкете должен быть вопрос: «Каков по Вашему мнению вклад каждого фактора (покупательная способность, себестоимость, затраты) в повышение рентабельности? Оцените вклад каждого фактора в десятибалльной шкале». Если эксперту сразу задать последний вопрос, то ответ на него вызовет значительные трудности или он вообще откажется отвечать. Организация анкеты по принципу логической увязки и усложнения вопросов позволяет эксперту самому глубже разобратся в проблеме и выдать обоснованную непротиворечивую информацию.

Помимо целевых вопросов анкета должна содержать информацию о правилах ее заполнения.

Способ обработки полученной экспертной информации определяется методом ее получения и типом представления (качественная, количественная).

Используя описанные методы, проводят опрос экспертов. Полученные результаты обрабатываются и анализируются. Группа управления по

результатам анализа принимает решение относительно целесообразности проведения дополнительных циклов экспертизы. Это означает, что опрос экспертов, как правило, осуществляется в несколько туров до получения согласованных результатов. Таким образом достигается регулируемая обратная связь, компенсирующая неточность первоначальной постановки задачи и повышающая надежность получаемых экспертных оценок.

ОБРАБОТКА И АНАЛИЗ РАНЖИРОВОК

При ранжировании эксперт должен расположить оцениваемые элементы в порядке возрастания (убывания) их предпочтительности и приписать каждому из них ранги в виде натуральных чисел. При прямом ранжировании наиболее предпочтительный элемент имеет ранг 1 (иногда 0), а наименее предпочтительный — ранг m .

Если эксперт не может осуществить строгое ранжирование из-за того, что, по его мнению, некоторые элементы одинаковы по предпочтительности, то допускается присваивать таким элементам одинаковые ранги. Чтобы обеспечить равенство суммы рангов сумме мест ранжируемых элементов, применяют так называемые стандартизированные ранги. *Стандартизированный ранг* есть среднее арифметическое номеров элементов в ранжированном ряду, являющихся одинаковыми по предпочтительности.

Пример 9. Эксперт упорядочил шесть элементов по предпочтению следующим образом:

$$d_1 \succ d_3 \sim d_4 \succ d_6 \succ d_2 \sim d_5.$$

Тогда стандартизированные ранги этих элементов будут

d_i	d_1	d_2	d_3	d_4	d_5	d_6
r_{d_i}	1	5,5	2,5	2,5	5,5	4

Таким образом, сумма рангов, приписанных элементам, будет равна сумме чисел натурального ряда.

Точность выражения предпочтения путем ранжирования элементов существенно зависит от мощности множества предъявления. Процедура ран-

жирования дает наиболее надежные результаты (по степени близости выявленного предпочтения и «истинного»), когда число оцениваемых элементов не более 10. Предельная мощность множества предъявления не должна превосходить 20.

Обработка и анализ ранжировок проводится с целью построения группового отношения предпочтения на основе индивидуальных предпочтений. При этом могут ставиться следующие задачи: а) определение тесноты связи между ранжировками двух экспертов на элементах множества предъявления; б) определение взаимосвязи между двумя элементами по индивидуальным мнениям членов группы относительно различных характеристик этих элементов; в) оценка согласованности мнений экспертов в группе, содержащей более двух экспертов.

В первых двух случаях в качестве меры тесноты связи используется коэффициент ранговой корреляции. В зависимости от того, допускается ли только строгое или нестрогое ранжирование, используется коэффициент ранговой корреляции либо Кендалла, либо Спирмена.

Коэффициент ранговой корреляции Кендалла для задачи а

$$\tau = \frac{2}{m(m-1)} \sum_{i < j} \text{sign}(r_{1j} - r_{1i}) \times \text{sign}(r_{2j} - r_{2i}), \quad (11.5)$$

где m — число элементов; r_{1i} — ранг, приписанный первым экспертом i -му элементу; r_{2i} — то же, вторым экспертом.

Для задачи б компоненты (11.5) имеют следующий смысл: m — число характеристик двух оцениваемых элементов; r_{1i} (r_{2i}) — ранг i -й характеристики в ранжировке первого (второго) элемента, выставленный группой экспертов.

При строгом ранжировании используется коэффициент ранговой корреляции ρ Спирмена:

$$\rho = 1 - \frac{6}{m(m^2-1)} \sum_{i=1}^m (r_{1i} - r_{2i})^2, \quad (11.6)$$

компоненты которого имеют тот же смысл, что и в (11.5).

Коэффициенты корреляции (11.5), (11.6) изменяются от -1 до $+1$. Если коэффициент корреляции равен $+1$, то это означает, что ранжировки одинаковы; если он равен -1 , то — противоположны (ранжировки обратны друг другу). Равенство коэффициента корреляции нулю означает, что ранжировки линейно независимы (некоррелированы).

Поскольку при таком подходе (эксперт — «измеритель» со случайной погрешностью) индивидуальные ранжировки рассматриваются как случайные, то возникает задача статистической проверки гипотезы о значимости полученного коэффициента корреляции. В этом случае используют критерий Неймана—Пирсона [2, 68]: задаются уровнем значимости критерия α и, зная законы распределения коэффициента корреляции, определяют пороговое значение c_α , с которым сравнивают полученное значение коэффициента корреляции. Критическая область — правосторонняя (в практике обычно сначала рассчитывают значение критерия и определяют по нему уровень значимости, который сравнивают с пороговым уровнем α).

Коэффициент ранговой корреляции τ Кендалла имеет при $m > 10$ распределение, близкое к нормальному с параметрами:

$$\left. \begin{aligned} M[\tau] &= 0; \\ D[\tau] &= \frac{2(2m+5)}{9m(m-1)}, \end{aligned} \right\} \quad (11.7)$$

где $M[\tau]$ — математическое ожидание; $D[\tau]$ — дисперсия.

В этом случае используют таблицы функции стандартного нормального распределения:

$$F_T(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt,$$

а граница τ_α критической области определяется как корень уравнения

$$\alpha = 1 - F_T\left(\frac{\tau_\alpha}{\sqrt{D[\tau]}}\right). \quad (11.8)$$

Если вычисленное значение коэффициента $\tau \geq \tau_\alpha$, то считается, что ранжировки действительно хорошо согласуются. Обычно значение α выбирают в пределах $0,01-0,05$. Для $m \leq 10$ распределение τ приведено в табл. 4.

Проверка значимости согласованности двух ранжировок с использованием коэффициента ρ Спирмена осуществляется в том же порядке с использованием таблиц распределения Стьюдента при $m > 10$.

В этом случае величина

$$t = \rho \sqrt{\frac{m-2}{1-\rho^2}} \quad (11.9)$$

имеет распределение, хорошо аппроксимируемое распределением Стьюдента [38] с $m-2$ степенями свободы. При $m > 30$ распределение величины ρ хорошо согласуется с нормальным, имеющим $M[\rho] = 0$ и $D[\rho] = \frac{1}{m-1}$.

Для $m \leq 10$ проверку значимости ρ осуществляют с помощью табл. 5.

Если ранжировки нестрогие, то коэффициент Спирмена

$$\rho_N = \frac{\rho + S_1 + S_2}{\sqrt{(1-S_1)(1-S_2)}}, \quad (11.10)$$

где ρ — вычисляют по (11.6);

$$\begin{aligned} S_1 &= \frac{3}{m(m^2-1)} \sum_{i=1}^{k_1} l_i(l_i-1), \\ S_2 &= \frac{3}{m(m^2-1)} \sum_{i=1}^{k_2} l_i(l_i-1), \end{aligned} \quad (11.11)$$

где k_1, k_2 — число различных групп нестрогих рангов в первой и второй ранжировках соответственно; l_i — число одинаковых рангов в i -й группе. При практическом использовании коэффициентов ранговой корреляции ρ Спирмена и τ Кендалла следует иметь в виду, что коэффициент ρ обеспечивает более точный результат в смысле минимума дисперсии.

Пример 10. Два эксперта провели ранжирование показателей эффективности $W_1 - W_4$ (см. пример 3). Необходимо установить степень близости

4. Распределение коэффициента ранговой корреляции Кендалла

S_α	m				S_α	m		
	4	5	8	9		6	7	10
0	0,625	0,592	0,548	0,540	1	0,500	0,500	0,500
2	0,375	0,408	0,452	0,460	3	0,360	0,386	0,431
4	0,167	0,242	0,360	0,381	5	0,235	0,281	0,364
6	0,042	0,117	0,274	0,306	7	0,136	0,191	0,300
8		0,042	0,199	0,238	9	0,068	0,119	0,242
10		0,0083	0,138	0,179	11	0,028	0,068	0,190
12			0,089	0,130	13	0,0083	0,035	0,146
14			0,054	0,090	15	0,0014	0,015	0,108
16			0,031	0,060	17		0,0054	0,078
18			0,016	0,038	19		0,0014	0,054
20			0,0071	0,022	21		0,0002	0,036
22			0,0028	0,012	23			0,023
24			0,0009	0,0063	25			0,014
26			0,0002	0,0029	27			0,0083
28				0,0012	29			0,0046
30				0,0004	31			0,0023
					33			0,0011
					35			0,0005

В таблице даны вероятности $\alpha = P \{ \hat{S}_\tau \geq S_\alpha | m \}$, где $S_\tau = \frac{m^2 - m}{2} \tau$.

мнений экспертов по Спирмену и Кендаллу и оценить значимость коэффициентов ранговой корреляции ρ , τ . Получены следующие результаты экспертизы (прямая нумерация рангов):

Эксперт	Оцениваемый элемент			
	w_1	w_2	w_3	w_4
1	1	2	4	3
2	2	1	4	3

По (11.6) вычисляют коэффициент

$$\rho = 1 - \frac{6}{4(16-1)} [(1-2)^2 + (2-1)^2 + (4-4)^2 + (3-3)^2] = 0,8.$$

Для того чтобы воспользоваться табл. 5, вычислим значение

$$S_\rho = \frac{m(m^2 - 1)}{6} (1 - \rho) = \frac{4(16 - 1)}{6} (1 - 0,8) = 2,0.$$

Так как критическая область определяется равенством $\rho > \rho_\alpha$, то равносильное неравенство для S_ρ согласно (11.6) имеет вид $S_\rho < S_\alpha$. Поэтому

$$\begin{aligned} \alpha &= P \{ \hat{S}_\rho < S_\alpha | m \} = \\ &= P \{ \hat{S}_\rho \geq S_{\max} - S_\alpha | m \} = \\ &= P \{ \hat{S}_\rho \geq 20 - 2 | m \} = 0,167 \end{aligned}$$

(см. примечание к табл. 5). Большее значение вероятности α свидетельствует о том, что, например, на 5 %-ном уровне значимости ($\alpha = 0,05$) гипо-

5. Распределение коэффициента ранговой корреляции ρ Спирмена

$m = 4$		$m = 5$		$m = 6$		$m = 7$		$m = 8$		$m = 9$		$m = 10$	
S_α	α	S_α	α	S_α	α	S_α	α	S_α	α	S_α	α	S_α	α
12	0,458	22	0,475	50	0,210	74	0,249	108	0,250	156	0,218	208	0,235
14	0,375	24	0,392	52	0,178	78	0,198	114	0,195	164	0,168	218	0,184
16	0,208	26	0,342	54	0,149	82	0,151	120	0,150	172	0,125	228	0,139
18	0,167	28	0,258	56	0,121	86	0,118	126	0,108	180	0,089	238	0,102
20	0,042	30	0,225	58	0,088	90	0,083	132	0,076	188	0,060	248	0,072
		32	0,175	60	0,068	94	0,055	138	0,048	196	0,038	258	0,048
		34	0,117	62	0,051	98	0,033	144	0,029	204	0,022	268	0,030
		36	0,067	64	0,029	102	0,017	150	0,014	212	0,011	278	0,017
		38	0,042	66	0,017	106	0,0062	156	0,0054	220	0,0041	288	0,0087
		40	0,0083	68	0,0083	110	0,0014	162	0,0011	228	0,0010	298	0,0036
				70	0,0014							308	0,0011
20		40		70		112		168		240		330	

Примечание. В таблице даны вероятности $\alpha = P \{ \hat{S}_\rho \geq S_\alpha | m \}$, где $S_\rho = \frac{m(m^2 - 1)}{6} (1 - \rho)$.

В последней строке указаны значения $S_{\max} = \frac{m(m^2 - 1)}{3}$. Для определения вероятности $P \{ \hat{S}_\rho < S_\alpha | m \}$ используют формулу $P \{ \hat{S}_\rho < S_\alpha | m \} = P \{ \hat{S}_\rho \geq S_{\max} - S_\alpha \}$.

тезу о независимости мнений экспертов не следует отвергать. Статистическая незначимость полученной оценки обусловлена главным образом малым числом оцениваемых элементов.

Используя коэффициент ранговой корреляции τ Кендалла, по (11.5) получаем

$$\begin{aligned} \tau &= \frac{2}{4(4-1)} (\text{sign}(2-1) \text{sign}(1-2) + \\ &+ \text{sign}(4-1) \text{sign}(4-2) + \\ &+ \text{sign}(3-1) \text{sign}(3-2) + \\ &+ \text{sign}(4-2) \text{sign}(4-1) + \\ &+ \text{sign}(3-2) \text{sign}(3-1) + \\ &+ \text{sign}(3-4) \text{sign}(3-4)) = \\ &= \frac{1}{6} (-1 + 1 + 1 + 1 + 1 + 1) = \frac{2}{3}. \end{aligned}$$

Величина $S_\tau = \frac{16-4}{2} \cdot \frac{2}{3} = 4$.

Вероятность выполнения неравенства $\tau \geq \tau_\alpha$ эквивалентна вероятности события $\hat{S}_\tau \geq S_\alpha$. Поэтому, воспользовавшись табл. 4, определим $\alpha = P \{ \hat{S}_\tau \geq S_\alpha \} = 0,167$.

Для иллюстрации задачи определения взаимосвязи между двумя элементами по индивидуальным мнениям группы экспертов приведем следующий пример.

Пример 11. Для решения задач повышения качества продукции и снижения ее себестоимости (см. пример 1 гл. 4), связанных с проблемой повышения рентабельности предприятия, эксперты предложили использовать шесть мероприятий (см. пример 3 гл. 4) и провели их ранжирование с учетом их эффективности по решению каждой из указанных задач. Требу-

ется установить степень влияния предложенных мероприятий на одновременное решение двух указанных задач.

Результаты нестрогого ранжирования после обработки индивидуальных мнений оказались такими:

Элементы	Мероприятия					
	Повышение качества сырья	Создание премиального фонда	Модернизация оборудования	Повышение квалификации рабочих	Сокращение управленческого аппарата	Закупка лицензий
Повышение качества продукции	1	5,5	2,5	2,5	5,5	4
Снижение себестоимости	1	6	4,5	2	4,5	3

По (11.6) коэффициент

$$\rho = 1 - \frac{6}{6(36-1)} (0 + 0,25 + 4 + 0,25 + 1 + 1) = 0,81.$$

По (11.10) и (11.11)

$$S_1 = \frac{3}{6(36-1)} [2(2-1) + 2(2-1)] = \frac{2}{35};$$

$$S_2 = \frac{3}{6(36-1)} [2(2-1)] = \frac{1}{35};$$

$$\rho_H = \frac{0,81 + \frac{2}{35} + \frac{1}{35}}{\sqrt{\left(1 - \frac{2}{35}\right)\left(1 - \frac{1}{35}\right)}} = 0,86.$$

Проверка значимости коэффициента ρ_H при использовании табл. 5 показывает, что $\alpha = 0,02$. Поэтому на 5 %-ном уровне значимости можно считать, что реализация всех шести мероприятий с высокой степенью ($\rho_H = 0,86$) обеспечивает одновременное решение двух указанных задач.

Приведенные в примере 11 нестрогие ранжировки мероприятий получены в результате формирования группового мнения экспертов. Групповое мнение

можно формировать только тогда, когда индивидуальные предпочтения экспертов хорошо согласуются. Для оценки согласованности мнений в группе применяют два способа. Первый способ основан на вычислении средних значений коэффициентов парной корреляции:

$$\bar{\tau} = \frac{2 \sum_{l < i} \tau_{li}}{n(n-1)}; \quad (11.12)$$

$$\bar{\rho} = \frac{2 \sum_{l < i} \rho_{li}}{n(n-1)}, \quad (11.13)$$

где l, i — номера экспертов ($l, i = 1, n$); τ_{li} и ρ_{li} — парные коэффициенты корреляции по Кендаллу и Спирмену соответственно.

Этот способ удобно применять при малом числе n экспертов. В случае полной согласованности мнений экспертов средние значения коэффициентов $\bar{\tau}$ и $\bar{\rho}$ принимают значения, равные 1.

Второй способ основан на вычислении специальных коэффициентов согласованности коэффициента конкордации S_K Кендалла и энтропийного коэффициента согласия S_E .

Коэффициент конкордации Кендалла имеет различный вид в зависимости

от типа ранжирования. При строгом ранжировании

$$C_K = \frac{12S_c}{n^2m(m^2-1)}, \quad (11.14)$$

где

$$S_c = \sum_{j=1}^m \left(\sum_{l=1}^n r_{lj} - \frac{n(m+1)}{2} \right)^2; \quad (11.15)$$

r_{lj} — ранг j -го элемента, присвоенный l -м экспертом; m — число оцениваемых элементов; n — число экспертов.

Для нестрогих ранжировок коэффициент конкордации Кендалла

$$C_K^H = \frac{12S_c}{n^2m(m^2-1) - n \sum_{l=1}^n S_l}, \quad (11.16)$$

где

$$S_l = \sum_{k=1}^{T_l} (t_k^3 - t_k); \quad (11.17)$$

t_k — число одинаковых рангов в k -й группе нестрогих рангов в ранжировке l -го эксперта; T_l — число групп одинаковых рангов в ранжировке l -го эксперта.

Коэффициенты (11.14) и (11.16) равны 1 в случае полного совпадения мнений экспертов и равны 0 в случае их полной несогласованности ($0 \leq C_K, C_K^H \leq 1$). Статистическая проверка значимости полученных коэффициентов конкордации осуществляется обычным порядком при использовании правосторонней критической области $C_K^H \geq C_\alpha$. В табл. 6 приведены

значения вероятностей: $\alpha = P(\hat{S}_c \geq S_\alpha | m, n) = P(\hat{C}_K^H > C_\alpha | m, n)$ для $m = 3 \div 10$ и $n = 3 \div 5$. Для $m > 7$ используют таблицы распределения χ^2 для величины $n(m-1)C_K^H$ с $m-1$ степенями свободы, которая линейно связана с коэффициентом конкордации.

Согласованность экспертов удовлетворительная, если $C_K^H \simeq 0,5$ и $\alpha \leq 0,01$, и хорошая, если $C_K^H \geq 0,7$ и $\alpha < 0,001$.

Энтропийный коэффициент согласия

$$C_E = 1 - \frac{H_E}{m \log m}, \quad (11.18)$$

где энтропия

$$H_E = - \sum_{j=1}^m \sum_{l=1}^n P_{lj} \log P_{lj}; \quad (11.19)$$

P_{lj} — оценка вероятности l -го ранга, присваиваемого j -му элементу. Эти оценки вероятностей вычисляют как отношение числа экспертов n_{lj} , присвоивших элементу j ранг l , к общему числу экспертов n :

$$P_{lj} = \frac{n_{lj}}{n}. \quad (11.20)$$

Энтропийный коэффициент согласия изменяется от 0 до 1.

Использование коэффициентов согласия C_K и C_E обеспечивает примерно одинаковые результаты при близких мнениях экспертов. Для вычислений проще использовать коэффициент C_K . Однако, если эксперты в своих мнениях разделились на две подгруппы, дающие противоположные ранжировки, то коэффициент C_K обратится в нуль, а C_E будет больше 0,5. Таким образом, использование энтропийного коэффициента согласия позволяет достаточно легко установить наличие двух противоположных в своих суждениях подгрупп.

Если в результате анализа выявлено, что согласованность экспертов неудовлетворительна (либо малы значения коэффициентов $\bar{r}$, $\bar{\rho}$, C_K^H , C_E , либо они незначимо отличны от нуля), необходимо произвести содержательный анализ причин расхождения мнений. Ими могут быть: наличие в группе экспертов с оригинальными мнениями (резко отличающимися от мнений остальных), разделение экспертов на несколько групп, придерживающихся своей точки зрения, и др. Одним из методов, позволяющих согласовать мнения или окончательно разделить экс-

6. Распределение коэффициента конкордации C_K^H

$m = 3$													
$n = 3$		$n = 4$		$n = 5$		$n = 6$		$n = 7$		$n = 8$		$n = 9$	
S_α	α	S_α	α	S_α	α	S_α	α	S_α	α	S_α	α	S_α	α
6	0,528	8	0,431	14	0,367	18	0,252	24	0,237	26	0,236	32	0,187
8	0,361	14	0,273	18	0,182	24	0,184	26	0,192	32	0,149	38	0,154
14	0,194	18	0,125	24	0,124	26	0,142	32	0,112	38	0,120	42	0,107
18	0,028	24	0,069	26	0,093	32	0,072	38	0,085	42	0,079	50	0,069
		26	0,042	32	0,039	38	0,052	42	0,051	50	0,047	56	0,048
		32	0,0046	38	0,006	42	0,029	50	0,027	56	0,030	62	0,031
				42	0,0009	50	0,012	56	0,016	72	0,0099	78	0,010
				50	0,0008	54	0,0081	62	0,0084	78	0,0048	86	0,0006
						56	0,0055	72	0,0036	86	0,0024	98	0,0029
						62	0,0017	78	0,0012	98	0,0009	104	0,0013
						72	0,0001	96	0,0003			114	0,0007

$m = 3$		$m = 4$								$m = 5$	
$n = 10$		$n = 3$		$n = 4$		$n = 5$		$n = 6$		$n = 3$	
S_α	α	S_α	α	S_α	α	S_α	α	S_α	α	S_α	α
32	0,222	19	0,342	32	0,200	41	0,210	46	0,218	46	0,213
42	0,135	21	0,300	36	0,158	43	0,162	52	0,163	50	0,163
50	0,092	25	0,207	40	0,105	51	0,107	62	0,108	56	0,096
56	0,066	27	0,175	46	0,068	57	0,075	68	0,073	60	0,063
62	0,046	29	0,148	50	0,052	61	0,055	74	0,056	62	0,056
74	0,026	33	0,075	54	0,033	67	0,034	80	0,037	66	0,038
86	0,012	35	0,054	62	0,012	81	0,012	100	0,010	74	0,015
96	0,0075	37	0,033	66	0,0062	85	0,0067	108	0,0061	78	0,0053
104	0,0034	41	0,017	70	0,0027	93	0,0023	118	0,0028	82	0,0028
122	0,0013	45	0,0017	74	0,0009	101	0,0014	128	0,0009	86	0,0009
126	0,0008					105	0,0006				

пертов на подгруппы, является метод Дельфи.

Проверив согласованность мнений экспертов (в том числе в выделенных подгруппах) и убедившись, что она достаточно высока, формируют групповое мнение. Его осуществляют либо с помощью средних рангов, набранных каждым элементом:

$$\bar{r}_j = \sum_{l=1}^n r_{lj},$$

(11.21)

либо отысканием некоторой ранжировки, которая наиболее тесно коррелирована с остальными;

$$r^* = \max_{r \in \{P\}} \sum_{l=1}^n K_{r,l}, r,$$

где r^* — искомая групповая ранжировка $r^* = (r_1, r_2, \dots, r_m)^T$; $\{P\}$ — множество всевозможных ранжировок (в том числе и нестрогих), построенных на m элементах множества предъ-

явления; $K_{r^l, r}$ — коэффициент корреляции Кендалла или Спирмена между ранжировкой r^l l -го эксперта и ранжировкой r из множества $\{r\}$, либо минимизацией специально выбранной метрики κ на множестве квази-порядков.

В последнем случае каждая ранжировка r^l l -го эксперта рассматривается как точка в m -мерном пространстве квазипорядков. Обычно используют метрику Хемминга:

$$\kappa(r^l, r^k) = \sum_{j=1}^m \sum_{i=1}^m |a_{ij}^l - a_{ij}^k|, \quad (11.22)$$

где a_{ij}^l — элементы матрицы попарных сравнений элементов d_i, d_j ; $i, j = \overline{1, m}$, построенные по ранжировке l -го эксперта;

$$a_{ij}^l = \begin{cases} 1, & \text{если } r_i^l < r_j^l; \\ 0, & \text{если } r_i^l = r_j^l; \\ -1, & \text{если } r_i^l > r_j^l, \end{cases} \quad (11.23)$$

r_i^l — ранг i -го элемента в ранжировке l -го эксперта в прямой нумерации рангов (см. пример 9).

С использованием введенной метрики κ групповая ранжировка определяется либо как медиана

$$r_m : \min_{r \in \{r\}} \sum_{l=1}^n \gamma_l \kappa(r^l, r), \quad (11.24)$$

либо как средняя ранжировка

$$r_c : \min_{r \in \{r\}} \sum_{l=1}^n \gamma_l \kappa^2(r^l, r), \quad (11.25)$$

где γ_l — коэффициент компетентности

$$l\text{-го эксперта } \left(\sum_{l=1}^n \gamma_l = 1 \right).$$

При ранжировании элементов по нескольким характеристикам вначале отыскивается «медианное мнение»

l -го эксперта по всем характеристикам (с учетом их весов):

$$r_{med}^l : \min_{r^l \in \{r\}} \sum_{t=1}^{t_0} q_t \kappa(r^{l, t}, r^l),$$

$$l = \overline{1, n}, \quad (11.26)$$

где q_t — коэффициент важности t -й характеристики ($t = \overline{1, t_0}$); $r^{l, t}$ — ранжировка элементов по t -му признаку l -м экспертом, а затем по (11.24) с подстановкой вместо r^l ранжировки r_{med}^l определяется групповая ранжировка r_m .

Задачи (11.22), (11.24)—(11.26) являются сложными для вычисления, поэтому групповую ранжировку определяют упорядочением по суммарным рангам (11.21). Для учета компетентности экспертов (11.21) трансформируют следующим образом:

$$\bar{r}_j = \sum_{l=1}^n \gamma_l r_{lj}. \quad (11.27)$$

Построение групповой ранжировки по суммарным взвешенным рангам (11.27) при хорошей согласованности экспертов обеспечивает практически одинаковые результаты с решениями задач (11.24) и (11.25).

Построить групповую ранжировку можно с помощью перехода от индивидуальных ранжировок к матрицам попарных сравнений и вычислений собственного вектора. Этот подход более обоснован в теоретическом отношении.

ОБРАБОТКА И АНАЛИЗ БАЛЛЬНЫХ И ТОЧЕЧНЫХ ОЦЕНОК

В некоторых случаях для выражения предпочтения удобнее сначала осуществить балльное оценивание элементов, а затем их ранжирование. Балльные оценки имеют промежуточную шкалу между качественной и количественной. Поэтому обработку и анализ экспертных суждений, выраженных в балльной шкале, обычно осуществляют комбинированным мето-

дом. Сначала балльные оценки переводят в ранги по правилу:

$$b_1 \geq b_2 \geq \dots \geq b_{m-1} \geq b_m \iff r_1 \leq r_2 \leq \dots \leq r_{m-1} \leq r_m \quad (11.28)$$

(где b_j — балл, присвоенный элементу $d_j \in D$, а r_j — ранг этого элемента) и обрабатывают их как ранжировки.

Затем баллы рассматривают как количественную меру выражения предпочтения эксперта. Если результаты обработки оказываются хорошо согласованными, то обоснованность группового мнения будет надежной.

Обработка и анализ балльных оценок как количественных показателей возможна лишь в том случае, если шкала балльных оценок непрерывна или имеет большое число градаций, а также установлены правила начисления баллов. Дальнейшая обработка и анализ проводятся статистическими методами оценивания в предположении, что разница в ответах экспертов объясняется лишь случайными погрешностями.

Каждому элементу $d_j \in D$ ставится в соответствие средний балл, определяемый как среднее арифметическое балльных оценок экспертов по рассматриваемому элементу:

$$b_j = \frac{1}{n} \sum_{l=1}^n b_{lj}, \quad j = \overline{1, m}, \quad (11.29)$$

где b_{lj} — балл, присвоенный j -му элементу l -м экспертом.

Степень согласованности экспертов в этом случае оценивается дисперсиями индивидуальных балльных оценок:

$$\sigma_j^2 = \frac{1}{n-1} \sum_{l=1}^n (b_{lj} - b_j)^2, \quad j = \overline{1, m}, \quad (11.30)$$

или (если балльные оценки положительные) коэффициентами вариации:

$$v_j = \frac{\sigma_j}{b_j}, \quad j = \overline{1, m}, \quad (11.31)$$

где $\sigma_j = +\sqrt{\sigma_j^2}$ — среднее квадратическое отклонение.

При анализе согласованности мнений экспертов обычно считают, что она удовлетворительная, если все $v_j \leq 0,3$, и хорошая, если все $v_j < 0,2$.

Если каждого из экспертов просят дать непосредственную оценку некоторой характеристики x элемента — точечную оценку (например, указать вероятность наступления некоторого события, срок его возможного наступления, оценить ожидаемые затраты или другие неопределенные факторы), то он должен поставить в соответствие каждому элементу точку на непрерывной числовой оси. В результате обработки таких оценок могут быть получены показатели среднего результата:

$$\bar{x} = \frac{1}{n} \sum_{l=1}^n x_l, \quad (11.32)$$

дисперсии:

$$\sigma_x^2 = \frac{1}{n-1} \sum_{l=1}^n (x_l - \bar{x})^2 \quad (11.33)$$

и вариации:

$$v_x = \frac{\sigma_x}{\bar{x}}, \quad (11.34)$$

где x_l — точечная оценка характеристики x , данная l -м экспертом.

Статистическая оценка полученных результатов позволяет определить не только интервал достоверных значений оцениваемой характеристики, но и противоречивость мнения конкретного эксперта. В первом случае используют методы интервального оценивания. Для симметричного закона распределения оценки

$$P(|\hat{x} - x^n| \leq \varepsilon) = q, \quad (11.35)$$

где x^n — неизвестное истинное значение x ; q — доверительная вероятность (обычно $q = 0,9 \dots 0,95$); ε — половина длины доверительного интервала.

В практике при числе экспертов в группе $n \geq 25 \div 30$ распределение оценки $\bar{x}$ обычно полагают нормальным. В противном случае величину ε в выражении (11.35) определяют с ис-

пользованием распределения Стюдента (с $n - 1$ степенью свободы):

$$\varepsilon = t_q \frac{\sigma_x}{\sqrt{n}},$$

где t_q — корень уравнения $q = 2 \int_0^{t_q} S(t, n-1) dt$; $S(t, n-1)$ —

плотность распределения Стюдента.

Наиболее удаленные от $\bar{x}$ оценки x_l проверяют на противоречивость (аномальность) следующим образом. Если проверяется оценка x_l , превышающая среднее значение $\bar{x}$, то по таблицам распределения Стюдента вычисляется вероятность

$$\alpha = 1 - \int_{-\infty}^{t_\alpha} S(t, n-1) dt, \quad (11.36)$$

где

$$t_\alpha = \frac{x_l - \bar{x}}{\sigma_x}.$$

Если же проверяемая оценка существенно меньше $\bar{x}$, то вероятность

$$\alpha = \int_{-\infty}^{t_\alpha} S(t, n-1) dt. \quad (11.37)$$

Оценка x_l l -го эксперта считается противоречивой при значении α , меньшем некоторого заданного значения $\alpha_{\text{тр}}$ (обычно $\alpha_{\text{тр}} = 0,05 \div 0,1$).

К суждению о противоречивости мнения «оригинального» эксперта при таком подходе следует подходить осторожно, дополняя вывод по соотношениям (11.36) и (11.37) логическим анализом, так как противоречивость мнения l -го эксперта может объясняться лучшим пониманием проблемы. Процедура формирования группового мнения в этом случае выглядит следующим образом:

определяют точечную групповую оценку $\bar{x}$; вычисляют дисперсию σ_x^2 и среднее квадратическое отклонение σ_x мнений экспертов; оценивают противоречивость «оригинальных» мнений с помощью (11.36), (11.37) и осу-

ществляют логический анализ; при противоречивых мнениях результаты опроса оформляют в виде точечной $\bar{x}$ и интервальной ($\bar{x} - \varepsilon$, $\bar{x} + \varepsilon$) оценок; при наличии «оригинальных» (противоречивых) мнений проводится второй тур экспертизы с коллективным обсуждением результатов предыдущего тура.

Более удобной формой выражения суждений экспертов являются не точечные, а диапазонные оценки. В этом случае эксперт указывает минимальное $x_{l\min}$ и максимальное $x_{l\max}$ значения, между которыми, по его мнению, находится истинное значение характеристики x^* . Полагая распределение оцениваемой характеристики x_l внутри указанного диапазона равномерным, вычисляют точечную оценку l -го эксперта

$$\bar{x}_l = 0,5 (x_{l\max} + x_{l\min}), \quad (11.38)$$

а затем — групповую оценку

$$\bar{x} = \frac{1}{n} \sum_{l=1}^n \bar{x}_l. \quad (11.39)$$

Дисперсия результатов экспертизы и коэффициент вариации определяют по (11.33) и (11.34) соответственно с подстановкой значения $\bar{x}_l$ вместо x_l .

Для уточнения распределения оценок x_l внутри диапазона $[x_{l\min}, x_{l\max}]$ эксперты должны дополнительно указать наиболее вероятную, по их мнению, оценку $x_{l\text{med}}$ характеристики x .

Хорошей аппроксимацией в этом случае является β -распределение [38]. По величинам $x_{l\min}$, $x_{l\text{med}}$, $x_{l\max}$ оценивают среднее значение $\bar{x}_l$ и дисперсию $\sigma_{x_l}^2$:

$$\bar{x}_l = \frac{x_{l\min} + 2x_{l\text{med}} + x_{l\max}}{4}, \quad l = \overline{1, n}; \quad (11.40)$$

$$\sigma_{x_l}^2 = \left(\frac{x_{l\max} - x_{l\min}}{4} \right)^2, \quad l = \overline{1, n}. \quad (11.41)$$

ОБРАБОТКА И АНАЛИЗ ПОПАРНЫХ СРАВНЕНИЙ

При выражении предпочтений на множестве предъявления эксперты достаточно просто решают эту задачу методом попарного сравнения. Задачи обработки и анализа в этом случае заключаются в получении группового суждения о элементах множества предъявления по результатам индивидуальных попарных сравнений. В отличие от выражения предпочтений методами ранжирования и балльного оценивания для попарного сравнения не требуется соблюдение условия транзитивности, что снимает психологические трудности восприятия экспертами различных характеристик всего множества элементов. Согласование мнений экспертов и противоречивость суждений отдельного эксперта являются первоочередными задачами обработки и анализа.

В результате попарного сравнения элементов $d_j \in D$, $j = \overline{1, m}$ каждый из n экспертов проводит оценку всех из $0,5m(m-1)$ пар элементов (d_j, d_k) по следующему правилу:

$$e_{jk}^{(l)} = \begin{cases} 1, & \text{если } d_j \succ d_k; \\ 0,5, & \text{если } d_j \sim d_k; \\ 0, & \text{если } d_j \prec d_k; \\ j, k = \overline{1, m}; l = \overline{1, n}. \end{cases}$$

Числа $e_{jk}^{(l)}$ представляют из себя элементы матрицы $E^{(l)}$ попарных сравнений каждого из n экспертов. Полученные матрицы $E^{(l)}$ осредняются с учетом мнений всех экспертов:

$$e_{jk}^r = \frac{1}{n} \sum_{l=1}^n e_{jk}^{(l)}. \quad (11.42)$$

Числа e_{jk}^r являются элементами матрицы E^r размером $m \times m$, причем $e_{jk}^r + e_{kj}^r = 1$. На основе матрицы E^r можно получить групповую ранжировку множества предъявления, определить весовые коэффициенты (коэффициенты важности) элементов $d_j \in D$ и оценить согласованность мнений

экспертов. Оценку согласованности мнений экспертов можно проводить с использованием коэффициентов вариации, вычисляемых по (11.33) — (11.34) с подстановкой вместо x_i величины $e_{jk}^{(l)}$, вместо $\bar{x}$ — величины e_{jk}^r . В общем случае противоречивость суждений отдельных экспертов оценивают с использованием (11.35) — (11.37). После выполнения указанных этапов построение групповой ранжировки и определение коэффициентов важности α_j , $j = \overline{1, m}$ осуществляют следующим образом.

Решается уравнение вида

$$|E^r - \eta I| = 0 \quad (11.43)$$

с целью отыскания собственного вектора α матрицы E^r . В (11.43) η есть собственное число матрицы E^r , а I — единичная диагональная матрица.

Поскольку в общем случае матрица E^r не удовлетворяет требованию транзитивности, то решение (11.43) осуществляется с помощью следующего итерационного алгоритма.

Алгоритм 2. 1. Задать критерий останова и требуемое количество итераций вычисления α_j , $j = \overline{1, m}$.

2. Положить $t = 0$ и все $\alpha_j = \frac{1}{m}$, $j = \overline{1, m}$.

3. Положить $t = t + 1$.

4. Вычислить

$$\alpha_j^{(t)} = \frac{\sum_{i=1}^m e_{ij}^r \alpha_i^{(t-1)}}{\sum_{i=1}^m \sum_{j=1}^m e_{ij}^r \alpha_i^{(t-1)}}.$$

5. Проверить условие $t \geq t^{\text{TP}}$: да — окончить вычисление; нет — перейти к п. 3.

Практика показывает, что $t^{\text{TP}} = 3 \div 5$. Компоненты собственного вектора матрицы E^r являются весовыми коэффициентами элементов d_j , измеренные

в шкале отношений, причем $\sum_{i=1}^m \alpha_j =$

$= 1$. Результирующая ранжировка определяется по правилу:

$$\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_m \Rightarrow r_1 \leq r_2 \leq \dots \leq r_m. \quad (11.44)$$

Если в экспертном опросе участвует сравнительно большое число специалистов ($n \geq 10$), то обработка и анализ попарных суждений проводится комбинированным методом, предложенным Л. Терстоуном. Этот метод предполагает предварительное ранжирование элементов, а затем получение групповой ранжировки методом попарных сравнений. В любом случае рабочей является матрица $P = \|p_{ij}\|$, $i, j = \overline{1, m}$, элементы которой суть доли экспертов, считающих элемент d_i предпочтительней элемента d_j . Элементы матрицы P удовлетворяют условию $p_{ij} + p_{ji} = 1$.

Строят матрицу $Z = \|z_{ij}\|$, элементы которой определяют с использованием табличной функции нормального распределения:

$$z_{ij} = F_T^{-1}(p_{ij}),$$

где $F_T^{-1}(p_{ij})$ — функция, обратная функции $F_T(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$.

Рассчитывают

$$z_j = \frac{1}{m} \sum_{i=1}^m z_{ij}, \quad j = \overline{1, m}$$

и определяют вероятности

$$p_j = F_T(z_j), \quad j = \overline{1, m}.$$

Полученные вероятности нормируют, в результате чего получают коэффициенты важности

$$\alpha_j = \frac{p_j}{\sum_{i=1}^m p_i}, \quad j = \overline{1, m}.$$

Групповую ранжировку получают по (11.44). Проверку непротиворечивости индивидуальных суждений экспертов осуществляют принятыми в статистике методами. Обычно рассчиты-

вают теоретическую долю случаев, когда элемент d_i предпочтительней элемента d_j , и сравнивают ее с фактической долей p_{ij} . Для этого определяют разности

$$\Delta z_{ij} = z_i - z_j, \quad i, j = \overline{1, m}, \quad i \neq j$$

и вычисляют

$$\tilde{p}_{ij} = F_T(\Delta z_{ij}), \quad i, j = \overline{1, m}, \quad i \neq j.$$

Рассчитывают отклонения

$$\Delta p_{ij} = \tilde{p}_{ij} - p_{ij}, \quad i, j = \overline{1, m}, \quad i \neq j.$$

Вычисляют среднее по модулю отклонение

$$\bar{\Delta} = \frac{2}{m(m-1)} \sum_{i \neq j} |\Delta p_{ij}|$$

и сравнивают с модулем максимального из Δp_{ij} отклонений. Если $|\Delta p_{ij}^{\max}| < 3\bar{\Delta}$, то полученные оценки непротиворечивы.

ОПРЕДЕЛЕНИЕ КОЭФФИЦИЕНТОВ ОТНОСИТЕЛЬНОЙ ВАЖНОСТИ

Существуют достаточно простые оперативные методы определения коэффициентов относительной важности.

Метод непосредственной численной оценки (M1). Каждому l -му эксперту предлагается непосредственно указать коэффициенты важности $\alpha_j^{(l)}$ элементов d_j , удовлетворяющих условию нормировки:

$$\sum_{j=1}^m \alpha_j^{(l)} = 1, \quad \alpha_j \geq 0, \quad j = \overline{1, m}.$$

За групповые оценки α_j в этом случае принимают средние значения

$$\alpha_j = \frac{1}{n} \sum_{l=1}^n \alpha_j^{(l)}. \quad (11.45)$$

Метод балльного оценивания (M2). В заданной балльной шкале каждый эксперт должен оценить элемент $d_j \in D$. Очень часто выбирают шкалу 0—100; при этом балл 100 соответствует

самому предпочтительному элементу. За групповые оценки принимают числа

$$\alpha_j = \frac{\sum_{l=1}^n b_{lj}}{\sum_{l=1}^n \sum_{j=1}^m b_{lj}}, \quad (11.46)$$

где b_{lj} — балл, присвоенный j -му элементу l -м экспертом.

Метод относительных частот рангов (МЗ). Экспертам предлагается строго упорядочить элементы множества предъявления. Элементам упорядоченного множества присваивают числа (обратные ранги), указывающие число элементов, доминируемых данным. Таким образом, наилучший элемент будет иметь ранг $m - 1$ (он лучше всех, кроме самого себя), а наихудший — ранг 0. Коэффициенты важности

$$\alpha_j = \frac{\sum_{l=1}^n r_j^l}{\sum_{l=1}^n \sum_{j=1}^m r_j^l}, \quad (11.47)$$

где r_j^l — ранг, присвоенный j -му элементу l -м экспертом.

Метод попарного сравнения с градациями (М4). Экспертам предлагается произвести попарное сравнение элементов d_i, d_j , руководствуясь заданной шкалой выражения предпочтения. Например, можно использовать шкалу со следующими градациями: 1 — элементы одинаковы по предпочтительности; 3 — имеются достаточные основания считать один элемент предпочтительнее другого; 5 — один элемент безусловно предпочтительнее другого. Если эксперт колеблется в оценке предпочтительности между указанными градациями, он ставит промежуточное целое число (2, 4). Матрица попарных сравнений в этом случае формируется следующим образом. Последовательно просматривая пары элементов (d_i, d_j) , соответствующие строкам i ($i = 1, m$) матрицы, эксперт выделяет и оценивает в заданной шкале только те пары, в которых элемент $d_i \geq d_j$. Значения

остальных элементов a_{ji} матрицы вычисляют формально по правилу:

$$a_{ji} = \frac{1}{a_{ij}},$$

где a_{ij} — оценка предпочтительности элемента d_i над элементом d_j по шкале с градациями 1, 2, 3, 4, 5. Это правило соответствует попарному выражению предпочтения как доли относительной интенсивности свойства.

Полученная матрица $A = \|a_{ij}\|$, $i, j = 1, m$ обрабатывается по алгоритму 2, в результате чего определяют коэффициенты α_j .

Метод последовательных сравнений (М5) предложен Черчменом и Акофом. Групповые коэффициенты важности определяют осреднением коэффициентов важности, получаемых в результате индивидуальной работы каждого эксперта. Индивидуальная экспертиза в этом случае состоит в следующем.

Алгоритм 3. 1. Упорядочить элементы $d_j \in D$, $j = 1, m$ в соответствии с убыванием их предпочтительности (пусть, например, $d_1 \succ d_2 \succ d_3 \succ \dots \succ d_m$).

2. Каждому элементу d_j приписать вес $\tilde{\alpha}_j$, причем $0 < \tilde{\alpha}_j < 1$, а $\sum_{j=1}^m \alpha_j$ не обязательно должна быть равна 1. Наиболее важному элементу d_1 соответствует $\tilde{\alpha}_1 = 1$.

3. Решить, оказывается ли элемент d_1 предпочтительнее всех остальных элементов $\cup d_j, j \neq 1$, вместе взятых¹:

а) если $d_1 \succ \cup d_j, j \neq 1$, то вес $\tilde{\alpha}_1$ корректируется так, чтобы выполнялось неравенство $\tilde{\alpha}_1 > \sum_{j \neq 1} \tilde{\alpha}_j$;

б) если $d_1 \sim \cup d_j, j \neq 1$, то $\tilde{\alpha}_1 = \sum_{j \neq 1} \tilde{\alpha}_j$;

¹ Суждение о предпочтительности над совокупностью остальных элементов делается с учетом диапазона ожидаемых значений характеристик рассматриваемых элементов.

в) если $d_1 \prec \bigcup_{j \neq 1} d_j$, то $\bar{\alpha}_1 < \sum_{j \neq 1} \bar{\alpha}_j$.

Далее сравнить d_1 с совокупностью $\bigcup_{j \neq 1} d_j \setminus d_m$ и выполнять п. 3 до тех пор, пока не будет справедливо условие а или б для совокупности оставшихся элементов $\bigcup_{(l+1)} d_j \setminus d_m \setminus d_{m-1} \setminus \dots \setminus d_{m-k}$.

4. Сравнить элемент d_2 с совокупностью элементов $\bigcup_{j \neq 1, 2} d_j$, п. 3 алгоритма.

5. Для элементов $d_3, d_4, \dots, d_{m-2}$ выполнить п. 4.

6. Определить коэффициенты важности элементов

$$\alpha_j^{(l)} = \frac{\bar{\alpha}_j}{\sum_{j=1}^m \bar{\alpha}_j}.$$

Если множество предъявления содержит более семи элементов, то непосредственное использование описанного метода вызовет серьезные затруднения из-за его громоздкости. Поэтому используют следующую модификацию метода.

Алгоритм 4. 1. Упорядочить элементы $d_j \in D$ множества предъявления. Пусть

$$d_1 \succ d_2 \succ d_3 \succ \dots \succ d_m.$$

2. Выбрать элемент d_k , номер которого k определяется с помощью датчика равномерно распределенных чисел $k =]r \cdot m + 1[$,

где $r \in (0, 1)$ — равномерно распределенное число, а $] \cdot [$ — целая часть числа.

3. Случайным образом разбить оставшееся множество элементов $d_j \in D$, $j \neq k$ на непересекающиеся подмножества так, чтобы каждое из них содержало не более шести элементов:

$$D_1 \cap D_2 \cap D_3 \cap \dots \cap D_{i_0} = \emptyset.$$

4. В каждое из подмножеств D_i , $i = \overline{1, i_0}$ включить элемент d_k , полученный в п. 2:

$$D_i = d_k \cup D_i, \quad i = \overline{1, i_0}.$$

5. Для каждого подмножества D_i , $i = \overline{1, i_0}$ применить алгоритм 3, приписав предварительно элементу d_k вес $\bar{\alpha}_k$ (например, 1, 10 или 100). Значения весов α_j остальных элементов подмножества корректируют, оставляя вес $\bar{\alpha}_k$ без изменения.

6. Сравнить веса $\bar{\alpha}_j$, $j = \overline{1, m}$ с предпочтениями, полученными в п. 1. Если получены непротиворечивые результаты, веса нормируют (см. п. 6 алгоритма 3). В противном случае выявленные противоречия сообщают эксперту и он корректирует значения $\bar{\alpha}_j$.

Графоаналитический метод (М6) применяют для определения коэффициентов относительной важности отдельных характеристик элементов и, в частности, показателей эффективности. Коэффициенты важности характеристик элементов используют для упорядочения самих элементов по предпочтительности. Например, для упорядочения стратегий по предпочтительности с помощью функции эффективности (11.1) последнюю часто аппроксимируют линейной сверткой вида

$$W_e(X) = \sum_{i=1}^{n_0} \alpha_i x_i, \quad (11.48)$$

где α_i — коэффициент важности i -го показателя эффективности W_i ; x_i — числовая оценка W_i ; n_0 — число частных показателей эффективности.

Коэффициенты важности α_i характеристик x_i определяют путем подсчета частот предпочтений следующим образом [41].

В пространстве оценок характеристик $X = X_1 \times X_2 \times X_3 \times \dots \times X_m$ ограничениями $X_i^{\min} \leq X_i \leq X_i^{\max}$ выделяется гиперкуб. Рассматривают всевозможные бинарные отношения R_{ij} на декартовом произведении характеристик $X_i \times X_j$, $i, j = \overline{1, n_0}$, $i < j$:

$$R_{ij} = X_i \times X_j / \{(X_i^{\min}, X_j^{\min}), (X_i^{\max}, X_j^{\max})\}, \quad (11.49)$$

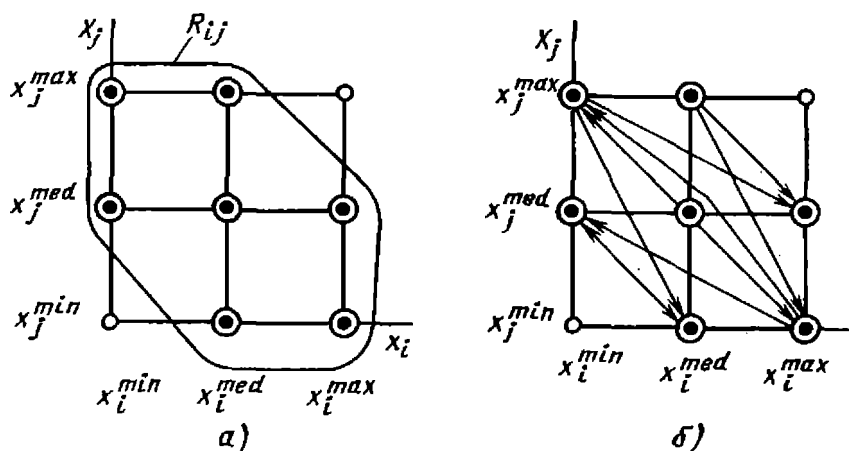


Рис. 5. Бинарные отношения R_{ij} (а) и граф предпочтений (б)

где $X_i = \{X_i^{min}, X_i^{med}, X_i^{max}\}$;

$$X_i^{med} = \frac{X_i^{min} + X_i^{max}}{2};$$

$X_i^{min}, X_i^{med}, X_i^{max}$ —

соответственно наихудшая, средняя и наилучшая оценки X_i характеристики X_i .

Таким образом бинарное отношение R_{ij} включает упорядоченные пары оценок характеристик (X_i, X_j) , имеющих по три градации $X_i^{min}, X_i^{med}, X_i^{max}$ без наилучшей и наихудшей пар. Бинарные отношения представляют графически в декартовой системе координат (рис. 5, а).

Коэффициенты важности вычисляют по следующему алгоритму.

Алгоритм 5. 1. Сравнить гипотетические элементы с характеристиками, принадлежащими R_{ij} , по предпочтительности и отразить суждение в виде стрелки, направленной от более предпочтительного элемента к менее предпочтительному¹ (см. рис. 5, б).

2. Подсчитать число $a_{ij}^{(l)}$ стрелок, направленных от характеристики X_i к характеристике X_j .

¹ Стрелки связывают только пары элементов, лежащих в направлении «северо-запад» — «юго-восток» рассматриваемой декартовой системы. Это обусловлено тем, что предпочтение между парами элементов в направлении «северо-восток» — «юго-запад» очевидно.

3. Определить частоту P_{ij} предпочтения характеристики X_i над характеристикой X_j :

$$P_{ij}^{(l)} = \frac{1}{9} a_{ij}^{(l)}; \quad i, j = \overline{1, n_0}, \quad i < j;$$

$$P_{ji} = 1 - P_{ij}. \quad (11.50)$$

4. Рассчитать коэффициенты относительной важности характеристик X_i :

$$\alpha_i^{(l)} = \frac{1}{n_0} \sum_{\substack{j=1 \\ (j \neq i)}}^{n_0} P_{ij}. \quad (11.51)$$

5. Вычислить групповые коэффициенты важности (если число n экспертов больше 1)

$$\alpha_i = \frac{1}{n} \sum_{l=1}^n \alpha_i^{(l)}. \quad (11.52)$$

Для примера, представленного на рис. 5, число a_{ij} стрелок, направленных от характеристики X_i к характеристике X_j , равно 3. Следовательно, $P_{ij} = \frac{1}{3}$.

Если используются шкалы оценок X_i с числом градаций k , большим трех, то в (11.50) вместо цифры 9 в знаменателе следует поставить величину $k^2(k-1)^2$.

Метод свертки (М7). Функция эффективности может быть задана в виде

некоторой свертки частных характеристик X_i (например, частных показателей эффективности W_i), взвешенных коэффициентами их важности. Если вид функции эффективности $W_e(X)$ определен (в простейшем случае в виде линейной свертки (11.48)), то, используя элементарные суждения относительно контрольных предъявлений, составляют соответствующие соотношения по правилу (11.1). В результате получают систему равенств и неравенств, решая которую определяют коэффициенты важности. Поскольку эта система соотношений может оказаться несовместной, то ее регуляризуют, подбирая такие коэффициенты α_i , чтобы обеспечить наименьшую степень ее несовместности.

Алгоритм 6. 1. В пространстве характеристик $X = X_1 \times X_2 \times X_3 \times X_{n_0}$ выделить гиперкуб, задаваемый системой неравенств $X_i^{\min} \leq X_i \leq X_i^{\max}$, $i = \overline{1, n_0}$.

2. Любым способом¹ отобрать из гиперкуба m_0 пар оценок $a, b \in X$, $c, d \in X$; $q, h \in X$, причем $m_0 \geq n_0$.

3. Сравнить полученные пары оценок по предпочтительности и составить систему соотношений по правилу:

$$\left. \begin{aligned} a \succ b &\Rightarrow W_e(a) > W_e(b); \\ c \sim d &\Rightarrow W_e(c) = W_e(d); \\ h \succ q &\Rightarrow W_e(h) > W_e(q), \end{aligned} \right\} \quad (11.53)$$

где $W_e(X) = \Psi(\alpha, X)$; $\Psi(\dots)$ — функция свертки.

4. Составить функцию невязки

$$f(\alpha) = \sum_{R_1} \min \{0, \Psi(\alpha, X) - \Psi(\alpha, Y)\} - \sum_{R_2} |\Psi(d, Z) - \Psi(\alpha, U)|, \quad (11.54)$$

где R_1 — множество пар, у которых $X \succ Y$; R_2 — $Z \sim U$.

¹ Обычно строят равномерную решетку либо с помощью датчика случайных чисел или ЛПГ-последовательностей генерируют необходимое число точек.

5. Решить задачу нелинейного программирования

$$\alpha^* : \max f(\alpha) \quad (11.55)$$

при ограничениях $\alpha \geq 0$, $\sum_{i=1}^{n_0} \alpha_i = 1$.

Если в экспертизе участвуют несколько экспертов, то коэффициенты важности, полученные для каждого из n экспертов по алгоритму 6, осредняют по (11.52).

При выборе конкретного метода определения коэффициентов относительной важности следует учитывать следующие обстоятельства: во-первых, ограничения по допустимому времени общения с экспертами, во-вторых, требуемую надежность получаемых оценок, в-третьих, наличие ЭВМ и математического обеспечения, позволяющих провести обработку результатов (сложность обработки).

По первому показателю эти методы можно упорядочить следующим образом:

$$M8 \sim M3 \succ M9 \succ M2 \sim M4 \succ M1 \succ M6 \succ M5 \succ M7,$$

где $M8$ — метод Терстоуна; $M9$ — метод попарного сравнения.

По второму показателю ряд предпочтений выглядит следующим образом:

$$M7 \succ M5 \succ M6 \succ M4 \succ M9 \succ M3 \succ M2 \sim M8 \succ M1.$$

По третьему показателю методы упорядочиваются так:

$$M1 \sim M2 \sim M3 \sim M6 \sim M8 \sim M4 \sim M9 \succ M5 \succ M7.$$

Методы $M3$, $M8$ используют только при групповой экспертизе, в то время как остальные можно применять и при индивидуальном экспертном опросе.

КОМПЛЕКСНЫЙ МЕТОД ВЫЯВЛЕНИЯ И ИЗМЕРЕНИЯ ПРЕДПОЧТЕНИЙ

Одним из наиболее распространенных методов выявления и измерения предпочтений, сочетающим процедуры

получения информации с эффективной обратной связью, позволяющей экспертам корректировать свои суждения, является метод Дельфи [6, 57]. Это метод групповой экспертной оценки, основанной на выявлении согласованного коллективного мнения при индивидуальном анкетном опросе экспертов в несколько туров (экспертизы по методу Дельфи чаще всего проводят в четыре тура). В каждом последующем туре предусматривается сообщение экспертам результатов предыдущего тура; в некоторых случаях экспертам предлагается обосновать свое мнение с учетом представленных ему оценок.

Метод Дельфи можно использовать только в том случае, если мнение эксперта может быть выражено в числовой форме и имеется достаточная информационная база для вынесения суждения экспертом в виде числовой оценки.

В первом туре после того, как экспертам разъяснена цель экспертизы и сообщены вопросы (персонально в виде анкеты, пояснительной записки), эксперты представляют свои ответы без аргументации. Вся экспертная информация поступает в группу управления. При ее обработке выделяют средние и крайние мнения — медиану, верхнюю и нижнюю квартили (значения оцениваемой характеристики, выше и ниже которых расположены 25 % численных значений оценок). Расстояние между квартилями в этом случае характеризует близость мнений экспертов. Экспертам сообщают эти сведения и проводят второй тур опроса. Иногда экспертам сообщают только квартили (считается, что это позволяет ослабить давление группового мнения на экспертов).

При втором туре опроса тех экспертов, чьи оценки оказались в крайних квартилях, просят представить обоснованные аргументы в пользу своих суждений. Подобные аргументы представляются анонимно и с ними знакомят остальных экспертов. Это позволяет всем экспертам учесть те обстоятельства, которые они по какой-либо причине пропустили или которыми пренебрегли. В результате этого эксперты могут пересмотреть свои суждения или подправить свои оценки. Скорректи-

рованная информация вновь поступает в группу управления экспертизы, реализуя обратную связь (см. рис. 4).

После обработки результатов второго тура скорректированные групповые (медиана) и «оригинальные» мнения (крайние квартили), а также всю аргументацию с сохранением анонимности сообщают экспертам и проводят очередной тур опроса.

В результате многотуровой процедуры опроса могут наблюдаться две ситуации: во-первых, от тура к туру растет согласованность мнений экспертов; во-вторых, — размежевание мнений с высокой согласованностью внутри подгрупп. Последнее может объясняться либо наличием представителей различных научных направлений, либо неодинаковой интерпретацией исходной информации специалистами различных профилей. В любом из этих случаев руководитель группы управления принимает решение о проведении дальнейшего опроса с целью подтверждения создавшегося положения или об окончании проведения экспертизы и составлении отчета по ней.

Таким образом, к основным особенностям метода Дельфи относятся: анонимность; полный отказ от личных контактов и коллективных обсуждений; обоснование ответов «оригинальных» экспертов по запросу группы управления; обратная связь, реализуемая в ходе многотурового опроса. При формировании группового мнения в каждом туре могут быть определены и в дальнейшем учитываться оценки коэффициентов компетентности экспертов. Это тем более важно, что предварительное определение коэффициентов компетентности по профессиональным показателям методом взаимной самооценки, результатам контрольных экспертиз и другими априорными методами не всегда надежно.

Основным недостатком метода Дельфи является то, что в процессе многотурового опроса, иногда достаточно разнесенного по времени, может сложиться лишь видимость высокой согласованности индивидуальных мнений вследствие стремления экспертов побыстрее отделаться от вопросов. Это приводит к поверхностному индиви-

дуальному анализу проблемы и стремлению к «групповому» мнению.

При реализации метода Дельфи комплексно используются процедуры выявления и измерения предпочтений. Существуют специальные комплексные методы, предназначенные для анализа сложных систем и процессов: планирование средств для фундаментальных исследований, управление научными и опытно-конструкторскими разработками, разработка сценариев развития при прогнозировании и др. К таким методам относятся: метод решающих матриц; метод прогнозного графа, метод взаимовлияний, методика ПАТТЕРН, КВЕСТ и др.

ФОРМИРОВАНИЕ И ФОРМАЛИЗАЦИЯ ЭКСПЕРТНОЙ ИНФОРМАЦИИ В ФОРМЕ НЕЧЕТКИХ СУЖДЕНИЙ

Формирование и формализация экспертной информации в форме нечетких суждений может проводиться как при индивидуальных, так и при групповых экспертизах. На этапе формирования экспертной группы необходимо подбирать экспертов, способных различать и оценивать градации свойств и характеристик элементов. При организации и проведении экспертизы (особенно при составлении анкет-опросников, вопросов интервью или бесед) следует учитывать особенности выражения и формализации нечетких характеристик элементов. Выделяют [32]:

квалификаторы — нечеткие категории, характеризующие признак элемента (например, «дешевый», «длинный», «перспективный» и т. п.);

квантификаторы — нечеткие категории, описывающие число элементов или повторяемость действий (например, «несколько этапов», «частая повторяемость» и т. п.);

модификаторы — нечеткие категории, уточняющие значение квалификаторов либо по интенсивности отдельной характеристики элемента (например, «очень», «приблизительно», «сильно» и т. п.), либо по степени сравнения с другим элементом или характеристикой (например, «более совершенный», «менее перспективный» и т. п.).

Для выражения градаций характеристик элементов модификаторы сочетаются как с квалификаторами, так и с квантификаторами (например, «очень дешевый», «очень частая повторяемость» и т. п.).

При проведении экспертного опроса суждения специалистов могут выражаться и представляться:

1) значением функции принадлежности $\mu(x)$ характеристики x ;

2) точкой на заданной шкале признака;

3) указанием типа функции принадлежности $\mu(x)$ из заданного их семейства на шкале характеристики;

4) в виде графика функции принадлежности;

5) в виде аналитической функции.

При реализации первого способа представления информации процедура построения функции принадлежности выглядит следующим образом:

1.1) определяют диапазон изменения характеристики и вводят численную шкалу;

1.2) вводят квалификатор;

1.3) каждому значению характеристики из заданной шкалы l -й эксперт ($l = \overline{1, n}$) ставит в соответствие определенную оценку в заданной балльной шкале (обычно 0, 1, 2, ..., 9) или значение функции принадлежности;

1.4) результаты опроса всех экспертов осредняют обычными методами.

Второй способ представления суждений обычно применяют в том случае, если квалификатор не имеет размерной шкалы, т. е. не может быть измерен в единицах физических величин. Функцию принадлежности в этом случае строят следующим образом:

2.1) вводят квалификатор и определяют его тип: антонимический («перспективно — неперспективно», «плохо — хорошо» и т. п.) и неантонимический, т. е. имеющий набор промежуточных значений;

для антонимических квалификаторов выполняют пп. 2.2—2.4;

2.2) вводят шкалу типа «сильно — средне — слабо — 0 — слабо — средне — сильно»;

2.3) l -й эксперт оценивает предъявляемый ему элемент по заданной харак-

геристике и ставит ему в соответствие оценку на введенной шкале;

2.4) осредняя результаты по п. 2.3, получают функции принадлежности для антонимов;

для неантонимических квалифика-
торов определяют:

2.5) положение границ, определяющих градации характеристики:

2.5.1) задают диапазон изменения характеристики;

2.5.2) экспертам предъявляют конкретный элемент;

2.5.3) l -й эксперт решает, по какую сторону от условного центра лежит характеристика элемента;

2.5.4) п. 2.5.3 повторяют для выбранной половины диапазона до тех пор, пока не будет определена граница градации;

2.6) внутри индивидуально установленных границ эксперт формирует функцию принадлежности (выполняется аналогично пп. 1.2, 1.3);

2.7) результаты опроса всех экспертов осредняют.

