Ю.Ю. Громов, Н.А. Земской,
А.В. Лагутин, О.Г. Иванова,
В.М. Тютюнник

СИСТЕМНЫЙ АНАЛИЗ В ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЯХ

• ИЗДАТЕЛЬСТВО ТГТУ •

Министерство образования и науки Российской Федерации

ГОУ ВПО «ТАМБОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Ю.Ю. Громов, Н.А. Земской, А.В. Лагутин,

О.Г. Иванова, В.М. Тютюнник

Системный анализ в информационных технологиях

Допущено УМО вузов по университетскому политехническому образованию в качестве учебного пособия для студентов высших учебных заведений, обучающихся по специальности 071900 – «Информационные системы и технологии».

Издание второе, стереотипное

ТАМБОВ

Издательство ТГТУ

2007

УДК 004(075)
ББК ←81я73
С40

Рецензенты:

Доктор технических наук, профессор
А.А. Безбогов

Доктор физико-математических наук, профессор
А.И. Булгаков

С40 Системный анализ в информационных технологиях : учеб. пособие / Ю.Ю. Громов, Н.А. Земской, А.В. Лагутин, О.Г. Иванова, В.М. Тютюнник. – 2-е изд., стереотип. – Тамбов : Изд-во Тамб. гос. техн. ун-та, 2007. – 176 с. – 110 экз. – ISBN 978-5-8265-0629-5.

Рассмотрены принципы и особенности системного подхода, включая методологию и проблемы моделирования, многокритериальные и иерархические системы с большим количеством конкретных примеров, элементы теории игр.

Ученым советом ТГТУ рекомендовано для студентов высших учебных заведений, обучающихся по специальностям 230201 «Информационные системы и технологии», 090105 «Комплексное обеспечение информационной безопасности автоматизированных систем», и для студентов среднего профессионального образования, обучающихся по специальности 230105 «Программное обеспечение вычислительной техники и автоматизированных систем».

УДК 004(075)
ББК ←81я73

ISBN 978-5-8265-0629-5

Учебное издание

ГРОМОВ Юрий Юрьевич,
ЗЕМСКОЙ Николай Александрович,
ЛАГУТИН Андрей Владимирович,
ИВАНОВА Ольга Геннадьевна,
ТЮТЮННИК Вячеслав Михайлович

Системный анализ в информационных технологиях

Учебное пособие

Издание второе, стереотипное

Редактор З.Г. Чернова
Инженер по компьютерному макетированию Т.А. Сынкова

Подписано к печати 15.10.2007.
Формат 60 × 84/16. 10,23 усл. печ. л.
Тираж 110 экз. Заказ № 655

Издательско-полиграфический центр
Тамбовского государственного технического университета
392000, Тамбов, ул. Советская, 106, к. 14

ВВЕДЕНИЕ

В научных исследованиях и технических разработках, на производстве, в социальных областях мы постоянно сталкиваемся с совокупностями объектов, которые принято называть сложными системами. Их отличительные особенности – это многочисленные и разные по типу связи между отдельно существующими элементами системы и наличие у системы функции (назначения), которой нет у составляющих ее частей. Связи (взаимодействия) между элементами сложной системы будут характеризоваться определенным порядком, внутренними свойствами, направленностью на выполнение функции системы. Такие особенности данной конкретной системы назовем ее организацией.

На первый взгляд, каждая сложная система имеет свою, только ей присущую организацию. Однако более глубокое рассмотрение способно выделить, например, общее в иерархической системе команд ЭВМ и в управлении экономикой, в процессе проектирования технического объекта и в создании художественного произведения, в управлении научными исследованиями и в военной стратегии, которой пользовались еще древние греки.

Что это означает?

Очевидно то, что организации присущи некоторые общие закономерности, и она может изучаться отдельно, независимо от конкретного содержания и назначения сложной системы. Типичные абстрагированные свойства организации – это наличие между элементами отношений подчиненности, чередование и упругая упорядоченность процедур, согласование событий и целей, своевременная передача информации и управления, влияние на направленность процессов, приемы учета неопределенностей и многое другое. Также возможно говорить о применении в системе различных знаний и технических средств, роли и месте человека, моделировании и упрощении, централизованном использовании информации.

Каковы же цели изучения организации? Понять функционирование системы? Да. Но задачей более высокого уровня выступает создание нужной нам системы и управление ею. Ведущей операцией при этом является принятие решения, т.е. некоторый формализованный или неформализованный выбор, позволяющий достичь фиксированной частной цели или продвинуться в ее направлении. Принятие решения в сложной системе производится техническим средством или человеком и основано на сравнении и оценке вариантов действий. Постановка как основной, так и частных целей в системе также, обычно, подлежит анализу и исследованию. И опять же, главной процедурой при этом выступает принятие решения.

Как известно, изучение процедур принятия решения и связанной с этим организации системы составляет актуальную проблему создания и эксплуатации сложных систем. Подчеркнем, что все это может осуществляться на основе специально разработанных приемов, методик, типовых моделей организации системы и принятия решений. Законы организации таковы, что допускают вывод следствий, конкретизацию; возможно эффективное применение формализации, в первую очередь, математического знания.

Таким образом, мы имеем новую, но уже широко известную и чрезвычайно обширную в приложениях междисциплинарную ветвь науки – системный анализ. Ей и посвящена данная книга.

Системный анализ как научная дисциплина

В современном понимании системный анализ – это научная дисциплина, занимающаяся проблемами принятия решения в условиях анализа большого количества информации различной природы.

Из этого определения следует, что целью применения системного анализа к конкретной проблеме является повышение степени обоснованности принимаемого решения, расширение множества вариантов, среди которых производится выбор, с одновременным указанием способов отбрасывания тех из них, которые заведомо уступают другим. В максимально упрощенном виде системный анализ – это некая методика, позволяющая не упустить из рассмотрения важные стороны и связи изучаемого объекта, процесса, явления.

В системном анализе могут быть выделены методология, аппаратная реализация, опыт применения в различных областях знания и практики. Последовательно рассмотрим эти три его составляющие.

Методология в определенном смысле есть базовое начало системного анализа. Она включает определения используемых понятий, принципы системного подхода, а также постановку и общую характеристику основных проблем организации системных исследований.

Определения в методологии обычно даются на словесно интуитивном уровне и, как правило, обладают свойством конструктивности. Общепринятые определения создают язык данной науки, влияют на научное мышление. В системном анализе процесс выработки единых определений не закончен и весьма актуален в связи с междисциплинарным характером исследований. Принципы системного подхода – это некоторые положения общего характера, являющиеся обобщением опыта работы человека со сложными системами. Часто их считают ядром методологии. Постановка и характеристика проблем системных исследований (например, целенаправленная структуризация, оптимальное чередование исполнительских и управленческих операций, задача о системе с плавающей границей между ней и внешней средой и др.) составляют в настоящее время наименее освещенную часть методологии системного анализа.

Под аппаратной реализацией будем понимать стандартные приемы моделирования принятия решения в сложной системе и общие способы работы с этими моделями. Модель строится в виде связных множеств (в простейшем случае – цепочек) отдельных процедур. Системный анализ исследует как организацию таких множеств, так и вид отдельных процедур, которые максимально приспосабливают для принятия согласующих и управленческих решений в сложной системе.

Модель принятия решения чаще всего изображается в виде схемы с ячейками, связями между ячейками и логическими переходами. Ячейки содержат конкретные действия-процедуры, которые могут иметь весьма разнообразный характер. Совместное изучение процедур и их организации вытекает из того, что без учета содержания и особенностей ячеек создание схем оказывается невозможным. Эти схемы определяют стратегию принятия решения в сложной системе. Именно с проработки связного множества основных процедур принято начинать решение конкретной прикладной задачи.

Говоря об отдельных процедурах, укажем, что основной из их многочисленных классификаций является деление на формализуемые и неформализуемые процедуры (операции). Важным тезисом является то, что в отличие от большинства научных дисциплин, стремящихся к формализации, системный анализ допускает, что отнюдь не все следует систематизировать и дополнять строгими правилами действий.

Утверждается, что в определенных ситуациях неформализуемые решения, принимаемые человеком, являются более предпочтительными, и активность человека внутри сложной системы может определять успех работы с ней. Таким образом, системный анализ рассматривает совместно, в совокупности, формализуемые и неформализуемые процедуры, и одной из его задач является определение их оптимального соотношения.

Формализуемые стороны отдельных операций, как правило, лежат в области прикладной математики и использования средств вычислительной техники. В ряде случаев математическими методами исследуется связное множество процедур, а иногда производится и само моделирование принятия решения. Все это позволяет говорить о математической основе системного анализа. Высокий уровень абстрагирования в математике приводит, в частности, к тому, что и фундаментально-прикладные, и даже чисто вычислительные исследования, как правило, выполняются безотносительно к тому, как их результаты будут использоваться дальше. Скажем, вопросы удобной записи и передачи данных, оценки количества информации и его уменьшения, передачи управления в другую задачу или человеку традиционно считаются лежащими вне математики, но представляют существенные званья системного анализа.

Наиболее близки к системной постановке вопросов такие области прикладной математики, как исследование операций и системное программирование. Изложение системного анализа в данной книге во многом основано именно на том, как рассматриваются и решаются системные задачи в этих областях. Следует отметить, что в системном анализе существует и другое направление, берущее свое начало в исследовании сложных, многоаспектных проблем социологии, философии, других гуманитарных наук. Эти два базовых начала практически едины в области методологии, но заметно расходятся в методах (аппаратной реализации) и тем более – в приложениях. Для первого из них характерны насыщение формализованными операциями, использование ЭВМ, математизация знания, низкая степень неопределенности в сочетании с конкретностью исходных данных и целей, относительно жесткая внутренняя структура системы.

Третья часть системного анализа – опыт его применения в различных областях – чрезвычайно обширна по содержанию. Важнейшими разделами являются научно-технические разработки и различные задачи экономики. Перечень лишь тех ветвей науки, где ссылки на системность исследований, анализа, подхода являются обычными, включает биологию, экологию, военное дело, психологию, социологию, медицину, управление государством и регионом, обучение и тренировку, выработку научного мировоззрения и многое другое. В рамках одного учебного пособия не представляется возможным даже прокомментировать использование системного анализа в этих разделах.

В данной книге в области применения авторы ориентируются на проектирование сложного технического объекта, а также на создание и совершенствование современных автоматизированных систем. Такой выбор определяется особой актуальностью этих задач в настоящий момент. Одновременно эти проблемы представляют хороший полигон для демонстрации системного анализа в действии и иллюстрации практически всех его положений и формальных структур. Наконец, и предназначение пособия для физико-математических специальностей университетов приводит к необходимости учитывать, что основная часть выпускников по окончании вуза будет иметь дело именно с этими проблемами.

Вычислительная техника в системном анализе

В настоящее время развитие системного анализа, прежде всего, характеризуется осмыслением широчайшего проникновения вычислительной техники в процесс принятия решения и сложной системе. Программные и технические средства различного уровня и масштаба выполняют значительное число отдельных процедур и начинают эффективно использоваться для составления наборов процедур и контроля за ходом решения задачи в целом. Особое место при анализе и принятии решения занимают такие сравнительно новые объекты, как информационная база (банки данных), диалоговые системы, имитационное моделирование. Эти объекты, обычно воспринимаемые как части автоматизированных систем или как специальные, использующие ЭВМ методы исследования, могут и должны рассматриваться и в качестве важных понятий системного анализа. Они отражают существенные и достаточно абстрактные стороны современного состояния аппаратной реализации системных исследований. С точки зрения системного анализа это некоторые классы операций, обладающие внутренней структурой, универсальностью использования и другими особенностями.

Ведущими среди этих объектов представляются диалоговые системы. Напомним, что их суть заключается в чередовании формализованных (ЭВМ) и неформализованных (человек) процедур и обычно характеризуется специальными средствами для организации диалога с ЭВМ и высокой оперативностью процедур, выполняемых ЭВМ и человеком.

Удачно организованные диалоговые системы эффективно усиливают возможности как машины, так и человеческого мозга и, в частности, позволяют решать задачи, недоступные только ЭВМ или только человеку. Диалог в виде вопросов и ответов присутствует в любой информационной базе, а также является удобным видом работы с имитационными моделями. Не останавливаясь здесь на других особенностях банков данных и имитации поведения системы, подчеркнем лишь общую основу этих понятий – взаимодействие человека и вычислительной техники.

Можно выделить три стороны этого взаимодействия, одна из которых уже затронута, – это партнерство в выполнении операций, названное диалогом с ЭВМ. Вторая сторона – человек является создателем программных средств, программного продукта, без которых вычислительная техника мертва. Многообразие программ и уровни их сложности даже трудно себе представить. Их спектр простирается от программы решения квадратного уравнения или программы засылки информации в данную ячейку памяти до программы расчета вибрации корпуса ракеты и управления работой вычислительной сети, охватывающей несколько стран. Два последних программных средства, во-первых, в качестве внутренних элементов насчитывают сотни и даже тысячи более простых программ, а во-вторых, способны организовать значительное число вариантов их работы. Для системного анализа наиболее существенно то, что программы, пакеты программ выступают как средство исследования сложной системы, средство, готовящее решение в ней. Применение отдельного программного средства является элементарной процедурой системного анализа.

Третья сторона взаимодействия человека и ЭВМ заключается в том, что именно человек оценивает решение или другую информацию, полученную с помощью вычислительной техники, и дает указание на использование результатов исследования на практике. В литературе по системным исследованиям привился термин «лицо, принимающее решение» (ЛПР). Именно на ЛПР, ответственном за всю систему или ее часть, замыкается выполнение совокупности операционных процедур. Роль ЛПР, рамки его действий, отделение или неотделение от исследователей системы составляют проблемы, которые в общем виде также относятся к области системного анализа.

Разнообразная считающая, управляющая, хранящая, преобразующая, советующая, изображающая и другая вычислительная техника является как неотъемлемой частью самих сложных систем, так и исследования. Владение аппаратом системного анализа невозможно без умения определять тактику и стратегию использования ЭВМ, баз данных вычислительных сетей. В конкретных же проблемах это умение часто вообще определяет успех системного исследования.

Смежные области

В настоящее время в научной литературе целый ряд терминов, имеющих отношение к исследованию сложных систем, употребляется в разных, нередко несогласованных или пересекающихся смыслах. Поэтому представляется необходимым и полезным дать определения ряда смежных понятий, которые, хотя и основаны на разнообразных литературных источниках, в ряде случаев переработаны с учетом мнения авторов данного пособия.

Термин «теория сложных систем» (а также «общая теория систем» или просто «теория систем») отнесем к всевозможным аспектам исследования систем, а не только к проблеме принятия в ней решения, как это имеет место в системном анализе. Таким образом, в этом варианте системный анализ составляет существенную, важную в прикладном отношении часть теории систем. Отметим, что в литературе широко встречается как смешение этих понятий, так. и попытки отделить их друг от друга.

Дадим самую краткую характеристику современного состояния теории систем. Достаточно сформировавшейся является математическая теория сложных систем (работы Р. Калмана [1], М. Месаровича [2, 3], У. Портера [4] и др.). Значительными достижениями по теории систем выступают теория организации русского экономиста А.А. Богданова и общая теория систем А. Берталакфи. В частности, биолог Берталакфи выделил направления теории систем, а в ее математической части дал содержательную классификацию. Однако в целом рассматривать ее как самостоятельную ветвь науки можно лишь с рядом оговорок.

Упомянутая интерпретация теории систем приводит и к тому, что ее частью следует считать кибернетику, которая традиционно определяется как наука об управлении и преобразовании информации. Ведь нетривиальные результаты этой области знания относятся именно к сложным системам. Понятие управления близко, но не совпадает с принятием решения. Условная граница между кибернетикой и системным анализом состоит еще и в том, что первая изучает отдельные и обычно строго формализованные процессы, а системный анализ – совокупность процессов и процедур. Стоит заметить, что системный анализ перенял у кибернетики значительное количество терминов. Упомянем такие, как входы и выходы в системе, модули, потоки информации, структурные схемы.

Одним из наиболее сложных для обсуждения является термин «системотехника». Он определяется и как применение теории систем и системного анализа к области техники [5, 6], и как применение техники, в первую очередь, вычислительной, при исследовании сложных систем [6], и еще более узко – как использование системного анализа для проектирования ЭВМ и сетей ЭВМ, а также создания их программного обеспечения (таково содержание квалификации «инженер-системотехник»). Дать определение системотехники, устраивающее хотя бы основной круг авторов, использующих этот термин, в настоящее время не представляется возможным.

Сравнительно новое понятие – информатика – чаще всего понимается как исследование проблем хранения, использования и преобразования информации при помощи средств вычислительной техники. Информатика имеет технический, программный, математический и системный аспекты. Эта ветвь знания является одной из основ при проведении системного анализа при помощи ЭВМ. Следует иметь в виду, что ряд авторов распространяют новый термин не только на информационные задачи, но и на все проблемы, связанные с использованием ЭВМ.

Весьма близким к термину «системный анализ» является понятие исследования операций. Однако мы будем избегать этого сочетания в упомянутом общем смысле в связи с тем, что в советской научной литературе оно традиционно обозначает достаточно обособленную математическую дисциплину, охватывающую исследование математических моделей для выбора величин (чисел, функций), оптимизирующих заданную математическую конструкцию (критерий). Системный анализ может сводиться к решению ряда задач исследования операций, но обладает свойствами, не охватываемыми этой дисциплиной. Здесь же отметим, что в литературе США термин «исследование операций» не является чисто математическим и приближается к термину «системный анализ» [5, 6]. Правда, более поздние работы различают эти термины в том смысле, что под исследованием операций понимаются системные исследования, ориентированные на количественное описание.

Рассмотрим системные понятия, не являющиеся научными направлениями. Системным подходом, понимаемым в данной книге как часть методологии системного анализа, называется применение ряда методологических положений (принципов) общего характера к исследованию систем. Известно около двух десятков таких принципов, связанных с необходимостью изучать систему комплексно, в ее разумной полноте, связности, организованности.

Прилагательное «системный» в применении к целому ряду понятий (метод, исследование, особенность, взгляд, модель и т.д.) означает учет в этих понятиях принципов системного подхода. Так, системные исследования – это акцентирование внимания на сложности конечной цели, единстве и расчлененности процесса исследования, наличии его внутренней структуры и т.д. Широкое и свободное употребление слова «системный», возможно, и неудобно в науке, но оно отражает то положение, когда системностью интересуются представители самых далеких друг от друга наук и часто с различных точек зрения. Корень слова «система» выдвинулся на одно из первых мест по частоте употребления даже в газетных текстах.

Возвращаясь к системному анализу, укажем, что он взаимодействует со всеми перечисленными понятиями, а наиболее тесно связан с теорией систем. Системный анализ в значительной мере опирается на такие ее части, как структуризация, иерархия в системе, законы протекания процессов в ней, связь системы с «не системой» (внешней средой), эволюция системы, в том числе самоорганизация. Здесь же полезно назвать и специфические части (разделы) самого системного анализа – это целеопределение, выделение действий и приемы работы с ними, сочетание формализованных и неформализованных процедур, действия ЛПР, системные вопросы информатики. Широкая опора системного анализа на исследование операций (которая имеет место, по крайней мере, в технике и экономике) приводит к таким его математизированным разделам, как постановка задач принятия решения, описание множества альтернатив, исследование многокритериальных задач, методы решения задач оптимизации, обработка экспертных оценок, работа с макромоделями системы. Указанные разделы могут быть отнесены и к исследованию операций.

Настоящее пособие охватывает те аспекты из перечисленных, которые, на взгляд авторов, являются наиболее существенными в обучении и способны составить учебник, обладающий определенной целостностью и взаимосвязью.

Контрольные вопросы

1. Перечислите основные свойства организации.
2. Какая операция является ведущей при создании нужной нам системы и управлению ей?
3. Что такое системный анализ?
4. Перечислите составляющие системного анализа.
5. Что такое принципы системного подхода?
6. В чем заключается связь системного анализа с теорией систем?

Глава 1
СИСТЕМНЫЙ ПОДХОД

1.1. ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

1.1.1. Элементы, связи, система

Введем достаточно обширный набор понятий, связанных с современным использованием слова «система». Большинство из этих понятий и ряд операций с ними запишем также в символьном виде, близком к употреблению сходных математических терминов.

Элементом назовем некоторый объект (материальный, энергетический, информационный), обладающий рядом важных для нас свойств, но внутреннее строение (содержание) которого безотносительно к цели рассмотрения.

Обозначим элементы через M, а всю их рассматриваемую (возможную) совокупность – через $\{M\}$. Принадлежность элемента совокупности принято записывать $M \in \{M\}$.

Связью назовем важный для целей рассмотрения обмен между элементами веществом, энергией, информацией.

Единичным актом связи выступает воздействие. Обозначая все воздействия элемента M_1 на элемент M_2 через x_{12}, а элемента M_2 на M_1 – через x_{21}, можно изобразить связь графически (рис. 1.1).

Системой назовем совокупность элементов, обладающую следующими признаками:

а) связями, которые позволяют посредством переходов по ним от элемента к элементу соединить два любых элемента совокупности;

б) свойством (назначением, функцией), отличным от свойств отдельных элементов совокупности.

Назовем признак а) связностью системы, б) – ее функцией. Применяя так называемое «кортежное» (т.е. «последовательность в виде перечисления») определение системы, можно записать:

$$\sum : \{\{M\}, \{x\}, F\}, \tag{1.1}$$

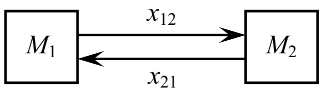

Рис. 1.1. Связь двух элементов

где Σ – система; $\{M\}$ – совокупность элементов в ней; $\{x\}$ – совокупность связей; F – функция (новое свойство) системы.

Будем рассматривать запись (1.1) как наиболее простое описание содержания системы. Существуют формы записи, включающие более 10 членов кортежной последовательности, соответствующих различным свойствам системы [5, 6].

Практически любой объект с определенной точки зрения может рассматриваться как система. Важно отдавать себе отчет – полезен ли такой взгляд или разумней считать данный объект элементом. Так, системой можно считать радиотехническую плату, преобразующую входной сигнал в выходной. Для специалиста по элементной базе системой будет слюдяной конденсатор на плате, а для геолога – и сама слюда, имеющая достаточно сложное строение.

Большой системой назовем систему, включающую значительное число однотипных элементов и однотипных связей.

Сложной назовем систему, состоящую из элементов разных типов и обладающую разнородными связями между ними.

Часто сложной системой считают только ту, которая является и большой. Разнородность элементов можно подчеркнуть записью

$$\{M\} : \{\{M^I\}, \{M^{II}\}, ..., \{M^R\}\}. \tag{1.2}$$

Допустима также некортежная запись $\underset{r=I}{\overset{R}{U}}\{M^r\}$. Аналогично может быть записана и разнородность связей.

Большой, но не сложной с точки зрения механики системой является собранная из стержней стрела крана или, например, труба газопровода. Элементами последней будут ее участки между сварными швами или опорами. Для расчетов на прогиб элементами газопровода скорее всего будут считаться относительно небольшие (порядка метра) участки трубы. Так поступают в известном методе конечных элементов. Связь в данном случае имеет силовой (энергетический) характер – каждый элемент действует на соседний.

Различие между системой, большой системой и сложной системой условно. Так, корпуса ракет или судов, которые, на первый взгляд, однородны, обычно относят к сложной системе – из-за наличия переборок разного вида, усилителей, слоистой конструкции. Типичными примерами сложных систем являются судно, самолет, ракета, системы управления ими, электронно-вычислительная машина, транспортная сеть города и многое другое.

В настоящее время важным классом сложных систем выступают так называемые автоматизированные системы. Слово «автоматизированный» указывает на участие человека, использование его активности внутри системы при сохранении значительной роли технических средств. Так, цех, участок, сборка могут быть как автоматизированными, так и автоматическими («цех-автомат»). Для сложной системы автоматизированный режим считается более предпочтительным. Например, посадка самолета выполняется при участии человека, а автопилот обычно используется лишь на относительно простых движениях. Также типична ситуация, когда решение, выработанное техническими средствами, утверждается к исполнению человеком.

Итак, автоматизированной системой называется сложная система с определяющей ролью элементов двух типов в виде:

1) технических средств;
2) действий человека.

Ее символьная запись [сравни с (1.1) и (1.2)]:

$$\sum{}^{A}: \{\{M^{\text{т}}\}, \{M^{\text{ч}}\}, \{M'\}, \{x\}, F\}, \qquad (1.3)$$

где $M^{\text{т}}$ – технические средства, в первую очередь, ЭВМ; $M^{\text{ч}}$ – решения и другая активность человека; M' – остальные элементы в системе.

В совокупности $\{x\}$ в этом случае могут быть выделены связи между человеком и техникой $\{x^{\text{т–ч}}\}$.

1.1.2. Структура и иерархия

Структурой системы называется ее расчленение на группы элементов с указанием связей между ними, неизменное на все время рассмотрения и дающее представление о системе в целом.

Указанное расчленение может иметь материальную (вещественную), функциональную, алгоритмическую и другие основы. Группы элементов в структуре обычно выделяются по принципу простых или относительно более слабых связей между элементами разных групп. Структуру системы удобно изображать в виде графической схемы, состоящей из ячеек (групп) и соединяющих их линий (связей). Такие схемы называются структурными.

Для символьной записи структуры введем вместо совокупности элементов $\{M\}$ совокупность групп элементов $\{\widehat{M}\}$ и совокупность связей между этими группами $\{\widehat{x}\}$. Тогда структура системы может быть записана как

$$\sum\sum: \{\{\widehat{M}\}, \{\widehat{x}\}\}. \qquad (1.4)$$

Структуру (1.4) можно получить из (1.1) объединением элементов в группы. Отметим, что функция (назначение) F системы в (1.4) опущена, поскольку структура может быть в определенной степени безотносительна к ней. Отметим также, что элементы M из какой-либо группы M будут, как правило, неоднородны [см. (1.2)]

Приведем примеры структур. Вещественная структура сборного моста состоит из его отдельных, собираемых на месте секций. Грубая структурная схема такой системы укажет только эти секции и порядок их соединения. Последнее и есть связи, которые здесь носят силовой характер. Пример функциональной структуры – это деление двигателя внутреннего сгорания на системы питания, смазки, охлаждения, передачи силового момента и др. Пример системы, где вещественные и функциональные структуры слиты, – это отделы проектного института, занимающиеся разными сторонами одной и той же проблемы. Типичной алгоритмической структурой будет алгоритм (схема) программного средства, указывающая последовательность действий. Также алгоритмической структурой будет инструкция, определяющая действия при отыскании неисправности технического объекта.

Примерами структур других типов являются календарь (временная структура) или деление книги на главы. Ситуация с книгой интересна тем, что здесь основа деления может быть информационной (в научной литературе), вещественной (для типографии глава – это количество бумаги и рабочего труда) или более сложной, например, основанной на наборе эстетических воздействий на читателя (для художественной литературы).

Структура системы может быть охарактеризована по имеющимся в ней (или преобладающим) типам связей. Простейшими из них являются последовательное, параллельное соединение элементов и обратная связь (рис. 1.2).

Поясним понятие обратной связи. Оно означает, что результат функционирования элемента влияет на поступающие на него воздействия. Как правило, обратная связь выступает важным регулятором в системе. Крайне редко встречается система без того или иного вида обратной связи.

Близким к понятию структуры является термин «декомпозиция».

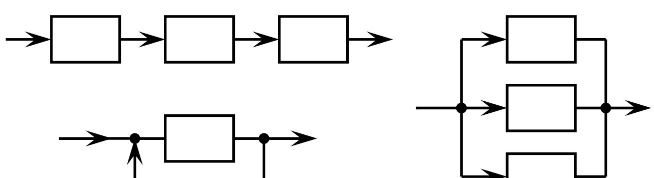

Рис. 1.2. Простейшие типы связей

Декомпозицией называется деление системы на части, удобное для каких-либо операций с этой системой. Примерами декомпозиции будут: рассмотрение физического явления или математическое описание отдельно для данной части системы; разделение объекта на отдельно проектируемые части, зоны обслуживания; другие частично или полностью независимые манипуляции с частями системы.

Важнейшим стимулом и сутью декомпозиции является упрощение системы, слишком сложной для рассмотрения целиком. Такое упрощение может:

а) фактически приводить к замене системы на некоторую другую, в каком-то смысле соответствующую исходной; как правило, это делается вводом гипотез об отбрасывании или ослаблении отдельных связей в системе;

б) полностью соответствовать исходной системе и при этом облегчать работу с ней; такая декомпозиция, называемая строгой, требует специальных процедур согласования и координации рассмотрения частей.

Иерархией назовем структуру с наличием подчиненности, т.е. неравноправных связей между элементами, когда воздействия в одном из направлений оказывают гораздо большее влияние на элемент, чем в другом.

Типичная иерархическая связь с воздействиями вида «информация» и «управление» изображена на рис. 1.3. Естественно, здесь доминирует элемент M_1.

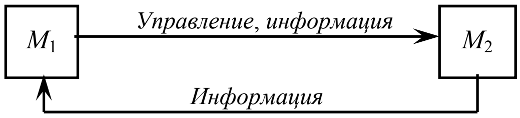

Рис. 1.3. Пример типичной иерархической связи

Виды иерархических структур разнообразны. Среди них встречаются такие экзотические, как кольцевые (первый элемент доминирует над вторым, второй – над третьим и т.д., но последний – над первым) или меняющие направление доминирования. Но основных, важных для практики иерархических структур, всего две – древовидная (веерная) и ромбовидная (рис. 1.4).

Древовидная структура наиболее проста для анализа и реализации. В ней почти всегда удобно выделять так называемые иерархические уровни – группы элементов, находящиеся на одинаковом (по числу промежуточных элементов) удалении от верхнего (главенствующего) элемента. Примеры таких структур в искусственных и живых системах чрезвычайно многочисленны:

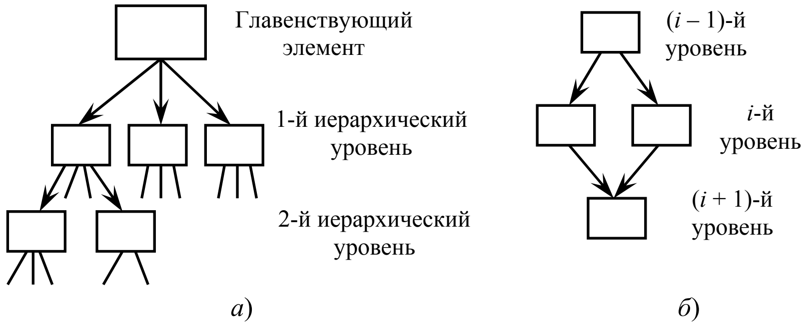

Рис. 1.4. Примеры иерархических структур:
а – древовидная (веерная); *б* – ромбовидная

а) цепочка «министерство–главк–завод–цех–бригада–звено»;

б) задача проектирования технического объекта – от его основных характеристик (верхний уровень) через проектирование основных частей, функциональных систем, групп агрегатов, механизмов до уровня отдельных деталей;

в) иерархия целей в задаче автоматизированного производства – от цели участка, состоящей в максимальном выпуске продукции, до программного обеспечения отдельной операции на станке (цель – операция);

г) в живой природе – иерархия по признаку управляемости процессов в организме, иерархия в стаде и др.

Ромбовидная структура ведет к двойной (иногда и более) подчиненности, отчетности, принадлежности нижнего элемента. В технике – это участие данного элемента в работе более чем одного узла, блока, использование одних и тех же данных или результатов измерений в разных задачах. Пример из проектирования (см. рис. 1.4, *б*): (i – 1)-й уровень – системы пожаротушения корабля в целом; i-й уровень – основной и дублирующей систем пожаротушения; (i + 1)-й уровень – насоса, который будет независимо поставлен и в основную, и в дублирующую системы.

Любая иерархия, в принципе, сужает возможности и, особенно, гибкость системы. Элементы нижнего уровня сковываются доминированием сверху, они способны влиять на это доминирование (управление) лишь частично и, как правило, с задержкой. Однако введение иерархии резко упрощает создание и функционирование системы, и поэтому ее можно считать вынужденным, но необходимым приемом рассмотрения сложных технических систем. Недаром та или иная степень иерархии наблюдается в подавляющем большинстве естественных систем.

Отрицательные последствия введения иерархии во многом могут быть преодолены предоставлением отдельным элементам возможности реагировать на часть воздействий без жесткой регламентации сверху.

1.1.3. Модульное строение системы и информация

Перейдем к введению следующей важной группы понятий. До сих пор мы называли связью воздействия одного элемента (или группы элементов) на другой элемент (группу). Ничто не мешает распространить понятие связи и на взаимодействие системы с «несистемой», которую обычно называют внешней средой. Следующий шаг в исследовании связей в системе состоит в выделении для данного элемента:

а) всех тех воздействий, которые он испытывает со стороны других элементов и «несистемы»;

б) воздействий, которые он оказывает на другие элементы и «несистему».

Первую группу воздействий принято называть входами (воздействия «на элемент»), а вторую – выходами (воздействия «от элемента»).

Как правило, выходы элемента определяются входами и его внутренним строением. В этом смысле говорят, что выход есть функция от входа и самого элемента.

Язык входов и выходов переносится на произвольную совокупность элементов, включая и всю систему целиком. И здесь можно говорить обо всех входящих и выходящих воздействиях. Это не просто удобный, но весьма плодотворный подход к рассмотрению системы, поскольку, характеризуя группу элементов только входами и выходами, можно получить возможность оперировать этой частью системы, не вникая, как связаны и взаимодействуют между собой ее элементы. Таким образом уйти от детализации в описании при сохранении основных особенностей системы.

Группа элементов системы, описываемая только своими входами и выходами и обладающая определенной цельностью, называется модулем.

Система может представляться набором модулей и сама рассматриваться как модуль. Модульное построение системы, как правило, определяет ее декомпозицию. Нередко оно определяет и структуру. Однако значение понятия модуля в системном анализе и смежных с ним дисциплинах еще шире. Деление системы на модули – это удобный и наиболее распространенный прием работы с искусственными системами, включая их создание (проектирование), проверку, настройку, усовершенствование. Именно модульное строение системы в сочетании с принципом введения все более крупных модулей при сохранении обозримого объема входов и выходов позволяет рассматривать в принципе сколь угодно сложные системы. Примерами реализации этого положения на практике являются создание из сотен тысяч элементов (материальных, информационных, энергетических) вычислительных машин четвертого поколения, а также создание информационных систем и вычислительных сетей, охватывающих целый ряд стран, включая их многоуровневое программное обеспечение. Разработка таких систем обычно идет «сверху», с продумыванием назначения, входов и выходов модулей верхнего уровня, и далее спускается вниз, все в большей степени детализируя систему.

Схематическое изображение модуля приведено на рис. 1.5.

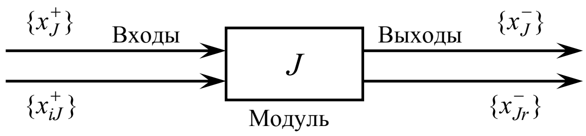

Рис. 1.5. Модуль

Здесь x_J^+ – внешние (от «не-системы») воздействия на элементы модуля J; связи от других элементов системы на элементы модуля J; x_{iJ}^+ – связи (воздействия) от элементов модуля J на другие элементы системы; x_{Jr}^- – связи (воздействия) от элементов модуля на не-систему, их также можно рассматривать как часть F_J функции системы F, которая реализуется модулем J. В этом случае имеем $\{x_J^-\} = F_J$. Теперь можно записать модуль в виде преобразования

$$\left(\{x_J^+\},\ \{x_{iJ}^+\},\ J\right) \to \left(\{x_J^-\},\ \{x_{Jr}^-\}\right).$$

Заметим также, что понятие модуля близко к концепции «черного ящика» в кибернетике – так называют объект, в котором известна только зависимость выходов от входов. Однако в отличие от такой крайней ситуации здесь, при исследовании сложных систем, обычно можно проанализировать, что же происходит внутри модуля, но удобно не делать этого на определенной стадии рассмотрения.

Важность понятий модуля, входа, выхода подчеркивается и большим количеством их синонимов в различных разделах науки и техники. Так, например, синонимом модуля являются «агрегат», «блок», «узел», «механизм» в технике; «подпрограмма», «программный модуль», «логический блок» – в программировании; «подразделение», «комиссия» – в организации и управлении. Типичными входами и выходами являются пары «сигнал – отклик», «воздействие (раздражение) – реакция», «запрос – ответ», «аргумент – решение», или, более широко, «информация – принятие решения», «управление – движение» и др.

Перейдем к анализу понятия информации.

Выше неоднократно подчеркивалось наличие трех видов связей (воздействий): материальных, энергетических, информационных. Для сложных искусственных систем следует особо выделить информационные. Во-первых, эти связи часто являются преобладающими, определяющими в системе; во-вторых, они как правило, сопровождают и два остальных вида – вещественные и энергетические воздействия в искусственной системе фиксируются и в качестве информации. Так, в гибкой производственной системе информационный характер носит основной системный элемент – комплексы управляющих программ и целый набор сопутствующих им программных средств. При этом в управляющие программы поступают сведения о материальных и силовых воздействиях на обрабатываемую деталь.

В целом информация в системе выступает как собирательный термин для обозначения всех нужных сведений.

Информация может изучаться с точки зрения ее получения, хранения, передачи, преобразования, свертки. В 1940-х гг. было введено универсальное количественное описание информации через ее влияние на вероятность того события, в котором она нужна. Однако на практике используется ряд других, более узких способов ее количественной оценки, которые могут быть как основаны, так и не основаны на универсальном описании, – через число сообщений, число операторов, файлов в программных средствах, объем информации в знаках или двоичных кодах и др.

В сложных системах особенно важна передача информации. Она может быть предметом специального рассмотрения; в этом случае выделяют потоки информации, которым обычно сопоставляют схемы типа структурных. В них указываются источники и потребители информации, направление передачи, возможно указание объема, формы представления и других ее характеристик. Такие схемы принято называть информационным графом или информационной структурой системы. Они могут в значительной степени соответствовать тому понятию структуры, которое мы употребляем в данной работе.

Информационный граф может быть исследован с целью минимизации потоков или сокращения их длины, с точки зрения отсутствия или наличия дублирования путей передачи информации и т.д.

Понятие информации обладает высокой степенью универсальности. В широком смысле функционирование системы можно трактовать как преобразование входной информации в выходную. Такая точка зрения особенно полезна при изучении принятия решений в системе, т.е. в системном анализе.

1.1.4. Процессы в системе

Введем понятия состояния и процесса в системе. Для этого сначала рассмотрим некоторый выделенный элемент. Что с ним может произойти? Он может быть помещен в систему, исключен из системы, перемещен в ней с одного места на другое. Кроме того, могут быть изменены его связи. Все эти ситуации относятся к изменению структуры системы.

Но возможны преобразования другого рода. Любой элемент обладает рядом свойств, характеристик, которые тоже могут меняться в процессе рассмотрения системы. Вследствие этого могут измениться свойства, характеристики группы элементов, модуля и системы целиком.

Зафиксируем все значения характеристик в системе, важных для целей рассмотрения. Такую ситуацию назовем состоянием системы.

Пусть теперь хотя бы одна такая характеристика изменилась. Это будет новое состояние системы. Аналогично можно рассмотреть третье и т.д. состояния, т.е. их набор. Но набор состояний – это еще не процесс. Пусть выбран некоторый физический параметр (чаще всего время) – такой, что различные состояния соответствуют разным его значениям.

Процессом назовем набор состояний системы, соответствующий упорядоченному непрерывному или дискретному изменению некоторого параметра, определяющего характеристики (свойства) системы.

Для пояснения определения сразу же приведем пример.

Состояние робота-манипулятора будем характеризовать положением его основного рабочего органа – схвата. Сделаем серию фотографий манипулятора, на которых схват будет находиться в разных точках пространства. Будет ли этот набор фотоснимков характеризовать какой-нибудь процесс? Без дополнительной информации это неизвестно. Если это последовательные во времени положения схвата, то – да (параметр процесса – время). Если же снимки сделаны наугад или перемешаны, то соответствующий набор состояний не будет процессом.

Процесс движения (изменения) системы во времени называют динамикой системы. Параметрами процесса могут также выступать температура, давление, другие физические величины. В качестве параметра иногда выступают линейные и угловые координаты (пример: процесс изменения атмосферы с высотой) и даже скорости. Однако более типично отнесение этих величин к характеристикам системы, которые сами зависят, например, от времени.

Для символьной записи процесса введем многомерную (по числу интересующих нас характеристик) величину y, описывающую их конкретные значения. Все множество этих возможных величин обозначим через Y: $y \in Y$. Введем параметр процесса t, множество его значений T и опишем y как функцию от этого параметра: $y = y(t)$. Тогда процесс S_{t_0t} есть некоторое правило перехода от ситуации со значением параметра t_0 к ситуации со значением $t > t_0$ через все его промежуточные непрерывные или дискретные значения:

$$S_{t_0t}(y(t_0)) = y(t),\ \ y \in Y,\ \ t \in T\ . \tag{1.5}$$

Этому же процессу будет соответствовать отображение множеств

$$T \times Y \to Y\ .$$

Процессы в системе могут играть различную роль. Так, в системе автоматизированного проектирования процесс проектирования как движение от технического задания до рабочих чертежей является основной функцией системы. И в целом функционирование (а также создание) сложной системы обычно является процессом. Однако в том же проектировании наверняка необходимо учитывать целый ряд внутренних процессов: если что-то двигается, то уравнения движения; если идут химические превращения, то ход реакции. Таким образом, типичен учет процессов в системе как способ получения зависимостей выходов от входов в модулях разных иерархических уровней. При этом, в принципе, неважно, способствует ли в целом данный процесс выполнению системой ее функции или препятствует этому. К последнему случаю относятся, например, процессы износа, старения, а также действия противоположной стороны в игровых ситуациях.

1.1.5. Целенаправленные системы и управление

В заключительной части данного подраздела обсудим понятия, связанные с постановкой перед системой некоторой сформулированной цели. Системы при этом называют целенаправленными. Такими почти всегда будут искусственные системы.

Понятие цели системы определим как задачу получения желаемого выходного воздействия или достижения желаемого состояния системы.

Подчеркнем, что двоякая трактовка цели – через выходное воздействие или через состояние системы – удобна в приложениях. В теории можно считать целью только выходные воздействия, а желаемое состояние включать в список этих воздействий. В конкретных же случаях такая интерпретация состояния системы может вносить дополнительные трудности, а иногда даже приводить к путанице.

Постановка цели перед системой (часто говорят: глобальной цели) влечет за собой необходимость:

а) формулировки локальных целей, стоящих перед элементами системы и группами элементов;

б) целенаправленного вмешательства в функционирование (строение, создание) системы.

Обе эти операции тесно связаны, хотя с точки зрения практики, обычно, сначала разбивают глобальную цель на набор локальных, а потом ищут пути достижения локальных целей.

Набор локальных целей, как правило, сам имеет иерархическое многоуровневое строение и в той или иной степени соответствует общей иерархии в системе. В этом случае понятие «локальные цели» есть собирательный термин для целей всех иерархических уровней; для любой из них можно указать, в какую цель более высокого уровня она входит и (кроме целей самого низшего уровня) на что она дробится сама. При модульном строении системы локальные цели выступают как требования к выходам (выходным характеристикам) модулей. Именно продуманные требования на выходы согласовывают модули так, что состоящая из них система выполняет глобальную цель. Таким образом, локальные цели выступают важным регулятором организации частей и элементов в целенаправленную систему, а их согласование направляет проводимые в системе изменения в единое русло.

Заметим, что согласование обычно является сложной, плохо формализуемой процедурой. При этом конкретная локальная цель может получаться и такой, что затруднит выполнение соседней цели, и лишь компромисс между ними даст продвижение к глобальной цели системы.

Перейдем теперь к обсуждению того, как и за счет чего может быть выполнена конкретная цель.

Целенаправленное вмешательство в процесс в системе назовем управлением. Управление – важнейшее понятие для целенаправленных систем. Оно естественным образом связано с постановкой целей: именно возможность вмешательства, выбора, альтернативы делает процесс в системе вариативным, а один или более из этих вариантов – ведущим к достижению цели.

Управление – универсальный термин в смысле огромного многообразия его конкретных реализаций: в математических моделях можно выбирать числа, функции, алгоритмы, графовые структуры; в технических системах – силы, геометрические размеры, различные сигналы, включая команды ЭВМ, физические величины – от температуры до жесткости материала, концентрации и перемещения веществ; в экономике – размеры финансирования, материальные ресурсы и сроки их поставки, расстановку кадров; в социальной области – приказы, советы, действия, влияние на общественное мнение, организацию новых коллективов. Подчеркнем, что здесь перечислена лишь малая доля того, чем в целях управления можно распоряжаться в сложной системе.

Управление – чрезвычайно широкий и свободный в употреблении термин. Строгий (рискнем сказать – научный) подход к управлению требует четкого, однозначного определения:

а) того, чем распоряжаемся;
б) каковы пределы, в которых можно выбирать;
в) каково влияние данного управления на процесс.

Но на практике по всем перечисленным требованиям могут быть неясности, а двумя последними иногда вовсе пренебрегают. Это может приводить, в частности, к тому, что управление не будет вести к цели. Такое положение возможно и в строгой трактовке управления – когда отсутствует описание процесса в системе. В этом случае просто набирается опыт работы с «черным ящиком».

Наконец, следует сказать, что в случае, когда исходят из цели (что чаще всего и бывает), может быть ситуация, при которой не существует управления, обеспечивающего ее выполнение. Тогда пробуют расширить пределы, в которых выбирается управление, ввести новые управляющие воздействия (т.е. еще что-то разрешить менять), иногда кардинально изменяют структуру системы. В этой ситуации цель не лежит в области достижимости, которая обеспечивается имеющимися управлениями, и надо либо расширять эту область, либо переместить в ее направлении цель.

Перейдем к символьной записи введенных понятий. Общий вид процесса $S^u_{t_0t}$ с управлением u из некоторой возможной совокупности управлений U есть [сравните с (1.5)]

$$S^u_{t_0t}(y(t_0)) = y(t,u),\ \ y \in Y,\ \ t \in T,\ \ u \in U\ . \tag{1.6}$$

Этому управляемому процессу будет соответствовать отображение множеств

$$U \times T \times Y \to Y\ .$$

В (1.6) отражена лишь управляемость, вариативность процесса, но не его цель. Для записи процесса, приводящего к выполнению цели, начнем с того, что введем специальное обозначение f для тех выходных воздействий, на которые можно влиять выбором управлений u. Таким образом, величины f, обычно называемые критериями, есть часть выходов x^-_{jr} и x^-_J рассматриваемого модуля (см. (1.4)) или системы целиком. Обозначим теперь желаемый вид выходных воздействий через f_G, где G есть символ поставленной цели. Критерии f, естественно, считаем зависящими от характеристик y: $f = f(y)$.

Пусть существует момент t_G (или он задан) и существует состояние характеристик y_G, позволяющие достичь цели f_G. Пусть состояние y_G может быть достигнуто управляемым процессом $S^u_{t_0t}$. Тогда управление u_G, позволяющее выполнить цель f_G, определяется как часть триады (t_G, y_G, u_G), удовлетворяющей соотношениям

$$S^u_{t_0t}(y(t_0)) = y(t,\,u),\ \ f(y) = f,\ \ y \in Y,\ \ t \in T,\ \ u \in U\ . \tag{1.7}$$

Перейдем к примеру и на его основе сделаем важные дополнения к соотношению (1.7).

Рассмотрим процесс распространения вибрации в машинном отделении судна. Пусть под параметром процесса t понимается удаление от источника вибрации – блока двигателей и механизмов. Характеристикой процесса $y(t)$ будем считать амплитуды скоростей вибрирующих поверхностей. Выбор управления u будет состоять в нахождении жесткостей рамной конструкции, поддерживающей блок двигателей. Критерий $f(t)$ – вибромеханическая мощность в фиксированных точках (т.е. параметр f_G задан). Цель f_G – ввод величин f в заданный диапазон.

Заметим, что такая задача вроде бы имеет вид, отличный от записи (1.7). Действительно, стандартная математическая запись нахождения многомерной величины f в заданном диапазоне есть

$$a_s \le f_s \le \beta_s, \; s = 1, 2, \ldots .$$

Пояснение состоит в том, что под целью f_G понимается любая точка множества, описываемого приведенными неравенствами.

Управлением может быть и параметр процесса t. Например, в данной задаче можно искать место с нужным уровнем вибрации. Если же параметром процесса является время, то определяется наиболее благоприятный момент. Во всех этих случаях величина t включается в список управлений u.

На примере с распространением вибрации удобно проиллюстрировать и введенную нами двоякую трактовку цели (напомним: через заданный вид выходных воздействий и через заданное состояние системы). Вибромеханическая мощность – это типичный случай выходного воздействия, рассчитываемого по характеристике процесса – скорости. Однако в ряде случаев удовлетворительной является борьба с вибрацией не путем снижения вибромеханической мощности, а непосредственно уменьшением амплитуды скорости. При этом $f(y) = y$, и мы имеем в качестве цели желаемое состояние системы. Надобность во втором равенстве в (1.7) отпадает, и управляемый процесс можно записать как

$$S^u_{t_0 t}(y(t)) = y_G$$

с известной заранее величиной y_G.