При выполнении п. 2.6 может быть проведен предварительный опрос экспертов с целью установления типа функции принадлежности (третий способ выражения суждений). Часто для этого семейство функций принадлежности задается в форме треугольника, трапеции и т. д. (табл. 7). В этом случае в п. 2.6 эксперт указывает характерные точки выбранного вида функции принадлежности для параметрического ее описания. Например, для функции принадлежности треугольного вида эксперт указывает три точки, из которых две крайние имеют нулевое значение функции принадлежности, а промежуточная — единичное значение.

Четвертый и пятый способы выражения суждений экспертов основаны на том, что каждый из экспертов способен графически или аналитически задать функцию принадлежности. Могут быть реализованы два подхода. Во-первых, из представленного группой управления семейства функций принадлежности эксперт может выбрать подходящую с его точки зрения и указать ха-

рактерные точки или параметры. Во-вторых, изобразить функцию принадлежности в указанном масштабе или дать ее аналитическое выражение. Если экспертиза групповая, то необходимо агрегировать индивидуальные функции принадлежности, установив ее осредненный тип и параметры. В любом случае необходимо задаться метрикой и критерием согласования индивидуальных суждений. В качестве метрики можно использовать либо сумму модулей, либо сумму квадратов отклонений индивидуальных функций принадлежности от искомой. В качестве критерия используется минимум одной из указанных сумм. Формально эта задача сводится к следующей задаче математического программирования:

$$\mu(x) : \min_{\mu(x) \in \{\mu_i(x)\}} \sum_{i=1}^n x(\mu(x), \mu_i(x)), \quad (11.56)$$

где x — метрика (либо $|\mu(x) - \mu_i(x)|$, либо $(\mu(x) - \mu_i(x))^2$); $\{\mu_i(x)\}$ — семейство типов функции принадлежности; n — число экспертов.

В табл. 7 представлено семейство функций принадлежности, широко используемых в практике [32]. С использованием этих функций можно построить нечеткие отношения $\tilde{R}$. При этом следует заменить значение x одномерной функции принадлежности значением расстояния между элементами упорядоченных пар, выраженного в принятой метрике. Обычно используют метрику типа $x_r = \sqrt{x_1^2 + x_2^2}$ либо типа $x_{mod} = |x_1 - x_2|$, где x_1, x_2 — характеристики элементов d_1, d_2 . Например, для функций принадлежности 1—14 отношение $\tilde{R}$ для характеристик x_1, x_2 можно трактовать как «величины x_1 и x_2 примерно равны» (так как величины x_r и x_{mod} «малы»). Для функции принадлежности 4 функция принадлежности нечеткого отно-

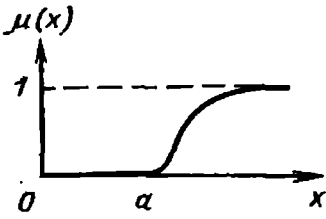
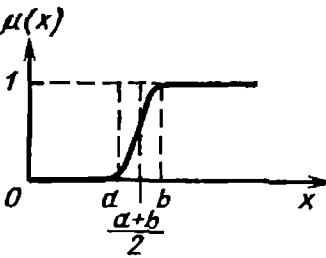
7. Типовые функции принадлежности суждения

№ по пор.	График	Функция
В е л и ч и н а х м а л а		
1		$\mu(x) = \begin{cases} 1, & 0 \leq x \leq a, \\ 0, & x > a \end{cases}$
2		$\mu(x) = e^{-kx}, \quad k > 0$
3		$\mu(x) = e^{-kx^2}, \quad k > 0$
4		$\mu(x) = \begin{cases} 1, & \text{если } 0 \leq x \leq a_1, \\ \frac{a_2 - x}{a_2 - a_1}, & \text{если } a_1 \leq x \leq a_2, \\ 0, & \text{если } a_2 \leq x \end{cases}$
5		$\mu(x) = \begin{cases} 1 - ax^k, & \text{если } 0 \leq x \leq \frac{1}{\sqrt[k]{a}}, \\ 0, & \text{если } \frac{1}{\sqrt[k]{a}} \leq x \end{cases}$
6		$\mu(x) = \frac{1}{1 + kx^2}, \quad k > 1$

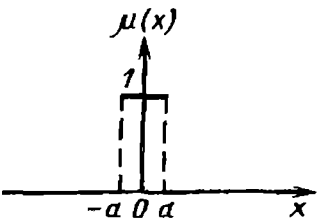
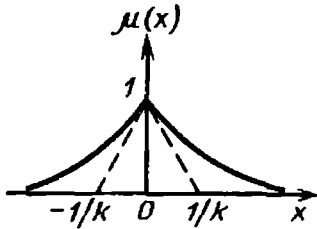
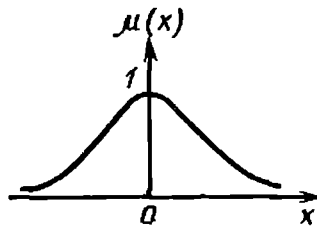
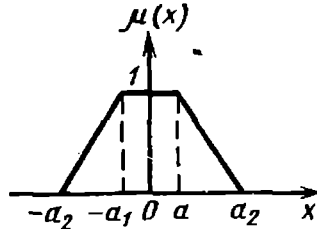
Продолжение табл. 7

№ по пор.	График	Функция
7		$\mu(x) = \begin{cases} 1, & \text{если } 0 \leq x \leq a, \\ \frac{1}{2} - \frac{1}{2} \sin \frac{\pi}{b-a} \left(x - \frac{a+b}{2} \right), & \text{если } a \leq x \leq b, \\ 0, & \text{если } b \leq x \end{cases}$
Величина x большая		
8		$\mu(x) = \begin{cases} 0, & \text{если } 0 \leq x < a, \\ 1, & \text{если } a \leq x \end{cases}$
9		$\mu(x) = \begin{cases} 0, & \text{если } 0 \leq x \leq a, \\ 1 - e^{-k(x-a)}, & \text{если } a \leq x, k > 0 \end{cases}$
10		$\mu(x) = \begin{cases} 0, & \text{если } 0 \leq x \leq a, \\ 1 - e^{-k(x-a)^2}, & a \leq x, k > 0 \end{cases}$
11		$\mu(x) = \begin{cases} 0, & \text{если } 0 \leq x \leq a_1, \\ \frac{x - a_1}{a_2 - a_1}, & \text{если } a_1 \leq x \leq a_2, \\ 1, & \text{если } a_2 \leq x \end{cases}$
12		$\mu(x) = \begin{cases} 0, & \text{если } 0 \leq x \leq a, \\ a(x-a)^k + \frac{1}{k\sqrt{a}}, & \text{если } a \leq x \leq a + \frac{1}{k\sqrt{a}}, \\ 1, & \text{если } a + \frac{1}{k\sqrt{a}} \leq x \end{cases}$

Продолжение табл. 7

№ по пор.	График	Функция
13		$\mu(x) = \begin{cases} 0, & \text{если } 0 \leq x \leq a, \\ \frac{k(x-a)^2}{1+k(x-a)^2}, & \text{если } a \leq x < \infty \end{cases}$
14		$\mu(x) = \begin{cases} 0, & \text{если } 0 \leq x \leq a, \\ \frac{1}{2} + \frac{1}{2} \sin \frac{\pi}{b-a} \left(x - \frac{a+b}{2} \right), & \text{если } a \leq x \leq b, \\ 1, & \text{если } a \leq x \end{cases}$

Величина | x | мала

15		$\mu(x) = \begin{cases} 0, & \text{если } -\infty < x < a, \\ 1, & \text{если } -a \leq x \leq a, \\ 0, & \text{если } a \leq x \end{cases}$
16		$\mu(x) = \begin{cases} e^{kx}, & \text{если } -\infty < x \leq 0, \\ e^{-kx}, & \text{если } 0 \leq x < \infty, \end{cases} \quad k > 1$
17		$\mu(x) = e^{-kx^2}$
18		$\mu(x) = \begin{cases} 0, & \text{если } -\infty < x \leq -a_2, \\ \frac{a_2 + x}{a_2 - a_1}, & \text{если } -a_2 \leq x \leq -a_1, \\ 1, & \text{если } -a_1 \leq x \leq a_1, \\ \frac{a_2 - x}{a_2 - a_1}, & \text{если } a_1 \leq x \leq a_2, \\ 0, & \text{если } a_2 \leq x < \infty \end{cases}$

Продолжение табл. 7

№ по пор.	График	Функция
19		$\mu(x) = \begin{cases} 0, & \text{если } -\infty < x \leq -\frac{1}{\sqrt[k]{a}}, \\ 1 - a(-x)^k, & \text{если } -\frac{1}{\sqrt[k]{a}} \leq x \leq 0, \\ 1 - a(x)^k, & \text{если } 0 \leq x \leq \frac{1}{\sqrt[k]{a}}, \\ 0, & \text{если } \frac{1}{\sqrt[k]{a}} \leq x < \infty \end{cases}$
20		$\mu(x) = \frac{1}{1 + kx^2}, \text{ если } k > 1$
21		$\mu(x) = \begin{cases} 0, & \text{если } -\infty < x \leq -b, \\ \frac{1}{2} + \frac{1}{2} \sin \frac{\pi}{b-a} \left(x + \frac{a+b}{2} \right), & \text{если } -b \leq x \leq -a, \\ 1, & \text{если } -a \leq x \leq a, \\ \frac{1}{2} - \frac{1}{2} \sin \frac{\pi}{b-a} \left(x - \frac{a+b}{2} \right), & \text{если } a \leq x \leq b, \\ 0, & \text{если } b \leq x < \infty \end{cases}$
Величина x большая		
22		$\mu(x) = \begin{cases} 1, & \text{если } -\infty < x < -a, \\ 0, & \text{если } -a < x < a, \\ 1, & \text{если } a < x < \infty \end{cases}$
23		$\mu(x) = \begin{cases} 1 - e^{kx}, & \text{если } -\infty < x \leq 0, \\ 1 - e^{-kx}, & \text{если } 0 \leq x < \infty, k > 1 \end{cases}$

Продолжение табл. 7

№ по пор.	График	Функция
24		$\mu(x) = 1 - e^{-kx^2}, \text{ если } k > 1$
25		$\mu(x) = \begin{cases} 1, & \text{если } -\infty < x < -a_2, \\ \frac{x + a_1}{a_2 - a_1}, & \text{если } -a_2 \leq x \leq -a_1, \\ 0, & \text{если } -a_1 \leq x \leq a_1, \\ \frac{x - a_1}{a_2 - a_1}, & \text{если } a_1 \leq x \leq a_2, \\ 1, & \text{если } a_2 \leq x < \infty \end{cases}$
26		$\mu(x) = \begin{cases} 1, & \text{если } -\infty < x < -\frac{1}{\sqrt[k]{a}}, \\ a(-x)^k, & \text{если } -\frac{1}{\sqrt[k]{a}} \leq x \leq 0, \\ ax^k, & \text{если } 0 \leq x \leq \frac{1}{\sqrt[k]{a}}, \\ 1, & \text{если } \frac{1}{\sqrt[k]{a}} \leq x < \infty \end{cases}$
27		$\mu(x) = \frac{kx^2}{1 + kx^2}, \text{ } k > 1$
28		$\mu(x) = \begin{cases} 1, & \text{если } -\infty < x \leq -b, \\ \frac{1}{2} - \frac{1}{2} \sin \frac{\pi}{b-a} \left(x + \frac{a+b}{2} \right), & \text{если } -b \leq x \leq -a, \\ 0, & \text{если } -a \leq x \leq a, \\ \frac{1}{2} + \frac{1}{2} \sin \frac{\pi}{b-a} \left(x - \frac{a+b}{2} \right), & \text{если } a \leq x \leq b, \\ 1, & \text{если } b \leq x \leq \infty \end{cases}$

шения «примерно равны» записывается следующим образом:

$$\mu_{\tilde{R}}(x_1, x_2) = \begin{cases} 1, & \text{если } 0 \leq \sqrt{x_1^2 + x_2^2} \leq a_1; \\ \frac{a_2 - \sqrt{x_1^2 + x_2^2}}{a_2 - a_1}, & \text{если } a_1 < \sqrt{x_1^2 + x_2^2} \leq a_2; \\ 0, & \text{если } \sqrt{x_1^2 + x_2^2} > a_2, \end{cases}$$

или

$$\mu_{\tilde{R}}(x_1, x_2) = \begin{cases} 1, & \text{если } 0 < |x_1 - x_2| \leq a_1; \\ \frac{a_2 - |x_1 - x_2|}{a_2 - a_1}, & \text{если } a_1 < |x_1 - x_2| \leq a_2; \\ 0, & \text{если } |x_1 - x_2| > a_2. \end{cases}$$

Большинство постановок задач принятия решения, вытекающих из модели проблемной ситуации (см. гл. 4), связано с выявлением и измерением предпочтений. Формальных методов решения большинства указанных задач не существует. Поэтому знание основных свойств предпочтений (способов их выражения, обработки анализа) позволяет ЛПР и исследователям правильно оценить возможности, достоинства и недостатки каждого метода с целью выбора такого, который окажется приемлемым для решения конкретной задачи. Рассмотренные методы выявления и измерения предпочтений являются теми элементами, с помощью которых можно сформировать различные комплексные процедуры обоснования решений, обеспечивающие компромисс между возможностями ЛПР и исследователей (время, ресурсы, уровень подготовленности) и их потребностями. Выявление и измерение предпочтений является в этом случае важным этапом при выработке решений в условиях как определенности, так и неопределенности различных типов.

Приложение

ЭЛЕМЕНТЫ ТЕОРИИ НЕЧЕТКИХ МНОЖЕСТВ

При анализе универсального множества D элементов произвольной природы ЛПР или исследователь выделяют некоторые подмножества $A \subseteq D$, элементы которых объединены некоторым общим для них свойством. Если это свойство качественного характера, описываемое лишь вербально (словесными оценками), то говорят, что на элементах подмножества A задана лингвистическая переменная.

Определение. Нечеткое подмножество A универсального множества D характеризуется функцией принадлежности $\mu_A : D \rightarrow [0, 1]$, которая ставит в соответствие каждому элементу $d \in D$ число $\mu_A(d)$ из отрезка $[0, 1]$, характеризующее степень принадлежности элемента d подмножеству A [24].

Обычные множества являются частными случаями нечетких множеств, так как для них функция принадлежности релейна, т. е. принимает значения:

$$\mu_A(d) = \begin{cases} 1, & \text{если } d \in A; \\ 0, & \text{если } d \notin A. \end{cases}$$

Нечеткое множество полностью описывается своей функцией принадлежности и поэтому ее обычно используют как обозначение нечеткого множества.

Нечеткое множество называется *пустым*, если его функция принадлежности равна нулю для всех элементов $d \in D$, т. е.

$$\mu_{\emptyset}(d) = 0, \quad \forall d \in D.$$

Носителем нечеткого множества $\mu_A(d)$ называется обычное множество, для которого функция принадлежности отлична от нуля.

Операции над нечеткими множествами определяются различными способами. Приведем лишь те из них, которыми наиболее часто пользуются в практике принятия решений.

Пусть A и B — нечеткие множества в D с функциями принадлежности $\mu_A(d)$ и $\mu_B(d)$.

1. Совпадение (эквивалентность)

$$A = B \Leftrightarrow \mu_A(d) \equiv \mu_B(d), \forall d \in D.$$

2. Включение

$$A \subseteq B \Leftrightarrow \mu_A(d) \leq \mu_B(d), \forall d \in D.$$

3. Объединение

$$A \cup B \Leftrightarrow \mu_{A \cup B}(d) =$$

$$= \max \{ \mu_A(d), \mu_B(d) \}, \forall d \in D.$$

4. Дополнение

$$\bar{A} \Leftrightarrow \mu_{\bar{A}}(d) = 1 - \mu_A(d), \forall d \in D.$$

5. Пересечение

$$A \cap B \Leftrightarrow \mu_{A \cap B}(d) =$$

$$= \min \{ \mu_A(d), \mu_B(d) \}, \forall d \in D.$$

6. Разность

$$A \setminus B \Leftrightarrow \mu_{A \setminus B}(d) =$$

$$= \begin{cases} \mu_A(d) - \mu_B(d), \\ \text{если } \mu_A(d) > \mu_B(d); \\ 0, \text{ если } \mu_A(d) \leq \mu_B(d), \end{cases} \forall d \in D.$$

7. Умножение A на число $\alpha \in [0, 1]$:

$$\alpha A \Leftrightarrow \mu_{\alpha A}(d) = \alpha \mu_A(d), \forall d \in D.$$

8. Возведение A в степень $\alpha > 0$:

$$A^\alpha \Leftrightarrow \mu_{A^\alpha}(d) = (\mu_A(d))^\alpha, \forall d \in D.$$

9. Концентрирование

$$\text{CON}(A) = A^2.$$

10. Растяжение

$$\text{DIL}(A) = A^{0.5}.$$

Множеством уровня α нечеткого множества A называется обычное множество $A_\alpha \in D$, для которого

$$\mu_{A_\alpha}(d) = \begin{cases} 1, \text{ если } \mu_A(d) \geq \alpha; \\ 0, \text{ если } \mu_A(d) < \alpha, \\ \forall d \in D. \end{cases}$$

Глава 12. Выработка решений в условиях определенности

1. ЗАДАЧИ ПРИНЯТИЯ РЕШЕНИЯ ПО СКАЛЯРНОМУ ПОКАЗАТЕЛЮ

Если установлено, что операция проводится в условиях определенности ($\Lambda \equiv \Lambda_F$), а исход G операции может быть описан адекватно единственным показателем эффективности, то процесс принятия решения сводится к установлению критерия эффективности K на основе цели операции A_0 и системы предпочтений $\mathcal{P}$ ЛПР и выбору наилучшего в соответствии с установленным критерием подмножества стратегий $U^* \subset U$. Среди эквивалентных по критерию K стратегий U^* выбор единственной стратегии $u^* \in U^*$ осуществляется либо произвольно, либо с учетом дополнительной информации о системе предпочтений ЛПР, т. е. можно ввести дополнительный показатель, характеризующий исход G , и провести аналогичный анализ на множестве стратегий U^* . Этим реализуется итерационная схема процесса принятия решения (см. рис. 4 гл. 10) и вложенность моделей предпочтений, описываемая (10.4).

Более того, многие задачи принятия решения по векторному показателю и (или) задачи в условиях неопределенности могут быть формально сведены к схеме задачи принятия решения по скалярному показателю в условиях определенности или декомпозированы на несколько таких задач.

ОБЩАЯ ПОСТАНОВКА ЗАДАЧИ

Постановка задачи принятия решения является результатом содержательного анализа модели проблемной ситуации (4.1), при котором последо-

вательно рассматриваются задачи, изложенные на с. 54. Если результатом решения задачи (4.5) является множество $\Lambda \equiv \Lambda_F$ и получены сведения θ_Λ о значениях характеристик X_Λ детерминированных факторов Λ_F , то задача относится к классу задач принятия решения в условиях определенности. К числу определенных относятся не только факторы, характеристики которых постоянны. В общем случае отнесение того или иного фактора к классу определенных основывается на представлениях ЛПР о влиянии степени разброса характеристик X_Λ фактора на степень вариации конечного результата Y операции. Если ЛПР считает, что этот разброс существенно не влияет на конечный результат Y , то такой фактор рассматривается детерминированным, а в качестве его характеристики часто используют среднее значение.

В зависимости от содержания обратной задачи исследования эффективности ее разделяют на задачу инженерного синтеза — установление тактико-технических характеристик системы (конструктивных параметров) и задачу оперативного синтеза — определение способа проведения операции. В этом случае результатом решения задачи формирования исходного множества стратегий (4.6) для задачи инженерного синтеза являются диапазоны возможных значений технических (конструктивных) параметров, удовлетворяющих целевым и физическим ограничениям, а для задач оперативного синтеза — множество способов использования активных средств и характеристики способов.

Так как операция проводится в условиях определенности, то задача моде-

лирования исходов операции (4.7) заключается в установлении однозначного функционального преобразования:

$$H: U \rightarrow Y, \quad (12.1)$$

а задача выбора показателя эффективности [см. (4.8)] сводится к установлению функции соответствия ρ :

$$W = \rho(y, y^{TP}), \quad (12.2)$$

где y, y^{TP} — детерминированные скалярные характеристики исхода G и требуемого результата операции.

В этом случае функция соответствия ρ задается обычным способом:

$$\rho(y, y^{TP}) = \begin{cases} 1, & \text{если } y \geq y^{TP}; \\ 0, & \text{если } y < y^{TP}, \end{cases} \quad (12.3)$$

если цель операции A_0 состоит в достижении требуемого результата, и

$$\rho(y, y^{TP}) = y, \quad (12.4)$$

если цель операции заключается в получении наибольшего (наименьшего) эффекта, который затем сравнивается с требуемым уровнем W^{TP} .

Используя (12.3), сначала можно найти допустимое применительно к достижению цели множество стратегий, а затем, используя (12.4), найти наилучшие стратегии. Или с помощью (12.4) найти наилучшие стратегии, а затем проверить их применительно к достижению цели по функции соответствия (12.3). Выбор конкретного из двух рассмотренных способов определяется тем, какая из задач проще: построить оптимизационную модель вида

$$H_{opt}: U \rightarrow (U^*, W_{opt}), \quad (12.5)$$

а затем сравнить величины W_{opt} с W^{TP} или вначале «сузить» множество U до подмножества U^{sat} , а затем выделить на нем подмножество U^* наилучших стратегий.

Таким образом, решая задачу выбора модели H и показателя эффективности W , исследователь всегда должен ориентироваться на достижение цели операции, а следовательно, задаваться определенным видом критерия эффективности. Именно поэтому задача мо-

делирования цели операции (4.8) решается комплексно с учетом предпочтений ЛПР. Систему предпочтений $\mathcal{P}$ ЛПР устанавливают либо на значениях результатов операции (характеристика y исхода G), либо на значениях показателя эффективности W в зависимости от используемой функции ρ . Если предпочтения ЛПР меняются скачком, т. е. либо значения $y(W)$ его полностью удовлетворяют, либо не удовлетворяют совсем, то формальной моделью его системы предпочтений служит критерий пригодности: любая стратегия, приводящая к желаемому значению результата (показателя эффективности), считается пригодной.

Если ЛПР указывает направление предпочтительного изменения значений результата (или показателя), то его система предпочтения формально задается критерием оптимальности, согласно которому наилучшей (оптимальной) считается та стратегия, для которой результат (показатель) достигает экстремального значения в указанном предпочтительном направлении. Например, если большие значения показателя эффективности предпочтительнее меньших, то критерием оптимальности будет критерий максимума результата

$$u^*: \max_{u \in U} W(u); \quad (12.6)$$

если наоборот — минимума результата

$$u^*: \min_{u \in U} W(u); \quad (12.7)$$

если направление предпочтений изменяется при переходе результата y через некоторое значение y_0 , то критерием оптимальности является минимум рассогласования

$$u^*: \min_{u \in U} W'(u), \quad (12.8)$$

где $W'(u) = |y(u) - y_0|$, либо $W'(u) = (y(u) - y_0)^2$.

Таким образом общая постановка задачи принятия решения по скалярному показателю в условиях определенности формально представляется следующим логическим высказыванием:

$$\langle \theta_\lambda, W, K, U; U^* \rangle. \quad (12.9)$$

Поскольку $W = \Psi(u)$ или $Y = H(u)$, то решение задачи (12.9) осуществляется путем сравнения показателей эффективности W или результатов операции u .

ФОРМАЛИЗАЦИЯ ЗАДАЧИ ВЫБОРА

Задача выбора, вытекающая из общей постановки задачи принятия решения (12.9), состоит в определении подмножества наилучших стратегий U^* из множества U :

$$\langle U; U^* \rangle. \quad (12.10)$$

Для формального разрешения этой задачи после того, как определен вид показателя эффективности W и задан критерий K , требуется функционально задать исходное множество стратегий U и оператор Ψ . Формализация задачи (12.10) основывается на использовании информации θ о механизме ситуации (связи между компонентами стратегии $u \in U$) и детерминированных факторах Λ_F, Y (параметрах модели). Поскольку каждая стратегия $u \in U$ в общем случае представляется совокупностью различных характеристик (конструктивных параметров, параметров структуры, последовательностью выполнения отдельных элементарных операций и т. п.), то каждую стратегию u будем описывать упорядоченным набором ее характеристик $X = (x_1, x_2, \dots, x_n)^T$. Детерминированные факторы Λ_F задают вид функционального преобразования Ψ :

$$W(X) = W(X, \Lambda_F) = \Psi(X), \quad (12.11)$$

позволяющего по значениям управляемых характеристик X однозначно определять показатель W при фиксированных условиях Λ_F .

Формальное представление допустимого множества U осуществляется заданием ограничений, определяемых содержательным анализом задачи принятия решения. Это означает, что модель Ψ формируется в рамках вполне определенных допущений (гипотез поведения системы) и ограничений, определяющих свободу выбора управляющих воздействий X . Допущения при построении модели (12.11) определяются системой более высокого

уровня по сравнению с рассматриваемой и адекватностью самой модели. В этом случае отображение Ψ представляет собой функциональную зависимость $\varphi(X)$, связывающую произвольные характеристики стратегий с показателями эффективности и определенную только на множествах допустимых стратегий, удовлетворяющих условиям $q_l(X) \leq 0$ и $h_k(X) = 0$, $l = \overline{1, l_1}$, $k = \overline{1, k_1}$. Кроме допущений в модели (12.11), при решении задачи (12.10) необходимо учесть ограничения на имеющиеся в распоряжении ЛПР ресурсы ($q_l(X) \leq 0$, $l = \overline{l_1 + 1, l_2}$, $h_k(X) = 0$, $k = \overline{k_1 + 1, k_2}$) и физические ограничения, исключающие «разрушение» системы, обеспечивающие гомеостазис ($q_l(X) \leq 0$, $l = \overline{l_2 + 1, l_0}$, $h_k(X) = 0$, $k = \overline{k_2 + 1, k_0}$). Таким образом, модель Ψ в задаче (12.10) формально представляется следующим образом:

$$\Psi : \varphi(X) \mid q(X) \leq 0, h(X) = 0, \quad (12.12)$$

где $q(X) = \{q_l(X), l = \overline{1, l_0}\}$; $h(X) = \{h_k(X); k = \overline{1, k_0}\}$.

С использованием функций (12.3) и (12.4) задачу (12.10) выбора по любому из критериев (пригодности или оптимальности) можно поставить как оптимизационную:

$$X^* : \text{extr } \varphi(X) \quad (12.13)$$

при ограничениях

$$\left. \begin{aligned} q(X) &\leq 0; \\ h(X) &= 0, \end{aligned} \right\} \quad (12.14)$$

где $\varphi(X)$ — целевая функция;

$$\text{extr} = \begin{cases} \text{max, если } \rho \text{ определяется по (12.3);} \\ \text{max или min (в зависимости от направления предпочтения), если } \rho \text{ определяется по (12.4);} \end{cases}$$

$X^* =$

$$= \begin{cases} (x_1, x_2, \dots, x_n)_{\text{SAT}}^T & \text{— характеристики пригодных стратегий, если используется функция (12.3);} \\ (x_1, x_2, \dots, x_n)_{\text{opt}}^T & \text{— характеристики оптимальных стратегий, если используется функция (12.4).} \end{cases}$$

Прежде чем переходить к решению задачи (12.13), (12.14), необходимо ее проанализировать. Если число независимых переменных $x_1, x_2, \dots, x_n$ не больше числа ограничений-равенств, то решением задачи может быть корень системы уравнений $h(X) = 0$. Полученное решение $\bar{X}$ системы уравнений проверяется на принадлежность области, задаваемой ограничениями-неравенствами $q(X) \leq 0$. В случае, если $q(\bar{X}) \leq 0$, то $X^* = \bar{X}$ есть решение задачи выбора. Если число независимых переменных больше числа ограничений-равенств, то необходимо убедиться в непустоте множества X_d , задаваемого ограничениями (12.14), и если оно непусто, то осуществить поиск экстремума (12.13).

В зависимости от способа представления модели Ψ и сложности ее составляющих $\varphi(X)$, $q(X)$, $h(X)$ решение задачи (12.13), (12.14) может быть осуществлено различными методами. Если Ψ представляет достаточно простую систему аналитических функций (размерность задачи невысока, функции непрерывны), то решение может быть получено классическими аналитическими методами — математического анализа, неопределенных множителей Лагранжа и др. (см. т. 2). Если число переменных в задаче не более двух (иногда трех), то решение можно отыскать графически. При большом числе переменных $x_i, i = 1, n$, большом числе уравнений связей и ограничений, а также сложном их виде используют

численные и экспериментальные методы, для которых требуется применять ЭВМ и специальное математическое обеспечение. К этой группе методов относятся процедуры и алгоритмы математического программирования. В более сложных случаях, когда получение даже одного результата $\varphi(x_1, x_2, \dots, x_n)$ сопряжено с большими затратами времени и ресурсов, используют методы вариантного анализа. Каждый из методов вариантного анализа состоит из серии перемежающихся этапов неформального и формального анализа, в ходе которых ЛПР и исследователи осуществляют последовательное сужение множества допустимых вариантов. На первом этапе ЛПР и эксперты эвристически формируют множество опорных стратегий, априори более предпочтительных по сравнению с остальными. В выбранных точках определяют значение показателя эффективности. Проанализировав полученную информацию, вновь формируют множество опорных точек в области предполагаемого экстремума. Процесс заканчивается, когда ЛПР удовлетворен полученным результатом.

ТИПЫ ЗАДАЧ ПРИНЯТИЯ РЕШЕНИЯ ПО СКАЛЯРНОМУ ПОКАЗАТЕЛЮ В УСЛОВИЯХ ОПРЕДЕЛЕННОСТИ

Выбор метода решения задачи (12.13), (12.14) представляет самостоятельную проблему. Существует множество методов, алгоритмов и процедур решения экстремальных задач. Каждый из них ориентирован на вполне определенный класс задач, а внутри каждого класса — на особенности целевой функции $\varphi(X)$ и ограничений $q(X)$, $h(X)$, способ их задания и др. Поэтому при определении возможных методов решения экстремальной задачи и выборе наиболее эффективного из них следует вначале определить тип задачи.

Типизация рассматриваемых задач принятия решений проводится по следующим направлениям.

Во-первых, устанавливается наличие или отсутствие зависимости характеристик X стратегии $u \in U$ от времени. Если $X = X(t)$, то задача относится к типу динамических, в противном

случае — статических. В динамических задачах целевая функция представляет собой характеристику эффективности процесса (в статических задачах — характеристику явления) и задается в виде функционала, зависящего от времени. Вектор переменных $X(t)$ в этом случае описывает управляющий процесс. Ограничения в динамической задаче обычно задают как дифференциальные уравнения связи (определяющие движение системы) и краевые условия (см. т. 2).

Во-вторых, устанавливают размерность вектора характеристик X стратегии. При этом выделяют одномерные и многомерные задачи. В некоторых случаях удается свести многомерную задачу принятия решения к одномерной. Это возможно в двух случаях: для статических задач — когда число ограничений-равенств на единицу меньше числа n независимых переменных x_i , $i = \overline{1, n}$ и можно применить метод исключения для выражения $(n - 1)$ переменной через одну оставшуюся; для динамических задач — когда целевая функция может быть представлена в аддитивной форме или является таковой, процесс может быть разбит на несколько шагов (см. метод динамического программирования, т. 2).

В-третьих, статические задачи принятия решения могут быть с ограничениями-равенствами, неравенствами, с ограничениями обоих типов или без ограничений. Первые три типа относятся к задачам условной оптимизации, последний — к задачам безусловной оптимизации. Иногда удобнее свести исходную задачу условной оптимизации к новой задаче безусловной оптимизации путем агрегирования функции $F(X)$ и ограничений $q(X)$, $h(X)$ в новую целевую функцию.

В-четвертых — по характеру изменения переменных x_i , $i = \overline{1, n}$ в области допустимых стратегий. При этом различают дискретные, непрерывные, дискретно-непрерывные (смешанные) задачи принятия решения. Техника решения дискретных и особенно дискретно-непрерывных задач сложна и трудоемка. Поэтому, если это возможно (по требуемой точности получаемого решения), исходную задачу ап-

роксимируют непрерывным ее аналогом.

В-пятых, по виду целевой функции и ограничений задачи оптимизации подразделяют на линейные и нелинейные. Если хотя бы одна из функций задачи (12.13)—(12.14) нелинейна, то задача относится к типу нелинейных. Достоинством линейных задач является то, что они являются наиболее полно теоретически обоснованными и для них разработан математический аппарат — линейное программирование (см. т. 2). Однако большинство реальных задач, адекватно описывающих моделируемые явления и процессы, оказываются нелинейными. К особенностям этих задач относится то, что в общем случае область допустимых решений, выделяемая ограничениями, оказывается не слишком «хорошей» (невыпукла, несвязна в математическом смысле), целевая функция $F(X)$ имеет не один, а несколько экстремумов как внутри, так и вне области. Гарантировать получение точного решения в этом случае нельзя. Однако очень часто решение может быть получено с требуемой точностью. В некоторых частных случаях можно гарантировать точное решение (задачи выпуклого программирования, см. т. 2).

На рис. 1 представлены основные типы задач принятия решения по скалярному показателю в условиях определенности. Каждому типу задач поставлены в соответствие наиболее эффективные методы их решения. Это позволяет легко сориентировать исследователя в выборе конкретного метода по характерным признакам сформулированной им задачи принятия решения. На рисунке приняты следующие обозначения (признаки): *Ст* — статические задачи принятия решения; *Дн* — динамические задачи принятия решения; *О* — одномерные задачи принятия решения; *М* — многомерные задачи принятия решения; *У* — задачи принятия решений при наличии ограничений (задачи условной оптимизации); *Б* — задачи принятия решений без ограничений (задачи безусловной оптимизации); *Н* — переменные x_i , $i = \overline{1, n}$ изменяются непрерывно (непрерывные задачи принятия решения); *Д* — дискретные (или дискретно-непрерывные)

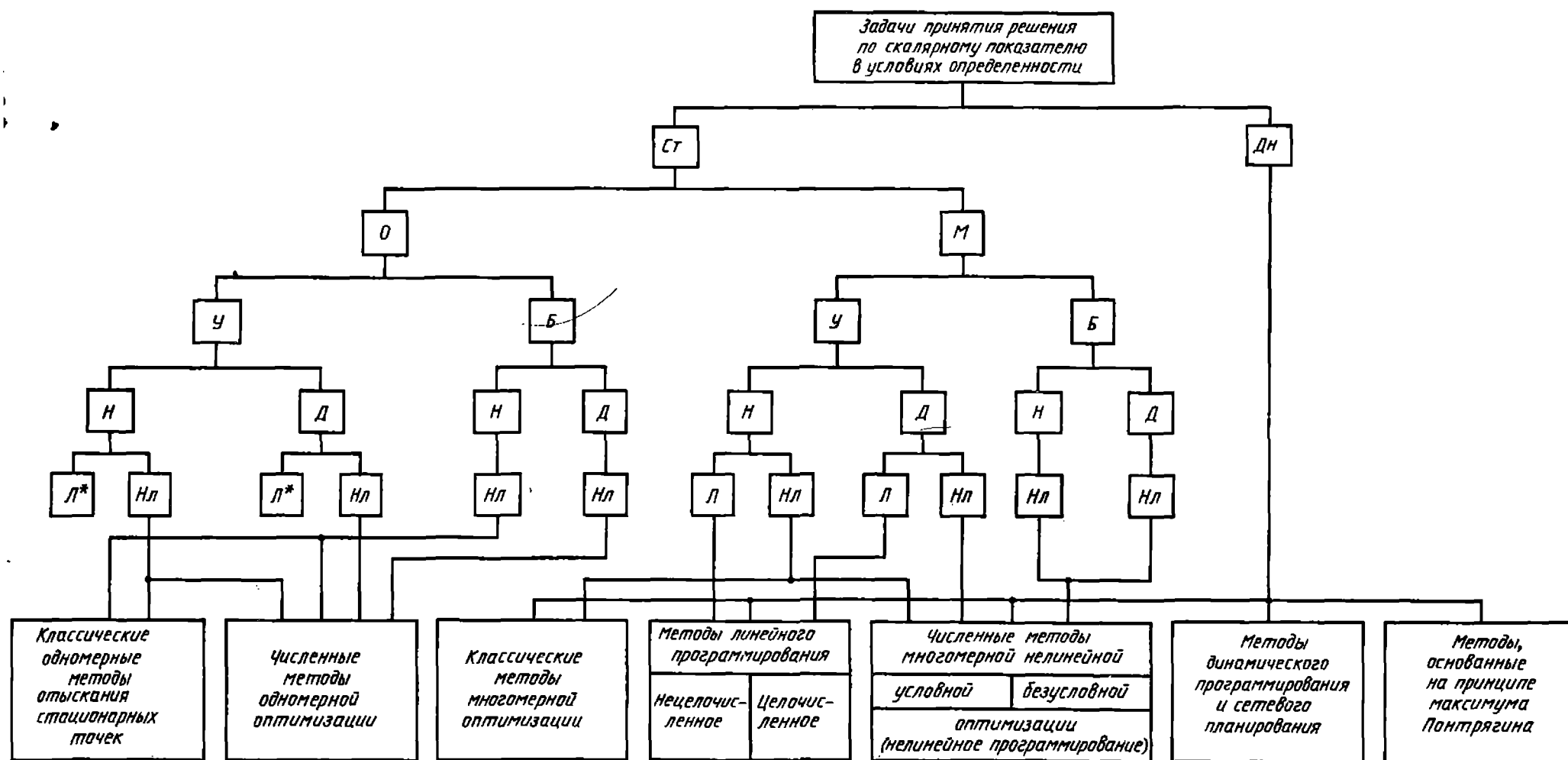


Рис. 1. Типы задач и методы принятия решений по скалярному показателю в условиях определенности

задачи принятия решения; L — линейные задачи принятия решения (задачи линейной оптимизации); NL — нелинейные задачи принятия решений (задачи нелинейной оптимизации).

Символом «*» на рис. 1 отмечены задачи одномерной линейной оптимизации, которые имеют тривиальное решение на одном из концов отрезка, задающего область ограничений. Например, статическая задача многомерной нелинейной условной оптимизации с параметрами x_i , $i = \overline{1, n}$ непрерывного типа может решаться либо классическими методами многомерной оптимизации (если задача достаточно проста и допускает аналитическое решение), либо численными методами нелинейной условной оптимизации (методами нелинейного программирования). Динамическая задача принятия решения может решаться либо путем сведения ее к задаче математического программирования (в общем случае нелинейного), либо специальными методами оптимизации процессов: классическими (вариационное исчисление) и современными методами (метод динамического программирования Беллмана; методы сетевого планирования и управления; методы, основанные на использовании принципа максимума Понтрягина).

Из методов, представленных на рис. 1, классические методы одномерной и многомерной оптимизации, методы линейного программирования, метод динамического программирования изложены в т. 2.

Классические методы условной и безусловной оптимизации основаны на отыскании стационарных точек целевой функции. В общем случае классический подход заключается в замене задачи на экстремум задачей поиска решений системы трансцендентных уравнений вида

$$\varphi'(X) = 0 \quad (12.15)$$

для задач безусловной оптимизации или задач

$$\left. \begin{aligned} \varphi'(X) + \sum_{i=1}^{k_0} c_i^h h_i'(X) &= 0; \\ h(X) &= 0, \end{aligned} \right\} \quad (12.16)$$

или

$$\left. \begin{aligned} \varphi'(X) + \sum_{i=1}^{l_0} c_i^q q_i'(X) &= 0; \\ c_i^q q_i(X) &= 0, \quad i = \overline{1, l_0}; \\ c_i^q &\geq 0, \quad i = \overline{1, l_0}, \end{aligned} \right\} \quad (12.17)$$

или

$$\left. \begin{aligned} \varphi'(X) + \sum_{i=1}^{l_0} c_i^q q_i'(X) + \\ + \sum_{j=1}^{k_0} c_j^h h_j'(X) &= 0; \\ c_i^q q_i(X) &= 0, \quad i = \overline{1, l_0}; \\ c_i^q &\geq 0, \quad i = \overline{1, l_0}. \end{aligned} \right\} \quad (12.18)$$

Система (12.15) есть необходимое условие существования экстремума функции $\varphi(X)$ в задаче безусловной оптимизации, система (12.16) — условие Лагранжа для задач с ограничениями-равенствами, а системы (12.17) и (12.18) получены из необходимого условия Куна — Таккера для задач с ограничениями-неравенствами и смешанного типа соответственно.

Рассмотрим в качестве примера задачу принятия решения, связанную с выбором конструктивных параметров кузова вагонетки для перевозки сыпучего груза.

Пример 1. Необходимо спроектировать кузов вагонетки (рис. 2) для перевозки и механизированной разгрузки сыпучего груза. При этом желательно обеспечить наибольшую вместимость кузова при заданных ограничениях на длину кузова (определяемую особенностями его эксплуатации) и его массу. В качестве конструктивных параметров, подлежащих определению, выберем длину кузова x_1 и радиус боковой стенки x_2 , т. е. стратегия и характеризуется вектором $X = (x_1, x_2)^T$. Предполагается, что мощность двигателя, развивающего крутящий момент M_B , достаточна для любых значений x_2 . Обозначим через l_1, l_2 — минимальную и максимальную длину кузова, определяемую конструктивными особенностями шасси вагонетки, а через P_0 —

максимально допустимую массу без груза. Толщину стального листа для изготовления кузова и удельную плотность стального листа обозначим через d и ρ соответственно.

Целевая функция в задаче проектирования в этом случае имеет вид

$$V = \frac{\pi}{2} x_1 x_2^2 \quad (12.19)$$

при ограничениях:

$$\left. \begin{aligned} (2\pi x_1 x_2 + \pi x_2^2) \rho d &\leq P_0; \\ l_1 &\leq x_1 \leq l_2; \\ x_2 &> 0. \end{aligned} \right\} \quad (12.20)$$

Преобразуем задачу проектирования в задачу оптимизации. Максимумы функций $V = \frac{\pi}{2} x_1 x_2^2$ и $\varphi(X) = x_1 x_2^2$ достигаются в одной и той же точке $X^* = (x_1^*, x_2^*)$, принадлежащей области ограничений. Содержательный анализ системы (12.19), (12.20) показывает, что максимум объема при фиксированных параметрах ρ, d, l_1, l_2, P_0 , задающих множество Λ_F , достигается при максимальной собственной массе $P = P_0$ кузова. Объем V есть монотонно возрастающая функция переменных x_1, x_2 , что позволяет первое неравенство системы (12.20) представить как равенство. Тогда задача оптимизации будет выглядеть следующим образом:

$$X^* : \max [\varphi(X) = x_1 x_2^2] \quad (12.21)$$

при ограничениях:

$$\left. \begin{aligned} h_1(X) &= 2x_1 x_2 + x_2^2 - b = 0; \\ q_1(X) &= -x_1 - l_1 \leq 0; \\ q_2(X) &= x_1 - l_2 \leq 0; \\ q_3(X) &= -x_2 < 0, \end{aligned} \right\} \quad (12.22)$$

где $b = P_0/(\pi d \rho)$.

Эта задача имеет следующий набор характерных признаков: Ст — М — У — Н — Нл. Этот набор признаков (см. рис. 1) определяет две группы методов решения задачи — классические методы многомерной оптимизации и методы нелинейного программирования. Поскольку задача достаточно проста (нет ни одного трансцендент-

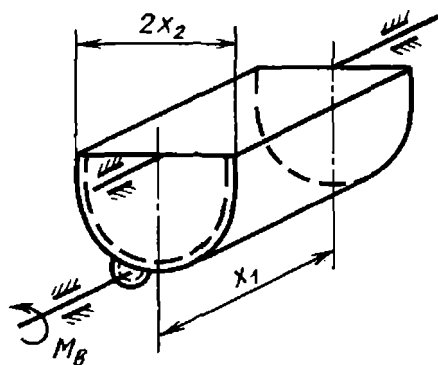


Рис. 2. Схема кузова вагонетки

ного уравнения), решим ее классическим методом. Формально решение можно получить, составив и разрешив систему (12.18). Однако есть более простой путь: отыскать решение задачи

$$\left. \begin{aligned} Z^* : \max \varphi(Z); \\ h_1(Z) &= 0, \end{aligned} \right\}$$

например, методом Лагранжа, проверить, удовлетворяет ли полученное решение ограничениям $q_i(Z) \leq 0, i = \overline{1, 3}$. Далее определить решение в одной из угловых точек области:

$$X = (x_1, x_2)^T \quad h_1(X) = 0, q_1(X) = 0, q_3(X) < 0,$$

$$Y = (y_1, y_2)^T : h_1(Y) = 0, q_2(Y) = 0, q_3(Y) < 0.$$

Та из точек X, Y, Z , которой соответствует максимальное значение φ , является оптимальной.

Запишем систему уравнений (12.16) для решаемой задачи:

$$\left. \begin{aligned} z_2^2 + c_1^h \cdot 2z_1 &= 0; \\ 2z_1 z_2 + c_1^h (2z_1 + 2z_2) &= 0; \\ 2z_1 z_2 + z_2^2 - b &= 0. \end{aligned} \right\}$$

Решение¹ этой системы $Z = \left(\sqrt{\frac{b}{3}}, \right.$

$\left. \sqrt{\frac{b}{3}} \right)^T$, а $\varphi(Z) = \left(\frac{b}{3} \right)^{3/2}$. Пола-

¹ Корни, имеющие отрицательные значения, отбрасываются как не соответствующие физическому смыслу задачи.

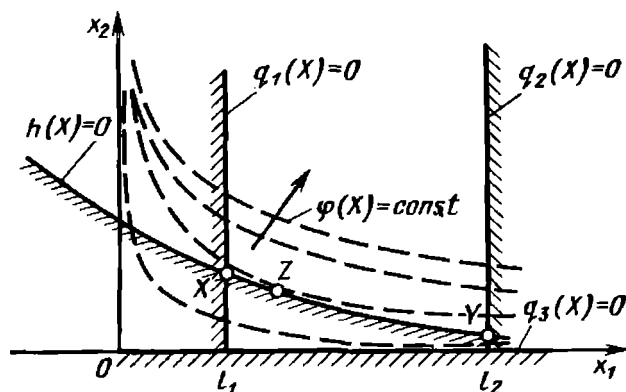


Рис. 3. Графическое решение задачи проектирования

гаем $x_1 = l_1$, тогда из $h_1(X) = 0$ находим $x_2 = \sqrt{l_1^2 + b} - l_1$, а $\varphi(X) = l_1(\sqrt{l_1^2 + b} - l_1)^2$. Аналогично для Y : $y_1 = l_2$, $y_2 = \sqrt{l_2^2 + b} - l_2$, $\varphi(Y) = l_2(\sqrt{l_2^2 + b} - l_2)^2$. Если $l_1 \leq z_1 \leq l_2$, то сравнивают величины $\varphi(X)$, $\varphi(Y)$, $\varphi(Z)$, в противном случае — только две первые. Оптимальным решением X^* является

$$X^* : \max \{ \varphi(X), \varphi(Y), \varphi(Z) \}.$$

Так как в данной задаче число ограничений-равенств на единицу меньше числа независимых переменных, то вместо метода Лагранжа на первом шаге можно было бы выразить одну переменную через другую и решить задачу безусловной одномерной оптимизации, после чего проверить это решение на принадлежность области, задаваемой ограничениями $q_i(X) \leq 0$,

$$i = \overline{1, 3}.$$

Графическое решение задачи проектирования представлено на рис. 3. Допустимая область решений отделена двойной штриховкой. Стрелкой указано направление возрастания целевой функции $\varphi(X)$, штриховыми линиями обозначены изокванты $\varphi(X)$ ($\varphi(X) = \text{const}$).

2. ЗАДАЧИ ПРИНЯТИЯ РЕШЕНИЙ ПО ВЕКТОРНОМУ ПОКАЗАТЕЛЮ

В отличие от задач принятия решения по скалярному показателю, результатом которых является оптимальная стратегия (с точностью до предпосылок и допущений модели Ψ), в задачах принятия решения по векторному показателю нельзя утверждать, что то или иное решение действительно оптимально. Это объясняется тем, что одно из решений может превосходить другое по одним показателям и уступать ему по другим. В таких условиях трудно определить, какая из стратегий более предпочтительна, не говоря уже об ее оптимальности. Если в задачах принятия решения по скалярному показателю основная трудность состоит в разработке или выборе формального метода (алгоритма) поиска экстремума, то в задачах принятия решения по векторному показателю главное внимание уделяется выработке решающего правила, основанного на компромиссе между значениями компонент векторного показателя. Таким образом, сложность проблемы принятия решений по векторному показателю даже в условиях определенности связана не столько с трудностями вычисления, сколько с концептуальной обоснованностью выбора «наилучшего» решения. Невозможно строго математически доказать, что выбранное решение наилучшее — любое решение из числа недоминируемых (неулучшаемых одновременно по всем показателям) может оказаться наилучшим для конкретного ЛПР в конкретных условиях. Это является основной аксиомой принятия решений по нескольким показателям.

ОБЩАЯ ПОСТАНОВКА ЗАДАЧИ

В отличие от постановки задачи принятия решения по скалярному показателю в условиях определенности модель Ψ задает отображение вида

$$\begin{aligned} \Psi : Y \rightarrow & (\rho_1(y_1, y_1^{\text{TP}}), \\ & \rho_2(y_2, y_2^{\text{TP}}), \dots, \\ & \rho_m(y_m, y_m^{\text{TP}})^T, \end{aligned} \quad (12.23)$$

где y_i, y_i^{TP} — детерминированные скалярные характеристики исхода G и требуемого результата ($i = 1, m$). Функция Ψ ставит в соответствие каждой стратегии $u \in U$ значение векторного показателя $W(u) = (W_1(u), W_2(u), \dots, W_m(u))^T$. Обычным порядком система предпочтений $\mathcal{P}$ ЛПР вводится либо на значениях результатов y_i , либо на значениях показателей $W_i, i = \overline{1, m}$ в зависимости от используемых функций ρ_i (см. (12.3), (12.4)). Поэтому в дальнейшем без потери общности рассуждений будем рассматривать предпочтения только на значениях W_i частных показателей эффективности и считать, что большие значения показателей эффективности предпочтительнее меньших¹, а предпочтения ЛПР не меняются скачком.

Так как каждому исходу $g \in G$, характеризующему вектором W , можно поставить в соответствие стратегию $u \in U$, приводящую к этому исходу по правилу

$$u = \Psi^{-1}(W), \quad (12.24)$$

то, выбирая более предпочтительную стратегию $u^* \in U$, ЛПР может ориентироваться на свои предпочтения, заданные на множестве значений $X = X_1 \times X_2 \times \dots \times X_m$ векторного показателя, где X_i — шкала (множество значений) частного показателя эффективности W_i .

Таким образом,

$$u \succsim v \Leftrightarrow W(u) \geq W(v), \quad \forall u, v \in U. \quad (12.25)$$

В общем случае отношение нестрогого предпочтения $\geq$ на множестве значений показателей эффективности X оказывается несвязным. Поэтому выбор наиболее предпочтительной стра-

тегии $u^* \in U$ следует осуществлять среди нехудших (недоминируемых), векторные оценки $W(u^*)$ которых составляют ядро отношения строгого предпочтения $\succ$, т. е.

$$u^* \in M \succ (u). \quad (12.26)$$

В частном случае, когда предпочтения ЛПР не изменяются скачком, а каждый исход операции $W(u)$ может быть оценен с помощью функции эффективности $W_e(X)$, можно выделить множество наилучших стратегий (полагая, что функция выбора моделируется функцией эффективности). Так как большие значения функции эффективности соответствуют более предпочтительным исходам операции, то выбор решения можно осуществлять с использованием критерия оптимальности

$$u^* : \max_{u \in U} W_e(X(u)). \quad (12.27)$$

Если построить функцию эффективности нельзя (либо предпочтения ЛПР меняются скачком в пространстве X^{*1} , либо измерить предпочтения ЛПР в интервальной шкале затруднительно), то при выборе решения следует руководствоваться лишь (12.26).

Общая постановка задачи принятия решения по векторному показателю в условиях определенности формально выглядит следующим образом:

$$\langle \theta_\lambda, \mathcal{P}, W, U; U^* \rangle. \quad (12.28)$$

В отличие от задачи (12.9), в логическом высказывании (12.28) критерий K явно не присутствует. Этим, во-первых, подчеркивается отличительная особенность данной задачи, аксиоматически предполагающей отсутствие оптимального в абсолютном смысле решения, а во-вторых, тем, что множество U во многих случаях уже включает в себя только пригодные по дости-

¹ Это всегда можно осуществить заменой знака на противоположный у показателя, значение которого желательно уменьшить. Если направление предпочтения меняется при переходе некоторого порогового значения y_0 , то осуществляется эквивалентное преобразование показателя, например, в следующем виде: $W_i(u) = -|y_i(u) - y_{0i}|$.

^{*1} ЛПР не может считать, что для любой точки $x \in X$ уменьшение значений одних компонент x_i может быть компенсировано увеличением других компонент x_j , так, что исходная оценка x и новая оказываются одинаковыми по предпочтительности.

жению цели операции стратегии. Таким образом,

$$U = \{u \mid W_i(u) \geq W_i^{TP}, i = \overline{1, m}\}, \quad (12.29)$$

т. е. в (12.28) множество U сформировано на основе критерия пригодности. Если удастся предпочтения ЛПР смоделировать стремлением к увеличению значений функции эффективности, то $U^* \subset U$ может быть найдено с помощью критерия оптимальности (12.27).

В общем случае задача (12.28) требует выявления системы предпочтений ЛПР на множестве X и построения на ее основе решающего правила. В зависимости от способа формирования решающего правила методы решения задачи (12.28) можно условно разделить на две группы. К первой группе методов относятся эвристические, т. е. такие, в которых ЛПР сразу устанавливает вид свертки компонент векторного показателя в скалярный, который затем максимизируется. При получении решения его анализируют; при необходимости ЛПР накладывает дополнительные ограничения на значения частных показателей эффективности, не удовлетворяющих ЛПР, или изменяет параметры свертки. Такая процедура повторяется до тех пор, пока не будет получено приемлемое с точки зрения ЛПР решение. Вторая группа методов (аксиоматические методы) основана на выдвижении и проверке аксиом, в случае удовлетворения которых указывается вполне определенный вид функции эффективности и способ определения ее параметров. С помощью построенной функции эффективности решается задача (12.27). Методы как первой, так и второй групп подразделяют на одношаговые и многошаговые (итеративные). Одношаговые методы позволяют отыскивать решение за один прием (шаг) на основе однократно сформированного решающего правила. В многошаговых процедурах реализуется, как правило, либо принцип вложенных отношений, либо одно и то же решающее правило применяют многократно (с соответствующей корректировкой параметров функции свертки).

Вид параметров функции свертки или функции эффективности определяют и уточняют на основе анализа комплексной информации θ о решаемой задаче и предпочтениях ЛПР. При этом мощность выделяемого решающим правилом множества стратегий $U^* \subset U$ определяется полнотой и качеством (достаточностью) информации θ .

Самая простая информация — независимость частных показателей по предпочтению, более сложная — значения «замещений» частных показателей W_i или информация Ω об относительной важности этих показателей с ее особыми случаями равноценных (S), строго упорядоченных (L) и симметрически лексикографических (SL) показателей эффективности. Самая сложная информация — значения коэффициентов важности частных показателей W_i .

При выборе конкретного метода построения решающего правила необходимо учитывать возможность получения той или иной информации и доступные способы ее обработки. Кроме того, каждой группе методов присущи свои достоинства и недостатки. Эвристические методы исследователи широко используют в практике. После того как осуществлена свертка частных показателей, решение задачи может быть сравнительно просто получено методами математического программирования. Однако проблема определения вида и параметров функции свертки не является в этом случае главной. Одна проблема подменяется другой, не менее сложной. Использование аксиоматических методов позволяет построить более адекватную модель системы предпочтений ЛПР и сформировать решающее правило. Однако эти методы требуют специальной подготовки ЛПР и исследователя, а реализация этих методов сопряжена со сравнительно большими затратами времени.

ЭФФЕКТИВНЫЕ СТРАТЕГИИ

Пусть известно, что предпочтения ЛПР на множестве векторных оценок X изменяются плавно и для любой произвольной точки $x \in X$ можно так

изменить ее компоненты, что исходная и новая точки будут одинаковы по предпочтительности для достижения цели операции. Для двумерного случая (рис. 4) это допущение означает, что в X имеется некоторое подмножество точек, эквивалентных по предпочтительности и всем им можно поставить в соответствие некоторую определенную степень (меру) эффективности. Если эту степень измерять с помощью функции эффективности $W_e(X)$, то в рассматриваемом двумерном пространстве X можно изобразить линии одинаковой предпочтительности — кривые равноценности $W_e(X) = \text{const.}$ Если $x \sim y$, то они расположены на одной кривой равноценности $W_e(x) = a$, а если точка $z \succ x$, то она расположена на кривой равноценности $W_e(z) = b > a$. Кривые равноценности полностью моделируют предпочтения ЛПР во множестве X . Выше было сделано допущение, что большие значения частных показателей эффективности (компонент оценок x) предпочтительнее меньших. Даже в этом случае может оказаться, что для числа частных показателей $m \geq 3$ направления предпочтения по одним показателям изменяются в зависимости от того, какие значения принимают другие.

Пример 1. При управлении деятельностью предприятия ЛПР руководствуется тремя показателями: W_1 — процент выполнения плана; W_2 — время, оставшееся до конца планового периода; W_3 — дополнительные затраты, связанные с обеспечением выполнения плана. Если процент выполнения плана высок и имеется большой резерв времени до отчетного периода, то стимулировать деятельность предприятия нет необходимости, т. е. ЛПР будет предпочитать большие значения W_3 меньшим. Если выполнение плана под угрозой (недопустимо низкие значения W_1 и W_2), то ЛПР скорее всего предпочтет уменьшение значений W_3 (увеличение дополнительных затрат), а не срыв плана. Таким образом, направление предпочтения по показателю W_3 зависит от того, какие значения принимают показатели W_1 и W_2 .

Показатель W_i , для которого направление предпочтения зависит от

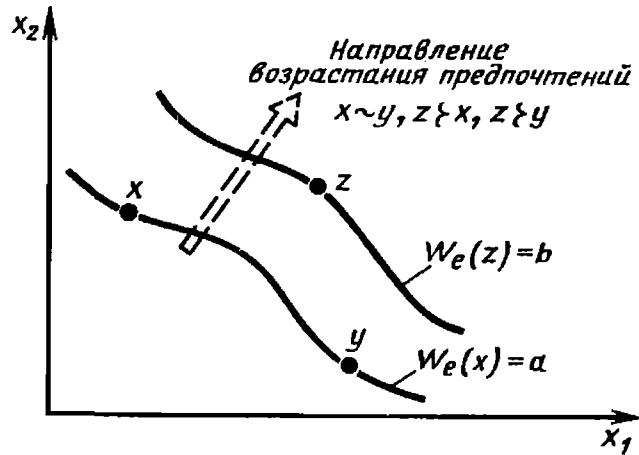


Рис. 4. Кривые равноценности

того, какие значения принимают другие показатели, называется *зависимым по предпочтению от остальных*. Если направления предпочтения по каждому частному показателю W_i не изменяются от того, какие значения принимают другие показатели, то такие показатели называются *взаимонезависимыми по предпочтению*. Случаи такого рода весьма часто имеют место в практике. Обозначим через Ω часть общей информации о задаче, относящуюся к предпочтениям ЛПР. На ранних этапах принятия решения информация $\omega_0 \in \Omega \subseteq \theta$ о взаимонезависимости частных компонент векторного показателя позволяет значительно сузить исходное множество стратегий U , не вскрывая полностью систему предпочтений ЛПР. Решающее правило в этом случае строится на основе аксиомы В. Парето [52].

Аксиома Парето. Если в задаче принятия решения частные показатели независимы по предпочтению и значение каждого из них желательно увеличивать, то из двух векторных оценок $x, y \in X$, для которых выполняются неравенства

$$x_i \geq y_i, \quad i = \overline{1, m}, \tag{12.30}$$

векторная оценка x не менее предпочтительнее оценки y :

$$x \overset{\omega_0}{\succsim} y. \tag{12.31}$$

Если хотя бы одно из неравенств (12.30) выполняется как строгое, то

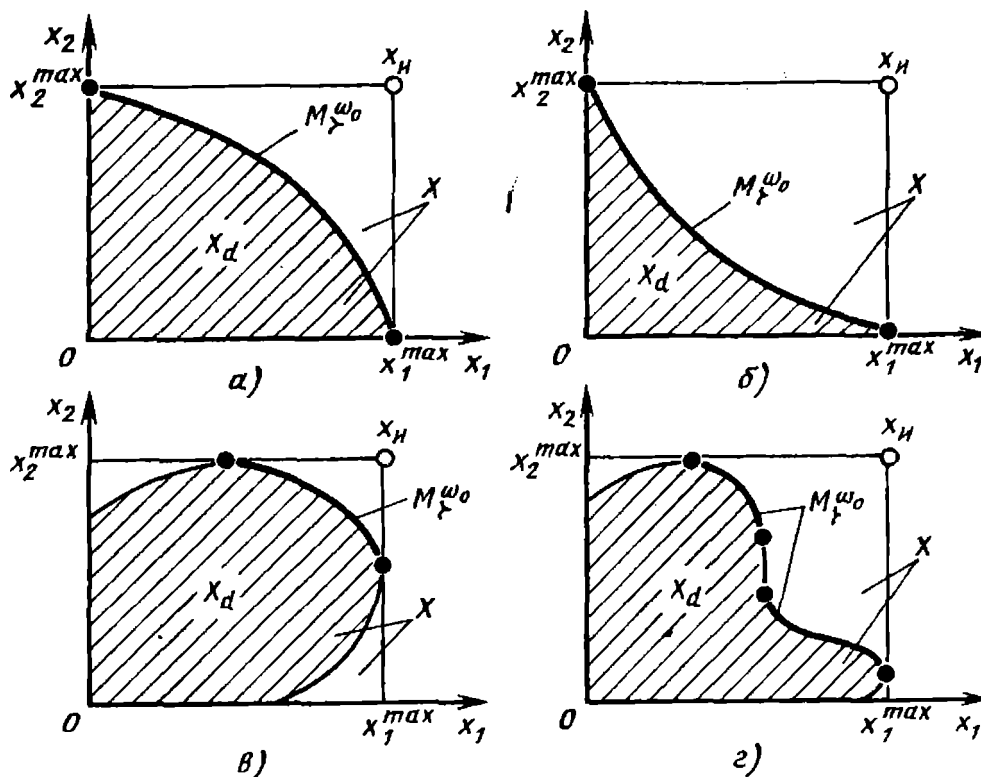


Рис. 5. Достижимое множество и ядро отношения

$x \succ_{\omega_0} y$. Если все неравенства выполняются как равенства, то $x \sim_{\omega_0} y$.

Построенные отношения (12.31) на множестве X векторных оценок порождают аналогичные отношения предпочтения на множестве стратегий:

$$u \succ_{\omega_0} v \Leftrightarrow x(u) \succ_{\omega_0} x(v). \quad (12.32)$$

Отношения (12.31) и (12.32) не являются связными, так как для произвольных векторных оценок $x, z \in X$ часть неравенств в (12.30) может выполняться в одну сторону ($x_i \geq z_i$, $i = \overline{1, k}$), а часть — в другую ($x_j \leq z_j$, $j = \overline{k+1, m}$). Такие векторные оценки называются *несравнимыми по Парето*. На основании отношения строгого предпочтения $\succ_{\omega_0}$ на множестве достижимых оценок X_d ¹ вводится

¹ Множество достижимых векторных оценок $X_d \subset X$ определяется структурой оператора Ψ , т. е. $\Psi: U \rightarrow X_d$, где U — множество допустимых стратегий.

ядро отношения M_{ω_0} Парето:

$$M_{\omega_0} = \{x \mid \exists y \in X_d : y \succ_{\omega_0} x, x \in X_d\}. \quad (12.33)$$

На рис. 5 для двумерного случая показаны множества X , X_d и M_{ω_0} ,

где $X = X_1 \times X_2$, $X_1 = [0, x_1^{\max}]$, $X_2 = [0, x_2^{\max}]$, $x_H = (x_1^{\max}, x_2^{\max})$ — идеальная точка. Ядро отношения M_{ω_0} выделено жирной линией; оно может быть как разрывным (рис. 5, г), так и неразрывным (рис. 5, а—в), содержать все точки границы X_d , лежащие в направлениях «север—восток» относительно начала координат (рис. 5, а и б), или только часть из них (рис. 5, в и г).

Стратегии U_{ω_0} , векторные оценки которых принадлежат ядру отношения M_{ω_0} , называются *эффективными* или *недоминируемыми по Парето*. Все остальные стратегии $U \setminus U_{\omega_0}$ называются *доминируемыми*. Если некоторая стратегия $v \in U$ доминируема,

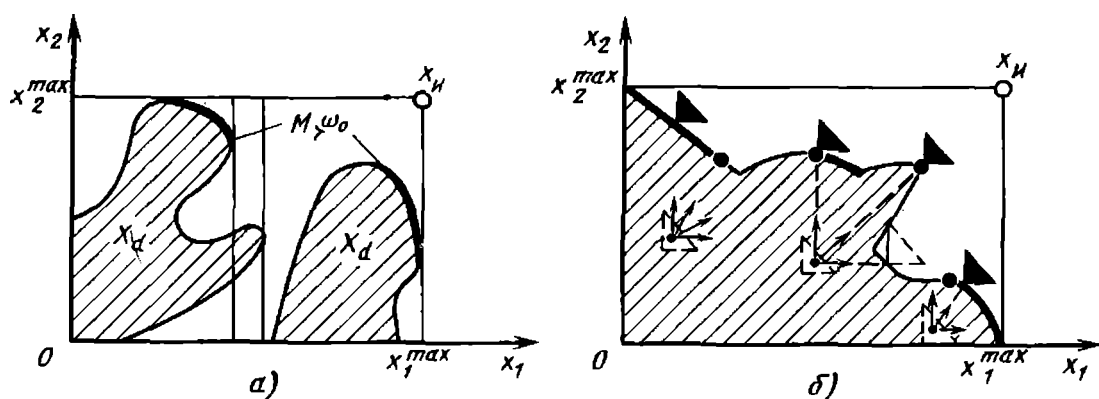


Рис. 6. Графические способы выделения эффективных стратегий

то ее нецелесообразно рассматривать в качестве возможного решения, так как существует (хотя бы одна) эффективная стратегия $u \in U_{\omega_0}$, которая дает большее значение хотя бы по одному из частных показателей W_i , $i = \overline{1, m}$ по сравнению со стратегией v [выполняется соотношение (12.30)]. Следовательно, решение следует искать лишь на множестве U_{ω_0} эффективных стратегий.

Эффективные стратегии не сравнимы между собой по информации ω_0 . Поэтому однозначный выбор по информации ω_0 можно осуществить лишь в крайне редком случае (множество M_{ω_0} состоит из одного элемента).

Для задачи принятия решения по двум, в редких случаях трем, частным показателям эффективности множество эффективных стратегий может быть найдено, исходя из определения (12.33). Для этого в пространстве X строится множество достижимых векторных оценок X_d , а затем выделяется та его «северо-восточная» часть, которая заключена между линиями $x = x_1^{\max}$ и $x = x_2^{\max}$. Это можно сделать, если область X_d строго выпукла (рис. 5, а и в) или строго вогнута (рис. 5, б). В противном случае (за исключением линейных границ) «северо-восточная» граница области X_d разбивается на строго выпуклые и строго вогнутые участки линиями, проходящими параллельно осям координат через точки перегиба или излома границы (рис. 6, а). В общем случае множество M_{ω_0}

можно выделить с помощью приема, показанного на рис. 6, б. Для этого необходимо прямоугольный треугольник расположить так, чтобы его катеты были параллельны осям координат, и перемещать его в области X_d по направлениям «север—восток» до тех пор, пока вершина прямого угла не коснется границы области X_d . Точка касания включается в множество M_{ω_0} , если в створ прямого угла ¹ не попадает ни одна из точек $x \in X_d$.

Если множество X_d дискретно, то эффективные стратегии отыскиваются на основе определения (12.33) путем попарного сравнения векторных оценок.

Если число компонент векторного показателя эффективности больше трех, то отыскать эффективные стратегии по определению (12.33) трудно. В этом случае эффективные стратегии выделяют с помощью специальных методов, в основе которых лежат теоремы о свойствах эффективных стратегий.

Теорема 1. Стратегия u^* является эффективной тогда и только тогда, когда для любого $i = \overline{1, m}$ она удовлетворяет условию

$$W_i(u^*) = \max_{u \in U} W_i(u); \quad (12.34)$$

$$W_j(u) \geq b_j, \quad j = \overline{1, m}, \quad j \neq i,$$

$$\text{где } b_j = W_j(u^*), \quad j = \overline{1, m}.$$

¹ Предполагается, что размеры катетов прямоугольника могут быть сколь угодно большими.

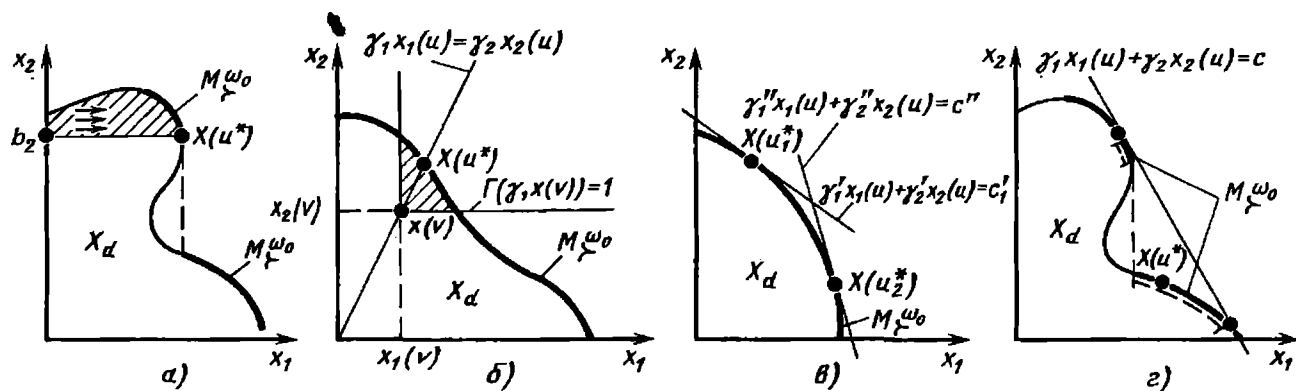


Рис. 7. Способы выделения эффективных стратегий

Теорема 2. Стратегия u^* эффективна тогда и только тогда, когда она при всех $W_i(u^*) > 0$ и $\gamma_i^* = \frac{1}{W_i(u^*)}$, $i = \overline{1, m}$ удовлетворяет условию

$$\Gamma(\gamma^*, W(u^*)) = \max_{u \in U} \Gamma(\gamma^*, W(u)), \quad (12.35)$$

где $\Gamma(\gamma, x) = \min_i \{\gamma_i \cdot x_i\}$.

Теорема 3. Если множество U выпукло, а все $W_i(u)$ — вогнуты, то для всякой эффективной стратегии u^* найдутся такие неотрицательные числа

$$\gamma_i, \quad \sum_{i=1}^m \gamma_i = 1, \quad \text{что}$$

$$F(\gamma, u^*) = \max_{u \in U} F(\gamma, u), \quad (12.36)$$

$$\text{где } F(\gamma, u) = \sum_{i=1}^m \gamma_i W_i(u).$$

Смысл теорем 1—3 для двумерного случая иллюстрируется на рис. 7, а—г. На всех рисунках показана достижимая область X_d и выделено ядро отношения M_{ω_0} . На рис. 7, а эффективная стратегия u^* соответствует точке $X(u^*)$, полученной в результате максимизации первой компоненты векторного показателя при условии, что множество значений второй компоненты принадлежит заштрихованной области. Линии уровня функции $\Gamma(\gamma, x)$ на рис. 7, б образуют прямой угол с осями координат. Для произвольной страте-

гии $v \in U$, векторная оценка которой соответствует точке $X(v)$ и $\gamma_i = \frac{1}{x_i(v)}$, $i = 1, 2$, функция $\Gamma(\gamma, v)$ принимает значение, равное 1. Для всех стратегий u , векторные оценки которых расположены в заштрихованной области, значения функции $\Gamma(\gamma, u) > 1$. Эффективной стратегии u^* соответствует максимум функции $\Gamma(\gamma, u)$.

Если условия теоремы 3 выполнены (множество X_d выпукло), то эффективной стратегии u^* будет соответствовать точка касания линии уровня функции $F(\gamma, u)$ границы области X_d (рис. 7, в). Требование выпуклости X_d (вогнутости всех W_i) существенно, что иллюстрирует рис. 7, г, т. е. эффективные стратегии u^* , векторные оценки которых принадлежат подчеркнутой штриховой линии части ядра, не могут быть выделены ни при каких коэффициентах γ_i , $i = 1, 2$.

С помощью (12.34)—(12.36), теорем 1—3 выделение эффективных стратегий из множества допустимых может быть формально осуществлено решением задач математического программирования. Такая процедура выделения эффективного множества может быть полностью осуществлена лишь для дискретного множества стратегий. Для непрерывного множества стратегий U его эффективное подмножество аппроксимируется по отдельным точкам и при необходимости уточняется.

Если задачи математического программирования, соответствующие (12.34)—(12.36), требуют больших затрат времени и ресурсов, то эффективные стратегии выделяют следующим

образом. Случайным образом генерируют стратегию $u \in U$, для нее вычисляют оценку $X(u)$ в соответствии с моделью Ψ . Генерируют следующую стратегию v , которую сравнивают со стратегией u по правилу (12.32). Доминируемую стратегию отбрасывают. Процедуру повторяют необходимое число раз; при этом каждую вновь сгенерированную стратегию сравнивают со всеми недоминируемыми на предыдущих шагах. Генерация стратегий должна осуществляться таким образом, чтобы обеспечить равномерность выборки на множестве U . Иногда используют специальные последовательности равномерно распределенных точек — ЛП $_{\tau}$ -последовательности.

ОМЕГА-ЭФФЕКТИВНЫЕ СТРАТЕГИИ

Как правило, множество эффективных стратегий включает достаточно большое число альтернатив, что не позволяет осуществить однозначный выбор. Например, для двумерного дискретного случая математическое ожидание мощности P множества M_{ω_0} стремится к величине $\lim_{N \rightarrow \infty} P(M_{\omega_0}) = \ln N$, где N — мощность множества допустимых стратегий.

Для дальнейшего сужения множества эффективных стратегий U_{ω_0} необходимо, чтобы информация Ω о предпочтениях ЛПР на множестве оценок X векторного показателя W была шире, чем простое сообщение ω_0 о взаимной независимости частных показателей. Такой информацией являются сведения об относительной важности частных показателей. Соизмеряя влияние изменения одного показателя по сравнению с другим (другими) на предпочтительность исходов операции, можно определить относительную важность частных показателей. Относительную важность показателей часто измеряют нормированными коэффициентами важности $\gamma_i \geq 0$,

$\sum_{i=1}^m \gamma_i = 1, i = \overline{1, m}$, которые определяют методами экспертного оценивания. Кроме того, относительную важ-

ность показателей W_i учитывают на основе качественной информации, представляющей собой совокупность сообщений $\omega \in \Omega$ о том, что некоторые показатели «равноценны» или же одни «важнее» других [59, 41]. Такая информация получается при предъявлении ЛПР векторных оценок $x, y \in X$ и выяснении у него, какую из них он предпочитает. При этом предъявляемые оценки x и y должны удовлетворять двум требованиям. Во-первых, все компоненты векторных оценок должны иметь общую шкалу (быть однородными), во-вторых, в предъявляемых оценках все компоненты, кроме тех, чья относительная важность выясняется, должны быть одинаковыми.

Из исходной оценки $x \in X$ путем перестановки компонент эти требования позволяют получить новую оценку z , которая также будет принадлежать X .

Определение 1. Показатели W_r и W_t имеют одинаковую важность («равноценны»), если эквивалентными по предпочтительности являются оценки x и $z = \tilde{x}^{rt}$, где символ « $\sim^{rt}$ » над x означает, что z получается из x перестановкой компонент с номерами r и t .

Качественная информация, полученная на основе определения 1, обозначается символом $r\omega_e t$.

Определение 2. Показатель W_r «важнее» показателя W_t , если оценка x , имеющая компоненту $x_r > x_t$, предпочтительнее оценки $z = \tilde{x}^{rt}$.

Качественную информацию, полученную на основе определения 2, обозначим символом $r\omega_{\text{в}} t$. Для того чтобы воспользоваться определениями 1 и 2, необходимо однородные частные показатели привести к однородному виду. В практике используют специальные приемы эквивалентного преобразования частных показателей к единому безразмерному виду. Так, если все исходные показатели $W_i^0, i = \overline{1, m}$ имеют количественную шкалу, применяют следующие приемы.

Если удастся установить требуемые значения $W_i^{\text{ТР}}$ показателей W_i^0 , или известен их некоторый эталонный уровень (например, международный стан-

дарт), то используют преобразование следующего вида:

$$W_i = \frac{W_i^0}{W_i^{\text{тр}}}, \quad i = \overline{1, m}, \quad (12.37)$$

причем $W_i^{\text{тр}} \neq 0$.

Второй прием основан на нормировании исходных показателей W_i^0 их максимально возможными значениями, которые могут быть установлены либо на основе логического анализа задачи, либо найдены путем решения частной оптимизационной задачи $W_i^{\text{max}} = \max_{u \in U} W_i(u)$. В этом случае

$$W_i = \frac{W_i^0}{W_i^{\text{max}}}, \quad i = \overline{1, m}, \quad W_i^{\text{max}} \neq 0. \quad (12.38)$$

Недостатком этого приема является то, что при поиске компромиссного решения тот показатель, который имеет максимальную ненормированную величину, может получить существенный приоритет перед остальными.

Следующий прием основан на нормировании исходных показателей диапазонами их возможного изменения, т. е. либо

$$W_i = \frac{W_i^0}{W_i^{\text{max}} - W_i^{\text{min}}}, \quad (12.39)$$

либо

$$W_i = \frac{W_i^0}{W_i^{\text{тр}} - W_i^{\text{min}}}, \quad i = \overline{1, m}, \quad (12.40)$$

где W_i^{min} — минимально допустимое значение i -го показателя. Этот прием нецелесообразно применять, если диапазоны изменения частных показателей сравнительно малы. В этом случае часто применяют такие преобразования:

$$W_i = \frac{W_i^0 - W_i^{\text{min}}}{W_i^{\text{max}} - W_i^{\text{min}}}; \quad (12.41)$$

$$W_i = \frac{W_i^0 - W_i^{\text{min}}}{W_i^{\text{тр}} - W_i^{\text{min}}}, \quad i = \overline{1, m}. \quad (12.42)$$

Если исходные показатели имеют качественную или смешанные шкалы, то преобразования (12.37)–(12.42) недопустимы. В этом случае используют следующий прием сведения показателей к единой порядковой шкале. Шкалу каждого показателя $W_i^0, i = \overline{1, m}$ разбивают на N качественных классов, а классы нумеруют. Экспертным путем осуществляют сортировку значений исходных показателей, относя их к тому или иному классу. В этом случае

$$W_i = N(W_i^0), \quad i = \overline{1, m}, \quad (12.43)$$

т. е. значением однородного показателя эффективности W_i будет номер класса, к которому принадлежит значение W_i^0 . Использовать такой прием надо достаточно осторожно, так как выделяемое решение существенно зависит от числа градаций порядковой шкалы N и адекватности проведенной сортировки.

На основе качественной информации ω_e о равноценности или ω_b о превосходстве по важности одних однородных показателей над другими строится отношение предпочтения на множестве векторных оценок X , которое затем можно трансформировать на множество стратегий U .

Модель системы предпочтений $\mathcal{P}_\Omega$ задается отношениями строгого предпочтения $\succ_\Omega$, нестрогого предпочтения $\succeq_\Omega$ и эквивалентности $\sim_\Omega$. При этом полагается, что информация $\Omega = \{\omega_0, \omega_e, \omega_b\}$. Построение модели $\mathcal{P}_\Omega$ на множестве значений оценок X однородных показателей $W_i, i = \overline{1, m}$ осуществляют одним из следующих приемов.

Произвольные оценки $x, y \in X$ непосредственно сравнивают по предпочтительности на основе определений 1, 2 и паретовского правила (12.30).

Пример 2. Оценки $x = (5, 2, 3)^T$ и $y = (2, 4, 3)^T$ являются эффективными. Получена информация $\Omega = 1 \omega_b 2$, т. е. первый показатель важнее второго. Тогда оценка $z = \tilde{y}^{12} = (4, 2, 3)^T \succ_{\omega_b} y = (2, 4, 3)^T$ Но $x = (5,$

$2, 3)^T \succ^{\omega_0} z = (4, 2, 3)^T$, поэтому вследствие транзитивности оценка x предпочтительнее оценки y по информации Ω . Если сравнить произвольные оценки $x, y \in X$ этим способом не удастся, то необходимо подобрать такие промежуточные оценки $z^1, z^2, \dots, z^t$, чтобы они образовали транзитивную цепочку отношений от x к y ^{*1}.

Пример 3. Оценки $x = (3, 2, 6, 4, 4)^T$ и $y = (2, 3, 2, 4, 5)^T$ являются эффективными; $\Omega = \{1 \omega_{\text{в}} 2, 3 \omega_{\text{в}} 5\}$. Построим следующие оценки $z^{(t)}$:

$$z^1 = (3, 2, 5, 4, 2)^T,$$

$$z^2 = (3, 2, 2, 4, 5)^T.$$

Тогда $x = (3, 2, 6, 4, 4)^T \succ^{\omega_0} z^1 = (3, 2, 5, 4, 2)^T \sim^{\omega_{\text{в}}} z^2 = (3, 2, 2, 4, 5)^T \succ^{\omega_{\text{в}}} y = (2, 3, 2, 4, 5)^T$, т. е. $x \succ y$.

Другой прием позволяет формализовать процедуру построения модели $\mathcal{P}_{\Omega}$ без эвристического подбора цепочек промежуточных оценок, z путем введения специальных (опорных) множеств оценок, построенных на исходных векторах x, y . Этот прием особенно удобен, когда размерность задачи и число сообщений велики. Его используют при построении человеко-машинных диалоговых процедур.

Алгоритм 1. 1. В качестве исходной выбрать одну из сравниваемых оценок $x, y \in X$ (например, x).

2. Для другой оценки (y) на основании информации $\omega_e \in \Omega$ построить опорное множество ^{*2}

$$Y^{\omega_e} = \{y, \tilde{y}^{\omega_e t}\}.$$

3. Сравнить оценку x с оценками из Y^{ω_e} по Парето. Если оценки несравни-

мы, то перейти к п. 4; в противном случае

$$x^{\Omega} \sim y, \text{ если } x^{\omega_0} \sim \tilde{y}^{\omega_e t};$$

$$x \succ y, \text{ если } x \succ^{\omega_0} \tilde{y}^{\omega_e t}$$

4. Используя опорное множество Y^{ω_e} и информацию $\omega_{\text{в}} \in \Omega$, построить опорное множество ^{*3}

$$Y^{\omega_{\text{в}}} = \{\tilde{y}^{\omega_{\text{в}} t}, a^{\omega_{\text{в}} t}, b^{\omega_e t}\},$$

$$\text{где } a = \tilde{y}^{\omega_e t}, b = \tilde{a}^{\omega_{\text{в}} t}.$$

5. Сравнить x по Парето с оценками из $Y^{\omega_{\text{в}}}$. Если оценки несравнимы, то выбрать за исходную оценку y и перейти к п. 1 ^{*4}.

Пример 4. Сравнить по информации $\Omega = \{1 \omega_{\text{в}} 2, 2 \omega_e 4\}$ векторные оценки $x = (3, 6, 3, 5)^T$, $y = (5, 3, 4, 7)^T$.

В качестве исходной принимаем оценку x . Тогда $Y^{\omega_e} = \{(5, 3, 4, 7)^T, (5, 7, 4, 3)^T\}$. Оценка x несравнима с оценками из Y^{ω_e} . Строим множество $Y^{\omega_{\text{в}}}$: $Y^{\omega_{\text{в}}} = \{(7, 5, 4, 3)^T, (7, 3, 4, 5)^T\}$. Оценка x несравнима с оценками из $Y^{\omega_{\text{в}}}$ по Парето, поэтому в качестве исходной выбираем оценку y . Тогда $X^{\omega_e} = \{(3, 6, 3, 5)^T, (3, 5, 3, 6)^T\}$. Оценка y несравнима с оценками из X^{ω_e} . Строим опорное множество

$$X^{\omega_{\text{в}}} = \{(6, 3, 3, 5)^T, (5, 3, 3, 6)^T, (6, 5, 3, 3)^T, (5, 6, 3, 3)^T\}.$$

^{*3} Множество $Y^{\omega_{\text{в}}}$ включает улучшенные по сравнению с Y^{ω_e} оценки, полученные путем перестановки компонент с номерами g и t , для которых есть информация о «превосходстве» по важности, и оценки, полученные из последних (и им эквивалентные) путем перестановки компонент в соответствии с информацией ω_e .

^{*4} Если это второе обращение к п. 5, то оценки x и y несравнимы по информации Ω .

^{*1} В предыдущем примере такая цепочка состояла из единственного элемента z .

^{*2} Множество Y^{ω_e} включает как саму оценку y , так и все эквивалентные по ω_e оценки, полученные перестановкой компонент с номерами g и t , для которых имеется информация о «равноценности».

Оценка y предпочтительнее по Парето второй оценки из множества X^{ω_B} , поэтому $y \succ_{\Omega} x$.

Использовать качественную информацию об относительной важности частных показателей для установления предпочтительности той или другой оценки из множества X можно лишь тогда, когда она непротиворечива и содержательна. Информация Ω противоречива, если построенное по ней отношение строгого предпочтения $\succ_{\Omega}$ оказывается рефлексивным. Необходимым и достаточным признаком, указывающим на противоречивость информации Ω , является существование цикла вида

$$r_1 \omega_1 r_2 \omega_2 r_3 \omega_3 \dots \omega_{l-1} r_l \omega_l r_1,$$

где r_i — номер компоненты показателя; ω_i — сообщения об относительной важности частных показателей, причем хотя бы одно из них сообщение типа ω_B .

Информация об относительной важности показателей эффективности содержательна, если с ее помощью можно сравнить хотя бы одну пару оценок, несравнимых по предыдущей информации.

Если непротиворечивая информация Ω содержит сведения об относительной важности всех компонент векторного показателя (полное число непротиворечивых сообщений о «равноценности» или «важности» равно $m(m-1)/2$), то процедуру построения отношения предпочтения [93] можно осуществить по следующему алгоритму.

Алгоритм 2. 1. По информации Ω составить систему неравенств и равенств для неизвестных $\alpha_i > 0$, $i = \overline{1, m}$ по следующему правилу:

$$\alpha_i > \alpha_j, \text{ если } i \omega_B j; \quad (12.44)$$

$$\alpha_r = \alpha_t, \text{ если } r \omega_e t.$$

2. Подобрать произвольные неотрицательные числа α_i , удовлетворяющие полученной системе неравенств.¹

¹ Если информация Ω противоречива, то система (12.44) несовместна.

3. Выбрать положительное число β из условия

$$\beta < \min_i \alpha_i, \quad i = \overline{1, m}. \quad (12.45)$$

4. Из компонент векторных оценок x и y сформировать упорядоченное множество $T(x, y) = (t_1, t_2, \dots, t_k)$, состоящее из чисел $t_1 > t_2 > t_3 > \dots > t_k$, таких, что они являются компонентами оценок x и y .²

5. Сформировать векторы $\alpha^j(x)$ с компонентами $\alpha_i^j(x)$, определяемыми по правилу:

$$\alpha_i^j(x) = \begin{cases} \alpha_i, & \text{если } x_i \geq t_j; \\ \beta, & \text{если } x_i < t_j, \end{cases} \quad (12.46)$$

и вектор $\alpha^j(y)$ — аналогично.

6. Построить вектор-функции $\varphi(\alpha^j(x))$ и $\varphi(\alpha^j(y))$ по правилу:

$$\varphi_1(\alpha^j(x)) = \min_i \alpha_i^j(x),$$

$$\dots, \quad \varphi_m(\alpha^j(x)) = \max_i \alpha_i^j(x);$$

$$\varphi_1(\alpha^j(y)) = \min_i \alpha_i^j(y),$$

$$\dots, \quad \varphi_m(\alpha^j(y)) = \max_i \alpha_i^j(y),$$

$$i = \overline{1, m}, \quad j = \overline{1, k-1}. \quad (12.47)$$

7. Для каждого номера j ($j = \overline{1, k-1}$) проверить справедливость неравенств:

$$\varphi_i(\alpha^j(x)) \geq \varphi_i(\alpha^j(y)), \quad i = \overline{1, m}. \quad (12.48)$$

8. Если хотя бы одно из неравенств системы (12.48) строгое, то $x \succ_{\Omega} y$; если все неравенства выполняются как равенства, то $x \sim y$. В противном случае x и y несравнимы.

Пример 5. Сравнить эффективные стратегии, векторные оценки которых $x = (0, 2, 6)^T$, $y = (1, 0, 5)^T$ Ин-

² Множество T получается из множества $\{x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_m\}$ исключением одинаковых элементов, кроме одного, и расположением оставшихся по убыванию.

формация Ω состоит из сообщений $\Omega = \{1 \omega_e 2, 1 \omega_B 3, 2 \omega_B 3\}$. Система (12.44) имеет вид:

$$\left. \begin{aligned} \alpha_1 &= \alpha_2; \\ \alpha_1 &> \alpha_3; \\ \alpha_2 &> \alpha_3. \end{aligned} \right\}$$

Выберем α_i , удовлетворяющие этой системе: $\alpha_1 = 3, \alpha_2 = 3, \alpha_3 = 2$; при этом β выберем равным нулю. Составляем множество из компонент оценок $x, y: \{0, 2, 6, 1, 0, 5\}$, из которого удаляем один нулевой элемент, а остальные упорядочиванием по убыванию. Получаем множество $T(x, y) = \{6, 5, 2, 1, 0\}$, $k = 5$. Формируем векторы $\alpha^j(x)$ и $\alpha^j(y)$:

$$\begin{aligned} \alpha^1(x) &= (0, 0, 2)^T, \alpha^2(x) = (0, 0, 2)^T, \\ \alpha^3(x) &= (0, 3, 2)^T, \alpha^4(x) = (0, 3, 2)^T; \\ \alpha^1(y) &= (0, 0, 0)^T, \alpha^2(y) = (0, 0, 2)^T, \\ \alpha^3(y) &= (0, 0, 2)^T, \alpha^4(y) = (3, 0, 2)^T \end{aligned}$$

Строим вектор-функции $\Phi(\alpha^j(x))$ и $\Phi(\alpha^j(y))$:

$$\begin{aligned} \Phi(\alpha^1(x)) &= (0, 0, 2)^T, \\ \Phi(\alpha^2(x)) &= (0, 0, 2)^T, \\ \Phi(\alpha^3(x)) &= (0, 2, 3)^T, \\ \Phi(\alpha^4(x)) &= (0, 2, 3)^T; \\ \Phi(\alpha^1(y)) &= (0, 0, 0)^T, \\ \Phi(\alpha^2(y)) &= (0, 0, 2)^T, \\ \Phi(\alpha^3(y)) &= (0, 0, 2)^T, \\ \Phi(\alpha^4(y)) &= (0, 2, 3)^T. \end{aligned}$$

Проверяем систему (12.48):

$$\begin{aligned} j = 1: 0 &= 0, 0 = 0, 2 > 0; j = 2: \\ 0 &= 0, 0 = 0, 2 = 2; \\ j = 3: 0 &= 0, 2 > 0, 3 > 2; j = 4: 0 = \\ &= 0, 2 = 2, 3 = 3. \end{aligned}$$

Так как система неравенств (12.48) выполнена и имеются строгие неравенства, то $x \succ y$.

Достоинством алгоритма 2 является то, что он может быть сравнительно просто реализован в диалоговой системе анализа решений и не требует от ЛПР получения сильной информации.

Отношения предпочтения, построенные по информации Ω , на множестве стратегий U задаются правилом:

$$u \succsim_{\Omega} v \iff x(u) \succsim_{\Omega} x(v). \quad (12.49)$$

Поскольку информация Ω о предпочтениях ЛПР в общем случае не позволяет получить связного отношения нестрогого предпочтения, то можно говорить лишь о множестве нехудших стратегий, векторные оценки которых образуют ядро M_{Ω} .

Такие стратегии U_{Ω} называются омега-эффективными.

Так как отношение $\succsim_{\Omega} \equiv \succsim_{\omega_0}$, то множество омега-эффективных стратегий не шире множества эффективных стратегий, т. е. $U_{\Omega} \subseteq U_{\omega_0}$. Чем более содержательна информация Ω , тем, как правило, меньше неопределенность выбора.

ОСОБЫЕ СЛУЧАИ ОМЕГА-ИНФОРМАЦИИ О ПРЕДПОЧТЕНИЯХ

Особую практическую значимость имеют следующие частные случаи омега-информации об относительной важности: сообщения об одинаковой важности всех частных показателей $\Omega \equiv S = \{r \omega_e t; \forall r, t\}$; сообщения об абсолютном превосходстве в важности каждого частного показателя над остальными $\Omega \equiv L = \{i_1 \omega_B i_2, i_1 \omega_B i_3, \dots, i_1 \omega_B i_m, i_2 \omega_B i_3, i_2 \omega_B i_4, \dots, i_2 \omega_B i_m, \dots, i_{m-1} \omega_B i_m\}$; сообщения о запрете компенсации уменьшения меньших значений всех «равноценных» показателей за счет сколь угодно большого увеличения больших $\Omega \equiv SL$; сообщения о равноценности или превосходстве в важности одних групп показателей над другими [59]. Как и ранее, предполагается, что все частные показатели однородны.

Информация S об одинаковой важности всех показателей. Основанием для вывода о равноценности всех пока-

зателей эффективности может служить тот факт, что для характеристики любой стратегии $u \in U$ ЛПР (эксперт) считает достаточным перечислить в произвольном порядке m значений оценок частных показателей $x_1(u)$, $x_2(u)$, ..., $x_m(u)$, не указывая при этом, какое конкретное из чисел является значением того или иного показателя W_i . Например, если для характеристики качества конкретного образца продукции достаточно указать, какие он набрал баллы по m характеристикам, то все эти характеристики имеют одинаковую важность; если для характеристики плана развития предприятия на ближайшую пятилетку достаточно перечислить ожидаемые значения показателей выполнения плана, то показатели выполнения плана в каждом году пятилетки являются одинаково важными.

Ввести отношение предпочтения на множестве X_d (а следовательно, и на U) можно значительно проще по сравнению с приемами, описанными выше. Поскольку информация $\Omega \equiv S$, то в соответствии с общей методикой можно построить лишь опорные множества эквивалентных векторных оценок X^{ω_e} и Y^{ω_e} для произвольных $x, y \in X$, а затем последовательно сравнивать элементы опорных множеств между собой. Однако можно сразу указать, какие элементы из X^{ω_e} и Y^{ω_e} надо сравнивать. Такими элементами являются:

$$\begin{aligned} \varphi(x) &= (\varphi_1(x), \varphi_2(x), \dots, \varphi_m(x))^T; \\ \varphi(y) &= (\varphi_1(y), \varphi_2(y), \dots, \varphi_m(y))^T, \end{aligned} \quad (12.50)$$

где $\varphi_1(z) = \min z_i$, ..., $\varphi_m(z) = \max(z)$; z — либо x , либо y ; а z_i — i -я компонента векторной оценки z . Тогда

$$x \overset{S}{\geq} y \iff \varphi(x) \overset{\omega_0}{\geq} \varphi(y), \quad (12.51)$$

где $\overset{S}{\geq}$ есть либо $\overset{S}{\succ}$, либо $\overset{S}{\sim}$; а $\overset{\omega_0}{\geq}$ — либо $\overset{\omega_0}{\succ}$, либо $\overset{\omega_0}{\sim}$.

Пример 6. Сравнить по предпочтительности оценки $x = (8, 2, 6, 5)^T$, $y = (1, 3, 7, 2)^T$, $z = (3, 5, 3, 6)^T$,

используя информацию $\Omega \equiv S$. Строим векторные оценки по (12.50):

$$\varphi(x) = (2, 5, 6, 8)^T,$$

$$\varphi(y) = (1, 2, 3, 7)^T,$$

$$\varphi(z) = (3, 3, 5, 6)^T$$

Так как $\varphi_i(x) > \varphi_i(y)$, $i = \overline{1, 4}$, то $x \overset{S}{\succ} y$; аналогично устанавливаем, что $z \overset{S}{\geq} y$, и $x \overset{S}{\geq} z$ (несравнимы).

Так как отношение $\overset{S}{\geq}$ вводится через отношение $\overset{\omega_0}{\geq}$ Парето, то оно в общем случае несвязно и решение задачи необходимо искать только среди тех стратегий, векторные оценки которых составляют ядро отношения $M_S \overset{S}{\succ}$ строгого предпочтения $\overset{S}{\succ}$.

Стратегии, векторные оценки которых принадлежат M_S , называются S -эффективными.

Множество векторных оценок $M_S \overset{S}{\succ}$ выделяется на основании суждения о строгой предпочтительности оценки $x \in M_S$ и оценки y , которая принадлежит X_d , но не принадлежит M_S . Это

суждение основано на сравнении компонент вектор-функций $\varphi(x)$ и $\varphi(y)$. Различные оценки [например, $y = (1, 2, 3, 4, 5)^T$ и $z = (5, 4, 3, 2, 1)^T$] могут иметь одинаковые вектор-функции $\varphi(x) = \varphi(y)$. Следовательно, множество $M_S \overset{S}{\succ}$ должно вклю-

чать в себя такие оценки $x \in M_{\omega_0} \overset{S}{\succ}$, которые предпочтительнее всех других оценок $y, z \in M_{\omega_0} \overset{S}{\succ}$. Такая про-

верка на строгое предпочтение может быть формально проведена, если множество X_d расширить до его симметричной оболочки $\text{Sym } X_d$, найти ее эффективное подмножество и множество $M_S \overset{S}{\succ}$ определить как пересечение $M_{\omega_0} \overset{S}{\succ}$ с эффективным подмножеством $\text{Sym } X_d$.

Симметричная оболочка получается из X_d путем добавления к нему векторов y , полученных из $x \in X_d$ произвольной перестановкой компонент. В этом случае M_S формально задается следующим образом:

$$M_S = \left\{ x \mid \exists y \in \text{Sym } X_d : y \succ^S x, x \in M_{\omega_0} \right\}. \tag{12.52}$$

На рис. 8 для двумерного случая представлены множества: X_d (заштрихованная область), $\text{Sym } X_d$ (граница его изображена штриховой линией), M_{ω_0} (жирная линия), M_S (жирная пунктирная линия). Штрихпунктирной линией выделено подмножество, симметричное X_d относительно диагонали $x_1 = x_2$. На этом же рисунке отмечены точки $y \in M_S$, $z \sim^S y$ и $x \in M_S$, $x \succ^{\omega_0} z \sim^S y$.

Мощность P множества M_S для двумерного случая стремится к величине $\lim_{N \rightarrow \infty} P(M_S) = 0,5 \ln N < \lim_{N \rightarrow \infty} P \times (M_{\omega_0}) = \ln N$, где N — мощность множества допустимых стратегий, т. е. информация об одинаковой важности частных показателей снижает неопределенность в выборе решения, так как в общем случае $M_{\omega_0} \supset M_S$, т. е. $U_S \subset U_{\omega_0}$, где U_S — множество Ω -эффективных стратегий.

Информация L об абсолютном превосходстве по важности. Основанием для вывода об абсолютном превосходстве одних показателей над другими является то, что при предъявлении ЛПР некоторых оценок для их сравнения оно, прежде всего, обращает внимание на значение конкретного, вполне определенного показателя, а затем, если значение этого показателя у нескольких оценок оказалось одинаково, ЛПР обращает внимание на значение другого (также вполне определенного) показателя и т. д., т. е. степень различия отдельных показателей по важ-

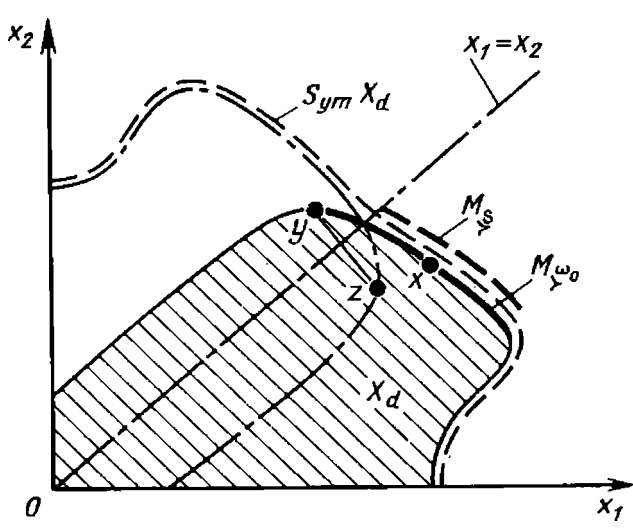


Рис. 8. Схема выделения ядра отношения

ности для ЛПР столь велика, что оно вначале сравнивает оценки только по самому важному показателю, не обращая внимания на остальные, потом только по второму и т. д.

Такая информация строго ранжирует показатели по важности. Это позволяет задать отношение предпочтения $\succ^L$ на множестве X_d следующим образом:

$$\begin{aligned} x \succ^L y \iff & (x_{r_1} > y_{r_1}) \vee (x_{r_1} = y_{r_1} \wedge x_{r_2} > y_{r_2}) \vee (x_{r_1} = y_{r_1} \wedge x_{r_2} = y_{r_2} \wedge x_{r_3} > y_{r_3}) \vee \dots \\ & \vee (x_{r_1} = y_{r_1} \wedge x_{r_2} = y_{r_2} \wedge x_{r_3} = y_{r_3} \wedge \dots \wedge x_{r_m} > y_{r_m}), \end{aligned} \tag{12.53}$$

где r_i — номер i -го показателя эффективности в ранжированном ряду; $\vee$ — логическое «ИЛИ»; $\wedge$ — логическое «И».

Отношение нестрогого по L -информации предпочтения $\succ^L$ порождает отношения эквивалентности $\sim^L$ и строгого предпочтения $\succ^L$. Если в (12.53) выполняются только m последних равенств, то отношение $\succ^L$ есть эквивалентность $\sim^L$.

Отношение $\succsim^L$ обладает свойством связности и не требует однородности частных показателей эффективности. Поэтому связными будут соответствующие отношения предпочтения во множестве стратегий U :

$$u \succsim^L v \iff x(u) \succsim^L x(v). \quad (12.54)$$

Связное отношение предпочтения $\succsim^L$ на множестве U позволяет сформировать функцию выбора и выделить подмножество наилучших стратегий $U_L \subseteq U$, которые называются *лексикографически оптимальными*. Поскольку все стратегии из U_L эквивалентны, то в качестве решения можно принять любую из них или привлекать дополнительную информацию при необходимости. Задачи такого типа называются *задачами лексикографической оптимизации*.

Пример 7. Необходимо сравнить по предпочтительности планы проведения спасательных работ на море. В качестве показателей эффективности плана выбраны: W_1 — вероятность спасения потерпевших за время, не превышающее срок их выживания в экстремальных условиях; W_2 — число привлекаемых средств специальной спасательной службы; W_3 — число иных привлекаемых к спасению средств.

ЛПР ранжирует указанные показатели следующим образом: W_1, W_3, W_2 , т. е. ЛПР, прежде всего, заинтересовано в увеличении показателя W_1 , невзирая на затраты средств, а если планы одинаковы по вероятности спасения потерпевших, то ЛПР стремится не привлекать иных средств, кроме спасательных.

Пусть представленные планы u_1, u_2, u_3, u_4, u_5 характеризуются следующими значениями показателей: $x(u_1) = (0,92; 15; 1)^T$, $x(u_2) = (0,92; 8; 6)^T$, $x(u_3) = (0,8; 4; 0)^T$, $x(u_4) = (0,915; 5; 1)^T$, $x(u_5) = (0,915; 6; 0)^T$. Сравнение планов по самому важному показателю приводит к выделению стратегий u_1 и u_2 . Сравнение недоминируемых по первому показателю планов u_1, u_2 требует отдать предпочтение плану u_1 .

Информация $\Omega \equiv L$ является сильной в том смысле, что для дискретных множеств стратегий U она дает возможность практически всегда выделять единственное решение.

Часто применяют прием сведения задачи принятия решения с векторным показателем эффективности, компоненты которого различаются по важности¹, к лексикографической задаче. В этом случае на значения частных показателей накладывают ограничения по минимально допустимым для них уровням $W_i \geq W_i^{\min}$, $i = \overline{1, m}$, а сами показатели ранжируют. Пусть самый важный показатель есть W_1 , затем $W_2, \dots$, а наименее важным является W_m . Тогда вначале стремятся достигнуть величины $W_1^{\min}$ по первому показателю (или немного превысить его, $W_1 = W_1^{\min} + \epsilon$), затем — величины $W_2^{\min}$ по второму показателю при условии, что $W_1 > W_1^{\min}$ и т. д. Этот прием формально соответствует преобразованию исходного показателя W с компонентами W_i , $i = \overline{1, m}$ в новый показатель $\tilde{W}$ с компонентами $\tilde{W}_i = \min\{W_i, W_i^{\min}\}$. Если для каждого из частных показателей помимо минимально допустимого уровня $W_i^{\min}$ можно указать и вполне удовлетворительное его значение W_i^{SAT} , то после достижения каждым из показателей минимального уровня $W_i^{\min}$ можно стремиться к поочередному их увеличению к уровню W_i^{SAT} . В этом случае вместо исходного показателя W размерности m используется новый показатель размерности $2 \cdot m$ с компонентами $\tilde{W}_i$ и $\hat{W}_i$, где $\hat{W}_i = \min\{W_i, W_i^{\text{SAT}}\}$, $i = \overline{1, m}$, причем частные показатели $\tilde{W}_i$ лексикографически важнее частных показателей $\hat{W}_i$.

На рис. 9 показаны прием сведения исходной задачи с векторным показателем W , на компоненты W_i которого

¹ Различие по важности понимается не как лексикографическое, а в смысле определения 12.2.

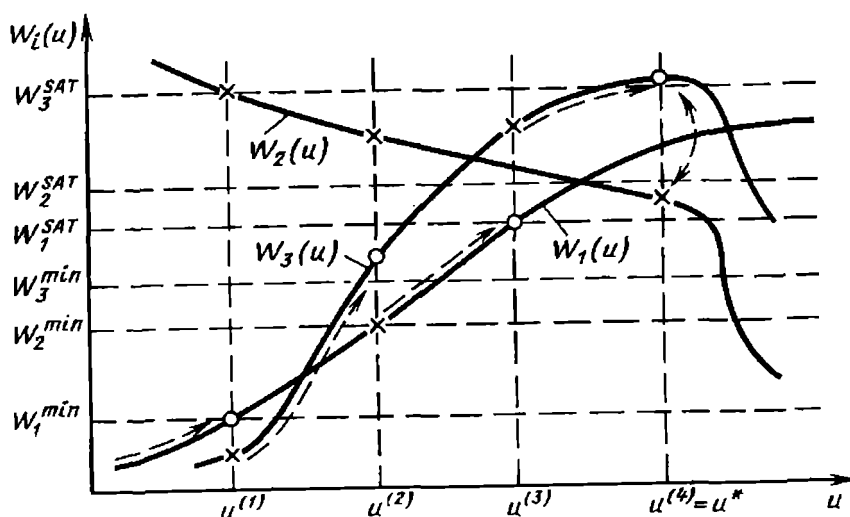


Рис. 9. Лексикографическая оптимизация

наложены ограничения по минимальному $W_i^{\min}$ и удовлетворительному W_i^{SAT} уровню, к задаче лексикографической оптимизации и последовательность получения решения u^* . При этом предполагается, что имеется информация Ω , позволяющая строго ранжировать исходные показатели в следующем порядке: W_1, W_2, W_3 . На первом шаге стратегия $u^{(1)}$ выбирается (движение показано штриховой линией со стрелкой) из условия обеспечения минимально допустимого уровня $W_1^{\min}$ по показателю W_1 ; при этом соответствующие значения $W_2(u^{(1)})$ и $W_3(u^{(1)})$ отмечены на рисунке крестиком. На втором шаге ищется стратегия $u^{(2)}$, обеспечивающая значение $W_3^{\min}$ показателя W_3 при условии, что $W_1 \geq W_1^{\min}, W_2 \geq W_2^{\min}$. Так как значение $W_2(u^{(2)}) > W_2^{\min}$, то на третьем шаге отыскивается стратегия $u^{(3)}$, обеспечивающая $W_1(u^{(3)}) = W_1^{\text{SAT}}$ при условии, что $W_2(u^{(3)}) \geq W_2^{\min}, W_3(u^{(3)}) \geq W_3^{\min}$. На четвертом шаге отыскивается стратегия $u^{(4)} = u^*$, доставляющая показателям W_1, W_2, W_3 следующие значения: $W_1(u^*) > W_1^{\text{SAT}}, W_2(u^*) = W_2^{\text{SAT}}, W_3(u^*) > W_3^{\text{SAT}}$, т. е. активным ограничением является

равенство W_2 его удовлетворительному значению (отмечено крестиком).

Из определения отношения нестро-
гого предпочтения $\succeq_L$ на множестве стратегий (12.54) с учетом правила (12.53) следует, что лексикографическая задача оптимизации

$$u^* : \text{lex max } W(u) \quad (12.55)$$

может быть декомпозирована на иерархию скалярных задач оптимизации типа (принято, что первый показатель важнее второго, второй — третьего и т. д.):

$$\begin{aligned} U_1^* &: \max_{u \in U} W_1(u); \\ U_2^* &: \max_{u \in U_1^*} W_2(u), \\ &\vdots \\ U_i^* &: \max_{u \in U_{i-1}^*} W_i(u); \end{aligned} \quad (12.56)$$

$$u^* : \max_{u \in U_{m-1}^*} W_m(u),$$

где U_i^* — множество стратегий, обеспечивающих максимум показателя W_i при условии, что значения показателей $W_j = W_j^*$, $j = \overline{1, i-1}$ (W_j^* — оптимальное значение j -го показателя); u^* — решение лексикографической задачи (12.55).

*1 Так как показатель $W_2(u^{(1)}) \gg W_2^{\min}$, а $W_3(u^{(1)}) < W_3^{\min}$, на втором шаге ведущим является показатель W_3 .

Последовательное решение задач оптимизации из системы (12.56) проводят до тех пор, пока для некоторого номера $i, i = \overline{1, m-1}$ мощность множества U_i^* не окажется равной единице. В этом случае последующие задачи оптимизации не решают, и $u^* = U_i^*$. Если последняя задача оптимизации имеет несколько решений U_m^* , эквивалентных между собой, то в качестве u^* выбирают либо любое из $u \in U_m^*$, либо проводят дополнительный содержательный анализ стратегий из U_m^* для установления предпочтений на них.

В общем случае задача (12.55) для множества U , обладающего мощностью континуума, может оказаться неустойчивой в том смысле, что небольшие изменения характеристик стратегии $u \in U$ приводят к выделению в качестве лексикографически оптимальных других стратегий, у которых значения первого показателя изменяются незначительно, а остальные — изменяются сильно. Поэтому в каждом конкретном случае необходимо учитывать эту особенность.

Лексикографические задачи с дискретным конечным множеством стратегий обычно устойчивы в указанном выше смысле. Для их решения используют следующий одношаговый метод [53]:

$$u^* : \max_{u \in U} \sum_{i=1}^m \gamma_i W_i(u), \quad (12.57)$$

где $\gamma_m > 0$ — произвольное положительное число;

$$\gamma_i > \frac{1}{\delta_i} \sum_{j=i+1}^m \gamma_j \Delta_j, \quad i = m-1, m-2, \dots, 2, 1; \quad (12.58)$$

δ_i — нижняя оценка минимально возможного приращения значения i -го показателя, определяемая из условия:

$$0 < \delta_i \leq \min \{ |W_i(u) - W_i(v)| \}, \quad u, v \in U, \quad W_i(u) \neq W_i(v); \quad (12.59)$$

Δ_i — верхняя оценка максимально возможного приращения значения i -го

показателя, определяемая из условия:

$$\Delta_i > \max_{u \in U} W_i(u) - \min_{u \in U} W_i(u). \quad (12.60)$$

Таким образом коэффициенты γ_i целевой функции в (12.57) выбирают из условия обеспечения минимально допустимой «скорости» их роста от γ_m к γ_1 .

Пример 8. Для условий примера 7 определить значения коэффициентов $\gamma_i, i = \overline{1, 3}$ и найти u^* .

Учитывая, что показатели W_2 и W_3 желательно минимизировать, проведем их эквивалентное в смысле задачи (12.57) преобразование: $\tilde{W}_2 = -W_2, \tilde{W}_3 = -W_3$.

Выберем $\gamma_2 = 1$, так как показатель W_2 в ранжированном ряду стоит последним.

Определим δ_i :

$$\begin{aligned} 0 < \delta_1 &\leq \min \{ |0,92 - 0,8|, |0,92 - 0,915|, |0,915 - 0,81| \} = \\ &= 0,05, \text{ примем } \delta_1 = 0,01; \\ 0 < \delta_2 &\leq \min \{ |-15 + 8|, |-15 + 4|, |-15 + 5|, |-15 + 6|, \\ &|-8 + 4|, |-8 + 5|, |-8 + 6|, |-4 + 5|, \\ &|-4 + 6|, |-5 + 6| \} = 1, \text{ примем } \delta_2 = 1; \\ 0 < \delta_3 &\leq \min \{ |-1 + 6|, |-1 + 0|, |-6 + 0| \} = 1, \\ &\text{примем } \delta_3 = 1. \end{aligned}$$

Определим Δ_i :

$$\begin{aligned} \Delta_1 &\geq \max \{ 0,92; 0,8; 0,915 \} - \min \{ 0,92; 0,8; 0,915 \} = \\ &= 0,92 - 0,8 = 0,12, \text{ примем } \Delta_1 = 0,2; \\ \Delta_2 &\geq \max \{ -15, -8, -4, -5, -6 \} - \min \{ -15, -8, -4, -5, -6 \} = \\ &= -4 + 15 = 11, \text{ примем } \Delta_2 = 20; \\ \Delta_3 &\geq \max \{ -1, -6, 0 \} - \min \{ -1, -6, 0 \} = 0 + 6 = 6, \\ &\text{примем } \Delta_3 = 10. \end{aligned}$$

Величины:

$$\gamma_3 > \frac{1}{\delta_3} (\gamma_2 \cdot \Delta_2) = 1 \cdot 1 \cdot 20 = 20,$$

примем $\gamma_3 = 25$;

$$\begin{aligned} \gamma_1 &> \frac{1}{\delta_1} (\gamma_2 \Delta_2 + \gamma_3 \Delta_3) = \\ &= 100 (1 \cdot 20 + 25 \cdot 10) = 27 \cdot 10^3, \end{aligned}$$

примем $\gamma_1 = 3 \cdot 10^4$.

Значения целевой функции $\varphi(u_j) =$

$$= \sum_{i=1}^3 \gamma_i W_i(u_j), \quad j = \overline{1, 5}:$$

$$\varphi(u_1) = 3 \cdot 10^4 \cdot 0,92 + 1(-15) + 25(-1) = 27\,560;$$

$$\varphi(u_2) = 3 \cdot 10^4 \cdot 0,92 + 1(-8) + 25(-6) = 27\,442;$$

$$\varphi(u_3) = 3 \cdot 10^4 \cdot 0,8 + 1(-4) + 25 \cdot 0 = 23\,996;$$

$$\varphi(u_4) = 3 \cdot 10^4 \cdot 0,915 + 1(-5) + 25(-1) = 27\,420;$$

$$\varphi(u_5) = 3 \cdot 10^4 \cdot 0,915 + 1(-6) + 25 \cdot 0 = 27\,444.$$

Таким образом, $\max_{u_j \in U} \varphi(u_j) = 27\,560$

и $u^* = u_1$.

В задачах проектирования технических систем иногда заранее неясно, можно ли удовлетворить одновременно всем ограничивающим требованиям технического задания. Например, требуется спроектировать техническое устройство максимальной производительности (единственный показатель $W = \varphi(X)$) при ограничениях на его массу $q_1(X) \leq 0$, габариты $h_1(X) = 0$, уровень звука $q_2(X) \leq 0$, где X — конструкционные параметры устройства. Заранее нельзя сказать, совместна ли система ограничений:

$$\left. \begin{aligned} q_1(X) &\leq 0; \\ h_1(X) &= 0; \\ q_2(X) &\leq 0. \end{aligned} \right\}$$

В этом случае задачу проектирования X^* : $\max \varphi(X)$;

$$q_i(X) \leq 0, \quad i = \overline{1, l};$$

$$h_j(X) = 0, \quad j = \overline{1, k} \quad (12.61)$$

целесообразно рассматривать в лексикографической постановке, стремясь вначале минимизировать возможную невязку:

$$\begin{aligned} f(X) &= - \sum_{i=1}^l \min(0, -q_i(X)) + \\ &+ \sum_{j=1}^k |h_j(X)|, \end{aligned} \quad (12.62)$$

а затем на множестве X^0 : $\min f(X)$ искать $\varphi(X)$.

На рис. 10, а для одномерного случая показано решение задачи проектирования в лексикографической постановке. Сплошными линиями изображены графики функций $\varphi(x)$, $q_1(x)$, $h(x)$, $q_2(x)$. Штрихпунктирной линией изображен график функции невязки (12.62). Область $X^0 = [x_1^0, x_2^0]$, соответствующая минимуму $f(X)$, выделена штриховой линией.

Решением задачи является $x^* \in [x_1^0, x_2^0]$. Функция $f(X)$ может и не иметь участков, соответствующих ее нулевому значению (рис. 10, б), т. е. система ограничений задачи (12.61) несовместна. В этом случае множество $X^0 = \{x : \min f(x)\}$ будет обеспечивать наименьшую невязку, а $x^* \in X^0$ — максимальное значение целевой функции на множестве X^0 . Рис. 10, б и в иллюстрируют некоторые отрицательные стороны лексикографической постановки задачи оптимизации. Например, в точке x^* на рис. 10, б достигается минимум функции невязки, а целевая функция имеет в этой точке значение, отмеченное звездочкой. Лексикографическая постановка не позволяет использовать в качестве решения точку $\tilde{x}^*$, в которой $\varphi(\tilde{x}^*) \gg \varphi(x^*)$, хотя разница между значениями функции невязки в этих точках может оказаться сколь угодно малой: $f(\tilde{x}^*) - f(x^*) \leq \varepsilon$. Если задача векторная [техническое устройство характеризуется, например, двумя показателями $W_1 = \varphi_1(x)$, $W_2 =$

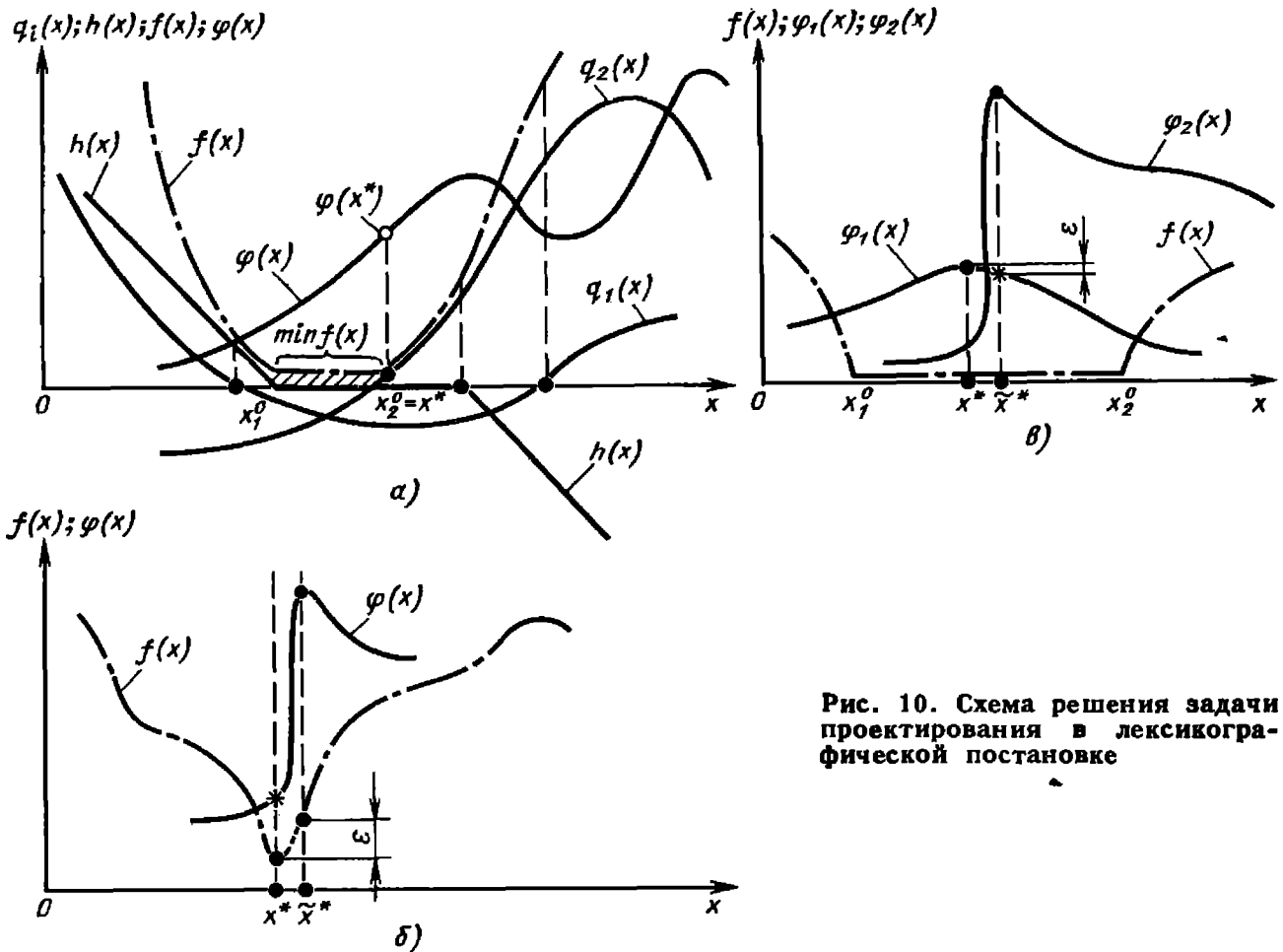


Рис. 10. Схема решения задачи проектирования в лексикографической постановке

$= \varphi_2(x)$ (см. рис. 10, в)], то даже при выполнении всех ограничений ($f(x) = 0$) назначение показателя W_1 в качестве самого важного не позволяет использовать в качестве решения точку $\tilde{x}^*$, в которой достигается наибольшее значение второго показателя, в то время как значение первого снижается лишь на величину ω , которая может быть сколь угодно мала.

Информация SL о недопустимости компенсации уменьшения меньших значений одинаково важных показателей сколь угодно значительным увеличением больших. Основанием для вынесения суждения о запрете компенсации уменьшения меньших компонент векторного показателя сколь угодно большим увеличением других компонент может служить тот факт, что при сравнении по предпочтительности двух произвольных стратегий, характеризуемых наборами равноценных частных показателей, ЛПР, прежде всего, обращает внимание на компоненты, имеющие самые малые значения. Если у векторных оценок самые малые значения компонент равны, то

ЛПР принимает во внимание следующие по величине компоненты и т. д.

Таким образом, информация SL является более сильной, чем информация о равноценности всех показателей, и более простой, чем информация L . Эта информация является достаточно полной для выделения единственного решения, поскольку на множестве векторных оценок отношение строгого предпочтения задается следующим образом:

$$x \succcurlyeq y \iff (\varphi_1(x) > \varphi_1(y)) \vee (\varphi_1(x) = \varphi_1(y) \wedge \varphi_2(x) > \varphi_2(y)) \vee (\varphi_i(x) = \varphi_i(y), i = \overline{1, j} \wedge \varphi_{j+1}(x) > \varphi_{j+1}(y), j = \overline{2, m}), \quad (12.63)$$

где $\varphi_i(x)$, $\varphi_j(y)$ задаются (12.50).

Если в (12.63) все неравенства выполняются как равенства, т. е. $\varphi_i(x) = \varphi_i(y)$, $i = \overline{1, m}$, то оценки x и y SL -эквивалентны. Объединение отношений строгого предпочтения и экви-

валентности приводит к отношению нестрогого предпочтения:

$$SL \quad SL \quad SL \quad SL \quad SL \\ \succsim = \sim \cup \succ, \quad \sim \cap \succ = \emptyset.$$

Поскольку из (12.63) следует, что

$x \succsim y \Leftrightarrow \varphi(x) \succsim \varphi(y)$, т. е. для его задания используется лексикографическое отношение (12.53), то симметрически лексикографическое отношение

нестрогого предпочтения $\succsim^{SL}$ оказывается связным и позволяет сформировать функции выбора, а следовательно, — выделить подмножество наилучших стратегий $U_{SL} \subseteq U$, т. е.

$$U_{SL} = \left\{ u \mid u \succsim^{SL} v \Leftrightarrow x(u) \succsim^{SL} x(v) \right\}. \quad (12.64)$$

Симметрически лексикографические задачи принятия решения являются обобщением максиминных задач принятия решения для случая векторного показателя эффективности. Это следует из того, что решение симметрически лексикографической задачи оптимизации

$$u^* : \text{lex max } \varphi(W(u)) \quad (12.65)$$

можно свести к последовательности m максиминных задач:

$$\left. \begin{aligned} U_1^*, i_1^* : \max_{u \in U} \min_{i \in I} W_i(u), \\ I = \{1, 2, 3, \dots, m\}; \\ U_2^*, i_2^* : \max_{u \in U_1^*} \min_{i \in I_1} W_i(u), \\ I_1 = I \setminus i_1^*; \\ U_3^*, i_3^* : \max_{u \in U_2^*} \min_{i \in I_2} W_i(u), \\ I_2 = I \setminus i_2^*; \\ \vdots \\ U_{m-1}^*, i_{m-1}^* : \max_{u \in U_{m-2}^*} \min_{i \in I_{m-2}} \\ \times W_i(u), \quad I_{m-2} = I_{m-3} \setminus i_{m-3}^*; \\ u^* : \max_{u \in U_{m-1}^*} \min_{i \in I_{m-1}} W_i(u), \\ I_{m-1} = I_{m-2} \setminus i_{m-2}^*. \end{aligned} \right\} \quad (12.66)$$

Пример 9. Для условий примера 6 и информации $\Omega \equiv SL$ сравнить оценки x, y, z с оценками $a = (7, 3, 4, 3)^T$, $b = (8, 2, 3, 8)^T$ и $c = (3, 3, 4, 5)^T$

Множеству U_1^* соответствуют векторные оценки z, a, c , у которых $W_{i_1}^* = 3$.