Теперь обозначим глобальную цель через G^0, набор локальных целей первого иерархического уровня – через $\{G^{\mathrm{I}}\}$, второго – через $\{G^{\mathrm{II}}\}$ и т.д. Иерархическая структура целей в системе запишется так:

$$G^0 \to \{G^{\mathrm{I}}\} \to \{G^{\mathrm{II}}\} \to \ldots .$$

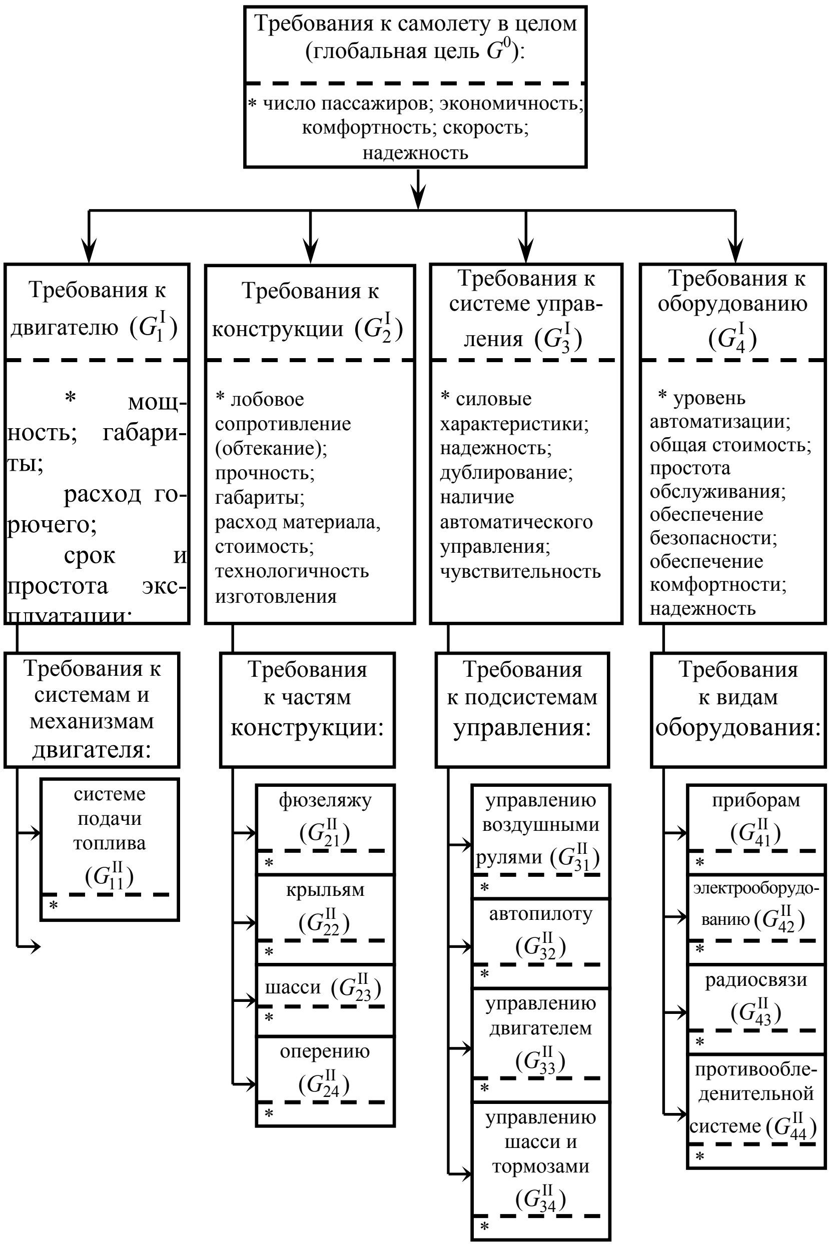

Рис. 1.6. Верхняя часть иерархической системы целей при проектировании пассажирского самолета

Графическое изображение этой структуры будет совпадать с рис. 1.4, *а*. Конкретный пример иерархии целей приведен на рис. 1.6.

Схема на рис. 1.6 демонстрирует, в частности, важное свойство управлений в сложной системе, состоящее в том, что собранные все вместе они сами образуют некоторую систему (подсистему), обладающую связями, структурой, иерархией. Такая система управления как бы накладывается на основную и обеспечивает ее превращение в целенаправленную систему. В связи с этим системы управления составляют предмет отдельного изучения, чему будет уделено внимание и в данной работе.

Краткий анализ понятия управления окончим указанием на то, что источником, формирующим управляющие команды, могут быть:

а) технические средства (управляющие и другие ЭВМ, микропроцессоры, программные устройства, регуляторы, следящие, стабилизирующие, компенсирующие системы и др.);

б) действия и решения человека (оператора, водителя, диспетчера, эксперта администратора, ответственного лица и др.).

Оба эти источника обладают как рядом общих свойств, определяемых их воздействующим характером на процессы в системе, так и существенными отличиями, которые будут рассмотрены ниже в подразделе, посвященном формализуемым и неформализуемым процедурам. Оба источника имеют свои достоинства и недостатки. Поэтому их полезно использовать совместно. Такие объекты принято вызывать автоматизированными системами управления (АСУ). Современную тенденцию в развитии АСУ можно определить как поручение техническим средствам формировать все те управляющие воздействия и выполнять все те сопутствующие операции, которые они делают качественнее и быстрее человека.

Контрольные вопросы

1. Что называется элементом, связью, системой?
2. Что называется структурой и иерархией?
3. Перечислите об основных типах связей.
4. Что такое состояние и процесс в системе?
5. Дайте характеристику целенаправленной системы.

1.2. ПРИНЦИПЫ СИСТЕМНОГО ПОДХОДА

1.2.1. Формулировка принципов

Зададимся вопросом – что же такое системность, что означает слово «системный», применяемое вместе с большим количеством терминов и понятий?

Поиски ответа на этот вопрос приводят нас к убеждению, что сложный объект надо рассматривать и как целое, и как состоящее из отдельных частей. Нужно исследовать предмет с разных сторон и точек зрения, вдаваться в его внутреннее строение и организацию... Но нельзя ли определить все это более четко достаточно полно и удобно для использования? Такая постановка вопроса приводит к формулировке положений, которые принято называть принципами системного подхода, Это – некоторые утверждения весьма общего характера, обобщающие опыт человека со сложными системами:

- *принцип конечной цели*: абсолютный приоритет конечной (глобальной) цели;
- *принцип единства*: совместное рассмотрение системы как целого и как совокупности частей (элементов);
- *принцип связности*: рассмотрение любой части совместно с ее связями с окружением;
- *принцип модульного построения*: полезно выделение модулей в системе и рассмотрение ее как совокупности модулей;
- *принцип иерархии*: полезно введение иерархии частей (элементов) и (или) их ранжирование;
- *принцип функциональности*: совместное рассмотрение структуры и функции с приоритетом функции над структурой;
- *принцип развития*: учет изменяемости системы, ее способности к развитию, расширению, замене частей, накапливанию информации;
- *принцип децентрализации*: сочетание в принимаемых решениях и управлении централизации и децентрализации;
- *принцип неопределенности*: учет неопределенностей и случайностей в системе.

Отметим, что хотя все перечисленные принципы так или иначе затрагиваются практически при любом изложении системного подхода, их формулировки пока не являются общепринятыми.

1.2.2. Обсуждение принципов системного подхода

Рассмотрим введенные принципы более подробно.

Первый из них – принцип конечной цели – означает, что в целенаправленной системе все должно быть подчинено глобальной цели. Любая попытка изменения, совершенствования и управления в такой системе должна оцениваться с точки зрения того, помогает или мешает она достижению конечной цели. Это накладывает особую ответственность на выбор цели и ее четкую трактовку. Расплывчатые, не полностью определенные конечные цели влекут за собой неясности в структуре и управлении системой, и, как следствие, неверные действия в системе. Такие действия могут быть и следствием неверия в конечную цель или в возможность ее достижения.

В несколько измененной трактовке принцип конечной цели применяют и к системам, которые не являются целенаправленными.

В этом случае понятие конечной цели заменяют понятиями основной функции, основного назначения, свойства системы. При этом принцип указывает, что изучение и работа с системой должны вестись на базе первоочередного уяснения этих понятий.

Следующие три принципа обладают довольно тесной взаимосвязью и иногда даже объединяются в один принцип единства и связи. Но, на наш взгляд, существуют причины, по которым их полезно рассматривать отдельно. Во-первых, принцип единства – это ориентация на «взгляд вовнутрь» системы или ее части, а принцип связи – на «взгляд изнутри». В разные моменты исследования полезна либо та, либо другая ориентация. Во-вторых, рекомендуемое в принципе единства расчленение системы с сохранением целостных представлений о ней на практике довольно резко отличается от процедуры выявления всевозможных связей, рекомендуемой в принципе связности. Наконец, процедура выявления связей, примененная ко всей системе целиком, приводит к принципу учета внешней среды, который нередко упоминают в литературе, но который, как следует из вышесказанного, можно не считать самостоятельным.

Принцип модульного построения указывает на возможность рассмотрения вместо части системы совокупности ее входных и выходных воздействий. Он утверждает полезность абстрагирования от излишней детализации при сохранении возможности адекватно описывать системы.

Принцип иерархии акцентирует внимание на полезности отыскания или создания в системе иерархического (доминирующего) характера связей между элементами, модулями, целями. Преимущества введения иерархии в системе обсуждались в п. 1.1.2. Здесь же поясним смысл слова «ранжирование» в формулировке принципа. Иерархические системы обычно исследуются и создаются «сверху», начиная с анализа модулей первого иерархического уровня. В случае отсутствия иерархии исследователь должен решить, в каком порядке он будет рассматривать части системы. Так, например, конструктор при создании нового образца выделяет в нем начальный элемент, к которому потом мысленно или на чертеже подгоняет второй, третий, следующие. Наладчик начинает поиск неисправности в системе с тестов, определяющих наиболее типичные отказы. Таким образом, они вводят порядок рассмотрения системы, который и называется ранжированием. Оно применимо и в сочетании с иерархией в системе, скажем, для введения очередности в модулях одного и того же уровня.

Для разбора принципа функциональности напомним, что мы определяли функцию системы как ее некоторое свойство. Функция для нас – это то, что система (модуль, элемент) «может делать» важного для целей рассмотрения. Полезно заметить, что понятие «функция» по отношению к системе является широко применяемым, но, к сожалению, не очень удобным. Дело в том, что оно имеет несколько различных смыслов. В рассматриваемом принципе «функция» означает свойство «может делать (влиять, обеспечивать)», которое отлично и от смысла «зависимость», и даже от смысла «назначение», поскольку последнее, вообще говоря, неприменимо к естественным системам, например, таким, как «атмосфера», «лес», «человек».

Принцип функциональности утверждает, что любая структура тесно связана с функцией системы и ее частей, и исследовать (создавать) структуру необходимо после уяснения функций в системе. На практике этот принцип, в частности, означает, что в случае придания системе новых функций полезно пересматривать ее структуру, а не пытаться втиснуть новую функцию в старую схему. Так, перестройка производства, связанная с введением автоматизации, ведет как к возникновению новых подразделений (вычислительный центр, группа системных программистов, группа создания и сопровождения банка данных), так и к перестройке структуры имеющихся. Эти изменения затрагивают, естественно, и систему управления.

Принцип развития достаточно хорошо пояснен в его формулировке. Понятие развития, изменяемости при сохранении качественных особенностей выделяется почти в любой естественной системе, а в искусственных возможность развития, усовершенствования, как правило, закладывается в основу создания системы. При модульном построении такое развитие обычно сводится к замене и добавлению модулей (частей). Так, возможности расширения функций и модернизации закладываются в принципы построения банков данных и знаний, программных комплексов, многоцелевых роботов и других сложных технических систем. Следует, однако, заметить, что пределы расширения функций обычно определены и достаточно ограничены. Вряд ли будет разумно создавать универсальное программное средство, способное управлять станком и играть в шахматы. Вряд ли кому-нибудь понадобится и робот, способный работать у плавильной печи и в квартире. Но вот замена частей, модернизация представляются нам безграничными. Практически безграничны и возможности запоминания информации, ведущие к самообучению, самоорганизации, искусственному интеллекту. Таким образом, использование принципа развития лежит в основе разработки этих направлений.

Принцип децентрализации рекомендует, чтобы управляющие воздействия и принимаемые решения исходили не только из одного центра (главенствующего элемента). Ситуация, когда все управления исходят из одного места, называется полной централизацией. Такое положение считается оправданным лишь при особой ответственности за все, происходящее в системе, и при неспособности частей системы самостоятельно реагировать на внешние воздействия. Система с полной централизацией будет негибкой, неприспосабливающейся, не обладающей «внутренней активностью». Весьма вероятно, что в такой системе каналы информации, ведущие к главному элементу, окажутся перегруженными, а сам этот элемент, будучи не в состоянии переработать такое количество информации, начнет выдавать неправильные управления.

Однако, чем выше степень децентрализации решений в системе, тем сложнее они согласовываются с точки зрения выполнения глобальной цели. Достижение общей цели сильно децентрализованной системой может обеспечиваться лишь каким-либо устойчиво работающим механизмом регуляции, не позволяющим сильно уклоняться от поведения, ведущего к выполнению цели. Такое положение встречается достаточно редко; во всех этих случаях имеет место ситуация с сильной обратной связью. (Таково функционирование рыночной экономики; в области живой природы – взаимодействие в системе, состоящей из акулы и маленьких рыбок лоцманов, которые наводят акулу на косяки рыб и питаются остатками ее пищи.)

В системах, где устойчивых механизмов регуляции нет, неизбежно наличие той или иной степени централизации. При этом возникает вопрос об оптимальном сочетании команд извне (сверху) и команд, вырабатываемых внутри данной группы элементов. Общий принцип такого сочетания прост: степень централизации должна быть минимальной, обеспечивающей выполнение поставленной цели.

Сочетание централизации и децентрализации имеет и еще один аспект. Его частным случаем будет передача сверху обобщенных команд, которые конкретизируются на нижних иерархических уровнях. Так, одной из команд верхнего уровня при управлении роботом-манипулятором будет: «Переместить схват в точку с такими-то координатами». Эта команда на следующем уровне управления в соответствии с имеющимися там алгоритмами будет разложена на необходимые для этого угловые повороты звеньев манипулятора, а на еще более низком уровне превращена в сигналы на включение и выключение электродвигателей, обеспечивающих отдельные повороты.

Принцип неопределенности утверждает, что мы можем иметь дело и с системой, в которой нам не все известно или понятно. Это может быть система с невыясненной структурой, с непредсказуемым ходом процессов, со значительной вероятностью отказов в работе элементов, с неизвестными внешними воздействиями и др. Частным случаем неопределенности выступает случайность – ситуация, когда вид события известен, но оно может либо наступить, либо не наступить. На основе этого определения можно ввести полное поле событий – это такое их множество, про которое известно, что одно из них наступит.

Как же оказывается возможным учесть неопределенность в системе?

Существует несколько способов, каждый из которых основан на информации определенного вида. Во-первых, можно оценивать «наихудшие» или в каком-то смысле «крайние» возможные ситуации и рассмотрение проводить для них. В этом случае определяют некое «граничное» поведение системы и на основе его можно делать выводы о поведении вообще. Этот способ обычно называют методом гарантированного результата (оценки). Во-вторых, по информации о вероятностных характеристиках случайностей (математическому ожиданию, дисперсии, другим оценкам) можно определять вероятностные характеристики выходов в системе. При этом, в связи со своеобразной трактовкой вероятностных результатов, можно получить сведения лишь об усредненных характеристиках совокупности однотипных систем. В-третьих, за счет дублирования и других приемов оказывается возможным из «ненадежных» элементов составлять достаточно «надежные» части системы. Математическая оценка эффективности такого приема также основана на теории вероятностей и носит название теории надежности.

Окончив обсуждение основных принципов системного подхода, подчеркнем, что именно основных, потому что в литературе встречается и ряд других принципов, одни из которых носят характер дублирования или уточнения приведенных (например, принцип внешней среды), другие имеют более узкую направленность или область применения, некоторые из них:

- *принцип полномочности*: исследователь должен иметь способность возможность (а в ряде случаев и право) исследовать проблему;
- *принцип организованности*: решения, действия, выводы в системе должны соответствовать степени ее детализации, определенности, организованности. (Поясним это крайней ситуацией, когда бессмысленно управлять системой, в которой команды не исполняются);
- *принцип чувствительности* (близок к принципу организованности): вмешательство в системе должно согласовываться с уровнем ее реакции на вмешательство. (Пример: различны навыки управления очень «послушным», чутким к командам механизмом и малочувствительным или с запаздывающей реакцией на управляющие воздействия);
- *принцип свертки*: информация и управляющие воздействия свертываются (укрупняются, обобщаются) при движении снизу вверх по иерархическим уровням.

1.2.3. Об использовании принципов системного подхода

Рассмотрим вопросы практического использования принципов системного подхода. Все они обладают очень высокой степенью общности, т.е. отражают отношения, сильно абстрагированные от конкретного содержания прикладных проблем. Такое знание нетипично для техники и естественных наук, в которых в основном используются утверждения и описания, пригодные для непосредственного применения.

Как же применять такое знание?

Для любой конкретной системы, проблемы, ситуации принципы системного подхода могут и должны быть конкретизированы: «Что это означает здесь?». Такая привязка к рассматриваемой проблеме производится исследователем. Он должен наполнять конкретным содержанием общие формулировки принципов. Опыт работы со сложными системами показывает, что это весьма полезно, потому что позволяет лучше увидеть существенные стороны проблемы, не забыть учесть важные взаимосвязи в ней. В ряде случаев продумывание конкретного содержания принципов системного подхода позволяет подняться на новый уровень осмысления системы в целом, выйти за рамки «узкого», «изнутри» отношения к ней.

Отметим, что интерпретация принципов для данного частного случая может приводить и к обоснованному выводу о незначимости какого-либо из принципов или об отсутствии условий для его применения. Так, в системе может не быть иерархии, она может считаться полностью определенной, связи могут быть заложены в самой математической модели и не требовать специального рассмотрения и т.д.

Многократное применение исследователем принципов системного подхода в различных системах приводит к тому, что у него развивается особый тип мышления, который принято называть системным. Такое мышление характеризуется умением более правильно (адекватно) ставить, а нередко и решать задачи, связанные со сложными системами.

Иногда утверждают, что принципы системного подхода удобны для критики уже имеющихся систем и менее удобны для создания новых. Такой взгляд связан с тем, что эти принципы возникли из многократно повторенного человеческого опыта. Однако имеется немало убедительных примеров того, как существенно новые проблемы решались именно на основе широкого использования системных принципов. Укажем на задачу создания многофункциональных вычислительных сетей и сред, которой, например, посвящена книга [5]. Близкий к этому характер имеют задачи создания автоматизированных систем проектирования (САПР), научных исследований (АСНИ), управления производством (АСУП).

Высокая общность принципов системного подхода во многом может быть преодолена их конкретизацией для фиксированных классов предметных задач. В ряде случаев это удобно выполнять в несколько приемов. Так, известны еще весьма общие, но уже предметно-ориентированные системные принципы проектирования, принципы создания программных комплексов. Они облегчают интерпретацию общих формулировок. Известны и примеры максимально конкретной трактовки принципов применительно к узким классам прикладных задач.

Контрольные вопросы

1. Сформулируйте принципы системного подхода.
2. Как практически используются принципы системного подхода?

1.3. СИСТЕМЫ И МОДЕЛИРОВАНИЕ

1.3.1. О понятии модели

Пусть имеется некоторая конкретная система. Лишь в единичных случаях мы имеем возможность провести с самой этой системой все интересующие нас исследования. В большинстве же ситуаций по разным причинам (сложность, громоздкость, недоступность и т.д.) мы вынуждены рассматривать не саму систему, а формальное описание тех ее особенностей, которые существенны для целей исследования. Такое формальное описание принято называть моделью.

Сначала приведем пример, когда можно рассматривать и саму систему, и ее модель. Для исследования радиотехнического элемента можно подать на его входы все интересующие нас комбинации сигналов и снять соответствующие выходные сигналы. Это будет полный натурный эксперимент. Если же описать прохождение сигнала внутри элемента формальным образом (дифференциальными уравнениями или оператором «вход–выход»), то мы сможем без самой системы определять выходные сигналы по входным. Это – работа с моделью. Что же лучше? Радиолюбителю, наверное, легче тестировать элемент, чем вводить и рассматривать уравнения. Но проектировщик радиоаппаратуры уже предпочтет хорошо описывающие элемент зависимости для того, чтобы с их помощью подобрать нужные параметры или даже саму структуру элемента.

Но с ростом сложности системы возможности натурного эксперимента резко падают. Он становится дорогим, трудоемким, длительным по времени, в слабой степени вариативным. Тогда предпочтительнее работа с моделью. В ряде же случаев мы вообще не имеем возможности наблюдать систему в интересующем нас состоянии. Например, разбор аварии на техническом объекте приходится вести по ее формальному (протокольному) описанию. Специалист по электронной технике будет изучать большинство типов ЭВМ по литературе и только часть из них опробует на практике. В этих примерах доступна лишь модель, но это не мешает нам эффективно познавать систему.

Рассмотрение вместо самой системы (явления, процесса, объекта) ее модели практически всегда несет идею упрощения. Мы огрубляем представления о реальном мире, так как оперировать категорией модели экономичнее, чем действительностью. Но вопрос выделения и формальной фиксации тех особенностей, которые существенны для целей рассмотрения, весьма непрост. Известно большое количество удачных моделей, составляющих предмет гордости человеческой мысли, – от конечно-элементной модели в прикладных задачах математической физики до модели генетического кода. Однако велико количество процессов и явлений, для которых на настоящий момент нет удовлетворительного описания. Правда, в области техники положение с моделированием можно считать удовлетворительным, но и здесь имеются «узкие» места, связанные с плохо определяемыми параметрами, коэффициентами, а также слишком грубые описания.

В разработке моделей различают три стадии: первую (основную) – построение модели; вторую – пробную работу с ней; третью – корректировку и изменение по результатам пробной работы. После этого модель считается готовой к использованию. Наиболее сложной и ответственной является первая стадия. Зачастую это в сильной степени неформализованный процесс, длительный путь проб и ошибок в поиске основной идеи. Построение принципиально новой модели носит характер открытия.

Достаточно сложным является и вопрос о том, кто должен создавать модель. Специалисту в данной практической области часто не хватает математических знаний, сведений о моделировании вообще, для сложных задач – знания системного анализа. Прикладному математику трудно хорошо ориентироваться в предметной области. Их совместная работа над моделью будет иметь смысл лишь при полном понимании друг друга.

Различают три основных вида моделей: вербальные (словесные, описательные); натурные (макетирование, физическое моделирование, масштабированные модели, модели части свойств и др.); знаковые. Среди знаковых моделей выделяется их важнейший класс – математические модели. Примеры других знаковых моделей – химические и ядерные формулы, графики, схемы, в том числе графовое изображение связей, информационных потоков в системе; с некоторой оговоркой (их относят и к макетам, т.е. натурным моделям) – чертежи, топографические карты.

Математическая модель – это описание протекания процессов (в том числе функционирования, движения), описание состояния, изменения системы на языке алгоритмических действий с математическими формулами и логических переходов.

Понятия действий с формулами и логических операций полезно дополнить. Так, к ним относят процедуру запоминания элемента, его вызова и подстановки в нужное место (это неявно присутствует при работе с любой формулой), операцию «следует за» в упорядоченной совокупности, операцию сравнения и идентификации совпадения элементов и др. Также традиционно математическая модель допускает работу с таблицами, графиками, номограммами, выбор из совокупности процедур и элементов. Последнее, в частности, требует операций предпочтения, частичной упорядоченности, включения, идентификации принадлежности и т.д. Логические переходы могут совершаться в схеме из вербально описываемых элементов (операций), что позволяет считать математической моделью даже жестко фиксированную последовательность действий человека. И это не есть какая-либо «натяжка», поскольку такая последовательность может эффективно изучаться математическими методами. Общий вывод состоит в том, что дать сколько-нибудь строгое и полное описание математической модели, по-видимому, невозможно. Упомянем здесь принадлежащую группе Бурбаки элегантную попытку уйти от этого множественного описания, провозгласив, что математическая модель (и математика в целом) – это просто аксиоматически охватываемые построения.

Вернемся к обсуждению знаковых моделей в целом.

Основное отличие этого типа моделей от остальных состоит в вариативности – в кодировании одним знаковым описанием огромного количества конкретных вариантов поведения системы. Так, линейные дифференциальные уравнения с постоянными коэффициентами описывают и движение массы на пружине, и изменение тока в колебательном контуре, и измерительную схему системы автоматического регулирования, и ряд других процессов. Однако еще более важно то, что в каждом из этих описаний одни и те же уравнения в буквенном (а вообще говоря, и в числовом) виде соответствуют бесконечному числу комбинаций конкретных значений параметров. Скажем, для процесса механических колебаний – это любые значения массы и жесткости пружины.

В знаковых моделях возможен дедуктивный вывод свойств, количество следствий в них обычно более значительно, чем в моделях других типов. Они отличаются компактной записью и удобством работы, возможностью изучения в форме, абстрагированной от конкретного содержания. Все это позволяет считать знаковые модели наивысшей ступенью и рекомендовать стремиться к такой форме моделирования.

Заметим, что деление моделей на вербальные, натурные и знаковые в определенной степени условно. Так, существуют смешанные типы моделей, скажем, использующие и вербальные, и знаковые построения. Можно даже утверждать, что нет знаковой модели без сопровождающей описательной – ведь любые знаки и символы необходимо пояснять словами. Часто и отнесение модели к какому-либо типу является нетривиальным. Так мы уже упоминали ситуацию с чертежами и с изучением словес но описываемого, но формализуемого поведения человека. В принципе, условность деления модели на типы означает, что это не более чем их удобная характеристика. Последнее отнюдь не опровергает приведенные выше утверждения о знаковом описании как наивысшей ступени моделирования, а лишь подчеркивает, что такая форма описания выступает желаемой, обладающей наибольшим числом достоинств характеристикой.

1.3.2. Общие и конкретные модели

Все типы моделей необходимо перед их применением к конкретной системе наполнить информацией, соответствующей используемым символам, макетам, общим понятиям. Наполнение информацией в наибольшей степени свойственно знаковым моделям, в наименьшей – натурным. Так, для математической модели – это численные (вместо буквенных) значения физических величин коэффициентов, параметров; конкретные виды функций и операторов, определенные последовательности действий и графовые структуры (там, где они не были фиксированы однозначно) и др. Наполненную информацией модель принято называть конкретной.

Модель без наполнения информацией до уровня соответствия единичной реальной системе называется общей (теоретической, абстрактной, системной).

Таким образом, если хотя бы часть параметров в модели не фиксирована, то она еще является общей. Практически всегда создаются и разрабатываются общие модели, описывающие классы, по крайней мере, близких однотипных систем. Но уровни их общности различны. Можно создать модель давления коленчатого вала на поддерживающие его подшипники, ограничившись при этом силовыми и геометрическими характеристиками, типичными, скажем, для автомобильных двигателей. Но можно рассмотреть модель реакций вращающегося твердого тела – и две модели будут очень близки, но, естественно, различны по уровню общности. Первую из них есть смысл рассматривать, если в ней учтены особенности, характеризующие именно данный узкий класс систем.

Наибольший интерес представляют общие модели с достаточно высоким уровнем абстракции. Такие модели могут самостоятельно изучаться, анализироваться, дополняться доказанными свойствами и утверждениями. Сведения, полученные при их теоретическом рассмотрении, будут применимы ко всем конкретным системам, содержащимся в них. Эти уровни общности или абстракции могут образовывать целые иерархические структуры, в которых переход к конкретной модели будет проходить в несколько этапов («спуск» к все более и более частному).

Особенно широко распространено и известно исследование абстрактных математических моделей. Типичными с точки зрения практики, являются модели в виде наборов формул, систем линейных и нелинейных алгебраических уравнений, дифференциальных уравнений, дискретных переходов, статистических описаний, аппроксимирующих представлений, описания игровых ситуаций и т.д. Можно говорить о ряде общих моделей в химии, физике, биологии, экономике. Абстрактное моделирование часто относят к области теории сложных систем [1 – 3]. В ней имеется ряд результатов, скажем, по теории линейных непрерывных нестационарных детерминированных или линейных дискретных стационарных стохастических систем. Каждая такая комбинация из четырех свойств (линейная–нелинейная, непрерывная–дискретная, стационарная–нестационарная, детерминированная–стохастическая) порождает общую модель высокого уровня с набором определенных свойств и своей методикой исследования.

Возвращаясь к превращению общей модели в конкретную, которое достигается наполнением ее информацией, отметим, что этот процесс не всегда прост, особенно при использовании неоднородной и объемной информации. Одновременно он настолько важен и ответствен, что ведет к самостоятельному исследованию понятий удобного хранения, выдачи и подготовки информации к непосредственному использованию. В настоящий момент эти исследования образуют отдельную, ориентированную на ЭВМ область знания, называемую организацией банков (баз) данных (знаний).

1.3.3. Формальная запись модели

Формальная запись модели традиционно занимает существенное место в общей теории систем, но полезна также и для анализа конкретной модели. Сначала обозначим:

- набор входных воздействий (входов) в системе – x^+ и всю их допустимую совокупность – X^+, $x^+ \in X^+$;
- набор выходных воздействий (выходов) в системе – x^- и всю их возможную совокупность – X^-, $x^- \in X^-$;
- набор параметров, характеризующих свойства системы, постоянные во все время рассмотрения, и влияющих на выходные воздействия системы, – a и всю их допустимую совокупность – A, $a \in A$;
- набор параметров, характеризующих свойства системы, изменяющиеся во время ее рассмотрения (параметры состояния), – y и всю их допустимую совокупность – Y, $y \in Y$;
- параметр (или параметры) процесса в системе (см. п. 1.1.4) – t и всю их допустимую совокупность – T, $t \in T$;

- правило S (функция, оператор) определения параметров состояния системы по входам x^+, постоянным параметрам a и параметру процесса t. Заметим, что мы всегда будем различать величины и правило их определения. Здесь запись $y = S(x^+, a, t)$ означает нахождение параметров по этому правилу, в то время как о величине y можно говорить и вне правила ее определения;
- правило V (функция, оператор) определения выходных характеристик системы по входам x^+, постоянным параметрам a, параметру процесса t и параметрам состояния y, т. е. $x^- = V(x^+, a, t, y)$;
- правило $\overline{V}$ (функция, оператор) определения выходных характеристик системы по входам x^+, постоянным параметрам a и параметру процесса t. Указанное правило $\overline{V}$ может быть получено подстановкой правила S в правило $\overline{V}$, что дает исключение из него параметров состояния: $x^- = \overline{V}(x^+, a, t)$.

На основе введения вышеуказанных воздействий, параметров и правил модель может быть записана как кортеж

$$\begin{gathered}\sum : \{x^+, x^-, a, t, y, S, V, \overline{V}\},\\ x^+ \in X^+, x^- \in X^-, a \in A, t \in T, y \in Y.\end{gathered} \tag{1.8}$$

Поясним определение модели (1.8) на ряде примеров. Сначала рассмотрим упрощенную схему работы дизельного двигателя. В этом случае имеем:

- входы (внешние воздействия): своевременная подача в камеру сгорания газовой смеси определенного состава; внешний момент (нагрузка) в точке вывода мощности;
- выход: мощность двигателя;
- неизменяемые параметры системы: объем камеры сгорания, число и расположение цилиндров, степень сжатия; размеры, массы и жесткость поршней, шатунов, коленвала, маховика и других частей силового механизма;
- параметр процесса: время или угол поворота коленвала;
- параметры состояния: температура и давление в камере сгорания, скорости (ускорения) движущихся частей, силы трения в двигателе;
- правило S (уравнения состояния): термодинамические уравнения, описывающие процесс сгорания газовой смеси, и механические уравнения, описывающие движение частей силового механизма;
- правило V: запись мощности двигателя в виде функции от скоростей движения частей силового механизма и внешнего момента; она равна произведению угловой скорости коленвала и внешнего момента;
- правило $\overline{V}$: запись мощности в виде функции от скорости подачи газовой смеси, ее состава и внешнего момента (нагрузки).

Во втором, математическом примере рассмотрим в качестве модели систему дифференциальных уравнений – $\dfrac{dy}{dt} + Ay = f(t)$, решаемую для различных начальных условий и различных правых частей.

В этом случае имеем:

- входы: начальные условия, вектор правых частей $f(t)$, значение t_1, до которого необходимо интегрировать систему;
- выход: значение $y(t_1) \stackrel{\Delta}{=} y_1$;
- неизменные параметры системы: матрица A;
- параметры состояния: вектор y;
- параметр процесса – t;
- правило S: решение дифференциального уравнения в зависимости от начальных условий, констант, правых частей и аргумента; $y = y(t_0, y_0, A, f(t), t)$;
- правило V: подстановка в решение дифференциального уравнения значения t_1; $y_1 = y \mid t = t_1$;
- правило $\overline{V}$: зависимость $y_1 = y(t_0, y_0, A, f(t), t_1)$.

Третий пример информационный. Рассмотрим модель длительности переработки человеком текста в резюме. В этом случае:

- входы: объем текста, численная оценка его сложности;
- выход: длительность τ составления резюме;
- неизменяемые параметры здесь будут соответствовать способностям данного человека: скорость осмысленного чтения текста и число повторных чтений в зависимости от его сложности, усредненное число переделок резюме;
- параметры процесса определяют объем проделанной работы на данный момент t: объем изученного текста, объем составленной части резюме, оставшееся число переделок резюме;
- параметр процесса: стадия работы или время;
- правило S: зависимость объема проделанной работы от объема и сложности текста, способностей человека, времени;
- правило V: зависимость величины от объема проделанной работы;
- правило $\overline{V}$: зависимость величины от объема текста, его сложности и способностей данного человека.

Обсудим определение модели (1.8) и приведенные примеры, в которых содержательно трактовались все восемь составляющих кортежа. Является ли это число универсальным, неизменным? Нет, это просто наиболее удобные на практике составляющие Их может быть как больше (см., например, ниже системы с управлением), так и меньше. Минимальное число составляющих имеет модель «черного ящика»:

$$\sum : \{x^{+}, x^{-}, \overline{V}\},\tag{1.9}$$

где $x^{-} = \overline{V}(x^{+})$.

Введение в рассмотрение «внутренности черного ящика» приводит к параметрам системы a, а типичное наличие процессов в системе – к параметрам состояния и процесса: y и t. На основе наличия процессов формулируются и правила S, V. Другими составляющими кортежа в определении модели могут быть входные случайные воздействия (представляющие собой часть входов x^{+}), характеристики структуры системы в отличие от характеристик элементов (выделенные из параметров a), некоторые свободные параметры модели, все множество значений которых должно быть учтено при расчете выходов (например, операциями взятия максимума, интегрированием), управления, введенные для целенаправленных систем.

Заметим также, что часто даже при незначительных изменениях постановки задачи происходит переход величин из одной составляющей кортежа в другую. Так, некоторую мало меняющуюся величину в системе можно отнести и к параметрам системы a (сделав условно постоянной), и к параметрам состояния. Математическим путем замены переменной нередко меняют местами параметр процесса и один m параметров состояния. В ряде случаев могут возникать трудности с отнесением данной величины к параметрам состояния или выходным воздействиям.

Так, в примере о двигателе интересно разобрать вопрос о месте сил трения в кортеже. Напомним, что они отнесены к группе параметров состояния. Однако при широко используемой записи сил трения через кинематические величины и постоянные коэффициенты трения они вообще могут быть выведены из рассмотрения с включением вместо них в список неизменяемых параметров системы указанных коэффициентов. Если же силы трения не зависят от кинематики, т.е. от состояния системы, то они могут считаться и входами. Наконец, при исследовании именно трения в двигателе эти силы станут выходами в системе.

Пример с моделью в виде системы дифференциальных уравнений интересен тем, что если считать выходом, не значение функции y в точке t_1, а саму функцию, то мы получаем совпадение операторов S и $\overline{V}$. Операторное равенство для V при этом является просто переобозначением: $x^{-} = y$. Такое положение дел, когда выходом в системе служит параметр состояния, достаточно типично. Аналогичная ситуация уже отмечалась нами при определении цели системы в п. 1.1.5. Для этого случая можно записать вместо (1.8) укороченный кортеж без правил S и V.

В примере с переработкой текста можно вполне обойтись без операторов S и V и строить сразу оператор $\overline{V}$. Такая ситуация, когда удобно сразу, без промежуточных стадий, искать основное правило $\overline{V}$, тоже встречается нередко и аналогично случаю с системой дифференциальных уравнений ведет к кортежу без S и V. Кстати, именно этим объясняется наличие на первый взгляд «лишней» составляющей $\overline{V}$ в (1.8), ведь еще в определении этого правила мы подчеркнули, что оно выводимо из предыдущих. Но именно типичность ситуации с отсутствием операторов S и V (или неудобство работы с ними) является основным оправданием практического удобства введения V в кортежную запись модели.

1.3.4. Общие свойства модели

Рассмотрим, как отражаются в записи (1.8) основные общие свойства системы.

Первое такое свойство – линейность или нелинейность. Оно обычно расшифровывается как линейная (нелинейная) зависимость от входов операторов S (линейность или нелинейность параметров состояния) или $\overline{V}$ (линейность или нелинейность модели в целом). Линейность может являться как естественным, хорошо соответствующим природе, так и искусственным (вводимым для целей упрощения) свойством модели.

Второе общее свойство модели – непрерывность или дискретность. Оно выражается в структуре множеств (совокупностей), которым принадлежат параметры состояния, параметр процесса и выходы системы. Таким образом, дискретность множеств Y, T, X^{-} ведет к модели, называемой дискретной, а их непрерывность – к модели с непрерывными свойствами. Дискретность входов (импульсы внешних сил, ступенчатость воздействий и др.) в общем случае не ведет к дискретности модели в целом. Важной характеристикой дискретной модели является конечность или бесконечность числа состояний системы и числа значений выходных характеристик. В первом случае модель называется дискретной конечной. Дискретность модели также может быть как естественным условием (система скачкообразно меняет свое состояние и выходные свойства), так и искусственно внесенной особенностью. Типичный пример последнего – замена непрерывной математической функции на набор ее значений в фиксированных точках.

Следующее свойство модели – детерминированность или стохастичность. Если в модели среди величин x^{+}, a, y, x^{-} имеются случайные, т.е. определяемые лишь некоторыми вероятностными характеристиками, то модель называется стохастической (вероятностной, случайной). В этом случае и все результаты, полученные при рассмотрении модели, имеют стохастический характер и должны быть соответственно интерпретированы (см. обсуждение принципа неопределенности в п. 1.2.2). Здесь же подчеркнем, что, с точки зрения практики, граница между детерминированными и стохастическими моделями выглядит расплывчатой. Так, в технике про любой размер или массу можно сказать, что это не точное значение, а усредненная величина типа математического ожидания, в связи с чем и результаты вычислений будут представлять собой лишь математические ожидания исследуемых величин. Однако такой взгляд представляется крайним. Удобный практический прием состоит в том, что при малых отклонениях от фиксированных значений модель считается детерминированной, а отклонение результата исследуется методами оценок или анализа ее чувствительности. При значительных же отклонениях применяется методика стохастического исследования.

Четвертое общее свойство модели – ее стационарность или нестационарность. Сначала поясним понятие стационарности некоторого правила (процесса). Пусть в рассматриваемом правиле присутствует параметр процесса, которым для удобства понимания будем считать время. Возьмем все внешние условия применения данного правила одинаковыми, но в первом случае мы применяем правило в момент t_0, а во втором – в момент $t_0 + \theta$. Спрашивается, будет ли результат применения правила одинаковым? Ответ на этот вопрос и определяет стационарность: если результат одинаков, то правило (процесс) считается стационарным, а если различен – нестационарным. Если все правила в модели стационарны, то стационарной называется и сама модель. Чаще всего стационарность выражается в неизменности во времени некоторых физических величин: стационарным является поток жидкости с постоянной скоростью, стационарна механическая система, в которой силы зависят только от координат и не зависят от времени.

Для отражения стационарности в формальной записи рассмотрим расширенный вид правила S, в которое введена его зависимость от начальных условий процесса t_0, y_0 и зависимость входов от параметра t:

$$y = S(x^+(t), a, t, t_0, y_0)\,.$$

Тогда для стационарного процесса имеет место равенство

$$S(x^+(t+\theta), a, t+\theta, t_0+\theta, y_0) = S(x^+(t), a, t, t_0, y_0)\,.$$

Аналогично можно определить стационарность правил V и $\overline{V}$.

Другим общим свойством модели является вид составляющих кортежа (1.8). Простейшим будет случай, когда входы, выходы и параметры a в системе – это числа, а правило $\overline{V}$ – математическая функция. Широко распространена ситуация, когда входы и выходы есть функции параметра процесса. Правила S, V, $\overline{V}$ тогда являются либо функциями, либо операторами и функционалами. Функциями, скажем, от параметров состояния могут быть и те параметры системы, которые мы ранее называли постоянными. Описанная выше ситуация еще достаточно удобна для исследования модели на ЭВМ.

Последним упомянем свойство модели (1.8), состоящее в конечности или бесконечности числа входов, выходов, параметров состояния, постоянных параметров системы. Теория рассматривает и тот, и другой тип модели, однако на практике работают лишь с моделями с конечно мерностью всех перечисленных составляющих.

1.3.5. Модели с управлением

Сделаем важное расширение формальной записи модели (1.8) – включим в нее управление. Пусть, как и в п. 1.1.5, мы рассматриваем управляемый процесс (правило перехода) S^u. Пусть это правило позволяет выбором управления u из некоторой фиксированной совокупности U достигать значения параметра состояния y_G, которое, в свою очередь, обеспечивает получение управляемых выходных воздействий f в виде f_G, соответствующем выполнению цели G [см. также (1.7)]. Кортежная запись управляемой модели имеет вид

$$\begin{gathered}\sum{}^u : \{x^+, x^-, f_G, a, u, t, y, S^u, V, \overline{V}\},\\ x^+ \in X^+,\ x^- \in X^-,\ a \in A,\ u \in U,\ t \in T,\ y \in Y.\end{gathered} \tag{1.10}$$

Все изменения в (1.10) по сравнению с (1.8) пояснены выше.

Составляющая u в (1.10) указывает на те величины, объекты, которыми мы можем распоряжаться для выполнения цели G. Напомним, что составляющая f_G в (1.10) есть сама цель G, записанная в виде требований на выходы модели.

Пусть теперь мы хотим превратить неуправляемую систему в управляемую. Из каких составляющих кортежа (1.8) выделит управление? Во-первых, из входов x^+. Часть из них может стать управляемыми, выбираемыми, контролируемыми. (Это, например, возможность выбора части сил, действующих на систему, посылки управляющих сигналов, допущение альтернативных решений.) Во-вторых, из параметров системы a. Это особенно типично для процесса проектирования. Мы получаем возможность выбирать размеры тел, массы, материал и тем самым создавать систему с нужными свойствами. В числе управлений, выделяемых из параметров a, могут быть и такие, которые описывают структуру системы. Их выбор будет означать изменение структуры с целью достижения заданного свойства системы.

Выбор структуры – весьма актуальная на практике, но, к сожалению, плохо формализуемая операция. Поясним это на примере. Пусть мы проектируем конструкцию, на которую ставиться некий прибор. Выберем стержневую форму конструкции – зафиксируем число стержней и их расположение (т.е. выберем структуру). Поставим задачу о выборе параметров стержней таким образом, чтобы, скажем, минимизировать вес конструкции при заданной прочности. Это – управление при заданной структуре. Но ведь мы сами себя ограничили формой конструкции. Возьмем теперь другое расположение стержней или допустим кратность данной модели реальной системе использование пластин. Весьма вероятно, что здесь удастся добиться еще меньшего веса. Мы стали управлять путем выбора структуры. Отметим, что в данном конкретном случае и, к сожалению, в целом практически не существует методов, которые позволили бы осмысленно перебирать структуры из достаточно широкого класса. Как правило, указанные задачи решаются с привлечением эвристических операций.

Возвращаясь к разбору перевода неуправляемой системы в управляемую, укажем и на обратную задачу – чем станут управления при переводе системы в неуправляемую? Ответ ясен: входами или неизменяемыми параметрами системы. Комплекс требований f_G просто исчезнет. Отсюда следует, что все утверждения и сведения о моделях вида (1.8) могут быть перенесены и на модели с управлением (1.10).

Рассмотрим теперь вопрос о практической полезности кортежных моделей (1.8) и (1.10). Уточнение математического вида совокупностей (множеств) X^+, X^-, A, T, U, Y, отнесение правил $S, V, \overline{V}$ к определенным математическим классам операторов и математическая формулировка требований f_G приводят к строго математической трактовке записей (1.8) и (1.10) и превращают эти модели в математические модели высокого уровня общности (см. также п. 1.3.2.). Напомним, что в целом мы рассматриваем кортежи (1.8) и (1.10), как и другие формальные записи в этой главе, лишь по форме близкими к математическим, а по сути просто удобной знаковой записью ряда понятий и операций, связанных с системами.

Теперь подчеркнем полезность этих кортежей для анализа конкретных моделей и моделей низкого уровня общности. Именно с такими моделями в основном приходится стакиваться на практике.

Разбор конкретной модели по схеме (1.8) и (1.10) состоит в отнесении различных величин, объектов, понятий, к приведенным составляющим кортежей и оказывается эффективным средством уяснения «внутренности» системы, составления и коррекции ее модели, выявления важнейших сторон моделирования. Еще более полезна эта процедура при введении управления в модель, ее перестройке и использовании в качестве элемента в более сложных моделях. Продумывание списков существенных входов, выходов, процессов, параметров в системе не всегда протекает гладко и беспроблемно. Но потраченный на это труд помогает не только эффективно строить операторы $S, V, \overline{V}$, но и выявлять избыточность или недостаточность величин и параметров модели, выяснять неправильное отнесение их к какой либо составляющей кортежа, учесть не принимавшиеся ранее во внимание обстоятельства, а то и в целом пересмотреть адекватность данной модели реальной системе.

1.3.6. Имитационное моделирование

Начнем рассмотрение моделирования с простого примера. Пусть моделью является некоторое дифференциальное уравнение. Решим его двумя способами. В первом получим аналитическое решение, которое позволило бы осмысленно перебирать структуры, запрограммируем найденный набор формул и просчитаем на ЭВМ ряд интересующих нас вариантов. Во втором воспользуемся одним из численных методов решения и для тех же вариантов последним изменением системы от начальной точки до заданной конечной. Если запись аналитического решения сложна, включает операции вычисления интеграла, то трудоемкость обоих способов будет вполне сравнима. Если ли принципиальная разница между двумя этими способами?

Оставим в стороне ряд известных преимуществ работы даже с громоздким аналитическом решением. Обратим внимание на то, что в первом способе решение в конечной точке дается как функция начала и постоянных коэффициентов дифференциального уравнения. Во втором для его нахождения приходится повторять путь, который система проходит от начальной до конечной точки. В ЭВМ осуществляется воспроизведение, имитация хода процесса, позволяющая в любой момент знать и при необходимости фиксировать его текущие характеристики, такие, как интегральная кривая, производные.

Мы подходим к понятию имитационного моделирования. Но, чтобы лучше разобраться в смысле этого термина, рассмотрим его применительно к той области, где он возник – в системах со случайными воздействиями и процессами. Для таких систем в 1960-х гг. стали моделировать на ЭВМ пошаговое протекание процессов во времени с вводом в нужный момент случайных воздействий. При этом однократное воспроизведение хода такого процесса в системе мало что давало. Но многократное повторение с разными воздействиями уже неплохо ориентировало исследователя в общей картине, позволяло делать выводы и давать рекомендации по улучшению системы.

Метод стали распространять на классы систем, где надо учесть возможно большее разнообразие в исходных данных, меняющиеся значения внутренних параметров системы, многовариантный режим работы, выбор управления при отсутствии четкой цели и др. Общим оставались специальная организация имитации поведения системы и многократное возобновление процесса по измененным сценариям.

Теперь дадим определение.

Моделирование процессов с многократным отслеживание хода их протекания каждый раз для различных условий называется имитационным моделированием.

Цель этого вида моделирования – получить представление о возможных границах или типах поведения системы, влияниях на нее управлений, случайных воздействий, изменений в структуре и других факторов.

Важной особенностью имитационного моделирования является удобное включение человека, его знаний, опыта, интуиции в процедуру исследования модели. Это делается между отдельными имитациями поведения системы или сериями имитаций. Человек изменяет сценарий имитации, что является важным звеном этого вида моделирования. Именно исследователь по результатам проведенных имитаций формирует следующие и, осмысливая полученные сведения, эффективно познает систему или двигается в ее исследовании к поставленной цели. Правда, следует заметить, что управлять процедурой многократной имитации может и ЭВМ. Однако наиболее полезным ее применение оказывается все-таки в сочетании с оперативным экспертным просмотром и оценкой отдельных имитаций.

Значительная роль человека в имитационном моделировании даже позволяет говорить об определенном противопоставлении методов чисто математического моделирования и имитации. Поясним это на примерах. Пусть мы имеем задачу оптимизации, которую решаем на ЭВМ при помощи некоторого запрограммированного алгоритма. В ряде сложных ситуаций алгоритм может остановиться или «зациклиться» далеко от оптимального решения. Если же весь путь решения шаг за шагом будет контролироваться исследователем, то это позволит, подправляя, изменяя и возобновляя работу алгоритма, достичь удовлетворительного решения. Второй пример возьмем из области систем со случайными воздействиями. Последние могут иметь такие «плохие» вероятностные свойства, что математическая оценка их влияния на систему практически невозможна. Вот тогда исследователь начинает машинные эксперименты с разными видами этих воздействий и постепенно получает хоть какую-то картину их влияний на систему.