Множеству U_2^* соответствуют оценки z, a, c , у которых $W_{i_2}^* = 3$. Множество U_3 состоит из единственной стратегии u^* , которой соответствует оценка z .

Дискретная симметрически лексикографическая задача может быть решена одношаговым методом, аналогичным методу, задаваемому (12.57)—(12.60):

$$u^* : \max_{u \in U} \sum_{i=1}^m \gamma_i \varphi_i(W(u)), \quad (12.67)$$

где $\gamma_m > 0$ — произвольное положительное число;

$$\gamma_i > \frac{\Delta}{\delta} \sum_{j=i+1}^m \gamma_j, \quad i = m-1, \\ m-2, \dots, 2, 1; \quad (12.68)$$

δ — нижняя оценка минимально возможного приращения любой из компонент функции $\varphi(W)$, т. е.

$$0 < \delta < \min \{ |W_i(u) - W_j(v)| \}, \\ u, v \in U, \quad W_i(u) \neq W_j(v); \quad (12.69)$$

Δ — верхняя оценка максимально возможного приращения любой из компонент функции $\varphi(W)$, т. е.

$$\Delta \geq \max_{u \in U} \max_{i \in I} W_i(u) - \\ - \min_{u \in U} \min_{i \in I} W_i(u), \\ I = \{1, 2, 3, \dots, m\}. \quad (12.70)$$

При числе компонент векторного показателя эффективности, большем четырех-пяти, вместо (12.67) удобнее

решать задачу в следующем виде [44, 72]:

$$u^* : \min_{u \in U} \sum_{i=1}^m (\bar{\Delta} - W_i(u))^P, \quad (12.71)$$

$$\text{где } \bar{\Delta} \geq \max_{u \in U} \max_{i \in I} W_i(u); \quad (12.72)$$

$$P \geq \frac{\ln m}{\ln \bar{\Delta} - \ln (\bar{\Delta} - \delta)}. \quad (12.73)$$

Если определенное по (12.72) значение $\bar{\Delta} = \delta$, то (12.73) не определено и величину $P > 0$ можно брать любой (удобнее брать $P = 1$).

Пример 10. В условиях примера 8 определить параметры для задачи (12.71) и (12.67).

Найдем параметры задачи (12.71):

$$\begin{aligned} \bar{\Delta} &\geq \max_{u \in U} \max_{i \in I} W_i(u) = \\ &= \max_{u \in U} \{8, 7, 6, 7, 8, 6\} = 8, \end{aligned}$$

примем $\bar{\Delta} = 10$;

$$\delta < \min \{ |W_i(u) - W_j(v)| \} = 1,$$

примем $\delta = 0,1$;

$$P \geq \frac{\ln 4}{\ln 10 - \ln (10 - 0,1)} = 138,6,$$

примем $P = 150$.

Параметры задачи (12.67):

$$\begin{aligned} \Delta &\geq \bar{\Delta} - \min_{u \in U} \min_{i \in I} W_i(u) = \\ &= 8 - \min \{2, 1, 3, 3, 2, 3\} = \\ &= 8 - 1 = 7, \text{ примем } \Delta = 10; \end{aligned}$$

$$\gamma_4 = 1; \quad \gamma_3 > \frac{10}{0,1} \cdot 1 = 100,$$

примем $\gamma_3 = 200$;

$$\gamma_2 > \frac{10}{0,1} (1 + 200) = 20\,100,$$

примем $\gamma_2 = 3 \cdot 10^4$;

$$\gamma_1 > \frac{10}{0,1} (1 + 200 + 3 \cdot 10^4) = 3\,020\,100,$$

примем $\gamma_1 = 4 \cdot 10^6$.

Информация о равноценности или превосходстве в важности одних групп показателей над другими. Определе-

ния 1 и 2 естественным образом распространяются на непересекающиеся группы показателей.

Определение 3. Непересекающиеся группы показателей с номерами из множеств:

$$\{r_i\}, \quad i \in I \quad \text{и} \quad \{t_j\}, \quad j \in J,$$

$$I \cup J \subseteq \{1, 2, \dots, m\},$$

$I \cap J = \emptyset$ имеют одинаковую важность («равноценны»), если эквивалентными по предпочтительности являются оценки x и $z = \bar{x}^{\{r_i\}\{t_j\}}$, где

$$x_{r_i} = c, \quad i \in I, \quad \bar{x}_{t_j} = b, \quad j \in J, \quad \text{а}$$

$\bar{x}^{\{r_i\}\{t_j\}}$ — вектор, полученный из x заменой каждой компоненты x_{r_i} , $i \in I$

на величину b , а каждой компоненты x_{t_j} , $j \in J$ на величину c .

Качественная информация, полученная на основе определения 3, обозначается символом $\{r_i\} \omega_e \{t_j\}$.

Например, если $\Omega = \{1, 3, 5\} \times \times \omega_e \{2, 4\}$, то эквивалентной оценкой для $x = (4, 2, 4, 2, 4, 7)^T$ будет оценка $z = (2, 4, 2, 4, 2, 7)^T$.

Определение 4. Группа показателей с номерами из множества $\{r_i\}$, $i \in I$ «важнее» непересекающейся с ней группы показателей с номерами $\{t_j\}$, $j \in J$, если векторная оценка x , имеющая компоненты $x_{r_i} > x_{t_j}$, $x_{r_i} = c$, $i \in I$, $x_{t_j} = b$, $j \in J$, предпочтительнее оценки $z = \bar{x}^{\{r_i\}\{t_j\}}$.

Качественную информацию, полученную на основе определения 4, обозначим символом $\{r_i\} \omega_B \{t_j\}$.

На основании определений 3, 4 вводится отношение нестрогого и строгого предпочтения:

$$\begin{aligned} &\text{по информации } \{r_i\} \omega_e \{t_j\} \\ &^{\Omega} x \succsim y \iff (x \succsim^{\omega_e} y) \vee \left(\min_{i \in I} x_{r_i} \geq \right. \end{aligned}$$

$$\left. \geq \max_{j \in J} y_{t_j} \wedge \min_{j \in J} x_{t_j} \geq \max_{i \in I} y_{r_i} \wedge \right.$$

$$\left. \wedge x_{l_k} \geq y_{l_k}, \right.$$

$$k \in \{1, 2, \dots, m\} \setminus (I \cup J)); \quad (12.74)$$

$$^{\Omega} x \sim y \iff (x = y) \vee (y = x^{\{r_i\}\{t_j\}});$$

по информации $\{r_i\} \omega_B \{t_j\}$

$$\begin{aligned} x \succ_{\Omega} y &\Leftrightarrow (x \succ_{\omega_0} y) \vee \left(\min_{i \in I} x_{r_i} \geq \right. \\ &\geq \max_{j \in J} y_{t_j} \wedge \min_{i \in J} x_{t_i} \geq \max_{i \in I} y_{r_i} \wedge \\ &\wedge x_{l_k} \geq y_{l_k}, \\ &k \in \{1, 2, 3, \dots, m\} \setminus (I \cup J) \wedge \\ &\wedge \left(\min_{i \in I} x_{r_i} > \max_{i \in I} y_{r_i} \right). \end{aligned} \quad (12.75)$$

Если информация Ω объединяет сообщения двух указанных типов $\Omega = \Omega_e \cup \Omega_B$, $\Omega_e = \{\{r_i\} \omega_B \{t_j\}, \dots, \{r_k\} \omega_e \{t_l\}\}$; $\Omega_B = \{\{r_{i_1}\} \omega_B \{t_{j_1}\}, \dots, \{r_{k_1}\} \omega_B \{t_{l_1}\}\}$, причем множества номеров, входящих в сообщения Ω_e и Ω_B , не пересекаются, то отношение $\succ_{\Omega}$ на векторных оценках x, y истинно в том случае, когда

$$\left. \begin{aligned} x(j) \succ_{\Omega(j)} y(j), \quad j = \overline{1, j_0}; \\ x(m \setminus j_0) \succ_{\omega_0} y(m \setminus j_0), \end{aligned} \right\} \quad (12.76)$$

где $x(j)$ — вектор, полученный из x отбрасыванием тех компонент, о номерах которых нет j -го сообщения $\Omega(j)$ о равноценности или превосходстве в важности; $\Omega(j)$ — элемент множества Ω_e , либо множества Ω_B ; j_0 — число сообщений типа ω_e, ω_B в множестве Ω ; $x(m \setminus j_0)$ — вектор, полученный из x отбрасыванием всех компонент с номерами, о которых есть сообщения из Ω .

Пример 11. Сравнить векторные оценки $x = (6, 3, 2, 5, 1, 4)^T$ и $y = (2, 1, 6, 3, 2, 6)^T$ по предпочтительности по информации

$$\Omega = \{\{2\} \omega_B \{5\}, \{1\} \omega_e \{3, 6\}\}.$$

Декомпозируем задачу на три подзадачи:

1. Установить истинность $x(1) \succ_{\Omega(1)} y(1)$, где

$$\begin{aligned} x(1) &= (3, 1)^T, \quad y(1) = (1, 2)^T, \\ \Omega(1) &= \{\{2\} \omega_B \{5\}\}. \end{aligned}$$

2. Установить истинность $x(2) \succ_{\Omega(2)} y(2)$, где

$$\begin{aligned} x(2) &= (6, 2, 4)^T, \quad y(2) = (2, 6, 6)^T, \\ \Omega(2) &= \{\{1\} \omega_e \{3, 6\}\}. \end{aligned}$$

3. Установить истинность $x(m \setminus j_0) \succ_{\omega_0} y(m \setminus j_0)$, где

$$x(m \setminus j_0) = 5, \quad y(m \setminus j_0) = 3.$$

Используя (12.75) без предпоследнего логического условия (это условие будет проверяться в задаче 3), получаем

$$\begin{aligned} x(1) \succ_{\omega_0} y(1) \vee \min \{3\} &= \\ = 3 \geq \max \{2\} = 2 \wedge \min \{1\} &= 1 = \\ = \max \{1\} \wedge \min \{3\} &= \\ = 3 > \max \{1\} = 1, \end{aligned}$$

следовательно, $x(1) \succ_{\Omega(1)} y(1)$. Аналогично по (12.74) без последнего логического условия получим

$$\begin{aligned} x(2) \succ_{\Omega(2)} y(2) \vee \min \{6\} &= 6 = \\ = \max \{6, 6\} = 6 \wedge \min \{2, 4\} &= 2 = \\ = \max \{2\} = 2, \end{aligned}$$

следовательно, $x(2) \succ_{\Omega(2)} y(2)$.

Сравнение по Парето $x(m \setminus j_0)$ и $y(m \setminus j_0)$ дает $x(m \setminus j_0) \succ_{\omega_0} y(m \setminus j_0)$. Так как условие (12.76) выполнено,

то $x \succ_{\Omega} y$. Если бы оказалось, что $x \not\succ_{\Omega} y$, то нужно было бы проверить истинность $y \succ_{\Omega} x$.

Мощность P множества M_{Ω} , являющегося ядром отношения строгого предпочтения по информации Ω об относительной важности групп частных показателей, значительно выше мощности множеств M_{ω_0} и M_S . Это объясняется тем, что сложные правила установления предпочтительности (12.74)–(12.76) в среднем нарушаются чаще, чем в случае информации о взаимной независимости по предпочте-

нию (она входит уже в условия) и информации об одинаковой важности всех частных показателей. Для трехмерного случая асимптотические оценки мощности множества M_{Ω} , где

$$\Omega = \Omega_e = \{\{1, 2\} \omega_e \{3\}\} \text{ или } \Omega = \Omega_b = \{\{1, 2\} \omega_b \{3\}\} \text{ составляют:}$$

$$\lim_{N \rightarrow \infty} P(M_{\Omega_e}) = \frac{9}{4} \ln N - 0,7,$$

$$\lim_{N \rightarrow \infty} P(M_{\Omega_b}) = 3 \ln N - 1,6.$$

ОДНОШАГОВЫЕ МЕТОДЫ РЕШЕНИЯ ЗАДАЧ С ВЕКТОРНЫМ ПОКАЗАТЕЛЕМ ЭФФЕКТИВНОСТИ

Сущность одношаговых методов состоит в следующем: векторный показатель эффективности тем или иным способом преобразуется в скалярную целевую функцию, а затем решается задача оптимизации. При этом предполагается, что от ЛПР может быть получена вся необходимая информация для построения обобщенной функции. Эта информация является весьма «сильной», а получить ее достаточно сложно. Поэтому использование одношаговых методов может считаться корректным лишь в том случае, когда ЛПР четко представляет, к каким последствиям приводит использование того или иного вида функции агрегирования, т. е. какова допустимая с его точки зрения степень компенсации уменьшения значений одних показателей увеличением других при использовании данной агрегирующей функции.

Метод обобщенного (агрегированного) показателя. Рассмотрим наиболее часто используемые виды функции агрегирования.

Если из сущности задачи может считаться допустимым, что абсолютное уменьшение одного из показателей может быть компенсировано суммарным абсолютным увеличением других (показатели однородны), то в качестве функции агрегирования может быть принята аддитивная функция

$$\Phi(W) = \sum_{i=1}^m \gamma_i W_i, \quad (12.77)$$

где γ_i — коэффициент относительной важности частного показателя W_i .

Задачи такого типа встречаются часто и связаны с показателями прибыли, денежных или временных затрат по годам, этапам жизненного цикла и т. д.

Иногда допустимой может считаться не абсолютная, а относительная компенсация изменения значений одних показателей другими, т. е. ЛПР согласно на то, что суммарный уровень относительного снижения одних показателей эквивалентен суммарному уровню относительного увеличения остальных. Это приводит к мультипликативной функции агрегирования вида

$$\Phi(W) = \prod_{i=1}^m W_i^{\gamma_i}. \quad (12.78)$$

Использование обобщенного показателя (12.78) приводит к тому, что недостаточная величина одного частного показателя компенсируется избыточной величиной другого, а выделяемые решения обладают примерно одинаковыми уровнями значений частных показателей (при одинаковых их весах γ_i).

Если из сущности задачи следует полная недопустимость компенсации значений одних показателей другими, т. е. требуется обеспечить равномерное «подтягивание» всех показателей к их наилучшему уровню, то используют агрегирующую функцию следующего вида:

$$\Phi(W) = \min_i \left\{ \frac{W_i}{\gamma_i} \right\}, \quad (12.79)$$

$$\gamma_i \neq 0, \quad i = \overline{1, m}.$$

Такой показатель используется в задачах планирования по «узкому месту».

Общим случаем функции агрегирования является средняя степенная функция

$$\Phi(W) = \left[\frac{1}{m} \sum_{i=1}^m W_i^p \right]^{\frac{1}{p}} \quad p \neq 0, \quad (12.80)$$

где показатель степени p отражает допустимую степень компенсации малых

значений одних равноценных показателей большими значениями других показателей (чем больше p , тем больше степень возможной компенсации).

Например, если $p \rightarrow -\infty$ (не допускается никакая компенсация и требуется равномерное «подтягивание»), то предельный вид агрегирующей функции (12.80) совпадает с (12.79); если $p \rightarrow 0$ (т. е. требуется обеспечение примерно одинаковых уровней частных показателей), то предельный вид агрегирующей функции совпадает с (12.78). При $p = 1$ получается аддитивная функция. Во всех этих случаях $\gamma_i = \frac{1}{m}$, $i = \overline{1, m}$. Важными частными случаями функции (12.80) являются:

$$\varphi(W) = \left(\frac{1}{m} \sum_{i=1}^m W_i^2 \right)^{\frac{1}{2}} \quad (12.81)$$

и эквивалентная ей

$$\varphi(W) = \sum_{i=1}^m W_i^2, \quad (12.82)$$

которые широко используются в задачах математической статистики и теории автоматического регулирования.

Если из существа задачи может считаться допустимым увеличение одного из показателей ценой любого уменьшения значений других (т. е. $p \rightarrow \infty$), то

$$\varphi(W) = \max_i \{\gamma_i W_i\}, \quad i = \overline{1, m}. \quad (12.83)$$

Функции агрегирования (12.78), (12.79), (12.83) могут быть использованы и для неоднородных показателей эффективности. В этом случае коэффициенты γ_i суть весовые размерные коэффициенты, показывающие «ценность» единицы i -го показателя. Однако получить размерные весовые коэффициенты γ_i трудно. Поэтому чаще всего используют однородные показатели и коэффициенты их относительной важности.

Если из существа задачи следует, что одни показатели желательно увеличивать, а другие — уменьшать, то иногда

используют функцию агрегирования в виде отношения одних показателей к другим, например,

$$\varphi(W) = \frac{\prod_{i=1}^{m_1} W_i}{\prod_{i=m_1+1}^m W_i}, \quad (12.84)$$

где $i = \overline{1, m_1}$ — номера показателей, значения которых желательно увеличивать, а $i = \overline{m_1+1, m}$ — уменьшать. Часто первая группа показателей отождествляется с целевым эффектом, а вторая — с затратами на его достижение. При этом показатели не обязательно должны быть однородными.

Метод «затраты—эффект». Особенности использования функции агрегирования (12.84) покажем на примере широко используемого метода «затраты — эффект». Каждый из рассматриваемых альтернативных вариантов $u \in U$ некоторой технической системы характеризуют векторным показателем W с компонентами: W_1 — полезный эффект; W_2 — затраты на создание системы и ее эксплуатацию. Тогда

показатель $\varphi(W) = \frac{W_1}{W_2}$ моделирует

предпочтения ЛПР, исходя из его стремления получить наибольший удельный эффект на единицу затрат. Такой показатель прост, логически непротиворечив, выделяет эффективное по Парето решение; однако он часто приводит к выделению альтернативы с низкими значениями частных показателей, т. е. «наилучшим» может считаться дешевый и низкоэффективный образец технической системы. На рис. 11 представлены векторные оценки $W(u) = (W_1(u), W_2(u))^T$ десяти образцов технической системы. Символами $\odot$ выделено эффективное по Парето множество альтернатив $U_{\omega_0} = \{u_1, u_3, u_6, u_7\}$. Так как большим значениям показателя (12.84) соответствует большее значение угла α , то в качестве наилучшего будет выделен образец u_1 . Чтобы исключить негативные стороны показателя (12.84) для отыскания наилучшей альтернативы, используют метод «затраты—эффект»

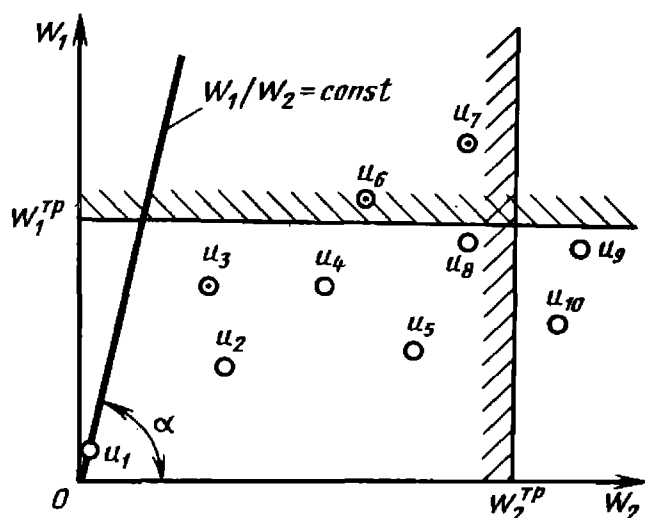


Рис. 11. Схема метода затраты—эффект

при ограничениях, сводя задачу выбора к одной из следующих задач оптимизации:

$$\left. \begin{aligned} u^* : \max_{u \in U} \varphi(W(u)); \\ W_1(u) \geq W_1^{TP}; \\ u^* : \max_{u \in U} \varphi(W(u)); \\ W_1(u) \geq W_1^{TP}, \\ W_2(u) \leq W_2^{TP}. \end{aligned} \right\} \quad (12.85)$$

В этом случае любая из задач системы (12.85) выделяет в качестве наилучшего вариант u_6 .

Метод целевого программирования. Основой этого метода является свертывание частных показателей эффективности в агрегирующую функцию, имеющую смысл расстояния до «идеальной» точки. Агрегирующая функция при этом задается метрикой определенного вида.

В качестве «идеальной» точки обычно выбирают либо физически недостижимую точку W^H с компонентами $W_i^H = \max_{u \in U} W_i(u)$, либо точку, отвечающую представлениям ЛПР об идеальном образце технической системы (перспективные уровни, мировые стандарты и т. п.).

Широкое применение получили следующие агрегирующие функции:

$$\begin{aligned} \rho(W, W^H) &= \varphi(W) = \\ &= \left[\frac{1}{m} \sum_{i=1}^m (W_i - W_i^H)^p \right]^{\frac{1}{p}}, \quad p \neq 0; \end{aligned} \quad (12.86)$$

$$\begin{aligned} \rho(W, W^H) &= \varphi(W) = \\ &= \sum_{i=1}^m |W_i - W_i^H|; \end{aligned} \quad (12.87)$$

$$\begin{aligned} \rho(W, W^H) &= \varphi(W) = \\ &= \sum_{i=1}^m \left[\frac{W_i^H - W_i}{W_i^H - W_i^{\min}} \right]^{p_i}; \end{aligned} \quad (12.88)$$

$$\begin{aligned} \rho(W, W^H) &= \varphi(W) = \\ &= \sum_{i=1}^m \gamma_i \frac{W_i^H - W_i}{W_i^H - W_i^{\min}}; \end{aligned} \quad (12.89)$$

$$\begin{aligned} \rho(W, W^H) &= \varphi(W) = \\ &= \max_i \{ \gamma_i |W_i^H - W_i| \}, \end{aligned} \quad (12.90)$$

где p — показатель, отражающий степень возможной компенсации малых значений одних показателей большими значениями других; p_i — весовой показатель степени; γ_i — коэффициент относительной важности.

Для функций (12.86), (12.87), (12.90) требуется однородность частных показателей.

Если в (12.86) $p = 2$, то вместо него часто используют функцию

$$\begin{aligned} \rho(W, W^H) &= \varphi(W) = \\ &= \left[\sum_{i=1}^m \gamma_i (W_i - W_i^H)^2 \right]. \end{aligned} \quad (12.91)$$

Задача отыскания наилучшей альтернативы в этом случае сводится к задаче оптимизации вида

$$u^* : \min_{u \in U} \rho(W(u), W^H).$$

Метод главного показателя. Если для некоторой области значений векторного показателя эффективности W ЛПР считает, что целевой эффект достигается в основном вследствие увеличения некоторого одного (так называемого главного) показателя, то исходная задача выбора может быть сведена к задаче оптимизации по единственному (главному) показателю при условии, что значения остальных — не ниже требуемых для них уровней:

$$u^* : \max_{u \in U} W_l(u) \tag{12.92}$$

$$W_i \geq W_i^{TP}, \quad i = \overline{1, m}, \quad i \neq l.$$

Выделяемая стратегия u^* является эффективной, что следует из теоремы 1.

Использование метода главного показателя для примера, изображенного на рис. 11, при условии, что $W_2 \leq W_2^{TP}$, приводит к выделению в качестве наилучшей стратегии альтернативы u_7 .

Использование рассмотренных методов в прямой постановке может привести к выделению неэффективной стратегии, если значения агрегирующей функции или значения главного показателя для нескольких стратегий одинаковы. Такая ситуация для метода главного показателя приведена на рис. 12. Стратегии, векторные оценки которых принадлежат отрезку XU , перпендикулярному оси OW_1 , имеют одинаковые значения главного показателя W_1 . Однако среди них эффективной является лишь стратегия с векторной оценкой X . Для того чтобы выделить такую стратегию, необходимо вместо задачи (12.92) решать задачу с целевой функцией, возрастающей по каждому показателю, т. е.

$$u_e^* : \max_{u \in U} \varphi^e(W(u)), \tag{12.93}$$

$$W_i \geq W_i^{TP}, \quad i = \overline{1, m}, \quad i \neq l,$$

где

$$\varphi^e(W(u)) = W_l(u) + \varepsilon \sum_{i=1}^m W_i(u). \tag{12.94}$$

Поскольку ε выбирают достаточно малым (на несколько порядков меньше значений показателей), то получаемое решение u_e^* практически не будет отличаться от решения u^* задачи (12.92).

Такая модификация проводится при решении всех задач, использующих одношаговые методы и функции агрегирования, т. е.

$$\varphi^e(W(u)) = \varphi(W(u)) + \varepsilon \sum_{i=1}^m W_i(u). \tag{12.95}$$

При реализации одношаговых методов главное внимание, таким образом, следует уделять выбору вида агрегируемой функции, или установлению пороговых значений «неглавных» показателей эффективности.

Методы, основанные на использовании функции эффективности $W_e(X)$, где X — оценки частных показателей эффективности, иногда относят к одношаговым методам. Однако, поскольку само построение функции эффективности представляет собой итерационный процесс, базирующийся на строгих аксиоматических предпосылках, методы, основанные на построении функции эффективности, будем относить к многошаговым.

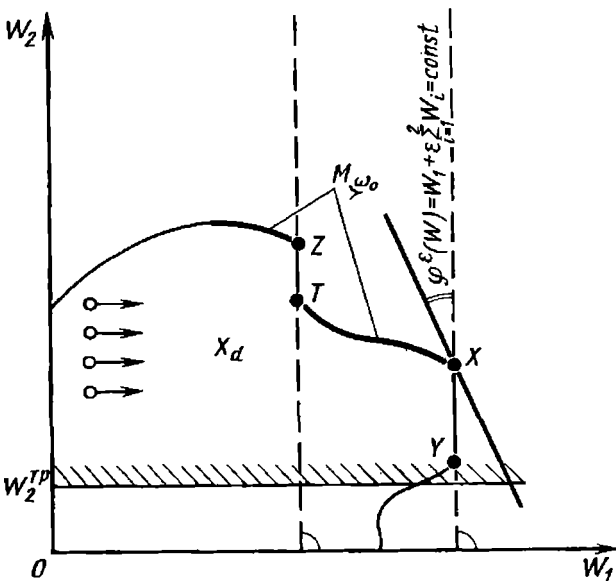


Рис. 12. Схема модификации метода главного показателя

**МНОГОШАГОВЫЕ МЕТОДЫ
РЕШЕНИЯ ЗАДАЧ С ВЕКТОРНЫМ
ПОКАЗАТЕЛЕМ ЭФФЕКТИВНОСТИ**

Многошаговые методы позволяют ослабить трудности решения задач с векторным показателем эффективности. Можно выделить два подхода. Первый состоит в декомпозиции исходной задачи на несколько более простых, при решении которых последовательно уточняют решение. Второй подход связан с последовательным усложнением получаемой информации или с последовательным решением задачи с использованием на каждом шаге однотипной информации. Оба подхода реализуются с помощью ЭВМ в интерактивном режиме.

К первой группе относится известный метод последовательных уступок [22]. В основу этого метода положена идея метода «главного» показателя.

На основании качественного анализа все частные показатели ранжируют и нумеруют в порядке убывания их важности (это можно сделать, например, на основе определения «один показатель важнее другого»). Затем максимизируют первый самый важный показатель и находят его наибольшее значение x_1^* . Исходя из практических соображений, назначают некоторую «уступку» Δ_1 от значения x_1^* , которую ЛПР согласно сделать, чтобы добиться увеличения значения второго по важности показателя W_2 до величины $x_2^* \times \times (\Delta_1)$. Снова назначают уступку Δ_2 по показателю W_2 , ценой которой стремятся увеличить значение показателя W_3 , и т. д. Формально этот

метод записывается следующей системой оптимизационных задач:

$$\left. \begin{aligned} x_1^* &= \max_{u \in U} W_1(u); \\ x_2^*(\Delta_1) &= \max_{u \in U} W_2(u); \\ W_1(u) &\geq x_1^* - \Delta_1 \\ x_3^*(\Delta_1, \Delta_2) &= \max_{u \in U} W_3(u); \\ W_1(u) &\geq x_1^* - \Delta_1 \\ W_2(u) &\geq x_2^* - \Delta_2 \\ \\ x_m^*(\Delta_1, \Delta_2, \dots, \Delta_{m-1}) &= \\ &= \max_{u \in U} W_m(u). \\ W_i(u) &\geq x_i^* - \Delta_i, \\ i &= \overline{1, m-1} \end{aligned} \right\} \quad (12.96)$$

Таким образом, уступки Δ_i назначают последовательно в результате анализа только попарной взаимосвязи показателей. Практически, выбирая уступку Δ_i для показателя W_i , ЛПР должно рассмотреть зависимость W_i от W_{i+1} (рис. 13). Анализ графика $W_1(W_2)$ показывает, что вначале даже незначительная уступка Δ_1^1 от значения x_1^* показателя W_1 приводит к существенному увеличению значения показателя W_2 , а затем с ростом уступки ($\Delta_1^1 < \Delta_1^2 < \Delta_1^3 < \dots$) маргинальное приращение в значениях показателя W_2 резко уменьшается. Сопоставляя выигрыш с потерями, ЛПР назначает уступку Δ_1 и определяет значение $x_2^*(\Delta_1)$.

Неформальное ранжирование показателей по важности позволяет ограничиваться назначением уступки предыдущего показателя только с учетом поведения последующего.

Поскольку вторая и последующие оптимизационные задачи (12.96) аналогичны формулировке метода главного показателя (12.92), метод последовательных уступок обладает всеми достоинствами и недостатками ме-

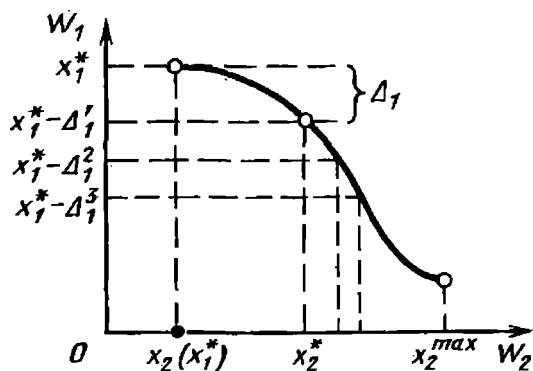


Рис. 13. Схема метода последовательных уступок

тогда главного показателя. Поэтому все задачи в системе (12.96), кроме первой, следует записывать в форме (12.93) с учетом (12.94).

Методы обобщенного, главного показателей и целевого программирования также можно применять при организации итеративных процедур решения векторных задач. При использовании метода обобщенного показателя вначале задают предварительный вид обобщенного показателя (как правило, в виде линейной свертки) и ориентировочные значения коэффициентов важности γ_i . Решается оптимизационная задача и отыскивается решение u^* . ЛПР анализирует полученные значения показателей $W_i(u^*)$ и при необходимости корректирует коэффициенты важности тех частных показателей, значения которых его не удовлетворяют. Если и вновь полученное решение неудовлетворительно (например, из-за недопустимо большой компенсации значений одних показателей значениями других), то подбирают другой вид функции агрегирования. Такую процедуру повторяют до тех пор, пока не будет получено приемлемое решение.

При использовании метода главного показателя обычно назначают серию пороговых значений W_i^{TP} , $i \neq l$, для каждой серии определяют величину $x_i^* = \max_{u \in U} W_i(u) \mid W_i(u) \geq W_i^{TP}$, $i \neq l$. Если величина x_i^* не удовлетворяет ЛПР, то уменьшают величины W_i^{TP} некоторых частных показателей и задачу решают вновь. Необходимость в изменении уровней W_i^{TP} возникает и тогда, когда система $W_i(u) \geq W_i^{TP}$, $i = \overline{1, m}$, $i \neq l$ оказывается несовместной. Иногда в качестве главного пробуют назначать другие частные показатели.

В практике особенно действенным оказывается совместное применение методов обобщенного и главного показателей. В этом случае сначала эвристически формируют область удовлетворительных решений $U^{SAT} = \{u \mid W_i \times (u) \geq W_i^{TP}, i = \overline{1, m}\}$, а затем, если

множество U^{SAT} не пусто, формируют подходящий вид агрегирующей функции в соответствии с рекомендациями метода обобщенного показателя (см. (12.77)–(12.83)) и решают оптимизационную задачу

$$u^* : \max_{u \in U^{SAT}} \Phi(W(u)). \quad (12.97)$$

Предварительное выделение множества удовлетворительных решений U^{SAT} вместо эвристического назначения вида агрегирующей функции позволяет использовать аксиоматические методы построения функции эффективности, которая будет адекватно отражать предпочтения ЛПР во множестве X^{SAT} удовлетворительных векторных оценок показателей эффективности W^* .

Итеративные методы построения функции эффективности. Функция $W_e(x)$, определенная на множестве векторных оценок X^{SAT} , называется *функцией эффективности*, если она упорядочивает по предпочтению любые векторные оценки $x, y \in X^{SAT}$, т. е.

$$x \geq y \Leftrightarrow W_e(x) \geq W_e(y).$$

При наличии функции эффективности задача выбора сводится к следующей оптимизационной задаче:

$$u^* : \max_{u \in U^{SAT}} W_e(x(u)).$$

Функцию эффективности строят на основе аксиоматического подхода. В большинстве случаев используют аддитивную форму представления функции эффективности:

$$W_e(x) = \sum_{i=1}^m f_i(x_i), \quad (12.98)$$

где x_i — оценка i -й компоненты векторного показателя эффективности;

*1 В достаточно малой области X^{SAT} пространства оценок X_d можно не опасаться, что будут нарушены допущения, в рамках которых может существовать функция эффективности, поскольку можно считать, что в этой области отношение предпочтения связно, транзитивно, асимметрично, а сами предпочтения не изменяются скачком.

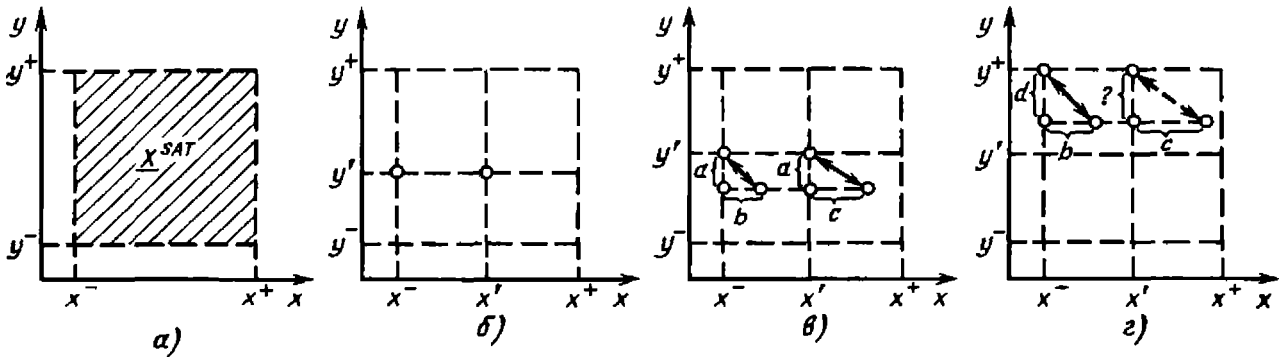


Рис. 14. Схема проверки условия соответственных замещений

$f_i(x_i)$ — частная функция эффективности по i -й компоненте векторного показателя.

Если частные показатели эффективности взаимонезависимы по предпочтению, число показателей не менее трех ($m \geq 3$), а их шкалы непрерывны, то функция эффективности может быть представлена в аддитивной форме. Для двумерной задачи ($m = 2$) необходимо выполнение условия соответственных замещений.

Проверка выполнения условия соответственных замещений может быть осуществлена с помощью следующего алгоритма.¹

Алгоритм 3 ($m = 2$). 1. Разбить области значений $[x^-, x^+]$ и $[y^-, y^+]$ на две примерно одинаковые части с границами x', y' .

2. Выбрать точки (x^-, y') и (x', y') .

3. Уменьшить значение y в каждой точке на фиксированную величину $a < y' - y^-$ (например, $a = 1$).

4. Получить информацию от ЛПР о величинах b, c — изменения значений x , компенсирующих уменьшение значения y так, чтобы выполнялись условия:

$$(x^-, y') \sim (x^- + b, y' - a);$$

$$(x', y') \sim (x' + c, y' - a).$$

5. Выбрать точки (x^-, y^+) и (x', y^+) .

6. Получить информацию от ЛПР о значении d уступки по оценке пока-

зателя y в точке (x^-, y^+) , компенсирующей увеличение значения x на величину b так, чтобы выполнялось условие $(x^-, y^+) \sim (x^- + b, y^+ - d)$.

7. При помощи ЛПР проверить выполнение условия

$$(x', y^+) \sim (x' + c, y^+ - d).$$

8. Если условие п. 7 выполняется, то для всей области X^{SAT} функция эффективности может быть задана в аддитивной форме². В противном случае следует установить подобласти пространства X^{SAT} , где это условие выполняется, и строить функцию эффективности для каждой из подобластей.

Последовательность выполнения действий по пп. 1—7 алгоритма иллюстрируется рис. 14, а—г.

Проверка взаимонезависимости частных показателей по предпочтению может быть осуществлена с помощью следующего алгоритма

Алгоритм 4 ($m \geq 3$). 1. Разбить все частные показатели на пары вида $\{W_1, W_i\}$, $i = \overline{2, m}$ и положить $i = 2$.

2. Проверить независимость пары $\{W_1, W_i\}$ от остальных показателей. Для этого выполнить следующие подпункты:

2.1. Зафиксировать каждый показатель W_j , $j \neq i$, $j = \overline{2, m}$ на соответствующем уровне x_j^* .

2.2. Разбить диапазоны значений показателей W_1 и W_i точками x_1' и x_i' , причем $x_1^- < x_1' < x_1^+$.

¹ Оценки по частному показателю W_1 обозначим через x , а по показателю W_2 — через y . При этом в области значений X^{SAT} векторного показателя эффективности наименее предпочтительное значение первой (второй) компоненты обозначим $x^-(y^-)$, а наиболее предпочтительное — $x^+(y^+)$.

² Более строго проверка по пп. 1—7 должна быть проведена не для двух, а трех-четырех участков значений оценок x и y .

2.3. Проверить условие соответственных замещений для оценок показателей W_1, W_i (п. 2—7 алгоритма 3);

2.4. Зафиксировать показатели W_j , $j \neq i$, $j = \overline{2, m}$ на соответствующих уровнях x_j^- и проверить условие соответственных замещений.¹

3. Положить $i = i + 1$ и проверить условие $i > m$: да — остановиться; нет — перейти к п. 2.

Для полной уверенности во взаимонезависимости частных показателей по предпочтению можно спросить ЛПР, как оно соотносит по предпочтению любые векторные оценки вида

$$X = (x_1, x_2, \dots, x_t,$$

$$y_{t+1}, y_{t+2}, \dots, y_m)^T,$$

$$Z = (z_1, z_2, \dots, z_t,$$

$$y_{t+1}, y_{t+2}, \dots, y_m)^T$$

и оценки $X' = (x_1, x_2, \dots, x_t, y'_{t+1}, y'_{t+2}, \dots, y'_m)^T$, $Z' = (z_1, z_2, \dots, z_t, y'_{t+1}, y'_{t+2}, \dots, y'_m)^T$. Если, например, из $X \geq Z$ всегда следует $X' \geq Z'$ при любых значениях компонент W_j , $j = \overline{t+1, m}$, то условие соответственных замещений выполнено.

В алгоритмах 3, 4 от ЛПР не требуется высокой точности ответов о величинах компенсирующих замещений a, b, c, d ; достаточно, если он назовет лишь их приблизительные значения.

Если с помощью алгоритмов 3 или 4 установлено, что частные показатели взаимонезависимы по предпочтению, то строится аддитивная функция эффективности. Так как функция эффективности задается с точностью до монотонного преобразования, то ее можно строить как в виде (12.98), так и в нормированном виде, когда ее значения изменяются в пределах от нуля до единицы:

$$W_e(x) = \sum_{i=1}^m \gamma_i f_i(x_i), \quad (12.99)$$

¹ Более строго проверка выполнения условия соответственных замещений должна проводиться для нескольких уровней x_j показателей W_j , $j \neq i$, $j = \overline{2, m}$.

где γ_i — шкалирующий коэффициент для i -й нормированной частной функции эффективности, удовлетворяющий условию:

$$\gamma_i > 0, \quad \sum_{i=1}^m \gamma_i = 1; \quad (12.100)$$

$f_i(x_i)$ — нормированная частная функция эффективности;

$$0 \leq f_i(x_i) \leq 1, \quad i = \overline{1, m}.$$

В соответствии с этим существуют два основных метода построения аддитивной функции эффективности — метод совместного шкалирования и метод половинного деления [30]. Первый из них целесообразно использовать при небольшой размерности задачи ($m \leq 3 \div 5$) применительно к функции эффективности в форме (12.98). Второй метод целесообразно использовать при числе частных показателей, большем трех-четырех, применительно к функции эффективности, задаваемой выражениями (12.99), (12.100). Этот метод хотя и экономнее первого, но требует более высокой подготовки ЛПР.

Метод совместного шкалирования может быть реализован с помощью следующего алгоритма.

Алгоритм 5. 1. Установить области изменения оценок по каждому из частных показателей W_i ; $[x_i^-, x_i^+]$, $i = \overline{1, m}$.

2. Положить $f_i(x_i^-) = 0$, $i = \overline{1, m}$.

3. Выбрать «небольшое» приращение a_1 по показателю W_1 ^{*1} и положить $f_1(y_1) = 1$, где $y_1 = x_1^- + a_1$.

4. Для каждого $j = \overline{2, m}$ при помощи ЛПР установить такие значения замещений $b_j^{(0)}$, для которых справедливо отношение

$$b_j^{(0)} : (x_1^-, x_2^-, \dots, y_j^{(0)}),$$

$$x_{j+1}^-, \dots, x_m^-)^T \sim$$

$$\sim (y_1, x_2^-, x_3^-, \dots, x_j^-, \dots, x_m^-)^T,$$

^{*1} Вместо показателя W_1 можно взять любой другой частный показатель.

где $\overset{0}{y}_j = x_j^- + b_j^{(0)}$; положить $f_j(y_j) = 1$.

5. Для каждого $j = \overline{1, m}$ выполнить следующие действия:

5.1. Положить $k = 1$ и для точки $(\overset{0}{y}_1, \overset{0}{y}_2, \overset{0}{y}_3, \dots, \overset{0}{y}_j, \dots, \overset{0}{y}_m)^T$ при помощи ЛПР определить значения компенсирующих замещений $b_j^{(k)}$, удовлетворяющих условиям:

$$b_j^{(k)} : (x_1^-, x_2^-, \dots, y_j^{(k)},$$

$$x_{j+1}^-, \dots, x_m^-)^T \sim$$

$$\sim (\overset{0}{y}_1, \overset{0}{y}_2, \dots, \overset{0}{y}_j, \dots, \overset{0}{y}_m)^T,$$

где $y_j^{(k)} = \overset{0}{y}_j + b_j^{(k)}$; положить

$$f(y_j^{(k)}) = 2.$$

5.2. Проверить условие $y_j^{(k)} \leq x_j^+$ и положить $k = k + 1$. Те компоненты W_j , для которых это условие не выполняется, исключают из дальнейшего рассмотрения.

5.3. Для точки $(\overset{0}{y}_1, \overset{0}{y}_2, \dots, y_j^{(k-1)}, \overset{0}{y}_{j+1}, \dots, \overset{0}{y}_m)^T$ при помощи ЛПР определить

$$b_j^{(k)} : (x_1^-, x_2^-, \dots, y_j^{(k)},$$

$$x_{j+1}^-, \dots, x_m^-)^T \sim$$

$$\sim (\overset{0}{y}_1, \overset{0}{y}_2, \dots, y_j^{(k-1)},$$

$$\overset{0}{y}_{j+1}, \dots, \overset{0}{y}_m)^T,$$

где $y_j^{(k)} = y_j^{(k-1)} + b_j^{(k)}$.

5.4. Положить $f(y_j^{(k)}) = k + 1$ и перейти к 5. 2.

6. Через полученные точки в пространстве $x_i 0 f_i(x_i)$, $i = \overline{1, m}$ провести плавные кривые частных функций эффективности и аппроксимировать их подходящими аналитическими функциями.

7. С помощью частных аналитических функций сформировать аналитическую зависимость для аддитивной функции эффективности вида (12.98).

Проиллюстрируем способ построения аддитивной функции эффективно-

сти на примере двумерного показателя эффективности.

Пример 12. Проектируемый самолет характеризуется векторным показателем W с компонентами: W_1 — дальность полета без дозаправки, W_2 — масса полезной нагрузки. Требуется построить функцию эффективности для принятия решения о выборе варианта проекта, который будет запущен в серию.

Так как частные показатели неоднородны, то предварительно осуществляется перевод их в единую безразмерную шкалу, например, с использованием выражения (12.37). Для эффективного множества недоминируемых проектов самолета значения показателей изменяются в процентах в пределах области $X^{\text{SAT}} = [x_1^-, x_1^+] \times [x_2^-, x_2^+]$, где $x_1^- = 60$, $x_1^+ = 120$, $x_2^- = 30$, $x_2^+ = 100$.

С помощью алгоритма 3 проверим выполнение условия соответственных замещений.

Разбиваем области значений показателей на две части граничными точками $x_1' = 80$, $x_2' = 70$. Выбираем величину $\alpha = 10$ и просим ЛПР указать, насколько необходимо увеличить значение первого показателя (дальность полета) для того, чтобы компенсировать уменьшение массы полезной нагрузки в точке $(60, 70)^T$ (рис. 15, а). ЛПР считает, что увеличение дальности примерно на 15 % эквивалентно снижению массы полезной нагрузки на 20 %. Эквивалентность точек $(60, 70)^T$ и $(75, 50)^T$ отмечена на рисунке двойной стрелкой.

Аналогичный вопрос задается ЛПР для точки $(80, 70)^T$ ЛПР считает, что в этой точке снижение массы полезной нагрузки на 20 % эквивалентно увеличению дальности полета на 20 %. Получаем точку $(100, 50)^T$.

Для точки $(60, 100)^T$ от ЛПР получена информация о том, что увеличение дальности полета на 15 % допустимо с его точки зрения в результате снижения грузоподъемности даже на 20 %, т. е. $(75, 80)^T \sim (60, 100)^T$.

ЛПР предлагают сравнить по предпочтительности точки $(80, 100)^T$ и

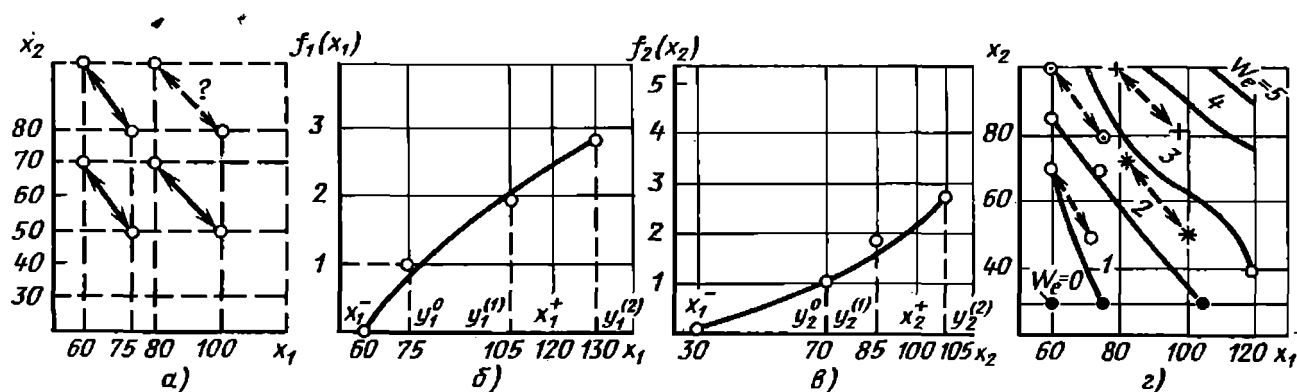


Рис. 15. Схема построения функции эффективности методом совместного шкалирования (случай двух показателей)

$(100, 80)^T$. ЛПР указывает, что эти точки приблизительно эквивалентны.

Таким образом, условие соответственных замещений выполнено и можно переходить к построению функции эффективности в аддитивной форме.

Для этого воспользуемся алгоритмом 5. Процесс построения функции эффективности приведен на рис. 15, б и в.

Полагаем $f_1(60) = f_2(30) = 0$. Выбираем величину $a_1 = 15$ и принимаем $f_1(75) = 1$. В соответствии с п. 4 алгоритма 5 при помощи ЛПР определяем величину замещения $b_2^{(0)}$, которая компенсирует уменьшение значения величины дальности с 75 % в точке $(75, 30)^T$ до 60 %. ЛПР считает, что $b_2^{(0)} \approx 40$ %. Принимаем $f_2(70) = 1$. Выполняем п. 5 алгоритма. Полагаем $k = 1$ и для точки $(75, 70)^T$ с помощью ЛПР определяем величину $b_1^{(1)}$, удовлетворяющую условию $(75 + b_1^{(1)}, 30)^T \sim (75, 70)^T$. ЛПР указало, что $b_1^{(1)} \approx 30$ %. Тогда $f_1(105) = 2$. Так как $y_1^{(1)} = 105 < x_1^+ = 120$, то принимаем $k = 2$.

При помощи ЛПР выбираем величину $b_1^{(2)}$ из условия $(105 + b_1^{(2)}, 30)^T \sim (105, 70)^T$. ЛПР указало, что $b_1^{(2)} \approx 25$. Принимаем $f_1(130) = 3$. Так как $y_1^{(2)} = 130 > x_1^+ = 120$, то переходим к построению частной функции по второму показателю ($j = 2$).

Аналогичным образом, выполняя п. 5 алгоритма для $j = 2$, получаем: $b_2^{(1)} = 15$ из условия $(60, 70 + b_2^{(1)})^T \sim (75, 70)^T$, $f_2(85) = 2$, $b_2^{(2)} = 20$ — из условия $(60, 85 + b_2^{(2)})^T \sim (75, 85)^T$, $f_2(105) = 3$.

Нанося полученные точки на графики зависимостей f_1 от $y_1^{(k)}$ и f_2 от $y_2^{(k)}$, $k = 0, 1, 2, 3$ и проводя через них гладкие кривые, получаем аппроксимации частных функций $f_1(x_1)$ и $f_2(x_2)$ (см. рис. 15, б, в). На рис. 15, г показаны изокванты полученной аддитивной функции эффективности.

Для проверки непротиворечивости суждений ЛПР воспользуемся информацией, полученной в ходе проверки условий соответственных замещений. Для этого эквивалентные точки, связанные двойной стрелкой на рис. 15, а, перенесены на график изоквант функции эффективности. Пары эквивалентных точек отмечены на рис. 15, г символами $\circ$, $\odot$, $*$, $+$. Анализ расположения этих точек показывает, что ранее выявленные предпочтения ЛПР достаточно хорошо согласуются с кривыми равноценности $W_e(X) = \text{const}$.

Построенная функция эффективности позволяет выделить наилучший проект из эффективного множества альтернативных вариантов. Пусть имеется четыре альтернативных варианта со значениями показателей эффективности: $W^I = (70, 90)^T$, $W^{II} = (90, 85)^T$, $W^{III} = (95, 70)^T$, $W^{IV} = (110, 60)^T$. Значения функции эффективности для указанных проектов равны:

$$W_e(70, 90) = f_1(70) + f_2(90) = 0,65 + 2,13 = 2,78;$$

$$W_e(90, 85) = f_1(90) + f_2(85) = 1,65 + 2,0 = 3,65;$$

$$W_e(95, 70) = f_1(95) + f_2(70) = 1,85 + 1,0 = 2,85;$$

$$W_e(110, 60) = f_1(110) + f_2(60) = 2,4 + 0,73 = 3,13;$$

поэтому наилучшим является вариант II.

Построение нормированной функции эффективности осуществляется методом половинного деления частных функций эффективности с помощью следующего алгоритма.

Алгоритм 6. 1. Установить области изменения оценок по каждому из частных показателей W_i $[x_i^-, x_i^+]$, $i = \overline{1, m}$.

2. Положить $f_i(\bar{x}_i) = 0$ и $f_i(x_i^+) = 1$, $i = \overline{1, m}$.

3. Разбить диапазоны изменения частных показателей на две подобласти граничными точками $\bar{x}_i$, $i = \overline{1, m}$.

4. Для каждого $j = \overline{1, m}$ выполнить следующие действия:

4.1. Присвоить $y^- = x_j^-$, $y^+ = x_j^+$ и положить $k = 1$.

4.2. Зафиксировать значения частных показателей W_i на уровнях $\bar{x}_i$, $i \neq j$ и положить $y_{\text{л}} = y^-$, $y_{\text{п}} = y^+$.

4.3. Вычислить $y^{(k)} = \frac{y_{\text{л}} + y_{\text{п}}}{2}$.

Получить от ЛПР информацию о предпочтительности увеличения W_j от значения y^- до значения $y^{(k)}$ над увеличением W_j от значения $y^{(k)}$ до значения y^+ вследствие некоторого уменьшения значений, показателей W_i , $i = \overline{1, m}$, $i \neq j$.

4.4. Если предпочтительнее переход $y^- \rightarrow y^{(k)}$, то положить $y_{\text{п}} = y^{(k)}$; если предпочтительнее переход $y^{(k)} \rightarrow y^+$, то $-y_{\text{л}} = y^{(k)}$; перейти к п. 4.3. Если переходы одинаковы по предпочтительности, то присвоить

$$f_j(y^{(k)}) = \begin{cases} 0,5, & \text{если } k = 1; \\ 0,25, & \text{если } k = 2; \\ 0,75, & \text{если } k = 3. \end{cases}$$

4.5. Положить $k = k + 1$ и проверить условие $k > 3$: да — окончить вычисления, нет — перейти к п. 4.6.

4.6. Присвоить

$$y^- = \begin{cases} x_j^-, & \text{если } k = 2; \\ y^{(1)}, & \text{если } k = 3; \end{cases}$$

$$y^+ = \begin{cases} y^{(1)}, & \text{если } k = 2; \\ x_j^+, & \text{если } k = 3 \end{cases}$$

и перейти к п. 4.2.

5. Через полученные точки в пространстве $x_i 0 f_i(x_i)$, $i = \overline{1, m}$ провести плавные кривые частных функций эффективности.

6. Для всех пар (t, l) с номерами из $j = \overline{1, m}$, $t \neq l$ выполнить следующие действия:

6.1. При помощи ЛПР установить предпочтительность при попарном сравнении оценок вида

$$X_+^l = (x_1^-, x_2^-, x_3^-, \dots,$$

$$x_l^+, x_{l+1}^-, \dots, x_m^-)^T \begin{pmatrix} \succ \\ \sim \\ \prec \end{pmatrix} \times$$

$$\times X_+^t = (x_1^-, x_2^-, x_3^-,$$

$$x_t^+, x_{t+1}^-, \dots, x_m^-)^T$$

6.2. Если $X_+^l \succ X_+^t$, то, уменьшая значение l -й компоненты оценки X_+^l , при помощи ЛПР нужно установить такое значение $x_l < x_l^+$, при котором оценки $X^l = (x_1^-, x_2^-, \dots, x_l, x_{l+1}^-, \dots, x_m^-)^T$ и X_+^t одинаковы по предпочтительности (если $X_+^l \prec X_+^t$, то, уменьшая значение t -й компоненты оценки X_+^t , аналогично определить $x_t < x_t^+$, приводящее к $X^t \sim X_+^l$).

Составить уравнение

$$\gamma_l f_l(x_l) = \gamma_t \quad (\text{или } \gamma_t \cdot f_t(x_t) = \gamma_l),$$

$$\gamma_l > \gamma_t \quad (\text{или } \gamma_t > \gamma_l).$$

Если при выполнении п. 6.2 имеет место

$$X_+^l \sim X_+^t, \quad \text{то } \gamma_l = \gamma_t.$$

7. Воспользовавшись построенными частными функциями эффективности

$f_j(x_j)$ и нормирующим условием $\sum_{j=1}^m \gamma_j = 1, \gamma_j > 0$, решить систему уравнений¹:

$$\left. \begin{aligned} \gamma_i f_i(x_i) &= \gamma_j, \quad i \neq j, \quad i, j = \overline{1, m}; \\ \sum_{i=1}^m \gamma_i &= 1 \end{aligned} \right\}$$

относительно γ_i .

Пример 13. В условия примера 12 вводится дополнительный частный показатель W_3 — экономичность самолета. При этом нормированные значения показателя W_3 изменяются в пределах от $x_3^- = 70$ до $x_3^+ = 130$. С помощью алгоритма 4 установлено, что частные показатели W_1, W_2, W_3 взаимонезависимы по предпочтению. С помощью алгоритма 6 построим нормированную функцию эффективности.

Полагаем $f_1(60) = f_2(30) = f_3(70) = 0, f_1(120) = f_2(100) = f_3(120) = 1$.

Выбираем граничные точки $x_1' = 80, x_2' = 70, x_3' = 100$. По п. 4 алгоритма строим частную функцию эффективности $f_1(x_1)$.

Полагаем $y^- = 60, y^+ = 120$ и $y_{\Pi} = 60, y_{\Sigma} = 120, k = 1$. Вычисляем $y^{(1)} = \frac{60 + 120}{2} = 90$. Показатели W_2

и W_3 фиксируют на уровнях $x_2' = 70$ и $x_3' = 100$. От ЛПР получена информация о том, что увеличение показателя W_1 от $y^- = 60$ до $y^{(1)} = 90$ предпочтительнее увеличения от $y^{(1)} = 90$ до $y^+ = 120$, так как ЛПР считает, что уступка по показателям W_2, W_3 от уровней x_2', x_3' может компенсироваться лишь первым переходом. Поэтому полагаем $y_{\Pi} = y^{(1)} = 90$ и вычисляем новое значение $y^{(1)} = \frac{60 + 90}{2} = 75$.

От ЛПР получена информация, что переход $y^- = 60 \rightarrow y^{(1)} = 75$ менее предпочтителен перехода $y^{(1)} = 75 \rightarrow y^+ = 120$. Поэтому полагаем $y_{\Pi} =$

$y^{(1)} = 75$ и вычисляем новое значение $y^{(1)} = \frac{75 + 90}{2} = 82,5 \approx 83$.

От ЛПР получена информация, что переходы $60 \rightarrow 83$ и $83 \rightarrow 120$ примерно эквивалентны. Тогда присваиваем $f_1 \times (y^{(1)}) = f_1(83) = 0,5$.

Полагаем $k = 2, y^- = x_1^- = 60, y^+ = y^{(1)} = 83$ и переходим к п. 4.2 алгоритма: $y_{\Pi} = y^- = 60, y_{\Sigma} = y^+ = 83$. Выполняя действия по пп. 4.3—4.4, определяем $y^{(2)} = 70$ и $f_1(70) = 0,25$.

Полагаем $k = 3, y^- = y^{(1)} = 83, y^+ = x_1^+ = 120$ и переходим к п. 4.2 алгоритма: $y_{\Pi} = y^- = 83, y_{\Sigma} = y^+ = 120$. В результате получаем $y^{(3)} = 105, f_1(105) = 0,75$. Действуя аналогичным образом по п. 4 алгоритма, получим:

$j = 2 : y^{(1)} = 78, y^{(2)} = 58, y^{(3)} = 93,$

$f_2(y^{(1)}) = 0,5, f_2(y^{(2)}) = 0,25,$

$f_2(y^{(3)}) = 0,75; j = 3 : y^{(1)} = 95,$

$y^{(2)} = 85, y^{(3)} = 110, f_3(y^{(1)}) = 0,5,$

$f_3(y^{(2)}) = 0,25, f_3(y^{(3)}) = 0,75.$

Аппроксимируем полученные точки плавными кривыми (рис. 16, а—в).

Выполняем п. 6 алгоритма. При помощи ЛПР сравниваем оценки:

$X_+^1 = (120, 30, 70)^T$ и

$X_+^2 = (60, 100, 70)^T;$

$X_+^1 = (120, 30, 70)^T$ и

$X_+^3 = (60, 30, 130);$

$X_+^2 = (60, 100, 70)^T$ и

$X_+^3 = (60, 30, 130).$

ЛПР указало, что $X_+^2 \succ X_+^1$, т. е. в этой ситуации (дальность — 60, грузоподъемность — 30, экономичность — 70) ЛПР считает более предпочтительным увеличение грузоподъемности до уровня $x_2^+ = 100$, нежели увеличение дальности полета до максимального уровня $x_1^+ = 120$. В этом случае ЛПР просят указать, насколько по его мнению допустимо уменьшить грузо-

¹ Так как система переопределена, а ответы ЛПР приближенные, то в практике последнее условие часто используют вместо одного из уравнений. Исключенное уравнение применяют для коррекции величины γ_i .

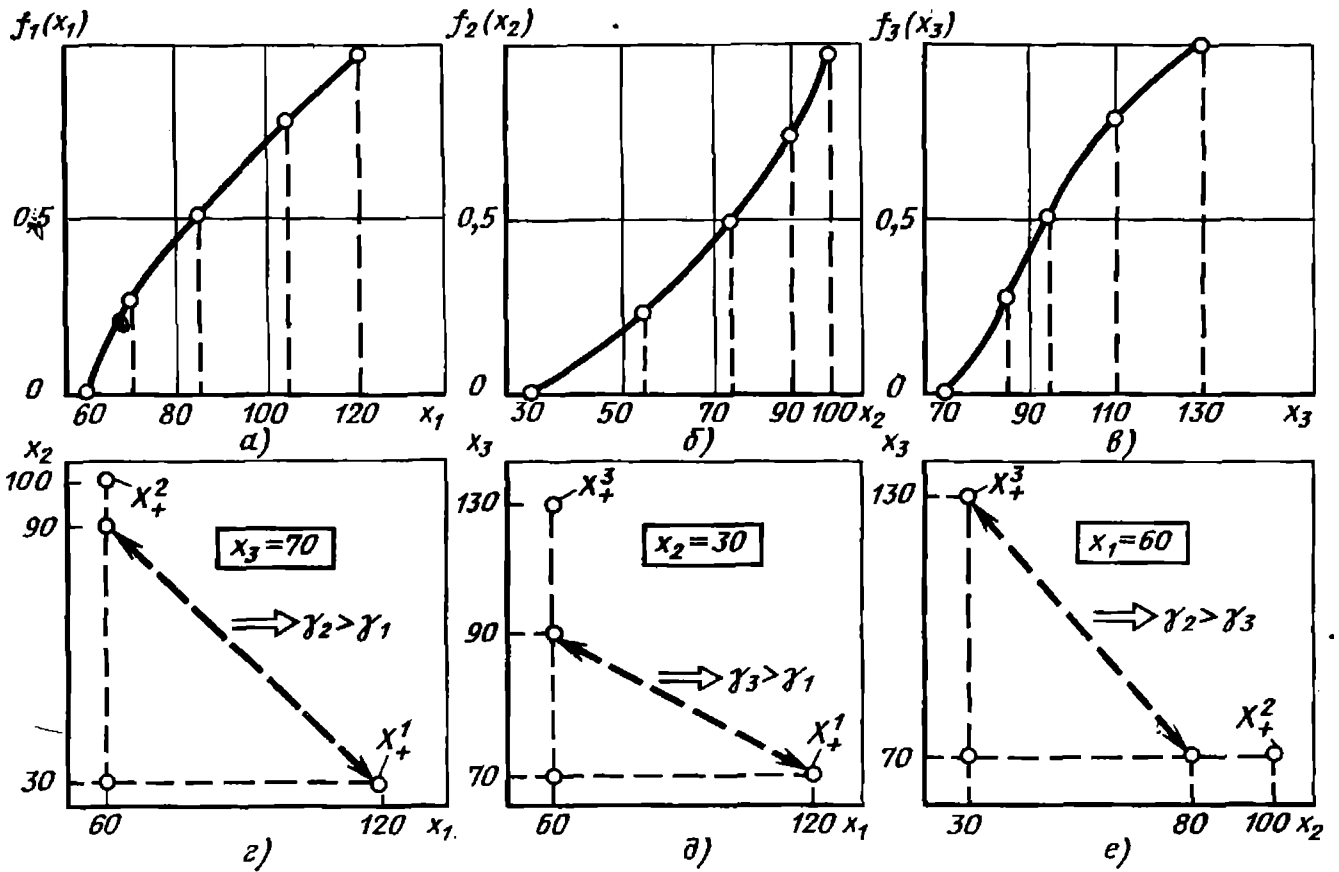


Рис. 16. Схема построения функции эффективности методом половинного деления (случай трех показателей)

подъемность, чтобы получить максимальную дальность. ЛПР указало, что с его точки зрения допустимо снижение грузоподъемности на 10 % от максимума, т. е. ЛПР считает одинаковыми по предпочтительности точки

$$(60, 90, 70)^T \sim (120, 30, 70)^T.$$

Составляем уравнение

$$\gamma_2 f_2(90) = \gamma_1.$$

Аналогичная информация получена для остальных пар:

$$(60, 30, 90)^T \sim (120, 30, 70)^T \Rightarrow$$

$$\Rightarrow \gamma_3 f_3(90) = \gamma_1;$$

$$(60, 30, 130)^T \sim (60, 80, 70)^T \Rightarrow$$

$$\Rightarrow \gamma_2 f_2(80) = \gamma_3.$$

На рис. 16, *г–е* эквивалентные точки соединены штриховыми линиями с двойной стрелкой.

Используя частные функции эффективности $f_1(x_1)$, $f_2(x_2)$, $f_3(x_3)$, решаем систему уравнений:

$$\left. \begin{aligned} \gamma_2 0,73 &= \gamma_1; \\ \gamma_3 0,37 &= \gamma_1; \\ \gamma_2 0,6 &= \gamma_3; \\ \Gamma &= \{\gamma_i | \gamma_1 + \gamma_2 + \gamma_3 = 1, \\ \gamma_i &> 0, i = \overline{1, 3}\} \end{aligned} \right\}$$

методом невязок, т. е. решаем оптимизационную задачу:

$$\begin{aligned} \min_{\gamma \in \Gamma} \varphi(\Gamma) &= \min_{\gamma_i} \{(\gamma_2 0,73 - \gamma_1)^2 + \\ &+ (\gamma_3 0,37 - \gamma_1)^2 + (\gamma_2 0,6 - \gamma_3)^2 + \\ &+ (\gamma_1 + \gamma_2 + \gamma_3 - 1)^2\}. \end{aligned}$$

Получаем $\gamma_1 = 0,22$, $\gamma_2 = 0,41$, $\gamma_3 = 0,37$. Так как из результатов сравнения точек X_+^i и X_+^j , $i < j$, $i, j = \overline{1, 3}$ по предпочтительности следует, что $\gamma_2 > \gamma_3 > \gamma_1$, то полученное решение непротиворечиво. Для уточнения по-

лученных коэффициентов и коррекции частных функций эффективности следует провести дополнительный анализ с участием ЛПР. При этом ЛПР просят сравнить по предпочтительности следующие пары оценок:

$$X_+^1 = (120, 30, 70)^T \text{ и}$$

$$X_+^{23} = (60, 100, 130)^T;$$

$$X_+^2 = (60, 100, 70)^T \text{ и}$$

$$X_+^{13} = (120, 30, 130)^T;$$

$$X_+^3 = (60, 30, 130)^T \text{ и}$$

$$X_+^{12} = (120, 100, 70)^T.$$

По результатам этого сравнения сопоставляют значения отдельных коэффициентов γ_i с суммой остальных, т. е.

$$X_+^i \left\{ \begin{matrix} > \\ \sim \\ < \end{matrix} \right\} X_+^k \Rightarrow \gamma_i \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} \gamma_j + \gamma_k.$$

При необходимости методом невязок корректируют полученные значения γ_i , $i = \overline{1, 3}$.

Кроме того, ЛПР предлагают сравнить по предпочтительности несколько векторных оценок $X = (x_1, x_2, x_3)^T$, Y , Z , имеющих одинаковое значение функции эффективности при вычисленных коэффициентах γ_i . По результатам его ответов корректируют вид частных функций эффективности (изменяя только одну из компонент векторной оценки) и значения коэффициентов.

Итеративные процедуры, основанные на вложенности отношений предпочтения (см. гл. 4), предполагают формирование на каждом этапе решения критерия оценивания альтернатив. Последовательность таких критериев есть итеративное решающее правило. Формирование решающего правила основывается, таким образом, на единой методологии, предполага-

ющей декомпозицию задачи выбора и получение последовательности усложняющейся непротиворечивой информации.

На первом шаге не требуется находить сразу наилучшие решения, так как для этого нет достаточной информации. Поэтому формируют наиболее очевидные допущения (например, о независимости частных показателей по предпочтению) и если они выполняются, то строят отношения предпочтения, позволяющие выделить эффективное множество решений (например, U_{ω_0}). На последующих шагах делают дополнительные допущения и получают минимально необходимую новую информацию, позволяющую сформировать расширенные отношения предпочтений, не изменяющие предыдущих. Поэтому эти отношения оказываются «вложенными» друг в друга, а следовательно, «вложенными» оказываются и соответствующие эффективные решения, т. е.

$$\left. \begin{aligned} \gamma_0 &\equiv \gamma_1 \equiv \dots \equiv \gamma_k \equiv \dots \equiv \gamma_N; \\ U &\equiv U_{\omega_0} \equiv U_{\omega_1} \equiv \dots \equiv U_{\omega_k} \equiv \dots \equiv U_{\omega_N}. \end{aligned} \right\} \quad (12.101)$$

Таким образом, процесс (12.101) заканчивается в одном из трех случаев:

1) все стратегии $u \in U_{\omega_N}$ эквивалентны;

2) множество U_{ω_N} содержит только одну стратегию u^* ;

3) дальнейшее сужение множества U_{ω_N} невозможно ввиду невозможности получения новой содержательной информации о предпочтениях.

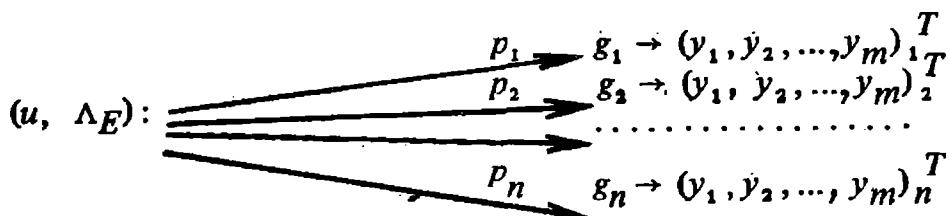
Такая итеративная процедура, реализующая методологию «вложенных» отношений, позволяет ЛПР лучше разобраться в проблеме и структуризовать собственную систему предпочтений.

Глава 13. Выработка решений в условиях неопределенности

1. ЗАДАЧИ ПРИНЯТИЯ РЕШЕНИЙ В УСЛОВИЯХ СТОХАСТИЧЕСКОЙ НЕОПРЕДЕЛЕННОСТИ ОБЩАЯ ПОСТАНОВКА ЗАДАЧИ

Так как операция проводится в условиях стохастической неопределен-

ности, то каждой стратегии $u \in U$ ставится в соответствие вероятностное распределение на множестве векторов характеристик $(y_1, y_2, \dots, y_m)^T$ исходов $g \in G$:



Таким образом, для установления предпочтений на множестве стратегий модель цели операции (показатель и критерий эффективности) должна учитывать вероятностный характер исхода. В этом случае

$$u \geq v \iff W_e^\lambda(u) \geq W_e^\lambda(v), \quad (13.1)$$

где $W_e^\lambda(u)$ — функция эффективности в условиях стохастической неопределенности, определяемая по правилу:

$$W_e^\lambda = M[f^{\theta_c}(\rho(Y, Y^{TP}))], \quad (13.2)$$

где $M[\cdot]$ — оператор математического ожидания; $\rho(Y, Y^{TP})$ — функция соответствия; $f^{\theta_c}(\cdot)$ — оценочная функция, построенная с учетом информации θ_c об отношении ЛПР к различным ситуациям в условиях стохастической неопределенности¹, или $f^{\theta_c}(\rho) = \rho(\cdot)$, если информация θ_c не используется.

¹ Информация θ_c рассматривается как сведения об отношении ЛПР к стохастическому риску. Поэтому в дальнейшем словосочетание «отношение к риску» понимается именно в этом смысле.

Если информация θ_c отсутствует или для установления предпочтения по правилу (13.1) достаточно использовать только «объективные» показатели эффективности, построенные как

$$W_i = M[\rho(y_i, y_i^{TP})], \quad i = \overline{1, m}, \quad (13.3)$$

то задача принятия решения в условиях стохастической неопределенности сводится к задаче принятия решения в условиях определенности в общем случае по векторному показателю W с компонентами $W_1, W_2, \dots, W_m$.

Для задачи принятия решений в условиях стохастической неопределенности функция $\rho(y, y^{TP})$ может учитывать «осреднение» по вероятности. В дополнение к функциям вида:

$$\rho(y, y^{TP}) = \begin{cases} 1, & y \geq y^{TP}; \\ 0, & y < y^{TP}; \end{cases} \quad (13.4)$$

$$\rho(y, y^{TP}) = y \quad (13.5)$$

часто используется квантиль

$$\rho = F_u^{-1}(1 - \alpha) = y_\alpha, \quad (13.6)$$

где α — заданная вероятность, а $F_u^{-1}(\cdot)$ — обратная функция к функции распределения результата операции при фиксированной стратегии $u \in U$; некоторые степенные функции результатов операции:

$$\rho(y, y^{TP}) = (y - y^{TP})^2; \quad (13.7)$$

$$\rho(y, y^{TP}) = (y - M[y])^2. \quad (13.8)$$

Таким образом,* в соответствии с (13.2) и (13.3) показатели эффективности являются определенными; поэтому для принятия решения относительно эффективности стратегий $u \in U$ можно использовать методы, описанные в гл. 12. Критерии эффективности в этом случае формулируются обычным порядком: критерий пригодности, задающий множество U^{SAT} :

$$U^{SAT} = \{u \mid W_i(u) \geq W_i^{TP}, i = \overline{1, m}\}; \quad (13.9)$$

критерий оптимальности

$$U^* = \{u^* \mid W_e^\lambda(u^*) = \max_{u \in U} W_e^\lambda(u)\}, \quad (13.10)$$

где $W_e^\lambda = \Phi(W)$ — функция агрегирования, построенная как свертка частных показателей $W_i, i = \overline{1, m}$ типа (13.3).

Если решение задачи выбора $\langle U; U^* \rangle$ не единственно, т. е. мощность множества «объективно» оптимальных или пригодных стратегий велика, то необходимо привлекать дополнительную информацию θ_c . Эта информация касается не частных характеристик распределения (математическое ожидание, дисперсия, квантиль и т. д.), порождаемого стратегией $u \in U$, а его исчерпывающей характеристики — закона распределения. При этом исследователь должен выяснить отношение ЛПР к различным вероятностным распределениям на множестве исходов, т. е.

$$u \geq v \iff F_u^{\theta_c} \geq F_v^{\theta_c}, \quad u, v \in U, \quad (13.11)$$

где F_u — закон распределения характеристик $(y_1, y_2, \dots, y_m)^T$ исходов $g \in G$ при фиксированной стратегии $u \in U$.

ЗАДАЧА ПРИНЯТИЯ РЕШЕНИЙ В УСЛОВИЯХ СТОХАСТИЧЕСКОЙ НЕОПРЕДЕЛЕННОСТИ ПО СКАЛЯРНОМУ ПОКАЗАТЕЛЮ

Рассмотрим случай, когда информация θ_c не учитывается, а предпочтения ЛПР описываются стремлением к увеличению значений скалярного показателя эффективности. Если при формировании показателя эффективности выбрана функция (13.4), то он будет представлять собой вероятность достижения цели операции

$$W(u) = P(g(u) \geq g^{TP}), \quad (13.12)$$

где символ « $\wedge$ » означает случайную величину.

Для использования показателя (13.7) необходимо найти закон распределения $z = g(u) - g^{TP}$. В этом случае отношение предпочтения на множестве стратегий задается отношением

$$u \geq v \iff P(z(u) > 0) \geq P(z(v) > 0). \quad (13.13)$$

Однако при исследовании эффективности технических систем встречаются задачи, когда требуемый результат определен или может рассматриваться таковым. Тогда в качестве показателя эффективности выступает вероятность получения результата не ниже требуемого:

$$W(u) = P(g(u) \geq y^{TP}); \quad (13.14)$$

он называется *показателем вероятностной гарантии*. Отношение предпочтения на множестве стратегий устанавливаются по (13.13).

Если используют функцию (13.5), то показатель эффективности есть математическое ожидание результата

$$W(u) = M[g(u)]; \quad (13.15)$$

он называется *показателем среднего результата*. Отношение предпочтения

на множестве стратегий устанавливают по соотношению

$$u \geq v \iff M [\vartheta (u)] \geq M [\vartheta (v)]. \quad (13.16)$$

С помощью (13.6) вводится показатель вероятностно-гарантированного результата

$$W(u) = M[F_u^{-1}(1 - \alpha)] = y_\alpha(u) \quad (13.17)$$

и

$$u \geq v \iff y_\alpha(u) \geq y_\alpha(v). \quad (13.18)$$

Показатели кучности и дисперсии результатов вводятся на основе функций соответствия (13.7), (13.8).

Для введения критериев эффективности значения только требуемого результата операции (как это было в задачах принятия решения в условиях определенности) недостаточно. Необходимо получить от ЛПР дополнительную информацию о требуемом уровне значения показателя эффективности W^{TP} .

На основании критерия пригодности: принцип приемлемого среднего результата

$$U^{SAT} = \{u \mid M[\vartheta(u)] \geq M^{TP}\};$$

принцип приемлемой вероятностной гарантии

$$U^{SAT} = \{u \mid P(\vartheta(u) \geq y_{TP}) \geq P^{TP}\};$$

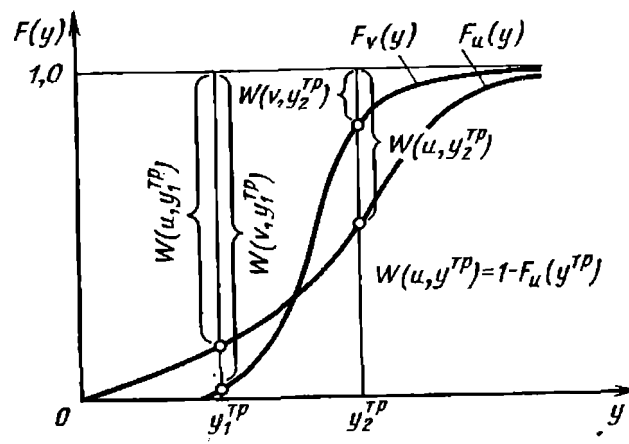


Рис. 1. Пример неустойчивости решения по критерию наибольшей вероятностной гарантии

принцип приемлемого вероятностно-гарантированного результата

$$U^{SAT} = \{u \mid y_\alpha(u) \geq y^{TP}\}$$

и другие;

на основании критерия оптимальности:

принцип наибольшего среднего результата

$$U^* = \{u^* \mid M[\vartheta(u^*)] = \max_{u \in U} M[\vartheta(u)]\};$$

принцип наибольшей вероятностной гарантии

$$U^* = \{u^* \mid P(\vartheta(u^*)) = \max_{u \in U} P(\vartheta(u))\};$$

принцип наибольшего вероятностно-гарантированного результата

$$U = \{u^* \mid y_\alpha(u^*) = \max_{u \in U} y_\alpha(u)\}$$

и другие.

Хотя в общем случае принципы вероятностной гарантии и вероятностно-гарантированного результата более информативны по сравнению с принципами, использующими отдельные моменты (математическое ожидание, дисперсия) результата, в вычислительном отношении они сложнее. Более того, получаемые с их помощью решения, как правило, неустойчивы (критичны к выбору требуемого уровня показателя эффективности и требуемого результата). Эта ситуация приведена на рис. 1. При назначении в качестве требуемого результата y_1^{TP} стратегии $v \succ u$, так как $W(v, y_1^{TP}) > W(u, y_1^{TP})$, а при требуемом результате y_2^{TP} — наоборот, $u \succ v$.

Принципы принятия решения на основе среднего результата проще, так как они не используют информацию о значениях требуемого результата y^{TP} , а при применении критерия оптимальности — информацию о требуемом значении показателя эффективности. Однако использование показателя среднего результата оправдано лишь в том случае, когда операция носит массовый характер и (или) обладает свойством повторяемости. Если операция

«уникальна», то целесообразно использовать показатели вероятностной гарантии и вероятностно-гарантированного результата либо, если это по каким-то причинам невозможно, использовать в дополнение к показателю среднего результата показатели его дисперсии на функциях (13.7) или (13.8). Таким образом вводят векторный показатель эффективности и используют методы принятия решения по векторному показателю в условиях определенности.

Для построения (13.13), (13.16), (13.18) необходимо знать закон распределения переменной $\vartheta(u)$ для всех $u \in U$. Установление вида и параметров закона распределения $F_u(y)$ представляет собой важную самостоятельную задачу. Если удастся поставить эксперимент, эта задача решается статистическими методами параметрического и непараметрического оценивания (см. т. 2). Если эксперимент ограничен или вовсе невозможен, то задачу выбора необходимо решать как задачу принятия решения в условиях нестохастической неопределенности либо восстанавливать априорное распределение методами экспертного оценивания (вводя отношения порядка на величинах вероятностей исходов) и уточнять их результатами ограниченного эксперимента по байесовскому правилу.

Особым случаем задач принятия решения в условиях стохастической неопределенности являются задачи статистических решений.

Задачи статистических решений возникают тогда, когда исход операции существенно зависит от того, какое значение принимает один (или несколько) неопределенный фактор. Относительно распределения этого фактора (если он стохастической природы) в лучшем случае имеются лишь предварительные (априорные) сведения. Поэтому ЛПР заинтересован в том, чтобы уточнить его распределение. Можно провести эксперимент (подоперацию в рамках рассматриваемой операции), измеряя некоторые характеристики Z , распределение которых зависит от того, какое значение принял неопределенный фактор λ .

Следовательно, задачу статистических решений можно рассматривать как задачу принятия решения относительно условий или результатов функционирования большой технической системы по результатам исследования ее сравнительно простых подсистем (подопераций).

Пример 1. При нормальном режиме технологического процесса любое изделие независимо от других оказывается годным с вероятностью p_1 . Изделия изготавливают партиями по N штук в каждой. Партия считается годной, если содержит не более n_d дефектных изделий. Качество контролируют путем проверки взятых из партии наудачу n изделий ($n < N$). Требуется установить пороговое значение z^0 , на основании которого по величине $\hat{z}$ (z — число дефектных изделий в контрольной выборке объема n) следует с минимальным риском считать всю партию объема N годной, т. е. содержащей не более n_d изделий. В этом случае степень риска связывается с возможными потерями в операции от принятых решений (правильных и ошибочных), а величины потерь выступают как внешнее дополнение для подоперации контроля. Решающее правило может быть как двух- (только принимать или только не принимать), так и многоальтернативным (принимать, отправить на доработку или произвести сплошной контроль, забраковать). В последнем случае необходимо установить не один порог z^0 , а два — z^1 , z^2 . Если $z < z^1$, принять партию, если $z > z^2$ — забраковать, в остальных случаях осуществить сплошной контроль.

Пример 2. Радиолокационная станция, находясь в режиме дежурного приема, должна зафиксировать факт появления объекта (летательного аппарата). Помимо полезного сигнала станция воспринимает помехи. В результате этого возможны четыре информационные ситуации: объект правильно распознан, помеха принята за объект, объект принят за помеху, помеха правильно распознана. Каждая информационная ситуация может быть охарактеризована потерями, связанными с принятым решением. При неизвестной априорной вероят-

ности появления объекта задача статистических решений может, например, формулироваться следующим образом: найти такое пороговое значение амплитуды сигнала, регистрируемого локатором, чтобы принимаемые по нему решения обеспечивали минимально возможные в среднем потери.

Таким образом, каждая из рассмотренных задач может быть сформулирована в следующей общей постановке.

Исследуемый фактор λ может принимать одно из m возможных значений (состояний) $\lambda_1, \lambda_2, \dots, \lambda_m$; априорные вероятности $P(\lambda_i)$, $i = \overline{1, m}$ удовлет-

воряют условию $\sum_{i=1}^m P(\lambda_i) = 1$. Про-

водится испытание, в котором осуществляется n независимых наблюдений какой-либо характеристики (признака) z фактора λ . Результаты наблюдений представляются системой случайных величин $(z_1, z_2, \dots, z_n)$ дискретного или непрерывного типа из-за стохастической природы фактора λ или измерения признака. Распределение этой системы зависит от конкретного значения λ_i и определяется функцией правдоподобия (см. т. 2). Функция правдоподобия $L(z_1, z_2, \dots, z_n/\lambda_i)$ известна для всех λ_i , $i = \overline{1, m}$ и задается в зависимости от типа случайной величины z следующими выражениями:

для дискретной случайной величины

$$\begin{aligned} L(z_1, z_2, \dots, z_n/\lambda_i) &= \\ &= P(z_1, z_2, \dots, z_n/\lambda_i) = \\ &= \prod_{j=1}^n P(z_j/\lambda_i); \end{aligned} \quad (13.19)$$

для непрерывной случайной величины

$$\begin{aligned} L(z_1, z_2, \dots, z_n/\lambda_i) &= \\ &= f(z_1, z_2, \dots, z_n/\lambda_i) = \\ &= \prod_{j=1}^n f_j(z_j/\lambda_i), \end{aligned} \quad (13.20)$$

где $f(\cdot/\lambda_i)$ — условная совместная плотность распределения системы случайных величин z_j ; $f_j(\cdot/\lambda_i)$ — частная условная плотность распределения.

По результатам эксперимента $(z_1, z_2, \dots, z_n)$ требуется осуществить выбор одного из k_0 решений $x_1, x_2, \dots, x_{k_0}$, каждое из которых основывается на оценке (по полученным данным) действительного состояния λ_i .

Последствия выбора любого решения x_k , $k = \overline{1, k_0}$ при условии, что объект находится в состоянии λ_i , $i = \overline{1, m}$, характеризуются значениями функции потерь $\Pi_{ki} = \Pi(x_k, \lambda_i)$, где Π_{ki} — потери, связанные с тем, что при истинном значении λ_i фактора λ было принято решение x_k . В этом смысле элементарные решения x_k являются характеристиками стратегии ЛПР, а их упорядоченный набор задает облик стратегии.

Сформулированная задача является общей задачей принятия решений при заранее фиксированном числе наблюдений (фиксированный эксперимент) в случае конечного множества возможных состояний λ_i *1.

Выбор стратегии (решающего правила) u_l , $l = \overline{1, l_0}$ в теории статистических решений осуществляется следующим образом. Пространство характеристик Z разбивается каким-либо образом на k_0 непересекающихся областей z^k , $k = \overline{1, k_0}$. На множестве Z определяется множество решающих правил u_l , предписывающих выбор того или иного решения x_k при попадании выборки в ту или иную область. Стратегии u_l могут быть рандомизированными (случайными) или нерандомизированными (детерминированными). Детерминированная стратегия устанавливает однозначное соответствие решений x_k областям z^k , на которые разбито множество Z .

При попадании выборки $(z_1, z_2, \dots, z_n)$ в ту или иную область z^l рандомизированное правило предусматривает случайный выбор одного из решений

*1. Общая задача принятия решения теоретически может быть сформулирована и для случая бесконечного множества состояний λ_i . Однако в практических приложениях такая задача возникает крайне редко и поэтому здесь не рассматривается. Не рассматривается также метод последовательного анализа Вальда, изложенный в т. 2.

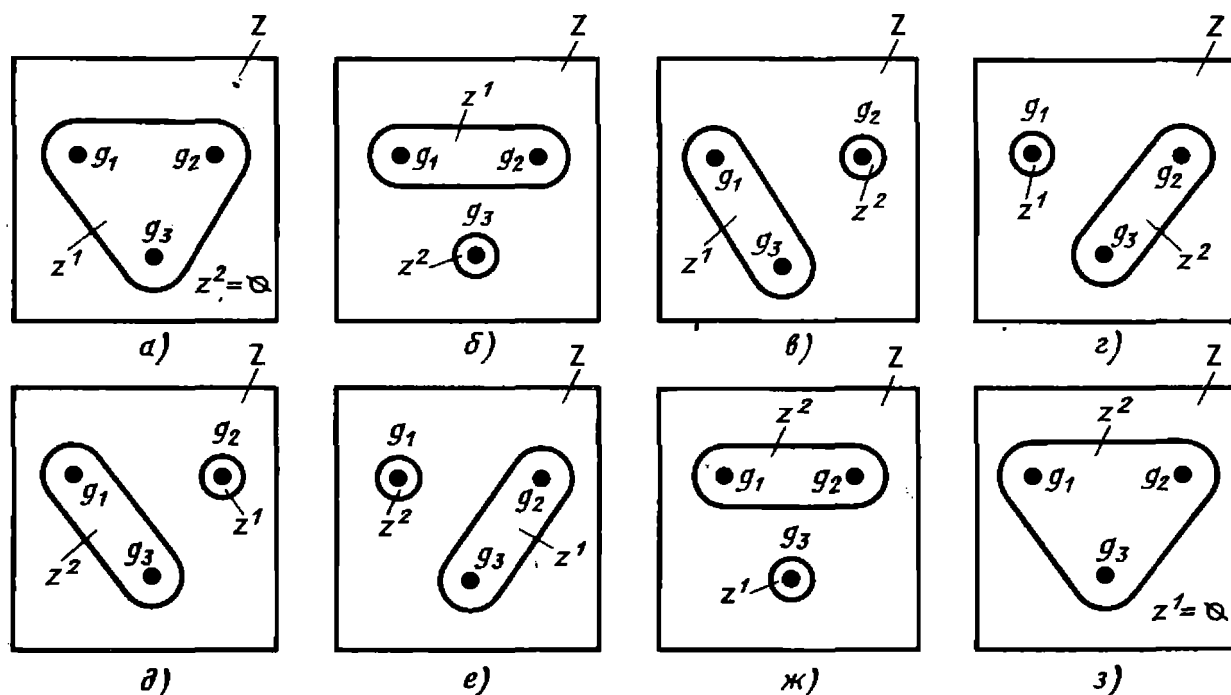


Рис. 2. Возможные разбиения пространства выборок Z на области z^1, z^2

x_k в соответствии с некоторым распределением вероятностей P_{tk} ($\sum_{k=1}^{k_0} P_{tk} = 1, t = \overline{1, k_0}$), т. е. по своему смыслу рандомизированное правило есть вероятностная смесь нерандомизированных правил.

В общем случае стратегия u_i сложнее правила x_k , т. е. наблюдатель может либо верить результатам эксперимента, либо поступать так, как считает целесообразным. Для иллюстрации этого рассмотрим следующий пример.

Пример 3. Для контроля качества мелкой партии из нее отбирают один образец и его размеры контролируют тремя шаблонами. В этом случае пространство z результатов измерений является дискретным и содержит три точки (исхода g_1, g_2, g_3), соответствующие трем возможным размерам по шаблонам. В соответствии с двумя возможными решениями (x_1 — принять партию, x_2 — забраковать партию) необходимо определить множество возможных нерандомизированных стратегий. В соответствии с решениями $x_k, k = \overline{1, k_0}$ разобьем пространство Z на k_0 непересекающихся подобластей (рис. 2). Тогда каждому из возможных разбиений можно поставить в соот-

ветствие стратегии, число которых $l_0 = 2^3 = 8$.

Содержательно стратегии предписывают следующий порядок действий:

u_1 — принять партию при любом исходе эксперимента (рис. 2, а);

u_2 — принять партию, если изделие соответствует размерам первого или второго шаблона, и забраковать в противном случае (рис. 2, б). Аналогично определяются стратегии $u_3, u_4, \dots, u_8$ (соответственно схемам разбиения на рис. 2, в—з).

Какого-либо общего метода определения функции потерь Π_{hi} теория статистических решений не дает. На практике значения Π_{hi} определяют из анализа внешнего дополнения с учетом реальных условий каждой задачи. Так, указанные значения оценивают либо убытками от выбора неправильных решений, либо наряду с убытками вводят в рассмотрение выигрыш от выбора правильных решений (отрицательные потери), либо одновременно с убытками и выигрышами учитывают затраты на эксперимент и т. д. Удобнее элементы матрицы потерь представлять относительными потерями:

$$\Delta \Pi_{hi} = \Pi_{hi} - \min_i \{\Pi_{hi}\}, \quad i = \overline{1, m}, \quad (13.21)$$

что обеспечивает неотрицательность всех ее элементов при любом подходе к определению потерь.

В качестве характеристики потерь, средних для многократного применения нерандомизированной стратегии u_l при условии, что действительное значение фактора λ равно λ_k , принимается условная функция потерь (риска)

$$\rho(u_l/\lambda_i) = \sum_{k=1}^{k_0} \Pi_{ki} P_{u_l}(\hat{Z}_n \in z^k/\lambda_i),$$

$$i = \overline{1, m}, \quad (13.22)$$

где $P_{u_l}(\hat{Z}_n \in z^k/\lambda_i)$ — вероятность попадания выборки $\hat{Z}_n = (z_1, z_2, \dots, z_n)^T$ в полученную согласно стратегии u_l область z^k (см. рис. 2).

Для установления предпочтения на множестве стратегий $U = \{u_l, l = \overline{1, l_0}\}$ вводится числовая функция (показатель эффективности подоперации), которая в зависимости от имеющейся в распоряжении ЛПР информации имеет вид

$$W(u_l) = \sum_{i=1}^m P(\lambda_i) \rho(u_l/\lambda_i) \quad (13.23)$$

в случае известных априорных вероятностей состояний λ_i , или

$$W(u_l) = \max_{\lambda_i \in \Lambda} \{\rho(u_l/\lambda_i)\}, \quad (13.24)$$

если априорные вероятности $P(\lambda_i)$ неизвестны, или

$$W(x_k) = \sum_{i=1}^m \Pi_{ki} P(\lambda_i/\hat{z}_n = g_k), \quad (13.25)$$

где $P(\lambda_i/\hat{z}_n = g_k)$ — апостериорная вероятность состояния λ_i , соответствующая попаданию выборки $\hat{z}_n$ в точку g_k пространства выборок Z (предполагается, что детерминированная стратегия u совпадает с решением x).

Выражение (13.23) представляет собой средние потери (риск) от применения стратегии u_l , выражение (13.24) — гарантированные средние потери, а (13.25) — апостериорные средние потери.

Если по тем или иным причинам нельзя определить матрицу потерь $\|\Pi_{ki}\|$, то в качестве показателя эффективности часто используют следующую функцию, заданную на множестве возможных значений λ_i :

$$W(\lambda_i) = P(\lambda_i/z_1, z_2, \dots, z_n), \quad (13.26)$$

значение которой суть апостериорная вероятность истинности состояния $\lambda = \lambda_i$ при условии, что реализовалась выборка $(z_1, z_2, \dots, z_n)$. Решение ЛПР в этом случае связывают с идентификацией (распознаванием) состояния по результатам эксперимента.

В качестве критерия эффективности при использовании показателей (13.23)–(13.25) используют критерий оптимальности вида

$$u^*: \min_U W(u_l) \quad \text{или} \quad x^*: \min_X W(x_k). \quad (13.27)$$

Если в (13.27) в качестве показателя эффективности используется (13.23), то такое правило называют байесовским, в случае (13.24) — минимаксным, а в случае (13.25) — байесовским апостериорным.

При использовании показателя (13.26) критерий оптимальности

$$\lambda^*: \max_{\Lambda} W(\lambda_i)$$

называют правилом максимума апостериорной вероятности.

Пример 4. В условиях примера 1 примем $p_1 = 0,5$, $N = 10$, $n_d = 5$, $n = 2$. Задана матрица потерь

$$\|\Pi_{ki}\| = \begin{pmatrix} 0 & 10 \\ 10 & 0 \\ 4 & 4 \end{pmatrix}$$

для трехальтернативного решающего правила: x_1 — принять партию; x_2 — забраковать, x_3 — провести сплошной контроль и двух состояний партии: λ_1 — годная, λ_2 — негодная.

В этом случае пространство Z представляется тремя точками: g_1 ($z = 0$), соответствующей числу z дефектных изделий в контрольной выборке, равному нулю; g_2 ($z = 1$) и g_3 ($z = 2$).

Определяем априорные вероятности состояний партии:

$$P(\lambda_1) = P(\hat{n}_d \leq n_d) = \\ = \sum_{t=0}^5 C_{10}^t (1-p_1)^t p_1^{10-t} = \\ = \sum_{t=0}^5 C_{10}^t 0,5^t \cdot 0,5^{10-t} = 0,623;$$

$$P(\lambda_2) = 1 - P(\lambda_1) = 1 - \\ - 0,623 = 0,377,$$

где $\hat{n}_d$ — случайное число дефектных изделий в партии.

Применим для выбора решений наиболее информативное байесовское апостериорное правило. Для этого необходимо вычислить апостериорные вероятности, входящие в выражение для показателя эффективности (13.25).

$$P(\lambda_i / \hat{Z}_n = z) = \\ = \frac{P(\lambda_i) P(\hat{Z}_n = z / \lambda_i)}{\sum_{i=1}^m P(\lambda_i) P(\hat{Z}_n = z / \lambda_i)}.$$

Вероятности $P(\hat{Z}_n = z / \lambda_i)$ определяются по формуле условной вероятности:

$$P(\hat{Z}_n = z / \lambda_1) = \frac{P(\hat{Z}_n = z, \hat{n}_d \leq n_d)}{P(\lambda_1)} = \\ = \frac{\sum_{j=1}^{n_d} P(\hat{Z}_n = z / \hat{n}_d = j) P(\hat{n}_d = j)}{P(\lambda_1)},$$

$z = 0, 1, 2;$

$$P(\hat{Z}_n = z / \lambda_2) = \\ = \frac{\sum_{j=n_d+1}^N P(\hat{Z}_n = z / \hat{n}_d = j) P(\hat{n}_d = j)}{P(\lambda_2)},$$

$z = 0, 1, 2.$

Входящие в два последних выражения составляющие определяются следующим образом:

$$P(\hat{Z}_n = z / \hat{n}_d = j) = \frac{C_j^z C_{N-j}^{n-z}}{C_N^n}, \\ j = 0, 1, 2, \dots, N; \quad z = 0, 1, 2; \\ P(\hat{n}_d = j) = C_N^j (1-p_1)^j p_1^{N-j}, \\ j = 0, 1, 2, \dots, N.$$

Подстановка исходных данных в эти выражения приводит к следующим результатам:

$$P(\lambda_1 / \hat{Z}_n = 0) = 0,856; \\ P(\lambda_1 / \hat{Z}_n = 1) = 0,636; \\ P(\lambda_1 / \hat{Z}_n = 2) = 0,364; \\ P(\lambda_2 / \hat{Z}_n = 0) = 0,144; \\ P(\lambda_2 / \hat{Z}_n = 1) = 0,364; \\ P(\lambda_2 / \hat{Z}_n = 2) = 0,636.$$

Тогда значения показателя (13.25) для различных z будут: при $z = 0$

$$W(x_1) = 1,44; \quad W(x_2) = 8,56; \\ W(x_3) = 4,00;$$

при $z = 1$

$$W(x_1) = 3,64; \quad W(x_2) = 6,36; \\ W(x_3) = 4,00;$$

при $z = 2$

$$W(x_1) = 6,36; \quad W(x_2) = 3,64; \\ W(x_3) = 4,00.$$

Анализ полученных результатов показывает, что $x^* = x_1$ при $z = 0$, $x^* = x_1$ при $z = 1$, $x^* = x_2$ при $z = 2$.

В этом случае стратегией ЛПР u^* (упорядоченный набор $(x_1, x_1, x_2)^T$) является следующее предписание действий по результатам эксперимента: принимать партию готовой продукции, если среди контрольных изделий окажется не более одной дефектной, и браковать, если оба контрольных изделия дефектны. В стратегии u^* частная ее характеристика (элементарное решение) x_3 отсутствует.

1. Множество возможных стратегий

u_i	0 =	1 =	2 =	u_i	0 =	1 =	2 =
u_1	x_1	x_1	x_1	u_{15}	x_2	x_2	x_3
u_2	x_1	x_1	x_2	u_{16}	x_2	x_3	x_1
u_3	x_1	x_1	x_3	u_{17}	x_2	x_3	x_2
u_4	x_1	x_2	x_1	u_{18}	x_2	x_3	x_3
u_5	x_1	x_2	x_2	u_{19}	x_3	x_1	x_1
u_6	x_1	x_2	x_3	u_{20}	x_3	x_1	x_2
u_7	x_1	x_3	x_1	u_{21}	x_3	x_1	x_3
u_8	x_1	x_3	x_2	u_{22}	x_3	x_2	x_1
u_9	x_1	x_3	x_3	u_{23}	x_3	x_2	x_2
u_{10}	x_2	x_1	x_1	u_{24}	x_3	x_2	x_3
u_{11}	x_2	x_1	x_2	u_{25}	x_3	x_3	x_1
u_{12}	x_2	x_1	x_3	u_{26}	x_3	x_3	x_2
u_{13}	x_2	x_2	x_1	u_{27}	x_3	x_3	x_3
u_{14}	x_2	x_2	x_2				

Если априорные вероятности $P(\lambda_i)$, $i = 1, 2$ неизвестны, то выбор стратегии осуществляют на основе минимаксного критерия с использованием (13.22), (13.24) и (13.27). Все возможные нерандомизированные стратегии u_i представлены в табл. 1.

Таблица содержит некоторые стратегии, являющиеся в условиях примера нелогичными: u_4 , u_{10} и др. Од-

2. Результаты вычисления условных функций потерь

u_i	$\rho(u_i/\lambda_1)$	$\rho(u_i/\lambda_2)$	u_i	$\rho(u_i/\lambda_1)$	$\rho(u_i/\lambda_2)$
u_1	0,00	10,00	u_{14}	10,00	0,00
u_2	1,45	5,79	u_{15}	9,13	1,68
u_3	0,58	7,47	u_{16}	5,48	6,14
u_4	5,11	5,18	u_{17}	6,93	1,93
u_5	6,56	0,97	u_{18}	6,06	3,61
u_6	5,69	2,65	u_{19}	1,38	9,42
u_7	2,04	7,11	u_{20}	2,83	5,21
u_8	3,49	2,90	u_{21}	1,96	6,89
u_9	2,62	4,58	u_{22}	6,49	4,60
u_{10}	3,44	9,03	u_{23}	7,94	0,39
u_{11}	4,89	4,82	u_{24}	7,07	2,07
u_{12}	4,02	6,50	u_{25}	3,42	6,53
u_{13}	8,55	4,21	u_{26}	4,87	2,32
			u_{27}	4,00	4,00

нако в интересах общности включим их в множество возможных стратегий.

Для каждой стратегии с помощью (13.22) вычислим условную функцию потерь. Поскольку исходная матрица потерь содержит нулевые элементы, то (13.22) примет вид:

$$\rho(u_i/\lambda_1) = 10P_{u_i}(\hat{Z}_n \in z^2/\lambda_1) + 4P_{u_i}(\hat{Z}_n \in z^3/\lambda_1);$$

$$\rho(u_i/\lambda_2) = 10P_{u_i}(\hat{Z}_n \in z^1/\lambda_2) + 4P_{u_i}(\hat{Z}_n \in z^3/\lambda_2).$$

Входящие в эти выражения условные вероятности $P_{u_i}(\hat{Z}_n \in z^k/\lambda_i)$ определяются соотношением

$$P_{u_i}(\hat{Z}_n \in z^k/\lambda_i) = \sum_{z \in z^k} P(\hat{Z}_n = z/\lambda_i),$$

где z^k — подобласти пространства Z , определяемые по аналогии с рис. 2, т. е., например, для стратегии u_1 имеем $z^1 = \{g_1, g_2, g_3\}$, а z^2 и z^3 пустые множества; для u_2

$$z^1 = \{g_1, g_2\}, \quad z^2 = \{g_3\}, \quad z^3 = \emptyset,$$

$$\text{для } u_6 - z^1 = \{g_1\},$$

$$z^2 = \{g_2\}, \quad z^3 = \{g_3\} \text{ и т. д.}$$

Результаты вычисления $\rho(u_i/\lambda_i)$ приведены в табл. 2. Рассматривая задачу выбора наилучшей стратегии u^* как задачу принятия решения с векторным показателем эффективности $W(u) = (\rho(u/\lambda_1), \rho(u/\lambda_2))^T$, выделим множества эффективных паретовских стратегий. Это множество состоит из стратегий:

$$U_{\gamma_{\omega_0}} = \{u_1, u_2, u_3, u_5, u_8, u_{14}, u_{23}\}.$$

Использование минимаксного правила приводит к тому, что оптимальной детерминированной стратегией u_d^* является стратегия u_8 , для которой средние гарантированные потери составляют

$$W(u_8) = \min_{u \in U_{\gamma_{\omega_0}}} \max_{\lambda_i \in \Lambda} \rho(u/\lambda_i) = 3,49.$$

Однако такая стратегия может оказаться излишне осторожной, так как может существовать рандомизированная стратегия u_p^* , для которой средние гарантированные потери меньше.

Поиск рандомизированной стратегии u_p^* можно осуществлять либо методом теории игр (этот метод будет рассмотрен позднее), либо основываясь на следующем положении, доказанном в теории статистических решений: рандомизированная минимаксная стратегия удовлетворяет условию

$$\rho(u_p^*/\lambda_1) = \rho(u_p^*/\lambda_2) = \dots \rho(u_p^*/\lambda_m), \quad (13.28)$$

причем u_p^* принадлежит смешанному расширению¹ множества U_{ω_0} .

Найдем рандомизированную стратегию u_p^* из условия (13.28). На рис. 3 представлены элементы множества U_{ω_0} и его смешанное расширение.

Рандомизированная стратегия u_p^* является вероятностной смесью детерминированных стратегий u_2 и u_8 .

Уравнение отрезка, соединяющего точки u_2 и u_8 , имеет вид

$$q\rho(u_2) + (1-q)\rho(u_8), \quad 0 \leq q \leq 1,$$

$$\text{где } \rho(u_i) = (\rho(u_i/\lambda_1), \rho(u_i/\lambda_2))^T.$$

Тогда из уравнения

$$q\rho(u_2) + (1-q)\rho(u_8) = \rho(u_p^*)$$

находим $q = 0,12$.

Таким образом, рандомизированная стратегия u_p^* есть смесь детерминированных стратегий u_2 и u_8 с вероятностями 0,12 и 0,88 соответственно. При этом гарантированные средние потери определяются выражением

$$W(u_p^*) =$$

$$= \min_{q \in Q} \max_{P(\lambda) \in P} \sum_{s=1}^{s_0} \sum_{i=1}^m q_s P(\lambda_i) \times \\ \times \rho(u_p^*/\lambda_i), \quad (13.29)$$

¹ Смешанное расширение дискретного эффективного множества представляет собой кусочно-линейную огибающую его точек (выпуклую комбинацию).

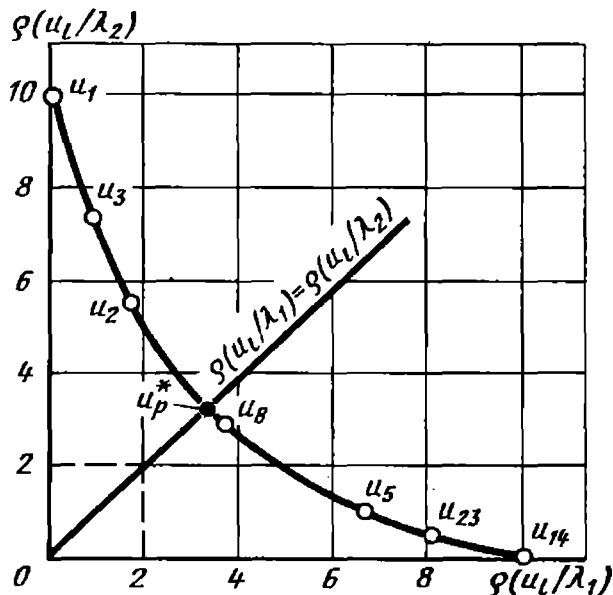


Рис. 3. Схема определения рандомизированной стратегии

где $q = (q_1, q_2, \dots, q_{s_0})^T$ — вектор вероятностей рандомизированной (смешанной) стратегии u_p ; Q — множество возможных распределений вероятностей; P — множество возможных априорных распределений состояний фактора λ ; $P(\lambda) = (P(\lambda_1), P(\lambda_2), \dots, P(\lambda_m))^T$ — вектор априорных вероятностей состояний фактора λ .

Для условий задачи $W(u_p^*) = 3,25$, что меньше $W(u_d^*) = 3,49$.

В практике исследования эффективности часто встречаются ситуации, когда неопределенный фактор λ , влияющий на исход операции, может принимать лишь два возможных значения. Такие задачи возникают при распознавании объектов и явлений, статистической проверке гипотез (см. т. 2), при проектировании автоматических систем и т. д. При этом измеряемая характеристика z фактора λ может рассматриваться как скалярная непрерывная случайная величина. В этом случае подоперация состоит в том, что по результатам измерения характеристики z необходимо принять решение о том, что фактор λ принял значение λ_1 (гипотеза H_1) или λ_2 (гипотеза H_2). Такое решение осуществляют с помощью заранее установленного порогового значения z^0 , например, следующим образом: если $z < z^0$ — принять гипотезу H_1 , если $z > z^0$ —

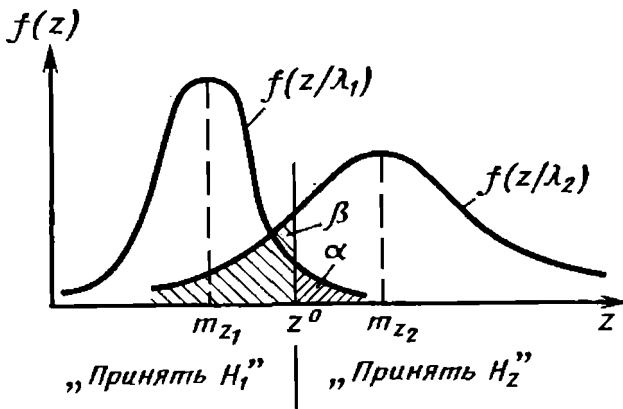


Рис. 4. Условные плотности распределения характеристики z

— H_2 . На рис. 4 изображены условные плотности распределения характеристики z — $f(z/\lambda_1)$ и $f(z/\lambda_2)$, положение которых определяется условными математическими ожиданиями m_{z_1} и m_{z_2} соответственно. Назначение порога z^0 приводит к тому, что можно совершить два рода ошибок: принять гипотезу H_2 при истинном состоянии фактора $\lambda = \lambda_1$ — ошибка первого рода; принять гипотезу H_1 при истинном состоянии $\lambda = \lambda_2$ — ошибка второго рода.

Часто эти ошибки называют «ложной тревогой» и «пропуском цели» соответственно.

Вероятности α и β ошибок первого и второго рода определяются выражениями:

$$\alpha = \alpha(z^0) = \int_{z^0}^{\infty} f(z/\lambda_1) dz = 1 - F(z^0/\lambda_1);$$

$$\beta = \beta(z^0) = \int_{-\infty}^{z^0} f(z/\lambda_2) dz = F(z^0/\lambda_2).$$

Отнесение ошибок к тому или иному роду достаточно условно. Однако во многих случаях гипотезы назначают таким образом, чтобы ошибка первого рода приводила к более серьезным последствиям, нежели ошибка второго рода.

При выборе порога руководствуются пятью основными критериями: Байеса, минимаксным, максимума апостериорной вероятности, максимума правдоподобия и Неймана—Пирсона.

Критерий Байеса применяется, когда известны априорные вероятности $P(\lambda_i)$ и матрица потерь $\| \Pi_{ki} \|$, $i = 1, 2$, $k = 1, 2$, где Π_{ki} — потери от решения о принятии гипотезы H_k при истинном состоянии $\lambda = \lambda_i$. Для определения порога z^0 используется отношение правдоподобия:

$$l = \frac{L(z/\lambda_2)}{L(z/\lambda_1)}.$$

Порог z^0 определяется как корень уравнения:

$$z^0: \frac{f(z^0/\lambda_2)}{f(z^0/\lambda_1)} = \frac{P(\lambda_1)(\Pi_{12} - \Pi_{11})}{P(\lambda_2)(\Pi_{21} - \Pi_{22})}. \quad (13.30)$$

Минимаксный критерий используется при неизвестных априорных вероятностях $P(\lambda_i)$ и заданной матрице потерь.

Порог z^0 определяется из уравнения

$$z^0: \Pi_{22} - \Pi_{11} + (\Pi_{11} - \Pi_{12}) \alpha(z^0) + (\Pi_{21} - \Pi_{22}) \beta(z^0) = 0. \quad (13.31)$$

Критерий максимума апостериорной вероятности применяется при известных априорных вероятностях $P(\lambda_i)$ и неизвестной матрице потерь:

$$z^0: \frac{f(z^0/\lambda_2)}{f(z^0/\lambda_1)} = \frac{P(\lambda_2)}{P(\lambda_1)}. \quad (13.32)$$

Критерий максимума правдоподобия используется тогда, когда не известны ни априорные вероятности, ни матрица потерь:

$$z^0: \frac{f(z^0/\lambda_2)}{f(z^0/\lambda_1)} = 1. \quad (13.33)$$

Критерий Неймана—Пирсона применяется в тех же случаях, что и критерий максимума правдоподобия, однако от ЛПР можно получить информацию о требуемом уровне α_0 вероятности ошибки первого рода:

$$z^0: 1 - F(z^0/\lambda_1) = \alpha_0. \quad (13.34)$$

Для решения практических задач и анализа различных ситуаций, связанных с выбором порога z^0 , удобно пользоваться так называемой рабочей характеристикой (рис. 5), которая представляет собой зависимость ве-

личины $1 - \beta$, называемой мощностью критерия, от величины α . С использованием рабочей характеристики помимо анализа зависимости величин α и β при использовании критерия Неймана—Пирсона можно определить ошибки α и β , соответствующие критерию Байеса и минимаксному критерию.

Для определения α и β по критерию Байеса на рабочей характеристике надо найти точку, касательная в которой имеет угол наклона:

$$\gamma = \arctg \frac{P(\lambda_1)(\Pi_{12} - \Pi_{11})}{P(\lambda_2)(\Pi_{21} - \Pi_{22})}.$$

Для определения α и β по минимаксному критерию необходимо найти точку пересечения рабочей характеристики с прямой:

$$(1 - \beta) = \frac{\Pi_{11} - \Pi_{12}}{\Pi_{21} - \Pi_{11}} \alpha + 1.$$

При неизвестных априорных вероятностях $P(\lambda)$ состояний фактора λ ЛПР вынуждено использовать либо минимаксный критерий (если задана матрица потерь), либо критерий Неймана—Пирсона (при неизвестной матрице потерь). Однако даже в этом случае у ЛПР могут быть некоторые субъективные представления о том, какие значения неопределенного фактора более возможны, а какие менее возможны. Поэтому вместо весьма «осторожного» минимаксного критерия появляется возможность использовать более информативные критерии (Байеса, максимума апостериорной вероятности), восстанавливая по информации от ЛПР субъективное априорное распределение вероятностей на множестве Λ_E .

Задача получения оценок субъективных априорных вероятностей родственна задаче выявления и измерения предпочтений на множестве Λ_E , так как, проводя операцию, ЛПР ориентируется на более вероятные (возможные) состояния λ_i по сравнению с состоянием (состояниями) λ_j . Формально это эквивалентно утверждению: «элемент λ_i важнее элемента λ_j или совокупности таких элементов».

Следовательно, для определения точечных оценок $P^*(\lambda_i)$ априорных вероятностей $P(\lambda_i)$ можно использо-

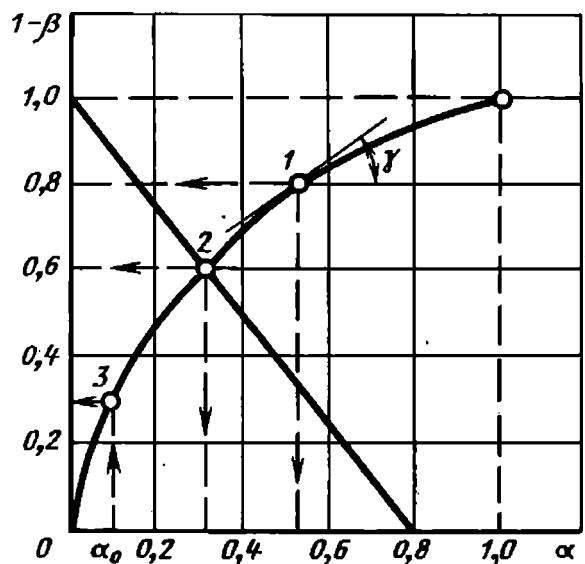


Рис. 5. Рабочая характеристика: 1 — байесовское решение; 2 — минимаксное решение; 3 — решение Неймана—Пирсона

вать аппарат, описанный на с. 206 и 211. В частности, оценки $P^*(\lambda_i)$ могут быть определены методами установления порядковых коэффициентов важности. Существуют и специальные методы оценки субъективных вероятностей. Рассмотрим те из них, которые основаны на использовании информации в форме элементарных суждений типа:

ранжирование

$$\begin{aligned} \lambda_{i_1} &\overset{p}{\geq} \lambda_{i_2} \overset{p}{\geq} \lambda_{i_3} \overset{p}{\geq} \dots \overset{p}{\geq} \lambda_{i_m} \Leftrightarrow \\ &\Leftrightarrow P(\lambda_{i_1}) \geq P(\lambda_{i_2}) \geq \\ &\geq P(\lambda_{i_3}) \geq \dots \geq P(\lambda_{i_m}); \end{aligned} \tag{13.35}$$

парное сравнение с группой

$$\begin{aligned} \lambda_{i_k} &\overset{p}{\geq} \bigcup_{j=1}^{m-k} \lambda_{i_{k+j}} \Leftrightarrow P(\lambda_{i_k}) \geq \\ &\geq \sum_{j=1}^{m-k} P(\lambda_{i_{k+j}}), \quad k = \overline{1, m-1}, \end{aligned} \tag{13.36}$$

где k — номер состояния в ранжированном ряду; парное сравнение с

двумя группами

$$\begin{aligned} \bigcup_{j=1}^{j_0+1} \lambda_{ik+j} &\stackrel{p}{\geq} \lambda_{ik} \stackrel{p}{\geq} \bigcup_{j=1}^{j_0} \lambda_{ik+j} \Leftrightarrow \\ \Leftrightarrow \sum_{j=1}^{j_0+1} P(\lambda_{ik+j}) &\geq P(\lambda_{ik}) \geq \\ \geq \sum_{j=1}^{j_0} P(\lambda_{ik+j}), &k = \overline{1, m}; \quad (13.37) \end{aligned}$$

а также при указании ЛПР интервальных оценок для каждой из априорных вероятностей

$$P_{\min}(\lambda_i) \leq P(\lambda_i) \leq P_{\max}(\lambda_i), \quad i = \overline{1, m}. \quad (13.38)$$

Здесь $\stackrel{p}{\geq}$ — отношение нестрогого предпочтения «не менее вероятен» на множестве факторов стохастической природы.

Для этих случаев получены следующие выражения для точечных оценок субъективных априорных вероятностей (точечные оценки Фишберна). Для информации (13.35) эти оценки образуют убывающую арифметическую прогрессию вида

$$P^*(\lambda_{ik}) = \frac{2(m-k+1)}{m(m+1)}, \quad k = \overline{1, m}; \quad (13.39)$$

для (13.36) они удовлетворяют рекуррентному соотношению

$$P^*(\lambda_{ik}) = \sum_{j=1}^{m-i_k} P^*(\lambda_{ik+j}) + C_h, \quad k = \overline{1, m}, \quad (13.40)$$

где $C_h = \frac{1}{2^m - 1}$, т. е. $P^*(\lambda_{im}) = C_h$;

для (13.37) они определяются рекуррентными соотношениями:

$$\left. \begin{aligned} P^*(\lambda_{ik}) &= \sum_{j=1}^{j_0} P^*(\lambda_{ik+j}) + \\ &+ 0,5P^*(\lambda_{ik+j_0+1}), \\ k &= \overline{1, m-j_0-1}; \\ P^*(\lambda_{ik}) &= \sum_{j=1}^{m-k} P^*(\lambda_{ik+j}) + \\ &+ C_q, \quad k = \overline{m-j_0, m}, \end{aligned} \right\} \quad (13.41)$$

где слагаемое C_q определяется из условия $\sum_{i=1}^m P^*(\lambda_i) = 1$; для (13.38)

точечные оценки определяются выражением

$$\begin{aligned} P^*(\lambda_i) &= P_{\min}(\lambda_i) + \\ &+ \frac{1 - \sum_{i=1}^m P_{\min}(\lambda_i)}{\sum_{i=1}^m (P_{\max}(\lambda_i) - P_{\min}(\lambda_i))} \times \\ &\times (P_{\max}(\lambda_i) - P_{\min}(\lambda_i)). \quad (13.42) \end{aligned}$$

Рассмотрим определение оценок субъективных вероятностей на следующих примерах.

Пример 5. Неопределенный фактор может принимать четыре возможных значения $\lambda_1, \lambda_2, \lambda_3, \lambda_4$. От ЛПР получена информация: $\lambda_4 \stackrel{p}{\geq} \lambda_1 \stackrel{p}{\geq} \lambda_3 \stackrel{p}{\geq} \lambda_2$. Следовательно, $P(\lambda_4) \geq P(\lambda_1) \geq P(\lambda_3) \geq P(\lambda_2)$. По (13.39) получаем

$$P^*(\lambda_4) = \frac{2(4-1+1)}{4(4+1)} = 0,4;$$

$$P^*(\lambda_1) = \frac{2(4-2+1)}{4(4+1)} = 0,3;$$

$$P^*(\lambda_3) = \frac{2(4-3+1)}{4(4+1)} = 0,2;$$

$$P^*(\lambda_2) = \frac{2(4-4+1)}{4(4+1)} = 0,1.$$

Пример 6. В условиях примера 5 получена информация $\lambda_4 \stackrel{p}{\geq} \{\lambda_1, \lambda_3, \lambda_2\}$, $\lambda_1 \stackrel{p}{\geq} \{\lambda_3, \lambda_2\}$, $\lambda_3 \stackrel{p}{\geq} \lambda_2$. Следова-

тельно, $P(\lambda_4) \geq P(\lambda_1) + P(\lambda_3) + P(\lambda_2)$, $P(\lambda_1) \geq P(\lambda_3) + P(\lambda_2)$, $P(\lambda_3) \geq P(\lambda_2)$. Применяя (13.40), получаем

$$P^*(\lambda_2) = \frac{1}{2^4 - 1} \approx 0,07;$$

$$P^*(\lambda_3) = P^*(\lambda_2) + \frac{1}{15} \approx 0,13;$$

$$P^*(\lambda_1) = P^*(\lambda_2) + P^*(\lambda_3) +$$

$$+ \frac{1}{15} \approx 0,27; \quad P^*(\lambda_4) = P^*(\lambda_2) + \\ + P^*(\lambda_3) + P^*(\lambda_1) + \frac{1}{15} \approx 0,53.$$

Пример 7. В условиях примера 5 получена следующая информация:

$\{\lambda_1, \lambda_3\} \stackrel{P}{\leq} \lambda_4 \stackrel{P}{\leq} \{\lambda_1, \lambda_3, \lambda_2\}, \{\lambda_3, \lambda_2\} \stackrel{P}{\leq} \lambda_1, \lambda_2 \stackrel{P}{\leq} \lambda_3$. Учитывая эту информацию и свойство транзитивности, упорядоченная последовательность отношений предпочтения на множестве Λ

имеет вид $\lambda_4 \stackrel{P}{\geq} \lambda_1 \stackrel{P}{\geq} \lambda_3 \stackrel{P}{\geq} \lambda_2$, т. е. $i_k = \{4, 1, 3, 2\}$ для $k = 1, 4$. Используя (13.41) и принимая во внимание, что $j_0 = 2$, находим:

$$k = 1: P^*(\lambda_4) = P^*(\lambda_1) + P^*(\lambda_3) + \\ + 0,5P^*(\lambda_2); \\ k = 2: P^*(\lambda_1) = P^*(\lambda_3) + P^*(\lambda_2) + C_q; \\ k = 3: P^*(\lambda_3) = P^*(\lambda_2) + C_q; \\ k = 4: P^*(\lambda_2) = C_q.$$

С учетом нормирующего условия

$$\sum_{i=1}^m P^*(\lambda_i) = 1 \text{ получаем}$$

$$P^*(\lambda_2) = C_q; \quad P^*(\lambda_3) = 2C_q; \\ P^*(\lambda_1) = 4C_q; \quad P^*(\lambda_4) = 6,5C_q; \\ C_q = \frac{2}{27},$$

т. е.

$$P^*(\lambda_2) \approx 0,07; \quad P^*(\lambda_3) \approx 0,15; \\ P^*(\lambda_1) \approx 0,30; \quad P^*(\lambda_4) \approx 0,48.$$

Пример 8. В условиях примера 5 получены следующие интервальные оценки априорных вероятностей:

$$0,2 \leq P(\lambda_1) \leq 0,4; \\ 0,05 \leq P(\lambda_2) \leq 0,15; \\ 0,1 \leq P(\lambda_3) \leq 0,3; \\ 0,35 \leq P(\lambda_4) \leq 0,55.$$

Используя (13.42), получаем

$$P^*(\lambda_1) \approx 0,28; \quad P^*(\lambda_2) \approx 0,1; \\ P^*(\lambda_3) \approx 0,19; \quad P^*(\lambda_4) \approx 0,43.$$

Информация типа (13.36) и (13.37) не всегда имеет место, так как весьма редко каждое из предыдущих состояний λ_{i_k} в ранжированном ряду будет более вероятно, чем хотя бы одно из последующих в этом ряду, а также, что сдвиг j_0 будет одинаков для всех k . В произвольной ситуации система предпочтений на множестве Λ_E порождает систему неравенств вида

$$AP(\lambda) + b \leq P(\lambda) \leq \\ \leq CP(\lambda) + d, \quad (13.43)$$

где A, C — неотрицательные матрицы; b, d — неотрицательные векторы.

Система (13.43) решается с помощью метода невязок.