Однако противопоставлять имитационное моделирование математическому в целом было бы методически неверно. Правильнее ставить вопрос об их удачном совмещении. Так, строгое решение математических задач, как правило, является составной частью имитационной модели. С другой стороны, исследователь крайне редко удовлетворяется однократным решением поставленной математической задачи. Обычно он стремится решить набор близких задач для выяснения «чувствительности» решения, сравнения с альтернативными вариантами задания исходных данных, а это не что иное, как элементы имитации.

Имитационное моделирование является одной из форм диалога человека с ЭВМ. При удачно организованной активности исследователя имитационное моделирование резко повышает эффективность изучения системы. Оно является особенно незаменимым, когда невозможна строгая постановка математической задачи (полезно попробовать разные постановки), отсутствует математический метод решения задачи (можно использовать имитацию для целенаправленного перебора), имеется значительная сложность полной модели (следует имитировать поведение декомпозиционных частей). Наконец, имитацией пользуются и в тех случаях, когда невозможно реализовать математическую модель из-за недостатка квалификации исследователя.

Кроме термина «имитационное моделирование» в литературе употребляется словосочетание «машинное моделирование». В него вкладывают весьма широкий смысл – от синонима имитации до указания на то, что в исследовании для каких-либо целей используется ЭВМ. Но, на наш взгляд, наиболее логичным является использование этого понятия в тех случаях, когда манипуляции с моделью целиком или почти целиком выполняются вычислительной техникой и не требуют участия человека.

1.3.7. Моделирование сложных систем

Выше, в пп. 1.3.1 – 1.3.6, мы рассматривали общие вопросы, связанные с моделированием систем. При этом речь шла о моделях, описывающих сразу всю систему целиком. Но годится ли этот путь для сложных систем? Единую модель для всей сложной системы принято называть макромоделью. Обычно такая модель достаточно проста и груба, она годна лишь для приблизительных оценок и самых общих выводов о системе. Попытки уточнить макромодель почти всегда ведут к такому росту ее сложности (размерности), что эффективное рассмотрение модели превышает возможности даже самых современных ЭВМ.

Чтобы выйти из этого положения, надо при моделировании системы вводить декомпозицию и деление на модули, при необходимости строить иерархию моделей, рассматривать потоки информации между отдельными моделями и т.д. Полученная совокупность моделей повторит структуру и иерархию самой системы.

Таким образом, фактически мы уже во многом касались проблем моделирования сложных систем. Ниже, в этом пункте, остановимся лишь на тех аспектах исследования взаимосвязанного набора моделей, которые не затрагивались выше.

Итак, основной спецификой моделирования сложной системы является учет связей между отдельными моделями (говоря также: согласование моделей). Как же это достигается? В самом общем плане можно указать, что строится схема взаимосвязанных моделей типа графовой структуры, в которой выходы одних моделей (модулей) являются входами других. При этом каждый отдельный модуль представляет собой модель в смысле определений (1.8) или (1.10). Для них фиксируется совокупность требований на входы. Работу с совокупностью моделей можно представить как «прохождение» задачи через эту совокупность, результатом которого является определение выходов сложной системы в целом.

Можно указать на две важные части описанной процедуры: первая – построение или выбор моделей для декомпозированных частей системы; вторая – согласование моделей. Эти части достаточно различны по содержанию. В первой, как известно, ищется формальное, чаще всего математическое описание. Во второй организуется совместное использование моделей. Первая требует, в основном, специальных знаний и навыков формализации. Вторая – прежде всего, системного подхода. В данном пункте мы интересуемся этой второй, специфичной для сложных систем частью.

Перейдем к примерам. Совокупности моделей будем изображать в виде графовой структуры с указанием стрелками переходов от одной модели к другой и передачи соответствующей информации. На рис. 1.7 представлена совокупность моделей для исследования влияния удара на работу механизмов, расположенных на поддерживающей их (опорной) конструкции. Моделью самого ударного воздействия является задание на коротком промежутке времени больших по величине внешних сил или ускорений характерных точек конструкции. Моделью конструкции считается стержневая, пластинчатая или другая система с фиксированной структурой. Используется декомпозиционный постулат (упрощение) о том, что движение опорной конструкции можно рассматривать отдельно от движения механизмов. Это, в частности, верно при массе конструкции, существенно превосходящей массу механизмов. Рассчитанные в итоге кинематические характеристики (ускорения, относительные перемещения) важнейших узлов механизмов сравниваются с моделью условий их разрушения.

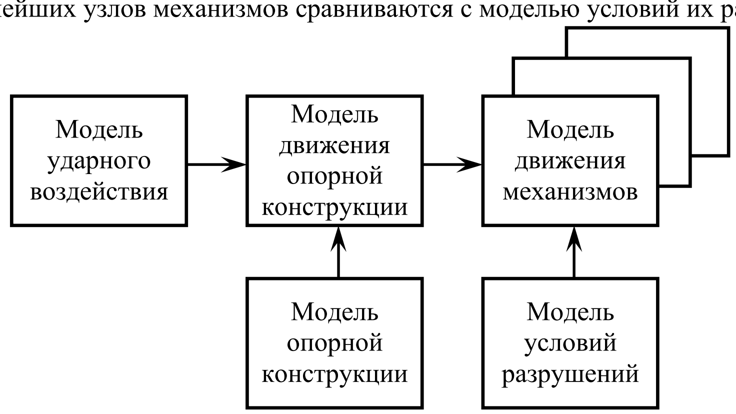

Рис. 1.7. Модель воздействия ударного воздействия

На рис. 1.8 изображена совокупность моделей для исследования функционирования подводного робота-манипулятора. Поясним, что собственно манипулятором называется выдвигаемый из корпуса робота механизм типа искусственной руки. Особенностью этой схемы является наличие верхнего иерархического уровня, на котором происходят моделирование режима использования робота и определение его производительности. Верхний уровень связан с собственно моделью робота передачей информации о необходимых погружениях и действиях манипулятора (стрелка с верхнего уровня на нижний), а также необходимыми для определения производительности сведениями о реальной динамике системы (стрелка с нижнего уровня на верхний).

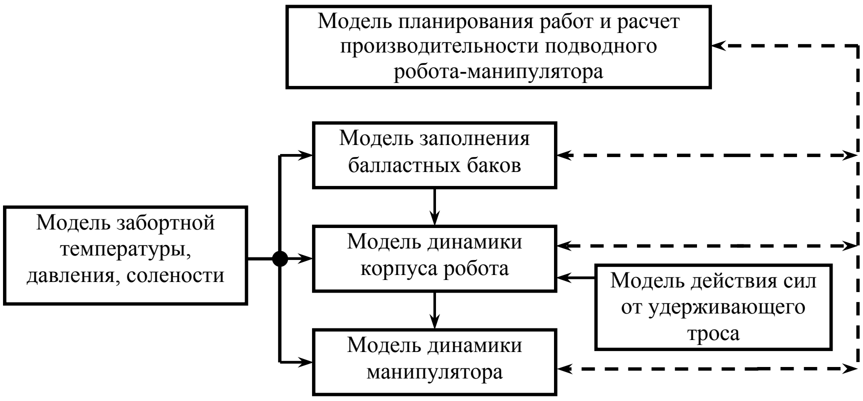

Рис. 1.8. Модель функционирования подводного робота-манипулятора

На рис. 1.9 показана не простая, но, тем не менее, еще достаточно грубая схема прогнозирования экологического состояния региона. В случае использования в ней математических моделей большинство функциональных зависимостей получают на основе статистической обработки наблюдений и применения интерполирующих представлений. Наиболее сложным для моделирования является блок «Модель изменения природных условий». Все основные компоненты этого блока – атмосфера, почва и растительность, водный бассейн, погода – связаны друг с другом зависимостями, которые на практике выявить не просто. Если же попытаться отбросить (не учитывать) часть этих связей, то мы рискуем упустить из виду возможные варианты резкого изменения экологического состояния.

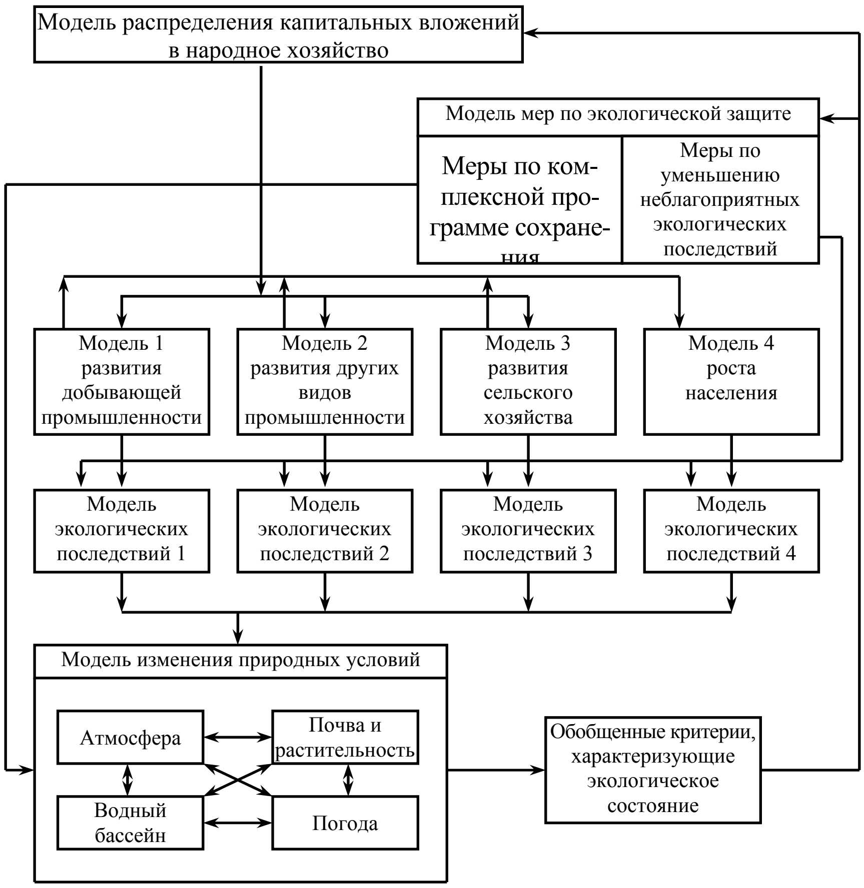

Рис. 1.9. Совокупность моделей для исследования развития экологического состояния региона

Практическое использование моделей сложных систем чаще всего носит характер имитации. В системе моделей организуется некоторый имитационный процесс, включающий своевременное подключение внешних воздействий и изменение условий протекания процессов, передачу информации от одной модели к другой, команды на начало и окончание работы данной модели и т.д. В приведенных примерах особенно выражен имитационный характер последней совокупности моделей, где на заданный промежуток времени будет имитироваться экологическое изменение состояния региона. Неоднократное повторение этого процесса по различным сценариям капитальных вложений и мер по защите окружающей среды позволит выбрать приемлемую стратегию реального развития.

Естественно, что и для имитации, и для моделирования сложных систем в целом характерно применение ЭВМ и других средств автоматизации. Однако имитация в ряде случаев может обходиться или, по крайней мере, не быть прямо основанной на использовании вычислительной техники. Таковы, например, деловые игры, где вся информация и фиксация действий могут проходить устно или на бумаге; военные игры на топографических картах; имитационное (двумя игроками) исследование неоконченных партий в шахматах. Имитация без ЭВМ лишний раз доказывает, что этот термин выходит за рамки определенного способа применения вычислительной техники и имеет более широкое значение.

Рассмотрим, что же может ЭВМ при моделировании сложных систем?

Прежде всего с помощью ЭВМ рассматриваются отдельные модели. Однако вычислительные машины удобны и для организации работы совокупности моделей. Они могут успешно осуществлять различные согласования по времени, проверять соответствие и достаточность данных, управлять потоками информации. Как правило, в сложной системе одна и та же информация нужна в разных местах (моделях). В этом случае насущной становится проблема организации банка данных. Отметим, что такой банк не только устраняет дублирующее хранение информации, но и гарантирует ее полную тождественность. Своеобразием сложных систем является и то, что при совместной работе разнородных моделей часто возникает проблема видоизменения данных и другой информации при передаче их от одной модели к другой или из банка данных в модель. Такое приспособление данных к задаче, в которой они будут использоваться, называется интерфейсной адаптацией.

В целом можно сделать вывод, что ЭВМ способна выполнять все основные операции при работе с моделями сложных систем. Совокупность нужных моделей, банка данных и разнообразных обслуживающих программных средств принято называть модельно-вычислительным комплексом.

Сам набор моделей, полностью готовых для использования, нередко называют библиотекой моделей. Чаще всего при этом речь идет о математической модели в виде набора формул, системы уравнений, алгоритма. Однако с полным основанием моделями, хранимыми с помощью вычислительной техники, можно назвать и тексты в библиотеке технологий и способов получения материала, библиотеке патентов и других инженерных решений, библиотеке диагнозов и типичных течений болезней и т.д. Из этого следует, что и вербальные модели могут являться основой при моделировании сложных систем. Специфична будет лишь работа с ними, которая в настоящее время, как правило, проводится человеком в диалоговом режиме общения с ЭВМ.

Рассмотрим вопрос, который достаточно важен при создании совокупности моделей. Как отлаживать работу отдельных моделей (модулей) в этой совокупности? Ведь режим и условия работы данной модели определяются ее связями в создаваемой схеме и, вроде бы, пока не заработают все остальные модели, мы не можем отлаживать и данную.

Эта проблема решается работой с фиктивными данными (связями), заменяющими настоящие. Для таких искусственно организованных входов в системном программировании даже возник специальный термин «заглушки». Его вполне можно применить к процессу отладки произвольной совокупности моделей, которые будут на начальных стадиях своей разработки совершенствоваться на «правдоподобных» – специально подобранных заглушках и лишь потом, окончательно доводиться на совместной работе всей совокупности моделей. При этом практика показывает, что чем удачней были выбраны заглушки (из опыта предыдущей работы, из схожих систем, по интуиции, просто перебором большого диапазона входных данных), тем меньше новых проблем возникает на последнем, системном этапе отладки.

1.3.8. Автоматизированное моделирование

При употреблении выражения «автоматизированное моделирование» речь идет о машинном построении модели, проводимом без участия или с минимальным участием человека. Это оказывается возможным, если сформулированы, формально реализованы и превращены в программные средства правила построения достаточно широкого класса моделей, к которому принадлежит и та конкретная, с которой мы хотим работать.

Поясним сказанное на двух типичных примерах. Первый из них – построение модели в виде системы линейных дифференциальных уравнений. Допустим для конкретности, что речь идет о системе упруго связанных тел. В этом случае построение дифференциальных уравнений может быть произведено по основанному на использовании законов механики алгоритму, исходя из задания взаимного расположения твердых тел и жесткостных (типа пружины) связей между ними. Исследователь специальным образом кодирует указанное расположение, а также физические параметры элементов системы, вводит эту кодировку в ЭВМ, после чего машина сама формирует модель в виде конкретной системы дифференциальных уравнений. Поскольку эти уравнения без надобности можно не выводить на печать, часто получается, что человек при автоматизированном моделировании работает с моделью, которой не видит и не знает.

Второй пример относится к рассмотрению напряженного состояния или динамики деформаций произвольной стержневой системы. Описывающие эти задачи уравнения также могут быть получены при помощи специальных алгоритмов на основе кодировки структуры системы, которая здесь заключается в задании координат концов стержней, их внутренних параметров и типа соединения между собой.

Таким образом, автоматизированное моделирование состоит в организации в ЭВМ последовательности действий по построению модели, которые инициируются введением в машину удобной для человека кодировки конкретной системы. Именно это по определенному и обычно не простому алгоритму делают программные средства автоматизированного моделирования данного класса задач.

Следует упомянуть важный раздел прикладной математики, в котором автоматизированное моделирование выступает одним из определяющих факторов. Это – метод конечных элементов, который в самом простом изложении сводится к делению системы на конечное число небольших элементов, внутри которых оказывается возможным применять грубые приближения. Основная задача в этом случае состоит во взаимосвязи элементов. Она достигается построением линейной алгебраической системы большой (до нескольких тысяч уравнений) размерности. Ясно, что ручное построение такой системы практически невозможно, и его следует поручать ЭВМ. Наиболее универсальные программные комплексы метода конечных элементов не только составляют и решают эту систему, но и сами производят разбиение на элементы, которое удовлетворяет требованиям по точности, а также ряду других условий. Можно утверждать, что на настоящий момент в методе конечных элементов реализован один из наивысших уровней автоматизированного моделирования:

Контрольные вопросы

1. Что такое модель и для чего они используются?
2. Перечислите типы моделей.
3. Расскажите об общих свойствах модели.
4. Дайте краткую характеристику моделям с управлением.
5. Расскажите об имитационном моделировании.
6. Дайте характеристику процесса моделирования сложных систем.
7. Что такое автоматизированное моделирование?

1.4. МЕТОДОЛОГИЯ СИСТЕМНЫХ ИССЛЕДОВАНИЙ

Ранее был введен целый ряд терминов и понятий, используемых при исследовании сложных систем. В этом заключительном разделе главы будет предложена некоторая общая методика по использованию рассмотренных понятий, представленная в виде стадий системного исследования. Подчеркнем, что эта абстрактная схема есть некая последовательность ориентирующих действий, и исследование каждой конкретной системы будет более или менее значительно отличаться от предложенной схемы.

1.4.1. Формирование общих представлений о системе

1. Выявление главных функций (свойств, целей, предназначения) системы. Формирование (выбор) основных предметных понятий, используемых в системе.

Речь идет об уяснении основных выходов в системе, именно с этого лучше всего начинать ее исследование. Должен быть определен тип выхода: материальный, энергетический, информационный; они должны быть отнесены к каким-либо физическим или другим понятиям [выход завода – продукция (какая?), выход системы управления – сигналы (для чего? в каком виде?), выход системы автоматизированного проектирования – конструкторская документация (чертежи чего?), выход двигателя – мощность (механическая? электрическая?) и т.д.].

2. Выявление основных частей (модулей) в системе и их функций. Понимание единства этих частей в рамках системы.

Происходит первое знакомство с внутренним содержанием системы, выявляется, из каких крупных частей она состоит и какую роль каждая часть играет в системе. Эта стадия – получение первичных сведений о структуре и характере основных связей. Такие сведения удобно представлять и изучать при помощи структурной схемы системы, где, например, выясняется наличие преимущественно последовательного или параллельного характера соединения частей, взаимная или преимущественно односторонняя направленность воздействий между частями и т.п. Уже на этой стадии следует обратить внимание на так называемые системообразующие факторы, т.е. на те связи, взаимообусловленности, которые и делают систему системой.

3. Выявление основных процессов в системе, их роли, условий осуществления; выявление стадийности, скачков, смен состояний и т.п. в функционировании системы; в системах с управлением – выделение основных управляющих факторов.

Изучается динамика важнейших изменений в системе, ход событий в ней, вводятся параметры состояния, рассматриваются факторы, изменяющие эти параметры, обеспечивающие течение процессов, а также условия начала и конца процессов и т.д. Изучается, управляемы ли процессы и способствуют ли они осуществлению системой своих главных функций. Для управляемых систем уясняются основные управляющие воздействия, их тип, источник и степень влияния на систему.

4. Выявление основных элементов «не-системы», с которыми связана изучаемая система. Выявление характера этих связей.

Решается ряд отдельных проблем. Исследуются основные внешние воздействия на систему (входы). Определяются их тип (вещественные, энергетические, информационные), степень влияния на систему, основные характеристики. Фиксируются границы того, что считается системой, определяются элементы «не-системы», на которые направлены основные выходные воздействия. Здесь же полезно проследить эволюцию системы, путь ее формирования. Нередко именно это ведет к пониманию структуры и особенностей функционирования системы В целом данная стадия позволяет лучше уяснить главные функции системы, ее зависимость и уязвимость или относительную независимость во внешней среде.

5. Выявление неопределенностей и случайностей в ситуации их определяющего влияния на систему и выбор способа их математической формализации.

Стадией 5 заканчивается формирование общих представлений о системе. Как правило, этих представлений достаточно, если речь идет об объекте, с которым мы непосредственно работать не будем. Если же речь идет о системе, которой надо заниматься с целью ее глубокого изучения, улучшения, управления, то нам следует пойти дальше по «спиралеобразному» пути углубленного исследования системы.

1.4.2. Формирование углубленных представлений о системе

6. Выявление разветвленной структуры, иерархии, формирование представлений о системе как о совокупности модулей, связанных входами-выходами.

7. Выявление всех элементов и связей, важных для целей рассмотрения. Их отнесение к структуре иерархии в системе. Ранжирование элементов и связей по их значимости.

Стадии 6 и 7 тесно связаны друг с другом, поэтому их обсуждение полезно провести вместе. Стадия 6 – это предел познания «внутрь» достаточно сложной системы для лица, оперирующего ею целиком. Более углубленные знания о системе (стадия 7) будет иметь уже только специалист, отвечающий за ее отдельные части. Для не очень сложного объекта уровень стадии 7 – знание системы целиком – достижим и для одного человека. Таким образом, стадии 6 и 7 говорят об одном и том же, но в первой из них мы ограничиваемся тем разумным объемом сведений, который доступен одному исследователю.

При углубленной детализации важно выделять именно существенные для рассмотрения элементы (модули) и связи, отбрасывая все то, что не представляет интереса для целей исследования. Познание системы предполагает не всегда простое отделение существенного от несущественного, а также удаление большего внимания более существенному. Детализация должна затронуть и уже рассмотренную в стадии 4 связь системы с «не-системой». На стадии 7 совокупность внешних связей считается проясненной настолько, что мы имеем право говорить о доскональном знании системы.

Стадии 6 и 7 подводят итог общему, цельному изучению системы. Дальнейшие стадии уже рассматривают только ее отдельные стороны. Поэтому важно еще раз обратить внимание на системообразующие факторы, на роль каждого элемента и каждой связи, на понимание, почему они именно таковы или должны быть именно таковы с точки зрения единства системы.

8. Учет изменений и неопределенностей в системе, входов и постоянных параметров в управляемые определен этой стадией. Исследуются недопустимые пределы управления и способы их реализации, детальное изменение свойств системы, которое принято называть «старением», а также возможность замены отдельных частей (модулей) на новые, позволяющие не только противостоять старению, но и повысить качество системы по сравнению с ее первоначальным состоянием. Такое совершенствование искусственной системы принято называть развитием. К нему также относят улучшение характеристик модулей, подключение новых модулей, накопление информации с целью ее лучшего использования, а иногда и перестройку структуры, иерархии, связей.

Основные неопределенности в стохастической системе считаются исследованными на стадии 5. Однако недетерминированность всегда присутствует и в системе, не предназначенной работать в условиях случайного характера входов и связей. О приемах учета неопределенностей и случайностей в системах, которые в целом считаются детерминированными, говорилось в п. 1.2.2. Здесь же добавим, что учет неопределенностей в этом случае обычно превращается в исследование чувствительности важнейших свойств (выходов) системы. Под чувствительностью понимают степень влияния изменения входов на изменение выходов.

9. Исследование функций и процессов в системе с целью управления ими. Введение управления и процедур принятия решения. Рассмотрение управляющих воздействий как систем управления.

Для целенаправленных и других систем с управлением данная стадия имеет большое значение. Основные управляющие факторы были уяснены при рассмотрении стадии 3. Однако это носило характер общей информации о системе. Для эффективного введения управлений или изучения их воздействий функции системы и процессы в ней необходимо глубокое знание системы. Именно поэтому мы говорим об анализе управлений только сейчас, после всестороннего рассмотрения системы. Напомним, что управление может быть чрезвычайно разнообразным по содержанию – от команд специализированной управляющей ЭВМ до министерских приказов (см. также п. 1.1.5). Однако возможность единообразного рассмотрения всех целенаправленных вмешательств в поведение системы позволяет говорить уже не об отдельных управленческих актах, а о системе управления, которая тесно переплетается с основной системой, но четко выделяется в функциональном отношении.

На данной стадии выясняется, где, когда и как (в каких точках системы, в какие моменты, в каких процессах, скачка, выборках из совокупности, логических переходах и т.д.) система управления воздействует на основную систему, насколько это эффективно, приемлемо и удобно реализуемо. При введении управлений в систему и постоянных параметров должны быть определены допустимые пределы.

Стадии 6 – 9 посвящены углубленному исследованию системы. Далее идет специфическая стадия моделирования. Напомним, что в этом разделе пока не говорилось о моделях, хотя, конечно, многие из упомянутых свойств системы удобно изучать именно на них. Однако противоречия здесь нет. О создании модели можно говорить только после полного изучения системы.

1.4.3. Моделирование системы как этап исследования

10. При введении совокупности моделей для описания системы мы опять, в очередной раз, повторяем ее рассмотрение – на этот раз с целью удобного отражения ее свойств. На стадиях 1 – 9 речь шла о регистрации в системе фактов и свойств, введении иерархии, модульности и т.д. Теперь же мы ставим цель – создать описание системы, пригодное для предсказания ее поведения и вывода неочевидных свойств. Если ранее в представлениях исследователя была допустима слитность модели и реальной системы, то на этой стадии необходимо отделять их друг от друга и четко представлять то огрубление и приближенность, которые несет в себе модель.

Важное отличие моделирования от стадий 1 – 9 состоит в том, что моделирование идет не сверху, от глобальной функции и выделения основных частей, а снизу, с построения моделей для отдельных процессов, для простых модулей нижних иерархических уровней. И далее, на основе разумного усложнения моделей перехода к их совокупностям моделируются все более крупные модули и, наконец, система в целом. Для последней, как отмечалось в п. 1.3.7, возможно, окажется полезным и построение макромоделей.

На практике наиболее распространены дедуктивное моделирование и близкие к нему методы. Это означает использование какой-либо общей (в смысле п. 1.3.2) модели для вывода из нее нужной конкретной. Такая процедура часто включает упрощение, эмпирическое или вполне обоснованное (теоретическое) уточнение коэффициентов, параметров, вида функций. Близким к дедукции является моделирование по аналогии – моделирование с взятием за основу сходной системы или ситуации. Хотя такой метод может быть подвергнут серьезной критике за часто необоснованное перенесение свойств другой системы на рассматриваемую, следует признать, что в целом он весьма продуктивен, а его недостатки преодолеваются критическим отношением к модели, которая используется в качестве основы.

Другой способ моделирования – индукционный, дающий в дословном переводе «выведенные из частного» модели. К ним относятся создание принципиально новых моделей, а также эмпирическое моделирование. Хотя точность таких аппроксимационных моделей чаще всего невелика, в ряде случаев удается успешно работать и с ними.

Представляется полезным здесь же сказать о точности моделирования в целом. Заметим, что она может быть как недостаточной для целей рассмотрения, так и чрезмерной. Общее утверждение о точности гласит, что она должна быть минимальной, обеспечивающей отражение всех важных особенностей системы. Уход от излишней детализации – это экономия времени и памяти ЭВМ, уменьшение объема исходных данных и даже рост надежности модели, связанный с уменьшением ее сложности. С другой стороны, слишком простая модель не опишет существенные качественные особенности системы и приведет к верным выводам о ее поведении. Найти грань разумной сложности часто нелегко, и она окончательно определяется только на этапе отладки модели при решении практических задач.

1.4.4. Сопровождение системы

11. Накопление опыта работы с системой и ее моделью. Уточнение сведений о системе, доводка и совершенствование моделей.

Данную стадию можно охарактеризовать как опытную эксплуатацию наших знаний о системе. Достаточны и верны ли они? Это проверяется только работой с системой и ее моделью. Если выявлено несоответствие между предсказанием поведения системы и результатами ее функционирования, то должны быть пересмотрены наши представления, возможно, заново произведен анализ структуры и иерархии для нахождения недостающих или неверно определенных элементов и связей.

Здесь же, еще раз после стадии 10 полезно акцентировать внимание на том, достаточно ли соответствие модели исходной системе. Обычно именно опыт эксплуатации выявляет «слабые» (как, впрочем, и «сильные») стороны модели. Накопление опыта имеет и психологический аспект. Нередко человеку трудно отрешиться от тех идей и представлений, которые, мягко говоря, не вполне выдерживают проверку на практике. Непросто признать ошибку и поменять суждение, но исследователю надо быть готовым к этому.

12. Оценка предельных возможностей системы. Исследование отказов, выходов из строя, отклонений от нормы.

Для эксплуатации системы важны сведения о ее работоспособности, потенциале, ресурсах, о возможных отказах, отклонениях, выходах из строя, незапланированных режимах, катастрофах. Работоспособность системы проверяется ее постоянным или периодическим тестированием. Аналогично проверяются ее ресурсы. Набор таких тестов может быть достаточно сложен сам образовывать некоторую систему, включающую в себя обработку и расшифровку данных тестирования, а также их комплексный анализ. Отказы и другие незапланированные (нетабельные) явления изучаются с точки зрения вероятности их возникновения, мер предупреждения и вариантов реагирования на них.

13. Расширение функций (свойств) системы, изменение требований к ней, новый круг задач, новые условия работы. Включение системы элементом в систему более высокого уровня.

Речь идет о частичной перестройке функционирования (назначения) системы или изменении задач ее исследования. При этом значительная часть сведений и знаний о системе останется нужной, и нет смысла относиться к такой измененной системе как к совершенно новой. Так же как и при моделировании по аналогии, здесь полезно «отталкиваться от близкого» и экономить на этом силы, время и средства. Схожими будут и проблемы: необходимо уяснить все отличия новой ситуации, проверить, не догматично ли следование старой структуре, иерархии и трактовке других системных понятий.

Включение системы элементом в некоторую макросистему само по себе требует лишь пересмотра стадии 4, где изучаются основные связи системы с «не-системой», и стадии 6, где рекомендуется обратить внимание уже не только на основные, а на все существенные связи с внешней средой. Однако выдвижение требований к системе со стороны макросистемы может привести к необходимости пересмотра всех основных системных понятий (выходы, управление и т.д.) и тем самым, затронуть все стадии исследования.

1.4.5. Особенности создания новой системы

Мы закончили описание стадий исследования сложной системы. В заключение остановимся на особенностях применения этой методики к процессу создания новой системы. Такой процесс также включает все перечисленные этапы, которые, однако, здесь носят характер предварительного проектирования (продумывания, детализации основного замысла). Далее следует период собственно создания системы. Следует подчеркнуть, что если продумывание идет от глобальной цели вниз к проектированию элементов, то создание, наоборот, начинается с подгонки друг к другу простейших элементов, далее включает согласование работы модулей все более высоких иерархических уровней и, наконец, обеспечение функционирования всей системы целиком.

Такая «разнонаправленность» процессов продумывания и создания (проектирования и конструирования) нередко приводит к тому, что проблемы согласования могут быть решены только повторным продумыванием (перепроектированием), включающим изменение требований к модулям или даже к системе целиком. В итоге процесс создания новой системы характеризуется многократным проходом по стадиям 1 – 13, а осуществление замысла имеет вид итеративного приближения к цели.

Контрольные вопросы

1. Перечислите основные этапы применения методологии системных исследований.
2. Дайте характеристику каждому этапу.

Глава 2

ИССЛЕДОВАНИЕ ДЕЙСТВИЙ И РЕШЕНИЙ

В главе 1 были рассмотрены вопросы, традиционно относящиеся к методологии системного анализа. При этом значительная часть из них пересекалась с проблематикой теории систем и моделирования процессов и явлений. Вторая глава посвящена системному анализу в узком смысле этого слова – методам и приемам организации процесса решения сложной задачи.

2.1. ДЕЙСТВИЯ И ИХ АНАЛИЗ

2.1.1. Процедуры и операции

Пусть имеется задача создания, совершенствования, эксплуатации или расширения функций сложной системы. Будем исходить из того, что процесс ее решения может быть разделен на некоторые элементарные акты, которые назовем процедурами. Очевидно, что процедуры должны быть определенным образом связаны друг с другом. В частности, необходимо определить их порядок, передачу информации, условия начала и окончания выполнения. Исполнение совокупности процедур приводит к тому, к чему каждая отдельная процедура привести не может, – к решению поставленной задачи. Таким образом, налицо все признаки системы (см. п. 1.1.1), в которой в качестве отдельных элементов выступают процедуры.

Это означает, что процесс решения задачи может рассматриваться как некоторая система процедур, обладающая внутренней организацией, структурой, иерархией, управлением. Такая система относится к классу целенаправленных систем (см. п. 1.1.5) с явно сформулированной целью – решением поставленной задачи. Для нее справедливы утверждения, приведенные в гл. 1.

В системе процедур модулем будет являться группа процедур, обладающая определенной целостностью и относительной независимостью. Для такой группы процедур введем термин «операция».

Итак, операция всегда состоит из отдельных процедур, но, как и модуль, может состоять из операций более низкого уровня. Такая иерархия (разномасштабность) понятия операции, которая на практике может достигать трех-четырех уровней, как правило, не вызывает неудобств, а, наоборот, позволяет акцентировать внимание то на единстве, то на делимости рассматриваемой совокупности. Приведем пример. Операция приобретения нового оборудования делится на операции его заказа, оплаты, перевозки, установки, наладки. Каждая из этих операций также раскладывается на более элементарные акты. Скажем, перевозка будет состоять из упаковки, одной или нескольких погрузок, разгрузок и транспортировок. Если нет необходимости вникать в эти акты более глубоко, то для нас они будут процедурами. Если же надо для каких-то целей рассматривать, из чего состоит погрузка, то она в нашем исследовании будет операцией. На этом примере хорошо видно, что взгляд на данный акт как на элементарный при необходимости может заменяться более детализированным – как на составной (операцию).

Для случая древовидной иерархии в процессе решения задачи (что далеко не всегда имеет место) деление на операции и процедуры может иметь вид, изображенный на рис. 2.1

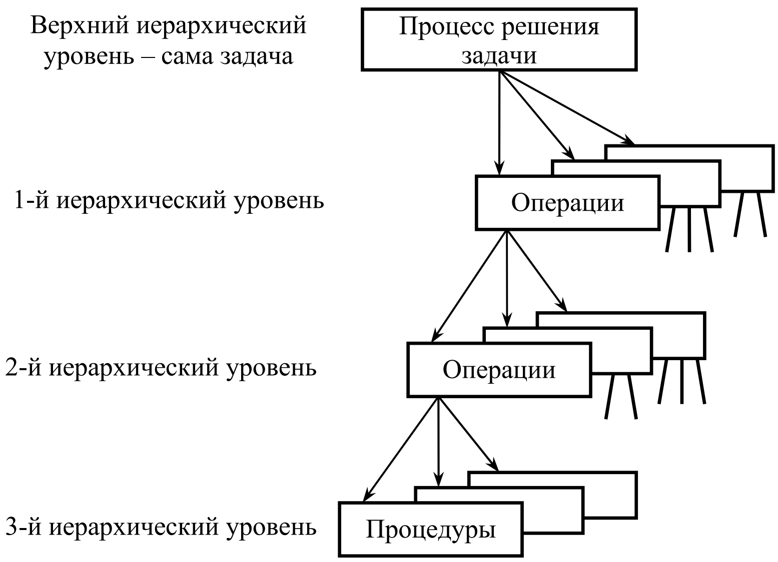

Рис. 2.1. Пример деления задачи на операции и процедуры

Приведем практический пример иерархического деления задачи на операции и процедуры. На рис. 2.2 дано описание действий по организации студенческого стройотряда в подготовительный период – от момента назначения руководства отряда и определения места дислокации до его отъезда на работу.

В дальнейшем для удобства изложения, когда не будет необходимости отличать друг от друга процедуру и операцию, будем использовать для них единый термин «действие». Процесс решения задачи при этом будет представлять собой систему действий, изучению которой и посвящена данная глава.

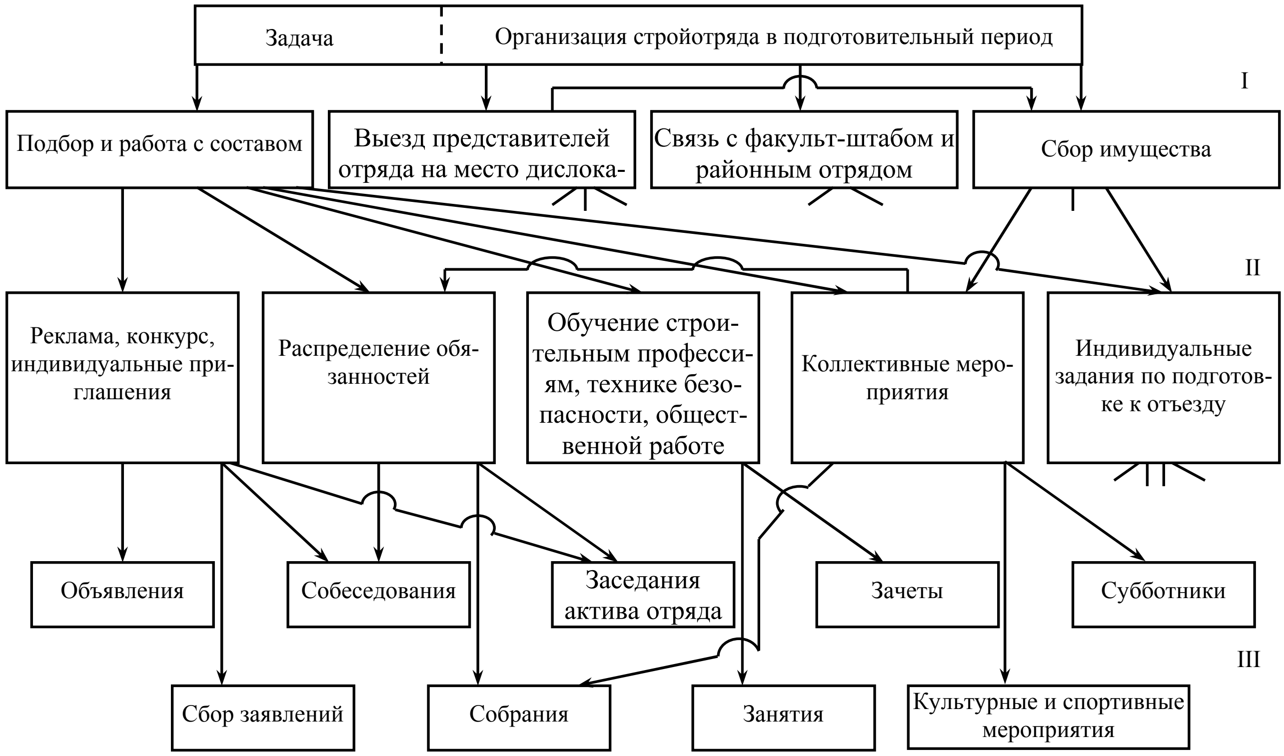

Рис. 2.2. Фрагмент иерархического деления задачи

В соответствии с общим видом системы (1.1) организация процесса решения задачи формально может быть записана как

$$R:\{\{M\},\{x\},F\}, \tag{2.1}$$

где $\{M\}$ – множество действий по решению задачи; $\{x\}$ – множество связей между действиями; F – формулировка поставленной задачи (цель).

Кортеж (2.1) обладает всеми особенностями записи произвольной системы. Он также условен в том смысле, что способ описания цели, действий и связей, их принадлежность определенным классам должны быть конкретизированы отдельно. В самом общем виде с составляющими кортежами $\{M\}$ и $\{x\}$ возможны лишь те операции, которые допустимы с множествами произвольной природы, например: дополнение, разделение, пересечение и др.

Обратим внимание также на то, что элемент M в (2.1) определен как действие, а не как простейший акт решения – процедура. Это сделано для большей вариативности формальной записи. Так, если $\{M\}$ – это операции верхнего иерархического уровня, то (2.1) представляет собой вполне обозримую, хотя и грубую схему решения задачи. Именно такие схемы чаще всего будут фигурировать в качестве примеров в данной книге. Если же $\{M\}$ – это все процедуры в решении, то для достаточно сложной задачи расшифровка всех элементов M и x может быть весьма объемной. Такая расшифровка, например, требуется при передаче (тиражировании) способа решения какой-либо задачи. Документация на стандартизированное описание даже среднего по сложности программного средства может занимать до сотни страниц текста и обозначений. Документация же, связанная с описанием всех процедур по строительству самолета или ракеты, достигает в весовом выражении десятков тонн бумаги.

В сложной системе, кроме введенных понятий процедуры, операции, действия, употребляются и другие термины. Так, большой комплекс действий, приводящий к выполнению в определенном смысле обособленной важной части всего решения задачи, называют «направлением работ». Для деления действий во времени употребляют термины «стадия» и «этап». Однако основным системным понятием в данной главе останется операция. Разложение процесса решения только на направления работ или стадии и этапы практически всегда есть еще недостаточное углубление в суть задачи.

2.1.2. Основные характеристики действий

Любое действие имеет три основные характеристики:

1) цель действия;
2) описание действия;
3) способ его выполнения.

Все эти характеристики можно представить в виде вопросов, которыми и будем пользоваться ниже. Цель (назначение) в зависимости от ситуации удобно обозначить вопросом «Зачем?» или «Что должно быть?» (т.е. каков должен быть результат). Описание действия или представление о его осуществлении обозначим вопросом «Что делать?», а способ выполнения действия (умение и возможность выполнить его) – вопросом «Как делать?». Коротко эти вопросы-характеристики будут звучать так: «Зачем?», «Что?», «Как?».

При продумывании решения задачи чаще всего используются три способа организации отдельных действий, которые удобно связать с ответами на поставленные вопросы (рис. 2.3). Первый из них отражает случай, когда вопросы «Зачем?» и

«Что делать?» рассматриваются исследователем совместно. Во втором ее случае отталкиваются от того, что, в принципе, можно делать. Третий вариант является наиболее строгим с точки зрения формальной логики.

Приведенные вопросы являются основными. Однако они не исчерпывают всех аспектов организации действий. Ее дополнительные стороны охватываются, например, вопросами «Удобно ли (технологично) выполнение данного действия?», «Каковы вторичные и неочевидные последствия данного действия?», «Где будет выполняться действие?», «Когда?», «Кем?».

Исследование вопросов «Зачем?», «Что?», «Как?» – это типовой и универсальный путь рассмотрения действий. Однако существует несколько распространенных исключений, когда можно избегать или не интересоваться ответами на один или даже два из них. Все эти ситуации оказываются в каком-то смысле крайними, особыми. Их рассмотрение лишний раз убеждает, что только продумывание ответов на все три вопроса ведет к уверенности в правильной организации действий.

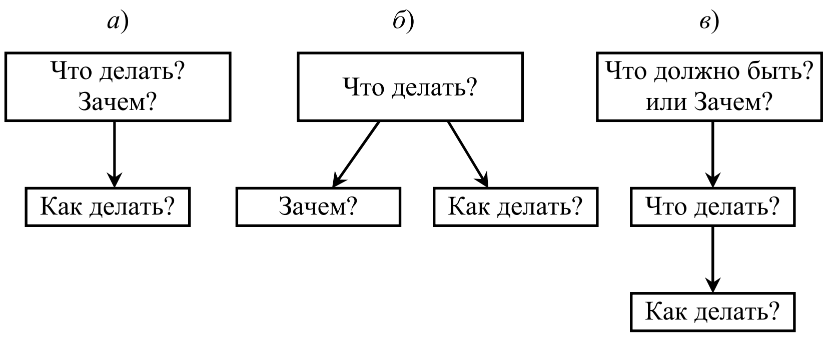

Рис. 2.3. Типовые приемы организации отдельных действий

Первый случай – это ограничение только вопросами типа «Как?». Таким способом можно решать задачи по не допускающей отклонения инструкции. Если имеется точное описание действий, и мы уверены, что оно подходит для решения данной задачи, то, вообще говоря, при исполнении отдельных действий можно не понимать ни что делается, ни зачем. Так можно выполнять регламентные работы, провести по описи химическую реакцию. Можно по командам инструктора управлять транспортным средством, а по методическому описанию вставить в программу для ЭВМ управляющие (системные) команды. Естественно, такие действия несовместимы с исследовательским подходом к проблеме. Они не годятся для решения новых или нестандартных задач. Подчеркнем, что всевозможные инструкции удобны только внутри жестко фиксированных и, как правило, очень детализированных ограничений. Их использование приводит к экономии усилий и времени в ситуациях, когда нет необходимости в понимании выполняемых процедур.

Второй случай – это организация действий на основе перебора, пренебрежение вопросом «Зачем?». Типичные примеры: случайные действия по обнаружению неисправности (пока не найдем); выпуск товара мелкими партиями для изучения спроса; последовательное использование различных программных средств, решающих одну и ту же задачу, с целью поиска наиболее приемлемого. На практике такой способ выбора действия используется, когда перебор относительно невелик и более удобен, чем исследование ответов на вопросы «Зачем?». Им пользуются и тогда, когда ответа на этот вопрос дать не удается.

Полезно иметь в виду, что выбор действия (решения) перебором обычно присутствует в любом исследовании, проводимом человеком: мы что-то пробуем и так, и эдак, выбираем... по-видимому, перебор типичен для человеческих действий вообще. Но его возможности всегда резко идут на убыль с ростом сложности задачи.

Третий вариант неполного списка наших вопросов – это пренебрежение анализом того, реализуемо ли данное действие на практике (т.е. пренебрежение вопросом «Как?»). Такие действия и их совокупности называют абстрактной схемой, неконструктивным подходом, оторванным от практики, или схоластическим решением. Прежде всего сюда относятся действия высокого уровня общности, в которых ответы на вопрос «Как?» исследуются дополнительно уже для данной конкретной задачи. О полезности таких усеченных, оторванных от непосредственной реализации действий мы будем говорить ниже (см. пп. 2.1.4 и 2.1.5). К действиям без «Как?» относятся и те, которые в настоящий момент никто не знает как выполнять – нет соответствующей технологии, материалов с требуемыми свойствами, ЭВМ с требуемыми ресурсами или соответствующего банка данных и т.д. В этом случае типичны поиск осуществимого набора действий, заменяющего данный, или изменение (как правило, упрощение) задачи.

2.1.3. Локальные цели

Наиболее распространенная схема организации действий по решению задачи изображена на рис. 2.4. В этом случае мы имеем дело с тремя уровнями организации решения. Легко заметить соответствие рис. 2.4 и 2.3, *в*. Обратим внимание и на то, что первая и вторая ячейки на рис. 2.4 названы системами. Действительно, ответы на вопросы «Что должно быть?» связаны друг с другом и только в совокупности определяют путь решения задачи. То же самое можно сказать и про ответы на вопросы «Что делать?». Нижняя же ячейка, вообще говоря, не является системой, так как ответы на вопросы «Как?» не обладают внутренними связями. Замена одного способа выполнения действия другим обычно не приводит к каким-либо изменениям в решении задачи, в то время как замена ответа на вопрос «Что?» или «Зачем?» влечет изменение и других действий и целей.

В этом пункте остановимся на обсуждении системы локальных целей, которая в определенной степени может быть, оторвана от всего процесса решения задачи и рассмотрена отдельно. Работа с системой локальных целей как первая ступень решения широко применяется на практике. Термин «целеопределение» используется при проектировании сложных технических объектов, в экономике, экологии, других областях. Типичный пример выделения системы локальных целей приведен на рис. 1.6, демонстрирующем целеопределение при проектировании самолета.

При разработке системы локальных целей для облегчения дальнейшего решения задачи полезно в общих чертах прикидывать и ответы на вопросы «Что делать для достижения цели?», «Как это делать?».

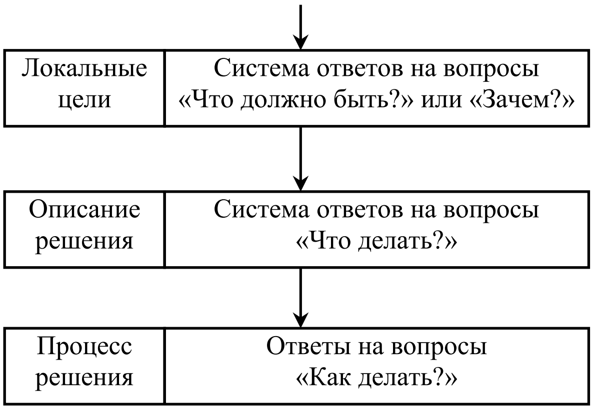

Рис. 2.4. Типичный вариант организации системы действий по решению задачи

Однако в сложных системах это позволило лишь значительно уменьшить, но не исключить вообще необходимость доработки (коррекции) системы локальных целей для перехода к следующей стадии – описанию действий. Может оказаться, что требуемые действия длительны, дорогостоящи, не удобны с других точек зрения. Как крайний случай – какая-либо локальная цель окажется невыполнимой вообще. Из этого следует, что окончательная доводка системы локальных целей достигается только после рассмотрения ответов и на вопрос «Что делать?», «Как делать?», а условный отрыв целей от задачи есть не более чем первая, предварительная стадия решения.

Понятие локальной цели будем относить и к операции, и к отдельным процедурам. Для операций, особенно высокого уровня общности, будем говорить о цели операции, а термин «глобальная цель» оставим только за решением всей задачи целиком.