Методами экспертного оценивания можно определять значения элементов матрицы потерь.

Построение функции эффективности в условиях стохастической неопределенности. Выше были рассмотрены принципы принятия решений в условиях стохастической неопределенности по скалярным «объективным» показателям. В этих случаях критерий (13.11) трансформировался в критерии (13.13), (13.16), (13.18) без учета информации θ_0 об отношении ЛПР к различным ситуациям в условиях стохастической неопределенности. Если получаемые с помощью «объективных» показателей решения образуют достаточно широкое подмножество, то необходимо расширять число показателей, вводя помимо математического ожидания и дисперсии другие моменты распределения и его квантили. В общем случае задача становится бесконечномерной. Поэтому целесообразно сразу установить предпочтения ЛПР на множестве распределений, а затем — стратегий с использованием (13.11). Для этого необходимо прежде всего сузить множество распределений $F_u(y)$, $u \in U$, выделив недоминируемое их подмножество¹:

$$U_0 = \{u \mid F_u(y) \leq F_v(y), \quad u, v \in U; \\ -\infty < y < \infty\}. \quad (13.44)$$

¹ Большие значения характеристики u исхода g предпочтительнее меньших.

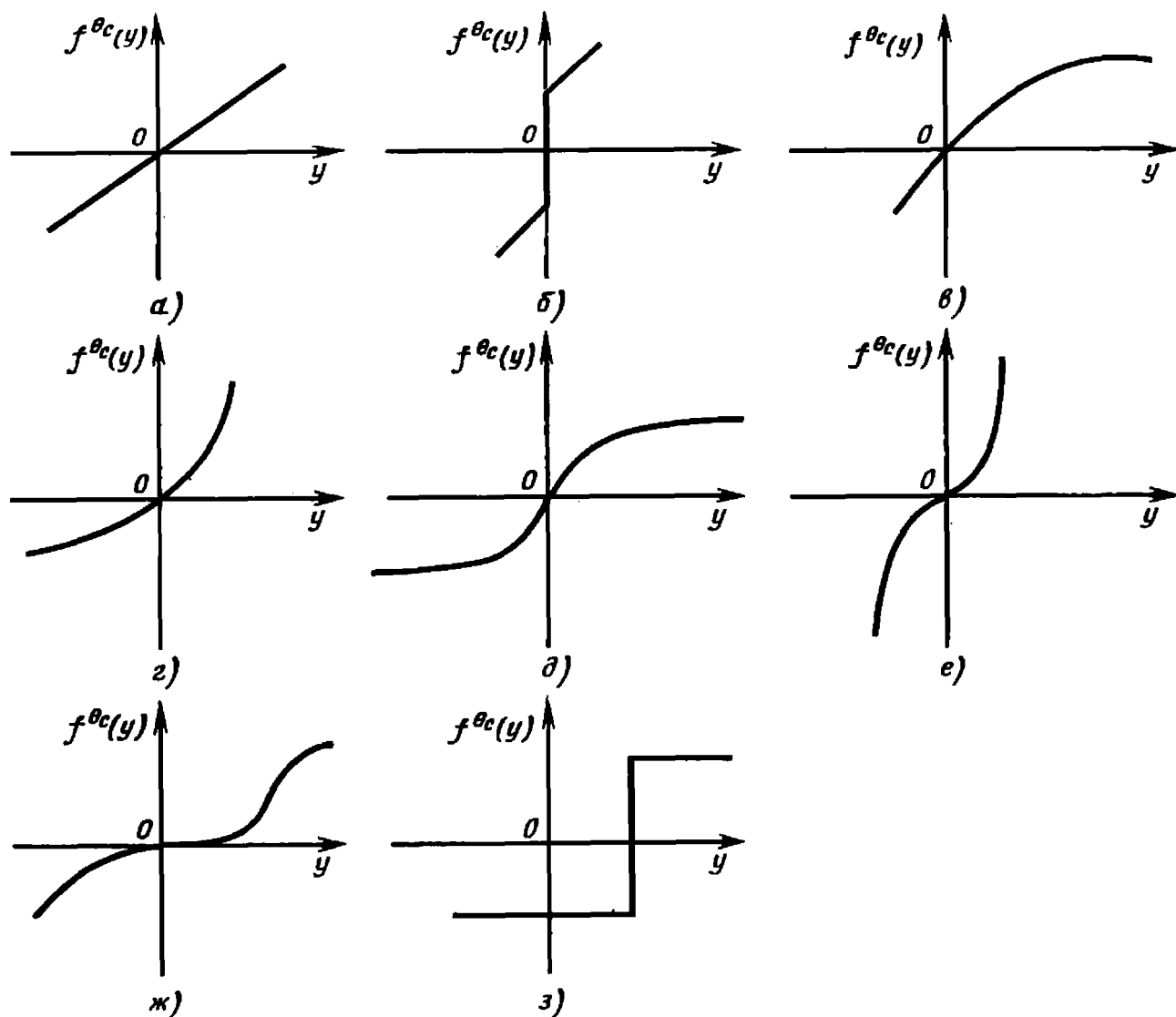


Рис. 6. Оценочные функции ЛПР с различной психологической доминантой:

а — «объективный»; б — «выигрывающий»; в — «осторожный»; г — «азартный»; д — «богач»; е — «бедняк»; ж — «заурядный»; з — «отчаянный»

Стратегии $u \in U_0$ называются *стохастически недоминируемыми*. В (13.44) информация θ_c является самой простой: «большее значение вероятности получения результата не ниже y предпочтительнее меньшего для всех значений характеристики y ». На рис. 1 стратегии $u, v \in U_0$ несравнимы по (13.44).

Как правило, множество U_0 оказывается весьма широким. Поэтому дальнейшее его сужение возможно лишь путем расширения информации θ_c . В сходных ситуациях принятия решения разные лица по-разному относятся к одному и тому же вероятностному распределению на множестве исходов. В этом случае говорят, что они имеют различную психологиче-

скую доминанту в ситуациях с риском. Формально это выражается в том, что оценочная функция $f^{\theta_c}(y)$ (принимается, что $p(y, y^{TP}) = y$) для лиц различных типов имеет различный вид. На рис. 6 представлены графики функции $f^{\theta_c}(y)$ для лиц с различной психологической доминантой.

Психологической доминанте «объективный» соответствует функция $f^{\theta_c}(y) \equiv y$, что приводит к установлению предпочтений на множестве стратегий по «объективным» показателям (13.3) или (13.44). Поясним сущность оценочной функции для ЛПР с различной психологической доминантой на следующем примере.

Пример 9. Руководство хозрасчетного объединения (ЛПР), занимающегося производством и сбытом бытовой радиоаппаратуры, должно решить вопрос: запускать ли в серийное производство в следующей пятилетке новую модель (стратегия u_1) или продолжать выпуск старой (стратегия u_2), пользуясь пока спросом. Результаты проведенного исследования (экспертного оценивания и конструкторских проработок) позволили определить облик новой модели и затраты y_{21} на ее производство, доход y_{22} объединения от продажи новой модели, если она будет пользоваться спросом, вероятность p_2 спроса, а также аналогичные величины y_{11} (потери от выпуска устаревшего образца), y_{12} (доход от продажи этого образца) и вероятность p_1 в случае продолжения выпуска «старой» продукции.

Предположим, что сравнение альтернатив u_1 и u_2 руководством предприятия по «объективным» показателям привело к следующим результатам:

$$W(u_1) = (W_1(u_1), W_2(u_1))^T = \\ = W(u_2) = (W_1(u_2), W_2(u_2))^T,$$

где $W_1(u_j) = p_j y_{j2} + (1 - p_j) y_{j1}$ — средний доход объединения по стратегии u_j , $j = 1, 2$; $W_2(u_j) = p_j [y_{j2} - W_1(u_j)]^2 + (1 - p_j) [y_{j1} - W_1(u_j)]^2$ — дисперсия дохода.

Таким образом, стратегии u_1 и u_2 с точки зрения «объективного» ЛПР (рис. 6, а) являются эквивалентными.

Если ЛПР обладает любой другой психологической доминантой, т. е. оценочная функция не является линейной ($f^{\theta_c}(y) \neq y$), то сравнение стратегий необходимо проводить следующим образом:

$$u_j \overset{\theta_c}{\geq} u_i \Leftrightarrow W(u_j) \geq W(u_i),$$

$$\text{где } W(u_j) = p_j f^{\theta_c}(y_{j2}) + (1 - p_j) \times \\ \times f^{\theta_c}(y_{j1}).$$

Это приводит к тому, что в общем случае можно указать на предпочтитель-

ность одной из стратегий. Например, одну и ту же альтернативу u со значениями равновероятных результатов $y_1 < 0$ и $y_2 > 0$ «объективный» оценит величиной $0,5(y_2 + y_1)$, а «азартный» $0,5(f^{\theta_c}(y_2) + f^{\theta_c}(y_1))$. Причем, поскольку $f^{\theta_c}(y_2) > y_2$, а $|f^{\theta_c}(y_1)| < |y_1|$, то субъективная ценность (полезность) этой альтернативы выше у «азартного» по сравнению с «объективным», т. е. если, например, $y_2 < |y_1|$, то для «объективного» ЛПР стратегия имеет отрицательную полезность (ему невыгодна), в то время как для «азартного» она может считаться выгодной (имеющей положительную полезность).

В рассмотренном примере предпочтения на множестве стратегий устанавливались путем сравнения математических ожиданий оценочной функции. Такой показатель, учитывающий психологические особенности ЛПР в ситуациях стохастической неопределенности (ситуации с риском), называется *функцией эффективности скалярного результата*¹ и определяется (13.2).

Если $f^{\theta_c}(y)$ — оценочная функция, то все ее положительные линейные преобразования

$$F^{\theta_c}(y) = \{af^{\theta_c}(y) + b, a > 0\} \quad (13.45)$$

также являются оценочными функциями характеристик исходов y для ЛПР с рассматриваемой психологической доминантой. Поэтому при построении оценочных функций $f^{\theta_c}(y)$ можно произвольно выбирать начало отсчета b и единицу измерения a , т. е. оценочная функция задается в интервальной шкале.

При исследовании эффективности почти всегда имеют дело с конечным множеством исходов G . Каждый исход $g \in G$ в рассматриваемом случае характеризуется скалярной величиной y . Поэтому в дальнейшем будем рассматривать только конечные распределения $P_i(u)$ на множестве значений $y_i \in Y$.

¹ Ее часто называют *одномерной функцией полезности*.

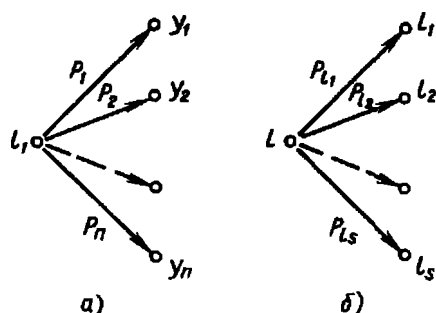


Рис. 7. Простая и сложная лотереи

Удобно описывать различные ситуации принятия решений в условиях стохастической неопределенности с помощью «лотерей».

Лотереей l называется пара (Y, P) , где $Y = \{y_1, y_2, \dots, y_n\}$ — множество характеристик исходов, $P = (p_1, p_2, \dots, p_n)$ — вероятностное распределение на нем.

На рис. 7, а приведен пример простой лотереи, а на рис. 7, б составной (сложной) лотереи, в которой исходами лотереи l являются лотереи l_j , $j = \overline{1, s}$. На основе выявленного отношения предпочтения на множестве лотерей $\{l(u) \mid u \in U\}$ строится отношение предпочтения на множестве стратегий U по правилу:

$$\begin{aligned} u \geq v &\Leftrightarrow l(u) \geq l(v) \Leftrightarrow \\ &\Leftrightarrow W_e^\lambda(l(u)) \geq W_e^\lambda(l(v)) \Leftrightarrow \\ &\Leftrightarrow M[f^{\theta_c}(y(u))] \geq M[f^{\theta_c}(y(v))]. \end{aligned} \quad (13.46)$$

Анализ (13.46) показывает, что задача сводится к получению оценок $f^{\theta_c}(y_i(u))$, $f^{\theta_c}(y_i(v))$, $i = \overline{1, n}$ и вычислению соответствующих математических ожиданий. Так как $f^{\theta_c}(\cdot)$ верна с точностью до положительного линейного преобразования (13.45), то для установления начала отсчета и единицы измерения удобно задать оценки для любых двух исходов, а затем остальные соизмерить с ними. В качестве таких исходов выбирают наиболее предпочтительный y_+ и наименее предпочтительный y_- и полагают

$$f^{\theta_c}(y_+) = 1, f^{\theta_c}(y_-) = 0.$$

При соизмерении произвольного исхода y с y_- и y_+ пользуются следующим допущением, которое называется *правилем замены*. Если в исходной лотерее l любой из исходов y заменить на эквивалентный ему по предпочтительности, то для ЛПР будет безразлично, в какой из лотерей (исходной или новой) участвовать.

Исход y в этом случае заменяют на лотерею вида y_+ с вероятностью $P(y)$ и y_- с вероятностью $1 - P(y)$. Это означает, что ЛПР должен ответить на вопрос: какова должна быть вероятность $P(y)$, чтобы для него было безразлично получение исхода y наверняка или получение наилучшего значения y_+ с вероятностью $P(y)$ и наихудшего y_- — с вероятностью $1 - P(y)$. Такое правило замены показано на рис. 8. В этом случае $f^{\theta_c}(y) = P(y)$. При числе исходов не более 10—20 описанный метод можно успешно применять для получения лотерей $l''(u)$, $u \in U$ и использовать (13.46) для установления предпочтения на множестве U стратегий. Эквивалентность лотерей l' и l'' на рис. 8, а следовательно, и $l \sim l''$ объясняется тем, что математические ожидания оценочной функции для этих лотерей одинаковы:

$$\begin{aligned} M_{l'}[f^{\theta_c}(y)] &= p_1 f^{\theta_c}(y_-) + \\ &+ p_2 [P(y_2) f^{\theta_c}(y_+) + \\ &+ (1 - P(y_2)) f^{\theta_c}(y_-)] + \dots + \\ &+ p_n [P(y_n) f^{\theta_c}(y_+) + \\ &+ (1 - P(y_n)) f^{\theta_c}(y_-)] = \\ &= \sum_{i=1}^n p_i P(y_i) = \sum_{i=1}^n p_i f^{\theta_c}(y_i) = \\ &= M_{l''}[f^{\theta_c}(y)], \end{aligned}$$

где $f^{\theta_c}(y_-) = 0$, $f^{\theta_c}(y_+) = 1$ по определению.

В задачах с большим числом исходов (более 20—30) установление предпочтительности лотерей удобнее произво-

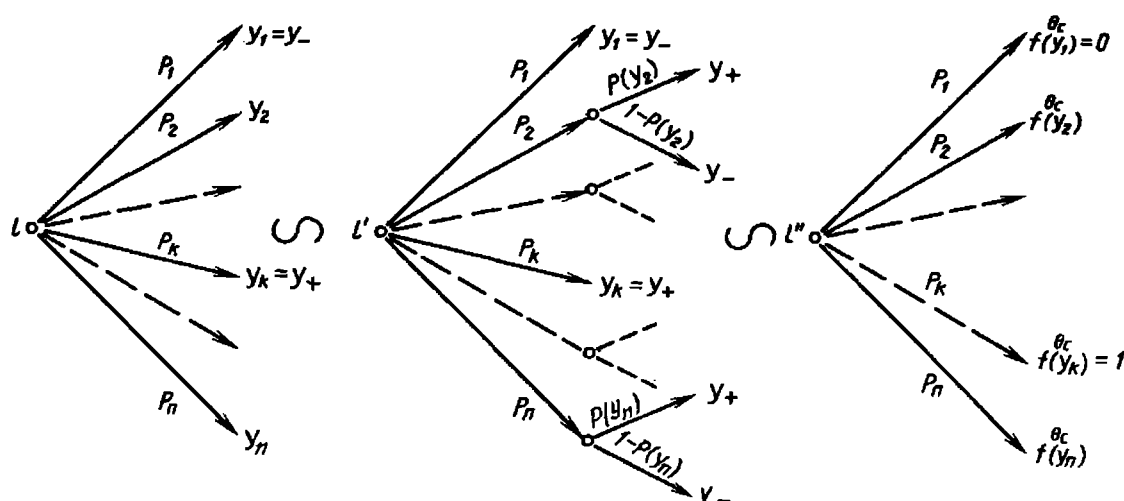


Рис. 8. Правило замены

дить другим методом, основанным на построении и аппроксимации оценочной функции $f^{\theta c}(y)$ по ограниченному числу точек. При этом вначале целесообразно установить тип отношения ЛПР к риску (несклонный, склонный, безразличный) для различных интервалов возможных значений характеристики y , а затем искать подходящую аппроксимацию оценочной функции. В общем случае на этом этапе может быть выявлена и психологическая доминанта ЛПР на всем интервале возможных значений y .

Тип отношения ЛПР к риску вводится на основе понятия достоверного (детерминированного) эквивалента лотереи [30].

Достоверным эквивалентом лотереи l называется величина y_F такая, что ЛПР безразлично получить ли исход y_F наверняка или участвовать в лотерее l :

$$y_F: f^{\theta c}(y_F) = M_l[f^{\theta c}(y)]. \quad (13.47)$$

ЛПР обладает несклонностью к риску, если

$$y_F < M_l[Y]. \quad (13.48)$$

Несклонность ЛПР к риску означает, что он всегда предпочитает наверняка получить средний выигрыш, нежели участвовать в лотерее.

ЛПР обладает склонностью к риску, если

$$y_F > M_l[Y], \quad (13.49)$$

т. е. склонное к риску ЛПР всегда предпочитает участие в лотерее получению среднего выигрыша наверняка.

ЛПР безразличен к риску, если

$$y_F = M_l[Y]. \quad (13.50)$$

Оценочная функция $f^{\theta c}(y)$ для ЛПР, несклонного к риску, строго вогнута, т. е.

$$f^{\theta c}(y_F) = M_l[f^{\theta c}(y)] < f^{\theta c}(M_l[Y]);$$

для ЛПР, склонного к риску, — строго выпукла, т. е.

$$f^{\theta c}(y_F) = M_l[f^{\theta c}(y)] > f^{\theta c}(M_l[Y]);$$

для ЛПР, безразличного к риску, — линейная, т. е.

$$f^{\theta c}(y_F) = M_l[f^{\theta c}(y)] = f^{\theta c}(M_l[Y]).$$

Так, для лотереи l с двумя исходами y_1 и y_2 и соответствующими вероятностями $p_1 = p$ и $p_2 = 1 - p$ в первом случае имеем

$$pf^{\theta c}(y_1) + (1 - p)f^{\theta c}(y_2) < f^{\theta c}(py_1 + (1 - p)y_2),$$

что соответствует определению строго вогнутой функции.

На рис. 9 приведены оценочные функции несклонного, склонного и безразличного к риску ЛПР.

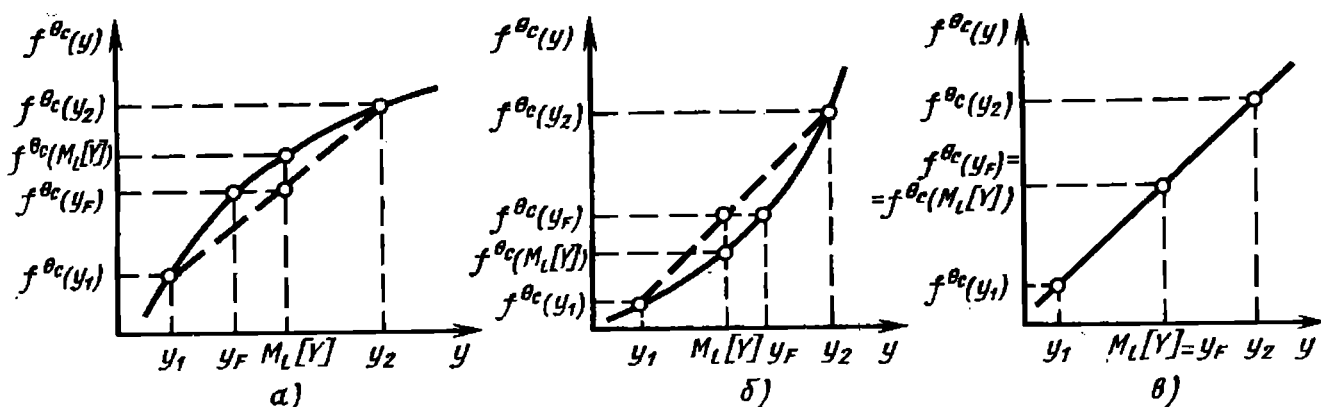


Рис. 9. Оценочные функции, учитывающие отношение ЛПР к риску:

а — «несклонный к риску»; б — «склонный к риску»; в — «безразличный к риску»

Установление типа отношений ЛПР к риску удобно осуществлять с помощью бинарных лотерей с равновероятными исходами: y_1 и y_2 , $p_1 = p_2 = 0,5$. Такую лотерею обозначим $l_{0,5}(y_1, y_2)$.

Разбиение области возможных значений характеристики y исхода на интервалы, соответствующие неизменному типу отношения ЛПР к риску, чаще всего осуществляют с помощью следующего алгоритма.

Алгоритм 1. 1. Установить наиболее предпочтительное y_+ и наименее предпочтительное y_- значения характеристики исходов.

2. Разбить интервал $[y_-, y_+]$ на k_0 одинаковых частей с граничными точками $y_0 = y_-$, $y_1, y_2, \dots, y_{k_0} = y_+$.

3. Положить $i = 1$.

4. Установить предпочтение ЛПР между лотереей $l_{0,5}(y_{i-1}, y_{i+1})$ и исходом y_i наверняка.

Если $y_i \geq l_{0,5}(y_{i-1}, y_{i+1})$, то ЛПР несклонен к риску на интервале $[y_{i-1}, y_{i+1}]$, если $l_{0,5}(y_{i-1}, y_{i+1}) \geq y_i$ — склонен к риску; в противном случае — безразличен к риску.

5. Положить $i = i + 1$ и проверить условие $i \leq k_0 - 1$; если условие вы-

полняется — перейти к п. 4; в противном случае — к п. 6.

6. Проанализировать полученные результаты с целью выявления интервалов $[y_i, y_{i+1}]$, на которых меняется отношение ЛПР к риску, т. е. интервалов, содержащих точки перегиба оценочной функции (рис. 10).

Выявление интервалов с известными типами отношения к риску позволяет построить оценочную функцию по ограниченному числу точек. Рассмотрим процедуру построения оценочной функции на интервале $[y_л, y_п]$, по пяти точкам, причем правая граница интервала $y_п$ — более предпочтительный исход, чем $y_л$.

Алгоритм 2. 1. Положить $f^{\theta c}(y_л) = 0$, $f^{\theta c}(y_п) = 1$; $a = y_л$, $b = y_п$, $i = 1$.

2. Определить достоверный эквивалент y_F лотереи $l_{0,5}(a, b)$.

3. Положить

$$y_i = y_F \text{ и } f^{\theta c}(y_i) = \begin{cases} 0,5, & \text{если } i = 1; \\ 0,25, & \text{если } i = 2; \\ 0,75, & \text{если } i = 3. \end{cases}$$

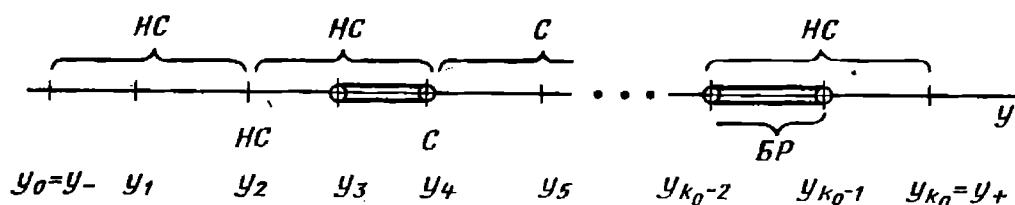


Рис. 10. Схема разбиения области возможных значений характеристики исходов на интервалы с постоянным типом отношений ЛПР к риску:

С — «склонный к риску»; НС — «несклонный к риску»; БР — «безразличный к риску»

3. Аналитические зависимости, аппроксимирующие оценочную функцию

«Несклонный к риску»		«Склонный к риску»	
$f^{\theta c}(y)$	Ограничения	$f^{\theta c}(y)$	Ограничения
$-ay^2 + by + c$	$a > 0, b \geq 0,$ $y \leq b/2a$	$a^2y^2 - by + c$	$a > 0, b \geq 0,$ $y > b/2a$
$y + a \ln(y + b)$	$y > -b, a > 0$		
$-ae^{-by} - ce^{-dy} + f$	$a \geq 0, b \geq 0,$ $c > 0, d > 0$	$ae^{by} + ce^{dy} + f$	$a \geq 0, b \geq 0,$ $c > 0, d > 0$
$\ln(y + b)$	$y > -b$		
$(y + a)^b$	$0 < b < 1,$ $y \geq -b$	$(y + a)^b$	$b > 1,$ $y > -a$
$c - (y + a)^{-b}$	$b > 0, c > 0,$ $y > -a$	$c - (y + a)^{-b}$	$b > 0, c > 0,$ $y < -a$

4. Положить $i = i + 1$ и проверить условие $i \leq 3$: да — перейти к п. 5; нет — к п. 6.

5. Положить $a = y_1$, если $i = 3$,
 $b = \begin{cases} y_1, & \text{если } i = 2; \\ y_{\Pi}, & \text{если } i = 3 \end{cases}$
и перейти к п. 2.

6. Проверить согласованность оценок, получаемых по пп. 2—5. Для этого выполнить следующие действия:

6.1. Определить с помощью ЛПР достоверный эквивалент y_F лотереи $l_{0,5}(y_2, y_3)$.

6.2. Сравнить величины y_F и y_1 . Если они отличаются незначительно, то перейти к п. 7. В противном случае указать ЛПР на выявленное противоречие и перейти к п. 2.

7. Нанести полученные точки на график и провести по ним плавную кривую, соответствующую типу ранее выявленного отношения ЛПР к риску.

8. Аппроксимировать полученную кривую одной из аналитических зави-

симостей, приведенных в табл. 3, или полиномом степени не выше пяти.

Определение достоверного эквивалента y_F лотереи $l_{0,5}(a, b)$ по п. 2 алгоритма 2 осуществляется с помощью следующей итерационной процедуры.

Алгоритм 3. 1. Положить $j = 1$, $y_{\text{л}} = a$, $y_{\text{п}} = b$.

2. Выбрать значение $y^{(j)}$, близкое к левой границе интервала $[y_{\text{л}}, y_{\text{п}}]$, так, чтобы можно было бы ожидать ответа ЛПР о том, что $l_{0,5}(a, b) \succ \succsim y^{(j)}$.

3. Проверить условия

$$l_{0,5}(a, b) \begin{cases} \succ \\ \sim \\ \prec \end{cases} y^{(j)}:$$

если $l_{0,5}(a, b) \succ y^{(j)}$ — перейти к п. 5;
если $l_{0,5}(a, b) \prec y^{(j)}$ — перейти к п. 4;
если $l_{0,5}(a, b) \sim y^{(j)}$ — перейти к п. 11.

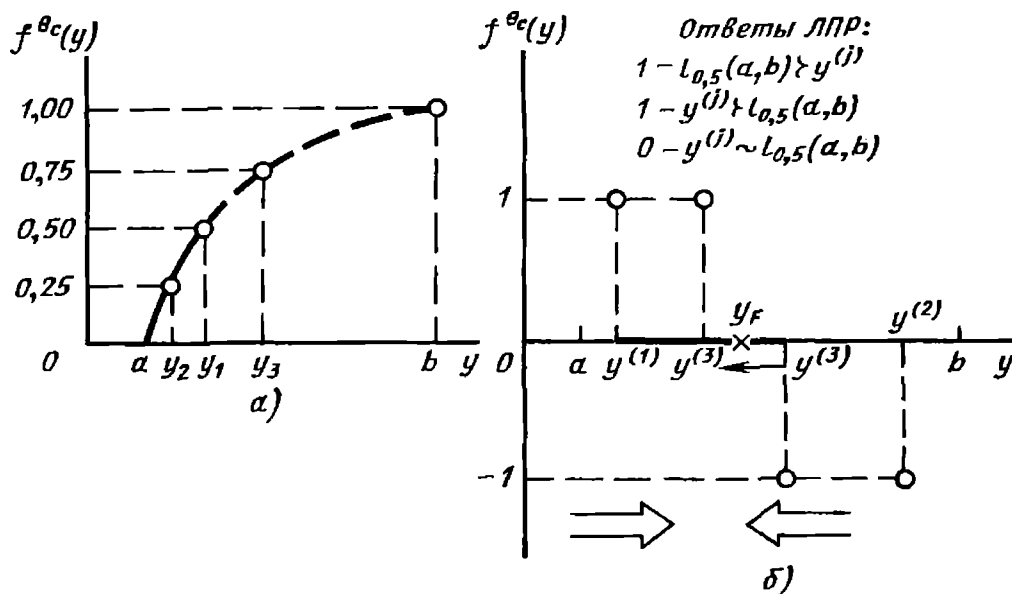


Рис. 11. Схема построения оценочной функции для выявленного типа отношения ЛПР к риску

4. Присвоить $y_{\Pi} = y^{(j)}$, $y^{(j)} = \frac{y_{\Pi} + y^{(j)}}{2}$ и перейти к п. 3.

5. Положить $j = j + 1$.

6. Выбрать значение $y^{(j)}$, близкое к правой границе интервала $[y_{\Pi}, y_{\Pi}]$ (так, чтобы можно было бы ожидать ответа ЛПР о том, что $y^{(j)} \succ l_{0,5}(a, b)$).

7. Проверить условия

$$y^{(j)} \begin{cases} \succ \\ \sim \\ \prec \end{cases} l_{0,5}(a, b):$$

если $y^{(j)} \succ l_{0,5}(a, b)$ — перейти к п. 9;

если $y^{(j)} \prec l_{0,5}(a, b)$ — перейти к п. 8;

если $y^{(j)} \sim l_{0,5}(a, b)$ — перейти к п. 11.

8. Присвоить $y_{\Pi} = y^{(j)}$, $y^{(j)} = \frac{y_{\Pi} + y^{(j)}}{2}$ и перейти к п. 7.

9. Присвоить $y_{\Pi} = y^{(j-1)}$, $y_{\Pi} = y^{(j)}$.

10. Положить $j = j + 1$ и перейти к п. 2.

11. Положить $y_F = y^{(j)}$ и окончить вычисления.

Работа алгоритмов 2 и 3 иллюстрируется рис. 11, а и б соответственно.

Используя алгоритмы 1—3, можно построить оценочные функции $f_m^{\theta_c}(y)$ для каждого m -го интервала ($m = \overline{1, m_0}$) с выявленным отношением ЛПР к риску, затем шкалировать полученные оценочные функции:

$$f^{\theta_c}(y) = \begin{cases} f_1^{\theta_c}(y), & \text{если } y_- \leq y \leq y^{(1)}; \\ f_2^{\theta_c}(y), & \text{если } y^{(1)} \leq y \leq y^{(2)}; \\ \dots \\ f_{m_0}^{\theta_c}(y), & \text{если } y^{(m_0-1)} \leq y \leq y_+ \end{cases}$$

а при необходимости перенормировать $f^{\theta_c}(y)$ таким образом, чтобы она изменялась в пределах 0—1 (рис. 12).

С помощью построенной оценочной функции $f^{\theta_c}(y)$ определяются зна-

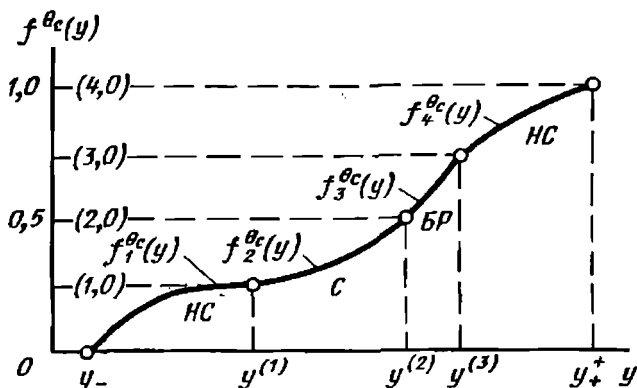


Рис. 12. Результирующая оценочная функция

чения функции эффективности для каждой стратегии $u \in U$ по (13.2), а затем сравниваются по (13.46).

ЗАДАЧИ ПРИНЯТИЯ РЕШЕНИЙ В УСЛОВИЯХ СТОХАСТИЧЕСКОЙ НЕОПРЕДЕЛЕННОСТИ ПО ВЕКТОРНОМУ ПОКАЗАТЕЛЮ

Подход к решению задач принятия решений в условиях стохастической неопределенности по векторному показателю определяется тем, какая информация о предпочтениях ЛПР и его отношении к риску имеется. Если информация θ_c об отношении ЛПР к риску отсутствует либо ЛПР безразличен к риску на всем диапазоне изменений характеристик $y_1, y_2, \dots, y_m$, то используют методы, основанные на применении «объективных» показателей. Выделенные таким образом «наилучшие» стратегии должны в обязательном порядке представляться заинтересованному лицу на рассмотрение и окончательное принятие решения. В случае, если ЛПР не удовлетворяется ни одним из представленных недоминируемых вариантов, необходимо провести анализ его отношения к риску. Обязательность такого анализа иллюстрируется следующим гипотетическим примером.

Пример 10. Пусть исход операции характеризуется вектором $Y = (y_1, y_2)^T$ и имеются две стратегии u_1 и u_2 , приводящие к следующим распределениям на исходах операции (рис. 13). Информация θ_c отсутствует, а предпочтения ЛПР выражаются стремлением ЛПР к увеличению каждой частной характеристики y_1, y_2 исхода. Пусть в качестве «объективного» показателя эффективности стратегии используется вектор математических ожиданий частных характеристик, т. е.

$$W(u_i) = (M_{u_i}[y_1], M_{u_i}[y_2])^T,$$

$$i = \overline{1, 2}.$$

Тогда стратегия u_1 явно доминирует стратегию u_2 , так как $M_{u_1}[y_1] = 5 > M_{u_2}[y_1] = 1 \wedge M_{u_1}[y_2] = 3 > M_{u_2}[y_2] = 1$.

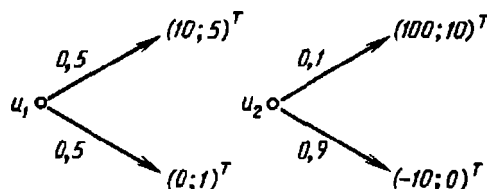


Рис. 13. Лотереи с векторными характеристиками исхода операции

Однако при представлении ЛПР стратегий u_1 и u_2 он может указать, что $u_2 \succ u_1$.

Тем не менее подход, основанный на использовании «объективных» показателей, весьма распространен. В этом случае принципы отыскания наилучших стратегий сходны с принципами принятия решения по векторному показателю в условиях определенности. Отличие состоит лишь в используемых при этом показателях. Использование критерия пригодности и функции соответствия типа (13.4) приводит к принципу приемлемого результата:

$$U^{SAT} = \{ u \mid P(g_1(u) \geq y_1^{TP}, g_2(u) \geq y_2^{TP}, \dots, g_m(u) \geq y_m^{TP}) \geq P_{TP} \} \quad (13.51)$$

или к принципу пригодности в форме (13.9) с подстановкой вместо $W_i(u)$ и W_i^{TP} моментов случайной величины g_i или ее квантилей.

Для использования критерия оптимальности необходимо предварительно «свернуть» векторный показатель в скалярный или использовать вместо векторной характеристики исхода скалярную (агрегирующую). Во всех случаях это можно осуществить лишь на основе дополнительной информации от ЛПР о важности частных показателей эффективности, или характеристик исхода (см. гл. 12).

Наиболее простая информация касается стремления ЛПР обеспечить как можно большее значение вероятности достижения цели операции

$$U^* = \left\{ u^* \mid \max_{u \in U} P(g_1(u) \geq y_1^{TP}, g_2(u) \geq y_2^{TP}, \dots, g_m(u) \geq y_m^{TP}) \right\}. \quad (13.52)$$

Если информация от ЛПР позволяет выделить одну из характеристик исхода, например y_1 , в качестве «главной», то для отыскания наилучших стратегий часто используют одну из следующих постановок задач:

$$u^*: \max_{u \in U_1} W_1(u), \quad (13.53)$$

где $U_1 = \{u \mid P(\vartheta_2(u) \geq y_2^{TP}, \hat{y}_3(u) \geq y_3^{TP}, \dots, \vartheta_m(u) \geq y_m^{TP}) > P_{TP}\}$;

$$u^*: \max_{u \in U_2} W_1(u), \quad (13.54)$$

где $U_2 = \{u \mid W_j(u) \geq W_j^{TP}, j = \overline{2, m}\}$.

В этих задачах при формировании показателей $W_i(u)$ можно применять функции соответствия типа (13.4)—(13.8).

Для лексикографического упорядочения характеристик y_i , $i = \overline{1, m}$ исхода операции и выбора для всех одной и той же функции $\rho(y_i, y_i^{TP})$ задача может иметь следующий вид:

$$u^*: \text{lex max}_{u \in U} (M[\rho(y_1(u), y_1^{TP})], M[\rho(y_2(u), y_2^{TP})], \dots, M[\rho(y_m(u), y_m^{TP})])^T; \quad (13.55)$$

если для строго упорядоченных характеристик также указан наиболее предпочтительный тип используемой функции соответствия, то

$$u^*: \text{lex max}_{u \in U} (W_1(u), W_2(u), \dots, W_m(u))^T; \quad (13.56)$$

если для строго упорядоченных частных характеристик y_i используется лексикографический набор частных показателей

$$W^i(u) = (W_1^i(u), W_2^i(u), \dots, W_i^i(u), \dots, W_m^i(u))^T,$$

то

$$u^*: \text{lex max}_{u \in U} \text{lex max}_i (W^i(u), W^2(u), \dots, W^i(u), \dots, W^m(u))^T. \quad (13.57)$$

Если в качестве частных показателей исхода используются показатели вероятностной гарантии или вероятно-гарантированного результата, то при получении информации об упорядочении этих показателей может быть использован метод последовательных уступок. Однако в этом случае необходимо участие ЛПР в процедуре отыскания «наилучшей» стратегии.

Иногда частные показатели эффективности (или частные характеристики исхода) удается свернуть в агрегированный показатель, используя, например, информацию о коэффициентах важности. Тогда

$$u^*: \max_{u \in U} \varphi(W_1(u), W_2(u), \dots, W_m(u)), \quad (13.58)$$

или

$$u^*: \max_{u \in U} M[\varphi(y_1(u), y_2(u), \dots, y_m(u))], \quad (13.59)$$

где $\varphi(\cdot)$ — функция агрегирования, задаваемая в общем случае (12.74).

Практически важным частным случаем является задача с двумя показателями эффективности $W_1^i(u) = M[\vartheta_i(u)]$, $W_2^i(u) = D[\vartheta_i(u)]$, $i = \overline{1, m}$, предпочтения между которыми не являются строго лексикографическими. В этом случае задачу часто решают с использованием различных функций агрегирования:

$$u^*: \text{lex max}_{u \in U} (\varphi^1(u), \varphi^2(u), \dots, \varphi^m(u)), \quad (13.60)$$

где $\varphi^i(u) = \varphi(M[\vartheta_i(u)], D[\vartheta_i(u)])$.

В качестве $\varphi^i(u)$ часто используют коэффициент вариации¹, т. е.

$$\varphi^i(u) = \frac{\sqrt{D[g_i(u)]}}{M[g_i(u)]}, \quad M[g_i(u)] \neq 0, \\ i = \overline{1, m}. \quad (13.61)$$

Однако эта свертка, хотя и имеет ясный физический смысл, в то же время ей присущи все недостатки функции (12.78).

Другим видом свертки является использование функции агрегирования вида

$$\varphi^i(u) = M \left[\left(g_i(u) - \max_{v \in U} M[g_i(v)] \right)^2 \right] \quad (13.62)$$

или

$$\varphi^i(u) = M \left[\left(y_i(u) - \frac{1}{s} \sum_{u_s \in U} M[g_i(u_s)] \right)^2 \right], \quad (13.63)$$

где s — число стратегий дискретного множества U .

Функции агрегирования (13.62) и (13.63) по смыслу являются моментами второго порядка, вычисляемыми относительно максимального среднего значения характеристики исхода и относительно среднего арифметического по стратегиям ее математических ожиданий; поэтому минимизация (13.62) или (13.63) будет одновременно обеспечивать максимизацию математического ожидания характеристики и минимизацию ее дисперсии.

Удобно учитывать предпочтения ЛПР на частных показателях $M[g_i(u)]$ и $D[g_i(u)]$ с помощью их однородной линейной свертки:

$$\varphi^i(u) = (1-c) M^2[g_i(u)] - c D[g_i(u)], \quad (13.64)$$

где $0 \leq c \leq 1$ — коэффициент, учитывающий предпочтения ЛПР. При

¹ В этом случае в (13.60) следует искать лексикографический минимум или положить $\varphi^i(u) = -\varphi^i(u)$.

$c = 0 \quad M[g_i(u)] \succ^{lex} D[g_i(u)]$, а при $c = 1 \quad D[g_i(u)] \succ^{lex} M[g_i(u)]$, что эквивалентно задаче (13.64).

Если эти предпочтения не носят лексикографического характера, то измерять их удобно с помощью следующих реперных значений коэффициента c :

$$c' = \min_{u \in U} \frac{M^2[g_i(u)]}{M[g_i^2(u)]}, \\ c'' = \max_{u \in U} \frac{M^2[g_i(u)]}{M[g_i^2(u)]}, \quad (13.65)$$

причем $0 \leq c' \leq c'' \leq 1$.

Если $M[g_i(u)] \succ D[g_i(u)]$, то $0 \leq c \leq c'$; если $D[g_i(u)] \succ M[g_i(u)]$, то $c'' \leq c \leq 1$; в противном случае $M[g_i(u)] \sim D[g_i(u)]$ $c' \leq c \leq c''$.

Для выбора c в интервале $[0, c']$ используют точечные оценки вида

$$c'(u, a) = \left(\frac{n}{n-1} \right)^{\frac{a}{2}} c' \rho^a(P(u)), \quad (13.66)$$

в интервале $[c'', 1]$

$$c''(u, a) = 1 - \left(\frac{n}{n-1} \right)^{\frac{a}{2}} \times \\ \times \rho^a(P(u)) (1 - c''), \quad (13.67)$$

в интервале (c', c'')

$$c(u, a) = c' + \left(\frac{n}{n-1} \right)^{\frac{a}{2}} \times \\ \times \rho^a(P(u)) (c'' - c'), \quad (13.68)$$

где a — произвольное неотрицательное число (обычно $a = 0, 1, 2$); $\rho(P(u))$ — расстояние от распределения $P(u) = (P_1(u), P_2(u), \dots, P_n(u))^T$ до равновероятного распределения $\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n} \right)^T$ определяемое по формуле

$$\rho(P(u)) = \sqrt{\sum_{i=1}^n \left(P_i(u) - \frac{1}{n} \right)^2}, \quad (13.69)$$

4. Точечные оценки коэффициента *c*

Предпочтение	Интервал (граница)	Параметр <i>a</i>		
		0	1	2
$M[g(u)] \succ^{lex} D[g(u)]$	0	Любой		
$M[g(u)] \succ D[g(u)]$	$[0, c']$	c'	$\sqrt{\frac{n}{n-1}} \rho(P(u)) c'$	$\frac{n}{n-1} \rho^2(P(u)) c'$
$M[g(u)] \sim D[g(u)]$	(c', c'')	c''	$c' + \sqrt{\frac{n}{n-1}} \times$ $\times \rho(P(u)) (c'' - c')$	$c' + \frac{n}{n-1} \times$ $\times \rho^2(P(u)) (c'' - c')$
$M[g(u)] \prec D[g(u)]$	$[c'', 1]$	c''	$1 - \sqrt{\frac{n}{n-1}} \times$ $\times \rho(P(u)) (1 - c'')$	$1 - \frac{n}{n-1} \times$ $\times \rho^2(P(u)) (1 - c'')$
$M[g(u)] \prec^{lex} D[g(u)]$	1	Любой		

n — число возможных значений величины *g_i*, *i* = 1, *m*.

Точечные оценки (13.66)—(13.68) коэффициента *c* для *a* = 0, 1, 2 приведены в табл. 4.

Иногда используется следующий прием. Рассматривая характеристики *y_i*, *i* = 1, *m* как неслучайные, строят функцию эффективности *W_e*(*Y*), например в форме (12.93), а затем определяют $W_e^\lambda(u) = M[W_e(\hat{Y}(u))]$. Однако использование этого приема теоретически не обосновано. Более того, учитывая предпочтения ЛПР на исходах, он предполагает безразличие к риску на множестве лотерей.

Общим по-прежнему остается подход, основанный на использовании информации θ_c об отношении ЛПР к риску и построении на ее основе функции эффективности $W_e^\lambda = M[f^{\theta_c}(\rho(Y))]$.

Построение оценочной функции

$f^{\theta_c}(Y)$. Если число возможных исходов операции *n* невелико (20—30), то функцию $f^{\theta_c}(Y)$ можно построить путем непосредственной оценки. Для этого среди всех исходов $Y = (y_1, y_2, \dots, y_m)^T$ выбирается наименее предпочтительный *Y₋* и наиболее предпочтительный *Y₊*. Устанавливается начало отсчета и масштаб измерения по правилу:

$f^{\theta_c}(Y_-) = 0, \quad f^{\theta_c}(Y_+) = 1.$

Для каждого исхода $Y^i, i = \overline{1, n}$ с помощью ЛПР устанавливается значение вероятности *P*(*Yⁱ*) из условия

$l_p(Y^-, Y^+) \sim Y^i, \tag{13.70}$

где *l_p*(*Y⁻*, *Y⁺*) — лотерея, приводящая с вероятностью *p* к исходу *Y⁺* и

с вероятностью $1 - p$ — к исходу Y^- , т. е. Y^i является достоверным эквивалентом лотереи. В этом случае

$$f^{\theta c}(Y^i) = P(Y^i).$$

Если число возможных исходов операции n больше 20—30, то функцию $f^{\theta c}(Y)$ иногда аппроксимируют по ее значениям, полученным для ограниченной представительной части исходов прямым методом оценивания. После этого в обязательном порядке должна осуществляться проверка полученной аппроксимации с помощью ЛПР. Однако этот метод в силу его трудоемкости не всегда приводит к успеху, так как трудно подобрать и в дальнейшем корректировать аппроксимирующую зависимость, не имея априорной информации о ее виде, и, кроме того, полученная зависимость не обязательно отражает структуру предпочтений ЛПР на всем множестве исходов операции.

Поэтому целесообразно (а обычно так и делают) вначале установить справедливость некоторых допущений о структуре предпочтений ЛПР на множестве лотерей, на основании которых можно вынести суждение о виде оценочной функции $f^{\theta c}(Y)$. Это позволяет использовать лишь ограниченное число точек для ее построения. Некоторые из таких допущений позволяют представить оценочную функцию в виде

$$f^{\theta c}(Y) = \Phi^{\theta c}(f_1^{\theta c}(y_1), f_2^{\theta c}(y_2), f_3^{\theta c}(y_3), \dots, f_m^{\theta c}(y_m)), \quad (13.71)$$

где $\Phi^{\theta c}(\cdot)$ — функция агрегирования частных оценочных функций $f_i^{\theta c}(y_i)$, $i = \overline{1, m}$, а также указать конкретный вид функции $\Phi^{\theta c}(\cdot)$.

Наибольшее применение получили функции $\Phi^{\theta c}(y)$ в аддитивной, мультипликативной и полилинейной форме соответственно:

$$f^{\theta c}(Y) = \sum_{i=1}^m \gamma_i f_i^{\theta c}(y_i), \quad (13.72)$$

$$f^{*\theta c}(Y) = (\text{sign } \gamma_0)^m \prod_{i=1}^m f_i^{*\theta c}(y_i); \quad (13.73)$$

$$\begin{aligned} f^{\theta c}(Y) = & \sum_{i=1}^m \gamma_i f_i^{\theta c}(y_i) + \\ & + \gamma_0 \sum_{i < j} \gamma_i \gamma_j f_i^{\theta c}(y_i) f_j^{\theta c}(y_j) + \\ & + \gamma_0^2 \sum_{i < j < k} f_i^{\theta c}(y_i) f_j^{\theta c}(y_j) f_k^{\theta c}(y_k) + \\ & + \dots + \gamma_0^{m-1} \prod_{i=1}^m \gamma_i f_i^{\theta c}(y_i), \end{aligned} \quad (13.74)$$

$$\begin{aligned} \text{где } f^{*\theta c}(Y) = \\ = \begin{cases} \gamma_0 f^{\theta c}(Y) + 1, & \text{если } \gamma_0 > 0; \\ -[\gamma_0 f^{\theta c}(Y) + 1], & \text{если } \gamma_0 < 0; \end{cases} \end{aligned}$$

$f_i^{*\theta c}(y_i)$ определяется по (13.79); γ_i , $i = \overline{0, m}$ — шкалирующие константы, причем $|\gamma_0| < 1$, $0 < \gamma_i < 1$, $i = \overline{1, m}$.

Аддитивная и мультипликативная формы представления являются частными случаями полилинейной оценочной функции.

Оценочная функция $f^{\theta c}(Y)$ может быть представлена в полилинейной форме (13.74), если выполняется условие взаимной независимости характеристик $y_1, y_2, \dots, y_m$ на лотереях.¹ Понятие взаимной независимости базируется на понятии независимости групп Y^1, Y^2 характеристик из множества $\{y_1, y_2, \dots, y_m\}$ на лотереях [30].

Характеристики Y^1 независимы на лотереях от характеристик Y^2 , если при фиксированных значениях компонент группы Y^2 предпочтения между лотереями, исходы которых различаются лишь значениями характеристик из Y^1 , не зависят от самих фиксированных значений характеристик из Y^2 .

Таким образом, понятие независимости по предпочтению является част-

¹ В зарубежной литературе этот вид отношения на характеристиках носит название взаимной независимости по полезности.

ным случаем понятия независимости на лотереях.

Характеристики $y_1, y_2, \dots, y_m$ взаимонезависимы на лотереях, если любое их подмножество Y^1 не зависит на лотереях от своего дополнения $Y^2 = Y \setminus Y^1$.

При большой размерности вектора характеристик $Y = (y_1, y_2, \dots, y_m)^T$ проверка взаимной независимости на лотереях непосредственно по определению практически невозможна. Поэтому обычно используют более простые условия, приводящие к взаимонезависимости характеристик на лотереях для $m \geq 3$:

характеристики $y_1, y_2, \dots, y_m$ взаимонезависимы на лотереях, если среди них имеется одна характеристика $Y^1 = y_h$, которая не зависит на лотереях от $Y^2 = Y \setminus y_h$, а все характеристики $y_1, y_2, \dots, y_m$ взаимонезависимы по предпочтению.

Проверка взаимонезависимости по предпочтению была рассмотрена выше (см. алгоритм 4, гл. 12).

Проверка условия независимости характеристики, например, y_1 на лотереях от остальных $Y^2 = Y \setminus y_1$ может быть осуществлена с помощью следующего алгоритма.

Алгоритм 4. 1. Зафиксировать характеристики группы Y^2 на уровне:

$$Y^2_- = (y_2^-, y_3^-, \dots, y_m^-).$$

2. С помощью алгоритма 3 установить достоверный эквивалент y_{1F}^- лотереи $l_{0,5}((y_1^-, Y^2_-), (y_1^+, Y^2_+))$.

3. Зафиксировать характеристики группы Y^2 на уровне:

$$Y^2_+ = (y_2^+, y_3^+, \dots, y_m^+).$$

4. С помощью алгоритма 3 установить достоверный эквивалент y_{1F}^+ лотереи $l_{0,5}((y_1^-, Y^2_+), (y_1^+, Y^2_+))$.

5. Зафиксировать характеристики группы Y^2 на уровне:

$$\bar{Y}^2 = \frac{Y^2_- + Y^2_+}{2}.$$

6. С помощью алгоритма 3 установить достоверный эквивалент $\bar{y}_{1F}$ лотереи $l_{0,5}((y_1^-, \bar{Y}^2), (y_1^+, \bar{Y}^2))$.

7. Сравнить величины $\bar{y}_{1F}$, y_{1F}^+ , y_{1F}^- . Если они примерно одинаковы, то y_1 не зависит на лотереях от Y^2 .

В алгоритме 4 в пп. 4, 6 при использовании алгоритма 3 следует выбирать разные значения $y^{(j)}$ (см. п. 2 алгоритма 3), отличающиеся от этих значений в п. 2.

Теорема 1. Если характеристики $y_1, y_2, \dots, y_m$ взаимонезависимы на лотереях, то оценочная функция $f^{\theta c}(Y)$ представима в форме (13.74), причем:

$$f^{\theta c}(Y_-) = 0, \quad f^{\theta c}(Y_+) = 1; \quad (13.75)$$

$$f_i^{\theta c}(y_i^-) = 0, \quad f_i^{\theta c}(y_i^+) = 1; \quad (13.76)$$

$$\gamma_i = f^{\theta c}(Y_-^i); \quad (13.77)$$

$$\gamma_0: 1 + \gamma_0 = \prod_{i=1}^m (1 + \gamma_0 \gamma_i), \quad (13.78)$$

где $Y_- = (y_1^-, y_2^-, \dots, y_m^-)^T$,

$$Y_+ = (y_1^+, y_2^+, \dots, y_m^+)^T;$$

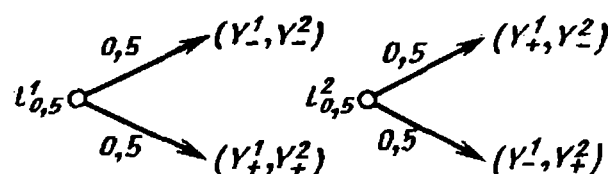
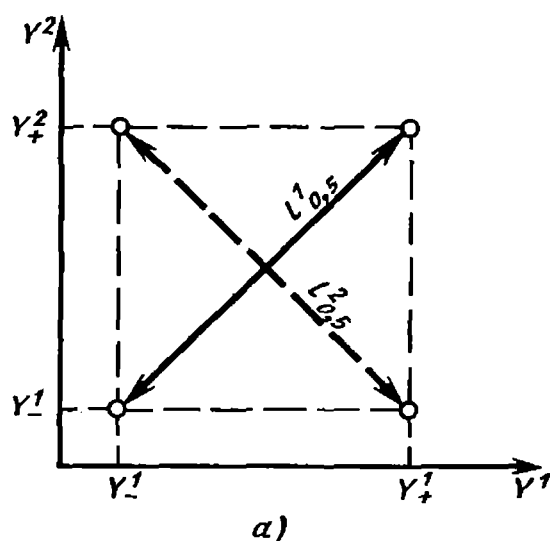
$$Y_-^i = (y_1^-, y_2^-, \dots, y_{i-1}^-, y_i^+, y_{i+1}^-, \dots, y_m^-)^T.$$

Если $\sum_{i=1}^m \gamma_i = 1$, то из (13.75) и (13.74) следует, что $\gamma_0 = 0$ и (13.74) принимает аддитивную форму (13.72).

Если $\sum_{i=1}^m \gamma_i \neq 1$, то $\gamma_0 \neq 0$ и (13.74) сводится к (13.73) со значениями:

$$\begin{aligned} f_i^{*\theta c}(y_i) &= \\ &= \begin{cases} 1 + \gamma_0 \gamma_i f_i^{\theta c}(y_i), & \text{если } \gamma_0 > 0; \\ -[1 + \gamma_0 \gamma_i f_i^{\theta c}(y_i)], & \text{если } \gamma_0 < 0. \end{cases} \end{aligned} \quad (13.79)$$

Для установления конкретного частного вида полилинейной оценочной функции (аддитивная или мультипликативная) после того, как выполнены условия теоремы 1, используют следующую процедуру, основанную на следствии из этой теоремы:



Частные распределения:

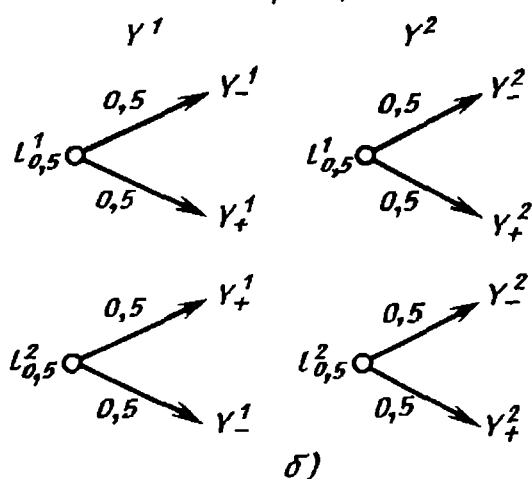


Рис. 14. Аддитивная независимость групп характеристик исходов

Алгоритм 5. 1. Выбрать две произвольные характеристики, например, y_1 и y_2 из $Y = (y_1, y_2, \dots, y_m)^T$
2. Зафиксировать остальные характеристики y_i , $i = \overline{3, m}$ на уровне

$$\bar{y}_i = \frac{y_i^+ + y_i^-}{2}.$$

3. Установить с помощью ЛПР равноценность лотерей

$$l_{0,5}((y_1^-, y_2^-, \bar{Y}^i), (y_1^+, y_2^+, \bar{Y}^i)) \text{ и } l_{0,5}((y_1^-, y_2^+, \bar{Y}^i), (y_1^+, y_2^-, \bar{Y}^i)),$$

где $\bar{Y}^i = (\bar{y}_3, \bar{y}_4, \dots, \bar{y}_m)$; если лотереи эквивалентны, то $f^{\theta c}(Y)$ — аддитивная, в противном случае — мультипликативная.

Другим условием, приводящим к аддитивной форме оценочной функции, является условие аддитивной независимости [30].

Характеристики $y_1, y_2, \dots, y_m$ являются аддитивно независимыми, если предпочтительность лотерей на $y_1, y_2, \dots, y_m$ зависит только от их частных

распределений и не зависит от их совместного распределения.

Для проверки аддитивной независимости характеристик удобно сравнивать по предпочтительности лотереи лишь с двумя равновероятными исходами и двумя группами характеристик Y^1 и $Y^2 = Y \setminus Y^1$. Например, если для ЛПР оказались одинаковыми по предпочтительности лотереи $l_{0,5}^1((Y_-^1, Y_-^2), (Y_+^1, Y_+^2))$ и $l_{0,5}^2((Y_+^1, Y_-^2), (Y_-^1, Y_+^2))$ (см. рис. 14, а) с произвольными значениями $Y_-^1, Y_+^1, Y_-^2, Y_+^2$, то все частные распределения каждой из характеристик (групп характеристик) совпадают (рис. 14, б), а следовательно, характеристики Y^1 и Y^2 являются аддитивно независимыми.

При этом группы характеристик Y^1 и Y^2 для проверки аддитивной независимости $y_1, y_2, \dots, y_m$ формируются следующим образом: положить $Y^1 = y_1$, $Y^2 = \{y_2, y_3, \dots, y_m\}$ и проверить условие аддитивной независимости групп Y^1 и Y^2 ; затем положить $Y^1 = y_2$, $Y^2 = \{y_3, y_4, \dots, y_m\}$ и проверить условие аддитивной независимости; процедуру продолжить до $Y^1 = y_{m-1}$, $Y^2 = y_m$.

Построение аддитивной оценочной функции $f^{\theta c}(Y)$ может быть осуществ-

влено с помощью следующего алгоритма, основанного на свойстве

$$\sum_{i=1}^m \gamma_i = 1.$$

Алгоритм 6. 1. С помощью алгоритма 2 построить частные оценочные функции $f^{\theta c}(y_i)$, $i = \overline{1, m}$, полагая $y_i = y_j$, $i \neq j$.

2. Подобрать $(m - 1)$ точку Y^{J_k} , $k = \overline{1, m - 1}$ пространства характеристик $Y = (y_1, y_2, \dots, y_m)$ исходов, причем $y_i^- \leq y_i \leq y_i^+$, где J_k — множество номеров характеристик, установленных на уровне y_j^+ , $j \in J_k$, причем остальные характеристики с номерами $l \notin J_k$ установлены на уровне y_l^- , а $J_1 \neq J_2 \neq \dots \neq J_{m-1}$.

3. С помощью ЛПР для каждой точки Y^{J_k} установить такие значения вероятностей P_{J_k} , для которых оказываются эквивалентными точка Y^{J_k} и лотерея $l_{P_{J_k}}(Y_+, Y_-)$, приводящая к Y_+ с вероятностью P_{J_k} .

4. Составить и разрешить систему уравнений:

$$\left. \begin{aligned} \sum_{i \in J_1} \gamma_i &= P_{J_1}; \\ \sum_{i \in J_2} \gamma_i &= P_{J_2}; \\ &\vdots \\ \sum_{i \in J_{m-1}} \gamma_i &= P_{J_{m-1}}; \\ \sum_{j=1}^m \gamma_j &= 1. \end{aligned} \right\}$$

5. Проверить согласованность полученных шкалирующих констант γ_i , $i = \overline{1, m}$. Для этого подобрать точку Y^{J_m} , у которой все компоненты, кроме одной, например y_s , зафиксированы на верхнем уровне. Оценить для нее по п. 3 вероятность P_{J_m} и вычислить $\gamma_s = 1 - P_{J_m}$.

Полученный результат сравнить с соответствующим результатом по п. 4.

При несогласованности полученных результатов указать ЛПР на выявленные противоречия и повторить процедуру вновь.

6. Записать оценочную функцию $f^{\theta c}(Y)$ в аддитивной форме.

В зависимости от подготовленности ЛПР алгоритм 6 может быть модифицирован. Например, если ЛПР в состоянии сравнить по предпочтительности оценки вида $Y_-^l = (y_1^-, y_2^-, \dots, y_{l-1}^-, y_i^+, y_{i+1}^-, \dots, y_m^-)$ с лотереей $l_{P_i}(Y_+, Y_-)$, т. е. $J_k = i$, то шкалирующие коэффициенты $\gamma_i = P_i$, $i = \overline{1, m - 1}$.

В зависимости от конкретно рассматриваемой задачи иногда удобнее оценивать шкалирующие коэффициенты путем последовательного сужения множества J номеров характеристик, установленных на уровне y_j^+ , $j \in J$. В этом случае сначала множество J принимают достаточно широким (вплоть до $J = \{1, 2, 3, \dots, m\}$). Оценивается вероятность P_J . Далее множество J сужается до $J_1 \subset J$ и оценивается вероятность $P_{J_1 \setminus J}$:

$$Y_{J_1} \sim l_{P_{J_1 \setminus J}}(Y_J, Y_-).$$

Тогда

$$\sum_{i \in J_1} \gamma_i = P_{J_1 \setminus J} P_J.$$

Построение мультипликативной оценочной функции $f^{\theta c}(Y)$ может быть осуществлено с помощью следующего алгоритма.

Алгоритм 7. 1. С помощью алгоритма 2 построить частные оценочные функции $f^{\theta c}(y_i)$, $i = \overline{1, m}$, полагая $y_i = y_j$, $i \neq j$.

2. Проранжировать константы γ_j в порядке их убывания, выполнив следующие операции:

2.1. Положить $i = 1$.

2.2. Сравнить по предпочтительности исходы:

$$Y_-^i \left\{ \begin{array}{l} \succ \\ \sim \\ \prec \end{array} \right\} Y_-^{i+1} \Rightarrow \gamma_i \left\{ \begin{array}{l} > \\ = \\ < \end{array} \right\} \gamma_{i+1}.$$

2.3. Положить $i = i + 1$ и проверить условие $i \leq m - 1$: да — перейти к п. 2.2; нет — к п. 2.4.

2.4. Построить транзитивное замыкание, приводящее к искомому ранжированию коэффициентов γ_i :

$$\gamma_{i_1} \geq \gamma_{i_2} \geq \gamma_{i_3} \geq \dots \geq \gamma_{i_m}.$$

3. С помощью ЛПР оценить коэффициент γ_{i_1} , исходя из условия:

$$Y_{i_1} \sim l_{P_{i_1}}(Y_+, Y_-),$$

и положить $\gamma_{i_1} = P_{i_1}$.

4. Для каждого $k = \overline{2, m}$ с помощью ЛПР установить такое значение характеристики с номером i_1 , чтобы выполнялось условие:

$$Y_{i_1}^k \sim (y_1^-, y_2^-, \dots, y_{i_1}^-,$$

$$y_{i_k}^-, \dots, y_m^-).$$

и найти γ_{i_k} из уравнения

$$\gamma_{i_1} f_{i_1}^{\theta c}(y_{i_1}) = \gamma_{i_k}.$$

5. Определить коэффициент γ_0 из уравнения¹

$$1 + \gamma_0 = \prod_{i=1}^m (1 + \gamma_0 \gamma_i).$$

6. Записать оценочную функцию $f^{\theta c}(Y)$ в мультипликативном виде (13.73) с использованием (13.79).

Если ЛПР в состоянии сравнивать любые точки из пространства характеристик $Y = (y_1, y_2, \dots, y_m)$, то вместо пп. 3 и 4 алгоритма 7 можно использовать другую их модификацию:

¹ При решении этого уравнения следует иметь в виду, что если $\sum_{i=1}^m \gamma_i > 1$,

то $-1 < \gamma_0 < 0$; если $\sum_{i=1}^m \gamma_i < 1$, то $0 < \gamma_0 < 1$.

с помощью ЛПР подобрать m пар эквивалентных по предпочтительности точек $A^{(k)}, B^{(k)}$, $k = \overline{1, m}$ и записать систему уравнений вида:

$$f^{\theta c}(A^{(k)}) = f^{\theta c}(B^{(k)}), \quad k = \overline{1, m};$$

$$1 + \gamma_0 = \prod_{i=1}^m (1 + \gamma_0 \gamma_i).$$

Решением указанной системы являются шкалирующие коэффициенты γ_i , $i = \overline{0, m}$ мультипликативной оценочной функции.

Если условия взаимной независимости частных характеристик или аддитивной независимости на лотереях не выполняются, то оценочная функция $f^{\theta c}(Y)$ в мультипликативной или аддитивной форме не существует. В этом случае для построения оценочной функции обычно используют следующие приемы.

Во-первых, в пространстве характеристик $(y_1, y_2, \dots, y_m)$ исходов $g \in G$ можно выделить подобласти, в наибольшей степени интересующие ЛПР, а в них — репрезентативные точки. В выбранных точках Y_i проводится прямое оценивание значений функции $f^{\theta c}(Y_i)$. Используя процедуры интерполяции и экстраполяции, аппроксимируют функцию $f^{\theta c}(Y)$ по полученным точкам $f^{\theta c}(Y_i)$ и проверяют согласованность полученной оценочной функции. При этом в качестве начального приближения оценочной функции $f^{\theta c}(Y)$ целесообразно использовать аддитивную или мультипликативную форму.

Во-вторых, если из содержательного анализа удастся установить, что условия взаимной независимости выполняются для некоторых непересекающихся групп характеристик $Y^1, Y^2, \dots, Y^l$, то, построив условные оценочные функции для групп характеристик (например, прямым оцениванием), можно искать $f^{\theta c}(Y)$ в полилинейной форме на группах.

2. ЗАДАЧИ ПРИНЯТИЯ
РЕШЕНИЙ В УСЛОВИЯХ
НЕСТОХАСТИЧЕСКОЙ
НЕОПРЕДЕЛЕННОСТИ

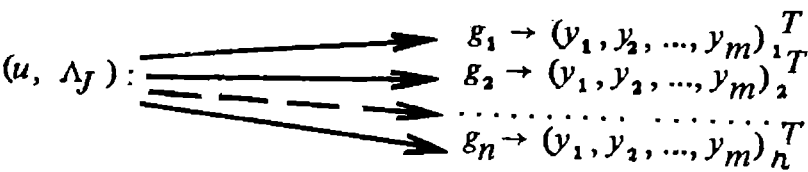
ОБЩАЯ ПОСТАНОВКА ЗАДАЧИ

Задача принятия решений в условиях нестохастической неопределенности возникает тогда, когда ЛПР, анализируя условия проведения операции, приходит к выводу, что некоторые факторы не могут быть описаны в рамках вероятностных моделей, так как либо отсутствует необходимая для этого информация, либо они вообще не являются случайными. В лучшем случае ЛПР располагает информацией лишь о диапазонах возможных значений этих факторов. Неопределенности такого рода рассматривают как «природные», т. е. вызванные недостаточной изученностью природы явлений. Кроме того, условия проведения операции могут формироваться в ре-

зультате деятельности других субъектов операции, преследующих собственные цели, которые не обязательно будут совпадать с целями ЛПР. Неопределенности такого рода, как правило, носят характер сознательного противодействия. Такую неопределенность рассматривают как «поведенческую».

На практике подобные ситуации возникают при проектировании, создании и производстве новых технических систем, при решении экономических, экологических и военных проблем, в социологии и других отраслях знаний.

При проведении операции в условиях нестохастической неопределенности («природной» Λ_J^N или «поведенческой» Λ_J^B) каждой стратегии $u \in U$ ставится в соответствие лишь множество возможных исходов $g \in G$ с характеристиками $(y_1, y_2, \dots, y_m) \in Y$:



Таким образом, для установления предпочтений на множестве стратегий модель цели операции (показатель и критерий эффективности) в качестве объективной информации должна учитывать лишь характеристики исходов. В этом случае в решающем правиле

$$u \geq v \triangleq W_e^\lambda(u) \geq W_e^\lambda(v)$$

функция эффективности $W_e^\lambda(u)$ в условиях нестохастической неопределенности должна выбираться специальным образом с учетом отношения ЛПР к риску нестохастической природы. Информацию об отношении ЛПР к нестохастическому риску в дальнейшем будем обозначать θ_H .

Тогда функция эффективности в условиях нестохастической неопреде-

ленности может быть представлена в виде

$$W_e^\lambda(u) = \varphi [\rho(Y(u), Y^{TP}, \theta_H)], \tag{13.80}$$

где $\varphi(\cdot)$ — функция агрегирования. Запись $\rho(Y(u), Y^{TP}, \theta_H)$ означает, что функция соответствия выбирается с учетом отношения ЛПР к нестохастическому риску и представляет в общем случае вектор различных функций соответствия, т. е.

$$\rho(Y(u), Y^{TP}, \theta_H) = (\rho_1(y_1(u), y_1^{TP}, \theta_H), \rho_2(y_2(u), y_2^{TP}, \theta_H), \dots, \rho_m(y_m(u), y_m^{TP}, \theta_H))^T. \tag{13.81}$$

Функция соответствия ρ_i может включать оператор математического ожи-

дания, что отражает то обстоятельство, что степень неопределенности во многих случаях может быть измерена с помощью субъективных вероятностей.

В задачах принятия решений в условиях нестохастической неопределенности при формировании конкретного вида функции ρ_i и ρ используются специальные принципы, учитывающие информацию θ_H .

К числу наиболее распространенных функций соответствия относятся следующие их виды¹:

$$\rho_i(u) = \min_{\lambda \in \Lambda_J^N} y_i(u, \lambda); \quad (13.82)$$

$$\rho_i(u) = \max_{\lambda \in \Lambda_J^N} \left(\max_{v \in U} y_i(v, \lambda) - y_i(u, \lambda) \right); \quad (13.83)$$

$$\rho_i(u) = \gamma \min_{\lambda \in \Lambda_J^N} y_i(u, \lambda) + (1 - \gamma) \max_{\lambda \in \Lambda_J^N} y_i(u, \lambda), \quad 0 \leq \gamma \leq 1; \quad (13.84)$$

$$\rho_i(u) = M[y_i(u, \lambda)]; \quad (13.85)$$

$$\rho_i(u) = \mu(y_i(u)). \quad (13.86)$$

Функция (13.82) используется как в условиях «природной», так и «поведенческой» неопределенности, в то время как функции (13.83)—(13.86) применимы лишь при отсутствии сознательного противодействия оперирующей стороне.

В дополнение к традиционным критериям пригодности и оптимальности в ситуациях принятия решения в условиях поведенческой неопределенности используют критерий адаптивности

$$u^*(t, \tau) : \sup_{u \in U(t, \tau)} \inf_{\lambda \in \Lambda_J} W_\varepsilon^\lambda(u, \lambda, t, \tau),$$

предполагающий формирование функции эффективности в зависимости от складывающейся ситуации на момент времени t и отыскание стратегии u^* , которая обеспечивает достижение верхней грани гарантированного ре-

зультата с учетом текущей и прогнозной (на момент времени $t + \tau$) информации об изменении условий.

ЗАДАЧИ ПРИНЯТИЯ РЕШЕНИЙ В УСЛОВИЯХ «ПРИРОДНОЙ» НЕОПРЕДЕЛЕННОСТИ

Для задач принятия решения этого типа основным является принцип наибольшего гарантированного результата. При этом в зависимости от информации θ_H гарантированный результат понимается по-разному.

Рассмотрим способы формирования гарантированного результата сначала для задач со скалярной характеристикой исхода.

Если ЛПР не склонно к нестохастическому риску, т. е. оно всегда предпочитает ориентироваться на самые неблагоприятные значения фактора λ , то в качестве показателя эффективности следует использовать функцию соответствия (13.82):

$$W(u) = \min_{\lambda \in \Lambda_J^N} y(u, \lambda).$$

В этом случае решение задачи отыскивается по максиминному критерию

$$u^* : \max_{u \in U} \min_{\lambda \in \Lambda_J^N} y(u, \lambda). \quad (13.87)$$

Максиминное правило ориентировано на «наихудшие» значения неопределенного фактора и в этом смысле является консервативным, т. е. его следует применять в том случае, когда «неуспех» операции крайне нежелателен независимо от того, каковы могут быть «наилучшие» исходы.

Пример 11. Значения возможных характеристик y исхода для двух стратегий ЛПР и двух состояний природы заданы таблицей исходов:

$$\begin{array}{cc} & \lambda_1 & \lambda_2 \\ \begin{array}{c} u_1 \\ u_2 \end{array} & \parallel \begin{array}{cc} a & 1000 \\ 1 & 1 \end{array} \parallel & \text{, причем величина} \end{array} \quad 0 < a < 1.$$

Тогда $W(u_1) = a$, $W(u_2) = 1$. Следовательно, наилучшей стратегией по критерию (87) является u_2 независимо от того, что разность $1 - a$ может быть сколь угодно малой, хотя стра-

¹ Большие значения y предпочтительнее меньших.

тегия u_1 может привести к исходу $y(u_1, \lambda_2) = 1000$.

Если информация θ_n указывает на то, что для ЛПР не безразлична величина возможного выигрыша (т. е. он «боится» мало выиграть), то в этом случае в качестве показателя следует использовать функцию (13.83). По смыслу эта функция задает максимальные потери, связанные с конкретным выбранным решением. Поэтому критерий выбора стратегии основывается на минимизации этих потерь:

$$u^* : \min_{u \in U} \max_{\lambda \in \Lambda_J^N} \left[\max_{v \in U} y(v, \lambda) - y(u, \lambda) \right]. \quad (13.88)$$

Это правило носит название *критерия минимаксного сожаления* (критерий Сэвиджа).

Пример 12. В условиях примера 11 матрица потерь («сожалений»), вычисляемых как $\max_{v \in U} y(v, \lambda) - y(u, \lambda)$,

имеет вид

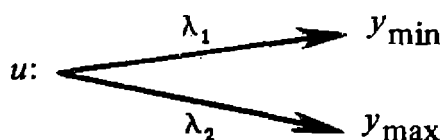
$$\begin{array}{cc} & \lambda_1 & \lambda_2 \\ \begin{array}{c} u_1 \\ u_2 \end{array} & \left\| \begin{array}{cc} 1-a & 0 \\ 0 & 999 \end{array} \right\| \end{array}$$

ЛПР не испытывает «сожалений» лишь в ситуациях (u_1, λ_2) и (u_2, λ_1) . А «сожаления» в других ситуациях для каждой из стратегий существенно различны. Следовательно, при выборе стратегии по критерию Сэвиджа ЛПР должно ориентироваться на наименьшее из максиминных по λ сожалений, т. е.

$$W(u_1) = 1 - a, \quad W(u_2) = 999 \quad \text{и} \quad u^* = u_1.$$

Критерий Сэвиджа чувствителен к исходному множеству стратегий U .

Пример 13. Пусть множество стратегий состоит из трех альтернатив, причем u_1 и u_2 соответствуют примеру 12, а u_3 приводит к ситуациям: $(u_3, \lambda_1) = 1000$, $(u_3, \lambda_2) = b$; причем $0 < b < a$.



Тогда матрица «сожалений» имеет вид

$$\begin{array}{cc} & \lambda_1 & \lambda_2 \\ \begin{array}{c} u_1 \\ u_2 \\ u_3 \end{array} & \left\| \begin{array}{cc} 1000-a & 0 \\ 999 & 999 \\ 0 & 1000-b \end{array} \right\| \end{array}$$

и критерий (13.88) приводит к $u^* = u_2$. Однако, если по какой-либо причине стратегия u_3 не может быть реализована, т. е. $U^{(1)} = U \setminus u_3$, то стратегия u_2 перестает быть оптимальной (см. предыдущий пример).

Таким образом, критерий Сэвиджа не обладает свойством независимости от дополнительных альтернатив.

Максиминный критерий и критерий Сэвиджа являются слишком категоричными в том смысле, что один ориентируется только на наихудший результат, а другой на максимальные потери. Поэтому, если информация θ_n указывает на то, что ЛПР «боится» не только мало выиграть, но и боится много проиграть, то его отношение к нестохастическому риску можно охарактеризовать как пессимистически-оптимистическое. В этом случае обычно используют функцию (13.84), в которой коэффициент γ носит название *коэффициента пессимизма—оптимизма Гурвица*.

Если $\gamma = 1$, то отношение ЛПР к риску характеризуется крайним пессимизмом и критерий Гурвица вырождается в максиминное правило (13.87). Если $\gamma = 0$, то ЛПР крайне оптимистично в своих предположениях о значениях λ , ориентируется только на наилучшие результаты, что приводит к максимаксному критерию

$$u^* : \max_{u \in U} \max_{\lambda \in \Lambda_J^N} y(u, \lambda).$$

При выборе коэффициента γ обычно используют эвристические методы, основанные на учете информации о приемлемости для ЛПР наихудших и наилучших состояний фактора λ . В качестве первого приближения может быть взята оценка коэффициента γ , полученная из условия:

$$v: \begin{array}{c} \bullet \longrightarrow \end{array} \lambda_1 \cup \lambda_2 \quad y_\gamma, \quad (13.89)$$

тогда

$$\gamma^* = 1 - \frac{y_{\gamma} - y_{\min}}{y_{\max} - y_{\min}},$$

где

$$y_{\min} = \min_{u \in U} \min_{\lambda \in \Lambda_J^N} y(u, \lambda);$$

$$y_{\max} = \max_{u \in U} \max_{\lambda \in \Lambda_J^N} y(u, \lambda).$$

Критерий Гурвица ставит в соответствие каждой стратегии линейную комбинацию только наихудшего и только наилучшего значений характеристики исхода. Это приводит к тому, что явно различающиеся даже по здравому смыслу стратегии могут иметь одинаковую оценку по критерию Гурвица.

Пример 14. Возможные исходы операции для двух стратегий и четырех возможных состояний λ характеризуются матрицей значений $y(u, \lambda)$:

$$\begin{array}{c|cccc} & \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \\ \hline u_1 & 0 & 1 & 2 & 1000 \\ u_2 & 1000 & 999 & 900 & 0 \end{array} \parallel;$$

$$\min_{\lambda \in \Lambda_J^N} y(u_i, \lambda) = 0, \quad i = 1, 2;$$

$$\max_{\lambda \in \Lambda_J^N} y(u_i, \lambda) = 1000, \quad i = 1, 2.$$

Следовательно, при любом γ стратегии u_1 и u_2 эквивалентны в смысле критерия Гурвица. Хотя здравый смысл подсказывает, что стратегия u_2 предпочтительнее u_1 .

Для устранения отмеченного недостатка можно рекомендовать два приема модификации критерия Гурвица. Во-первых, каждой стратегии можно поставить в соответствие свой коэффициент пессимизма-оптимизма, т. е. $\gamma_k = \gamma(u_k)$. При этом γ_k определяется из условия (13.89) с подстановкой $y_{\min} = \min_{\lambda \in \Lambda_J^N} y(u_k, \lambda)$ и $y_{\max} = \max_{\lambda \in \Lambda_J^N} y(u_k, \lambda)$. Во-вторых, рассмотреть задачу как векторную и воспользоваться, например, методом главного показателя. В этом случае в качестве главного показателя используется (13.84), а в качестве дополнительного—

рандомизированный средний выигрыш, на который накладывается ограничение, т. е.

$$u^*: \max_{u \in U^{(1)}} \left[\gamma \min_{\lambda \in \Lambda_J^N} y(u, \lambda) + (1 - \gamma) \max_{\lambda \in \Lambda_J^N} y(u, \lambda) \right], \quad (13.90)$$

где

$$U^{(1)} = \left\{ u \left| \frac{1}{n} \sum_{i=1}^n y(u, \lambda_i) \leq b \right. \right\}. \quad (13.91)$$

В (13.91) в качестве порогового значения b целесообразно выбирать величину y из условия (13.89).

Сужение исходного множества U до более узкого подмножества $U^{(1)}$ можно осуществлять, если имеется качественная информация о степени возможности состояний $\lambda \in \Lambda_J^N$. Эта информация задается в форме элементарных суждений типа: «состояние λ_i более возможно, чем λ_j », «состояния λ_i и λ_j возможны в одинаковой степени», «группа состояний Λ_1 более возможна по сравнению с группой состояний Λ_2 » и т. п. В этом случае формально каждой стратегии $u \in U$ ставится в соответствие векторная оценка $y(u) = (y(u, \lambda_1), y(u, \lambda_2), \dots, y(u, \lambda_n))^T$, компоненты которой различаются по важности (в смысле большей или меньшей возможности реализации состояний λ_i). Такой прием позволяет использовать технику выделения омега-эффективных стратегий, которые образуют подмножество $U^{(1)}$.

Если ЛПР считает, что ни одно из состояний $\lambda \in \Lambda_J^N$ не следует рассматривать как более возможное по сравнению с любым другим, то омега-информация есть информация о равноценности (равновозможности) состояний λ . В этом случае можно ввести порядковые коэффициенты «важности» состояний λ и, учитывая нормирующее условие $\sum_{i=1}^n \alpha_i = 1$, положить их равными $\alpha_i = \frac{1}{n}$. Тогда

правило выбора наилучшей стратегии формулируется следующим образом:

$$u^*: \max_{u \in U} \left\{ \sum_{i=1}^n \alpha_i y(u, \lambda_i) = \right. \\ \left. = \frac{1}{n} \sum_{i=1}^n y(u, \lambda_i) \right\} \quad (13.92)$$

Правило (13.92) может быть интерпретировано как критерий максимизации функций соответствия (13.85) при условии, что априорные вероятности $p^*(\lambda_i) = \frac{1}{n}$. В такой интерпретации правило (13.92) соответствует известному принципу «недостаточного основания», впервые сформулированному Бернулли: если нет никаких оснований считать одно состояние природы $\lambda \in \Lambda_J^N$ более вероятным, чем любое другое, то априорные вероятности $p(\lambda_i)$ следует принять одинаковыми.

В том случае, когда ЛПР, выбирая решения, ориентируется не только на величины максиминных сожалений (13.83), но и учитывает промежуточные их значения при фиксированном состоянии $\lambda \in \Lambda_J^N$, равные $\max_{v \in U} y(v, \lambda_i) -$

$- y(u, \lambda_i)$, то оно может приписывать различные веса состояниям λ_i , пропорциональные суммарному сожалению $\sum_{u \in U} \left[\max_{v \in U} y(v, \lambda_i) - y(u, \lambda_i) \right]$.

Нормированные суммарные сожаления $p^*(\lambda_i)$ состояний λ_i рассматриваются как «потенциальное распределение вероятностей» на состояниях λ_i [98]:

$$p^*(\lambda_i) = \frac{\sum_{u \in U} \left[\max_{v \in U} y(v, \lambda_i) - y(u, \lambda_i) \right]}{\sum_{i=1}^n \sum_{u \in U} \left[\max_{v \in U} y(v, \lambda_i) - y(u, \lambda_i) \right]}, \\ i = \overline{1, n}. \quad (13.93)$$

Тогда правило выбора наилучшей стратегии формулируется следующим образом:

$$u^*: \max_{u \in U} \sum_{i=1}^n p^*(\lambda_i) y(u, \lambda_i). \quad (13.94)$$

Правило выбора (13.94) совместно с (13.93) предложено В. Хоменюком. Рассмотрим следующий иллюстративный пример.

Пример 15. Возможные исходы операции для трех стратегий и четырех возможных состояний λ характеризуются матрицей значений $y(u, \lambda)$:

$$\begin{array}{c|cccc} & \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \\ \hline u_1 & 5 & 12 & 6 & 1 \\ u_2 & 10 & 2 & 2 & 10 \\ u_3 & 15 & 9 & 0 & 0 \end{array}$$

При использовании критерия Сэвиджа (13.88) стратегии u_1, u_2, u_3 эквивалентны, так как

$$W(u_1) = W(u_2) = W(u_3) = \\ = \max_{\lambda \in \Lambda_J^N} \left(\max_{v \in U} y(v, \lambda_i) - y(u, \lambda_i) \right) = 10.$$

К аналогичному результату приводит использование и критерия Бернулли (13.92):

$$W(u_1) = W(u_2) = W(u_3) = \\ = \frac{1}{4} \sum_{i=1}^4 y(u, \lambda_i) = 6.$$

Для критерия Гурвица построим функцию соответствия (13.84) с постоянным коэффициентом γ для всех стратегий. Оценка γ^* коэффициента γ по условию (13.89): ЛПР указало, что $y_\gamma = 4,5$.

Тогда

$$\gamma^* = 1 - \frac{4,5 - 0}{15 - 0} = 0,7$$

и

$$W(u_1) = 0,7 \times 1 + (1 - 0,7) 12 = 4,3;$$

$$W(u_2) = 0,7 \times 2 + (1 - 0,7) 10 = 4,4;$$

$$W(u_3) = 0,7 \times 0 + (1 - 0,7) 15 = 4,5.$$

Таким образом $u^* = u_3$ и $u_3 \succ u_2 \succ u_1$.

Пусть теперь ЛПР указало y_γ для каждой стратегии:

$$y_\gamma(u_1) = 5, y_\gamma(u_2) = 6, y_\gamma(u_3) = 4,5.$$

Тогда

$$\gamma_1 = 1 - \frac{5-1}{12-1} = 0,63;$$

$$\gamma_2 = 1 - \frac{6-2}{10-2} = 0,5;$$

$$\gamma_3 = 1 - \frac{4,5-0}{15-0} = 0,7$$

и

$$W(u_1) = 0,63 \times 1 + (1 - 0,63) 12 = 5,07,$$

$$W(u_2) = 0,5 \times 2 + (1 - 0,5) 10 = 6,0.$$

$$W(u_3) = 0,7 \times 0 + (1 - 0,7) 15 = 4,5,$$

т. е. $u^* = u_2$, но $u_2 \succ u_1 \succ u_3$.

При использовании критерия Хомеюка потенциальное распределение вероятностей, подсчитанное по (13.93), имеет вид

$$p^*(\lambda_1) = 0,26, p^*(\lambda_2) = 0,23,$$

$$p^*(\lambda_3) = 0,18, p^*(\lambda_4) = 0,33.$$

Тогда $W(u_1) = 5,44$, $W(u_2) = 6,77$,

$W(u_3) = 4,17$, т. е. $u^* = u_2$ и $u_2 \succ u_1 \succ u_3$.

Таким образом критерий Гурвица и его модификация и критерий Хомеюка указывают на то, что $u_2 \succ u_1$. В то же время критерий Гурвица в качестве оптимальной выделяет стратегию u_3 , а два последних — стратегию u_2 . Минимаксный критерий в качестве оптимальной также выделяет стратегию u_2 .

Применение функции соответствия (13.85) возможно и тогда, когда в качестве априорного распределения используются субъективные вероятности, заданные на множестве Λ_J^N .

Если информация θ может быть представлена в форме нечетких суж-

дений, то при отыскании наилучших стратегий используется функция соответствия типа (13.86).

В этом случае на универсальном множестве Λ_J^N с помощью нечеткого отношения принадлежности задается нечеткое подмножество A . Нечеткое отношение описывается функцией $\mu_A(\lambda)$, значения которой характеризуют степень принадлежности состояния $\lambda \in \Lambda_J^N$ событию A . В этом случае правило выбора наилучшей стратегии может быть сформулировано в следующем виде:

$$u^*: \max_{u \in U} W_1(\mu, u), \quad (13.95)$$

где $W_1(\mu, u) =$

$$= \frac{1}{\sum_{i=1}^n \mu_A(\lambda_i)} \sum_{i=1}^n \mu_A(\lambda_i) y(u, \lambda_i). \quad (13.96)$$

При решении задачи (13.95) событие A должно выбираться таким образом, чтобы его наступление предопределяло интересующий ЛПР исход операции.

Формально величины $\frac{\mu_A(\lambda_i)}{\sum_{i=1}^n \mu_A(\lambda_i)}$,

входящие в показатель эффективности (13.96), рассматриваются как субъективные вероятности состояний λ_i (или как субъективные вероятности компонент функции принадлежности). Поэтому сам показатель эффективности по смыслу является математическим ожиданием функции соответствия (13.86).

В качестве дополнительного показателя при решении задачи (13.95), (13.96) может быть использован аналог дисперсии результата $y(u, \lambda)$:

$$W_2(\mu, u) = D_y(\mu, u) = \frac{1}{\sum_{i=1}^n \mu_A(\lambda_i)} \sum_{i=1}^n \mu_A(\lambda_i) \times [y(u, \lambda_i) - W_1(\mu, u)]^2. \quad (13.97)$$

Тогда решение может быть получено по скалярному показателю с использованием сверток типа (13.61)—(13.63).

Задача нахождения «наилучшей» стратегии u^* в условиях нестохастической неопределенности может быть решена с использованием аналога показателя вероятностной гарантии:

$$\begin{aligned} W(\mu, u) &= P(y(u, \lambda) \geq y^{TP}) = \\ &= P(\lambda \in \Lambda_J^N(y^{TP}, u)) = \\ &= \sum_{\lambda \in \Lambda_J^N(y^{TP}, u)} \frac{\mu_A(\lambda)}{\sum_{i=1}^n \mu_A(\lambda_i)}, \end{aligned} \quad (13.98)$$

где

$$y^{TP} \in \left[\min_u \min_{\lambda} y(u, \lambda), \max_u \max_{\lambda} y(u, \lambda) \right];$$

$$\Lambda_J^N(y^{TP}, u) = \{\lambda_i \mid y(u, \lambda_i) \geq y^{TP}\}.$$

Пример 16. В условиях примера 9 руководством предприятия принято решение о запуске в серийное производство новой модели. Необходимо принять решение об объеме серии. Прогноз показывает, что спрос на новую модель составит примерно 3000 комплектов при условии, что розничная цена C будет не выше 750 р. На основе проведенного математического моделирования получена регрессионная модель полных затрат, связанных с производством и продажей x комплектов изделий:

$$z(x) = 250x \left(1 + e^{-\frac{(x-1000)}{2000}} \right), \quad 1000 < \lambda < 10\,000.$$

На множестве Λ_J^N возможных значений спроса введено нечеткое событие A —

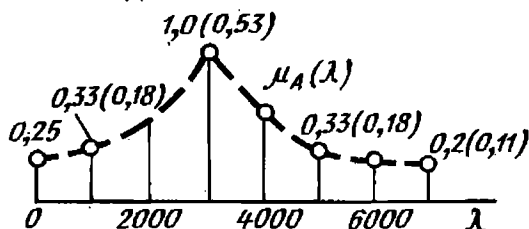


Рис. 15. Функция принадлежности уровня спроса

спрос примерно 3000 комплектов — с функцией принадлежности вида

$$\mu_A(\lambda) = \frac{1000}{1000 + (\lambda - 3000)},$$

$$1000 \leq \lambda \leq 10\,000.$$

В качестве альтернативных стратегий рассматриваются объемы выпусков x : $u_1 = 1000$, $u_2 = 3000$, $u_3 = 5000$, $u_4 = 7000$ комплектов.

В качестве критерия эффективности используется правило (13.95), (13.96) с характеристикой исхода:

$$y(u, \lambda) = C \min \{x(u), \lambda\} - z(x(u)),$$

который характеризует чистую прибыль предприятия. На рис. 15 представлен график функции принадлежности $\mu_A(\lambda)$ уровня спроса и выде-

лены значения $\frac{\mu_A(\lambda_i)}{\sum_{i=1}^n \mu_A(\lambda_i)}$ для λ_i ,

соответствующих рассматриваемым стратегиям (указаны в скобках). В табл. 5 приведены результаты расчетов. Анализ этой таблицы показывает, что «наилучшей» стратегией является стратегия u_3 .

Иногда не удастся интерпретировать значения функции принадлежности как элементы субъективного распределения вероятностей на состояниях $\lambda_i \in \Lambda_J^N$. В этом случае может быть применен подход, основанный на построении функции принадлежности отображения $H: U \times \Lambda \rightarrow Y(G)$ с последующим ее содержательным анализом для принятия решений. В этом случае отображение H (при условии, что λ — нечетко заданный фактор) есть отображение нечетких множеств.

Нечеткое отображение формально описывается как отображение, при котором элементу $\lambda \in \Lambda_J^N$ ставится в соответствие не конкретный элемент множества Y , а его нечеткое подмножество. Если при фиксированной стратегии $u \in U$ отображение H можно представить как $y = H(\lambda)$, то функ-

5. Результаты расчетов

Стратегия	Объем выпу- ска $\times 10^{-3}$	Полные за- траты $\times 10^{-5}$	Выручка от продаж 10^{-6}	Чистая прибыль $y(u, \lambda) \times 10^{-5}$ предприятия				Значение показателя $W(u) \times 10^{-5}$ эффектив- ности
				$\lambda_1 =$ $= 1000$	$\lambda_2 =$ $= 3000$	$\lambda_3 =$ $= 5000$	$\lambda_4 =$ $= 7000$	
u_1	1	5	7,5	2,5	2,5	2,5	2,5	2,5
u_2	3	10,2	22,5	−2,7	12,3	12,3	12,3	9,6
u_3	5	14,2	37,5	−6,7	8,3	23,3	23,3	9,95
u_4	7	18,3	52,5	−10,8	4,2	19,2	34,2	7,5

ция принадлежности $\mu(y)$ определяется по правилу:

$$\mu(y) = \max_{\lambda \in H^{-1}(y)} \mu_A(\lambda), \quad \forall y \in Y,$$

(13.99)

где

$$H^{-1}(y) = \{\lambda \mid H(\lambda) = y\}.$$

(13.100)

В случае, если y является функцией от нескольких нечетких переменных (например, $x_1, x_2, \dots, x_n$), то

$$\mu(y) = \max_{x \in X} \min \{\mu_1(x_1), \mu_2(x_2), \dots, \mu_m(x_m)\},$$

(13.101)

где

$$x = (x_1, x_2, \dots, x_m)^T;$$

(13.102)

$$X = \{x \mid H(x) = y\}.$$

(13.103)

Однако нахождение функции принадлежности $\mu(y)$ по правилу (13.101) трудоемко, поэтому обычно декомпозируют задачу (13.101) на ряд подзадач, использующих лишь бинарные и унарные (13.99) операции. Например, если $y = x_1^{x_2} + x_3$, то, обозначив $x_1^{x_2} = y_1, x_3 = y_3$, получим $y = y_1 + y_2$.

Тогда, определив $\mu_1(y_1)$ по правилу (13.99), получим

$$\mu(y) = \max_{y_1} \min \{\mu_1(y_1), \mu_2(y - y_1)\},$$

$$\forall y \in Y,$$

$$\text{где } \mu_2(\cdot) = \mu(x_3).$$

Функция принадлежности преобразования $y = kx$ совпадает с функцией принадлежности переменной x в масштабе, увеличенном в k раз, а $\mu(y)$ при $y = x + a$ эквивалентна исходной функции $\mu(x)$, сдвинутой на величину a .

Выбор наилучшей стратегии осуществляется на основе анализа $\mu(y(u))$. В настоящее время не существует общепринятого подхода к осуществлению такого выбора. Если, например, большие значения исхода $y(u)$ предпочтительнее меньших, то можно говорить о доминировании одних стратегий над другими в том случае, если носители¹ $\mu(y(u))$ не перекрываются:

$$u \succ v \iff \min \{\text{Supp } y(u)\} > \min \{\text{Supp } y(v)\};$$

(13.104)

$$\text{Supp } y(u) \cap \text{Supp } y(v) = \emptyset.$$

Для несравнимых стратегий, выделенных по правилу (13.104), необходимо провести дополнительный содержательный анализ, включающий идентификацию типа отношения принадлежности («не очень велико», «не больше ...», «не меньше ...», «не очень мало», «близко к ...» и т. д.) и информацию о цели операции, которая может быть и нечеткой. В этом случае рассматриваются два в общем случае нечетких события A и B , описываемые

¹ Носителем нечеткого множества A с функцией принадлежности $\mu_A(x)$ называется обычное множество $\text{Supp } A$ вида: $\text{Supp } A = \{x \mid \mu_A(x) > 0, x \in X\}$.

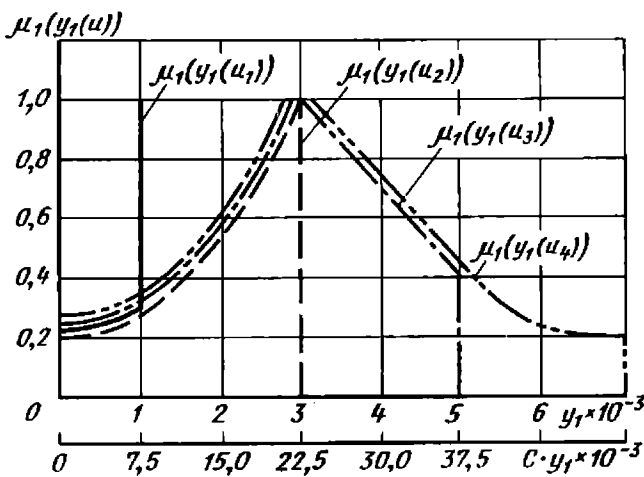


Рис. 16. Функции принадлежности преобразования $y_1(x(u), \lambda)$

функциями принадлежности $\mu_A(y)$ и $\mu_B(y)$, причем функция $\mu_B(y)$ задает отношение принадлежности результата y к y^{TP} . Тогда пересечение D нечетких множеств A и B :

$$\mu_D(y) = \mu_{A \cap B}(y) = \min\{\mu_A(y), \mu_B(y)\}$$

содержательно описывает принадлежность получаемых результатов требуемым, и, следовательно, правило выбора лучшей стратегии может иметь вид

$$u^*: \max_{u \in U} \mu_D(y(u)). \quad (13.105)$$

Для иллюстрации изложенного рассмотрим пример.

Пример 17. В условиях примера 16 цель операции сформулирована в виде нечеткого суждения «чистая прибыль предприятия должна быть не ниже примерно $20 \cdot 10^5$ ».

Для решения данной задачи необходимо вначале построить функции принадлежности преобразования $y(u) = C \min\{x(u), \lambda\} - z(x(u))$ для каждой стратегии $u \in U$. Преобразование $y(u)$ представляет собой произведение четкой константы C на нечеткую функцию $\min\{x(u), \lambda\}$ и сдвиг на величину $z(x(u))$. Поэтому вначале отыскиваются функции принадлежности $\mu_1(y_1(u))$ преобразования $y_1(u) = \min\{x(u), \lambda\}$ по правилу:

$$\begin{aligned} \mu_1(y_1(u)) &= \\ &= \max_{x(u), \lambda \mid \min\{x(u), \lambda\} = y_1} \times \\ &\times \min\{1, \mu(\lambda)\}, \end{aligned}$$

представленные на рис. 16, затем изменяется масштаб и получаются значения функции принадлежности преобразования $y_2(u) = Cy_1$. Результирующие графики функций принадлежности $\mu(y(u))$ преобразования $y(u)$ после сдвига $y_2(u)$ на величину $z(x(u))$ представлены на рис. 17. На этом же

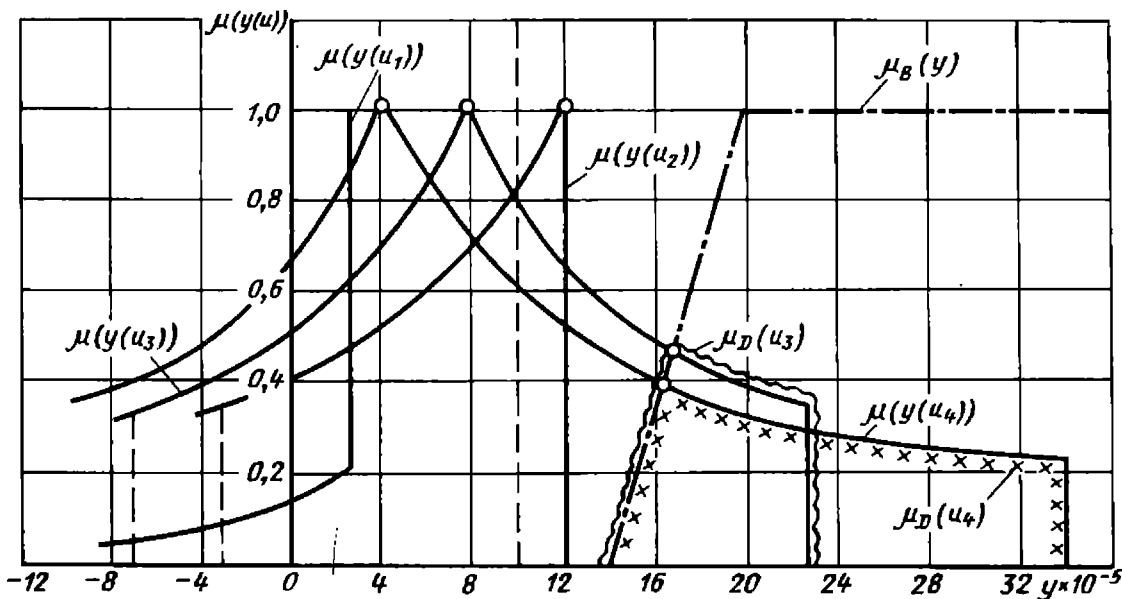


Рис. 17. Функции принадлежности преобразования $y(u)$ и нечетких событий B и D

рисунке штрихпунктиром изображена функция принадлежности нечеткого события B — «чистая прибыль предприятия не ниже примерно $20 \cdot 10^5$ », волнистой линией и знаками $\times$ выделены функции принадлежности $\mu_D(u_3)$ и $\mu_D(u_4)$ соответственно. Сравнение максимальных значений функций $\mu_D(\cdot)$, выделенных кружками, показывает, что $u^* = u_3$. Если бы требуемый результат y^{TP} был другим (например, четким и равным 10^6), то в качестве наилучшей стратегии u^* могла оказаться стратегия u_2 .

Если бы цель операции была задана в форме «получить неотрицательную прибыль», то предлагаемое правило приводило бы к неопределенности (все стратегии имеют одинаковые максимальные значения функции $\mu_D(u)$), поэтому здесь целесообразно выбирать u^* , исходя из минимума значения функции принадлежности результата к нулевому, что приводит к $u^* = u_1$. Если бы цель операции формулировалась в виде «получить максимально возможный результат», то в качестве наилучшей стратегии следовало бы выбирать альтернативу, дающую наибольший результат y с ненулевым значением функции принадлежности $\mu(y(u))$:

$$u^*: \mu(y^*) \neq 0, y^* = \max_{u \in U} y(u).$$

В примере 17 коэффициент C , характеризующий продажную цену комплекта, мог быть задан и нечетко. Например, «цена комплекта будет около 750 р». Кроме того, стратегия $u \in U$, задающая уровень выпуска, также могла быть нечеткой из-за наличия нечетких ограничений на возможности предприятия по отдельным видам ресурсов. Если бы все параметры задачи были четкими, то такую задачу выбора наилучшей стратегии можно было бы сформулировать как задачу математического программирования [см. задачи (12.13), (12.14)], в противном случае — как задачу нечеткого математического программирования.

Заранее не следует вводить четкие ограничения (четкое множество допустимых стратегий U), так как может

оказаться, что стратегии, не удовлетворяющие четким ограничениям, будут более предпочтительными для ЛПР, как дающие больший целевой эффект. Поэтому нечеткое описание более адекватно, так как позволяет ЛПР оперировать решениями, лежащими на грани дефицита ресурсов. Для того чтобы учесть указанную особенность, используют соответствующую технику решения нечетко поставленных задач.

В зависимости от формы нечеткого описания исходной информации существуют различные математические формулировки и методы решения задач нечеткого математического программирования (НМП).

Некоторые постановки задач нечеткого математического программирования. Если целевая функция $\varphi(x)$ — обычная четкая, а множество допустимых альтернатив характеризуется функцией принадлежности $\mu_U(x)$, причем $x \in \text{Re}^m$, то для решения задачи

$$u^*: x^*: \max_{x \in U} \varphi(x)$$

используют следующий прием:

решается задача четкого математического программирования

$$\varphi^* = \max_{x \in \text{Supp } \mu_U(x)} \varphi(x);$$

целевая функция нормируется полученным максимальным значением

$$\bar{\varphi}(x) = \frac{\varphi(x)}{\varphi^*};$$

выбираются такие характеристики x стратегии u , что

$$u^*: x^*: \max_{x \in X(U)} \min \{\mu_U(x), \bar{\varphi}(x)\}. \quad (13.106)$$

Если цель операции состоит в достижении некоторого требуемого результата y^{TP} и ЛПР в состоянии сопоставить степень нарушения ограничений по дефицитным ресурсам с получаемым дополнительным целевым эффектом сверх y^{TP} или указать предельно возможное отступление от y^{TP} (в сторону его уменьшения) для рационального использования сверхли-

митного ресурса, то задача НМП может формулироваться в виде

$$u^*: x^*: \max_{x \in X(U)} \min \{ \mu_\varphi(x), \mu_U(x) \}, \quad (13.107)$$

где нечеткие множества цели μ_φ и ограничений μ_U задаются следующими выражениями:

$$\mu_\varphi(x) = \begin{cases} 0, & \varphi(x) < y^{TP} - \Delta y; \\ \mu_y(x, \Delta y), & y^{TP} - \Delta y < \varphi(x) < y^{TP}; \\ 1, & \varphi(x) \geq y^{TP}; \end{cases}$$

$$\mu_U(x) = \begin{cases} 0, & q(x) \geq \Delta q; \\ \mu_x(x, \Delta q), & 0 < q(x) < \Delta q; \\ 1, & q(x) \leq 0, \end{cases}$$

где $\mu_y(x, \Delta y)$, $\mu_x(x, \Delta q)$ — функции принадлежности, описывающие степени выполнения соответствующих целевых и ресурсных ограничений с точки зрения ЛПР.

В практике часто встречаются задачи, которые сводятся к задачам максимизации целевой функции на области допустимых решений, причем параметры как целевой функции $\varphi(x)$, так и ограничений $q(x)$ описываются нечетко. Для решения таких задач разработаны специальные методы и алгоритмы [50]. Однако в большинстве они оказываются неконструктивными, так как для решения даже весьма простых задач требуется чрезвычайно большой объем вычислений. Более того, для получаемого множества возможных решений требуется дополнительный трудоемкий анализ с привлечением ЛПР.

Формально задача НМП может быть представлена в виде

$$u^*: x^*: \max_{x \in X(U)} \varphi(x, a); \quad (13.108)$$

$$X(U) = \left\{ x \mid \begin{matrix} q(x, b) \leq 0 \\ h(x, c) = 0 \end{matrix} \right\}, \quad x \in \text{Re}^n,$$

где $a = (a_1, a_2, \dots, a_n)^T$ — вектор нечетких параметров целевой функции с функциями принадлежности $\mu_\varphi(a_i)$; $b = \|b_{ij}\|$, $i = \overline{1, n}$, $j = \overline{1, l_0}$ — матрица нечетких коэффици-

ентов ограничений — неравенств с функциями принадлежности $\mu_q(b_{ij})$; $c = \|c_{ik}\|$, $k = \overline{1, k_0}$ — матрица нечетких коэффициентов ограничений — равенств с функциями принадлежности $\mu_h(c_{ik})$. С практической точки зрения, при решении задачи (13.108) можно руководствоваться лишь пессимистической и оптимистической оценками возможного решения, являющихся основой для выбора конкретных альтернатив в исходной задаче НМП. Обычно в качестве оптимистической оценки рассматривается решение α -уровня следующей задачи НМП:

$$u^*: x^*: \max_{x, a} \varphi(x, a) \quad (13.109)$$

при ограничениях:

$$\left. \begin{aligned} q(x, b) &\leq 0, \\ h(x, c) &= 0, \\ \mu_\varphi(a_i) &\geq \alpha; \\ \mu_q(b_{ij}) &\geq \alpha; \\ \mu_h(c_{ik}) &\geq \alpha, \quad i = \overline{1, n}, \\ j &= \overline{1, l_0}, \quad k = \overline{1, k_0}. \end{aligned} \right\} \quad (13.110)$$

Решение x^* задачи (13.109), (13.110) можно рассматривать как альтернативу, степень недоминируемости которой не меньше α .

В качестве наиболее осторожной (пессимистической) оценки для задачи (13.108) рассматривается решение α -уровня следующей задачи НМП:

$$u^*: x^*: \max_x \min_a \varphi(x, a) \quad (13.111)$$

при ограничениях (13.110).

Последовательно задавая α -уровни в задачах (13.109) — (13.111), получают интервалы возможных значений целевой функции $\varphi(\cdot)$, анализ которых позволяет осуществить выбор решения и установить степень его недоминируемости.

Если задача принятия решения в условиях природной неопределенности связана с анализом альтернатив, описываемых векторной характеристикой исхода $y(u, \lambda) = (y_1(u, \lambda), y_2(u, \lambda), \dots, y_m(u, \lambda))^T$, то предварительно сужают исходное множество страте-

гий, выделяя их недоминируемое подмножество по правилу:

$$u \geq v \iff W_i(u) \geq W_i(v), \quad i = \overline{1, m}, \quad (13.112)$$

где $W_i(u)$ определяются как функции типа (13.82)—(13.86), конкретный выбор которых зависит от содержательного смысла информации θ_H (в случае использования функции соответствия (13.83) знак неравенства в (13.112) меняется на противоположный).

Выделение наилучшей стратегии из полученного эффективного множества альтернатив, удовлетворяющих (13.112), осуществляется по критерию оптимальности, например, совместно с тем или иным способом агрегирования векторной характеристики исхода в скалярную (см. гл. 12):

$$\Phi(u, \lambda) = \Phi[y_1(u, \lambda), y_2(u, \lambda), \dots, y_m(u, \lambda)].$$

Далее используются функции (13.82)—(13.86) применительно к «характеристике» исхода $\Phi(u, \lambda)$. Например, согласно принципу наибольшего гарантированного результата оптимальная стратегия u^* будет выделена в результате решения задачи:

$$u^*: \max_{u \in U} \min_{\lambda \in \Lambda_J^N} \Phi(u, \lambda).$$

Иногда из содержательной постановки задачи следует, что целесообразнее сразу свести исходную задачу к задаче принятия решения в условиях определенности по векторному показателю, т. е. заменить $y(u, \lambda) = (y_1(u, \lambda), y_2(u, \lambda), \dots, y_m(u, \lambda))^T$ на вектор $W(u) = (W_1(u), W_2(u), \dots, W_m \times(u))^T$ с компонентами $W_i(u)$, определяемыми по правилам (13.82)—(13.86). При этом вид функции соответствия един для всех компонент $y_i(u, \lambda)$. Тогда выделение наилучшей стратегии u^* можно проводить различными методами решения векторных задач в условиях определенности (см. гл. 12) в зависимости от имеющейся информации об относительной важности частных показателей $W_i \times(u)$, $i = \overline{1, m}$: независимость по предпочтению, качественная информация типа «один показатель важнее

другого» и ее особые случаи (L, S, SL — информация) и др.; при этом могут применяться как одно-, так и многошаговые методы. Например, если частные показатели определяются как гарантированный результат, т. е.

$$W_i(u) = \min_{\lambda \in \Lambda_J^N} y_i(u, \lambda),$$

а от ЛПР получена информация о том, что гарантированные уровни частных характеристик для него являются одинаково важными (равноценными) показателями, то задача может решаться как симметрически лексикографическая:

$$u^*: SL \max_{u \in U} W(u),$$

которая сводится к задаче

$$u^*: \max_{u \in U} \Phi(W(u)),$$

где

$$\Phi(W(u)) = (\Phi_1(W(u))),$$

$$\Phi_2(W(u)), \dots, \Phi_m(W(u))^T;$$

$$\Phi_1(W(u)) = \min_i W_i(u), \dots,$$

$$\Phi_m(W(u)) = \max_i W_i(u).$$

При простом упорядочении показателей по важности может использоваться, например, метод главного показателя:

$$u^*: \max_{u \in U} W_l(u);$$

$$W_j(u) \geq b_j, \quad j \neq l$$

или метод последовательных уступок.

При выражении степени уверенности ЛПР в «истинности» состояний $\lambda_i \in \Lambda_J^N$ с помощью субъективных вероятностей решение может быть получено путем максимизации среднего значения оценочной функции $f^{\theta c}(y)$.

ЗАДАЧИ ПРИНЯТИЯ РЕШЕНИЙ В УСЛОВИЯХ СТОХАСТИЧЕСКОЙ И «ПРИРОДНОЙ» НЕОПРЕДЕЛЕННОСТИ

Во многих практических задачах неопределенности различного характера наблюдаются совместно. Например, каждая стратегия $u \in U$ по-

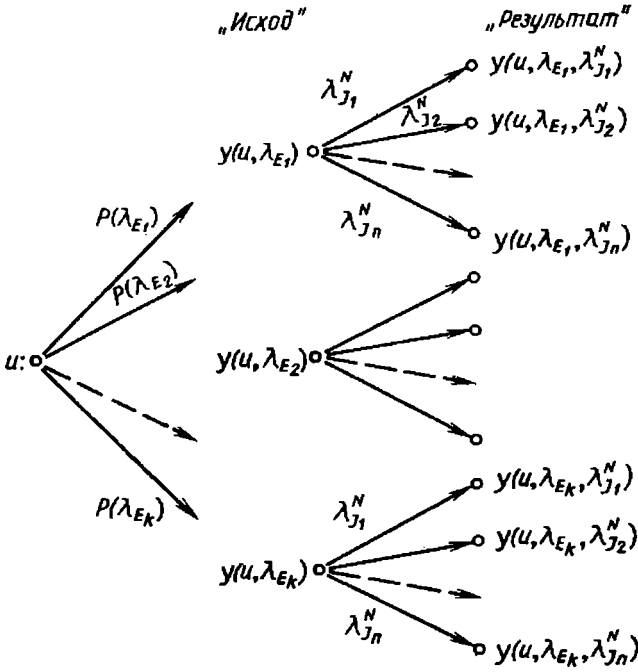


Рис. 18. Схема формирования результата операции в условиях стохастической и «природной» неопределенностей

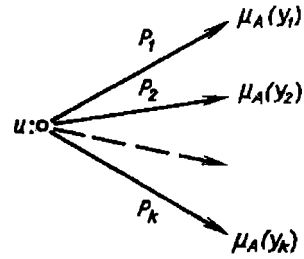


Рис. 19. Нечеткий случайный результат операции

рождает вполне определенное распределение вероятностей на множестве Y характеристик исходов, а результат операции формируется под действием факторов нестохастической природы (рис. 18). Если зафиксировать некоторое состояние «природы» $\lambda \in \Lambda_J^N$, то задача ничем не будет отличаться от задачи принятия решений в условиях стохастической неопределенности. Установленная таким образом условная система предпочтений ЛПР, отражаемая величинами соответствующих условных показателей эффективности $W(u, \lambda) \Big|_{\lambda \in \Lambda_J^N}$, подвергается

дальнейшему анализу с использованием аппарата выработки решений в условиях только «природной» неопределенности.

Если «природная» неопределенность может быть описана в терминах нечетких отношений принадлежности, то может рассматриваться следующая постановка задачи принятия решения. На множестве возможных результатов $y \in Y$ вводится в рассмотрение нечеткое событие A , задаваемое с помощью некоторого нечеткого отношения принадлежности данного исхода к интересующему ЛПР результату операции. Значения y характеристики исхода операции есть реализация случайной величины, распределение ко-

торой зависит от выбранной стратегии $u \in U$, следовательно, множество Λ_J^N состояний «природы» заменяется нечетким множеством $\mu_A(y) = \mu_A \times (y, y_{tr})$. В этом случае результат операций есть нечеткое случайное событие, а каждая стратегия $u \in U$ порождает распределение вероятностей нечеткого случайного результата операции (рис. 19). Тогда в качестве основных числовых характеристик нечеткого случайного результата операции можно использовать вероятность нечеткого случайного события

$$P(\mu_A(y(u))) = \sum_{i=1}^k \mu_A(y_i) P_i(u); \quad (13.113)$$

математическое ожидание нечеткого результата

$$M[Y(u, \lambda)] = \frac{\sum_{i=1}^k y_i \mu_A(y_i) P_i(u)}{\sum_{i=1}^k \mu_A(y_i) P_i(u)} \quad (13.114)$$

и дисперсию этого результата

$$D[Y(u, \lambda)] = \frac{\sum_{i=1}^k (y_i - M[Y(u, \lambda)])^2 \mu_A(y_i) P_i(u)}{\sum_{i=1}^k \mu_A(y_i) P_i(u)} \quad (13.115)$$

Формально величины

$$\frac{\mu_A(y_i) P_i(u)}{\sum_{i=1}^k \mu_A(y_i) P_i(u)},$$

входящие в (13.113)—(13.115), по-прежнему можно рассматривать как субъективное априорное распределение вероятностей результата операции.

Наилучшая стратегия u^* может отыскиваться путем решения задачи u^* : $\max_{u \in U} M[Y(u, \lambda)]$,

являющейся аналогом критерия наибольшего среднего результата. Если в качестве дополнительного показателя использовать выражение (13.115), то задача отыскания наилучшего решения может быть сведена к скалярной с использованием сверток типа (13.61)—(13.63).

Если ЛПР в состоянии определить гарантированный желаемый результат y^{TP} , то наилучшим будет решение, отыскиваемое по правилу:

$$u^*: \max_{u \in U} P(y(u, \lambda) \geq y^{TP}) = \frac{\sum_{i \in I(u)} \mu_A(y_i) P_i(u)}{\sum_{i \in I(u)} \mu_A(y_i)}, \quad (13.116)$$

где $I(u) = \{i \mid y_i(u) \geq y^{TP}\}$;

$$y^{TP} \in \left[\min_u \min_{\lambda} y(u, \lambda), \right.$$

$$\left. \max_u \max_{\lambda} y(u, \lambda) \right].$$

Пример 18. Производственное объединение планирует выпуск пластинчатых теплообменных аппаратов (ПТА) для криогенно-гелиевых установок с целью поставок на международный рынок. Для обеспечения работоспособности ПТА его термодинамический коэффициент полезного действия (КПД) должен быть не ниже $\eta_T^{расч} = 0,97$. Величина η_T стохастически зависит от класса точности σ изготовления теплообменного аппа-

рата, что при заданной точности изготовления σ (стратегия производства) приводит к вполне определенному вероятностному распределению $F_{\sigma}(\eta_T) = P(\hat{\eta}_T(\sigma) < \eta_T)$ (рис. 20, а). Коммерческая стоимость (средняя) одного ПТА при заданной стратегии σ определяется в этом случае из следующих соображений. Если будет сделано N аппаратов стоимостью $C_1(\sigma)$ за один аппарат, то эти расходы распределятся между $N[1 - F(\eta_T^{расч})]$ кондиционными ПТА, которые поступят на рынок. Следовательно, коммерческая цена ПТА должна составить не менее

$$C_k(\sigma) = \frac{C_1(\sigma) N}{N[1 - F_{\sigma}(\eta_T^{расч})]} = \frac{C_1(\sigma)}{1 - F_{\sigma}(\eta_T^{расч})}.$$

Для того чтобы на мировом рынке теплообменные аппараты пользовались «неограниченным спросом» (нечеткое событие A), их показатели η_T и $C_k(\eta_T)$ должны быть определенным образом сбалансированы. Кроме того, чем выше КПД, тем меньше габариты аппарата. Это означает, что функция принадлежности нечеткого события $\mu_A(\eta_T)$ может иметь вид, представленный на рис. 20, б, т. е. ПТА с небольшим КПД, имеющие низкую стоимость и большие габариты, будет пользоваться меньшим спросом, так же как и аппараты с очень высоким КПД, а вследствие этого с высокой стоимостью и малыми габаритами.

Среди возможных значений η_T из интервала $[0,97; 1]$ эксперты выделили строго различающиеся по предпочтительности классы пластинчатых теплообменных аппаратов со значениями η_T :

класс 0: $\eta_T < 0,97$;

класс I: $\eta_T \in [0,97; 0,98]$;

класс II: $\eta_T \in [0,98; 0,99]$;

класс III: $\eta_T > 0,99$.

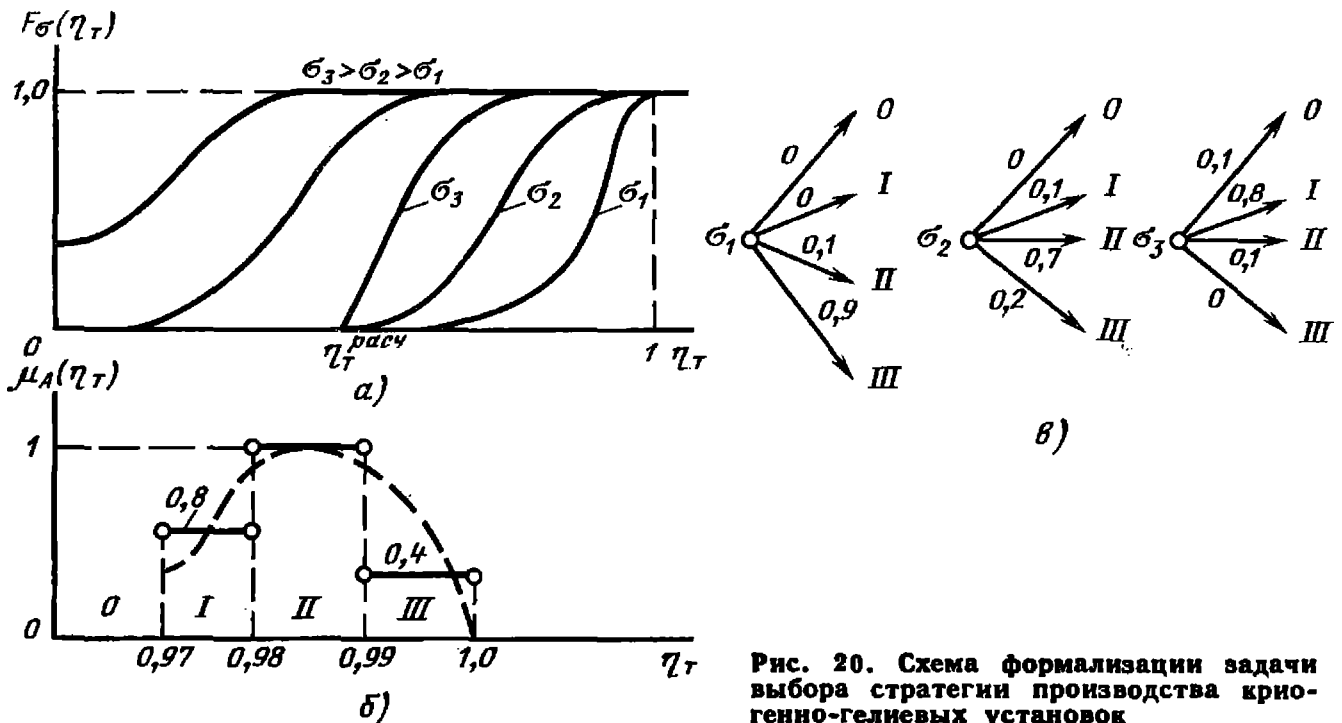


Рис. 20. Схема формализации задачи выбора стратегии производства криогенно-гелиевых установок

Анализ зависимости $F_{\sigma}(\eta_T)$ показал, что поиск решения проблемы следует осуществлять среди стратегий $\sigma_1, \sigma_2, \sigma_3$. На рис. 20, в представлены распределения вероятностей на классах ПТА в зависимости от выбранных стратегий.

Определим значения показателей эффективности для выбранных стратегий производства.

Если в качестве показателя эффективности используется вероятность наступления нечеткого случайного события «спрос на установку неограничен», то согласно (13.113) имеем

$$W_1(\sigma_1) = P(\mu_A(\eta_T(\sigma_1))) = 0,1 \times 1 + 0,9 \times 0,4 = 0,46;$$

$$W_1(\sigma_2) = P(\mu_A(\eta_T(\sigma_2))) = 0,1 \times 0,8 + 0,7 \times 1 + 0,2 \times 0,4 = 0,86;$$

$$W_1(\sigma_3) = P(\mu_A(\eta_T(\sigma_3))) = 0,8 \times 0,8 + 0,1 \times 1 = 0,74,$$

т. е. $\sigma_2 \succ \sigma_3 \succ \sigma_1$.

Для показателя вероятностно гарантированного нечеткого результата при

значении $\eta_{\text{треб}} = 0,985$ имеем [см. (13.116)]:

$$W_2(\sigma_1) = P(\eta_T(\sigma_1, \mu_A) \geq \eta_{\text{треб}}) = \frac{1 \times 0,1 + 0,9 \times 0,4}{1 + 0,4} = 0,329;$$

$$W_2(\sigma_2) = \frac{1 \times 0,7 + 0,4 \times 0,2}{1 + 0,4} = 0,557;$$

$$W_2(\sigma_3) = \frac{1 \times 0,1}{1 + 0,4} = 0,07,$$

т. е. $\sigma_2 \succ \sigma_1 \succ \sigma_3$.

Если же цель операции моделируется стремлением увеличить нечеткий средний результат (показатель $W_3(\sigma_i)$ вычисляется по (13.113)), то

$$W_3(\sigma_1) = \frac{0,985 \times 0,1 \times 1 + 0,995 \times 0,9 \times 0,4}{0,46} = 0,993;$$

$$W_3(\sigma_2) = \frac{0,975 \times 0,1 \times 0,8 + 0,985 \times 0,7 \times 1 + 0,995 \times 0,2 \times 0,4}{0,86} = 0,985;$$

$$W_3(\sigma_3) = \frac{0,975 \times 0,8 \times 0,8 + 0,985 \times 0,1 \times 1}{0,74} = 0,976,$$

т. е. $\sigma_1 \succ \sigma_2 \succ \sigma_3$.