Выделим совокупность локальных целей $\{g_J^s\}$, обеспечивающих выполнение цели G_J операции J,

$$\{g_J^s\} \to G_J,\ s = 1, 2, ..., \sigma. \tag{2.2}$$

Данная запись означает лишь достаточность выполнения $\{g_J^s\}$ для осуществления цели G_J; может существовать и другая совокупность локальных целей, влекущая ту же цель G_J. Основной смысл записи (2.2) состоит в том, что можно «забыть» про цель G_J и выполнять более простые цели $g_J^1, g_J^2, ..., g_J^\sigma$.

Приведем п р и м е р . Задача организации САПР в отделе, выпускающем конструкторскую документацию, стандартно разбивается на следующие цели:

1) техническое оснащение отдела вычислительной техникой и графопостроителями;
2) заимствование и создание программного обеспечения вычислительной техники;
3) создание банка данных;
4) обучение и переквалификация кадров;
5) дополнительные локальные цели – стыковка нового способа работы с другими отделами, изменение нормирования труда, введение нового объема загрузки на отдел и др.

Здесь имеем разбиение на цели по функциональному признаку. Каждая из четырех первых целей будет в терминах п. 2.1.1 направлением работ; далее они разбиваются на отдельные локальные цели для операций и групп операций. Но в этой же задаче возможно разделение на цели на основе другого, например, временного (стадийного) признака. В этом случае система целей $\{g\}$ будет такой:

1) сбор информации о возможности применения вычислительной и другой техники в отделе;
2) концептуальная организация банка данных (какая именно информация будет храниться, ее деление на группы, требуемые формы вопросов и ответов, принципы поиска информации и др.);
3) обучение и переквалификация кадров, поставка техники, создание банка данных;
4) внедрение ЭВМ и графопостроителей в режиме выполнения изолированных работ с частичным использованием банка данных;
5) системное применение технических средств на отдельных заданиях – пробная эксплуатация;
6) основная эксплуатация с непрерывным усовершенствованием и развитием.

Такие (на практике более развитые и детализированные) типовые схемы выделения целей значительно облегчают решение новых задач. Это особенно ощутимо, если схема учитывает специфику данной организации, т.е. носит типовой характер для данной отрасли, вида выпускаемого изделия, научного направления.

Систему локальных целей принято создавать сверху, с введения набора целей 1-го иерархического уровня. Декомпозиция целей должна сопровождаться их согласованием, чтобы выполненные все вместе, они привели к достижению глобальной цели. Обсуждению проблем согласования, т.е. связи между отдельными целями, и посвящен следующий пункт.

2.1.4. Связи между локальными целями

В общем случае структура связей между локальными целями имеет произвольный характер. Как крайние ситуации назовем:

а) случай, когда выполнение любой цели связано с выполнением каждой из остальных (при отсутствии иерархии);

б) случай полной независимости достижения локальных целей: каждая цель выполняется самостоятельно и их связь друг с другом проявляется лишь в том, что выполненные все вместе они решают поставленную задачу.

Однако простейших и одновременно основных типов связей между целями всего три. Удобнее всего пояснить их на примере деления цели операции на две локальные цели:

$$\{g^1, g^2\} \to G_J ,$$

(индекс J у локальных целей для простоты опущен). При этом возможно:

а) последовательное выполнение – только достижение одной из целей дает возможность выполнить другую;

б) параллельное выполнение – цели могут выполняться независимо;

в) циклическое выполнение – частичное выполнение одной из целей позволяет частично выполнить другую, что, в свою очередь, позволяет вернуться к выполнению первой, и так до полного выполнения обеих целей (рис. 2.5).

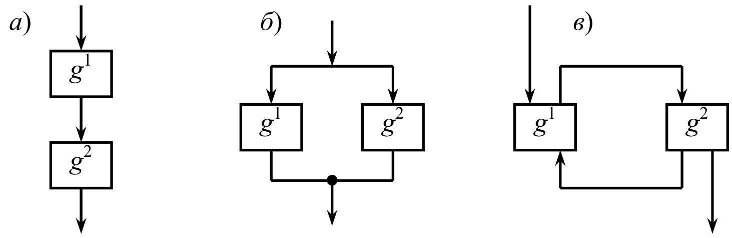

Рис. 2.5. Способы достижения двух целей

Цели на рис. 2.5 называются также связанными (*а*), несвязанными (*б*) и сложно связанными (*в*). Последнее с точки зрения кибернетики является примером системы с обратной связью.

Типичным примером циклического способа (рис. 2.5, *в*) является организация цикла в программном средстве. В этом случае перед ЭВМ ставятся две локальные цели: перебрать все параметры цикла (цель g^1) и выполнить для каждого значения параметра определенные действия (g^2). Циклическое выполнение целей весьма многообразно и вне области программирования. По этой схеме представимо, например, любое управление, требующее постоянной выработки команд: цель g^1 – определение управляющего воздействия, цель g^2 – исполнение этого воздействия. Строительство можно рассматривать как цикл: завоз материала и механизмов и собственно строительные работы. Процесс обучения для студента преследует две циклические цели – усвоение знаний и сдачу зачетов.

Способ достижения каждой из целей g^1 и g^2 в отдельности может быть дискретным (порциями, скачками) и непрерывным. В первом случае схему рис. 2.5, *в* также называют итеративной, а каждый переход от цели g^1 к g^2 и обратно – итерацией, шагом, циклом.

Для более чем двух локальных целей связь между ними будет комбинированием приведенных выше типов. Схематические примеры некоторых из них для случая трех локальных целей изображены на рис. 2.6.

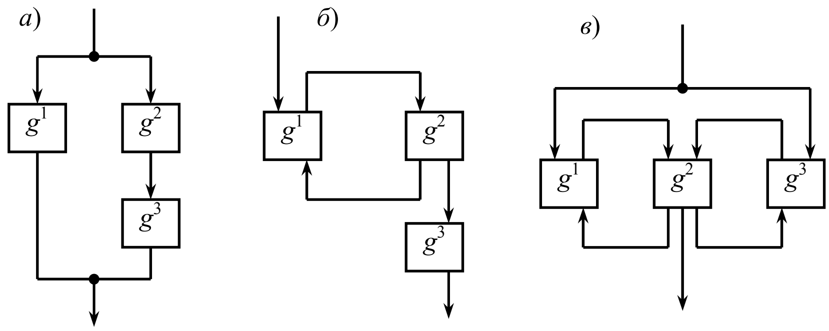

Рис. 2.6. Некоторые способы достижения трех целей

Приведем пример теоретического использования знаний о последовательном и параллельном способах выполнения целей. Среди математических шуток имеются задачи, подобные следующей: четверо рабочих могут собрать щитовой домик за 10 ч, за сколько соберут этот дом 400 рабочих? Неужели за 6 мин? Серьезный ответ на встречающиеся на практике аналогичные вопросы гласит, что стоящая перед нами цель (в данном случае – сборка дома) в весьма ограниченной степени делится на параллельно реализуемые цели.

Нередко выполнение одной локальной цели может затруднить и даже исключить выполнение другой. Такие цели (две и более) называют антагонистическими. В сложных системах практически не удается избавиться от той или иной степени антагонистичности локальных целей. Проблема является наиболее острой для целей одного и того же иерархического уровня. В этом случае задачу принято называть многоцелевой или многокритериальной. Этот вид задач является весьма актуальным и в настоящее время активно изучается [7 – 11]. Обычно это делается в строго формализованной постановке, переводящей центр тяжести исследования на математический аппарат. В этой книге значительное место таким задачам уделено в гл. 3.

Отталкиваясь от типовых порядков выполнения локальных целей, проанализируем наиболее распространенные на практике схемы достижения глобальной цели.

Типовая схема I. Общая характеристика: в задаче явным образом присутствуют стадийность, этапы, очередность операций или процедур.

Примеры: последовательность операций по обработке детали на станке; этапы изготовления проектной документации; выращивание растений; очередность изучения разделов учебника.

В схеме I преобладает последовательный переход от одной локальной цели к другой. Стадийность решения обычно обеспечивает простоту условий перехода от одной цели к другой и, таким образом, облегчает их согласование.

Типовая схема II. Общая характеристика: задача имеет сильно связанные и резко различающиеся стороны и аспекты, которые должны рассматриваться одновременно.

Примеры: проектирование сложного технического объекта, требующее участия различных исследовательских коллективов и приложения больших научных, инженерных и организационных усилий; задача наземного сопровождения космического полета; конвейерная сборка, в которой одновременно участвует высококвалифицированный персонал различных специальностей; задача экологической защиты; задача управления обществом. Основой решения таких задач является параллельное достижение набора локальных целей в условиях их тесной связи и антагонистичности. Выделение целей здесь, как правило, довольно очевидно по функциональному признаку. Но согласование является серьезной, часто трудно решаемой проблемой. Удачный (т.е. хорошо согласованный) выбор локальных целей в задаче, скажем, создания сложной автоматизированной системы может являться значительным достижением, имеющим характер изобретения или даже открытия.

Типовая схема III. Общая характеристика: большая однородная задача, подлежит делению на части в связи с ее громоздкостью, значительным объемом входной информации или ограничением времени на решение.

Примеры: распределение срочного заказа по заводам (в основе решения задачи – параллельное выполнение слабо связанных целей); создание разветвленной информационной системы на основе сети ЭВМ (в основе – параллельное выполнение сильно связанных целей); поиск информации в банке данных (в основе – последовательный переход вниз по древовидной системе признаков); математическая декомпозиция линейной системы уравнений большой размерности (в основе – циклическое решение систем меньшей размерности). Из примеров видно, что локальные цели здесь связываются с выделением достаточно однородных частей. В случае выполнения целей одним коллективом (одним человеком, одной ЭВМ) преобладает последовательное или последовательно-циклическое достижение целей (типа рис. 2.6, *б*). Если же цели выполняются различными коллективами (людьми, машинами), то преобладает параллельное или параллельно-циклическое достижение (типа рис. 2.6, *в*).

В этой типовой схеме сложным может быть как выделение, так и согласование целей. (Пример простого деления и согласования – это разбивка годового задания по кварталам, месяцам и т.д.).

2.1.5. Система действий. Операционные модели

Снова обратимся к рис. 2.4 и напомним, что в предыдущих пунктах говорилось о выделении локальных целей как о первом шаге построения системы действий. Теперь обсудим построение системы ответов на вопрос «Что делать для выполнения локальных целей?» (средняя ячейка на рис. 2.4). Эти ответы составляют описание действий.

Уже отмечалось, что существует тесная связь между содержанием средней и верхней ячеек (рис. 2.4). Во-первых, опытный разработчик мыслит категориями только принципиально осуществимых целей, чем сводит к минимуму проблемы выбора действий после фиксации согласованных целей. Во-вторых, нередко ответ на вопрос «Что делать для осуществления данной цели?» ищется непосредственно после выбора локальной цели, что позволяет говорить об одновременном создании системы целей и организации действий. В-третьих, формулировка цели часто сама уже указывает на действия по ее выполнению (как это имело место у нас в примере с созданием САПР в п. 2.1.3).

То же самое можно сказать и о последней стадии построения системы действий – ответах на вопрос «Как?» (нижняя ячейка на рис. 2.4). Способы выполнения действий (процесс решения) также полезно продумывать на начальной стадии построения системы действий. Все это позволяет перейти к обсуждению совокупности действий в целом.

Создание системы действий в достаточно сложной задаче представляет собой в значительной степени неформализованный процесс. В нем необходимо учитывать как специфику задачи, ее предметно-понятийную (техническую) и научную сферы, как и сведения о системном применении знания, моделировании в целом, математической и другой формализации. Можно утверждать, что общих приемов, позволяющих составлять подробную систему действий в любой конкретной задаче, не существует. Различные системы действий, безусловно, обладают рядом общих, безотносительных к характеру задачи свойств, но эти действа лишь в самых общих чертах определяют организацию действий. Между такими системными сведениями и их практическим применением существует значительный разрыв. Он преодолевается работой исследователя, выступающего интерпретатором обобщенного знания и одновременно носителем конкретного, нужного в данной прикладной проблеме.

Построение системы действий облегчается использованием типовых схем действий, разработанных для отдельных узких, а иногда и достаточно широких классов задач. Такие схемы называют операционными моделями (операционными диаграммами, схемами, технологическими линиями, маршрутами). Эти модели, состоящие из набора связанных операций (процедур), представляют собой описания типовых путей решения задач.

Операционными моделями являются всевозможные методики, инструкции, программы и алгоритмы действий, указанные последовательности операций. Высокий уровень общности демонстрируется в таких операционных моделях, как типовой САПР отрасли или главка, типовая АСУ «Бухгалтерия», типовой ГАП инструментального цеха.

Как любое типовое (усредненное) решение, операционная модель требует к себе критического, системного отношения исследователя, нуждается в «настройке» на данный конкретный случай. С другой стороны, такая модель ориентирует в ситуации, позволяет использовать имеющийся опыт, заимствовать удачно подобранные и согласованные операции. При обмене опытом воспринимается именно операционная модель или ее элементы.

Однако еще раз укажем на опасность бездумного использования предложенного пути решения. Требование «подгонки» способа действий относится даже к такому строго оговоренному виду операционных моделей, каким является алгоритм для ЭВМ. Современные алгоритмы включают в себя внутренние настроечные параметры, разумный выбор которых в значительной, а иногда и решающей степени, приспосабливает алгоритм к задаче.

Итак, операционная модель – это определенный, достаточно общий вид системы действий, годный для передачи и тиражирования. Практически все, что говорится о системе действий, относится и к операционным моделям. Однако в целях общности будем впредь в основном употреблять термин «действие».

2.1.6. Запись структуры действий

Первым, наиболее распространенным способом записи структуры действий является ее изображение в виде графической схемы (см., например, рис. 2.2). Элементами в такой схеме являются ячейки и соединяющие их линии. Ячейки, как правило, соответствуют действиям, но они могут указывать и на используемые источники или приемники данных, на технические средства, документы, могут представлять собой краткие комментарии, быть указателями межстраничных переносов схемы и др. В ячейке или около нее могут быть записаны ее важнейшие характеристики. Различным типам и видам действий могут соответствовать разные по форме ячейки. Так, в ГОСТе на изображение символов в схемах алгоритмов и программ для ЭВМ введено более 20 типов действий: пуск и останов, действие-процесс, логический выбор (решение), ручная операция, вспомогательное действие, ввод-вывод данных, использование дисплея, оперативной памяти и др.

Соединительные линии могут соответствовать очередности действий, путям передачи информации и управления, согласованности ячеек и др. Они могут быть разными по виду (толщине, цвету, конфигурации), чем кодируется интенсивность передачи информации, значимость связи, особые варианты работы и т.д. Возле линий иногда пишут числовые или вербальные характеристики связи. Стрелками на линии можно указывать строгую подчиненность или последовательность во времени и предпочтительное (доминирующее) направление связи.

Отдельно остановимся на таком важном виде действий, как логический выбор (условный переход). Именно он позволяет записывать структуру действий, пригодную одновременно для ряда способов ее осуществления. Этот элемент принято изображать в виде ромба. Его значимость и способ употребления хорошо иллюстрирует рис. 2.7, изображающий схему действий по решению квадратного уравнения с вещественными коэффициентами. Эта, на первый взгляд, простая операция требует пяти условных переходов, а в результате ее получается шесть видов выходной информации (нижние ячейки).

Вообще говоря, логический выбор может иметь более чем два исхода. Такой случай изображен на рис. 2.8. (Заметим, что тройной и более выбор всегда может быть заменен набором обычных двойных условных переходов. В данном примере можно сначала поставить условие «частота ≤20 Гц», а затем, при его неудовлетворении, перейти к условию «частота ≤100 Гц»).

Графические схемы есть развитие понятия графа (граф с различными и неформальными понятиями элементов и связей). Поэтому схемы могут быть упрощены до уровня графов и исследоваться математическими методами, их терминология во многом совпадает. Можно говорить об ориентированных и неориентированных (со стрелками и без) связях, о тупиковых и начальных элементах (из которых никуда нельзя попасть или в которые нельзя попасть из схемы действий), о среднем числе связей на один элемент, о петлях и контурах в схеме, о сильно связанных частях схемы и т.д.

Графические схемы могут иметь различные степени детализации – от весьма грубых для первого знакомства с системой до максимально подробных для использования специалистами разработчиками. Постоянная работа с графическими схемами (алгоритмами действий) вырабатывает навык оперировать категориями схемы, создавать сначала именно структуру. Считается, что в сложных задачах это наиболее эффективный способ работы с системой. В частности, он проявляется в том, что именно на структурной схеме удобно найти место для внесения изменений, тут же выяснить, на что это повлияет, а затем воплощать задуманное изменением текста программного средства, добавлением новых элементов и связей в радиотехническую схему, переналадкой управляющих механизмов и т.д.

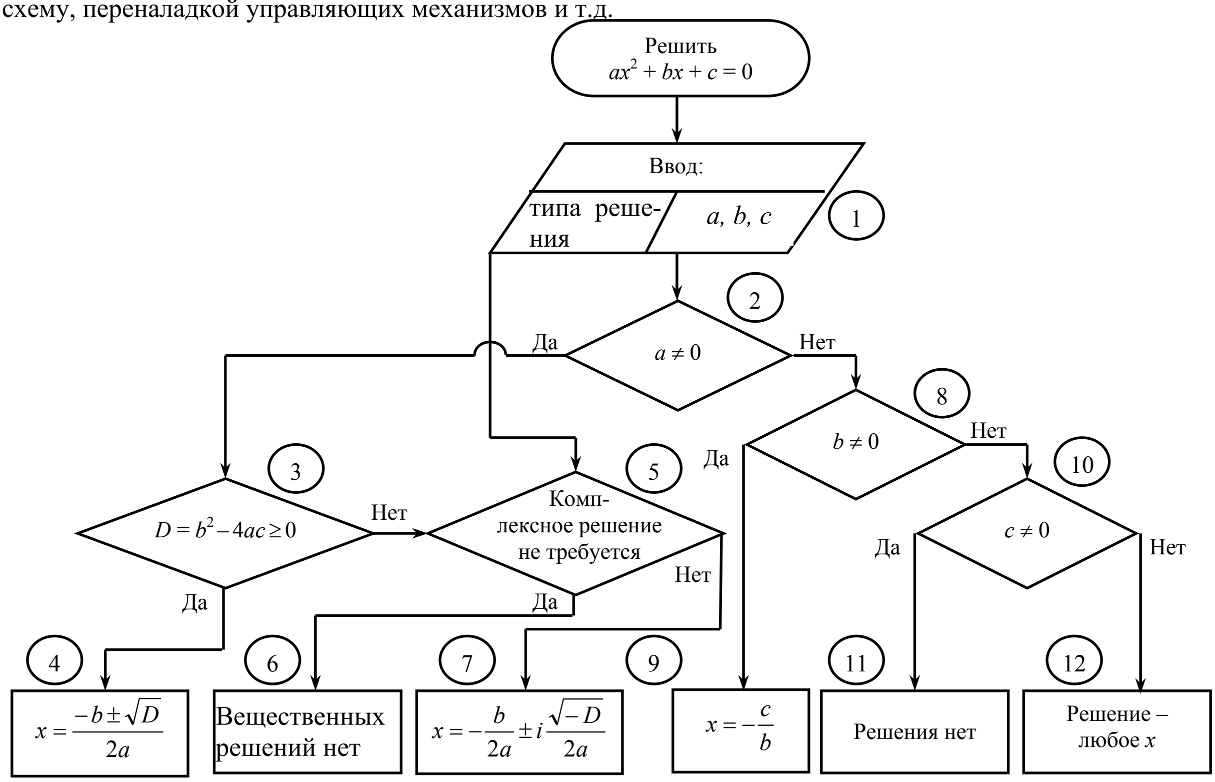

Рис. 2.7. Схема действий по решению квадратного уравнения

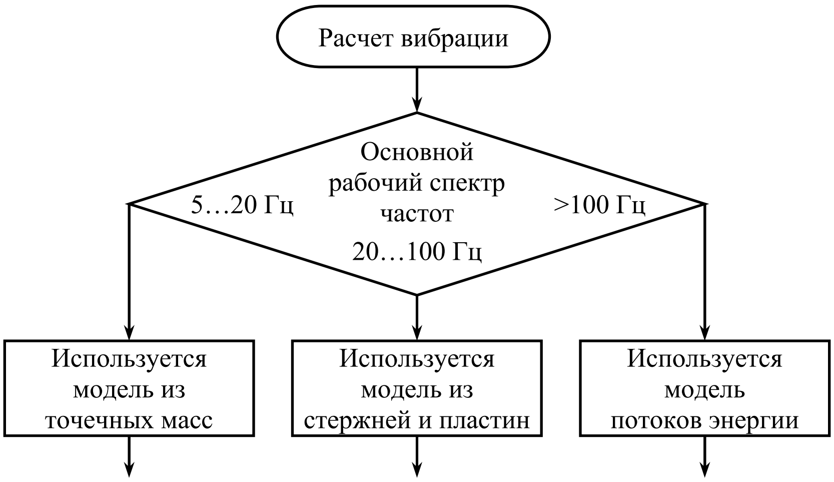

Рис. 2.8. Пример тройной логической связи

Второй способ записи структуры действий состоит в построении квадратной матрицы специального вида. Размерность матрицы равна числу элементов структуры. Рассмотрим этот способ на примере действий по решению квадратного уравнения. Напомним, что в графическом виде эти действия изображены на рис. 2.7. Их матричная запись представлена на рис. 2.9.

Нумерация действий (см. рис. 2.7) соответствует номерам строк и столбцов в матрице. На главной диагонали помещены условные обозначения типов элементов: ВВ – ввод данных, У – условный переход, ВЫЧ – вычисления, И – информационное сообщение. Если от элемента *i* имеется связь к элементу *j*, то на пересечении *i*-й строки и *j*-го столбца стоит символ, которым обычно шифруется основное содержание связи. У нас символ со стрелкой обозначает передачу (вызов) данных, символы «+» и «–» – выполнение или невыполнение условий перехода («да» или «нет» на схеме рис. 2.7).

Матрица на рис. 2.9 будет полностью описывать действия по решению квадратного уравнения, если дополнительным перечислением раскрыть содержание ее элементов. Таким образом, матричная запись указывает, как правило, на основные характеристики системы действий. Проанализируем некоторые из них. Наша матрица на рис. 2.9 получилась, во-первых, разреженной. В ней мало символов, соответствующих связям: из возможных 132 имеется всего 12. Это говорит о простоте схемы действий. Во-вторых, – треугольной. Правда, при неудачной нумерации элементов мы бы получили символы под главной диагональю, но возможность треугольной записи говорит об отсутствии контуров (замкнутых путей) в схеме действий. В-третьих, – с одним столбцом, содержащим более двух символов (столбец № 5). Это говорит о том, что система действий не является строго древовидной, но весьма близка к ней. В-четвертых, – с шестью строчками, содержащими только диагональные символы, что означает наличие шести выходов из системы действий. Этим далеко не исчерпывается анализ матричной записи.

	1	2	3	4	5	6	7	8	9	10	11	12
1	ВВ	↴			↴							
2		У	+					–				
3			У	+	–							
4				ВЫЧ								
5					У	+						
6						И						
7							ВЫЧ					
8								У	+	–		
9									ВЫЧ			
10										У	+	–
11											И	
12												И

Рис. 2.9. Матричная запись структуры действий по решению квадратного уравнения

Особо примечательна возможность хранения и анализа матрицы действий с помощью ЭВМ. В машине может содержаться и полная расшифровка ее элементов. Графическая запись традиционно считается более наглядной, однако исследователи, имеющие навык работы с матричной записью, нередко оспаривают это мнение, что выглядит особенно аргументированным при сильной связанности системы действий.

Третий способ записи структуры состоит в перечислении всех действий с указанием приходящих и исходящих связей. Он называется стратовым описанием системы (описанием через слои, срезы) и является основным при фиксировании или передаче разработанной системы действий. Графическое и матричное изображения при этом обычно выступают в качестве огрубленной иллюстрации или указателя (оглавления) приводимых далее стратов. С другой стороны, стратовое описание обычно присутствует и тогда, когда основными считаются схемы или матрицы. Оно состоит в дополнительных пояснениях к изображенным элементам и связям.

Примеры. Стратами являются стандартные списки спецификаций, используемые в проектировании (указания, куда, что и в какой форме идет, откуда и в каком виде поступает). Аналогично обстоит дело при создании программных комплексов. При их тиражировании или сдаче в фонды алгоритмов и программ все отдельные программные средства должны быть стандартно описаны и привязаны к их месту в комплексе. В частности, должно быть указано, какие данные нужны в каждом отдельном средстве и откуда они берутся.

Полное стратовое описание включает в себя формулировки условий осуществления каждого действия, его окончания и передачи результатов. Последнее может потребовать специальных действий по преобразованию выходов к нужному виду. Такие процедуры могут как выделяться из системы действий, так и включаться в нее. В любом случае их принято называть (как мы уже отмечали в главе 1) интерфейсной адаптацией.

Контрольные вопросы

1. Приведите пример деления задачи на процедуры и действия.
2. Перечислите основные характеристики действий.
3. Охарактеризуйте локальные цели.
4. Расскажите о структуре связей между локальными целями.
5. Ответьте на вопрос, что такое операционные модели?
6. Приведите пример записи структуры действий.

2.2. ПРОБЛЕМА ПРИНЯТИЯ РЕШЕНИЯ

Особое место среди всевозможных действий по решению задачи занимает их специальный вид – принятие решений. Его можно определить как преодоление альтернатив.

Достаточно очевидно, что такие действия типичны при разработке или усовершенствовании какой-либо системы. Однако не менее важно принятие решений и при функционировании уже созданной и отлаженной системы (как, например, в автоматизированных системах).

Подчеркнем, что принятие решения – отнюдь не прерогатива человека. Оно может быть произведено и ЭВМ, и целым рядом других технических средств – от простейшего регулятора до следящих радиолокационных сетей.

Однако все многочисленные ситуации принятия решений могут быть охвачены единым подходом. Именно этому посвящен данный раздел.

2.2.1. Постановки задачи принятия решений

С преодолением альтернатив связаны два фундаментальных понятия: множество альтернатив (вариантов действий), которое обозначим через $\{\chi\}$, и принцип выбора, который обозначим через Φ. Задача принятия решения может быть записана как

$$\{\{\chi\}, \Phi\} \to \chi^*, \tag{2.3}$$

где χ^* – выбранные альтернативы (одна или более).

В зависимости от степени формализации введенных понятий различают три задачи.

1. Задача оптимального выбора: если множество $\{\chi\}$ однозначно определено (фиксировано), а принцип выбора Φ формализован, т.е. может быть описан, передан и результаты его применения к элементам из $\{\chi\}$ не зависят от субъективных условий.

2. Задача выбора: если множество $\{\chi\}$ однозначно определено, но принцип выбора Φ не может быть формализован или даже фиксирован. В этом случае выбор зависит от того, кто и на основе какой информации его делает.

3. Общая задача принятия решения: если множество $\{\chi^*\}$ не имеет определенных границ (может дополняться и видоизменяться), а принцип выбора Φ неформализован или даже не фиксирован. В этом случае разные субъекты могут выбирать в качестве решения те альтернативы, которые другими субъектами и не рассматривались, а один и тот же субъект при использовании одного и того же принципа выбора (неформализованного, но для него существующего) может изменять свое решение при обнаружении им новой альтернативы.

С формальной точки зрения может показаться, что последняя задача является настолько расплывчатой, что теряет смысл. Можно утверждать, что мы не знаем, ни из чего выбирать, ни чем при этом руководствоваться. Однако именно эта задача с некоторыми естественными ограничениями наиболее типична для практики.

Каковы же эти естественные ограничения?

Во-первых, в реальной задаче, как правило, существует так называемое начальное множество альтернатив $\{\chi^{(0)}\}$, на основе которого приступают к принятию решения. В дальнейшем это множество изменяется, но можно считать, что на любой момент процесса принятия решения мы имеем дело с фиксированным множеством $\{\chi^{(i)}\}$:

$$\{\chi^{(0)}\} \to \{\chi^{(1)}\} \to \ldots \to \{\chi\}.$$

Во-вторых, подразумевается, что любая альтернатива $\hat{\chi}$ из множества всех мыслимых альтернатив $\{\hat{\chi}\}$ может быть оценена с точки зрения полезности ее включения в $\{\chi\}$. Это делается при помощи некоторого вспомогательного принципа выбора $\hat{\Phi}$. Чаще всего этот принцип неформализован. Таким образом, и само множество $\{\chi\}$, вообще говоря, является итогом задачи принятия решения:

$$\{\{\hat{\chi}\}, \hat{\Phi}\} \to \{\chi\}. \tag{2.4}$$

В-третьих, считается, что существуют хотя бы неформализованные принципы выбора, относящиеся к принимаемому решению. Часто (но не всегда) есть уверенность, что применение таких принципов различными субъектами дает пересекающиеся или в каком-то смысле близкие результаты.

В перечисленных условиях общая задача принятия решения 3 становится обозримой и пригодной для попыток решить ее в той или иной степени обоснованно.

Практические пути решения не полностью определенных задач 3 и 2 состоят в использовании для этой цели ряда задач с фиксированным, но меняющимся от задачи к задаче множеством $\{\chi\}$ и фиксированным (хотя необязательно формализованным) принципом выбора Ф. Это происходит с применением ряда приемов. Первый из них – организация итерактивного процесса решения набора задач вида 1. Она состоит в начальном решении одной или нескольких формализованных задач, экспертного анализа их решения, назначения измененных множеств альтернатив $\{\chi\}$ и измененных принципов выбора Ф, нового решения набора задач и т.д. до достижения удовлетворительного решения. Другой прием заключается в решении ослабленного варианта задачи 1, когда принцип выбора формализован не полностью, а допускает участие экспертов, каждый из которых по-своему, обычно неформальным образом, фиксирует принцип Ф. В этом случае любой из экспертов порождает свою задачу типа 1, а решение исходной задачи формируется на основе их решений. Следующий прием близок к первому. Здесь задаче 3 или 2 сопоставляется ее некоторый аналог, выбранный среди задач 1, а полученное решение служит основой для неформального поиска решения требуемой задачи.

В целом можно сказать, что ядром задачи принятия решения остается задача оптимального выбора 1. Полезно считать, что в общей задаче принятия решения отнюдь нет «абсолютной свободы» для множества $\{\chi\}$ и принципа Ф, а есть лишь допущение разумности выхода за пределы формализмов, которые использовались на стадиях решения задач.

2.2.2. Декомпозиция задачи принятия решения и оценка свойств альтернатив

Общепринятым принципом, который облегчает принятие решения, является переход от сравнения альтернатив в целом к сравнению их отдельных свойств (аспектов, характеристик, признаков, преимуществ). Основная идея такого перехода состоит в том, что в отношении отдельного свойства существенно легче сказать, какая из альтернатив предпочтительней.

Так, обращаясь к рис. 1.6 (п. 1.1.5) и понимая его на этот раз как задачу выбора наилучшего проекта самолета, можно гораздо более уверенно говорить, что проект *A* лучше проекта *B* по свойству комфортности или надежности, нежели о том, что проект *A* вообще лучше проекта *B*.

Сразу же заметим, что сравнение по отдельным свойствам порождает серьезные проблемы обратного перехода к требуемому сравнению альтернатив в целом. Эти проблемы мы будем обсуждать в следующем пункте.

Выделение свойств альтернатив является не чем иным, как декомпозицией. Свойства первого иерархического уровня могут делиться на следующие наборы свойств и т.д. Глубина такого деления определяется стремлением дойти до тех свойств, которые удобно сравнивать друг с другом. Так, в примере с самолетом из п. 1.1.5 говорить о комфортности, конечно, проще, чем о самолете в целом, однако такое свойство для сравнения также неудобно и требует дальнейшей декомпозиции. Ее пример будет приведен несколько ниже.

Сравнение альтернатив по отдельным свойствам может быть выполнено тремя способами:

а) на основе попарного (реже – группового) сравнения альтернатив по данному свойству;

б) на основе введения естественных числовых характеристик данного свойства;

в) на основе введения искусственных числовых характеристик данного свойства.

Разберем важнейшие свойства этих сравнений.

Попарное сравнение. Считаем, что для двух альтернатив χ_1 и χ_2 из $\{\chi\}$ мы каким-то образом можем произвести выбор наиболее предпочтительной по данному свойству. Способ выбора в общем случае не конкретизируется. Если он связан с использованием числовых характеристик, то такая ситуация относится к способу б) или в). Можно задаться вопросом: а существует ли объективный способ выбора, не связанный с числами? В строгой постановке этот вопрос, возможно, останется спорным. Но с практической точки зрения, мы считаем вполне объективными и не основанными на числовых характеристиках утверждения, что «это кресло более удобно», «этот вариант более способствует безавостной работе программного средства», «этот человек более удачно справится с поставленной задачей» и т.д. В реальных (в том числе технических) системах при принятии решения нередко приходится иметь дело именно с подобными сравнениями.

С формальной точки зрения, для альтернатив χ_1, χ_2 из $\{\chi\}$ вводится бинарная операция сравнения по признаку (свойству) *R*. Запись

$$\chi_1 R \chi_2 \tag{2.5}$$

означает, что альтернатива χ_1 предпочтительней (или, в несколько измененной трактовке, «не хуже») альтернативы χ_2 по признаку *R*. Указанная операция может быть применена как к любой паре (χ_1, χ_2) из $\{\chi\} \times \{\chi\}$, так и не ко всем из них. В последнем случае допускаем, что относительно некоторых пар нельзя сделать выбор. При этом говорится, что элементы множества $\{\chi\}$ лишь частично сравнимы по признаку *R*.

Операция бинарного сравнения для небольшого числа элементов удобно интерпретируется и анализируется с помощью графов. Вершинами графа является свойство *R* различных альтернатив, а дуги со стрелками указывают на предпочтения. Для операции R естественной является аксиома транзитивности: из $\chi_1 R \chi_2$ и $\chi_2 R \chi_3$ следует $\chi_1 R \chi_3$. Дополнительно могут быть введены аксиомы антисимметричности и антирефлексивности. Антисимметричность: из $\chi_1 R \chi_2$ и $\chi_2 R \chi_1$ верно лишь одно. Антирефлексивность: из $\chi_1 R \chi_2$ следует несовпадение альтернатив χ_1 и χ_2. Естественное отношение предпочтения антисимметрично и антирефлексивно. Естественное отношение «не хуже» этими свойствами не обладает. Для обозначения операции сравнения вместо записи (2.5) может использоваться запись $\chi_1 \succ \chi_2$ (предпочтение) и $\chi_1 \succeq \chi_2$ («не хуже»). На основе бинарного сравнения может быть выполнена специальная операция ранжирования (упорядочения). В результате ее выполнения альтернативы в зависимости от их свойства R располагаются в определенном порядке: от наиболее до наименее предпочтительной. Математически эта операция эквивалентна определенной перестановке.

Введение числовых характеристик. Сравнение элементов на основе сопоставления их числа представляется наиболее аргументированным способом выбора. Необходима лишь уверенность, что выполненное сопоставление объективно. Как правило, это имеет место, если числовая характеристика обладает физическим смыслом. Объективно сравнение по массе, размерам, скорости передачи информации, числу связей, времени готовности и многому другому. Так, в упомянутой выше задаче о проекте пассажирского самолета проведем декомпозицию свойства комфортности на:

а) уровень шумности в салоне;
б) уровень вибрации пола;
в) расстояние между креслами;
г) чувствительность и предельные условия работы системы искусственного климата и др.

Все эти характеристики выражаются в числах и объективны. Можно утверждать, что в задаче принятия решений следует стремиться довести композицию до уровней, на которых возможны численные оценки. Свойства, для которых существуют объективные численные характеристики, принято называть критериями. Таким образом, получение набора критериев – это наилучший итог декомпозиции. Он настолько привлекателен для практической реализации, что к его аналогу прибегают и тогда, когда естественные числовые характеристики отсутствуют. В этом случае вводят искусственные оценки типа баллов. Они проставляются экспертами (судьями, оценщиками, проверяющими, дегустаторами и др.), каждый из которых может исходить из своего неформального принципа выбора. Таким образом, решается задача количественной оценки качественных сторон явления или проблемы. Примеры искусственных числовых оценок весьма многочисленны. Они простираются от коэффициента трудового участия до баллов и суммы мест в фигурном катании, от разрядной сетки рабочих специальностей до экспертного определения процента износа механизмов.

Искусственные оценки практически непрерывно переходят в естественные. Так, процент износа может определяться на основе измерения зазоров, остаточного напряжения, времени наработки и других физических величин. В этом случае он будет естественной оценкой, подвергнутой специальному преобразованию в проценты. Такова же ситуация при выведении оценки экзаменационным автоматом, который превращает в баллы процент правильных ответов. В ряде случаев требования, запреты и рекомендации не являются такими, что определяют оценку полностью, и эксперт обладает определенной свободой выбора. Это имеет место при присвоении рабочих раз, рядов, назначении коэффициентов в эмпирически подобранные зависимости, определении внутренних параметров программного средства и т.д.

Дополнительным приемом, который в ряде случаев облегчает все три приведенных выше способа сравнения, является распределение элементов по подмножествам. В этом случае любая альтернатива χ из $\{\chi\}$ в целом или по своему свойству R относится к одному из фиксированных подмножеств $\{\chi^{\mathrm{I}}\}, \{\chi^{\mathrm{II}}\}, \ldots$. Такая задача называется задачей классификации и может как сводиться к перечисленным способам сравнения, так и быть самостоятельной. Частным случаем классификации выступает деление свойств альтернатив на группы по их важности в данной задаче принятия решения. Выделяются свойства, которые наиболее важны для учета, просто важные, менее важные и т.д. Смысл этого приема состоит в сужении числа свойств, принимаемых во внимание в первую очередь.

2.2.3. Композиция оценок и сравнений

В предыдущем пункте речь шла о сравнении и оценках отдельных свойств альтернатив. Но как от всего этого вернуться к сравнению альтернатив в целом? Указанная операция называется композицией и рассматривается в данном пункте.

Сначала проанализируем ситуацию, когда все свойства альтернатив имеют численную оценку, т.е. являются критериями. Обозначим их через $C_i(\chi),\ i=\overline{1},\overline{n}$. В этом случае любой альтернативе может быть сопоставлена точка n-мерного пространства E^n, координаты которой есть значения соответствующих критериев. Такое пространство называется критериальным. Будем для определенности считать, что чем больше значение i-го критерия $C_i(\chi)$, тем предпочтительнее данная альтернатива по свойству i. Рассмотрим две произвольные альтернативы. Возможны две ситуации:

1) одна альтернатива не хуже другой по всем критериям:

$$C_i(\chi) \geq C_i(\chi_l),\ i=\overline{1},\overline{n} \tag{2.6}$$

(причем хотя бы одно неравенство выполняется как строгое);

2) этого утверждать нельзя.

Условие (2.6) – это естественное условие предпочтения альтернативы χ_2 перед альтернативой χ_1. Таким образом, переход от χ_1 к χ_2 улучшает наш выбор. Существуют ли неулучшаемые альтернативы? Да, и практически всегда – для этого требуется лишь ограниченность значений критериев $C_i(\chi),\ i=\overline{1},\overline{n}$. Для демонстрации важнейших идей по композиции оценок воспользуемся удобной графической интерпретацией критериальной пространства при $n = 2$.

Обратимся к рис. 2.10. На нем в осях C_1, C_2 точками или звездочками изображены альтернативы. Неулучшаемой альтернативой на рис. 2.10, *а* очевидным образом является та, которая расположена выше и правее всех. Проверить ее неулучшаемость можно так: провести из данной точки лучи параллельно положительному направлению осей и убедиться, что в образованном углу других альтернатив нет. Это свойство неулучшаемости легко доказать от противного.

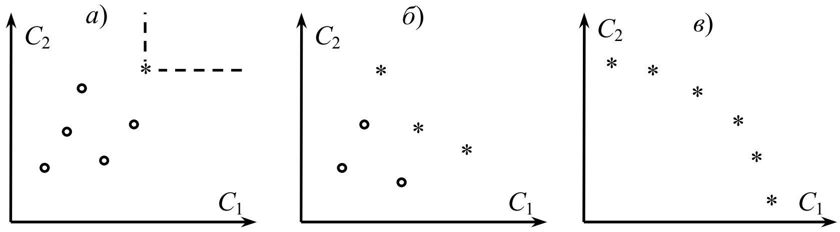

Рис. 2.10. Критериальное пространство. Множество Парето

Итак, в ситуации рис. 2.10, *а* мы нашли единственную неулучшаемую альтернативу, которую естественно выбрать в качестве наилучшей. Однако уже рис. 2.10, *б* демонстрирует, что таких альтернатив может быть более одной, а рис. 2.10, *в* показывает, что возможен случай, когда все альтернативы будут неулучшаемы. Однако типичен именно вариант рис. 2.10, *б*, на котором число неулучшаемых альтернатив меньше (зачастую – значительно) числа исходных альтернатив.

Множество неулучшаемых альтернатив называется множеством Парето для данной задачи.

Ясно, что точки, не принадлежащие множеству Парето, не претендуют на то, чтобы считаться лучшей альтернативой. Выделение множества Парето – это первый шаг в сравнении альтернатив. Можно вообще ограничиться этим и считать лучшими все те альтернативы, которые попали в это множество. Однако в абсолютном большинстве практических задач требуется в итоге выбрать только одну альтернативу. Как же выбирать на множестве Парето?

Приемов такого выбора, основанных на столь же естественных предположениях, как и те, которые привели к выделению множества Парето, к сожалению, не существует. Для дальней формализации выбора вводятся более специфические и часто достаточно спорные приемы.

Приведем наиболее распространенные из них.

1. Выбирают альтернативу, у которой сумма значений критериев максимальна. Развитие этой идеи сравнения значений различных критериев ведет к максимизации некоторой выбранной функции от критериев $f(C_1, C_2, ..., C_n)$. Вид $f = \sum_{i=1}^{n} \alpha_i C_i$ «наиболее употребителен и называется линейной сверткой критериев с весами α_i. На рис. 2.11 альтернативой с максимальной суммой критериев (свертка с $\alpha_i = 1$) будет точка χ_5.

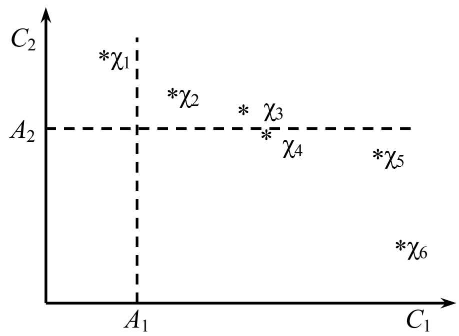

Рис. 2.11. Примеры выбора на множестве Парето

Сложение критериев друг с другом и другие операции с ними редко бывают физически обоснованными. Весьма искусственной выглядит, скажем, сумма массы и прочности, стоимости и эффективности. Введение функции от критериев – в большинстве случаев вынужденная мера, ведущая к необходимости экспертного определения весов отдельных критериев.

2. Фиксируют набор чисел (уровней) $A_i,\ i = \overline{2, n}$, и ищут альтернативу, у которой на все критерии, кроме одного, наложены ограничения $C_i(\chi) \geq A_i$, а оставшийся критерий C_1 максимален. Естественно, что взятие в качестве основного, главного критерия именно C_1 условно; он, как и важные в этой задаче уровни A_i, подлежит специальному выбору. На рис. 2.11 при закреплении уровня A_1 для первого критерия в качестве решения получим альтернативу χ_2, а при уровне A_2 для второго – альтернативу χ_3. Такой прием называется методом главного критерия или методом критериальных ограничений.

Приемы 1, 2 обладают важным свойством – предварительное выделение множества Парето в них не обязательно. Доказывается, что использование этих приемов на всем множестве альтернатив при весьма общих условиях дает тот же самый результат, что и на множестве Парето. Другими словами, методы свертки и главного критерия приводят к альтернативам, принадлежащим множеству Парето. Хотя назначение этих методов – выделять единственную альтернативу, сильная зависимость решения от весов и уровней, вида свертки и выбора главного критерия приводит к тому, что на практике предпочитают решить набор задач с различным выбором всего перечисленного. Полученный набор решений в случае их значительного несовпадения далее обрабатывается аналогично приводимому ниже приему 4.

3. Точки множества Парето оцениваются по некоторому дополнительному свойству, которое не учитывалось ранее. Это свойство (одно или более) может иметь физический характер или быть просто математическим приемом. Так, альтернативы можно сравнивать по вторичным последствиям, по специальным образом определенной устойчивости решений, по та-

кой геометрической характеристике, как «серединность». На рис. 2.11 точкой, наименее удаленной от всех остальных, будет χ_4.

4. Точки множества Парето поступают на экспертную оценку, по результатам которой на основе баллов, системы приоритетов, ранжирования, правила вето и т.д. выделяется единственная альтернатива. Если точек множества Парето слишком много, то предварительно проводят их отбор, в котором также пользуются и формальными, и неформальными драмами. Формальные способы обычно связаны с какой-либо «равномерной представимостью» точек, а экспертные могут быть основаны на выборе интересных комбинаций значений критериев и других соображениях.

Таким образом, видно, что даже для случая, когда все свойства альтернатив являются критериями, ее выбор достаточно сложен. Рассмотрим теперь ситуацию, когда для части, или даже для всех свойств альтернатив можно ввести не численную оценку, а лишь отношение сравнения. Допустим, что любая из альтернатив имеет n свойств, по каждому из которых может быть задана операция сравнения вида (2.5). Обозначим эти операции через $R_1, R_2, ..., R_n$. Пусть они транзитивны и антирефлексивны. Допустим пока, что по любому отношению R_i, $i = \overline{1, n}$, сравнимы две любые альтернативы из $\{\chi\}$**.** Тогда (от противного) по каждому свойству может быть выполнено полное ранжирование альтернатив. Это – весьма полезная операция, которая далее может быть использована различными путями. Отметим, что ее результатом будет набор перестановок из альтернатив, который иногда записывается в виде матрицы из n столбцов (по числу свойств) и N строк (по числу альтернатив). Поясним это на следующем примере. Пусть мы имеем задачу с четырьмя альтернативами и двумя свойствами. Ранжирование альтернатив по свойствам дало

$$\begin{aligned}&\chi_1 R_1 \chi_4;\ \ \chi_4 R_1 \chi_3;\ \ \chi_3 R_1 \chi_2;\\&\chi_4 R_2 \chi_3;\ \ \chi_3 R_2 \chi_2;\ \ \chi_2 R_2 \chi_1.\end{aligned} \tag{2.7}$$

Матрица ранжирования имеет вид $\begin{pmatrix} 1 & 4 \\ 4 & 3 \\ 3 & 2 \\ 2 & 1 \end{pmatrix}$. (2.8)

Напомним, что в первой строчке помещены наиболее предпочтительные альтернативы по первому и второму свойствам.

Одним из способов работы с такой матрицей является введение условного пространства свойств. В нем в проекции на ось i альтернативы будут располагаться в соответствии с ранжированием по операции R_i. Эквивалент записей (2.7) и (2.8) показан на рис. 2.12. Неулучшаемые альтернативы выделяются аналогично тому, как это происходило в критериальном пространстве. На рис. 2.12 это альтернативы χ_1 и χ_4.

Более сложный случай составляет частичное ранжирование. Пусть вместо второй строчки в (2.7) нам известно только то, что $\chi_4 R_2 \chi_3$ и $\chi_2 R_2 \chi_1$. Как здесь определить неулучшаемые альтернативы? Общий метод состоит в выделении из всех пар альтернатив (χ_R, χ_L) таких, что $\chi_R R_i \chi_L$, $i = \overline{1, n}$. Как только такая пара выделяется, альтернатива χ_1 убирается из дальнейшего рассмотрения, так как альтернатива χ_R предпочтительней. В нашем примере таким способом удается вывести из сравнения альтернативу χ_3. Большего мы не знаем и обязаны считать неулучшаемыми оставшиеся альтернативы χ_1, χ_2, χ_4.

Отсюда следует, что частичное ранжирование (упорядочение) ведет к росту числа неулучшаемых альтернатив. При частичном ранжировании не существует ни матрицы ранжирования, ни условного пространства свойств. Дальнейший выбор среди неулучшаемых альтернатив, в основном, производится методом экспертизы.

Возвратимся к случаю полного ранжирования. Здесь с неулучшаемыми альтернативами работают аналогично тому, как это происходило с точками множества Парето в критериальном пространстве. Ведь в условном пространстве свойств мы фактически ввели искусственную оценку – место альтернативы в столбце матрицы ранжирования.

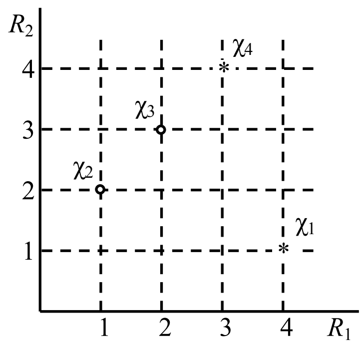

Рис. 2.12. Пример условного пространства свойств

Аналогом свертки будет сумма мест в столбцах. В примере на рис. 2.12 по этому признаку следует считать наилучшей альтернативу χ_4, она занимает второе и первое места, их сумма (с единичными весами) равняется трем. Аналогом уровней A_i будут места, ниже которых данная альтернатива не опускается в столбце i.

Еще один способ выделения неулучшаемых альтернатив состоит в использовании для этой цели графов. Наиболее распространен вариант, когда вершины графа соответствуют альтернативам, а ориентированные (со стрелкой) дуги между χ_R и

χ_L проводятся только в том случае, если $\chi_R R_i \chi_L$, $i=\overline{1,\ n}$. Такой способ пригоден и для полной, и для частичной упорядоченности. Он часто используется в упрощенном случае, когда имеется всего одна операция сравнения R, т.е. возможно непосредственное, а не по свойствам, сравнение альтернатив друг с другом. В любом случае неулучшаемыми будут альтернативы, из которых не выходят дуги.