Таким образом, различные цели операции даже в нечеткой формулировке приводят к выделению различных стратегий в качестве оптимальных. Стратегия σ_2 является инвариантной к целям, моделируемым первым и вторым показателями W_1 и W_2 .

Выше рассматривалась задача статистических решений. При этом предполагалось, что ЛПР в общем случае известны значения матрицы потерь Π_{ki} , априорные вероятности состояний факторов λ_i , имеющих стохастическую природу, и условные распределения характеристик z факторов. Теория статистических решений не дает каких-либо рекомендаций по назначению величин Π_{ki} . Тем не менее анализ внешнего дополнения часто позволяет установить диапазоны изменения значений потерь. Используя аппарат нечеткого описания потерь, задачу статистических решений с использованием критерия Байеса можно сформулировать в нечеткой постановке. Отличие подобной постановки двухальтернативной задачи от аналогичной задачи минимизации среднего (четкого) риска состоит в следующем:

$$z^*: \max_{z^0 \in Z} \min \{ \mu_W(W(z^0)), \mu_0(W^{TP}) \}, \quad (13.117)$$

где z^0 — пороговое значение измеряемой характеристики z фактора λ_i ;

$$W(z^0) = \sum_{i=1}^2 P(\lambda_i) \sum_{k=1}^2 \Pi_{ki} P \times \times (z \in Z^k / \lambda_i); \quad (13.118)$$

$$Z^k = \begin{cases} z < z^0, & k = 1; \\ z \geq z^0, & k = 2; \end{cases}$$

Π_{ki} — нечеткая переменная с функцией принадлежности $\mu_{\Pi}(\Pi_{ki})$; W^{TP} — предельно допустимое для ЛПР значение среднего риска; $\mu_0(W^{TP})$ —

функция принадлежности W^{TP} ; $\mu_W \times \times (W(z^0))$ — функция принадлежности среднего нечеткого риска (13.118), определяемая обычным порядком:

$$\mu_W(W(z^0)) = \max \min \{ \mu_{\Pi}(\Pi_{ki}) \} = \left\{ \Pi_{ki} \left| \sum_{i=1}^2 P(\lambda_i) \sum_{k=1}^2 \Pi_{ki} P \times \times (z \in Z^k / \lambda_i) = W(z^0) \right. \right\}.$$

Для установления величины W^{TP} и ее функции принадлежности используют следующий прием. Находят W как решение четкой задачи

$$\min_{z^0 \in Z} \max_{\Pi_{ki} \in \text{Supp } \mu_{\Pi}(\Pi_{ki})} W(z^0)$$

и величину W^+ из задачи

$$\min_{z^0} \min_{\Pi_{ki} \in \text{Supp } \mu_{\Pi}(\Pi_{ki})} W(z^0),$$

а затем полагают

$$W^{TP} = \frac{W^- + W^+}{2} \quad \text{или} \quad W^{TP} = W^+,$$

считая W^{TP} нечеткой переменной с функцией принадлежности типа «примерно равно» или «не менее» с носителем $[W^-, W^+]$. Указанный подход может быть распространен и на многоальтернативный случай принятия решений при нечеткой исходной информации, и на случай описания природных неопределенностей (величин Π_{ki}) субъективными вероятностями.

3. ЗАДАЧИ ПРИНЯТИЯ РЕШЕНИЙ В УСЛОВИЯХ ПОВЕДЕНЧЕСКОЙ НЕОПРЕДЕЛЕННОСТИ

Во многих практических задачах, связанных с оценкой эффективности, результат операции определяется не только действием оперирующей стороны (ЛПР), но и действиями других субъектов системы, которые в общем случае преследуют собственные цели,

не всегда совпадающие с целями ЛПР. Неопределенность такого рода при принятии решений относят к классу поведенческих неопределенностей Λ_J^B . Анализом ситуаций принятия решений в условиях поведенческой неопределенности занимается специальная наука — теория игр. Объектом исследования в теории игр являются конфликтные ситуации, т. е. ситуации, в которых цели и интересы отдельных субъектов не совпадают. Случай совпадающих интересов всех субъектов системы можно рассматривать как задачу принятия решений коллективным согласованным органом.

Основной постулат теории игр — любой субъект системы по меньшей мере так же разумен (рационален), как и оперирующая сторона (ЛПР) и делает все возможное, чтобы достичь своих целей. От реального конфликта игра (математическая модель конфликта) отличается тем, что она ведется по определенным правилам, которые устанавливают порядок и очередность действий субъектов системы, их информированность, порядок обмена информацией, формирование результата игры (кто и сколько выиграл или проиграл).

При формализации реального конфликта обычно различают следующие основные типы игр:

1) стратегические и нестратегические; первые характеризуются тем, что каждый субъект системы действует независимо от остальных, преследуя свои цели, а вторые — тем, что субъекты выбирают единую для всех стратегию;

2) парные игры и игры N-лиц;

3) коалиционные и бескоалиционные;

4) кооперативные и некооперативные; в первых разрешен обмен информацией о возможных стратегиях субъектов системы (игроков);

5) со строгим и нестрогим соперничеством;

6) конечные и бесконечные (в первых — конечное число стратегий игроков) и другие.

При решении задач, связанных с оценением эффективности в технике, особую роль играют парные стратегические бескоалиционные конечные некооперативные игры. Модель проблемной ситуации в этом случае имеет вид $\langle U; V, W_1, W_2, \mathcal{P}_1, \mathcal{P}_2 \rangle$, (13.119)

где U, V — множество стратегий игроков¹; W_1 и W_2 — их показатели эффективности (функции выигрыша), которые могут быть получены с учетом действия факторов Λ_E и Λ_J^N ; $\mathcal{P}_1, \mathcal{P}_2$ — системы предпочтений игроков. Пара стратегий $(u, v) \in U \times V$ называется *ситуацией игры*. Отношение предпочтения на ситуациях вводится обычным порядком:

$$\left. \begin{aligned} \text{для первого игрока } (u_1, v_1) &\geq \\ &\geq (u_2, v_2) \Leftrightarrow W_1(u_1, v_1) \geq \\ &\geq W_1(u_2, v_2); \\ \text{для второго игрока } (u_1, v_1) &\geq \\ &\geq (u_2, v_2) \Leftrightarrow W_2(u_1, v_1) \geq \\ &\geq W_2(u_2, v_2). \end{aligned} \right\} \quad (13.120)$$

Формальное выражение систем $\mathcal{P}_i$ предпочтений игроков ($i = 1, 2$) основывается на двух ведущих принципах рационального поведения: принципе наибольшего гарантированного результата и принципе равновесия.

Принцип наибольшего гарантированного результата основан на том, что рациональным выбором одного из субъектов системы должен считаться такой, при котором он рассчитывается на самую неблагоприятную для него реакцию другого субъекта системы. В соответствии с этим наилучшей стратегией, например, первого игрока является та, для которой значение $\min_{v \in V} W_1(u, v)$ является наибольшим.

Принцип равновесия (равновесие по Нэшу) основан на том, что рациональ-

¹ В дальнейшем везде полагаем, что игрок I имеет множество стратегий U , а игрок II — множество стратегий V .

ным выбором любого субъекта считается такая стратегия (u^0 для первого или v^0 для второго), для которой ситуация (u^0, v^0) обоюдовыгодна: любое отклонение от ситуации (u^0, v^0) не является выгодным ни для одного из игроков.

После выявления структуры предпочтений игрока на множестве $U \times V$ ситуаций игры с целью применения того или иного принципа рационального поведения необходимо, прежде всего, установить, различаются ли интересы игроков на множестве ситуаций. Если, например, предпочтения игроков строго противоположны на всех парах ситуаций, т. е.

$$\begin{matrix} I & & II \\ (u_i, v_j) \geq (u_k, v_l) & \Leftrightarrow & (u_i, v_j) \leq (u_k, v_l), \forall (u, v) \in U \times V, \end{matrix} \quad (13.121)$$

то игра называется игрой со строгим соперничеством (с противоположными интересами). Если к тому же показатели $W_1(u, v)$ и $W_2(u, v)$ однородны и выполняется условие

$$W_1(u, v) + W_2(u, v) = 0, \quad (13.122)$$

$$\forall (u, v) \in U \times V,$$

то игра называется *антагонистической*, или *игрой с нулевой суммой*. Таким образом, для задания антагонистической игры достаточно получить информацию о том, что один игрок выигрывает ровно столько, сколько ему проигрывает другой. Поэтому модель проблемной ситуации (13.119) в этом случае принимает вид

$$(U, V, W), \quad (13.123)$$

причем везде далее полагаем, что $W(u, v) = W_1(u, v)$.

При ведении антагонистической игры каждый игрок естественно следует принципу наибольшего гарантированного результата, т. е. наилучшей стратегией первого игрока считается

$$u^*: \max_{u \in U} W_e^I(u), \quad (13.124)$$

где

$$W_e^I(u) = \min_{v \in V} W(u, v).$$

Наилучшая стратегия u^* первого игрока, определяемая по правилу (13.124), называется *максиминной*. Соответствующий ей наибольший гарантированный результат W^- носит название *нижней цены игры*. Наилучшая стратегия второго игрока $v^* \in V$, определяемая по правилу

$$v^*: \min_{v \in V} W_e^{II}(v), \quad (13.125)$$

$$\text{где } W_e^{II}(v) = \max_{u \in U} W(u, v),$$

называется *минимаксной*, так как она обеспечивает ему наименьший гарантированный проигрыш W^+ , который представляет собой верхнюю цену игры.

Таким образом, если игроки I и II выбирают наилучшие стратегии по правилам (13.124), (13.125) соответственно, то первый не может «выиграть» меньше W^- , а второй — проиграть ему больше, чем W^+ . Отсюда следует важное соотношение

$$W^- \leq W^+. \quad (13.126)$$

Стратегии, определенные по правилам (13.124), (13.125), будут удовлетворять принципу равновесия лишь в том случае, если нижняя цена игры совпадает с верхней, т. е.

$$W^- = W^+ = W^0. \quad (13.127)$$

В этом случае говорят, что игра имеет седловую точку (u^0, v^0) , для которой

$$\begin{aligned} W(u, v^0) &\leq W(u^0, v^0) \leq \\ &\leq W(u^0, v). \end{aligned} \quad (13.128)$$

Выражение (13.128) представляет собой формальную запись принципа равновесия по Нэшу для антагонистической игры. Седловая точка существует, если множества U и V замкнуты, выпуклы и ограничены, функция $W(u, v)$ непрерывна, вогнута по $u \in U$ и выпукла по $v \in V$ (см. т. 2).

МАТРИЧНЫЕ ИГРЫ

Парная антагонистическая игра с конечными множествами U и V стратегий игроков называется *матричной*.

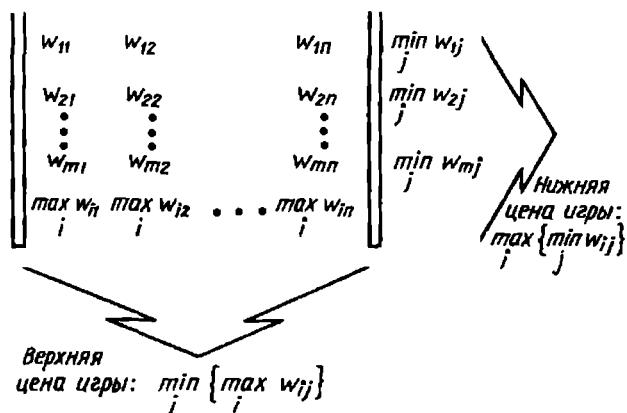


Рис. 21. Схема отыскания нижней и верхней цены игры $m \times n$

Термин «матричная игра» объясняется тем, что все ситуации игры могут быть представлены как элементы матрицы

$$W = \|w_{ij}\|, \quad i = \overline{1, m}, \quad j = \overline{1, n},$$

где $w_{ij} = W(u_i, v_j)$.

Приведение игры к матричной форме является специфической задачей принятия решений, связанной с формированием исходных множеств стратегий игроков и моделированием исходов операции. Чаще всего представление реального конфликта матричной игрой проводится на ранних этапах исследования операций, когда не требуется находить сразу наилучшие стратегии, а достаточно определить лишь области возможных решений для проведения дальнейших обобщенных и детальных исследований (см. гл. 10). Таким образом в рамках концептуальных исследований как строгих, так и нестрогих конфликтов проблема описания ситуации принятия решений в виде матричной игры имеет два аспекта. Во-первых, представление множеств U и V стратегий конечным множеством репрезентативных альтернатив и, во-вторых, — формирование адекватных в рамках выбранного уровня исследования моделей целей и исходов операций (которые описываются скалярным показателем $W(u, v)$, отражающим строгий антагонизм возможно и в нестрогом конфликте сторон).

Для существования в матричной игре седловых точек необходимо и достаточно, чтобы

$$i^*, j^*: \max_i \min_j w_{ij} = \min_j \max_i w_{ij}.$$

В силу конечности множеств стратегий игроков решение минимаксной задачи не представляет затруднений, так как может быть осуществлено последовательно по схеме, представленной на рис. 21. Однако равенство нижней и верхней цены матричной игры в чистых стратегиях, как правило, не выполняется, что приводит к нарушению условия равновесия по Нэшу. Это объясняется взаимной возможностью сторон адаптироваться к поведению противника, т. е. повышать ранг рефлексии: прогнозировать возможную ответную реакцию противника и отклоняться от своей чистой стратегии i^* (или j^*) с целью улучшить значение выигрыша. Однако для того чтобы использовать критерий адаптивности, игрок должен получать текущую информацию, а следовательно, игра должна повторяться многократно. Таким образом, использование критерия адаптивности придает дополнительные стратегические возможности игрокам, позволяет им находить рациональное сочетание чистых стратегий, использование которых в определенной пропорции дает в среднем дополнительный выигрыш. Если бы такая игра повторялась достаточно большое число раз (в пределах бесконечное), а стороны осуществляли выбор решений на основе критерия адаптивности, то их наилучшей стратегией в смысле этого критерия оказались бы так называемые смешанные стратегии.

Смешанной стратегией игрока называется вероятностное распределение на множестве его чистых стратегий.

Таким образом, случайная переменная $\hat{a}(\theta)$, значениями которой являются чистые стратегии i (соответственно j), является смешанной стратегией первого (второго) игрока с функцией вероятности $P^I = (P_1^I, P_2^I, \dots, P_m^I)$ (соответственно $P^{II} = (P_1^{II}, P_2^{II}, \dots, P_n^{II})$).

В этом случае математическое ожидание выигрыша каждого из игроков определяется при их фиксированных смешанных стратегиях $\hat{u}$ и $\hat{v}$ обычным порядком:

$$M[W(\hat{u}, \hat{v})] = \sum_{i=1}^m \sum_{j=1}^n w_{ij} P_i^I P_j^{II} \quad (13.129)$$

Использование смешанных стратегий приводит к тому, что средний выигрыш (13.129) — непрерывная функция на $U \times V$, а сами множества $\hat{U}$ и $\hat{V}$ — замкнутые многогранники $\left(\sum_{i=1}^m P_i^I = 1, \sum_{j=1}^n P_j^{II} = 1 \right)$. Следовательно, существует пара стратегий $\hat{u}^0, \hat{v}^0$, которые удовлетворяют условию (13.128) с заменой $W(u, v)$ на $M \times [W(\hat{u}, \hat{v})]$, т. е. ситуация $(\hat{u}^0, \hat{v}^0)$ является седловой точкой (основная теорема матричных игр).

Теорема. Какова бы ни была матрица $W(u, v)$, всегда найдется хотя бы одна ситуация равновесия (возможно в смешанных стратегиях).

$M[W(\hat{u}^0, \hat{v}^0)]$ удовлетворяет условию

$$W^- \leq M[W(\hat{u}^0, \hat{v}^0)] \leq W^+.$$

Отыскание решения игры осуществляется в общем случае методом сведения матричной игры к задаче линейного программирования. Для этого вначале на исходных множествах U и V чистых стратегий игроков выделяют недоминируемые их подмножества по правилам:

$$\left. \begin{aligned} u_i \geq u_l &\iff W(u_i, v) \geq W(u_l, v), \forall v \in V; \\ v_j \geq v_k &\iff W(u, v_j) \leq W(u, v_k), \forall u \in U. \end{aligned} \right\} \quad (13.130)$$

Эта операция приводит к редуцированной матрице, получаемой из начальной путем «вычеркивания» доминируемых (отношение доминирования касается только стратегий одного и того же игрока) строк и столбцов. Далее

осуществляется линейное преобразование:

$$c: a_{ij} = w_{ij} + c, \quad c > 0, \quad a_{ij} \geq 0,$$

приводящее к новой матрице $A = \|a_{ij}\|$ с положительными компонентами.

Решаются прямая и двойственная задачи линейного программирования (см. т. 2):

$$\left. \begin{aligned} x^*: \min \sum_{i=1}^m x_i; \\ x \in X = \{x \mid A^T x \geq 1, x_i \geq 0\}; \end{aligned} \right\} \quad (13.131)$$

$$\left. \begin{aligned} y^*: \max \sum_{j=1}^n y_j; \\ y \in Y = \{y \mid Ay \leq 1, y_j \geq 0\}. \end{aligned} \right\} \quad (13.132)$$

Тогда оптимальные смешанные стратегии определяются по правилу: для первого игрока

$$P_i^I = \frac{x_i}{\sum_{i=1}^m x_i},$$

для второго игрока

$$P_j^{II} = \frac{y_j}{\sum_{j=1}^n y_j}.$$

Если один из игроков имеет только две стратегии (игра $2 \times n$ или $m \times 2$), то оптимальные смешанные стратегии могут быть найдены графоаналитическим методом.

Рассмотрим игру $2 \times n$, матрица которой имеет вид:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \end{pmatrix}.$$

Обозначим P_1^I через x , тогда $P_2^I = 1 - x$. Учитывая, что

$$\begin{aligned} M[W(\hat{u}^0, v)] &= \min_j M[W(\hat{u}^0, j)] = \\ &= \max_{x \in [0; 1]} \min_j M[W(x, j)], \end{aligned}$$

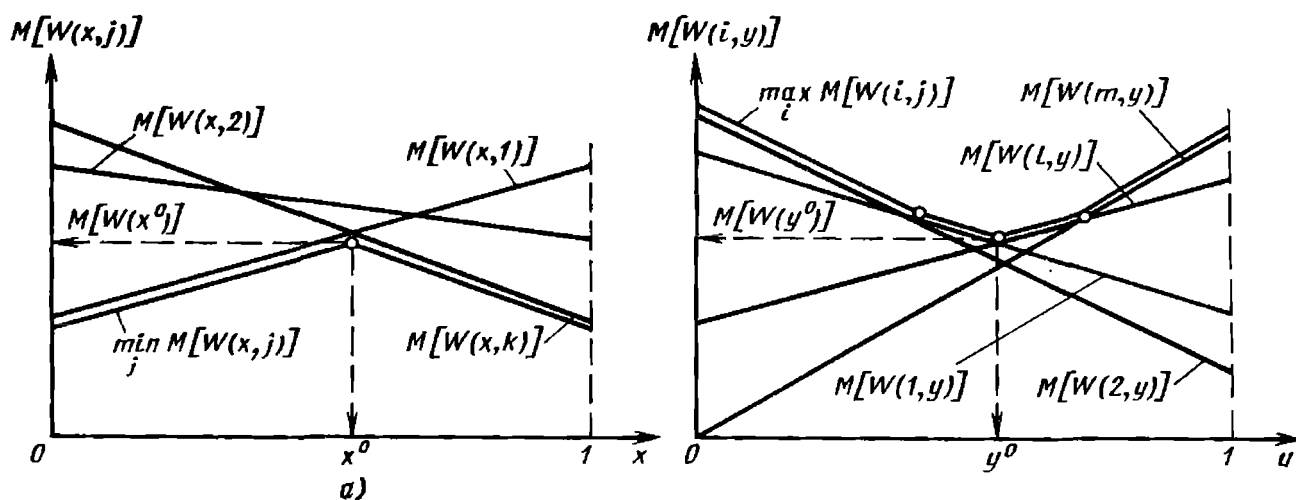


Рис. 22. Графический метод решения матричной игры: а — $2 \times n$; б — $m \times 2$

составим уравнения для определения величины $M[W(x, j)]$:

$$j = 1: M[W(x, 1)] = a_{11} \cdot x + a_{21} \times (1 - x);$$

$$j = 2: M[W(x, 2)] = a_{12} \cdot x + a_{22} \times (1 - x);$$

$$j = n: M[W(x, n)] = a_{1n} \cdot x + a_{2n} \times (1 - x).$$

Графиками уравнений $M[W(x, j)]$ являются прямые, определенные на значениях $x \in [0; 1]$ (рис. 22, а). График функции $\min_j M[W(x, j)]$ является нижней огибающей графиков $M[W(x, j)]$. Оптимальной смешанной стратегией первого игрока в этом случае является распределение

$$P_1^I = x^0, P_2^I = 1 - x^0,$$

дающее наибольший средний гарантированный результат. Чтобы найти оптимальную стратегию θ^0 , необходимо поступить следующим образом:

определить номера j_1 и j_2 чистых стратегий второго игрока, для которых линии $M[W(x, j_1)]$ и $M[W(x, j_2)]$ пересекаются в точке с абсциссой x^0 и ординатой $M[W(x^0)]$; положить $P_{j_1}^{II}$ и $P_{j_2}^{II}$ отличными от нуля, а остальные P_j^{II} , $j \neq j_1, j \neq j_2$ равными нулю; решить уравнение относительно y :

$$y^0: a_{1j_1} \cdot y + a_{1j_2} (1 - y) = a_{2j_1} \cdot y + a_{2j_2} (1 - y);$$

$$\text{положить } P_{j_1}^{II} = y^0, P_{j_2}^{II} = 1 - y^0.$$

Игра типа $m \times 2$ представляется матрицей вида

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ \vdots & \vdots \\ a_{m1} & a_{m2} \end{pmatrix};$$

обозначив через y величину P_1^{II} и учитывая, что

$$M[W(u, \theta^0)] = \max_i M[W(i, \theta^0)] = \min_{y \in [0; 1]} \max_i M[W(i, y)],$$

получаем

$$i = 1: M[W(1, y)] = a_{11}y + a_{12}(1 - y);$$

$$i = 2: M[W(2, y)] = a_{21}y + a_{22}(1 - y);$$

$$i = m: M[W(m, y)] = a_{m1}y + a_{m2}(1 - y).$$

Графики зависимостей $M[W(i, y)]$ и их верхняя огибающая $M[W(i, y)]$ представлены на рис. 22, б. Оптимальной смешанной (рандомизированной) стратегией второго игрока является следующее распределение вероятностей на его чистых стратегиях:

$$P^{II} = (y^0, 1 - y^0).$$

Смешанная стратегия первого игрока определяется из уравнения

$$x^0: a_{i_1,1}x + a_{i_2,1}(1-x) = \\ = a_{i_1,2}x + a_{i_2,2}(1-x),$$

где i_1 и i_2 — номера чистых стратегий первого игрока, для которых прямые $M[W(i_1, y)]$ и $M[W(i_2, y)]$ пересекаются в точке с абсциссой $y = y^0$ и ординатой $M[W(y^0)]$.

Пример 19. В условиях примера 4 поиск рандомизированной (смешанной) стратегии u_p^* осуществим графоаналитическим методом решения матричной игры $m \times 2$. Выделенное в примере 4 недоминируемое множество стратегий $\{u_1, u_2, u_3, u_5, u_8, u_{14}, u_{23}\}$ приводит к редуцированной матрице игры 7×2 :

$$A = \begin{pmatrix} 0 & 10 \\ 1,45 & 5,79 \\ 0,58 & 7,47 \\ 6,56 & 0,97 \\ 3,49 & 2,90 \\ 10 & 0 \\ 7,97 & 0,39 \end{pmatrix}.$$

Обозначим через y вероятность состояния λ_1 ; тогда

$$u_1: i = 1: 0 \cdot y + 10(1-y) = 10 - 10y;$$

$$u_2: i = 2: 1,45y + 5,79(1-y) = \\ = 5,79 - 4,34y;$$

$$u_3: i = 3: 0,58y + 7,47(1-y) = \\ = 7,47 - 6,89y;$$

$$u_5: i = 4: 6,56y + 0,97(1-y) = \\ = 0,97 + 5,59y;$$

$$u_8: i = 5: 3,49y + 2,90(1-y) = \\ = 2,90 + 0,59y;$$

$$u_{14}: i = 6: 10y + 0(1-y) = 10y;$$

$$u_{23}: i = 7: 7,94y + 0,39(1-y) = \\ = 0,39 + 7,55y.$$

На рис. 23 представлены графики $M[W(i, y)]$, характеризующие зависимость средних потерь ЛПР от при-

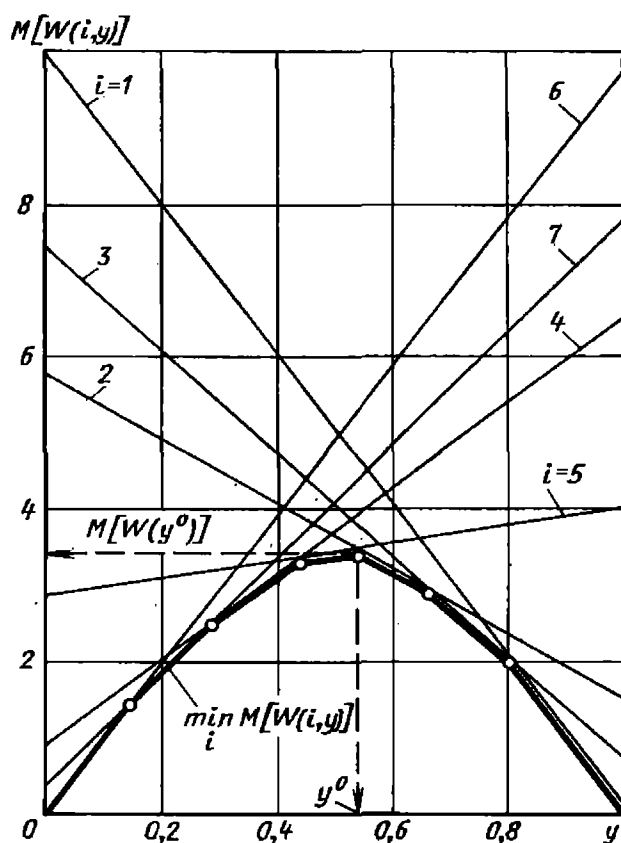


Рис. 23. Схема определения рандомизированной стратегии контроля

нятия решения $u \in U_{\omega_0}^*$, и их нижняя огибающая $\min_i M[W(i, y)]^{*1}$.

Точка с координатами $(y^0; M[W(y^0)])$ является точкой пересечения графиков $M[W(2, y)]$ и $M[W(5, y)]$, что соответствует применению рандомизированной смеси чистых стратегий u_2 и u_8 . Поэтому полагаем отличными от нуля вероятности P_2^I и P_5^I , значения которых определяются из условия:

$$1,45x + 3,49(1-x) = 5,79x + 2,90 \times \\ \times (1-x).$$

Решением полученного уравнения является $x^0 = 0,12$. Тогда $P_2^I = 0,12$ и $P_5^I = 0,88$, обеспечивающие гарантированные средние потери, равны 3,25.

*1 Считаем, что природа «настроена антагонистически по отношению к ЛПР», т. е. действует таким образом, чтобы уличить гарантированные потери.

БИМАТРИЧНЫЕ ИГРЫ

Представление реального конфликта матричной игрой — лишь начальный этап исследования. Адекватность такого представления определяется степенью антагонизма реального конфликта. При более тонком анализе необходимо учитывать то обстоятельство, что интересы субъектов не являются строго противоположными, т. е. (13.121) выполняется не для всех ситуаций игры. Такие конфликты могут быть описаны неантагонистическими играми.

Как и ранее, основным принципом выбора игроками своих стратегий является принцип наибольшего гарантированного результата. Однако он не исчерпывает всех стратегических возможностей игроков. В отличие от антагонистических игр, в которых равновесная ситуация (в чистых или смешанных стратегиях) является недоминируемой по наибольшему гарантированному результату, в неантагонистических играх равновесная ситуация может доминироваться другой (может быть даже и неравновесной) ситуацией.

Для неантагонистических игр равновесная ситуация (α^0, β^0) (по Нэшу) определяется условиями:

$$M[W_1(\alpha^0, \beta^0)] \geq M[W_1(\alpha, \beta^0)];$$

$$M[W_2(\alpha^0, \beta^0)] \geq M[W_2(\alpha^0, \beta)].$$

(13.133)

Любая конечная антагонистическая игра имеет хотя бы одну равновесную ситуацию (13.133), и, кроме того, максиминные выигрыши игроков не превосходят их равновесных выигрышей. Поэтому в случае неантагонистической игры использование игроками максиминных стратегий может считаться рациональным лишь в том случае, если максиминный выигрыш совпадает по величине с равновесным. Следовательно, решение неантагонистических игр сопряжено с дополнительными трудностями: необходимо отыскивать не только равновесные, но и максиминные точки. Общим правилом рационального поведения в не-

антагонистической некооперативной игре является следование:

максиминной стратегии, если ее выигрыш совпадает с равновесным; стратегии, приводящей к равновесному выигрышу, если его величина превышает максиминный и есть уверенность, что противник также будет придерживаться равновесной стратегии;

максиминной стратегии, если величины выигрышей противника известны лишь приблизительно или вообще неизвестны.

Общие методы решения конечной антагонистической игры N лиц в равновесных стратегиях неизвестны. Однако для конечных игр двух лиц, которые наиболее распространены в практике оценивания эффективности и надежности технических систем, решение может быть найдено просто.

Конечная бескоалиционная игра двух лиц называется *биматричной*. Модель такой игры описывается проблемной ситуацией (13.119) в предположении, что наиболее предпочтительными стратегиями игроков могут быть как максиминные (α^*, β^*) , так и равновесные (α^0, β^0) . Описывать игру удобно двухэлементной матрицей $\|w_{1ij}, w_{2ij}\|$ (биматрицей). Алгоритм для нахождения недоминируемой равновесной ситуации (i^0, j^0) может быть составлен на основе системы (13.133), так как пара чистых стратегий i^0, j^0 удовлетворяет условиям:

$$\begin{aligned} i^0: \max_i w_{1ij^0} &= w_{1i^0j^0}; \\ j^0: \max_j w_{2i^0j} &= w_{2i^0j^0}. \end{aligned} \quad (13.134)$$

Алгоритм 8. 1. Положить $j = 1$.

2. Найти $l_j: \max_l w_{1lj}$.

3. Проверить условие $j > n$: да — перейти к п. 5, нет — к п. 4.

4. Положить $j = j + 1$ и перейти к п. 2.

5. Положить $i = 1$.

6. Найти $k_i: \max_k w_{2ik}$.

7. Проверить условие $i > m$: да — перейти к п. 9; нет — к п. 8.

8. Положить $i = i + 1$ и перейти к п. 6.

9. Сформировать множество пар индексов (I_0, J_0) , для которых выполняется условие:

$$(I_j, j)^T = (i, k_i)^T, \quad i = \overline{1, m}, j = \overline{1, n};$$

если множество пар (I_0, J_0) не пусто, то перейти к п. 10, иначе — к п. 11.

10. Определить недоминируемую равновесную ситуацию (i^0, j^0) по правилу

$$(i^0, j^0): w_{1 \ i^0 j^0} \geq w_{1 \ i j} \wedge w_{2 \ i^0 j^0} \geq w_{2 \ i j}, \quad \forall (i, j) \in (I_0, J_0), \quad (13.135)$$

причем хотя бы одно из неравенств в (13.135) должно быть нестрогим. Если стратегия, удовлетворяющая (13.135), не найдена, то перейти к п. 11, иначе — к п. 13.

11. Недоминируемая равновесная ситуация отсутствует.

12. Окончить вычисления.

Если равновесная по Нэшу ситуация в чистых стратегиях отсутствует, то, руководствуясь условием (13.133), можно отыскать ситуацию равновесия (i^0, j^0) в смешанных стратегиях, которая в биматричной игре всегда существует. Идея алгоритма нахождения пары (i^0, j^0) основана на следующем положении. Равновесной смешанной стратегией одного из игроков является такая, при которой средний выигрыш другого не зависит от применяемой им стратегии.

Пример 20. Реальная ситуация с нестрогим конфликтом моделируется игрой, заданной матрицей:

$$\begin{pmatrix} (4; 2) & (1; 6) \\ (0; 3) & (3; 1) \end{pmatrix}.$$

Нетрудно убедиться, что игра не имеет ситуаций равновесия по Нэшу в чистых стратегиях (см. алгоритм 8). Для того чтобы определить рациональную стратегию каждого игрока, необходимо найти максиминный и равновесный выигрыши и сравнить их.

Для обеспечения полноты анализа предварительно найдем чистые максиминные стратегии игроков:

$$i^*: \max_i \min_j w_{1 \ i j}; \quad i^* = 1;$$

$$j^*: \max_j \min_i w_{2 \ i j}, \quad j^* = 1,$$

которые дают им наибольшие гарантированные результаты $W_1^-(i^*) = 1$ и $W_2^-(j^*) = 2$ соответственно. Если игра проводится лишь 1 раз или игроки не знают платежных матриц противника, то, выбирая свои максиминные стратегии, они придут к ситуации (i^*, j^*) , которая реально обеспечивает выигрыши $W_1(i^*, j^*) = 4$, $W_2(i^*, j^*) = 2$. Однако эта ситуация неустойчива, так как если второй игрок захочет улучшить свое положение на основе информации об устремлениях первого игрока (узнав его матрицу выигрышей), то он может выбрать стратегию $j = 2$, что приведет к ситуации $(i^*, 2)$, невыгодной для первого игрока. В свою очередь первый игрок, узнав платежную матрицу второго и предполагая описанную выше его реакцию (т. е. рефлексировав его поведение), может также отклониться от стратегии i^* , что приведет к ситуации $(2, 2)$, невыгодной для второго игрока, и т. п. Подобная многоуровневая рефлексия (т. е. использование критерия адаптивности) приводит в итоге к смешанным максиминным стратегиям игроков.

Максиминные смешанные стратегии игроков определяются по их собственным платежным матрицам порядком, описанным на стр. 314. Поскольку рассматриваемая в примере 20 игра имеет размерность 2×2 , ее можно решить аналитическим методом. Например, стратегия i^* определяется следующим образом: первый игрок должен выбрать такие вероятности P_1^{*I} и P_2^{*II} , которые обеспечивают ему неизменный выигрыш независимо от того, какую стратегию выбирает второй, т. е., если $P_1^{*I} = x$, то $P_2^{*II} = 1 - x$ и, следовательно,

$$4x + 0(1 - x) = 1x + 3(1 - x),$$

$$\text{откуда } x = \frac{1}{2}. \text{ Таким образом } P_1^{*I} = \frac{1}{2}, \quad P_2^{*II} = \frac{1}{2}, \quad M[W_1(i^*)] = 2.$$

Аналогично для второго игрока

$$2y + 6(1 - y) = 3y + 1(1 - y),$$

$$\text{откуда } y = \frac{5}{6}, P_1^{*II} = \frac{5}{6}, P_2^{*II} = \frac{1}{6}, M[W_2(\theta^*)] = 2 \frac{2}{3}.$$

Сравнение полученных результатов показывает, что смешанные максиминные стратегии обеспечивают игрокам больший средний гарантированный результат.

Найдем равновесную по Нэшу ситуацию (a^0, θ^0) в смешанных стратегиях. Для этого первый игрок может определить вероятности P_1^{0I} и P_2^{0I} , взяв за основу матрицу выигрышей второго игрока, выбрав такие их величины, чтобы выигрыш второго игрока не зависел от выбираемой им стратегии. Это приводит к следующему уравнению:

$$2x + 3(1 - x) = 6x + (1 - x) \cdot 1,$$

$$\text{откуда } x = \frac{1}{3}; \text{ следовательно, } P_1^{0I} = \frac{1}{3}, P_2^{0I} = \frac{2}{3}, M[W_1(\theta^0)] = 2 \frac{2}{3}.$$

Аналогично для второго игрока

$$4y + 1 \cdot (1 - y) = 0 \cdot y + 3(1 - y),$$

$$\text{откуда } y = \frac{1}{3}, P_1^{0II} = \frac{1}{3}, P_2^{0II} = \frac{2}{3}, M[W_1(\theta)] = 2.$$

Таким образом, выигрыши игроков в ситуациях (a^*, θ^*) и (a^0, θ^0) совпадают и соответственно равны

$$M[W_1(a^*, \theta^*)] = M[W_1(a^0, \theta^0)] = 2;$$

$$M[W_2(a^*, \theta^*)] =$$

$$= M[W_2(a^0, \theta^0)] = 2 \frac{2}{3};$$

поэтому применять стратегии a^0 и θ^0 игрокам не обязательно. Это связано с тем, что отклонение одного из игроков от равновесной по Нэшу ситуации (a^0, θ^0) не может улучшить его положение, но может существенно повлиять на выигрыш другого игрока. Действительно, если второй игрок применит, например, максиминную стратегию θ^* в то время, как первый

придерживается стратегии a^0 , то средний выигрыш первого в ситуации (a^0, θ^*) составит

$$\begin{aligned} M[W_1(a^0, \theta^*)] &= 4 \frac{1}{3} \times \frac{5}{6} + 0 + \\ &+ 1 \times \frac{1}{3} \times \frac{1}{6} + 3 \times \frac{2}{3} \times \frac{1}{6} = \\ &= 1 \frac{1}{2} < M[W_1(a^0, \theta^0)]. \end{aligned}$$

Если от равновесной ситуации будет отклоняться первый игрок и применит свою стратегию a^* , то второй получит выигрыш

$$\begin{aligned} M[W_2(a^*, \theta^0)] &= 2 \times \frac{1}{2} \times \frac{1}{3} + \\ &+ 6 \times \frac{1}{2} \times \frac{1}{3} + 3 \times \frac{1}{2} \times \frac{1}{3} + \\ &+ 1 \times \frac{1}{2} \times \frac{2}{3} = 3 \frac{1}{6} > \\ &> M[W_2(a^0, \theta^0)]. \end{aligned}$$

Таким образом, для первого игрока целесообразно придерживаться максиминной, а для второго — равновесной смешанных стратегий.

Рассмотренный пример показывает, что при решении биматричных игр необходим весьма полный и подробный анализ, включающий ряд неформальных этапов. При этом необходимо определить максиминные и равновесные ситуации как в чистых, так и в смешанных стратегиях, а выбор решения осуществлять на основе анализа предпочтительности для игроков возможных отклонений от равновесных ситуаций.

ИГРЫ В НЕЧЕТКОЙ ПОСТАНОВКЕ

Модель конфликта, в которой цели сторон и их допустимые стратегии описаны в терминах нечетких множеств, называется *игрой в нечеткой постановке*.

На множествах U и V стратегий игроков введены нечеткие события с функциями принадлежности $\mu_1(u)$ и $\mu_2(v)$, ставящие в соответствие каж-

дой из стратегий значение функции принадлежности:

$$\mu_1(u): U \rightarrow [0; 1];$$

$$\mu_2(v): V \rightarrow [0; 1].$$

Нечеткие события на множествах U и V могут вводиться с помощью нечетких высказываний (лингвистических переменных) типа «вариант дешевый», «стратегия является перспективной» и т. д. На множестве $U \times V$ ситуаций игры заданы нечеткие события, соответствующие достижению цели первым и вторым игроком:

$$\mu_{W_1}(u, v): U \times V \rightarrow [0; 1];$$

$$\mu_{W_2}(u, v): U \times V \rightarrow [0; 1].$$

Можно рассматривать следующие нечеткие цели: «результат близок к желаемому», «исход удовлетворительный» и т. д. Степень достижения цели при фиксированных стратегиях сторон в этом случае характеризуется значениями функций принадлежности, являющихся пересечением нечетких событий $\mu_i(u)$ и $\mu_{W_i}(u, v)$:

$$\begin{aligned} \text{для первого игрока } \mu_{A_1}(u, v) &= \\ &= \min \{ \mu_1(u), \mu_{W_1}(u, v) \}; \end{aligned} \quad (13.136)$$

$$\begin{aligned} \text{для второго игрока } \mu_{A_2}(u, v) &= \\ &= \min \{ \mu_2(v), \mu_{W_2}(u, v) \}. \end{aligned}$$

Формально можно полагать, что каждый игрок стремится максимизи-

ровать значения соответствующей функции принадлежности $\mu_{A_i}(u, v)$, $i = 1, 2$. В этом случае модель проблемной ситуации (13.119) принимает вид:

$$\langle U, V, \mu_{A_1}, \mu_{A_2} \rangle, \quad (13.137)$$

т. е. игра в нечеткой постановке сводится к обычной четкой игре. При решении игр типа (13.137) возможны два подхода. Во-первых, их можно решать как обычную биматричную игру, руководствуясь принципом наибольшего гарантированного результата (возможно в смешанных стратегиях):

$$\begin{aligned} u^*: \max_u \min_v \mu_{A_1}(u, v); \\ v^*: \max_v \min_u \mu_{A_2}(u, v) \end{aligned} \quad (13.138)$$

или принципом равновесия по Нэшу:

$$\begin{aligned} \mu_{A_1}(u^0, v^0) &\geq \mu_{A_1}(u, v^0); \\ \mu_{A_2}(u^0, v^0) &\geq \mu_{A_2}(u^0, v). \end{aligned} \quad (13.139)$$

Во-вторых, на множествах μ_{A_1} и μ_{A_2} можно выделить подмножества α - и β -уровней:

$$\begin{aligned} U_\alpha &= \{u \mid \mu_{A_1}(u, v) \geq \alpha, \forall v \in V\}; \\ V_\beta &= \{v \mid \mu_{A_2}(u, v) \geq \beta, \forall u \in U\} \end{aligned}$$

и решать обычную игру на множествах стратегий U_α и V_β , обеспечивая тем самым степень достижения целей игроков не ниже заданной.

Список литературы

1. Автоматизация проектирования вычислительных систем. Языки моделирования и базы данных/Под ред. М. Брейера. М.: Мир, 1979. 463 с.
2. Айвазян С. А., Енюков И. С., Мешалкин С. Д. Прикладная статистика. Основы моделирования и первичная обработка данных. М.: Финансы и статистика, 1983. 471 с.
3. Акофф Р. Искусство решения проблем: Пер. с англ. М.: Мир, 1982. 224 с.
4. Атогин В. В., Згуровский М. З. Моделирование на цифровых, аналоговых и гибридных ЭВМ. Киев: Вища школа, 1983. 279 с.
5. Бендат Дж., Пирсол А. Применение корреляционного и спектрального анализа: Пер. с англ. М.: Мир, 1983. 312 с.
6. Бешелев С. Д., Гурвич Ф. Г. Математико-статистические методы экспертных оценок. 2-е изд., перераб. и доп. М.: Статистика, 1980. 263 с.
7. Бир Ст. Кибернетика и управление производством. М.: Наука, 1965. 392 с.
8. Брукс Ф. П. Как проектируются и создаются программные комплексы. М.: Наука, 1979. 151 с.
9. Бусленко В. Н. Автоматизация имитационного моделирования сложных систем. М.: Наука, 1977. 240 с.
10. Бусленко Н. П. Моделирование сложных систем. М.: Наука, 1978. 400 с.
11. Васильев Ф. П. Методы решения экстремальных задач. М.: Наука, 1981. 400 с.
12. Вельбицкий И. В., Ходаковский В. Н., Шолмов Л. И. Технологический комплекс производства программ на машинах ЕС ЭВМ и БЭСМ-6. М.: Статистика, 1980. 263 с.
13. Вентцель Е. С. Исследование операций. М.: Советское радио, 1972. 551 с.
14. Вилкас Э. Й., Майминас Е. З. Решения: теория, информация, моделирование. М.: Радио и связь, 1981. 328 с.
15. Волков Л. И. Управление эксплуатацией летательных комплексов. М.: Высшая школа, 1981. 368 с.
16. Гермейер Ю. Б. Введение в теорию исследования операций. М.: Наука, 1971. 384 с.
17. Горстко А. Д. К вопросу о содержании понятия «имитационное моделирование»//Имитационное моделирование экономических систем. М.: Наука, 1978. С. 21—28.
18. Гроп Д. Методы идентификации систем: Пер. с англ. М.: Мир, 1979. 302 с.
19. Глушко В. М., Иванов В. В., Яненко В. М. Моделирование развивающихся систем. М.: Наука, 1983. 350 с.
20. Горохов В. Г. Методологический анализ системотехники. М.: Радио и связь, 1982. 160 с.
21. Дружинин В. В., Конторов Д. С. Системотехника. М.: Радио и связь, 1985. 200 с.
22. Евланов Л. Г., Кутузов В. А. Экспертные оценки в управлении. М.: Экономика, 1978. 133 с.
23. Ефимов В. М., Комаров В. Ф. Введение в управленческие имитационные игры. М.: Наука, 1980. 272 с.
24. Заде Л. А. Понятие лингвистической переменной и его применение к принятию приближенных решений: Пер. с англ./Пер. Н. И. Ринго. М.: Мир, 1976. 165 с.
25. Ивахненко А. Г. Индуктивный метод самоорганизации моделей слож-

ных систем. Киев: Наукова думка, 1982. 296 с.

26. Ильичев А. В., Волков В. Д., Грущанский В. А. Эффективность проектируемых элементов сложных систем. М.: Высшая школа, 1982. 280 с.

27. Имитационное моделирование производственных систем/Под общ. ред. А. А. Вавилова. М.: Машиностроение; Берлин: Техника, 1983. 416 с.

28. Иодан Э. Структурное проектирование и конструирование программ: Пер. с англ. М.: Мир, 1979. 415 с.

29. Калашников В. В. Качественный анализ поведения сложных систем методом пробных функций. М.: Наука, 1978. 248 с.

30. Кини Р. Л., Райфа Х. Принятие решений при многих критериях: предпочтения и замещения. М.: Радио и связь, 1981. 560 с.

31. Клыков Ю. И., Горьков П. Н. Банки данных для принятия решений. М.: Сов. радио, 1980. 208 с.

32. Кофман А. Введение в теорию нечетких множеств: Пер. с франц./Под ред. Травкина М.: Радио и связь, 1982. 432 с.

33. Куропаткин П. В. Оптимальные и адаптивные системы. М.: Высшая школа, 1980. 287 с.

34. Липаев В. В. Надежность программного обеспечения АСУ. М.: Энергоиздат, 1981. 240 с.

35. Литвак Б. Г. Экспертная информация. Методы получения и анализ. М.: Радио и связь, 1982. 184 с.

36. Лотов А. В. Введение в экономико-математическое моделирование. М.: Наука, 1984. 392 с.

37. Мартин Дж. Организация баз данных в вычислительных системах. М.: Мир, 1980. 664 с.

38. Математическая энциклопедия. Ред. коллегия: И. М. Виноградов и др. Т. 1. М.: Сов. энциклопедия, 1977. 1152 с.

39. Математика и кибернетика в экономике. Словарь-справочник/Под ред. Н. П. Федоренко. М.: Экономика, 1975. 700 с.

40. Миркин Б. Г. Проблема группового выбора. М.: Наука, 1974. 256 с.

41. Многокритериальные задачи принятия решений/Под ред. Д. Н. Гви-

шиани и С. В. Емельянова. М.: Машиностроение, 1978. 192 с.

42. Моделирование систем полуавтоматического управления космических кораблей/Под ред. А. И. Яковлева. М.: Машиностроение, 1986. 280 с.

43. Моисеев Н. Н. Математические задачи системного анализа. М.: Наука, 1981. 487 с.

44. Морозов В. П., Дымарский Л. С. Элементы теории управления ГАП. Л.: Машиностроение, Ленинградское отделение, 1984. 333 с.

45. Налимов В. В., Голикова Т. Н. Логические основы планирования эксперимента. М.: Металлургия, 1976. 128 с.

46. Нейлор Т. Машинные эксперименты с моделями экономических систем: Пер. с англ. М.: Мир, 1975. 500 с.

47. Нечипоренко В. И. Структурный анализ систем (эффективность и надежность). М.: Сов. радио, 1977. 216 с.

48. Николаев В. И., Брук В. М. Системотехника: методы и приложения. Л.: Машиностроение, Ленинградское отделение, 1985. 199 с.

49. Опойцев В. И. Равновесие и устойчивость в моделях коллективного поведения. М.: Наука, 1977. 248 с.

50. Орловский С. А. Проблемы принятия решений при нечеткой исходной информации. М.: Наука, 1981. 203 с.

51. Отнэс Р., Эноксон Л. Прикладной анализ временных рядов. Основные методы: Пер. с англ. М.: Мир, 1982. 428 с.

52. Подиновский В. В., Ногин В. Д. Парето-оптимальные решения многокритериальных задач. М.: Наука, 1982. 254 с.

53. Подиновский В. В., Гаврилов В. М. Оптимизация по последовательно применяемым критериям. М.: Сов. радио, 1975. 192 с.

54. Поляк Б. Т. Введение в оптимизацию. М.: Наука, 1983. 384 с.

55. Попов Е. П. Теория линейных систем автоматического регулирования и управления. М.: Наука, 1978. 256 с.

56. Пфанцagl И. Теория измерений. М.: Мир, 1976. 248 с.

57. Рабочая книга по прогнозированию. Ред. кол.: И. В. Бестужев-

Лада (отв. ред.). М.: Мысль, 1982. 430 с.

58. Саридис Дж. Самоорганизующиеся стохастические системы: Пер. с англ. М.: Наука, 1980. 400 с.

59. Современное состояние теории исследования операций/Под ред. Н. Н. Моисеева. М.: Наука, 1979. 464 с.

60. Советов Б. Я., Яковлев С. А. Моделирование систем. М.: Высшая школа, 1985. 271 с.

61. Стронгин Р. Г. Численные методы в многоэкстремальных задачах (информационно-статистические алгоритмы). М.: Наука, 1978. 239 с.

62. Супес П., Зинес Р. Основы теории измерений//Психологические измерения. М.: Мир, 1967. С. 9—110.

63. Трухаев Р. И. Модели принятия решений в условиях неопределенности. М.: Наука, 1981. 258 с.

64. Трухаев Р. И. Инфлюентный анализ и принятие решений (Детерминированный анализ). М.: Наука, 1984. 235 с.

65. Уемов А. И. Логические основы метода моделирования. М.: Мысль, 1971. 311 с.

66. Федоров В. В. Теория оптимального эксперимента. М.: Наука, 1971. 312 с.

67. Флейшман Б. С. Основы системологии. М.: Радио и связь, 1982. 368 с.

68. Хан Г., Шапиро С. С. Статистические модели в инженерных задачах: Пер. с англ. М.: Мир, 1981. 520 с.

69. Характеристики качества программного обеспечения/Б. Бозм, Дж. Браун, Х. Каспар и др.: Пер. с англ. М.: Мир, 1981. 208 с.

70. Цыпкин Я. З. Основы информационной теории идентификации. М.: Наука, 1984. 320 с.

71. Четвериков В. Н., Баканович Э. А., Меньков А. В. Вычислительная техника для статистического моделирования. М.: Сов. радио, 1978. 310 с.

72. Шапиро Д. И. Принятие решений в системах организационного управления: использование расплывчатых категорий. М.: Энергоатомиздат, 1983. 184 с.

73. Шеннон Р. Имитационное моделирование систем//Искусство и наука. М.: Мир, 1978. 418 с.

74. Шеффе Г. Дисперсионный анализ: Пер. с англ. М.: Наука, 1980. 512 с.

75. Шрейдер Ю. А. Равенство, сходство, порядок. М.: Наука, 1971. 256 с.

Предметный указатель

А

Автомат вероятностный 97

— конечный 97

Агрегат кусочно-линейный 98

Адекватность 76

Анализ проблемный 11

— системный 11

Б

Балльные оценки 203

Бинарное отношение на множестве элементов 173—177

— Геометрическая интерпретация 174

— Графическое представление 173, 174

— Свойства 174—177

Г

Гарантия вероятностная 32

Градиент функции отклика 146

Граф задач 66

— предпочтений 210

— целей 66

Д

Диаграмма принятия решения при исследовании эффективности и качества концептуальная 154, 155

Диапазонные оценки 205

Дисперсия адекватности 145

— воспроизводимости 138 — Проверка однородности 143, 144

З

Задача анализа неопределенности 55

— моделирования исходов операции 56, 57

— моделирования предпочтений 59

— моделирования цели операции 57—59

— сравнительной оценки вариантов 15

— структуризации исходной инфор-

мации 54

— формирования исходного множества стратегий 55, 56

Задачи лексикографической оптимизации 244

Задачи принятия решений в условиях неопределенности 59, 60, 266—311

— нестохастической 296—311

— поведенческой 311—321

— стохастической 266—295

Задачи принятия решений динамические 61

— по векторному показателю 230—265

— по скалярному показателю 222—230

— статические 61

Задачи принятия решения группового 59

— индивидуального 59

Задачи с показателем эффективности векторным 59

— скалярным 59

И

Игра матричная 313—317

— с нулевой суммой 313

— со строгим соперничеством 313

Иерархия принципов выработки решений при исследовании эффективности и качества 160—162

— уровней принятия решений 165, 166

Измерения показателей первичные 29

— производные (косвенные) 29

Интервал варьирования фактора 139

Информация об абсолютном превосходстве по важности 243—248

— об одинаковой важности всех показателей 241—243

— о недопустимости уменьшения меньших значений одинаково важных показателей сколь угодно значительным увеличением больших 248—250

— о равноценности или превосходстве в важности одних групп показателей над другими 250—252

— экспертная — формирование и формализация в форме нечетких суждений 213—220

Исследование концептуальное 47

— операциональное 47

Исследования имитационные — Требования к программному обеспечению 109—112

— по уровням вертикальной декомпозиции — Особенности выработки

решений 151—156

— эффективности системные детали-
ные 47 — обобщенные 47

К

Качества экспертов профессиональные — Оценка 191

Качество моделей — Оценка 127—129

Квазипорядок 175

Контраст определяющий 143 — обобщающий 143

Конфликт нестрогий 92

— строгий 92

Коэффициент аргументации эксперта — Оценка 191, 192

— пессимизма-оптимизма Гурвица 298

— регрессии — Значимость 145—147

Критерий Байеса 276

— допустимого гарантированного результата 39

— допустимой гарантии 39

— максимума апостериорной вероятности 276

— максимума правдоподобия 276

— минимаксного сожаления 298

— минимаксный 276

— наибольшего гарантированного результата 39

— наибольшего результата 39

— наибольшего среднего результата 39

— наибольшей вероятностной гарантии результата 39

— Неймана—Пирсона 276—279

— эффективности 37 — Классификация 40

Л

Лотереи с векторными характеристиками исхода операции 287

Лотерея простая 282

— сложная 282

Лицо, принимающее решение, безразличное к риску 283, 284

— несклонное к риску 283, 284

— склонное к риску 283, 284

М

Марковский случайный процесс управляемый 91

Метод выявления и измерения предпочтений комплексный 211—213

Методы моделирования — Классификация 76, 77

Методы определения коэффициентов относительной важности 207—211

— балльного оценивания 207, 208

— графоаналитический 209—211

— непосредственной численной оценки 207

— относительных частот рангов 208

— попарного сравнения с градациями 208

— последовательных сравнений 208

Методы планирования эксперимента 136, 137

Методы повышения качества оценивания эффективности активные 133, 134 — оптимального отбора 134 — пропорционального отбора 134 — косвенные 134—136 — значимой выборки 135 — сопутствующих переменных 135

— пассивные 131, 133 — Компенсации 133 — регрессионной выборки 132, 133

Методы понижения дисперсии 131 — принятия решений по скалярному показателю в условиях определенности 226—228

Методы решения задач с векторным показателем эффективности — многошаговые 256—265 — одношаговые 252—255

Методы экспертного оценивания — Классификация 187—189

Множества нечеткие — Элементы теории 220, 221

Множество наилучших стратегий лексикографически оптимальные 244

— уровня нечеткого множества 221

Модели имитационные автоматические 101

— Виды программной реализации 112—117

— диалоговые 101

Модели оценивания эффективности операций — Классификация 80—82

Моделирование дискретных случайных величин 103, 107—109

— непрерывных случайных величин 102—106

— случайных событий 102

Модель 76 — адекватная 74

— проблемной ситуации 52—54

— цели операции 57

Н

Неопределенность поведенческая 23

— природная 23

— целевая 23

Нетранзитивность в суждениях лица, принимающего решения 176

Нижняя цена игры 313

Носитель нечеткого множества 220

О

Область эксперимента 137

Омега-информация о предпочтениях — Особые случаи 241—252

Оператор сопряжения элементов 96

Операция — Понятие 11

Оптимизация лексикографическая 244, 245

Организация экспертного оценивания 189

Отношение — Понятие 173

— антирефлексивное 179

— асимметричное 175

— бинарное 25

— тернарное 25

— п-нарное 25

Отношения нечеткие 177—179

— Объединение отношений 178

— Рефлексивность 178

— Симметричность 178

— Транзитивность 178, 179

Отношения предпочтения 179, 180—184

— модельные 181

— нечеткие 180—184

Оценивание экспертное — Взаимосвязь этапов 190

Оценки векторные несравнимые по Парето 233, 234

Ошибка опыта 137

П

Пакеты программ моделирования

проблемно-ориентированные 115, 116

Переменная лингвистическая 22, 23

Показатели взаимонезависимые по предпочтению 233

— качественные 29

— количественные 29

— объективные 31

— субъективные 31

— технологические 111, 112

— функциональные 110, 111

— эксплуатационные 112

Показатель вероятностно-гарантированного результата 268

— вероятностной гарантии 267

— дисперсии результатов 268

— зависимый по предпочтению от остальных 233

— кучности 268

— среднего результата 267

— эффективности операции 30

Полуреплики 143

Помехоустойчивость системы 18

Порядок — Понятие 175

Правило замены 282
 Предпочтение 167—170 — Свойство направленности 168
 Принцип внешнего дополнения 46
 — вещественно-энергетического баланса 42
 — выбора решений 42
 — гомеостазиса 42
 — декомпозиции систем 46, 47
 Принцип моделирования операции в технике агрегирования 79
 — информационной достаточности 79
 — множественности моделей 80
 — осуществимости 79
 — параметризации 79
 — рационального факторного пространства 79, 80
 Принцип наибольшего гарантированного результата 312
 — перспективной активности 43
 — равновесия 312, 313
 — рефлексии 43
 — самообучения 40
 — свободы выбора решений 40
 — селекции 40
 Принципы концептуального уровня принятия решений 160, 161
 Проблема адекватности 29
 — принятия решения 51
 Программное обеспечение имитационной системы 117—121
 — автоматизации разработки и конструирования программ 120
 — информационно-справочное 119, 120
 — Модели и расчетные модули 120, 121
 — Модули-макроопределения 120
 — Подмодели 120
 — Программные модули 120.
 Проектирование системы внешнее 62
 Процесс выработки решений при исследовании эффективности и качества
 — Вертикальная декомпозиция процесса 148—151 — Требования к процессу 156—162
 — моделирования итерационный 75

Р

Работы 99
 Рандомизация 23
 Ранжирование 171 — Обработка и анализ 196—203
 Регрессия 137
 Результат вероятно-гарантированный 34
 Реплики 143

С

Свойства модели модификационные 127
 — целевые 127
 — эксплуатационные 127
 Свойства системы интегративные (эмерджентные) 10
 Семиотика 76
 Система динамическая 93—98
 — массового обслуживания 97
 — общая 93—98
 Система имитационная — Диалоговое обеспечение 123, 124
 — Информационное обеспечение 121—123
 — проблемно-ориентированная —
 Структура 117
 — Состав 116, 117
 — Специальное системное программное обеспечение 119—121
 Система кибернетическая 83, 84
 — обеспечения информацией — Оценка рациональных вариантов 69—71
 — поведенчески неоднозначная 84
 — со стохастической структурой 96
 Система числовая изоморфная 26
 — полная 26
 Системотехника 11
 Системы программирования для моделирования расширенные 114, 115
 — простые 10
 — сложные 10
 Ситуация игры 312
 — операции 13
 События 99
 Соотношения генерирующие 143
 Состояние системы — Понятие 13
 Способность системы — Понятие 19
 Способы выражения предпочтения 170—173
 — коэффициентом важности 172
 — лингвистическими переменными 172
 — Парное выражение 172
 — Парное сравнение 171
 — Сортировка 171
 Средства операции активные 12
 Стадии внешнего проектирования 63, 64
 — внутреннего проектирования 64
 Стандартизованный ранг 196
 Стратегии доминируемые 234
 — омега-эффективные 237—241
 — стохастические недоминируемые 280
 Стратегии эффективные — Графические способы выделения 235

- недоминируемые по Парето 234
- Способы выделения 236
- Стратегия максиминная** 313
- минимаксная 313
- рандомизированная 85 — Схема определения 275
- Строгий частичный порядок** 175
- Структура отношений субъектов системы** 86
- Схема выделения ядра отношения** 243
- методологических уровней анализа систем 44
- многоцелевой системы 14
- модели операции структурная 73, 74
- направленности свойств системы 17, 18
- обобщенной операции 64—69
- одноэтапной операции частная 88, 89
- операции общая 87
- построения функции эффективности методом половинного деления (случай трех показателей) 263—265
- процесса принятия решений на заданном уровне декомпозиции исследования — Общая схема 162
- решения задачи проектирования в лексикографической постановке 247, 248
- решения задач средствами имитационной системы технологическая 124—126
- сопряжения модели 94—96
- факторов, определяющих эффективность технических систем 20, 21
- формирования исходного множества стратегий 56

Т

- Типы задач принятия решений по скалярному показателю в условиях определенности** 226—228
- Толерантность** 176
- Точечные оценки** 204
- Транзитивное замыкание отношения** 176, 177 — нечеткого 179

У

- Управление** — Понятие 12
- Управляемость системы** 18, 19
- Уравнение регрессии** — Проверка адекватности 145
- Условия обстановки** 13
- Устойчивость системы** 18

Ф

- Фактор** — Понятие 17
- Факторы неуправляемые** 20
- определяющие эффективность тех-

нических систем — Классификация 21, 22 — Схема 20, 21

Форма сопряжения элементов каноническая 96

Формы функции соответствия и показателя эффективности 35—37

Функции оценочные лица, принимающего решения, с различной психологической доминантой оценочные 280

— принадлежности суждения типовые 214—219

— учитывающие отношение лица, принимающего решения, к риску оценочные 283—286

Функция выбора 184

— отклика 137

— эффективности 186, 187 — в условиях стохастической неопределенности — Построение 279—287 — скалярного результата 281

Ц

Цель 11, 12

Центр распорядительный основной 84

— эксперимента 139

Ч

Числа псевдослучайные 101, 102

Число степеней свободы 138

Ш

Шкала 26—29

Э

Эквивалент лотереи достоверный 283

Эквивалентность 175

Эксперимент активный 136

— дробный факторный 142, 143

— насыщенный 138

— пассивный 137

— полный факторный 139—142

— сверхнасыщенный 138

Элемент наилучший во множестве 184

— недоминируемый по отношению строгого предпочтения 184

Этап оценивания эффективности операции 72

Этапы жизненного цикла системы 62, 63

— исследования эффективности операции основные 49

— экспертного оценивания 190

Эффективность операции 13—15

Я

Ядро отношения 184

Языки программирования машинно-ориентированные 113, 114

— специализированные 114

— универсальные 113, 114