2.2.4 Организация принятия решения

Если при принятии решения не пользоваться ни выделением и сравнением отдельных свойств (декомпозицией), ни формализованными методами композиции, то решение такой задачи называется простым. Оно осуществляется либо на основе сравнения какого-либо очевидного (главного) свойства решения, либо на основе чистой интуиции. Во всех остальных случаях требуется организация решения. Она предполагает:

а) декомпозицию альтернатив на свойства, удобные для сравнения;

б) возможное ранжирование этих свойств по важности;

в) выбор числовых характеристик свойств (критериев) и операций предпочтения, утверждение экспертных процедур для искусственной оценки свойств;

г) выбор методов композиции;

д) выбор вида информации для окончательного решения;

е) окончательное решение.

Примечание: в этапы а) – д) могут входить дополнительные экспертные процедуры для выполнения поставленных в них задач.

К организации принятия решения по схеме а) – е) привлекаются, вообще говоря, пять видов специалистов:

1. Лицо принимающее решение (ЛПР) – полностью отвечает за решение задачи. Он утверждает организацию решения по этапам а) – д) и единолично принимает окончательное решение. Коллегиальное решение (типа голосования) здесь не рассматривается.

2. Консультанты (помощники) ЛПР – совместно со специалистами по системному анализу (см. 5) участвуют в организации решения, обсуждают промежуточные и окончательные результаты, могут быть назначены как защитниками, так и оппонентами вариантов решений. Только они могут давать советы ЛПР.

3. Эксперты – в заданных, жестко очерченных рамках производят оценку, сравнение, ранжирование представленных им на экспертизу отдельных сторон альтернатив (как исключение – альтернатив целиком), могут привлекаться к оценке организации решения, однако сами решений не принимают.

4. Специалисты по использованию технических средств (в первую очередь ЭВМ) в задаче принятия решения – программисты, специалисты по банкам данных, постановщики задач на ЭВМ. Последние отвечают за выбор метода решения формализованных частей задачи и полученную при этом информацию. Как и эксперты, они действуют в строго ограниченных рамках предложенных им задач. Отметим, что применение ЭВМ, в основном, сосредоточено в этапах в) и г). Организация этих этапов должна включать фиксацию формальных задач наиболее распространенными, из которых являются вычисление критериев, применение методов свертки и главного критерия, нахождение множества Парето.

5. Специалисты по системному анализу – совместно с консультантами организуют процедуру принятия решения. Проверяют ее на соответствие общим положениям системного анализа. Детально конкретизируют передачу информации в задаче условия экспертиз и другие моменты решения задачи. По результатам процесса принятия решения вносят предложения об усовершенствовании процесса.

В приведенном перечне функции специалистов 1 – 5 разделены. На практике они могут совпадать. Например, довольно естественным выглядит совпадение функций консультанта и специалиста по системному анализу. Если это не так, то последнему придется вникать в целый ряд вопросов, специфичных именно для данной задачи.

На практике нередко встречаются ситуации, когда:

1) требуется принять решение при отсутствии возможности или времени для четкого определения цели, выбора альтернатив и поиска нужной информации;

2) окончательная фиксация цели (условий), альтернатив и требуемой информации считается частью принятия решения.

В первой ситуации обычно приходится принимать решения, близкие к простым. Здесь можно говорить лишь о грубой декомпозиции альтернатив, ранжировании свойств по важности, об эвристической композиции и окончательном решении с использованием интуиции и опыта. Применение математических методов затруднительно, так как отсутствие четкой постановки задачи и точной исходной информации ведет к недостоверности результатов. Во второй ситуации необходимо расширить список стадий а) – е), включив в него: конкретизацию условий принятия решения, сбор альтернатив, сбор и переоценку информации. Принятие решения будет тогда характеризоваться неоднократным возвратом к этим стадиям.

В любом случае полезно предусматривать стадию выработки предварительного решения и его проверку на дополнительных оценках, экспертизах, исследовании неочевидных и вторичных последствий, широкое обсуждение и др. Это помогает уяснить положительные и отрицательные стороны, а также увидеть направления улучшения решения. Указанные обстоятельства в сложных системах нередко приводят к итеративности процесса решения задачи. Как правило, первый вариант (варианты) не удовлетворяет ЛПР. Он выдвигает требования учесть дополнительные свойства, переаранжировать их, решать частные задачи в другой постановке, изменять ограничения, весовые множители, расширить или сузить используемую информацию и т.д. Все это следует считать естественным. Такие изменения в задаче позволяют лучше понять ее динамику, оценить «чувствительность» и в итоге принять более обоснованное решение.

Следует указать, что в определенных рамках по созданным человеком алгоритмам решение может самостоятельно, без человека, приниматься вычислительными машинами и другими техническими средствами. Организация такого решения может быть весьма сложной и практически повторять все этапы схемы а) – е). Важным вопросом здесь является доверие к такому, не просмотренному человеком, решению. Широкие классы задач, когда такое решение незаменимо, – это очевидные схемы обработки большого количества информации, случай, когда человек не успевает обработать информацию за требуемое время (нередко – доли секунд), принятие большого количества однотипных и достаточно тривиальных решений.

В заключение этого пункта обсудим вопросы, касающиеся организации экспертиз. Мы уже видели, что они могут быть достаточно разнообразны и многочисленны. Классические формы работы с экспертами – это заполнение анкет (таблиц), интервью, запрос аналитического текста.

Первая из них является наиболее распространенной. В вопросники следует включать простые вопросы, которые для ответа не надо разбивать на отдельные части. Интервью предпочтительнее анкеты, если оно проводится высококвалифицированным специалистом, способным подстроиться под интервьюируемого, помочь ему выбрать более обоснованные ответы, но одновременно не привнести в них свое мнение.

Следует также отметить, что хотя обычно существует желание получить оценку именно в числах (процентах, вероятности, баллах и т.п.), специальные исследования показывают, что человек более обоснованно приводит качественные ответы, чем количественные.

Экспертизы различаются и по форме взаимодействия экспертов. Обмен мнениями может быть свободным, регламентированным и недопустимым. Все эти способы имеют свои преимущества и недостатки. При свободном общении ряд экспертов может доминировать над другими. Чье-то мнение может оказаться неучтенным. Регламентируемое общение требует более сложной организации; его известный вид – это метод «мозговой атаки», когда сначала мнения высказываются без обсуждения, а лишь через некоторое время дискутируются, как правило, под руководством хорошо подготовленного ведущего. Изолированная работа с экспертами чревата попаданием в дальнейшую обработку искаженных или просто неверных оценок, которые могли бы быть выявлены и изменены при свободном или регламентированном способе. Причинами неудовлетворительных ответов может, быть нарушение целого ряда требований к экспертам – от неполной компетентности и предвзятости до неспособности решать нестандартные задачи и предвидеть неочевидные последствия.

Контрольные вопросы

1. Приведите пример постановки задачи принятия решений.
2. Ответьте на вопрос, что такое декомпозиция.
3. Расскажите об организации процесса принятия решений.

2.3. СОЧЕТАНИЕ ФОРМАЛИЗОВАННЫХ И НЕФОРМАЛИЗОВАННЫХ ДЕЙСТВИЙ

Выше неоднократно употреблялись термины формализованных (формализуемых), а также необладающие этими свойствами процедур и операций. Считалось, что было достаточно интуитивного понимания их смысла. Данный раздел посвящен углубленному рассмотрению этих важных видов действий.

2.3.1. Понятие формализованных и неформализованных действий

Назовем процедуру (операцию) формализованной, если определена и однозначно понимаема (человеком, вычислительной машиной, другим техническим устройством) последовательность элементарных актов по ее реализации.

Обычно формализация предполагает возможность многократного повторения процедуры (неуникальность), ее пригодность для некоторого множества исходных данных (вариативность входов), возможность фиксации последовательности действий на каком-либо носителе для хранения, передачи, тиражирования. Упомянутая в определении однозначная понимаемость имеет своим следствием совпадение (на практике чаще приближенное) результатов применения одной и той же процедуры к одним и тем же исходным данным.

Назовем процедуру (операцию) неформализованной, если она производится с использованием интуиции человека, т.е. с неполным осознанием аргументов и приемов выбора действий.

Типовыми примерами формализованных операций являются работа программных средств для ЭВМ, действия рабочего на конвейере, ответы справочной службы, обработка результатов эксперимента по определенной методике, работа следящих или компенсирующих технических средств и многое другое.

Неформализованными операциями будут составление нового программного средства, исправление ошибок в нем, экспертизы, действия водителя в нестандартной ситуации, научно-техническое творчество.

Могут быть формализованы, но чаще всего останутся неформализованными выбор метода решения задачи, составление зависимостей, описывающих задачу, декомпозиция и выделение иерархии в системе, анализ результатов исследования и т.д.

Будет ли любое действие либо формализованным, либо неформализованным в смысле этих определений? Строго говоря, нет. Но можно рекомендовать всегда исследовать возможность формализации данного действия. Общей тенденцией является, что все формализованные действия следует стараться поручать вычислительной и другой технике, разгружая человека в творческой деятельности. Однако абсолютизировать это утверждение не стоит. Приведем следующий пример. Пусть надо решить, какая кирпичная кладка в углу стены прочнее – внахлест или со сплошным швом? Эта задача может быть формализована и решена на основе, скажем, метода конечных элементов. При этом она будет считаться средней или даже значительной по сложности. Но любой каменщик уверенно ответит, что стена со сплошным швом только чудом не развалится. Нужна ли здесь формализация? Она с лихвой перекрывается человеческим опытом. Именно таков наиболее распространенный аргумент в пользу отказа от введения моделей и других формальных структур. И к этому аргументу, конечно, надо прислушаться. Человек при помощи интуиции может решать задачу быстрее, дешевле и даже (при плохой модели, неполной информации и др.) надежнее.

Но вернемся еще раз к рассмотренному примеру, чтобы показать полезность более широкого взгляда на проблему. Да, для единичного акта оценки прочности стены программное средство по методу конечных элементов создавать бесполезно. Но, будучи созданным, такое средство поможет в нахождении варианта наиболее прочной кладки. И это тоже типично – формализация, которая не нужна при решении данной проблемы, оказывается полезной при решении несколько измененной задачи и тем самым оправдывает себя как средство решения класса задач. Все это говорит о том, что отказ от формализованного описания там, где оно в принципе возможно, должен быть хорошо продуманным исключением.

2.3.2. Совместные действия человека и ЭВМ

В настоящее время важнейшими носителями действий являются человек и вычислительная техника. При этом человек может выполнять как формализованные, так и неформализованные процедуры и операции, а ЭВМ – только формализованные. Рассмотрим систему действий, в которой имеются как те, так и другие процедуры.

Именно на примере взаимодействия человека и ЭВМ было впервые осознано значение сочетания формализованных и неформализованных действий. Перейдем к важному примеру, демонстрирующему различные аспекты этого взаимодействия. Пусть имеется комплекс, состоящий из ЭВМ с пультом оператора в виде дисплея, подсоединенного к ЭВМ банка данных с нужной информацией, библиотеки программных средств и библиотеки математических моделей, также хранящихся во внешней памяти комплекса, дополнительных технических средств в виде графопостроителя (чертежной машины), устройства для печатания и размножения текстовых результатов, средств для обработки результатов эксперимента и др. Напомним, что и дополнительные средства также подсоединены к основной ЭВМ. Такой комплекс в советской литературе принято называть АРМом – автоматизированным рабочим местом.

Высококвалифицированный оператор, который свободно владеет этим комплексом, начиная решать на нем научную, инженерную, конструкторскую или другую задачу, вызывает модель из библиотеки моделей, наполняет ее информацией с помощью библиотеки данных и «запускает» на расчет или оптимизацию с помощью имеющихся в его распоряжении программ. При этом он предусматривает остановки работы в определенный момент времени, по завершении стадий процесса, при возникновении особых ситуаций. В эти остановки он оперативно оценивает полученную информацию и принимает решение на изменения в модели, данных, выборе программы и ее внутренних параметров и т.д. Далее продолжаются новые формализованные операции до следующей остановки.

Так специалист двигается к поставленной перед ним цели. Наверное, излишне говорить, что такой сложный, трудно налаживаемый комплекс используется для решения нетривиальных задач, которые вряд ли будут решены в автоматическом, без вмешательства человека, режиме. Речь идет о разумном разделении труда между человеком и ЭВМ, в котором он делает все то, что может сделать лучше машины. Такой режим работы принято называть диалоговым. Но в каком смысле человек выступает равноправным партнером ЭВМ в ходе самого процесса решения? Что же он может делать лучше машины? Ведь и она, в принципе, может сама изменять модель, выполнять расчеты для наборов данных, даже адаптировать программу к задаче, т.е., говоря современным языком, самообучаться.

Однако человек (специалист) сделает это лучше. Он привлечет свой прошлый опыт, накопит на решаемой задаче новый, использует свою интуицию, способность принимать решение при недостатке информации, поступать нетрадиционно. Таков главный побудительный мотив создания диалоговых систем.

Более того, опыт работы на АРМах и просто больших вычислительных машинах с развитым программным и другим обеспечением (сервисом) показывает, что общение с ЭВМ может приводить к качественно новому уровню проникновения в проблему. Тесный контакт с вычислительной техникой приводит к тому, что человек начинает «ощущать» задачу. Быстрая реакция машины на запросы информации и изменения в процессе работы позволяет часто неосознанно предвидеть последствия своих решений и в результате существенно быстрее достигать цели.

Диалог между человеком и ЭВМ имеет целый ряд аспектов, из которых мы рассмотрели здесь только один, главный – чередование формализованных и неформализованных действий.

2.3.3. Интерактивные системы

Рассмотрим диалог как некоторый общий режим сочетания формализованных действий и действий человека. Развитие техники привело к разгрузке человека от физического труда, а в настоящее время – и от рутинного умственного, однако можно утверждать, что неформализуемые действия останутся всегда и будут прерогативой человека. Таким образом, возникает задача об оптимальном сочетании формализованных и неформализованных процедур. Коснемся некоторых подходов к этой задаче.

В будущем, по крайней мере, в технике основной объем действий будет формализован. Действия человека останутся «островками» в море контролируемой формализации. Для этих «островков», «горячих точек» в системе действий должно быть определено свое место, должно быть обосновано, что именно здесь необходимо участие человека. По-видимому, в большинстве случаев действия человека будут носить характер принятия решения. При этом важно изучать и оценивать качество информации, на основе которой принимается решение. В ряде случаев как эта информация, так и перечень возможных действий подлежат строгой регламентации. Желательно, чтобы решения человека были простыми (см. разд. 2.2); все решения, для которых необходимо учитывать (продумывать) несколько сторон проблемы, менее оперативны и обоснованны.

Основные формы участия человека:

а) вмешательство – оно обычно называется управлением;

б) визирование работы технических средств (т.е. просмотр и согласие с их действиями);

в) запрос информации и ее направление на обработку (сюда же входят тестирование технических средств, поиск неисправностей).

Разновидностью вмешательства в случае создания новой системы является выдвижение новых идей и представлений, проверка различных гипотез и вариантов.

Примеры описанных систем существуют уже в наши дни. Это развитые системы автоматизированного проектирования (САПР), гибкие автоматизированные производства (ГАП), автоматизированные системы научных исследований и др. В них человек решает лишь узловые, неформализуемые или плохо реализуемые проблемы.

Общим результатом размещения человека в сложной системе является наделение системы внутренней активностью, способностью совершенствоваться, приспосабливаться, изменяться не на основе заранее заложенных алгоритмов и механизмов самообучения, а на основе неформализованных решений. Такая активность является естественным продолжением обычной активности человека, ее отличие лишь в многократном расширении возможностей, которые предоставляет техника. Уже сейчас мы можем говорить о быстром доступе к огромным объемам информации, переработке ее гигантских массивов, быстрой смене моделей, принципов и методов управления. Это сочетается с такими качествами человека, как гибкость действий, быстрая смена уровней охвата проблемы, способность оценки с различных, в том числе новых, неожиданных точек зрения и, наконец, возможность интуитивных действий. В программировании режим непосредственного вмешательства оператора в работу ЭВМ принято называть интерактивным. Идея интерактивности применительно к произвольно автоматизированной системе состоит в интенсификации работы за счет активности человека внутри сложной системы. Именно она позволяет объединить все достоинства формализованных и неформализованных действий и успешно решать широкие классы задач, недоступных только техническим средствам или только человеку.

Контрольные вопросы

1. Расскажите, что такое формализованные и неформализованные действия.
2. Дайте характеристику интерактивной системе.
3. Охарактеризуйте достоинства и недостатки интерактивных систем.

Глава 3
МНОГОКРИТЕРИАЛЬНЫЕ И ИЕРАРХИЧЕСКИЕ СИСТЕМЫ

3.1. ПОСТАНОВКА ЗАДАЧ ВЫБОРА ОПТИМАЛЬНОГО РЕШЕНИЯ

3.1.1. Общая постановка задачи в многокритериальных и иерархических системах

При постановке задач оптимизации требуется прежде всего определить цель (или несколько целей), преследуемую субъектами управления, и установить, какими характеристиками (переменными) системы (или процесса) можно оперировать, т.е. какие переменные можно рассматривать в качестве управляющих параметров.

Под целью будем понимать тот конечный результат, который рассчитывают получить субъекты управления с помощью выбора управляющих воздействий на исследуемую систему.

После того как цель определена, необходимо найти оптимальный способ действий каждого субъекта управления, обеспечивающий ее достижение. Если для системы определены несколько целей развития, то требуется указать принцип оптимальности, который позволял бы выделять решения, наилучшие в смысле достижения этих целей.

Следующим важным моментом является задание множества допустимых воздействий на систему со стороны субъекта управления – множества управляемых переменных. Любой набор допустимых воздействий будем называть решением.

При построении математических моделей функционирования и развития даже сравнительно небольших реальных систем мы сталкиваемся с необходимостью учета сложных взаимосвязей компонент модели, оказывающих действенное влияние на реализацию альтернатив развития и достижение поставленных целей. Внутренние межкомпонентные связи системы могут быть описаны с помощью некоторого конечного графа $G = (Z, G)$, вершинами которого служат компоненты модели.

Графом называют пару $G = (Z, G)$, в которой множество вершин обозначено через Z, а множество дуг (ребер) задано выражением $G : Z \to Z$. Для упрощения анализа графов вершины их обычно нумеруют. В этом случае, если i и j – номера смежных вершин, то ребро графа может быть задано парой (i, j). Если все ребра графа заданы упорядоченными парами, то такой граф называют ориентированным.

На графе межкомпонентных связей G, исходя из описания системы, целесообразно выделить основные компоненты модели, на которые субъект управления может оказывать непосредственное воздействие, и сопутствующие компоненты, состояние которых однозначно определяется состоянием основных компонент. Рассмотрим некоторые примеры управления сложными системами.

Пример 1. Трехуровневая система управления гибким автоматизированным участком.

В общем случае гибкое автоматизированное производство (ГАП) представляет собой систему, включающую следующие компоненты:

- автоматизированные технологические модули (станки, линии, участки);
- автоматизированный транспорт;
- автоматизированные склады.

Управление работой этих компонент и осуществление связей между ними обеспечивает система управления ГАП (СУ ГАП). С ее помощью осуществляются запуск, управление и контроль за работой технологического оборудования, синхронизация выполняемых работ, оптимизация загрузки оборудования, формируется график работы транспортных средств, автоматизированных складов и т.п.

Как правило, СУ ГАП имеет иерархическую структуру. Наиболее распространенной структурой системы управления гибким автоматизированным участком является трехуровневая (рис. 3.1).

Верхний уровень решает задачи организационно-экономического характера и принимает долгосрочные решения; проводит расчет сменно-суточных заданий по каждой единице станочного оборудования, заданий по технологической подготовке участка; учитывает запас заготовок, инструмента и приспособлений на складе; накапливает информацию для различных служб цеха,

Средний уровень осуществляет контроль за работой микропроцессорных систем; принимает оперативные решения в соответствии с поступающей от подсистем нижнего уровня информацией; вырабатывает управляющие воздействия на эти подсистемы.

Нижний уровень обеспечивает с помощью микропроцессорных систем непосредственное управление технологическим процессом.

Анализ и управление работой такого участка требует решения значительного количества оптимизационных многокритериальных задач, задач сетевого планирования, транспортных задач, задач размещения и т.д.

Пример 2. Проектирование оптимального программного комплекса.

При проектировании программного комплекса необходимо обеспечить выполнение ряда требований: увеличить точность задания входных воздействий, сократить объем оперативной памяти, уменьшить время работы программ, уменьшить загрузку каналов связи между ЭВМ и внешними запоминающими устройствами и т.д.

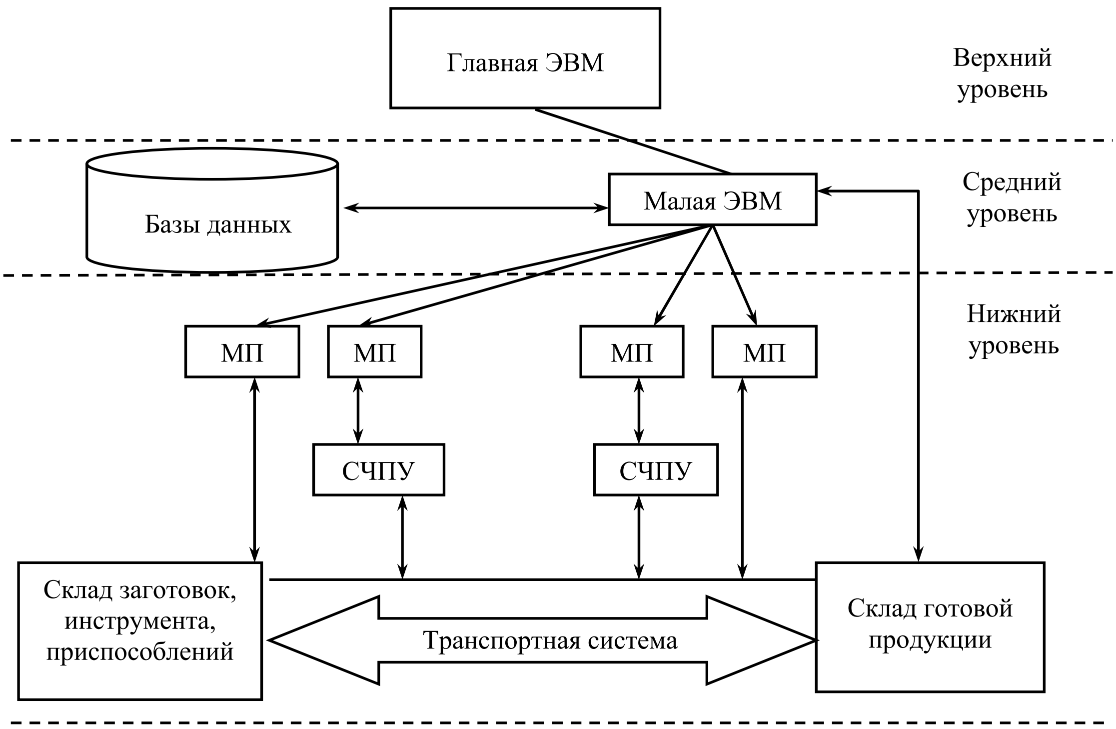

Рис. 3.1. Принципиальная схема трехуровневой системы управления гибким автоматизированным участком

Предположим, что программный комплекс должен реализовать множество операций $\sigma = \sigma_1, \sigma_2, ..., \sigma_m$. Под операцией мы понимаем, например, решение системы дифференциальных или алгебраических уравнений конкретного вида, нахождение экстремума некоторой функции, поиск информации в заданном массиве и т.п.

Каждая операция $\sigma_i \in \sigma$ может быть реализована любой из программ некоторого заданного множества $\pi_i = (\pi_{i1}, \pi_{i2,} ..., \pi_{ij})$ $i = 1, 2, ..., M$.

Программы отличаются друг от друга по своим характеристикам, влияющим на выполнение требований к комплексу в целом. Программный комплекс представляет собой упорядоченный набор программ

$$\Pi =< \pi_{1l_1}, \pi_{2l_2}, ..., \pi_{Ml_M} > .$$

При его разработке учитывается вектор критериев $H = (H_1, H_2, \ldots, H_m)$ поэтому оценка качества комплекса является векторной величиной.

Множеством допустимых управлений является множество $\pi = \pi_1 \times ... \times \pi_m$. Отображение $H : \pi \to E_m$ задает правило, но которому каждому набору программ (или варианту программного комплекса) соответствует векторная оценка. Таким образом, задача проектирования оптимального программного комплекса является многокритериальной задачей оптимального управления.

Пример 3. Планирование развития региона [12, 13]. В рамках программы «Человек и биосфера», осуществляемой в Международном институте системных исследований, была поставлена и решена задача оптимального развития высокогорной деревни Обергурль (Австрия). Находясь на высоте 2000 м над уровнем моря, она привлекает к себе постоянно увеличивающийся поток туристов как зимой, так и летом. Естественно, что к 1970-м гг. здесь стали проявляться первые признаки экологических последствий бурного развития, принимающего форму строительства гостиниц и подъемников, резкого увеличения числа людей и транспортных потоков в деревне и ее окрестностях. Стала реальной опасность того, что в конце концов туристы перестанут посещать деревню, поскольку она потеряет для них всякую привлекательность. Ограничивающим фактором развития туризма является нехватка площадей под строительство новых гостиниц и недостаток безопасных от снежных лавин площадей. По инициативе Комитета Австрии по реализации программы «Человек и биосфера» деревня Обергурль была выбрана в качестве объекта для интенсивных исследований. При тесном взаимодействии ученых разных специальностей: метеорологов, ботаников, зоологов, микробиологов, географов, экономистов, социологов и даже антропологов – были разработаны различные динамические модели развития этой деревни. На графе (рис. 3.2) легко выделить основные компоненты, развитие которых может управляться извне путем привлечения дополнительных ресурсов и капиталовложений, и сопутствующие компоненты. Изменение хотя бы одной из компонент может оказать зачастую неожиданное воздействие на другие компоненты.

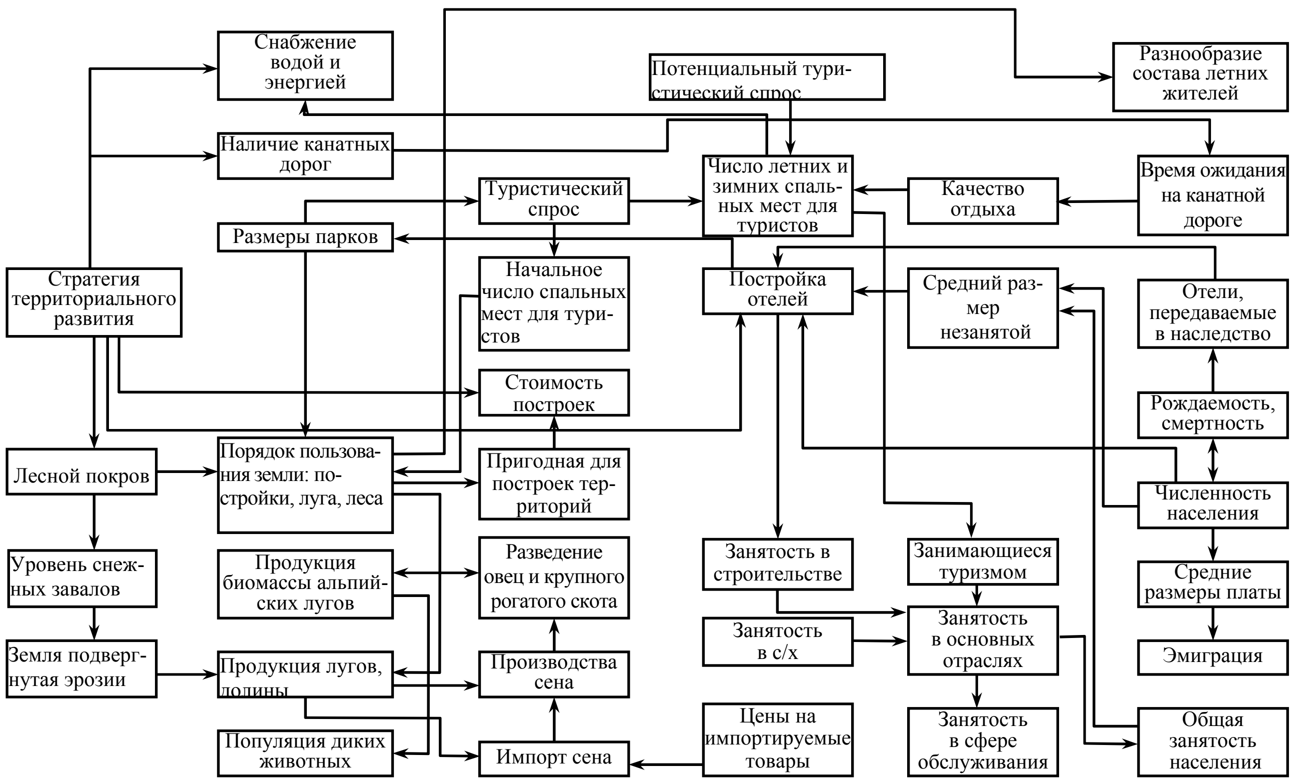

Рис. 3.2. Важнейшие компоненты модели Обергурля

П р и м е р 4. Нефтехимическое производство [12 – 14]. Рассмотрим процесс производства этилена на нефтеперегонном заводе, имеющем интегрированную систему управления. Обобщенная блок-схема производства, принятая на заводе, показана на рис. 3.3. Производство можно разбить на три основных подпроцесса: крекинг, компрессия и разделение.

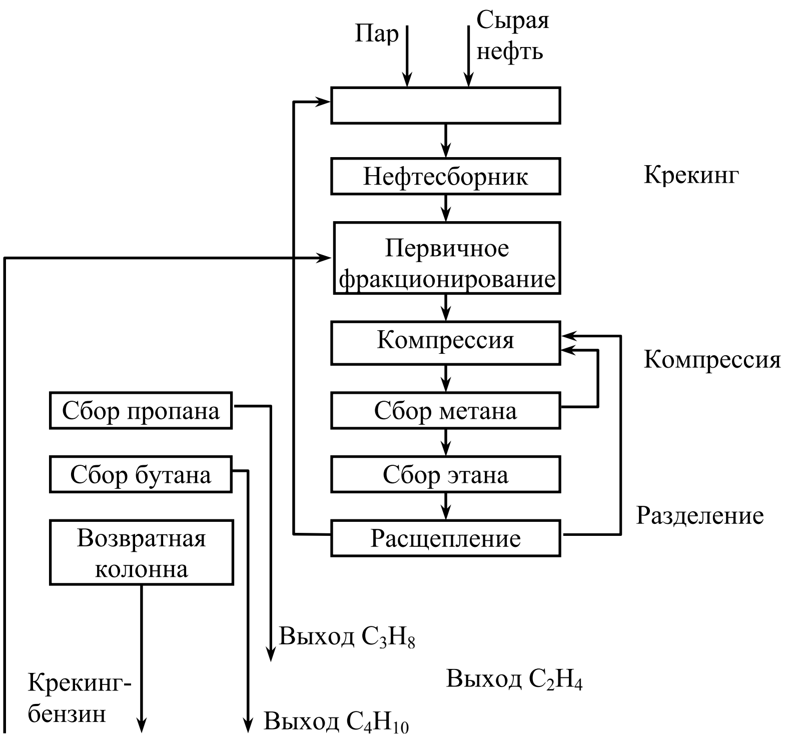

Рис. 3.3. Процесс крекинга нефти с получением этилена

Блок-схема интегрированной системы управления состоит из двух частей – производственной и управляющей (рис. 3.4). Приведенная здесь система управления имеет иерархическую структуру. Имеется три основных уровня управляющей системы. На верхнем уровне вырабатывается плановое задание по производству этилена с учетом максимизации прибыли. На втором уровне производство ведется с учетом минимизации затрат. На третьем уровне обеспечивается управление ходом процесса.

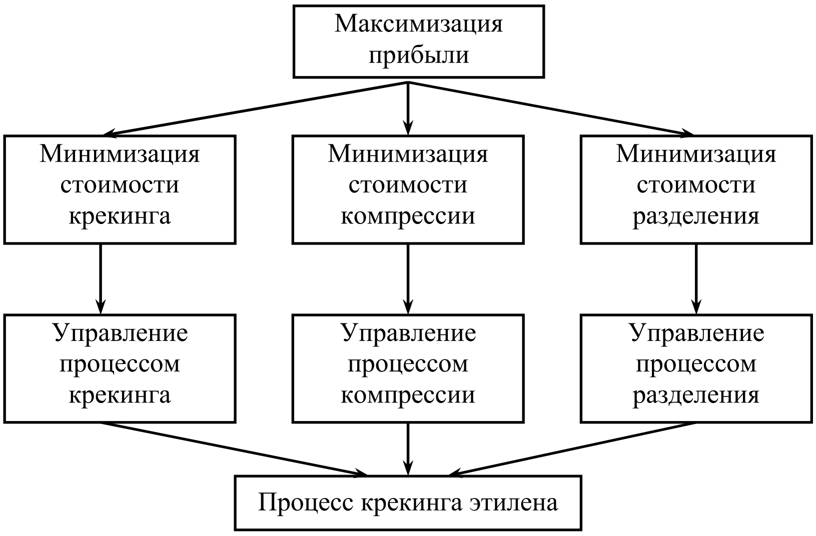

Рис. 3.4. Блок-схема интегрирования системы управления крекингом этилена

Перейдем теперь к формализации задач управления в многокритериальных и иерархических системах.

Рассмотрим общий конечный граф (Z, G), определяющий взаимосвязь компонент системы. В этом графе узлы $z \in Z$ представляют собой компоненты, а отображение G определяет зависимость компонент между собой.

Предположим, что множество Z разбито на два множества X, Y ($X \cup Y = Z$, $X \cap Y = \varnothing$) основных управляемых извне компонент X и сопутствующих компонент Y. Перенумеруем компоненты множества X индексами $i = 1, 2, \ldots, n$, а компоненты множества Y – индексами $j = 1, 2, ..., m$. Количественное состояние компоненты $i \in X$ определяется вектором $x^i \in R^n$ и количественное состояние компоненты $j \in Y$ – вектором $y^i \in R^n$.

Введем следующие обозначения:

$$x^{G^{-1}(j)} = \{x^k : k \in G^{-1(j)}\},\ y^{G^{-1}(j)} = \{y^k : k \in G^{-1}(j)\},$$

т.е. $x^{G^{-1}(j)}$, $y^{G^{-1}(j)}$ – векторы, координаты которых представляют собой количественные состояния компонент из множеств $G^{-1}(j) \cap X$ и $G^{-1}(j) \cap Y$, соответственно.

Обозначения $h_j(x^{G^{-1}(j)}, y^{G^{-1}(j)})$, $f_i(x^{G^{-1}(j)}, y^{G^{-1}(j)}, u_i)$, $u_i(x^{G^{-1}(j)}, y^{G^{-1}(j)})$ будем использовать для выражения зависимостей соответствующих функций от количественных состояний компонент из множеств $G^{-1}(i)$, $G^{-1}(j)$. При построении модели предполагается, что можно определить соотношения между компонентами:

$$y^j = (x^{G^{-1}(j)}, y^{G^{-1}(j)}), j \in Y;$$

$$x^i = (x^{G^{-1}(j)}, y^{G^{-1}(j)}), \quad (3.1)$$

где h_j, f_i – вещественные вектор-функции размерности n; u_i – управляющий параметр, выбираемый из множества: $U_i(x^{G^{-1}(j)}, y^{G^{-1}(j)}) \subset R^{l_i}$, структура которого определяется количественными соотношениями компонент, оказывающих влияние на изменения компоненты x^i.

Для каждой компоненты $i \in X$ определим ее полезность. Эта полезность, вообще говоря, определяется количественным состоянием (значением) не только компоненты i, но и других компонент. Обозначим ее через $H_i(x, y)$, $i \in X$, где наборы векторов $x = \{x^i, i \in X\}$, $y = \{y^j, j \in Y\}$ определяют количественные состояния компонент всей системы. Поскольку вектор состояний x явным образом, а вектор y неявно зависят от выбора управлений u_i, $i \in X$, можем определить полезность M_i по формуле

$$M_i(u_1, u_2, \ldots, u_n) = H_i(x, y), \quad i \in X. \quad (3.2)$$

Таким образом, мы получаем вектор полезностей

$$M = [M_1(u_1, u_2, \ldots, u_n), M_2(u_1, u_2, \ldots, u_n), \ldots, M_n(u_1, u_2, \ldots, u_n)]. \quad (3.3)$$

Если субъектом управления является единственный управляющий центр, который принимает решение о выборе управлений u_1, u_2, …, u_n, то задачу нахождения оптимального управления (решения) с векторным критерием качества (3.3) будем называть задачей многокритериальной оптимизации.

Если в процессе управления участвуют n различных сторон, выбирающих соответственно управления u_1, u_2, …, u_n и максимизирующих свои собственные критерии качества $M_1(u_1, u_2, \ldots, u_n)$, $M_2(u_1, u_2, \ldots, u_n)$, …, $M_n(u_1, u_2, \ldots, u_n)$, то получаем математическую модель принятия решений в условиях несовпадающих интересов участников. Такие модели называются играми, процесс принятия решения в таких условиях – конфликтом, а стороны, принимающие решения, – игроками. Каждый игрок i имеет возможность выбрать любое допустимое управление $u_i \in U_i(x^{G^{-1}(j)}, y^{G^{-1}(j)})$. Множество $U_i(x^{G^{-1}(j)}, y^{G^{-1}(j)})$ называют также множеством стратегий игрока i. Особенность такого подхода к анализу системы заключается в том, что множества стратегий игроков $U_i(x^{G^{-1}(j)}, y^{G^{-1}(j)})$ зависят от количественных состояний компонент, влияющих на компоненту x^i, а следовательно, и от управлений других игроков, влияющих на изменение состояний компонент из множества $G^{-1}(i)$. Поэтому здесь нельзя говорить о том, что игроки выбирают свои стратегии одновременно и независимо друг от друга, поскольку выбор может привести к возникновению противоречивых ситуаций. Действительно, для выбора своей стратегии u_i, игрок i должен знать множество $U_i(x^{G^{-1}(j)}, y^{G^{-1}(j)})$, а следовательно, количественные состояния $x^{G^{-1}(j)}, y^{G^{-1}(j)}$. Это возможно при конкретизации информационной структуры и порядка выбора стратегии игроками, управляющими основными компонентами системы.

Изучение оптимального поведения в конфликтных иерархических системах представляет собой достаточно серьезную проблему, и этот вопрос является предметом специального рассмотрения в последующих параграфах настоящей главы.

3.1.2. Основные понятия, определения и свойства

Аппаратная реализация системного анализа предполагает выработку стандартных приемов моделирования процесса принятия решений в сложной системе и общих способов работы с построенными моделями.

Большинство ситуаций, связанных с проблемой принятия решения в сложной системе, заключается в том, что из имеющегося множества вариантов решения (допустимых управлений) необходимо выделить некоторое подмножество вариантов, являющихся более предпочтительными. Правило, по которому устанавливается предпочтительность в множестве решений, называется ***принципом оптимальности***.

Указанные элементы – множество вариантов решения и принцип оптимальности – позволяют формализовать процесс принятия решения. Отсутствие одного из этих элементов полностью лишает задачу смысла.

Обозначим множество вариантов решения через Ω. Элементы множества называют иногда также ***альтернативами***.

Пусть задано отображение φ: $\Omega \to E_1$, которое каждой альтернативе ставит в соответствие некоторое вещественное число. Число $\varphi(x)$ называют оценкой альтернативы x. Как правило, на практике оценка альтернативы, которая выражается числом, не в полной мере характеризует качество альтернативы в сложной системе. Поэтому для определения качества альтернативы часто пользуются сразу несколькими оценками, которые составляют вектор оценок, или векторную оценку.

В общем случае задача принятия решений сводится к решению двух последовательных задач: выбора множества допустимых альтернатив и выбора оптимального множества альтернатив, которое часто называют решением.

В дальнейшем будем рассматривать методы решения задач оптимального управления системами, поэтому в качестве множества допустимых альтернатив обычно будем использовать множество допустимых управлений [9, 11, 15 – 19].

Пусть задано множество допустимых управлений, которое обозначим через U. Управления могут иметь различную природу: непрерывные и дискретные функции, стратегии в игре, правило остановки и т.п. Для общей постановки задачи вид управлений и структура множества U несущественны. Каждому управлению ставится в соответствие векторный критерий $H(u) = [H_1(u), H_2(u), \ldots, H_n(u)]$, где $H_i(u)$ – заданные функции; E_n – евклидово векторное пространство. В задачах многокритериальной оптимизации сравнение решений (управлений) по предпочтительности осуществляется для заданного в пространстве критериев E_n отношения предпочтения. Пространство E_n называют также пространством оценок.

Пусть на пространстве E_n задано бинарное отношение R. Бинарные отношения могут применяться для описания предпочтений и попарных связей различного характера между компонентами системы или объектами произвольной природы.

Отношением R на множестве E_n называется подмножество множества $E_n \times E_n$, т.е. $R \subset E_n \times E_n$. Содержательный смысл состоит в том, что отношением R является совокупность упорядоченных пар $<a, b>$, где $a, b \in E_n$. Если пара $<a, b>$ входит в R, то пишут aRb и говорят, что a находится в отношении R с b.

Отношение R называется рефлексивным, если $<a, b> \in R$ для любого $a \in E_n$.

Отношение R называется симметричным, если из $<a, b> \in R$ следует, что $<b, a> \in R$; асимметричным, если из $<a, b> \in R$ следует $<b, a> \overline{\in} R$; антисимметричным, если из $<a, b>, <a, b> \in R$ вытекает $a = b$.

Отношение R называют транзитивным, если из aRb и bRc следует aRc.

Отношение R называется полным, если для любых $a, b \in R$ справедливо aRb или bRa. Отношение, не являющееся полным, называется частичным.

Например, отношение $\geq$ («не меньше») на множестве действительных чисел рефлексивно, антисимметрично, транзитивно и полно.

Определим на множестве E_n отношения $\geqq$, $\geq$, $>$ следующим образом:

$$a \geqq b \Leftrightarrow a_i \geq b_i,\ i = 1, 2, \ldots, n,\ a \geqq b \Leftrightarrow a \geq b \text{ и } a \neq b$$

(т.е. хотя бы одно из n неравенств $a_i \geq b_i$ строгое);

$$a > b \Leftrightarrow a_i > b_i,\ i = 1, 2, \ldots, n,$$

где $a = (a_1, a_2, \ldots, a_n)$, $b = (b_1, b_2, \ldots, b_n)$.

Обозначим через (μ, a) скалярное произведение векторов. Справедливо следующее утверждение.

Л е м м а 1. Для любых $a, b \in E_n$ из неравенства $a \geqq b$ следует, что существует такой вектор $\lambda \in M = \{(\lambda_1, \lambda_2, \ldots, \lambda_n)\}$.

Доказательство. Предположим, что $a \geq b$. Если $a = b$, то очевидно $(\lambda, a) = (\lambda, b)$ для любого $\lambda \in M$. Пусть для некоторого номера $j \leq n$ имеет место строгое неравенство $a_j \geq b_j$. Обозначим $p = \max\{A, B\}$, где $A = \sum_{i \neq j} |a_i|$, $B = \sum_{i \neq j} |b_i|$. Если $p = 0$, то $(\lambda, a) = \lambda_i a_i > \lambda_i b_i = (\lambda, b)$ для любых $\lambda_i > 0$.

Если $p > 0$, положим $r = (a_j - b_j) / 2p > 0$.

Поскольку $p \geq (A + B) / 2$, то $(a_j - b_j) = 2pr \geq r (A + B)$, откуда $a_j - rA \geq b_j - rB$.

Учитывая

$$\sum_{i \neq j} a_i \geq -\sum_{i \neq j} |a_i| = -AB = \sum_{i \neq j} |b_i| \geq \sum_{i \neq j} b_i$$

получим $a_j + r\sum a_i \geq b_j + \sum b_i$.

Выберем $\lambda_j = 1/(1 + r (n - 1))$, $\lambda_i = r / (1 + r (n - 1))$ при $i \neq j$, тогда окончательно получим $(\lambda, a) \geq (\lambda, b)$.

Для произвольного бинарного отношения R часто возникает задача: среди элементов множества E_n найти недоминируемые по бинарному отношению R. Элемент a называется недоминируемым по бинарному отношению R, если не существует $b \in E_n$, такого, что bRa. Для решения задачи могут быть использованы свойства этих отношений. Одним из основных свойств такого рода является отделимость.

Пусть $\lambda \in E_n$. Отношение R на E_n называется λ-отделимым, если $aRb \Rightarrow (\lambda, a) > (\lambda, b)$.

Если неравенство заменить на нестрогое, то получим понятие нестрогой λ-отделимости.

Пусть отношение λ_1-отделимо и λ_2-отделимо, т.е.

$$aRb \Rightarrow (\lambda_1, a) > (\lambda_1, b),\ (\lambda_2, a) > (\lambda_2, b).$$

Тогда, очевидно, отношение R будет $(\lambda_1 + \lambda_2)$-отделимым. Если R λ-отделимо и k – положительная константа, то R является
$k\lambda$-отделимым. Таким образом, множество векторов λ, для которых R λ-отделимо, представляет собой конус.

Рассмотрим векторный критерий $H(u)$. Для каждого $u \in U$, $H(u)$ есть вектор пространства E_n. Сформулируем понятие оптимальности для произвольного бинарного отношения R. Управление $u^* \in U$ называется оптимальным, если не существует такого $u \in U$, что вектор $H(u)$ более предпочтителен, чем $H(u^*)$, т.е. $H(u)\,\overline{R}\,H(u^*)$ для всех $u \in U$. Такие управления иногда называют также R-оптимальными.

Решением $\sigma(U, H, R)$ многокритериальной задачи оптимального управления называется множество всех R-оптимальных управлений
$u^* \in U$.

Обозначим $N = H(U)$ множество всевозможных исходов, получаемое при использовании всех допустимых управлений $u^* \in U$. Справедливо следующее утверждение.

Т е о р е м а 1. Если N – замкнутое, ограниченное множество в E_n, R-отделимое отношение, то многокритериальная задача оптимального управления имеет решение, т.е. $(U, H, R) \neq \varnothing$.

Доказательство. В силу замкнутости и ограниченности множества $\tilde{\lambda}$ множество Arg max $(\lambda, H) \neq 0$. Пуст $H^* \in$ Arg max(λ, H) и $u^* \in U$, такое, что $H^* = H(u^*)$. Здесь Arg max $(\lambda, H) = \{H^*\colon H^* \in \tilde{\lambda}, (\lambda, H^*) > (\lambda, H), \forall H \in \tilde{\lambda}\}$. Покажем, что $u^* \in \sigma(\tilde{\lambda}, H, R)$. Действительно, предположим, что некоторое $u \in U$ более предпочтительно, чем u^*, т.е. $H(u)RH(u^*)$. Тогда в силу λ-отделимости выполняется условие

$$(\lambda, H) = (\lambda, H(u^*)) > (\lambda, H(u^*)) = (\lambda, H^*).$$

Это противоречит тому, что $H^* \in$ Arg max (λ, H). Теорема доказана.

Основным способом нахождения решений задач многокритериальной оптимизации является использование λ-сверткой исходной задачи. Пусть $\lambda \in E_n$. Будем называть λ-сверткой многокритериальной задачи $<U, H, R>$ однокритериальную задачу $<U(\lambda, H), R'>$, где отношение R' определено как > («больше») Следующая теорема позволяет найти R-оптимальные управления в многокритериальной задаче.

Т е о р е м а 2. Если u^* – любое оптимальное управление в задаче $< U, (\lambda, H), R' >$, где $\lambda \in E_n$, то $u^* \in \sigma(U, H, R)$ для любого λ-отделимого отношения R.

Доказательство. Достаточно заметить, что $H(u^*) \in$ Arg max (λ, H), и далее повторить схему доказательства теоремы 1.

3.1.3. Эффективные и слабоэффективные оценки и решения

При исследовании сложных систем критерий для оценки управленческих решений, как правило, является вектором, поэтому выбор наилучшего решения – нетривиальная задача. В многокритериальной задаче максимизация векторной оценки довольно легко установить предпочтительность одной оценки другой, если эти оценки отличаются только одной компонентой. Для этого достаточно сравнить несовпадающие компоненты по отношению > («не меньше») и отдать предпочтение той оценке, у которой соответствующая компонента больше.

Если отношение нестрогого предпочтения R транзитивно, то для любых двух векторных оценок y и y', таких, что $y_i > y_i'$, $(i = 1, 2, ..., n)$, используя отношение >, для их компонент можно написать:

$$
\begin{aligned}
&(y_1, y_2, ..., y_n)R(y_1', y_2, ..., y_n);\\
&(y_1', y_2, ..., y_n)R(y_1', y_2', ..., y_n);\\
&... ;\\
&(y_1', y_2', ..., y_{n-1}, y_n)R(y_1', y_2', ..., y_n).
\end{aligned} \tag{3.4}
$$

Из этих соотношений и транзитивности R следует, что верно yRy', т.е. векторная оценка y не менее предпочтительна, чем y'.

Предположение о транзитивности отношения R является настолько естественным, что часто формулируется в виде аксиомы, называемой аксиомой Парето.

Пусть U – множество допустимых альтернатив (управлений). Каждое из управлений $u \in U$ оценивается с помощью векторного критерия $H(u) = \{H_1(u), ..., H_i(u), ..., H_n(u)\}$ (предположим, что степень предпочтительности управления увеличивается с возрастанием компонент вектора H). Введем в пространстве оценок E_n отношение строгого (>) и нестрогого (≥) предпочтения. Будем говорить, что вектор $H' = \{H'_i\} > H'' = \{H_i''\}$, тогда и только тогда, когда $H_i' > H_i''$ для всех $i = 1, 2, ..., n$.

Вектор $H' > H''$ тогда и только тогда, когда $H'_i >= H_i''$ для всех $i = 1, 2, ..., n$. и хотя бы для одного i верно $H_i' > H_i''$.

Два решения u^1 и u^2 называются эквивалентными, если $H(u^1) = H(u^2)$.

Обозначим через $\lambdabar = \{H(u)$: $u \in U\}$ множество оценок для всех возможных значений $u \in U$. Довольно очевидно, что если найдется такой вектор $H^* \in \lambdabar$, что $H^* \geq H$ для всех $H \in \lambdabar$, то решение u^*, для которого $H(u^*) = H^*$, следует считать наилучшим (поскольку оно явится наилучшим по всем компонентам векторного критерия H среди решений $u \in U$).

Векторная оценка $H^* \in \lambdabar$ называется максимальной для отношения ≥ (>), если не существует оценки $H \in \lambdabar$, такой, что $H \geq H^*$ $(H > H^*)$. Оценка, максимальная по ≥, называется оптимальной по Парето или эффективной оценкой, а соответствующее решение u^* – оптимальным по Парето или эффективным.

Таким образом, оптимальное по Парето решение обладает тем свойством, что если u^* – Парето оптимальное решение, то из условия $H_i\{u'\} \geq H_i(u^*)$, $i = l, ..., n$, должно следовать $H(u') = H(u^*)$.

Множество оценок $H^p \in H$, удовлетворяющих этому условию, называется множеством Парето, или эффективным, а множество соответствующих решений $P(U) \in U$ называется множеством эффективных решений, или Парето оптимальным множеством, т.е.

$$P(U) = \{u : u \in U, H(u) \in H^p\}.$$

Векторная оценка $H^s \in \lambdabar$, максимальная по >, называется слабоэффективной или слабооптимальной по Парето, или оптимальной по Слейтеру, а соответствующее решение u – оптимальным по Слейтеру или слабоэффективным. Таким образом, оптимальное по Слейтеру решение обладает тем свойством, что не существует никакого другого решения $u' \neq u \in U$, которое превосходит его в смысле порядка > по всем компонентам критерия H. Иными словами, если u оптимальное по Слейтеру, то не существует такого $u' \in U$, что $H_i\{u'\} > H_i(u^*)$, $i = 1, 2, ..., n$.

Множество оценок $\lambdabar^s \subset \lambdabar$, оптимальных по Слейтеру [7, 8, 11], называется слабоэффективным множеством, а множество соответствующих решений $S(u) \subset U$ – слабоэффективным множеством решений, т.е. $S(U) = \{u$: $u \in U$, для которых не существует $u' \in U$, таких, что $H_i(u') > H_i(u)$, $t = 1, 2, ..., n\}$.

Поскольку из $H > H'$ следует $H \geq H'$, то всякая эффективная оценка слабоэффективна, т.е. $\lambdabar^p \subset \lambdabar^s$, $P(u) \subset S(U)$.

Проиллюстрируем введенные понятия на простом примере. Пусть множество $\lambdabar \subset R^2$ имеет вид, как на рис. 3.5 (случай двух критериев).

Множество $\lambdabar^p$ совпадает с северо-восточной границей множества $\lambdabar$ (на рис. 3.5 оно состоит из кривых ab, cd (без точки c), fg, а множество H^s состоит из отрезков кривой $abed$ и etg).

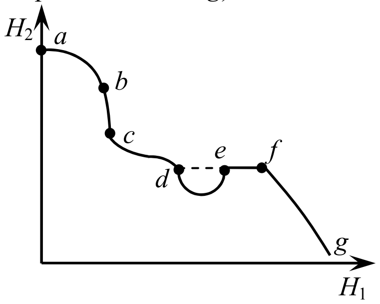

Рис. 3.5. Вид множеств $\lambdabar^p$, $\lambdabar^s$

Основной задачей многокритериальной оптимизации является выделение оптимального решения из множества всех решений. Естественно, что хорошим следует считать метод, когда выделенное решение оказывается эффективным или слабоэффективным. Рассмотрим некоторые методы выбора оптимального решения, основанные на скаляризации многокритериальной задачи. Первый метод состоит в том, что вначале находят точки максимума u^i для каждого из критериев в отдельности, а затем в качестве оптимального решения выбирают такое значение u^*, которое минимизирует максимальное отклонение оценки $H_i(u)$ от соответствующих максимальных значений $H_i(u')$. Обозначим

$$Y_i^* = \max_{u \in U} H_i(u). \tag{3.5}$$

Рассмотрим выражение

$$\max_{1\le i\le n}\frac{y_i^*-H_i(u)}{\left|y_i^*\right|}, \tag{3.6}$$

оценивающее максимальное отклонение оценки $H(u)$ произвольного решения $u \in U$ от вектора $y^* = \{y_1,\ y_2^*,\ ...,\ y_n^*\}$, представляющего собой вектор максимумов по каждому критерию. В качестве оптимальной точки $u^* \in U$ предлагается выбрать точку u^*, минимизирующую выражение (3.6), т.е. u^* выбирается из условия

$$\max_{1\le i\le n}\frac{y_i^*-H_i(u^*)}{\left|y_i^*\right|}=\min_{u\in U}\ \max_{1\le i\le n}\frac{y_i^*-H_i(u)}{\left|y_i^*\right|}.$$

Решение u^* всегда слабоэффективно, а если оно единственно (с точностью до эквивалентности), то и эффективно. Действительно, предположим, что u^* не является слабоэффективным. Тогда существует решение $u^0 \in U$, такое, что $H_i(u^0) > H_i(u^*)$ для всех $i = 1, 2, ..., n$. Отсюда следует, что

$$\frac{y_i^*-H_i(u^*)}{\left|y_i^*\right|}>\frac{y_i^*-H_i(u^0)}{\left|y_i^*\right|},\qquad i=1,\,2,\,\ldots,\,n;$$

$$\max_{1\le i\le n}\frac{y_i^*-H_i(u^*)}{\left|y_i^*\right|}=\min_{u\in U}\ \max_{1\le i\le n}\frac{y_i^*-H_i(u)}{\left|y_i^*\right|}>\max_{1\le i\le n}\frac{y_i^*-H_i(u^0)}{\left|y_i^*\right|}.$$

Последнее неравенство означает, что u^* не является решением, минимизирующим выражение (3.6). Полученное противоречие доказывает, что решение u^* является слабоэффективным. Если u^* – единственное решение, минимизирующее выражение (3.6), то для любого $u \in U$ выполняется неравенство

$$\max_{1\le i\le n}\frac{y_i^*-H_i(u^*)}{\left|y_i^*\right|}<\max_{1\le i\le n}\frac{y_i^*-H_i(u)}{\left|y_i^*\right|}.$$

Это означает, что для любого $u \in U$ найдется номер i_0, такой, что $H_{i_0}(u^*) > H_{i_0}(u)$, т.е. не существует решения $u \in U$, для которого $H(u) \ge H(u^*)$, и, следовательно, u^* является эффективным.

Характерным методом решения задач многокритериальной оптимизации является метод главного критерия. Он состоит в том, что исходная многокритериальная задача сводится к задаче оптимизации по одному критерию с заданными ограничениями на остальные. Выбранный критерий H_j называется главным. Для остальных критериев задаются некоторые пороговые значения h_i, т.е. многокритериальная задача сводится к задаче

$$\max_{u\in U} H_j(u),$$

$$H_i(u) \ge h_i,\ i = 1, 2, ..., n,\ i \ne j. \tag{3.7}$$

Любое решение u^0 этой задачи считается оптимальным решением многокритериальной задачи. Заметим, что такое решение всегда оказывается слабоэффективным, а если оно единственно (с точностью до эквивалентности), то и эффективным. Действительно, если существует решение $\bar{u}$ такое, что $H_i(\bar{u}) > H(u^0)$ для всех $i = 1, 2, ..., n$, тогда очевидно для $i \ne j$ $H_i(\bar{u}) > h_i$, а для главного критерия $H_j(\bar{u}) > \max H_j(u) = = H_j(u^0)$, но это противоречит тому, что u^0 является решением задачи (3.7), т.е. решение u^0 является слабоэффективным. Нетрудно также убедиться, что это решение эффективно, если оно единственно (с точностью до эквивалентности).

При использовании метода главного критерия на практике обычно берут несколько наборов пороговых значений $\{h_i\}$ и для каждого набора решают задачу (3.7). Может оказаться, что для некоторых наборов система ограничений в задаче (3.7) окажется несовместной. Это означает, что пороговые значения слишком высоки. После решения задачи (3.7) для каждого набора пороговых значений производят окончательное назначение величин h_i и определяют оптимальное решение.

Такой подход к нахождению оптимального решения носит эвристический характер. Поэтому метод главного критерия целесообразно применять в том случае, если имеются какие-либо соображения о подходящих значениях пороговых величин h_i. Преимущество метода главного критерия заключается в том, что он позволяет ограничиться рассмотрением сравнительно небольшой части всего множества эффективных решений.

Еще одним методом выбора оптимального решения являются так называемые арбитражные схемы. Метод формулируется при некоторых предположениях о структуре множества λ и функциях $H_i(u)$, $i = 1, 2, ..., n$. Однако он может быть применен и в более общем случае.

Будем считать, что множество λ всевозможных оценок выпукло и компактно в E_n. Рассмотрим некоторое исходное допустимое решение $u^0 \in U$, которое называется консервативным и подлежит улучшению при решении данной многокритериальной задачи. Значения вектора оценок $H(u)$ в точке $u^0 \in U$: $H(u^0) = \{H_1(u^0), H_2(u^0), ..., H_n(u^0)\}$ называется точкой статус-кво.

Под арбитражной схемой понимается правило φ, которое каждой паре (λ, $H(u^0)$) ставит в соответствие некоторую пару ($\bar{u}$, $\bar{H}$) = φ (λ, $H(u^0)$), где $\bar{H} \in \lambda$, $u \in U$ и $H = H(\bar{u})$ ($\bar{u}$ интерпретируется как оптимальное решение).

Таким образом, применение арбитражной схемы для определения оптимального решения предполагает, что правило φ позволяет, отталкиваясь от какого-либо допустимого решения, перейти к оптимальному. Ясно, что выбор оптимального решения зависит от выбора точки статус-кво. Она интерпретируется как некоторое решение, удобное из каких-либо соображений в качестве начального приближения к оптимальному. Например, в методе главного критерия, который может быть сведен к арбитражной схеме [7, 8, 11], точка статус-кво может быть задана с использованием пороговых значений $\{h_i\}$, а величина h_j (пороговое значение для главного критерия) выбрана как допустимое решение задачи (3.7).

Для того чтобы правило φ приводило к оптимальному решению, оно должно удовлетворять некоторым требованиям, которые мы сформулируем в виде аксиом. Пусть заданы выпуклое замкнутое подмножество λ, точка $H(u^0) \in \lambda$ и пара $(u, H) = \varphi(\lambda, H(u^0))$.

А к с и о м а 1. Реализуемость: $\overline{H} \in \lambda$, $\overline{u} \in U$, $\overline{H} = H(\overline{u})$.

А к с и о м а 2. Индивидуальная рациональность: $\overline{H} \geq H(u^0)$.

А к с и о м а 3. Оптимальность по Парето: если $H \in \lambda$ и $H \geq \overline{H}$, то $\overline{H} = H$.

А к с и о м а 4. Независимость от посторонних альтернатив: если $\overline{H} \in A \subset \lambda$ и $(\overline{H}, \overline{u}) = \varphi(\lambda, H, (u^0))$, то $(\overline{H}, \overline{u}) = \varphi(A, H(u^0))$.

Первая аксиома означает, что решение, полученное в результате применения правила φ, должно быть допустимым. Аксиома 2 предполагает, что полученное решение должно быть не менее предпочтительно по каждому из критериев, чем консервативное решение. Третья аксиома говорит о том, что полученное решение должно быть эффективным. Аксиома 4 означает, что если даже имеются большие возможности для выбора $(\overline{H}, \overline{u})$, можно выбрать это решение при меньших возможностях, если этот вектор реализуем.

Будем для простоты считать, что в множестве λ существует вектор H, каждая i-я координата которого строго больше $H_i(u^0)$, т.е. решение u^0 не является эффективным.

Имеет место следующая теорема.

Т е о р е м а 3. Функция φ может быть выбрана в виде

$$\varphi(\lambda, H(u^0)) = \{(\overline{H}, \overline{u}): \max_{\substack{H \geq H(u^0) \\ H \in \lambda}} g(H, \lambda, H(u^0))\} =$$

$$= g(\overline{H}, (\overline{u}), \lambda, H(u^0)),$$

где $g(\overline{H}, \lambda, H(u^0)) = \prod_{i=1}^{n} (H_i - H_i(u^0))$ удовлетворяет аксиомам 1 – 4.

Построение отображения φ в виде (3.8) означает не что иное, как максимизацию скаляризованного критерия, который в данном случае представляет собой произведение приращений по каждому из критериев. Поскольку функция $g(H, \lambda, H(u^0))$ удовлетворяет аксиомам 1 – 4, то применение арбитражной схемы с функцией φ, выбранной в виде (3.8), позволяет получить оптимальное решение многокритериальной задачи.

В предыдущем параграфе мы уже говорили о том, что основным методом решения многокритериальных задач является скаляризация векторного критерия и решение λ-свертки многокритериальной задачи.

Для того чтобы использовать этот метод для нахождения эффективных и слабоэффективных решений, необходимо установить, для каких значений $\lambda \in E_n$ отношение Парето и Слейтера λ-отделимо. Справедливы следующие леммы.

Л е м м а 2. Отношение Парето (≥) λ-отделимо при любых $\lambda \in E_n$ с положительными компонентами.

Доказательство. Выберем в множестве λ две любые оценки H^1 и H^2, такие, что $H^1 \geq H^2$. Это означает, что для всех $j = 1, 2, ..., n$ $H^1 \geq H^2$ и существует j, такое, что $H^1 > H^2$. Отсюда для любых $\lambda_i > 0$ справедливо неравенство

$$\sum_{i=1}^{n} \lambda_{iH_i^1} > \sum_{i=1}^{n} \lambda_i H_i^2 ,$$

т.е. отношение Парето λ-отделимо.

Л е м м а 3. Отношение Слейтера (>) λ-отделимо для всех неотрицательных $\lambda \in E_n$.

Доказательство. Необходимо повторить схему доказательства леммы 2 с учетом свойств отношения Слейтера.

Пусть λ^P и λ^S – соответственно, множества эффективных и слабоэффективных оценок. Справедлива следующая теорема.

Т е о р е м а 4. Для любых $\lambda \in E_n$ и $\mu \in E_n$ таких, что $\lambda > 0$, $\mu \geq 0$, справедливо

$$\operatorname{Arg}\max_{H \in \lambda} (\lambda, H) \subset \lambda^P ,$$

$$\operatorname{Arg}\max_{H \in \lambda} (\mu, H) \subset \lambda^S .$$

Доказательство непосредственно следует из лемм 2 и 3.

Теорема 4 позволяет находить эффективные и слабоэффективные решения с помощью максимизации свертки критериев.

Можно показать, что при некоторых ограничениях на множество $\tilde{\lambda}$ любая эффективная или слабоэффективная оценка является точкой максимума некоторой свертки.

Определение. Множество S называется выпуклым, если для любых $x', x'' \in S$ точка $kx' + (1 - k)x'' \in S$ при всех $k \in [0, 1]$. Множество S называется слабовыпуклым, если выпуклым будет множество $S + A$, где $A = \{y : y : y \le 0\}$.

Теорема 5. Пусть $\tilde{\lambda}$ слабовыпукло. Оценка $H^* \in \tilde{\lambda}$ эффективна (слабоэффективна) тогда и только тогда, когда существует такой вектор λ с положительными (неотрицательными) коэффициентами, $\sum_{i=1}^{n} \lambda_i = 1$, что $(\lambda, H) \ge (\lambda, H^*)$ для всех $H \in \tilde{\lambda}$.

Контрольные вопросы

1. Дайте характеристику общей постановки задачи выбора в многокритериальных и иерархических системах.
2. Что такое альтернатива?
3. Дайте характеристику отношению.
4. Что такое эффективные и слабоэффективные оценки?

3.2. МНОГОКРИТЕРИАЛЬНЫЕ ЗАДАЧИ УПРАВЛЕНИЯ

3.2.1. Многокритериальные задачи оптимального управления

Большинство моделей сложных систем характеризуются тем, что описываемые ими процессы носят динамический характер, т.е. компоненты системы характеризуются величинами, изменяющимися во времени. Функцией времени является также и управление в этих системах [16, 19].

Рассмотрим следующую систему управления.

Пусть состояние системы описывается вектором $x \in E_m$. В начальный момент t_0 система находится в состоянии $x(t_0) = x^0$. Предположим, что динамика изменения компонент системы на отрезке времени $[t_0, T]$ описывается векторным дифференциальным уравнением

$$\dot{x} = f(x, u),\ x \in E_m,\ u \in U \subset \mathrm{Comp}\, E_n, \tag{3.9}$$

где u – управляющий параметр, имеющий смысл внешних воздействий, с помощью которых происходит управление развитием.

Будем считать, что параметр u выбирается непрерывно во времени и получившаяся в результате функция $u(t)$, $t \in [t_0, T]$, $u(t) \subset U$, измерима по t. Будем также считать выполненными все условия, гарантирующие существование, продолжимость и единственность решения дифференциального уравнения (3.9) при любом измеримом управлении $u(t)$ на отрезке времени $[t_0, T]$ и при начальном условии $x(t_0) = x^0$. Управление $u(t)$, являющееся только функцией времени, называется программным. Каждое программное управление $u(t)$, $t \in [t_0, T]$, определяет некоторую траекторию движения $x(t)$, $t \in [t_0, T]$, получаемую как решение уравнения (3.9) при начальном условии $x(t_0) = x^0$.

Обозначим через U множество допустимых управлений. Выбирая различные управления из множества U, получим различные траектории. Пусть $C^{T-t_0}(x^0)$ – множество достижимости уравнения (3.9), т.е. множество точек E_m, в которые может попасть решение уравнения (3.9) из начального состояния в момент времени T при использовании всевозможных программных управлений $u(t) \subset U$, $t \in [t_0, T]$. Иными словами, множество достижимости есть множество концов траекторий дифференциального уравнения (3.9) $\{x(T)\}$, исходящих из начального состояния x^0 при всевозможных программных управлениях $u(t) \in U$, $t \in [t_0, T]$.

Предположим далее, что качество траектории определяется точкой $x(T)$, в которую переходит система в результате этого развития в конечный момент T. Таким образом, будем считать, что на множестве достижимости $C^{T-t_0}(x^0)$ задан векторный критерий $H(x(T))$, $x(T) \in C^{T-t_0}(x^0)$, определяющий качество траектории $x(t)$ и соответствующего управления $u(t)$. Приходим к динамической многокритериальной задаче оптимизации, в которой множество вариантов решения или исходов обозначим

$$\chi(x^0, T - t_0) = \{x(T) : x(T) \in C^{T-t_0}(x^0)\}, \tag{3.10}$$

а множество оценок будет иметь вид

$$\tilde{\lambda}(x^0, T - t_0) = \{H(x(T)) : x(T) \in C^{T-t_0}(x^0)\}. \tag{3.11}$$

Множества $\tilde{\lambda}$ и χ зависят от параметров x^0, t_0, представляющих собой начальные условия для задачи (3.9). Поскольку решение сформулированной динамической многокритериальной задачи оптимизации зависит от x^0, $T - t_0$, будем обозначать ее через $\Gamma(x^0, T - t_0)$.

Итак, мы имеем динамическую многокритериальную задачу оптимизации $\Gamma(x^0, T - t_0)$, множество всевозможных исходов $\chi(x^0, T - t_0)$, определенных в (3.10), и множество всевозможных оценок $\tilde{\lambda}(x^0, T - t_0)$, определенных в (3.11). Пусть даны: $\tilde{\lambda}^P(x^0, T - t_0) \in \tilde{\lambda}(x^0, T - t_0)$ – множество эффективных оценок (Парето-оптимальных), $\tilde{\lambda}^S(x^0, T - t_0) \subset \tilde{\lambda}(x^0, T - t_0)$ – множество слабоэффективных оценок (оптимальных по Слейтеру) и $P(\chi(x^0, T - t_0))$ и $S(\chi(x^0, T - t_0))$ – соответствующие множества оптимальных управлений.

Пусть $\overline{u}(t)$ – некоторое оптимальное управление, а $\overline{x}(t)$ – соответствующая этому управлению оптимальная траектория. Рассмотрим в каждый момент времени $t \in [t_0, T]$ задачу многокритериальной оптимизации с начальными условиями t, $\overline{x}(t)$, которую обозначим через $\Gamma(\overline{x}(t), T-t)$.

В общем случае траектория $\overline{x}(\tau)$, $t \le \tau \le T$ не обязательно является оптимальной в задаче $\Gamma(\overline{x}(t), T-t)$. Такое свойство оптимальных решений называется динамической неустойчивостью. С другой стороны, если $\overline{x}(\tau)$ является оптимальной траекторией в текущей задаче многокритериальной оптимизации $\Gamma(\overline{x}(t), T-t)$ $\forall\ t \in [t_0, T]$, то оптимальное управление $\overline{u}(t)$ и траекторию $\overline{x}(t)$ называют динамически устойчивыми.

Если динамически устойчивыми оказываются все оптимальные управления, то говорят о динамической устойчивости решения и принципа оптимальности.

Вопрос динамической устойчивости тесно связан с выбором принципа оптимальности. Существует целый ряд принципов оптимальности, для которых оптимальные решения оказываются динамически устойчивыми. К таким принципам относятся оптимальность по Парето и Слейтеру, равновесие по Нэшу и ряд других. Существуют также принципы оптимальности, которые не обладают свойством динамической устойчивости.

В рамках данного параграфа мы покажем динамическую устойчивость Парето-оптимальных решений. Действительно, пусть $\tilde{\lambda}^P(x^0, T-t_0)$ – Парето-оптимальное множество оценок и $P(\chi(x^0, T-t_0))$ – Парето-оптимальное множество решений (управлений) в многокритериальной динамической задаче оптимизации $\Gamma(x^0, T-t_0)$ для начального состояния x^0 с предписанной продолжительностью $T - t_0$. Пусть $\{H_i^*\} = H^*$ – вектор оценок из множества $\tilde{\lambda}^P(x^0, T-t_0)$. Предположим, что выбраны управление $u(t)$ и соответствующая траектория $x(t)$, при которых в конце процесса реализуется оценка $H^* = \{H_i^*\}$. Это означает, что управление $u(t)$ таково, что $x(t)$ в момент времени T (в момент окончания процесса) проходит через точку $x(T)$, в которой $H(x(T)) = \{H_i(x(T))\}$ равно как раз вектору полезностей $H^* = \{H_i^*\}$. Пусть C^{T-t_0} – множество достижимости управляемой системы из начального состояния x^0. Рассматривая изменение этого множества вдоль траектории $x(\tau)$, можно заметить, что

$$C^{T-\tau_1}(x(\tau_1)) \supset C^{T-\tau_2}(x(\tau_2)), \quad t_0 \le \tau_1 < \tau_2 \le T. \tag{3.12}$$

Из (3.12) имеем

$$\tilde{\lambda}(x(\tau_1), T-\tau_1) \supset \tilde{\lambda}(x(\tau_2), T-\tau_2). \tag{3.13}$$

Поскольку вектор $H^* = \{H_i^*\}$ принадлежит Парето-оптимальному множеству, то не существует такого вектора $H' \ne H^*$, принадлежащего $\tilde{\lambda}(x^0, T-t_0)$, что $H_i' \ge H_i^*$ для всех $i = 1, 2, ..., n$. Поэтому из (3.13) следует, что тем более это имеет место для множества $\tilde{\lambda}(x(\tau), T-\tau)$, $t_0 \le \tau \le T$. Следовательно, вектор оценок $H^* = \{H_i^*\}$ не доминируется ни одним из векторов множества $\tilde{\lambda}(x(\tau), T-\tau)$, или, что то же, принадлежит Парето-оптимальному множеству текущей задачи с начальным условием $x(\tau)$ и продолжительностью $T - \tau$. Таким образом, вектор H^* во всех текущих задачах остается Парето-оптимальным при движении системы вдоль оптимальной траектории $x(\tau)$. Поскольку вектор полезностей H^* был выбран произвольно из множества $\tilde{\lambda}^P(x^0, T-t_0)$, то это означает динамическую устойчивость любого Парето-оптимального решения, а следовательно, и Парето-оптимального множества в целом.

3.2.2. Принцип максимума в многокритериальных задачах

Мы уже говорили, что основным методом решения многокритериальных задач является их скаляризация и рассмотрение λ-свертки исходной задачи. Это позволяет заменить многокритериальную задачу оптимального управления на задачу с одним критерием, а затем воспользоваться для отыскания оптимальных решений принципом максимума Понтрягина. Рассмотрим систему дифференциальных уравнений

$$\dot{x} = F(x,u),\ x \in E_m,\ u \in U,$$

$$x(t_0) = x^0,\ t \in [t_0, T]. \tag{3.14}$$

Критерий качества управления выберем в виде

$$K(x,u) = \int_{t_0}^{T} f_0(x,u)dt + H_0(x(T)). \tag{3.15}$$

Однокритериальной задачей оптимального управления называется задача

$$K(x,u) \to \max,$$

$$\dot{x} = f(x,u),\ x(t_0) = x^0,\ u \in U. \tag{3.16}$$

Если в правой части выражения (3.15) $f_0(x,u,t) \equiv 0$, то критерий качества называется терминальным, а если $H_0(x(T)) \equiv 0$ – интегральным.

Существует стандартный способ сведения интегрального критерия к терминальному с помощью увеличения числа фазовых переменных. В однокритериальных задачах поступают следующим образом. Положим

$$x_0(t) = \int_{t_0}^{t} f_0(x(\tau), u(\tau))d\tau .$$

Тогда $x_0(t)$ удовлетворяет дифференциальному уравнению

$$\dot{x}_0(t) = f_0(x(t), u(t)) \tag{3.17}$$

при начальном условии $x(t_0) = 0$.

Добавим уравнение (3.17) с начальным условием к системе (3.14). Критерий качества (3.15) примет вид

$$K(x,u) = x_0(T) + H_0(x(T)) .$$

Очевидно, что полученный критерий является терминальным. Поэтому, не умаляя общности, рассмотрим однокритериальную задачу с терминальным критерием

$$H_0(x(T)) \to \max ,$$

$$\dot{x} = f(x,u),\ \ x(t_0) = x^0,\ \ u \in U . \tag{3.18}$$

Сформулируем для задачи (3.18) принцип максимума ***Понтрягина***. Пусть $u^*(t)$, $x^*(t)$ – оптимальные управления и траектория в задаче (3.18). Тогда существует вектор-функция $\psi(t) = (\psi_1(t), ..., \psi_m(t))$, такая, что если мы определим функцию Гамильтона

$$B(x,u,\psi) = \sum_{i=1}^{m} \psi_i(t) f_i(x,u) , \tag{3.19}$$

то будут выполняться следующие необходимые условия:

$$\psi_i(t) = -\sum_{j=1}^{m} \psi_j(t) \frac{\partial f_j(x^*, u^*)}{\partial x_i} , \quad i = 1, 2, \ldots, m; \tag{3.20}$$

$$\psi_i(T) = \frac{\partial H_0(x^*(T))}{\partial x_i} , \quad i = 1, 2, \ldots, m; \tag{3.20}$$

$$B(x^*, u^*, \psi(t)) = \max_{u \in U} B(x, u, \psi(t)) . \tag{3.22}$$

Принцип максимума дает нам необходимые условия оптимальности и позволяет найти управления, которые могут оказаться оптимальными. Для того чтобы найти эти управления, необходимо сделать следующее:

1) записать функцию Гамильтона в виде (3.19), подставив соответствующие функции $f_i(x,u)$ из системы дифференциальных уравнений (3.14);
2) выразить с помощью условия (3.22) управление и как функцию остальных переменных:

$$u = u(x, \psi) ; \tag{3.23}$$

3) записать систему дифференциальных уравнений

$$\dot{\psi}_i = -\frac{\partial B}{\partial x_i} , \quad \psi_i(T) = \frac{\partial H_0(x(T))}{\partial x_i} , \quad i = 1, 2, \ldots, m;$$

$$\dot{x}_i = \frac{\partial B}{\partial \psi_i} , \quad x_i(t_0) = x_i^0 , \quad i = 1, 2, \ldots, m; \tag{3.24}$$

4) подставить в систему (3.24) управление (3.23) и найти ее решение $x^*(t)$, $\psi^*(t)$;
5) подставить найденные функции $x^*(t)$, $\psi^*(t)$ в (3.23) и найти управление $u^*(t)$.

Перейдем теперь к рассмотрению многокритериальной задачи. Пусть динамика системы описывается векторным уравнением

$$\dot{x} = f(x,u),\ \ x(t_0) = x^0,\ \ u \in U . \tag{3.25}$$

Относительно управления u выполнены все предположения, обеспечивающие существование и единственность решения системы (3.25). Введем вектор оценок с компонентами

$$H_i(x, u) = H_i(x(T)), \quad i = 1, 2, ..., n, \tag{3.26}$$

где функции $H_i(x(T))$ непрерывно дифференцируемы. Пусть $C^{T-t}(x^0)$ – множество достижимости системы (3.25), а

$$\tilde{\lambda}\ (x(t_0), T - t_0) = \{H(x(T)): x(T) \in C^{T-t}(x^0)\} -$$

множество оценок.

Для отыскания решения многокритериальной задачи (3.25), (3.26), которую обозначим $T(x^0, T - t_0)$, воспользуемся ее представлением в виде λ-свертки, т.е. рассмотрим однокритериальную задачу

$$\sum_{i=1}^{n} \lambda_i H_i(x(T)) \to \max ,$$

$$\dot{x} = f(x, u), \; x(t_0) = x^0, \; u \in U . \tag{3.27}$$

Выберем в качестве принципа оптимальности оптимальность по Парето. Из леммы 3 нам известно, что отношение Парето будет λ-отделимо при положительных λ. Поэтому будем считать, что в задаче (3.27) $\lambda_i > 0$ ($i = 1, 2, ..., n$).

Тогда из теоремы 4 получим, что любое оптимальное решение (3.27) будет Парето-оптимальным. Принцип максимума для задачи (3.27) формулируется следующим образом.

Пусть $u^*(t)$, $x^*(t)$ – оптимальные управление и траектория в задаче (3.27). Тогда существует вектор-функция $\psi(t) = (\psi_1(t), \psi_2(t), ..., \psi_m(t)$, такая, что если мы определим функцию Гамильтона

$$B(x,u,\psi) = \sum_{i=1}^{m} \psi_i(t) f_i(x,u) , \tag{3.28}$$

то будут выполняться следующие необходимые условия:

$$\dot{\psi}_i(t) = -\sum_{j=1}^{m} \psi_i(t) \frac{\partial f_i(x^*, u^*)}{\partial x_i}, \quad i = 1, 2, ..., m , \tag{3.29}$$

$$\psi_i(T) = \sum \lambda_i \frac{\partial H_j(x^*(T))}{\partial x_i}, \quad i = 1, 2, ..., m , \tag{3.30}$$

$$B(x^*, u^*, \psi(t)) = \max_{u \in U} B(x, u, \psi(t)) . \tag{3.31}$$

Алгоритм использования принципа максимума тот же, что и для однокритериальной задали.

Если в задаче (3.27) коэффициенты λ_i неотрицательны, то полученное оптимальное решение λ-свертки многокритериальной задачи в соответствии с теоремой 4 будет оптимальным по Слейтеру (или слабо эффективным).

Контрольные вопросы

1. Назовите последовательность действий, позволяющих найти управление на основе использования принципа максимума.
2. Какие условия оптимальности дает принцип максимума.

3.3. ЗАДАЧА СБЛИЖЕНИЯ С НЕСКОЛЬКИМИ ЦЕЛЕВЫМИ ТОЧКАМИ

В этом параграфе мы рассмотрим метод отыскания Парето оптимальных решений в многокритериальной задаче сближения с несколькими целевыми точками. Приложения этой задачи разнообразны – от чисто технических до экономических.

Пусть состояние некоторой системы описывается в начальный момент времени t_0 вектором $x^0 \in E_m$. Динамика развития системы на отрезке времени $[t_0, T]$ описывается векторным дифференциальным уравнением

$$\dot{x} = f(x,u), \; x \in E_m, \; u \in U \subset E_i . \tag{3.32}$$

Каждому начальному условию $x(t_0) = x^0$ и измеримому программному управлению $u(t)$ при $t \in [t_0, T]$ соответствует (при выполнении некоторых требований к правой части уравнения (3.32)) единственная траектория, определенная на отрезке $[t_0, T]$.

Пусть $C^{T-t_0}(x^0)$ – множество достижимости системы (3.32), т.е. множество состояний $x(T)$, в которых может оказаться система в момент времени T при всевозможных допустимых управлениях. Предположим, что это множество выпукло и компактно и имеет гладкую границу.

Будем считать для удобства изложения, что управлением в модели распоряжается некоторый субъект управления, который называем центром A_0, а конечное состояние системы $x(T)$ оценивается несколькими «экспертами» $B_1, B_2, ..., B_n$. Каждый из экспертов в соответствии со своим пониманием стоящих перед системой целей развития определяет целевую точку M_i. Целевые точки M, могут находиться достаточно близко друг от друга, однако предположение об их совпадении было бы слишком сильной идеализацией.

В результате полезность каждой точки $x(T) \in C^{T-t_0}(x^0)$ может быть оценена центром с точки зрения ее близости к целевым точкам $M_1, ..., M_n$. Поэтому получаем связанный с каждой точкой $x(T)$ вектор оценок, который называется также *вектором экспертных оценок*:

$$\{-\rho(x(T),\ M_i)\} = \{H_i(x(T))\}\ . \quad (3.33)$$

где ρ – евклидово расстояние между точками $x(T)$, M_i.

Математическая задача сводится к нахождению оптимальных траекторий развития в смысле векторного критерия $H(x)$. Напомним, что управление $\overline{u}(t)$ называется оптимальным по Парето, если не существует такого управления $u(t)$, что

$$H_i(\overline{x}(T),\ \overline{u}(t)) \le H_i(x(T),\ u(t)), \quad i = 1,\ 2,\ ...,\ n\ ,$$

и хотя бы для одного i_0

$$H_{i_0}(\overline{x}(T),\ \overline{u}(t)) < H_{i_0}(x(T),\ u(t))\ ,$$

где $H(x(T),\ u(t))$ – вектор оценок при использовании управления $u(t)$, а $x(t)$ – соответствующая траектория.

Оптимальное по Парето управление мы будем обозначать через u^P, а соответствующую траекторию – через $x^P(t)$.

Определим структуру множества Парето в рассматриваемой задаче.

В множестве достижимости $C^{T-t_0}(x^0)$ у каждого эксперта B_i имеется своя наилучшая точка $x(T)$, такая, что

$$\rho(x(T), M_i) = \min_{\xi \in C^{T-t_0}(x^0)} \rho(\xi, M_i)\ .$$

Однако, вообще говоря, ни один из B_i, не может гарантировать достижения своего наилучшего результата. Можно ожидать, и это естественно, что в результате анализа экспертных оценок в начале процесса центр ограничится рассмотрением множества терминальных точек, заключенных в некотором смысле «между» наилучшими точками для каждого из B_i.

Покажем, что именно множество точек такой структуры соответствует множеству оптимальных по Парето решений.

Каждой точке $x(T) \in C^{T-t_0}$ соответствует вектор

$$H(x(T)) = (-\rho(x(T), M_1),\ ...,\ -\rho(x(T), M_n))\ .$$

Обозначим через

$$\tilde{\lambda}(t_0, \lambda^0) = \{H(x(T)) : x(T) \in C^{T-t_0}(x^0)\}$$

множество всевозможных реализуемых в момент T полезностей, а подмножество множества $\tilde{\lambda}(t_0, x^0)$, соответствующее множеству оптимальных по Парето управлений, обозначим через $P(x^0)$:

$$P(x^0) = \{H(x^p(T))\}\ .$$

Пусть $\hat{M} = \mathrm{conv}\{M_i,\ i = 1, 2, ..., n\}$, где $\mathrm{conv}\{M_i,\ i = 1, 2, ..., n\}$ – выпуклая оболочка точек M_i.

Обозначим через π оператор ортогонального проектирования из пространства E_m на некоторое выпуклое компактное множество B. Под ортогональной проекцией точки $x \in E_m (x \overline{\in} B)$ на B будем понимать точку $\pi_B x \in B$, такую, что

$$\rho(x, \pi_B, x) = \min_{y \in B} \rho(x, y)\ . \quad (3.34)$$

Данную точку назовем образом, а точку x – прообразом оператора проектирования. Под ортогональной проекцией точки $x \in B$ на B будем понимать саму точку x, а под ортогональной проекцией $\pi_B A$ некоторого множества A на множество B – множество ортогональных проекций, входящих во множество A точек на B.

Имеют место следующие вспомогательные утверждения, которые мы приведем здесь без доказательств.

Утверждение 1. Пусть B – замкнутое выпуклое множество в E^m и $x \in E^m$ – некоторая точка, не принадлежащая B. Тогда для всех $y \in B$ справедливо неравенство

$$\rho(\pi_B x,\ y) \le \rho(x,\ y)\ .$$

Рассмотрим множество $\omega_\lambda = \lambda y^1 + (1-\lambda) y^2$, $y^1 \in E_m$, $y^2 \in E_m$, $\lambda \in [0,\ 1]$, и выберем в пространстве E^m точки x^1 и x^2, которые не принадлежат множеству ω_λ. Введем функцию

$$F(\lambda) = \rho(x^1, \omega_\lambda) - \rho(x^2, \omega_\lambda)\ .$$

Утверждение 2. Из $F(\lambda) \ge 0$ при λ, равных 0 и 1, следует, что $F(\lambda) \ge 0$ при всех $\lambda \in (0,\ 1)$.

Утверждение 3. Пусть Q^r – r-мерный замкнутый выпуклый многогранник в E_m с вершинами $Q_1, Q_2, ..., Q_q$. Тогда для всех $y \in Q^r$ и $x \overline{\in} Q^r$ существует хотя бы одно $i \in \{1,\ 2,\ ...,\ q\}$, такое, что $\rho(y, Q_i) < \rho(x, Q_i)$.

Утверждение 4. Пусть Q^r – r-мерный замкнутый многогранник в E_m с вершинами $Q_1, Q_2, ..., Q_q$, а x, y – некоторые точки пространства E_m, не принадлежащие Q^r. Если найдется точка $\xi \in Q^r$ для которой $\rho(y, \xi) < \rho(x, \xi)$, то хотя бы для одного $i \in \{1,\ 2,\ ...,\ q\}$ выполнено неравенство $\rho(y, Q_i) < \rho(x, Q_i)$.

В справедливости данных утверждений нетрудно убедиться, рассмотрев их геометрическую иллюстрацию в пространствах E_2 или E_3.

Рассмотрим различные случаи расположения точек $M_1, M_2, ..., M_n$ и области достижимости $C^{T-t_0}(x^0)$.

1. $\hat{M} \cap C^{T-t_0}(x^0) = \varnothing$. В этом случае целевые точки недостижимы (рис. 3.6).

Введем функцию

$$F(\xi) = \rho(y', \xi) - \rho(x, \xi)$$

и множество

$$Y = \{y \in \hat{M};\ y' \in \pi_{C^{T-t_0}(x^0)} y\}.$$

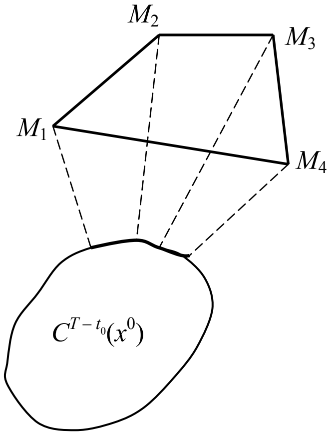

Рис. 3.6. Случай 1

Для определения возьмем $\bar{y} \in Y$:

$$\rho(y', \bar{y}) = \min_{y \in Y}(\rho(y', y)).$$

Такая точка существует, так как Y компактно. По определению точки $\bar{y}$ для всех x из (3.34) справедливо неравенство $F(\bar{y}) < 0$. Тогда из утверждения 3 вытекает неравенство

$$\rho(y', M_i) < \rho(x, M_i)$$

хотя бы для одного $i = 1, ..., n$, т.е. $H_i(y') > H_i(x)$. А это означает, что $H(y') \in P(x^0)$.

Покажем, что, кроме точек множества $\pi_{C^{T-t_0}(x^0)} \hat{M}$, в множестве $C^{T-t_0}(x_0)$ нет точек, обладающих свойством (3.35). Это значит, что для любых $x \in C^{T-t_0}(x^0) \setminus \pi_{C^{T-t_0}(x^0)} \hat{M}$ найдется точка $\tilde{y} \in \pi_{C^{T-t_0}} \hat{M}$, в которой

$$H_i(\tilde{y}) \geq H_i(x), \quad i = 1, 2, ..., n. \tag{3.35}$$

Рассмотрим множества

$$\hat{M}_1 = \mathrm{conv}(\hat{M}, \pi_{C^{T-t_0}(x^0)} \hat{M}); \qquad \hat{M}_2 = \mathrm{conv}\ \pi_{C^{T-t_0}(x^0)} \hat{M}.$$

Очевидно, что $\hat{M}_2 \subset \hat{M}_1$, согласно утверждению 3 для любых $x \in C^{T-t_0}(x^0) \setminus \hat{M}_1$ существует $\tilde{y} \in \pi_{C^{T-t_0}(x^0)} \hat{M}$, удовлетворяющий неравенству (3.35).

Пусть теперь $x \in \hat{M}_2 \setminus \pi_{C^{T-t_0}(x^0)} \hat{M}$, а $\pi_{\hat{M}} x$ – ее образ на $\hat{M}$. Поскольку множество $C^{T-t_0}(x^0)$ по определению выпукло и $\pi_{C^{T-t_0}(x^0)} \hat{M} \subset \overline{C}^{T-t_0}(x^0)$ (ввиду пустоты множества $\hat{M} \cap C^{T-t_0}(x^0)$, где $\overline{C}^{T-t_0}(x^0)$ – граница множества $C^{T-t_0}(x^0)$), то существует точка $x^1 \in \pi_{C^{T-t_0}(x^0)} \hat{M} \cap \omega_\lambda$, $\omega_\lambda = \lambda x + (1-\lambda)\pi_{\hat{M}} x$, $\lambda \in [0, 1]$. Эта точка является искомой, ибо она удовлетворяет неравенству (3.35).

Итак, только точки множества $\pi_{C^{T-t_0}(x^0)} \hat{M}$ обладают свойством (3.35). Следовательно,

$$P(x_0) = \{H(y) \mid y \in \pi_{C^{T-t_0}(x^0)} \hat{M}\}.$$

Таким образом, в этом случае Парето-оптимальные точки $x^P(t)$ являются проекциями выпуклой оболочки целевых точек.

2. $\hat{M} \subset C^{T-t_0}(x^0)$. Все целевые точки достижимы (рис. 3.7).

Из утверждения 3 следует, что в множестве $C^{T-t_0}(x^0)$, кроме точек $y \in \hat{M}$, нет точек, обладающих свойством

$$H_i(y) \geq H_i(x), \quad i = 1, 2, \ldots, n, \quad x \overline{\in} \hat{M},$$

ибо для любого $x \overline{\in} \hat{M}$ существует $\xi \in \hat{M}(\xi \in \pi_{\hat{M}} x)$, такое, что $H_i(\xi) \geq H_i(x)$, $i = 1, 2, \ldots, n$.

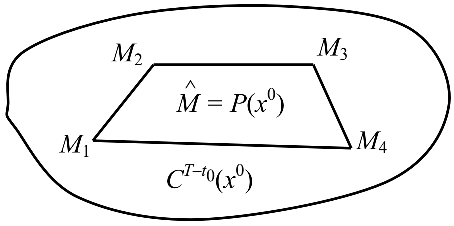

Рис. 3.7. Случай 2

Кроме того, для любого $x \overline{\in} \hat{M}$ и всех $y \in \hat{M}$ неравенство $H_i(y) > H_i(x)$ выполняется хотя бы для одного i.

Мы установили, что Парето-оптимальными могут быть только векторы выигрышей на множестве M. Покажем, что все векторы этого множества являются таковыми.

Пусть x, $y \in \hat{M}$, $x \neq y$. Отрезок с концами M_1 и M_2. Имеем

$$\rho(M_1, M_2) = \rho(M_1, y) + \rho(y, M_2) = \rho(M_1, x) + \rho(x, M_2),$$

но так как $x \neq y$, то

$$\rho(M_1, y) < \rho(M_1, x) \quad \text{либо} \quad \rho(M_2, y) < \rho(M_2, x).$$

Поэтому $H(y) \in P(x^0)$.

Пусть утверждение справедливо для $r = k - 1$. Рассмотрим луч z_x с началом в точке x, проходящий через y, если $r = k$.

Тогда существует точка $\bar{z} \in Z_x$, такая, что

$$\rho(x, \bar{z}) = \max_{z \in \hat{M} \cap Z_x} \rho(x, z).$$

Очевидно, что $\bar{z}$ лежит на границе $\hat{M}$, т.е. принадлежит $(k-1)$-мерной грани k-мерного многогранника $\hat{M}$. По индукции $H(z) \in P(x^0)$. Тогда по определению Парето-оптимального множества существует хотя бы одно i_0, для которого $H_{i_0}(\bar{z}) > H_{i_0}(x)$, или, что то же, $\rho(\bar{z}, M_{i_0}) < \rho(x, M_{i_0})$.

Учитывая, что $y \in \omega_\lambda = \lambda x + (1-\lambda)\bar{z}$, $\lambda \in [0, 1]$, имеем неравенство

$$\rho(y, M_{i_0}) = \rho(\bar{\lambda} x + (1-\bar{\lambda})\bar{z}, M_{i_0} \leq \bar{\lambda}\rho(x, M_{i_0}) +$$
$$+ (1-\bar{\lambda})\rho(\bar{z}, M_{i_0}) < \bar{\lambda}\rho(x, M_{i_0}) + (1-\bar{\lambda})\rho(x, M_{i_0}) = \rho(x, M_{i_0}),$$

т.е. $\rho(y, M_{i_0}) < \rho(x, M_{i_0})$.

Поменяв местами x и y, аналогично получим, что $H(x) \in P(x^0)$.

Таким образом, для всех $y \in \hat{M} \equiv \pi_{C^{T-t_0}(x^0)} \hat{M}$ и только для них $H(y) \in P(x^0)$. Следовательно, множество $P(x^0)$ имеет вид

$$P(x^0) = \{H(x) \mid x \in \pi_{C^{T-t_0}(x^0)} \hat{M} \equiv \hat{M}\}.$$

3. Пусть точки M_i, $i = 1, 2, \ldots, n$, расположены таким образом, что $\hat{M} \supset C^{T-t_0}(x^0)$ (рис. 3.8). Тогда

$$P(x^0) = \{H(x) \mid x \in \pi_{C^{T-t_0}(x^0)} \hat{M} = C^{T-t_0}(x^0)\}$$

(в этом случае все целевые точки недостижимы, но цели экспертов сильно отличаются друг от друга).

Доказательство этого факта следует из предыдущего случая, если множества $\hat{M}$ и $C^{T-t_0}(x^0)$ поменять местами.

4. Рассмотрим общий случай расположения точек M_i, $i = 1, 2, \ldots, n$ относительно множества $C^{T-t_0}(x^0)$ (рис. 3.9).

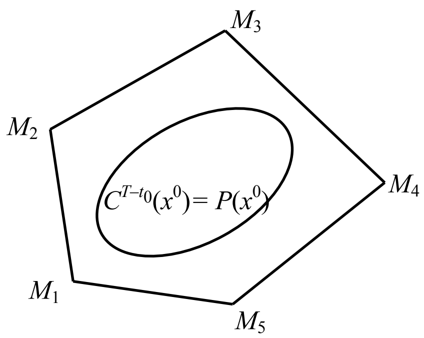

Рис. 3.8. Случай 3

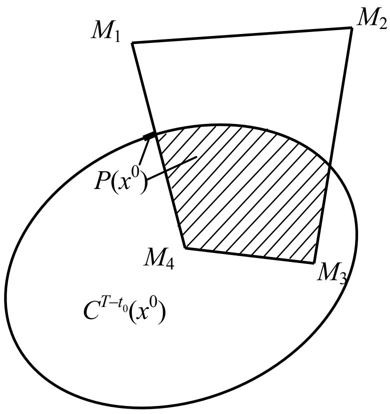

Рис. 3.9. Случай 4

Рассмотрим два подмножества множества индексов целевых точек

$$N_1 N_2 : N_1 \cup N_2 = N\,, \quad N_1 \cap N_2 = 0, \quad N = \{1, 2, ..., n\}\,,$$

$$N_1 = \{i \in N \mid M_i \in C^{T-t_0}(x^0)\}\,, \quad N_2 = \{i \in N \mid M_i \;\overline{\in}\; C^{T-t_0}(x^0)\}\,.$$

Для этого случая множество Парето-оптимальных оценок имеет вид

$$P(x^0) = \{H(x) \mid x \in \pi_{C^{T-t_0}(x^0)} \hat{M}\}\,.$$

Покажем это. Если $N_1 = 0$, то мы находимся в условиях случаев 1 или 3 соответственно, когда

$$\hat{M} \cap C^{T-t_0}(x^0) = 0 \quad \text{или} \quad C^{T-t_0}(x^0) \subset \hat{M}\,,$$

если $N_2 = \varnothing$, то $\hat{M} \subset C^{T-t_0}(x^0)$, и мы находимся в условиях случая 2.

Будем считать, что $N_1, N_2 \neq 0$. Выведем обозначения:

$$\hat{M}_1 = \hat{M}_1 \cap C^{T-t_0}(x^0)\,, \quad \hat{M}_2 = \hat{M} \setminus \hat{M}\hat{M}_1\,.$$

Тогда

$$\pi_{C^{T-t_0}(x^0)} \hat{M} = \hat{M}_1 \cup \pi_{C^{T-t_0}(x^0)} \hat{M}_2\,.$$

Пусть

$$y' \in \pi_{C^{T-t_0}(x^0)} \hat{M}\,, \quad x \in C^{T-t_0}(x^0)\,, \quad x \neq y'\,. \tag{3.36}$$

Нужно показать, что

$$H_i(y') > H_i(x)\,, \tag{3.37}$$

хотя бы для одного $i \in N$.

Для этого рассмотрим два варианта:

1) $y' \in \hat{M}_1$. Так как $\hat{M}_1 \subset \hat{M}$, тогда доказательство неравенства (3.37) проводится таким же образом, как в случае 2.

2) $y' \in \pi_{C^{T-t_0}(x^0)}(\hat{M} \setminus \hat{M}_1) = \pi_{C^{T-t_0}(x^0)}\hat{M}_2$. Тогда для любых x из (3.36) и $\bar{y} \in Y$, где $Y = \{y \in \hat{M}_2 \mid \pi_{C^{T-t_0}(x^0)} y = y'\}$, выполняется неравенство $\rho(y', \bar{y}) < \rho(x, \bar{y})$. Поэтому неравенство (3.37) вытекает из утверждения 4. Таким образом, множество $\pi_{C^{T-t_0}(x^0)}\hat{M}$ состоит только из точек, удовлетворяющих условию (3.37).

Покажем, что, кроме точек множества $\pi_{C^{T-t_0}(x^0)}\hat{M}$, в множестве $C^{T-t_0}(x^0)$ нет точек, обладающих свойством (3.37), т.е. что для любых $x \in C^{T-t_0}(x^0) \setminus \pi_{C^{T-t_0}(x^0)}\hat{M}$ найдется такая точка $\tilde{y} \in \pi_{C^{T-t_0}(x^0)}\hat{M}$, для которой $\rho(\tilde{y}, M_i) < \rho(x, M_i)$ для всех $i \in N$.

Рассмотрим множество $\hat{M}_3 = \operatorname{conv} \pi_{C^{T-t_0}(x^0)}\hat{M}$. Для любых $x \in C^{T-t_0}(x^0) \setminus \hat{M}_3$ существует точка $\tilde{y} \in \pi_{C^{T-t_0}(x^0)}\hat{M}_3$, такая, что $\rho(\tilde{y}, M_i) < \rho(x, M_i)$ для всех $i \in N$.

Пусть теперь $x \in \hat{M}_3 \setminus \pi_{C^{T-t_0}(x^0)}\hat{M}$. Если $\pi_{\hat{M}} x \in \hat{M}_1$, то точка $\tilde{y}$ – искомая (см. утверждение 2), т.е. $\tilde{y} \in \pi_{\hat{M}} x$. Если же $\pi_{\hat{M}} x \in \hat{M}_2$, то $\pi_{\hat{M}} x \,\overline{\in}\, \pi_{C^{T-t_0}(x^0)}\hat{M}$, но существует точка $x' \in \pi_{C^{T-t_0}(x^0)}\hat{M}_2 \cap \omega_\lambda$, где $\omega_\lambda = \lambda x + (1-\lambda)\pi_{\hat{M}} x$, и в соответствии с утверждением 3

$$\rho(x', M_i) < \rho(x, M_i) \text{ для всех } i = 1, 2, ..., n,$$

т.е. x' есть искомая точка $\hat{y}$.

Проведенный анализ всех случаев взаимного расположения целевых точек и множеств достижимости дает конструктивный способ построения Парето-оптимального множества в задаче сближения с несколькими целевыми точками и позволяет определить геометрическую структуру множества Парето-оптимальных концов траекторий движения системы.

Контрольные вопросы

1. Перечислите возможные случаи расположения точек M и множества достижимости.

3.4. ОПТИМИЗАЦИЯ В СИСТЕМАХ С ИЕРАРХИЧЕСКОЙ СТРУКТУРОЙ

Понятие иерархической структуры невозможно определить одной сжатой формулировкой. Однако можно выделить ряд существенных характеристик, присущих всем иерархическим системам. Во-первых, совокупность подсистем, составляющих данную систему, имеет последовательное вертикальное расположение. Во-вторых, устанавливается приоритет действий, принятия решения. В-третьих, результаты действий подсистем верхнего уровня зависят от действий нижнего уровня.

Таким образом, любая иерархия состоит из вертикально соподчиненных подсистем.

На деятельность подсистемы любого уровня (кроме верхнего) непосредственное воздействие оказывают подсистемы, расположенные на более верхних уровнях. Хотя такое воздействие направлено сверху вниз, успех действия системы в целом и каждого уровня зависит от поведения всей элементов системы. Понятие приоритета действий предполагает, что вмешательство подсистем верхнего уровня предшествует действиям более низких уровней. Поэтому успешность работы подсистем вышестоящих уровней зависит не только от собственных действий, но и от реакций подсистем нижних уровней на вмешательство.

Подсистему самого верхнего уровня будем называть центром, а подсистемы более низких уровней – элементами.

В системах управления элементам предоставлено право вырабатывать определенные управляющие воздействия принимать решения. Поэтому наряду с иерархией системы говорят об иерархической структуре управления.

Иерархическая структура управления в сложной системе представляет собой совокупность уровней управления, следующих друг за другом в порядке определенного приоритета. Между элементами различных уровней иерархии существуют как вертикальные, так и горизонтальные связи.

Появление иерархической структуры в системах управления и принятия решений обусловлено наличием большого объема информации об управляемых процессах в системе, невозможностью обработки этой информации и принятия решения одним управляющим центром, а также существующей в реальных системах децентрализацией процесса принятия решений, когда элементы, подчиненные центру, вырабатывают управляющие воздействия исходя из указаний центра и с учетом собственных интересов.

Теория иерархических систем управления в настоящее время достаточно хорошо освещена в литературе. Одними из первых работ, посвященных систематическому исследованию иерархических систем управления, являются, например, работы [1 – 3].

Рассмотрим математическую модель простейшей двухуровневой иерархической системы управления. Пусть центру A_0 подчинены элементы системы управления $B_1, B_2, \ldots, B_n$, которые в дальнейшем будем называть подсистемами. Центр вырабатывает управляющее воздействие $u = \{u_1, u_2, \ldots, u_n\}$ и сообщает его подсистемам нижнего уровня $B_1, B_2, \ldots, B_n$, которые, получив информацию о решении центра, выбирают собственные управления $\{v_i\}$ из некоторых множеств допустимых управлений $V_1(u), V_2(u), \ldots, V_n(u)$, зависящих от выбора управления u игроком A_0.

Обозначим через U множество допустимых управлений центра. Управление u будем называть допустимым, если для любого $i = 1, 2, \ldots, n$ множества $V_i(u)$ не являются пустыми.

Если для любого $u \in U$ все множества $V_i(u)$ состоят из единственных управлений, то в этом случае центр обладает полной информацией о реакции подсистем нижнего уровня на свое управление.

Пусть $H_0(u, \mathrm{v})$ – критерий оптимальности центра, а $H_i(u_i, \mathrm{v}_i)$ – критерии оптимальности подсистем $B_1, B_2, \ldots, B_n$. Каждый из элементов стремится максимизировать свой функционал. Если множества $V_i(u)$ состоят из единственных управлений, т.е. $V_i(u) = \{\mathrm{v}_i(u)\}$, то центр выбирает свое управление u^* так, чтобы

$$H_0(u^*, \mathrm{v}(u^*)) = \max_{u \in U} H_0(u, \mathrm{v}(u)), \tag{3.38}$$

а значения функционалов H_i будут равны

$$H_1(u_1^*, \mathrm{v}_1^*),\ H_2(u_2^*, \mathrm{v}_2^*),\ \ldots,\ H_n(u_n^*, \mathrm{v}_n^*),\ \mathrm{v}_i^* = \mathrm{v}_i(u^*).$$

В рассмотренном нами случае выбор центром управления $u \in U$ позволяет однозначно определить и конечный результат, т.е. значения функционалов центра и подсистем. Это является следствием того, что множества $V_i(u)$ состоят из единственных элементов. Однако в общем случае выбор управления u не определяет единственные значения управлений подсистем, а лишь позволяет вышестоящему уровню оценить возможный выбор подсистемы на некотором множестве допустимых управлений.

Будем называть множество управлений i-й подсистемы $R_i(u)$ множеством оптимальных реакций этой подсистемы, если

$$R_i(u) = \{\mathrm{v}_i \in V_i(u) \mid H_i(u, \mathrm{v}_i) \ge H_i(u, \mathrm{v}_i') \forall \mathrm{v}_i' \in V_i(u)\}.$$

Если множества $R_i(u)$ не являются одноэлементными, то центр при принятом решении оказывается в условиях неопределенности. Для решения задачи в этих условиях требуется сделать дополнительные предположения о поведении подсистем нижнего уровня. В зависимости от характера предположений получим те или иные постановки задач.

Одной из возможных гипотез поведения является предположение о том, что подсистемы выбирают управление $\mathrm{v} = (\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_n)$ такое, что

$$H_0(u, \mathrm{v}) \le H_0(u, \mathrm{v}') \tag{3.39}$$

для любых $\mathrm{v}_i' \in R_i(u)$, т.е. подсистемы выбирают управления наихудшим для центра образом. В этом случае естественным является выбор центром управления u^0 такого, что

$$H_0(u^0, \mathrm{v}^0) = \min_{\mathrm{v} \in R(u^0)} H_0(u^0, \mathrm{v}) \ge \min_{\mathrm{v} \in R(u)} H_0(u, \mathrm{v}) \forall u \in U, \tag{3.40}$$

где $R(u) = \prod_{i=1}^{n} R_i(u)$.

Выбор решения центром в соответствии с процедурой (3.39) – (3.40) называется принципом гарантированного результата.

Изменив гипотезу поведения, получим другие постановки задачи оптимизация. Например, предположим, что подсистемы, проявляя доброжелательность по отношению к центру, выбирают управление $\hat{\mathrm{v}} \in R(u)$ таким образом, что

$$H_0(u, \hat{\mathrm{v}}) = \max_{\mathrm{v} \in R(u)} H_0(u, \mathrm{v}), \tag{3.41}$$

т.е. подсистемы строят свое решение в виде функции $\hat{\mathrm{v}} = \hat{\mathrm{v}}(u)$ в соответствии с условием (3.41). Тогда центр выберет управление $\hat{u}$, доставляющее максимум функция $H_0(u, \hat{\mathrm{v}}(u))$. Таким образом,

$$H_0(\hat{u}, \hat{\mathrm{v}}) = \max_{u \in U} H_0(u, \hat{\mathrm{v}}(u)) = \max_{u \in U} \max_{\mathrm{v} \in R(u)} H_0(u, \mathrm{v}). \tag{3.42}$$

В условиях предположения о доброжелательности центр, очевидно, добьется большего значения функционала, чем при реализации принципа гарантированного результата. Это следует из неравенства

$$\max_{u \in U} \max_{\mathrm{v} \in R(u)} H_0(u, \mathrm{v}) \ge \max_{u \in U} \min_{\mathrm{v} \in R(u)} H_0(u, \mathrm{v}).$$

Заметим, что от управления центра зависят как функционалы подсистем, так и множество их допустимых управлений и оптимальных реакций. Это позволяет центру осуществлять руководство подсистемами посредством воздействия на множества допустимых управлений и значения функционалов.

Рассмотрим два примера иерархических систем управления.

Пример 1 (распределение ресурсов) [12, 16]. Рассматривается следующая идеализированная экономическая ситуация. Административный центр распределяет ограниченный объем ресурсов между подчиненными ему подразделениям i, которые, в свою очередь, расходуют полученный ресурс для производства продукции с учетом собственного критерия.

Проведем формализацию этой задачи. Центр A_0 выбирает систему из n векторов:

$$(u_1, u_2, \ldots, u_n) : u_i \ge 0, \quad u_i \in E_i, \quad i = 1, 2, \ldots, n,$$

$$\sum_{i=1}^{n} u_i \le b, b \ge 0.$$

Вектор u будем интерпретировать как набор ресурсов из l наименований, выделяемых центром A_0 для i-го производственного подразделения. Каждое из подразделений B_i, зная выбор A_0, выбирает вектор $\mathrm{v}_i \in E_m$ из множества векторов, удовлетворяющих ограничениям

$$\mathrm{v}_i A_i \le u_i + \alpha_i, \quad \mathrm{v}_i \ge 0, \quad \alpha_i \ge 0, \quad A_i \ge 0\,. \tag{3.43}$$

Здесь вектор v_i интерпретируется как производственная программа производственного подразделения B_i, по m видам продукции; A_i – производственная или технологическая матрица i-го производственного подразделения; α_i – вектор наличных ресурсов подразделения B_i.

Определим критерии участников. Для центра положим

$$H_0(u,\ \mathrm{v}_1(u),\ ...,\ \mathrm{v}_n(u)) = \sum_{i=1}^{n} a_i \mathrm{v}_i(u), a_i \ge 0\,,\ a_i \in E_m\,,\ i = 1, 2, ..., n,$$

где $u = (u_1, u_2, \ldots, u_n)$ – управление центра A_0; $\mathrm{v}_i(u)$ – производственная программа подразделения B_i, удовлетворяющая условию (3.31); a_i – вектор полезности центра A_0 от продукции, выпускаемой i-м производственным подразделением; $a_i\mathrm{v}_i(u)$ – скалярное произведение векторов a_i и $\mathrm{v}_i(u)$. Для производственного подразделения B_i критерий будет иметь вид

$$H_i(u,\ \mathrm{v}_i(u)) = c_i \mathrm{v}_i(u)\,, \quad c_i \ge 0\,, \quad c_i \in E_m\,, \quad i = 1, 2, ..., n,$$

где c_i – вектор полезности предприятия i от своей продукции.

Целью каждого является максимизация своего критерия.

Рассмотрим следующую процедуру принятия решения. Пусть $\mathrm{v}_i^*(u)$ – решение задачи параметрического программирования (параметром является вектор u)

$$\max_{\mathrm{v}_i \in R_i(u)} c_i \mathrm{v}_i\,,$$

$$R_i(u) = \{\mathrm{v}_i \mid \mathrm{v}_i \ge 0,\ \mathrm{v}_i A_i \le u_i + \alpha_i,\ u_i \ge 0,\ \alpha_i \ge 0\}\,,$$

а $u^* = (u_1^*, u_2^*, ..., u_n^*)$ – решение задачи

$$\max_{u \in U} \sum_{i=1}^{n} a_i \mathrm{v}_i^*(u)\,,$$

$$U = \left\{ u \mid u_i \ge 0;\ \sum_{i=1}^{n} u_i \le b \right\}. \tag{3.44}$$

Покажем, что построенное решение удовлетворяет неравенствам

$$H_0(u^*,\ \mathrm{v}^*) \ge H_0(u,\ \mathrm{v}^*(u)),\ \ u \in U\,,$$

$$H_i(u_i^*,\ \mathrm{v}_i^*(u^*)) \ge H_i(u^*,\ \mathrm{v}_i),\ \ \mathrm{v}_i \in V_i(u)\,,\ \ i = 1, 2, ..., n. \tag{3.45}$$

Действительно,

$$H_0(u^*,\ \mathrm{v}^*(u^*)) = \sum_{i=1}^{n} a_i \mathrm{v}_i^*(u^*) \ge \sum_{i=1}^{n} a_i \mathrm{v}_i^*(u^*) = H_0(u,\ \mathrm{v}^*(u))$$

и для всех $i = 1, 2, ..., n$

$$H_i(u_i^*,\ \mathrm{v}_i^*(u^*) = c_i \mathrm{v}_i^*(u^*) \ge c_i \mathrm{v}_i(u^*) = H_0(u^*,\ \mathrm{v}_i(u^*))\,.$$

Это означает, что ни одному производственному подразделению B_i, а также центру A_0 невыгодно одностороннее отклонение от ситуации (u^*, $\mathrm{v}_1{}^*(u^*)$, …, $\mathrm{v}_n^*(u^*)$). Такая ситуация в теории игр называется ситуацией равновесия по Нэшу.

В этом примере центр оказывает воздействие только на множество допустимых управлений подчиненных подразделений, не влияя никак на их функционалы.

П р и м е р 2. Задача о нормировании выбросов. Предположим, что уровень загрязнения в промышленном районе характеризуется скалярной величиной

$$q = \sum_{i=1}^{n} a_i \mathrm{v}_i\,, \quad a_i > 0\,, \quad 0 \le \mathrm{v}_i \le b_i\,, \quad i = 1, 2, ..., n,$$

где v_i – объем выброса вредных веществ i-м предприятием.

Зависимость между объемами выбросов и затратами предприятий на переработку несброшенных отходов выражается функцией

$$h_i(\mathrm{v}_i) = c_i(b_i - \mathrm{v}_i), \quad c_i > 0, \quad i = 1, 2, ..., n.$$

Если уровень загрязнения в районе превышает величину Q, то на предприятия накладываются штрафы $s_i > 0$. Таким образом, функция затрат предприятия i имеет вид

$$H_i(b_i;\, s_i;\, \mathrm{v}_1, ..., \mathrm{v}_n) = \begin{cases} c_i(b_i - \mathrm{v}_i), & \sum\limits_{i=1}^{n} a_i \mathrm{v}_i \le Q; \\ c_i(b_i - \mathrm{v}_i) + s_i, & \sum\limits_{i=1}^{n} a_i \mathrm{v}_i > Q. \end{cases} \tag{3.46}$$

Предприятия заинтересованы в минимизации своих затрат. Центру поручено осуществлять контроль за уровнем загрязнения и предоставлено право ограничивать выбросы предприятий и налагать штрафы за загрязнение, т.е. устанавливать значения величин $b_1, b_2, \ldots, b_n; s_1, s_2, \ldots, s_n$. Критерий центра зададим в виде

$$H_0(\mathrm{v}_1, \mathrm{v}_2, ..., \mathrm{v}_n) = \begin{cases} 1, & \sum\limits_{i=1}^{n} a_i \mathrm{v}_i \le Q; \\ 0, & \sum\limits_{i=1}^{n} a_i \mathrm{v}_i > Q. \end{cases} \tag{3.47}$$

Целью центра является максимизация функции $H_0(\mathrm{v}_1, \ldots, \mathrm{v}_n)$ посредством выбора величин $b_1, \ldots, b_n; s_1, \ldots, s_n$. При этом их необходимо выбрать таким образом, чтобы предприятиям было невыгодно отклоняться от значений $\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_n$

Пусть $\mathrm{v} = (\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_n)$ таково, что $\sum\limits_{i=1}^{n} a_i \mathrm{v}_i = Q$. В этом случае

$$H_i(b_i;\, s_i;\, \mathrm{v}_1, ..., \mathrm{v}_n) = c_i(b_i - \mathrm{v}_i), \quad i = 1, 2, ..., n,$$

$$H_0(\mathrm{v}_1, ..., \mathrm{v}_n) = 1.$$

Найдем, при каких значениях b_i и s_i указанная точка является точкой минимума функций $H_i(b_i;\, s_i;\, \mathrm{v}_1, ..., \mathrm{v}_n)$ по аргументу v_i. Для этого фиксируем значение $\mathrm{v}_i, \ldots, \mathrm{v}_{i-1}, \mathrm{v}_{i+1}, \ldots, \mathrm{v}_n$ и положим $\mathrm{v}_i = b_i$, тогда

$$H_i(b_i;\, s_i;\, \mathrm{v}_i, ..., \mathrm{v}_{i-1}, b_i, \mathrm{v}_{i+1}, ..., \mathrm{v}_n) = s_i.$$

Заметим, что для $\mathrm{v}_1' \in (\mathrm{v}_i, b_i)$ выполнено неравенство

$$H_i(b_i;\, s_i;\, \mathrm{v}_1, ..., \mathrm{v}_i', ..., \mathrm{v}_n) > H_i(b_i;\, s_i;\, \mathrm{v}_1, ..., b_i, ..., \mathrm{v}_n),$$

а для $\mathrm{v}_i' \in [0, \mathrm{v}_i]$

$$H_i(b_i;\, s_i;\, \mathrm{v}_1, ..., \mathrm{v}_i, ..., \mathrm{v}_n) < H_i(b_i;\, s_i;\, \mathrm{v}_1, ..., \mathrm{v}_i', ..., \mathrm{v}_n).$$

Следовательно, для того чтобы $(\mathrm{v}_1, \ldots, \mathrm{v}_n)$ являлась точкой минимума по каждому из аргументов, достаточно, чтобы выполнялось неравенство $c_i(b_i - \mathrm{v}_i) < s_i$.

Таким образом, решением задачи являются все значения $b_1, \ldots, b_n; s_1, \ldots, s_n, \mathrm{v}_1, \ldots, \mathrm{v}_n$, удовлетворяющие условиям

$$\sum_{i=1}^{n} a_i \mathrm{v}_i = Q, \tag{3.48}$$

$$c_i(b_i - \mathrm{v}_i) < s_i, \quad b_i > 0, \quad s_i > 0, \quad i = 1, 2, ..., n.$$

В этой иерархической системе управления центр может оказывать воздействие как на область допустимых решений подсистем (предприятий), так и на их критерии (функции затрат предприятий). Заметим, что в данном примере мы предполагали, что предприятия действуют изолированно друг от друга, и не учитывали возможность их объединения для совместного принятия решения об установлении величин $\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_n$. Этот случай потребовал бы дополнительных исследований.

Как видно из приведенных примеров, иерархические системы управления предполагают наличие в них подсистем, каждая из которых стремится к достижению собственной цели, т.е. образуется ситуация, которая в теории игр характеризуется как конфликт. Поэтому для исследования таких систем целесообразно использовать аппарат теории игр. Характерной особенностью игр, которые служат моделями иерархических систем управления является присутствие в них хотя бы одного игрока, который принимает решение независимо от решений других игроков, ориентируясь лишь на знание их функций выигрыша. Естественно, что рассмотренные нами системы являются лишь простейшими представителями обширного класса иерархических систем, в которых каждая из подсистем может; в свою очередь, являться центром для других подсистем или быть связана с подсистемами того же уровня иерархии.

Контрольные вопросы

1. Приведите примеры постановки задачи управления для системы с иерархической структурой.
2. Дайте характеристику задачи.

3.5. ЭЛЕМЕНТЫ ТЕОРИИ ИГР В СИСТЕМНОМ АНАЛИЗЕ

В изучении организации систем можно выделить два важных направления: исследование структуры и исследование поведения подсистем (элементов или участников организации). При исследовании структуры, как правило, участники организации рассматривают как некоторый элемент системы, выполняющий порученную задачу наиболее эффективным образом. Исследование поведения участников предполагает изучение мотивов их действий при условии, что участнику предоставлено право принимать решения и самостоятельно оценивать их результат в соответствии с собственными критериями [20, 21].

Наличие сложной иерархической структуры, значительное количество участников организации неизбежно приводит к тому, что интересы участников вступают в противоречия с интересами друг друга и системы в целом. Таким образом, создается ситуация, которую можно охарактеризовать как конфликтную. Наличие конфликта и стремление к его оптимальному разрешению являются характерными чертами функционирования многих сложных систем. Математическая проблематика, связанная с таким направлением исследований, состоит в установлении связи между формальными характеристиками конфликта и формальными характеристиками оптимального поведения участников этого конфликта.

3.5.1. Основные элементы теории игр

Раздел математики, посвященный изучению математических моделей принятия оптимальных решений в условиях конфликтов, называется теорией игр. Участников конфликта в теории игр называют игроками.

Пусть $I = \{1, 2, ..., n\}$ – множество игроков. Каждый из игроков может совершать различные действия, которые обычно называют стратегиями. Множество X_i всевозможных действий игроков называется множеством стратегий. Вектор $x = (x_1, x_2, \ldots, x_n)$, где x_i принадлежит множеству X_i, стратегий игрока i, называется ситуацией в игре. В каждой ситуации x игрок i получает выигрыш, который обозначим через $H_i(x)$. Отображение $H_i : X_1 \times X_2 \times ... \times X_n \to E_1$ называется функцией выигрыша игрока i.

О п р е д е л е н и е . Бескоалиционной игрой называется система

$$\Gamma =< I,\ \{X_i\}_{i \in I},\ \{H_i\}_{i \in I} >,$$

где $I = \{1, 2, ..., n\}$ – множество игроков; X_i – множество стратегий игрока i; H_i – функция выигрыша игрока i, заданная на множестве $X = X_1 \times X_2 \times ... \times X_n$ стратегий игры.

Каждый из игроков в игре Γ старается максимизировать свой выигрыш. Поэтому естественно, что любую ситуацию игрок пытается изменить с помощью своей стратегии таким образом, чтобы его выигрыш был максимальным. Обозначим ситуацию x, в которой игрок i изменил свою стратегию x_i, на x_i' через

$$x \parallel x_i' = (x_1,\ ...,\ x_{i-1},\ x_i',\ x_{i+1},\ ...,\ x_n)\ .$$

Будем говорить, что ситуация приемлема для игрока i, если $H_i(x) \ge H_i(x \parallel x_i')$ для любой стратегии $x_i' \in X_i$. Если некоторая ситуация является приемлемой для одного игрока, но не является приемлемой для другого игрока, то второй игрок будет стремиться изменить ситуацию. Если же ситуация приемлема для всех игроков, то ни один из них не будет этого делать, т.е. в игре устанавливается равновесие.

О п р е д е л е н и е . Ситуация x называется ситуацией равновесия (по Нэшу), если для любого $i \in I$ и любой стратегии $x_i' \in X_i$ выполнено неравенство $H_i(x) \ge H_i(x \parallel x_i')$.

П р и м е р 1. Охрана окружающей среды.

Каждое из трех предприятий (игроки 1, 2, 3), использующее для технических целей воду некоторого водоема, располагает двумя стратегиями:

1) использовать очистные сооружения для очистки отработанной воды;
2) сбрасывать ее без очистки.

Предполагается, что в случае, когда неочищенную воду сбрасывает не более одного предприятия, вода в водоеме остается пригодной для использования и предприятия убытков не несут. Если же неочищенную воду сбрасывают два и более предприятий, то каждое несет убытки в размере трех единиц. Стоимость эксплуатации очистных сооружений обходится каждому предприятию в одну единицу.

Построим куб ситуаций для описанной игры и укажем выигрыши игроков в этих ситуациях (рис. 3.10).

Опишем множество приемлемых ситуаций для игроков:

1: (2, 1, 1), (2, 2, 2), (1, 1, 2), (1, 2, 1);
2: (1, 2, 1), (2, 2, 2), (1, 1, 2), (2, 1, 1);
3: (1, 1, 2), (2, 2, 2), (2, 1, 1), (1, 2, 1).

Отсюда получаем, что ситуациями равновесия в этой игре являются все приемлемые ситуации. При этом ситуация (2, 2, 2) наименее выгодна как с точки зрения охраны природы, так и с точки зрения величины выигрыша игроков, поскольку в этой ситуации загрязнение больше, а выигрыш игроков меньше, чем в других приемлемых ситуациях. Использование в качестве принципа оптимальности равновесия по Нэшу имеет несколько недостатков. Во-первых, ситуации равновесия существуют далеко не в каждой игре. Во-вторых, в различных ситуациях равновесия выигрыши игроков различны. И, в-третьих, ситуации равновесия могут оказаться неустойчивыми относительно отклонения группы игроков, которые одновременно изменяют свои стратегии с целью увеличения выигрышей. Любое подмножество S множества I будем называть коалицией. Обозначим через $x \parallel x_S'$ ситуацию, в которой все игроки, входящие в S, одновременно изменили свои стратегии x_i на x_i'.

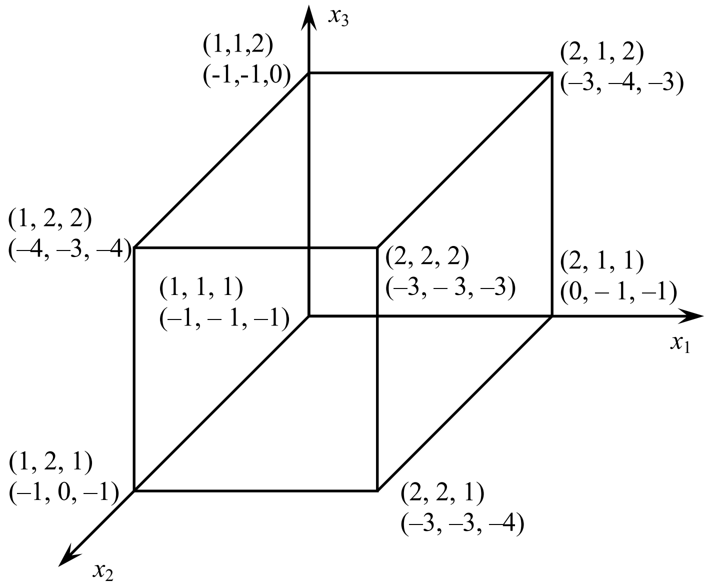

Рис. 3.10. К примеру 1

О п р е д е л е н и е . Ситуация x называется сильно равновесной, если не существует такой коалиции $S \subset I$ и такой стратегии $x_S \in \prod_{i \in S} X_i$ что $H_i(x \| x_S) > H_i(x)$ для всех $i \in S$.

В рассмотренном примере сильно равновесными будут три ситуации: (1, 1 ,2), (1, 2, 1), (2, 1, 1). Ситуация (2, 2, 2) сильно равновесной ситуацией не является.

Множество сильно равновесных ситуаций, очевидно, содержится во множестве ситуаций равновесия.

П р и м е р 2. Два предприятия получают электроэнергию от энергосистемы ограниченной мощности. Известно, что максимальный объем электроэнергии, который могут потребить предприятия за одни сутки, равен для первого предприятия трем единицам, а для второго – четырем единицам. Недостаток электроэнергии приводит к убыткам предприятий, которые выражаются для первого предприятия величиной $3 - x_1 + x_2$, а для второго $4 - x_2 + x_1$, где x_1 – объем электроэнергии, потребляемой первым предприятием, а x_2 – вторым. Если общий объем потребляемой энергии превышает величину пять единиц, т.е. $x_1 + x_2 > 5$ то в энергосистеме происходит авария, которая обходится каждому предприятию в одну единицу. Необходимо решить вопрос об ограничении объемов потребляемой каждым предприятием энергии.

Считая выигрыш игроков (предприятий) равным убыткам со знаком минус, запишем функции выигрыша в следующем виде:

$$H_1(x_1, x_2) = \begin{cases} x_1 - 3 - x_2, x_1 + x_2 \le 5; \\ x_1 - 3 - x_2 - 1, x_1 + x_2 > 5, \end{cases}$$

$$H_2(x_1, x_2) = \begin{cases} x_2 - 4 - x_1, x_1 + x_2 \le 5; \\ x_2 - 4 - x_1 - 1, x_1 + x_2 > 5, \end{cases}$$

где $0 \le x_1 \le 3$, $0 \le x_2 \le 4$.

Множество приемлемых ситуаций (рис. 3.11):

- для игрока 1:

$$P_1 = \{x \mid x_1 = 3, x_2 \in [0, 2] \cup [3, 4]\} \cup \{x \mid x_1 + x_2 = 5, 2 \le x_2 \le 3\};$$

- для игрока 2:

$$P_2 = \{x \mid x_2 = 4, x_1 \in [0, 1] \cup [2, 3]\} \cup \{x \mid x_1 + x_2 = 5, 1 \le x_1 \le 2\}.$$

Пересечение этих множеств дает нам две точки $x^1 = (2, 3)$ и $x^2 = (3, 4)$, которые являются ситуациями равновесия и одновременно сильно равновесными ситуациями. Таким образом, в этом примере множество сильно равновесных ситуаций совпадает с множеством ситуаций равновесия. При этом ситуация x^1 является более предпочтительной с точки зрения предотвращения загрязнения водоема.

Пусть в игре T участвуют два игрока. Если при этом $H_1(x) = -H_2(x)$, то такая игра называется антагонистической.

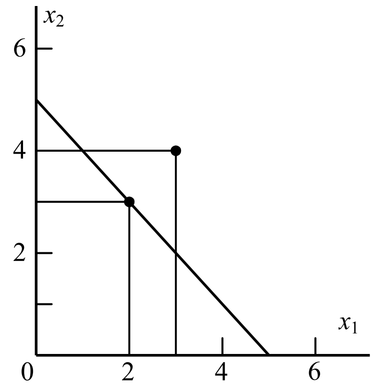

Рис. 3.11. К примеру 2

О п р е д е л е н и е . Пусть заданы два множества произвольной природы P, E. Множество P будем называть множеством стратегий первого игрока, а множество E – множеством стратегий второго игрока. Пусть на декартовом произведении $P \times E$ задана вещественная функция K. Тройку $T = <P, E, K>$ будем называть антагонистической игрой в нормальной форме.

В такой игре игроки одновременно выбирают стратегии $a \in P, b \in E$. После этого второй игрок получает выигрыш, равный
$K(a, b)$, а первый игрок – выигрыш, равный $K(a, b)$. Величина

$$\overline{V} = \inf_{a\in P} \sup_{b\in E} K(a,b)$$

называется верхним значением, а величина

$$\underline{V} = \sup_{b\in E} \inf_{a\in P} K(a,b) \ –$$

нижним значением игры T.

Из определения величин $\overline{V}$ и $\underline{V}$ вытекает следующее неравенство: $\overline{V} \ \ge \ \underline{V}$. Действительно, для любых стратегий $a \in P,\ b \in E$

$$\sup_{b\in E} K(a,b) \ge K(a,b) .$$

Отсюда заключаем, что

$$\overline{V} = \inf_{a\in P} \sup_{b\in E} \ge \inf_{a\in P} K(a,b) .$$

Поскольку это неравенство справедливо при всех $b \in E$, то

$$\overline{V} \ge \sup_{b\in E} \inf_{a\in P} K(a,b) = \underline{V} .$$

Говорят, что игра имеет значение, если выполняется равенство $\overline{V} = \underline{V}$. Значение игры обычно обозначают символом $V = \mathrm{val} T = \overline{V} = \underline{V}$.

П р и м е р 3. Предположим, что сельскохозяйственное предприятие может посеять одну из трех культур A_1, A_2, A_3. Урожайность каждой культуры зависит от погодных условий. Необходимо выбрать для посева культуру, которая даст максимальный урожай. Таким образом, с одной стороны, сельскохозяйственное предприятие (назовем его игроком II) заинтересовано в том, чтобы выбрать для посева культуру, дающую максимальный урожай, с другой – природа (назовем ее игроком I) может максимально повредить сельскохозяйственному предприятию, если условия погоды будут неблагоприятны для той или иной культуры, т.е. природа как бы преследует противоположные интересы.

Будем считать, что погода может быть засушливой, нормальной и дождливой, т.е. игрок I (природа) имеет только три стратегии. У предприятия также три стратегии: посеять культуру A_1, A_2 или A_3. Зададим урожайность культур в зависимости от погодных условий матрицей $A = \{a_{ij}\}$ где a_{ij} – урожайность культуры A_i, при погодных условиях типа i; $i, j = 1, 2, 3$. Пусть A имеет вид

$$A = \begin{pmatrix} 3 & 6 & 10 \\ 5 & 9 & 6 \\ 6 & 8 & 4 \end{pmatrix}.$$

Функция выигрыша $K(i, j)$ имеет вид $K(i, j) = a_{ij}$. Вычислим верхнее и нижнее значения игры:

$$\overline{V} = \min_i \max_j K(i,j) = \min_i \max_j a_{ij} = a_{32} = 8 ,$$

$$\underline{V} = \max_j \min_i K(i, j) = \max_j \min_i a_{ij} = 6 .$$

Стратегия $j = -2$ второго игрока обеспечивает ему при любом выборе стратегии игроком I выигрыш не меньше V и называется максиминной стратегией. Стратегия $1 = 3$ первого игрока обеспечивает, ему при любых действиях второго игрока проигрыш не больше V и называется минимаксной стратегией.

Таким образом, если сельскохозяйственное предприятие выберет для посева культуру A_2, то при самых неблагоприятных условиях урожайность будет не меньше шести единиц.

Вернемся теперь к рассмотрению игр n лиц.

Пусть игроки из множества I находятся в таких условиях, что совокупный выигрыш, который в состоянии получить любая из коалиций $S \subset I$, может быть произвольным образом распределен между членами коалиции S. В этом случае говорят, что выигрыши трансферабельны. Обозначим через $\text{v}(S)$ максимальный выигрыш, который может гарантировать себе коалиция S. Функция v, ставящая в соответствие каждой коалиции $S \subset I$ максимальный гарантированный ею выигрыш $\text{v}(S)$, называется характеристической функцией, которая является основополагающим понятием теории кооперативных игр.

Происхождение характеристической функции может иметь различную природу. Рассмотрим следующий пример.

П р и м е р 4. Три предприятия могут осуществлять реконструкцию производства как совместно, так и каждое в отдельности. При этом реконструкция, осуществляемая самостоятельно, обойдется предприятию с номером i в сумму $w(\{i\})$, а для любого объединения предприятий S – в сумму $w(S)$. Если считать, что объединение предприятий снижает затраты на реконструкцию, то предприятия получат от объединения выигрыш

$$\text{v}(s) = \sum_{i \in S} w(\{i\}) - w(S), \;\; S \subset \{1, 2, 3\} ,$$

который может быть распределен между предприятиями из S в виде «премии». Поскольку $u(S)$ представляет собой максимальный гарантированный выигрыш, который получает коалиция S, то ее можно назвать характеристической функцией.

При построении характеристической функции наиболее часто пользуются следующим ее определением.

О п р е д е л е н и е . Характеристической функцией игры n лиц называется вещественная функция u, определенная на подмножествах множества I и ставящая в соответствие любому $S \subset I$ значение (для S) антагонистической игры двух лиц, которую разыграли бы S и $I \setminus S$, если бы эти две коалиции действительно возникли; при этом под функцией выигрыша коалиции S понимается сумма выигрышей участников коалиции

$$K^S = \sum_{i \in S} H_i(x) ,$$

а функция выигрыша коалиции $I \setminus S$ полагается равной $- K^S$.

Пустую коалицию, как и всякое пустое множество, будем обозначать символом $\varnothing$. Из приведенного определения следует $\text{v}(\varnothing) = 0$. Если S и R – непересекающиеся коалиции, то, очевидно, объединив свои усилия, они могут получить выигрыш не меньший, чем если бы они действовали отдельно. Следовательно, $\text{v}(S) + \text{v}(R) <= (S \cup R)$, если $S \cap R \neq \varnothing$. Такое свойство характеристической функции называется супераддитивностью. Докажем, что оно имеет место.

Обозначим через x_s вектор стратегий игроков, входящих в коалицию S, который определим в качестве стратегии коалиции S. Представим $\text{v}(S)$ в виде

$$\text{v}(S) = \sup \inf \sum_{i \in S} H_i(x), \;\; x_S \in X_S, x_{I \setminus S} \in X_{I \setminus S},$$

где $X_s = \prod_{i \in S} X_i$ – множество стратегий коалиции S; $X_{I \setminus S} = \prod_{i \in I \setminus S} X_i$ – множество стратегий коалиции $I \setminus S$.

Имеем

$$\text{v}\left(S \bigcup R\right) = \sup_{x_{S \,\cup R} \in X_{S \,\cup\, R}} \; \inf_{x_{I \setminus (S \cup R)} \in X_{I \setminus (S \cup R)}} H_i(x_{S \cup R}, x_{I \setminus (S \cup R)}) .$$

Если супремум в правой части взять по множеству X_s, а затем по X_R, то он разве лишь уменьшится, т.е.

$$\text{v}(S \bigcup R) \geq \sup_{x_S \in X_S} \; \sup_{x_R \in X_R} \; \inf_{x_{I \setminus (S \cup R)} \in X_{I \setminus (S \cup R)}} \sum_{i \in S \cup R} H_i(x_S, x_R, x_{I \setminus (S \cup R)}) ,$$

и тем более

$$\begin{aligned} \text{v}\left(S \bigcup R\right) \geq & \inf_{x_{I \setminus (S \cup R)} \in X_{I \setminus (S \cup R)}} \sum_{i \in S \cup R} H_i(x_S, x_R, x_{I \setminus (S \cup R)}) \geq \\ \geq & \inf_{x_{I \setminus (S \cup R)} \in X_{I \setminus (S \cup R)}} \sum_{i \in S} H_i(x_S, x_R, x_{I \setminus (S \cup R)}) + \\ + & \inf_{x_{I \setminus (S \cup R)} \in X_{I \setminus (S \cup R)}} \sum_{i \in R} H_i(x_S, x_R, x_{I \setminus (S \cup R)}) . \end{aligned}$$

Возьмем инфимум первого слагаемого по $x_R \in X_R$, а инфимум второго по $x_S \in X_S$. Тогда

$$\mathrm{v}(S \cup R) \geq \inf_{x_R \in X_R} \inf_{x_{I\setminus(S\cup R)} \in X_{I\setminus(S\cup R)}} \sum_{i \in S} H_i(x_S, x_R, x_{I\setminus(S\cup R)}) +$$
$$+ \inf_{x_S \in X_S} \inf_{x_{I\setminus(S\cup R)} \in X_{I\setminus(S\cup R)}} \sum_{i \in R} H_i(x_S, x_R, x_{I\setminus(S\cup R)}).$$

Переход в первом слагаемом от инфимума по множествам X_R и $X_{I\setminus(R\cup S)}$ инфимуму по множеству $X_{I\setminus S}$ всевозможных пар (X_R, $X_{I\setminus(R\cup S)}$) и во втором слагаемом от инфимума по множествам X_S и $X_{I\setminus(R\cup S)}$ к инфимуму по множеству $X_{R\setminus S}$ всевозможных пар (X_S, $X_{I\setminus(S\cup R)}$) может лишь уменьшить правую часть. Поэтому

$$\mathrm{v}(S \cup R) \geq \inf_{x_{I\setminus S} \in X_{I\setminus S}} \sum_{i \in S} H_i(x_S, x_{I\setminus S}) + \inf_{x_{I\setminus R} \in X_{I\setminus R}} \sum_{i \in R} H_i(x_R, x_{I\setminus R})$$

для всех $x_S \in X_S$ и $x_R \in X_R$. Отсюда

$$\mathrm{v}(S \cup R) \geq \sup_{x_S \in X_S} \inf_{x_{I\setminus S} \in X_{I\setminus S}} \sum_{i \in S} H_i(x_S, x_{I\setminus S}) + \sup_{x_R \in X_R} \inf_{x_{I\setminus R} \in X_{I\setminus R}} \sum_{i \in R} H_i(x_R, x_{I\setminus R}),$$

что и требовалось доказать.

Из свойства супераддитивности характеристической функции следует, что игроки, объединяясь в коалицию I, получают наибольший выигрыш $\mathrm{v}(I)$. Поскольку коалиции I никто не противостоит, то ее гарантированный выигрыш $\mathrm{v}(I)$ совпадает с максимальным

$$\mathrm{v}(I) = \sup_{x \in X} \sum_{i \in I} H_i(x).$$

В такой ситуации проблема конфликтного выбора стратегии отсутствует. Основной задачей игроков становится достижение справедливого дележа общего выигрыша $\mathrm{v}(I)$. Эти вопросы являются предметом исследования теории кооперативных игр. Для того чтобы определить кооперативную игру, необходимо задать множество игроков I и характеристическую функцию.

Игра $T_{\mathrm{v}} = \langle I, \mathrm{v} \rangle$, называется кооперативной игрой в форме характеристической функции.

Определим понятие дележа в кооперативной игре.

Определение. Дележом для игры n лиц с характеристической функцией v называется вектор $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n)$, удовлетворяющий условиям:

1) $\sum_{i=1}^{n} \alpha_i = \mathrm{v}(I)$;
2) $\alpha_i >= \mathrm{v}(\{i\})$ для всех $i \in I$.

Первое условие означает, что весь полученный выигрыш распределяется между игроками. Это условие называется условием коллективной рациональности. Второе условие означает, что игрок в результате объединения должен получить выигрыш не меньший, чем он может себе уверенно обеспечить самостоятельно, и называется условием индивидуальной рациональности.

Пусть α' и α'' – дележи и S – некоторая коалиция.

Определение. Дележ α' доминирует дележ α'' по коалиции S ($\alpha' > \alpha''$), если:

а) $\alpha'_i > \alpha''_i$ для всех $i \in S$;

б) $\sum_{i \in S} \alpha'_i \leq \mathrm{v}(S)$.

Дележ α' доминирует α'' ($\alpha' \subset \alpha''$), если существует такая коалиция $S \subset I$, что $\alpha' \subset \alpha''$.

Условие а) в определении доминирования означает, что все члены коалиции S предпочитают α'; условие б) – что они в состоянии реализовать дележ α'. Отметим, что доминирование по коалиции, состоящей из единственного игрока, а также по множеству всех игроков I невозможно. Как легко видеть, в этих случаях нарушаются первое и второе свойства дележа.

Дележ, который не доминируется никаким другим дележом, можно считать в известном смысле «вполне устойчивым».

Определение. Множество всех недоминируемых дележей в кооперативной игре с характеристической функцией v называется
С-ядром.

Любой дележ из С-ядра устойчив в том смысле, что ни одна из коалиций не имеет ни желания, ни возможности изменить исход игры. Но С-ядро часто оказывается пустым.

Приведем важное необходимое и достаточное условие принадлежности дележа С-ядру.

Утверждение. Для того чтобы дележ α принадлежал С-ядру кооперативной игры, необходимо и достаточно, чтобы выполнялось неравенство

$$\sum_{I \in S} \alpha_i \geq \mathrm{v}(S)$$

для любой коалиции $S \subset I$.

Доказательство. Необходимость. Пусть дележ α принадлежит С-ядру. Предположим, что найдется некоторая коалиция $S \subset I$, для которой

$$\sum_{i\in S}\alpha_i < \mathrm{v}(S).$$

Функция v супераддитивна, поэтому

$$\mathrm{v}(I) \ge \mathrm{v}(I \setminus S) + \mathrm{v}(S) \ge \sum_{i\in I\setminus S}\mathrm{v}(\{i\}) + \mathrm{v}(S).$$

Обозначим

$$\xi = \mathrm{v}(I) - \mathrm{v}(S) - \sum_{i\in I\setminus S}\mathrm{v}(\{i\}) \ge 0$$

и составим вектор $\overline{\alpha}$:

$$\overline{\alpha}_i = \begin{cases} \alpha_i + \dfrac{\mathrm{v}(S) - \sum_{i\in S}\alpha_i}{|S|}, & i \in S, \\ \mathrm{v}(\{i\}) + \dfrac{\xi}{|I \setminus S|}, & i \in I \setminus S. \end{cases}$$

Из положительности ξ имеем $\overline{\alpha}_i \ge \mathrm{v}\,(\{i\})$. Непосредственной проверкой получаем, что $\sum_{S,i\in I}\overline{\alpha}_i \ge \mathrm{v}(I)$, поэтому вектор $\overline{\alpha}$ является дележом, более того, $\overline{\alpha} > \alpha$. Последнее противоречит принадлежности дележа к С-ядру. Следовательно, $\sum_{i\in S}\alpha_i \ge \mathrm{v}(S)$.

Достаточность. Предположим, что условие утверждения выполняется, а α не принадлежит С-ядру. Тогда существуют некоторая коалиция S и дележ $\overline{\alpha}$, такой, что

$$\sum_{i\in S}\overline{\alpha}_i \le \mathrm{v}(S),\quad \alpha_i < \overline{\alpha}_i$$

для всех $i \in S$. Отсюда получаем $\sum_{i\in S}\alpha_i < \sum_{i\in S}\overline{\alpha}_i < \mathrm{v}(S)$. Это противоречие доказывает утверждение.

Сформулированные теоретико-игровые понятия и свойства применяются для анализа систем, в которых процесс принятия решений носит однократный характер, а компоненты системы описываются статическими величинами. Поэтому рассмотренные игры называются статическими. На практике реальные системы, как правило, описываются параметрами, которые изменяются во времени, т.е. для более адекватного описания систем требуется задать динамику изменения ее параметров. Математическими моделями систем в этом случае служат системы дифференциальных уравнений или дискретные процессы, а допустимыми решениями являются некоторые функции времени. Игры, предметом исследования которых являются конфликтные задачи об управлении объектами, динамика которых описывается дифференциальными уравнениями, называются дифференциальными.

Пусть динамика компонент системы описывается векторным дифференциальным уравнением

$$\dot{x} = f(x(t), u_1(t), ..., u_n(t)),\quad t_0 \le t \le T, \tag{3.49}$$

при начальном условии

$$x(t_0) = x^0,$$

где $x \in E_m$, $f = (f_1, f_2, \ldots, f_m)$, $u_i \in U_i$, $i \in I$, U_i – множество допустимых управлений.

Если управление $u(t) = (u_1(t), u_2(t), ..., U_n(t))$ зависит только от значения параметра t, то такое управление называется программным. Управление $u = u(t, x)$, которое в каждый момент времени зависит от состояния системы, называется синтезированным или управлением в синтезированной форме.

Стратегией игрока называется отображение, которое в каждый момент времени t ставит в соответствие этому моменту времени или состоянию $x(t)$ системы некоторое значение управления u. Стратегии в зависимости от отображения также называются программными или позиционными.

Обозначим через D множество стратегий игрока i. Будем предполагать, что множество допустимых управлений $U = (U_1 \times U_2 \times \ldots \times U_n)$ таково, что применение любой стратегии из множества $D = D_1 \times D_2 \times \ \ldots \times D_n$ для каждого начального состояния порождает единственное измеримое управление $u(i) \in U$ и единственную траекторию $x(t)$. Набор стратегий $\varphi = (\varphi_1, \varphi_2, \ldots, \varphi_n)$, где ($\varphi_i \in D_i$) называется ситуацией в игре. Каждой ситуации в игре соответствует значение функции выигрыша, которую в общем случае можно определить следующим образом:

$$K^i_{t_0,x^0}(\varphi_1, \varphi_2, ..., \varphi_n) = \int_{t_0}^{T} h_i(x(t))dt + H_i(x(T)), \tag{3.50}$$

где $x(t)$ – траектория системы, реализующаяся в ситуации $\varphi = (\varphi_1, \varphi_2, \ldots, \varphi_n)$.

Дифференциальная игра, для которой задан момент ее окончания T, называется игрой с предписанной продолжительностью $T - t_0$.

Описанную дифференциальную игру n лиц с предписанной продолжительностью $T - t_0$ динамикой (3.49) и функционалами выигрышей (3.50) называют дифференциальной игрой в нормальной форме и обозначают

$$\Gamma(t_0, x^0) = \left\langle D_i, K^i_{t_0, x^0},\ i \in I \right\rangle .$$

В этом параграфе изложены лишь некоторые основные понятия теории игр и возможности ее использования в системном анализе. Для более глубокого ознакомления с этой теорией рекомендуем обратиться к книгам [7, 8, 11, 16, 17, 20, 21].

3.5.2. Принципы оптимальности в иерархических теоретико-игровых моделях

Мы уже упоминали о таких принципах оптимальности, используемых при принятии решений, как гарантированный результат и равновесие по Нэшу [20]. Здесь более подробно остановимся на понятиях равновесия по Нэшу и по Штакельбергу для дифференциальных иерархических игр.

Рассмотрим двухуровневую иерархическую игру, в которой игрок A_0, находящийся на первом уровне иерархии, называется центром. На втором уровне иерархии находятся игроки B_1, B_2, …, B_n. Динамика игры описывается системой дифференциальных уравнений:

$$\dot{x} = f(x, u), \quad \dot{y}_i = g_i(y_i, \mathrm{v}_i, u), \quad i = 1, 2, ..., n,$$
$$t \in [0, T], \quad x(0) = x^0, \quad y_i(0) = y_i^0, \quad u \in U, \quad \mathrm{v}_i \in V_i ,$$

где u – управление центра; $\{\mathrm{v}_i\}$ – управления игроков, выбираемые из областей допустимых управлений U и V соответственно.

Для каждого игрока определены функционалы выигрышей:

$$H_i(u, \mathrm{v}) = \int_0^T h_0(x, y, u, \mathrm{v})dt;$$
$$H_i(u, \mathrm{v}_i) = \int_0^T h_i(x, y_i, u, \mathrm{v}_i)dt, \qquad (3.51)$$

где h_i, h_0 – интегрируемые функции своих аргументов.

Целью каждого игрока является максимизация своего выигрыша.

Центр, используя программное управление, сообщает $u(t)$ игрокам нижнего уровня, а те максимизируют свои функционалы, т.е. их реакция на управление центра следующая:

$$R(u) = \{\operatorname{Arg}\max_{\mathrm{v}_i \in V_i(u)} H_i(u, \mathrm{v}) = \{\bar{\mathrm{v}}_i(t)\} = \bar{\mathrm{v}}(u(t)), \qquad (3.52)$$

$\bar{\mathrm{v}}_i(t)$ есть управление, максимизирующее функцию $H_i(u, \bar{\mathrm{v}}_i)$ при заданном управлении центра.

Центр, зная реакцию игроков нижнего уровня, выбирает управление $\bar{u}(t)$, которое максимизирует функционал $H_0(u, \bar{\mathrm{v}}(u))$.

Ситуация $(\bar{u}, \bar{\mathrm{v}})$ является ситуацией равновесия по Нэшу в рассматриваемой иерархической игре. Действительно,

$$H_0(\bar{u}, \bar{\mathrm{v}}) = \int_0^T h_0(x, y, \bar{u}, \bar{\mathrm{v}})dt \geq \int_0^T h_0(x, y, u, \bar{\mathrm{v}}(u))dt = H_0(u, \bar{\mathrm{v}}(u)).$$

Для всех $i = 1, 2, ..., n$

$$H_i(\bar{u}, \bar{\mathrm{v}}_i) = \int_0^T h_i(x, y_i, \bar{u}, \bar{\mathrm{v}}_i)dt \geq \int_0^T h_i(x, y_i, \bar{u}_i, \bar{\mathrm{v}}_i)dt = H_i(\bar{u}, \mathrm{v}_i) .$$

Таким образом, ни одному из игроков невыгодно отклоняться от ситуации $(\bar{u}, \bar{\mathrm{v}})$, т.е. она является равновесной но Нашу,

Дадим определение равновесия по Штакельбергу. Предположим, что центр выбрал программное управление $u(t)$, тогда игроки нижнего уровня определяют множество оптимальных реакций следующим образом:

$$R(u) = \{\mathrm{v} \mid \mathrm{v} \in V_1 \times V_2 \times ... \times V_n, H_i(u, \mathrm{v}_i) \geq H_i(u, \mathrm{v}'_i) \forall \mathrm{v}'_i \in V_i\} .$$

Игрок A_0, зная множество оптимальных реакций игроков нижнего уровня, выбирает такую стратегию u^*, чтобы

$$\min_{\mathrm{v} \in R(u^*)} H_0(u^*, \mathrm{v}_1, \mathrm{v}_2, ..., \mathrm{v}_n) \geq \min_{\upsilon \in R(u)} H_0(u, \mathrm{v}_1, \mathrm{v}_2, ..., \mathrm{v}_n) \forall u \in U .$$

Управление u^* называется оптимальным иерархическим решением центра. Пара (u^*, v^*) где $\mathrm{v}^* \in R(u^*)$, называется ситуацией равновесия по Штакельбергу в иерархической дифференциальной игре, а множество таких всевозможных пар – решением по Штакельбергу.

Предположим, что центр выбрал программное управление $u(t)$. Множество оптимальных реакций $R(u)$, согласно определению решения по Штакельбергу, есть множество ситуаций равновесия в бескоалиционной дифференциальной игре игроков нижнего уровня при фиксированном управлении центра. При выбранных функционалах выигрышей (3.52) множество оптимальных реакций есть

$$R(u) = \{\operatorname{Arg}\max_{\mathrm{v}_i \in U_i} H_i(u, \mathrm{v}_i)\} = \overline{\mathrm{v}}_i(u(t)).$$

Тогда $u^* = \operatorname{Arg}\max_{u \in U} H_0(u, \overline{\mathrm{v}}(u))$. Обозначим $\overline{\mathrm{v}}(u^*(t)) = \mathrm{v}^*(t)$.

Пара (u^*, v^*) является ситуацией равновесия по Штакельбергу и совпадает в данном случае с ситуацией равновесия по Нэшу $(\overline{u}, \overline{\mathrm{v}})$. Использование таких принципов оптимальности, как равновесия по Нэшу и Штакельбергу, для двухуровневых иерархических игр является вполне естественным, и они могут быть применены в процессе исследования большого числа систем, имеющих подобную структуру управления.

Однако это возможно и в том случае, когда система имеет и более сложную многоуровневую структуру управления.

Рассмотрим многоуровневую иерархическую систему, в которой иерархия управления определена следующим образом (рис. 3.12). Центр (игрок A_0) находится на первом уровне иерархии.

На *k-м* уровне иерархии располагаются игроки, входящие в непересекающиеся множества $S_k \subset I$ ($k = (1, 2, ..., l)$; $I = \{1, 2, ..., n\}$). Все игроки первого уровня подчинены игроку A_0, игрок i k-го уровня иерархии ($k = 1, 2, ..., l - 1$) имеет в своем подчинении множество игроков $F_1 \in S_{k+1}$. Множества F_i и F_j не пересекаются для любых i и j, неравных между собой. Динамика игры описывается системой дифференциальных уравнений

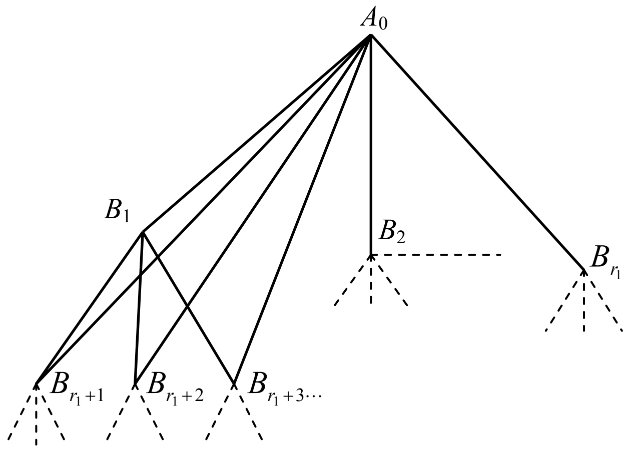

Рис. 3.12. Пример многоуровневой иерархической системы

$$\dot{x} = f(x,u), \quad \dot{y}_i = g_i(y_i, u, \mathrm{v}_{L_i}, \mathrm{v}_{F_i}, \mathrm{v}_i),$$
$$u \in U, \quad \mathrm{v}_i \in V_i, \quad i = 1, 2, ..., n,$$

где v_{L_i} есть управление игрока L_i, которому подчинен игрок i; v_{F_i} – вектор управлений игроков, подчиненных игроку i.

Функционалы выигрышей зададим следующим образом:

$$H_i(u, \mathrm{v}) = \int_0^T h_0(x, y, u, \mathrm{v})dt,$$

$$H_i(u, \mathrm{v}_{L_i}, \mathrm{v}_{F_i}, \mathrm{v}_i) = \int_0^T h_0(x, y_{i,} \mathrm{v}_{L_i}, \mathrm{v}_{F_i}, \mathrm{v}_i)dt, \quad i = 1, 2, ..., n.$$

Центр, используя программное управление, сообщает игрокам нижних уровней управление $u(t)$. Игроки всех других уровней, за исключением последнего, также сообщают игрокам, нижних уровней свои программные управления. Игроки нижнего уровня определяют свою реакцию на управление центра и игрока более верхнего уровня, которому они подчинены:

$$R_i(u, \mathrm{v}_{L_i}) = \operatorname{Arg}\max_{\mathrm{v}_i \in V_i} H_i(u, \mathrm{v}_{L_i}, \mathrm{v}_i) = \overline{\mathrm{v}}_i(u, \mathrm{v}_{L_i}), \quad i \in S_l.$$

Затем игроки более высоких уровней последовательно определяют свою реакцию:

$$R_i(u, \mathrm{v}_{L_i}) = \operatorname{Arg}\max_{\mathrm{v}_i \in V_i} H_i(u, \mathrm{v}_{L_i}, \overline{\mathrm{v}}_{F_i}(\cdot), \mathrm{v}_i) = \overline{\mathrm{v}}_i(u, \mathrm{v}_{L_i}).$$

Центр, зная реакцию игроков нижних уровней, выбирает управление $\overline{u}(t)$, максимизирующее функционал $H_0(u, \overline{\mathrm{v}})$. Построенная таким образом ситуация $(\overline{u}, \overline{\mathrm{v}})$ является ситуацией равновесия по Нэшу в многоуровневой иерархической игре.

Равновесие по Штакельбергу в многоуровневой игре определяется следующим образом. Множество оптимальных реакций игроков из S_i имеет вид

$$R_i(u, \mathrm{v}_{L_i}) = \{\mathrm{v}_i \in V_i \mid H_i(u, \mathrm{v}_{L_i}, \mathrm{v}_i) \ge H_i(u, \mathrm{v}_{L_i}, \mathrm{v}'_i) \forall \mathrm{v}'_i \in V_i\}.$$

Для игроков более высоких уровней множество оптимальных реакций задается следующим образом:

$$R_i(u, \mathrm{v}_{L_i}) = \{\mathrm{v}_i \in V_i \mid \min_{\substack{\mathrm{v}_{F_i} \in \prod R_j(u, \mathrm{v}_i) \\ j \in F_i}} H_i(u, \mathrm{v}_{L_i}, \mathrm{v}_{F_i}, \mathrm{v}_i) \ge$$
$$\ge \min_{\substack{\mathrm{v}_{F_i} \in \prod R_j(u, \mathrm{v}'_i) \\ j \in F_i}} H_i(u, \mathrm{v}_{L_i}, \mathrm{v}_{F_i}, \mathrm{v}'_i) \forall \mathrm{v}'_i \in V_i\}.$$

Оптимальным решением центра в многоуровневой иерархической дифференциальной игре называется $u^* \in U$, такое, что

$$\min_{\substack{\mathrm{v}_i \in R_i(u^*) \\ i \in S_1}} \min_{\substack{\mathrm{v}_i \in R_i(u^*, \mathrm{v}_{L_i}) \\ i \in S_2}} \ldots \min_{\substack{\mathrm{v}_i \in R_i(u^*, \mathrm{v}_{L_i}) \\ i \in S_l}} H_0(u^*, \mathrm{v}_1, \ldots, \mathrm{v}_n) \ge$$
$$\ge \min_{\substack{\mathrm{v}_i \in R_i(u) \\ i \in S_1}} \min_{\substack{\mathrm{v}_i \in R_i(u, \mathrm{v}_{L_i}) \\ i \in S_2}} \ldots \min_{\substack{\mathrm{v}_i \in R_i(u, \mathrm{v}_{L_i}) \\ i \in S_l}} H_0(u, \mathrm{v}_1, \ldots, \mathrm{v}_n) \forall u \in U.$$

Любой вектор $(u^*, \mathrm{v}_1^*, \ldots, \mathrm{v}_n^*)$ называется ситуацией равновесия по Штакельбергу, если $\mathrm{v}_\mathrm{i}^* \in R_i(u^*, \mathrm{v}_{L_i}^*)$ для любого $i \in I$.

Рассмотрим случай, когда оптимальные реакции являются одноэлементными. Построим множество оптимальных реакций для игроков различных уровней. Для игроков нижнего уровня будем иметь

$$R_i(u, \mathrm{v}_{L_i}) = \mathrm{Arg} \max_{\mathrm{v}_i \in V_i} H_i(u, \mathrm{v}_i, \mathrm{v}_{L_i}) = \overline{\mathrm{v}}_i(u, \mathrm{v}_{L_i}),\ i \in S_l.$$

Для игроков более верхних уровней получим

$$R_i(u, \mathrm{v}_{L_i}) = \mathrm{Arg} \max_{\mathrm{v}_i \in V_i} H_i(u, \mathrm{v}_{L_i}, \overline{\mathrm{v}}_{F_i}(u, \mathrm{v}_i), \mathrm{v}_i) = \overline{\mathrm{v}}_i(u, \mathrm{v}_{L_i}).$$

Оптимальным решением центра в рассматриваемой иерархической дифференциальной игре будет

$$u^* = \mathrm{Arg} \max_{u \in U} H_0(u, \overline{\mathrm{v}}_1(\cdot), \ldots, \overline{\mathrm{v}}_n(\cdot)),$$

а ситуация $(u^*, \mathrm{v}_1^*, \ldots, \mathrm{v}_n^*)$, где $\mathrm{v}_i^*(u^*) = \mathrm{v}_i^*$ – равновесием по Штакельбергу.

Таким образом, решение по Штакельбергу и решение по Нэшу в многоуровневой дифференциальной иерархической игре с указанными функционалами выигрышей совпадают. Это обусловлено предположением о единственности точек максимума функционалов выигрышей при всех значениях параметров. В общем случае решение по Нэшу не совпадает с решением по Штакельбергу.

Для нахождения ситуаций равновесия по Нэшу и по Штакельбергу, отражающих оптимальное поведение как центра, так и подсистем, требуется решить значительное количество задач линейного и нелинейного параметрического программирования. Причем, чем выше игрок находится в иерархической структуре, тем больший объем информации для принятия решения ему необходим, поскольку в силу иерархической структуры принятия решений игрок, делающий первый ход, для нахождения своей оптимальной стратегии должен вычислить сначала оптимальные стратегии всех прямо или опосредованно подчиненных ему игроков.

3.5.3. Двухуровневые и ромбовидные иерархические структуры управления

Рассмотрим двухуровневую иерархическую игру, моделирующую процесс принятия решений в системе управления, для которой известны:

1) центр A_0 и множество $I = \{1, 2, \ldots, n\}$ игроков нижнего уровня. Центр обладает правом первого хода, т.е. выбирает первым свое управление и сообщает его игрокам нижнего уровня;
2) множество стратегий центра U и игроков нижнего уровня $V_1 \times V_2 \times \ldots \times V_n$;
3) на множестве $U \times V_1 \times V_2 \times \ldots \times V_n$ определены функции выигрышей

$$H_0 = H_0(u, \mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_n), \quad (3.53)$$

$$H_i = H_i(u, \mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_n), \quad i = 1, 2, \ldots, n. \quad (3.54)$$

Все игроки стремятся максимизировать свои выигрыши. Будем предполагать, что для любого $u \in U$ множество оптимальных реакций $R(u)$ игроков нижнего уровня не пусто.

Рассмотрим частный случай описанной игры – двухуровневую древовидную игру, которую иногда называют «веерной», характеризующуюся тем, что выигрыши игроков нижнего уровня описываются функциями

$$H_i = H_i(u, \mathrm{v}_i). \quad (3.55)$$

Обозначим эту игру через Γ_1. Из вида функций выигрышей (3.55) игроков нижнего уровня можно сделать вывод, что выбор управления v_i любым игроком i зависит только от управления центра. Поэтому множество оптимальных реакций игроков нижнего уровня можно представить в виде прямой суммы множеств оптимальных реакций каждого из игроков.

Справедлива следующая теорема.

Т е о р е м а 6. Решение двухуровневой игры Γ_1 эквивалентно для центра решению двухуровневой игры Γ_2 двух лиц, в которой одним из игроков является центр, а вторым – игрок с функцией выигрыша

$$H(u,\mathrm{v}) = \sum_{i\in I} H_i(u,\mathrm{v}_i)\,, \tag{3.56}$$

где $\mathrm{v} = (\mathrm{v}_1, \mathrm{v}_2, ..., \mathrm{v}_n)$.

Доказательство. В игре Γ_2 множество стратегий центра A_0 есть U, а второго игрока – $V = V_1 \times V_2 \times \ldots \times V_n$. Для любого значения управления $u \in U$ справедливо равенство

$$\max_{\mathrm{v}\in V} H(u,\mathrm{v}) = \sum_{i\in I} \max_{\mathrm{v}_i\in V_i} H_i(u,\mathrm{v}_i)$$

и, следовательно, множество оптимальных реакций игрока A_1 $R^1(u)$ в игре Γ_2 представимо в виде

$$R^1(u) = R_1(u)\times R_2(u)\times \ldots \times R_n(u)\,,$$

где $R_i(u) = \operatorname{Arg}\max_{\mathrm{v}_i\in V_i} H_i(u,\mathrm{v}_i)$.

Но поскольку множество оптимальных реакций игроков нижнего уровня в игре Γ_1 есть также

$$R(u) = R_1(u)\times R_2(u)\times \ldots \times R_n(u)\,,$$

то $R(u) = R^1(u)$. Следовательно, множества оптимальных реакций игроков нижнего уровня в играх Γ_1 и Γ_2 совпадают. Учитывая, что функции выигрыша центра в обеих играх одинаковы и $R(u) = R^1(u)$, получим, что множества оптимальных решений центра в них также совпадают, а, следовательно, решения игр Γ_1 и Γ_2 эквивалентны.

Замечание. Теорема остается справедливой и для функций $H = \sum_{i\in I} \lambda_i H_i(u,\mathrm{v}_i)$, где параметр $\lambda_i > 0$.

Рассмотрим пример из области математической экономики [12], иллюстрирующий проблему установления рациональных цен на товары и ресурсы. В качестве модели используется веерная иерархическая система.

Пусть известен общий объем товаров Q, производимых промышленностью за фиксированный промежуток времени (Q – вектор с положительными компонентами). Все потребители товаров разбиты на n однородных групп, каждая из которых имеет свою функцию полезности (функцию выигрыша) $H_i(\mathrm{v}^i)$, заданную на пространстве товаров, где вектор v^i характеризует объем товаров, закупаемых i-й группой. Суммарный объем товаров, закупаемых всеми группами, ограничен векторной величиной Q, т.е.

$$\sum_{i=1}^{n} \mathrm{v}^i \le Q,\ Q \in E_m\,. \tag{3.57}$$

Управление центра $u = (p, s)$ заключается в выборе вектора цен на товары $p = (p_1, p_2, \ldots, p_n)$ и установлении уровня дохода (заработной платы) каждой группы потребителей, т.е. общего количества денег s_i, получаемого всеми потребителями i-й группы за данный промежуток времени. Множество допустимых управлений i-й группы опишем следующим образом:

$$V_i(p,s_i) = \{\mathrm{v}^i \mid \mathrm{v}^i \ge 0, < p,\ \mathrm{v}^i > \le s_i\},\quad i = 1,\ 2,\ ...,\ n\,.$$

где $<p, \mathrm{v}_i>$ – скалярное произведение векторов p и v^i.

Оптимальные стратегии (управления) i-й группы образуют множество ее оптимальных реакций

$$R_i(p,s_i) = \operatorname{Arg}\max_{\mathrm{v}^i\in V_i(p,s_i)} H_i(\mathrm{v}^i)\,.$$

При этом для всех $\mathrm{v}^i \in R_i(p,s_i)$ должно выполняться условие (3.57). Для этого зададим множество допустимых управлений центра следующим образом:

$$U = \left\{(p,s) \mid p \ge 0, s \ge 0, \sum_{i=1}^{n} \mathrm{v}^i \le Q \forall \mathrm{v}^i \in R_i(p,s_i),\ i = 1,\ 2,\ ...,\ n\right\}.$$

Критерий эффективности (функцию выигрыша) центра зададим в виде

$$H_0(\mathrm{v}^1,\mathrm{v}^2,...,\mathrm{v}^n) = \sum_{i=1}^{n} \alpha_i H_i(\mathrm{v}^i)\,,$$

где α_i – положительные константы.

Такой критерий является довольно естественным, так как представляет собой некоторый обобщенный показатель среднего уровня потребления.

Если множества $R_i(p, s_i)$ состоят для всех значений (p, s) из единственного элемента $\mathrm{v}_0^i(p,s)$, то задача выбора оптимального управления центра (p^o, s^o) будет иметь вид

$$\max_{(p,s)\in U} \sum_{i=1}^{n} \alpha_i H_i(v_0^i(p,s_i)) .$$

Если же множества $R_i(p, s_i)$ не являются одноэлементными, то для определения оптимального управления центра можно воспользоваться принципом гарантированного результата, т.е. решить задачу

$$\max_{(p,s)\in U} \min_{v_0^i \in R_i(p,s_i)} \sum_{i=1}^{n} \alpha_i H_i(v_0^i) .$$

Рассмотрим ромбовидную систему управления. Схема простейшей ромбовидной системы представлена на рис. 3.13. На первом уровне располагается центр (игрок A_0), на втором – игроки B_1 и B_2, которые подчинены центру, на третьем уровне – игрок B_3, который подчинен всем трем игрокам. Центр выбирает свою стратегию (управление) из множества U. Множества альтернатив игроков B_1 и B_2 обозначим $V_1(u)$ и $V_2(u)$ Игрок B_3 выбирает управление из множества $V_3(u, v_1, v_2)$, где $v_1 \in V_1(u)$, $v_2 \in V_2(u)$. Функции выигрыша игроков заданы в виде $H_0(u, v_1, v_2, v_3)$, $H_1(u, v_1, v_3)$, $H_2(u, v_2, v_3)$, $H_3(u, v_1, v_2)$. Таким образом, мы определили бескоалиционную игру в нормальной форме:

$$\Gamma = < U, V_1, V_2, V_3, H_0, H_1, H_2, H_3 >$$

Опишем процесс построения ситуаций равновесия по Нэшу и по Штакельбергу в этой игре. Предположим, что в качестве принципа оптимальности в игре выбрано равновесие по Нэшу, и построим множество оптимальных реакций игроков

$$R_3(u, v_1, v_2) = \mathrm{Arg}\max_{v_3 \in V_3} H_3(u, v_1, v_2, v_3) . \qquad (3.58)$$

Выберем некоторое управление $v_3^*(u, v_1, v_2) \in R_3(u, v_1, v_2)$. Функции выигрыша игроков первого уровня (B_1, B_2) представим в виде

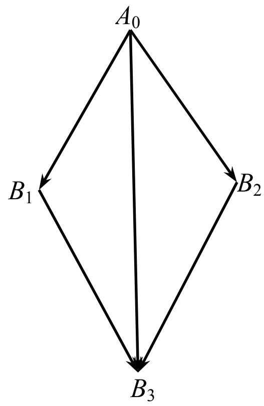

Рис. 3.13

$$H_1^*(u, v_1, v_2) = H_1(u, v_1, v_3^*(u, v_1, v_2)) ,$$

$$H_2^*(u, v_1, v_2) = H_2(u, v_2, v_3^*(u, v_1, v_2)) .$$

Множество оптимальных реакций игроков первого уровня при известном выборе $v_3^*(u, v_1, v_2)$ получим следующим образом:

$$R_{1,2}^* = \{(v_1, v_2) | v_1 \in V_1, \quad v_2 \in V_2 : H_1^*(u, v_1, v_2) \ge H_1^*(u, v_1, v_2),$$
$$H_2^*(u, v_1, v_2) \ge \forall v_1' \in V_1,\ v_2' \in V_2\}.$$

Предположим, что множество $R_{1,2}^*(u)$ не пусто и выберем некоторые управления, $v_1^*(u)$, $v_2^*(u)$ такие, что, ($v_1^*(u)$, $v_2^*(u)$) $\in$ $R_{1,2}^*(u)$. Рассмотрим задачу

$$\max_{u\in U} H_0(u, v_1^*(u), v_2^*(u), v_3^*(u, v_1^*(u), v_2^*(u))) . \qquad (3.59)$$

Пусть u^* есть решение задачи (3.59). Тогда ситуация (u^*, $v_1^*(u^*)$, $v_2^*(u^*)$, $v_3^*(u^*, v_1^*, v_2^*)$) является ситуацией равновесия по Нэшу в бескоалиционной игре четырех лиц A_0, B_1, B_2, B_3.

Действительно, в силу того, что u* является решением задачи (3.59), для любых $u \in U$ выполняется неравенство

$$H_0(u^*, v_1^*(u^*), v_2^*(u^*), v_3^*(u^*, v_1^*(u^*), v_2^*(u^*))) \ge$$

$$\ge H_0(u, v_1^*(u), v_2^*(u), v_3^*(u, v_1^*(u), v_2^*(u))) .$$

Учитывая определение множества $R^*_{1,2}(u)$, получим, что для любых $v_1 \in V_1$, $v_2 \in V_2$ справедливы неравенства

$$H_1(u^*, v_1^*(u^*), v_3^*(u^*, v_1^*, v_2^*)) \geq H_1(u^*, v_1, v_3^*(u^*, v_1, v_2^*)),$$

$$H_2(u^*, v_2^*(u^*), v_3^*(u^*, v_1^*, v_2^*)) \geq H_2(u^*, v_2, v_3^*(u^*, v_2, v_3^*)).$$

И, наконец, из определения множества $R_3(u, v_1, v_2)$ следует

$$H_3(u^*, v_1^*(u^*), v_2^*(u^*), v_3^*(u^*, v_1^*, v_2^*)) \geq H_3(u^*, v_1^*(u^*), v_2^*(u^*), v_3)$$

для любого управления $v_3 \in V_3$. Таким образом, построенная ситуация является ситуацией равновесия по Нэшу.

Рассмотрим теперь процесс построения ситуации равновесия по Штакельбергу в ромбовидной игре. В соответствие с определением решения по Штакельбергу множество оптимальных реакций игрока B_3 будет следующим:

$$R_3(u, v_1, v_2) = \{v_3 \mid v_3 \in V_3, H_3(u, v_1, v_2, v_3) \geq H_3(u, v_1, v_2, v_3')\forall v_3' \in V_3\}.$$

Нетрудно заметить, что в данном случае множество оптимальных реакций игрока B_3 такое же, что и при построении ситуации равновесия по Нэшу.

Множество оптимальных реакций игроков B_1, B_2 будем строить, исходя из того, что они должны обеспечить себе гарантированный результат при наихудших для каждого из них действиях игрока B_3. Обозначим

$$H_1^u(v_1, v_2) = \min_{v_3 \in R_3(u, v_1, v_2)} H_1(u, v_1, v_3), \tag{3.60}$$

$$H_2^u(v_1, v_2) = \min_{v_3 \in R_3(u, v_1, v_2)} H_2(u, v_2, v_3). \tag{3.61}$$

Эти функции задают значения выигрышей в ситуации, когда игроки B_1 и B_2 придерживаются стратегий v_1 и v_2, а игрок B_3 выбирает управления, наихудшие соответственно для B_1 или B_2. Множество оптимальных реакций игроков B_1, B_2 определим следующим образом:

$$R^*_{1,2}(u) = \{(v_1, v_2) \mid v_1 \in V_1, v_2 \in V_2, H_1^u(v_1, v_2) \geq H_1^u(v_1', v_2),$$

$$H_2^u(v_1, v_2) \geq H_2^u(v_1, v_2')\forall v_1' \in V_1, v_2' \in V_2\},$$

или, что то же самое,

$$R^*_{1,2}(u) = \{(v_1, v_2) \mid v_1 \in V_1, v_2 \in V_2, \min_{v_3 \in R_3(u, v_1, v_2)} H_1(u, v_1, v_3) \geq$$

$$\geq \min_{v_3 \in R_3(u, v_1', v_2)} H_1(u, v_1', v_3), \min_{v_3 \in R_3(u, v_1, v_2)} H_2(u, v_2, v_3) \geq$$

$$\geq \min_{v_3 \in R_3(u, v_1, v_2')} H_2(u, v_2', v_3)\forall v_1' \in V_1, v_2' \in V_2\}.$$

Множество оптимальных управлений центра в соответствие с определением решения по Штакельбергу запишем в виде

$$PR = \{u \mid u \in U, \min_{(v_1, v_2) \in R_{1,2}(u)} \min_{v_3 \in R_3(u, v_1, v_2)} H_0(u, v_1, v_2, v_3) \geq$$

$$\geq \min_{(v_1, v_2) \in R_{1,2}(u')} \min_{v_3 \in R_3(u', v_1, v_2)} H_0(u', v_1, v_2, v_3)\forall u' \in U\}.$$

Решением по Штакельбергу ромбовидной игры будет любой вектор $(u^0, v_1^0, v_2^0, v_3^0)$, такой, что $u^0 \in PR$, $(v_1^0, v_2^0) \in R^*_{1,2}(u^0)$, $v_3^0 \in (R_3(u^0, v_1^0, v_2^0))$.

Здесь, очевидно, требует некоторого обсуждения построение множества оптимальных реакций $R_{1,2}(u)$ игроков B_1 и B_2. Если множество $R_3(u, v_1, v_2)$ для всех значений u, v_1, v_2 является одноэлементным множеством или минимумы выражений (3.60) и (3.61) достигаются на одних и тех же управлениях игрока B_3, то множество $R_{1,2}(u)$ является множеством ситуации равновесия по Нэшу в неантагонистической игре двух лиц (игроков B_1 и B_2) с функциями выигрышей $H_1(u, v_1, v_3(u, v_1, v_2))$ и $H_1(u, v_1, v_3(u, v_1, v_2))$, где

$$v_3(u, v_1, v_2) = \mathrm{Arg} \min_{v_3 \in R_3(u, v_1, v_2)} H_1(u, v_1, v_3) = \mathrm{Arg} \min_{v_3 \in R_3(u, v_1, v_2)} H_2(u, v_2, v_3).$$

В противном случае множество $R_{1,2}(u)$ является множеством ситуаций равновесия по Нэшу в неантагонистической игре игроков B_1 и B_2 с функциями выигрыша $H_1^u(v_1, v_2)$ и $H_2^u(v_1, v_2)$. Любая ситуация $(v_1, v_2) \in R_{1,2}(u)$ характерна тем, что каждый из игроков B_1, B_2 в этой ситуации гарантирует себе максимальный выигрыш при фиксированной стратегии другого и наихудшем выборе управления игроком B_3.

Для иллюстрации описанного процесса нахождения оптимального по Штакельбергу решения в ромбовидной игре рассмотрим пример.

П р и м е р . Пусть A_0 – центр, B_1 и B_2 – игроки первого уровня, игрок B_3 расположен на третьем уровне иерархии. Функции выигрыша игроков имеют вид

$$H_0(u, \text{v}_1, \text{v}_2, \text{v}_3) = u\text{v}_1\text{v}_2\text{v}_3, \;\; H_1(u, \text{v}_1, \text{v}_3) = u\text{v}_1\text{v}_3 ,$$

$$H_2(u, \text{v}_2, \text{v}_3) = u\text{v}_2\text{v}_3, \;\; H_3(u, \text{v}_1, \text{v}_2, \text{v}_3) = (u + \text{v}_1 + \text{v}_2)\text{v}_3 ,$$

$$U = \{-1, 1\}, \;\; V_1 = \{-1, 1\}, \;\; V_2 = \{-1, 1\}, \;\; V_3 = \{-1, 1\} .$$

Множество оптимальных реакций игрока B_3 для каждого набора управлений игроков A_0, B_1 и B_2 состоит из единственного элемента, а именно,

$$R_3(u, \text{v}_1, \text{v}_2) = \{\text{sign}\,(u + \text{v}_1 + \text{v}_2)\} = \{\text{v}_3^0\} .$$

Опишем множество оптимальных реакций игроков B_1 и B_2. Для этого запишем сначала выражения для функций H_1^u и H_2^u:

$$H_1^u(\text{v}_1, \text{v}_2) = u\text{v}_1\text{sign}\,(u + \text{v}_1 + \text{v}_2) ,$$

$$H_2^u(\text{v}_1, \text{v}_2) = u\text{v}_2\text{sign}\,(u + \text{v}_1 + \text{v}_2) .$$

В соответствие с определением множества $R_{1,2}(u)$ для различных значений управления u получим

$$R_{1,2}(1) = \{(-1, 1), (1, -1), (-1, -1), (1, 1)\},$$
$$R_{1,2}(-1) = \{(-1, 1), (1, -1), (1, 1)\}.$$

Вычислим значения функции

$$\min_{(\text{v}_1, \text{v}_2)\in R_{1,2}(u)} H_0(u, \text{v}_1, \text{v}_2, \text{v}_3^0): \min_{(\text{v}_1, \text{v}_2)\in R_{1,2}(1)} H_0(1, \text{v}_1, \text{v}_2, \text{v}_3^0) =$$

$$= \min_{(\text{v}_1, \text{v}_2)\in R_{1,2}(1)} (|\,u\,|\,\text{v}_1\text{v}_2 + u\,|\,\text{v}_1\,|\,\text{v}_2 + u\text{v}_1\,|\,\text{v}_2\,|) = -1,$$

$$\min_{(\text{v}_1, \text{v}_2)\in R_{1,2}(-1)} H_0(-1, \text{v}_1, \text{v}_2, \text{v}_3^0) = -1 .$$

Таким образом, $PR = U$. Ситуациями равновесия по Штакельбергу являются следующие ситуации: (1, –1, 1, 1), (1, 1, –1, 1), (1, –1, –1, –1), (1, –1, –1, 1), (–1, –l, l, –l), (–1, 1, –1, –1), (–1, 1, 1, 1).

3.5.4. Динамические модели иерархических систем

В предыдущих параграфах мы исследовали в основном статические модели, или модели, в которых динамика систем не оказывала влияния на процесс принятия решений и служила лишь иллюстрацией возможной постановки задачи. Однако изучение динамических систем управления представляет собой интерес, поскольку здесь возникает целый ряд специфических проблем. Как правило, развитие (движение) системы во времени приводит к тому, что изменяется и процесс принятия решений. Если в начальный момент игроки (подсистемы) выбирают оптимальные управления, ориентируясь на начальные условия и интервал времени $[t_0, T]$, то по истечении некоторого времени меняются состояние системы, а также множества допустимых управлений, могут меняться и функционалы выигрышей, появляется новая информация о процессе в целом.

Динамику иерархической системы будем описывать с помощью обыкновенных дифференциальных уравнений

$$\dot{x} = f(x, u, \text{v}_1, \text{v}_2, ..., \text{v}_n), \;\; x(t_0) = x^0 ; \tag{3.62}$$

здесь $x \in E_m$ – вектор фазовых переменных, описывающий состояние системы в момент t.

Изменение системы во времени происходит под воздействием управления центра $u(t)\in U$ и управлений подсистем $\text{v}_1(t)$, $\text{v}_2(t)$, ..., $\text{v}_n(t)$, $\text{v}_i(t) \in V_i$; множества U, V_1, V_2, ..., V_n будем называть множествами допустимых управлений. На множестве траекторий системы заданы критерии эффективности (функционалы) подсистем

$$H_i(u, \text{v}_1, \text{v}_2, ..., \text{v}_n) = \int_{t_0}^{T} h_i(x, u, \text{v}_1, \text{v}_2, ..., \text{v}_n)dt , \quad i = 0, 1, 2, ..., n . \tag{3.63}$$

Будем считать, что параметры управляемой динамической системы удовлетворяют следующим условиям:

1) множества U, V_1, ..., V_n компактны в соответствующих векторных пространствах;

2) вектор-функция $f \in E_m$ непрерывно дифференцируема по своим переменным;
3) существует константа $x > 0$, такая, что при любых $u \in U$, $v_i \in V_i$ выполнено неравенство

$$\left\| f(x, u, v_1, v_2, \ldots, v_n) \right\| \le x(1 + \|x\|);$$

4) функции $h_i(x, u, v_1, v_2, \ldots, v_n)$ положительны и интегрируемы, $i = 1, 2, \ldots, n$.

Выполнения этих условий достаточно для существования единственного решения задачи Коши при любых кусочно-непрерывных допустимых управлениях. Анализ такой иерархической системы может быть сведен к исследованию решений дифференциальной иерархической игры. Под стратегиями игроков (центра и подсистем) мы будем понимать выбор ими управлений $u(t)$, $v_1(t)$, ..., $v_n(t)$, т.е. в данном случае можно считать, что стратегиями игроков являются их управления.

Рассмотрим дифференциальную иерархическую игру

$$\Gamma(t_0, x^0) = \left\langle \{A_0, I\},\ \{U, V_i\}_{i \in I}, \{H_0, H_i\}_{i \in I} \right\rangle,$$

динамика которой описывается системой дифференциальных уравнений (3.62) с функционалами выигрышей (3.63).

Решением дифференциальной игры называется множество всех управлений, оптимальных в смысле выбранного принципа оптимальности. Обозначим решение игры $T(t_0, x^0)$ через $M(t_0, x^0)$.

Выберем оптимальное управление ($u(t)$, $v_1(t)$, ..., $v_n(t)$) из множества $M(t_0, x^0)$ и обозначим соответствующую оптимальную траекторию через $x(t)$.

Важным свойством всякого оптимального решения является его динамическая устойчивость. Оказывается, что далеко не все принципы оптимальности обладают этим свойством. Напомним, в чем состоит понятие динамической устойчивости решений дифференциальных игр. Предположим, что система развивается вдоль оптимальной траектории $x(t)$ под воздействием оптимальных управлений $u(t)$, $v_1(t)$, ..., $v_n(t)$. В каждый момент времени будем рассматривать игру $T(t, x(t))$, в которой множества допустимых управлений представляют собой сужения множеств U, V_1, V_2, ..., V_n на интервал времени $[t, T]$. Обозначим эти множества через U^t V_1^t, ..., V_n^t. Они будут включать в себя измеримые функции, заданные на интервале $[t, T]$. Обозначим через

$$H_i^t\left(u^{t,T}, v_1^{t,T}, \ldots, v_n^{t,T}\right) = \int_t^T h_i(x, u^{t,T}, v_1^{t,T}, \ldots, v_n^{t,T})\,d\tau,$$
$$i = 0, 1, 2, \ldots, n.$$

функционалы выигрышей в игре $T(t, x(t))$, где $u^{t,T} \in U^t$, $v^{t,T} \in U^t$. Обозначим также через u^t v_1^t, ..., v_n^t сужения оптимальных управлений игроков на отрезок $[t, T]$, т.е., например, $u^t(\tau) = u(\tau)$, $t < \tau < T$, и т.д. Игра $T(t, x, \{t\})$ называется *текущей игрой*.

Пусть $M(t, x(t))$ – решение текущей игры, которая развивается вдоль оптимальной траектории $x(t)$, а $u(t)$, $v_1(t)$, ..., $v_n(t)$ – оптимальные управления игроков.

Говорят, что ситуация $(u, v_1, v_2, \ldots, v_n) \in M(t_0, x_0)$ *динамически устойчива*, если в любой момент времени сужения оптимальных управлений игроков образуют ситуацию в текущей игре (u^t v_1^t, ..., v_n^t), которая принадлежит решению игры $\Gamma(t, x(t))$, т.е. (u^t v_1^t, ..., v_n^t) $\in M(t, x(t))$.

Определение. Решение $M(t_0, x^0)$ дифференциальной игры $T(t_0, x^0)$ называется динамически устойчивым, если для любого $t \in [t_0, T]$ и любой ситуации $(u, v_1, v_2, \ldots, v_n) \in M(t_0, x^0)$ выполнено условие

$$(u^t, v_1^t, v_2^t, \ldots, v_n^t) \in M(t, x(t)),$$

где $x(t)$ – оптимальная траектория системы (3.62), соответствующая оптимальным управлениям u, v_1, v_2, ..., v_n.

Свойство динамической устойчивости является очень важной характеристикой решения дифференциальной игры. Если какая-либо ситуация $(u, v_1, v_2, \ldots, v_n) \in M(t_0, x^0)$ не является динамически устойчивой (т.е. решение $M(t_0, x^0)$ не является динамически устойчивым), то это означает, что в некоторый момент времени t управления u^t v_1^t, ..., v_n^t не будут оптимальными в текущей игре $T(t_0, x^0)$, и игроки перестанут придерживаться этих управлений в дальнейшем. Если же решение динамически устойчиво, то у игроков не будет оснований изменять свои управления до конца игры.

Свойство динамической устойчивости присуще далеко не всем принципам оптимальности, например, равновесие по Нэшу является динамически устойчивым, а равновесие по Штакельбергу таковым не является.

Практическая ценность свойства динамической устойчивости решений дифференциальных игр состоит в том, что если игроки договариваются в начале игры о реализации некоторой оптимальной ситуации в течение всей игры, то эта договоренность для динамически устойчивых принципов оптимальности сохраняется до конца игры.

Покажем, что равновесие по Нэшу в дифференциальной игре $T(t_0, x^0)$ является динамически устойчивым. На первом уровне иерархии находится игрок A_0. на втором – игроки B_1, B_2, .., B_n, входящие в множество I. Обозначим через $v = (v_1, v_2, \ldots, v_n)$ вектор управлений игроков нижнего уровня. Пусть оптимальные управления являются программными, т.е. являются функциями времени: $u = u(t)$, $v = v(t)$.

Предположим, что для некоторого момента времени $\tau(u^\tau, v^\tau) \overline{\in} M(\tau, x(\tau))$. Следовательно, найдутся игрок нижнего уровня j и управление $\overline{v}_j^{\tau,T}$ или управление центра $\overline{u}^{\tau,T}$ такие, что выполнено одно из неравенств

$$H_j^\tau(u^\tau, u^\tau \| \overline{u}_j^{\tau,T}) > H_j^\tau(u^\tau, v^\tau); \tag{3.64}$$

$$H_0^{\tau}(\overline{u}^{\tau,T}, \mathrm{v}^{\tau}) > H_0^{\tau}(u^{\tau}, \mathrm{v}^{\tau}) \ . \qquad (3.65)$$

Пусть выполнено неравенство (3.64). Обозначим

$$\mathrm{v}_j^{\tau}(t) = \begin{cases} \mathrm{v}_j(t), & t_0 \le t \le \tau; \\ \overline{u}_j^{\tau,T}(t), & \tau < t \le T. \end{cases}$$

Тогда из неравенства (3.64) будет следовать неравенство

$$H_j(u, \mathrm{v}) < H_j(u, \mathrm{v} \| \mathrm{v}') \ ,$$

а это означает, что ситуация (u, v) не является ситуацией равновесия по Нэшу. Аналогичный вывод можно сделать, если выполнено неравенство (3.65). Следовательно, предположение о динамической неустойчивости равновесия по Нэшу неверно, и, значит, равновесие по Нэшу в программных стратегиях в игре $T(t_0, x^0)$ динамически устойчиво.

Читатель, желающий глубже ознакомиться с использованием понятия динамической устойчивости решений в сложных системах, может сделать это, прочитав дополнительно, например, книги [2, 4, 19, 21].

Контрольные вопросы

1. Что такое теория игр?
2. Обоснуйте необходимость использования теории игр в системном анализе.
3. Сформулируйте принципы оптимальности в иерархических теоретико-игровых моделях.
4. Даете характеристику двухуровневым и ромбовидным иерархическим структурам управления.
5. Дайте характеристику динамическим моделям иерархических систем.

ЗАКЛЮЧЕНИЕ

В системном анализе, понимаемом как исследование проблемы принятия решения в сложной системе, обращают на себя внимание чрезвычайно широкие и разнообразные области приложений. Они простираются от техники до экологии, от математики до социального планирования, от космических исследований до процессов обучения. Казалось бы, системный анализ давно должен иметь какое-то общее изложение, удовлетворяющее все эти области. Однако, такое изложение, где были бы систематизированы те принципы, рассуждения и методики, на применении которых основано множество прикладных работ, неизвестно, что дает основание на дальнейшую их разработку.

Системный анализ как знание существует, но формулировки его положений и приемов крайне фрагментарны и, за редким исключением, нацелены на конкретные классы задач. Вдобавок укоренившееся представление об основах системного анализа состоит в разобщенном наборе методологических положений и математизированных структур, что в равной степени относится к отечественной и зарубежной литературе.

Предлагаемое учебное пособие – это попытка сделать шаг в оформлении новой научной дисциплины, попытка выйти за типичное в настоящее время узкоприкладное изложение системного анализа.

Авторы настоящего пособия не останавливаются на достигнутом и продолжают свои исследования в этой области.

СПИСОК ЛИТЕРАТУРЫ

1. Калман, Р. Очерки по математической теории систем / Р. Калман, П. Фалб, М. Арбиб. – М. : Мир, 1971. – 400 с.
2. Месарович, М. Общая теория систем: математические основы / М. Месарович, Я. Такахара. – М. : Мир, 1978. – 312 с.
3. Месарович, М. Теория иерархических многоуровневых систем / М. Месарович, Д. Мако, Я. Такахара. – М. : Мир, 1973. – 344 с.
4. Портер, У. Современные основания общей теории систем / У. Портер. – М. : Наука, 1971. – 556 с.
5. Клир, Дж. Системотология. Автоматизация решения системных задач / Дж. Клир. – М. : Радио и связь, 1990. – 544 с.
6. Раскин, Л.Г. Анализ сложных систем и элементы теории управления / Л.Г. Раскин. – М. : Сов. радио, 1976. – 344 с.
7. Айзерман, М.А. Выбор вариантов: основы теории / М.А. Айзерман, Ф.Т. Алексеев. – М. : Наука, 1990. – 240 с.
8. Теория выбора и принятия решений / И.М. Макаров, Т.М. Виноградская, А.А. Рубчинский, В.Б. Соколов. – М. : Наука, 1982. – 328 с.
9. Ногин, В.Д. Принятие решений в многокритериальной среде / В.Д. Ногин. – М. : ФИЗМАТЛИТ, 2002. – 176 с.
10. Горелова, В.Л. Основы прогнозирования систем / В.Л. Горелова, Е.Н. Мельников. – М. : Высшая школа, 1986. – 287 с.
11. Дубов, Ю.А. Многокритериальные модели формирования выбора вариантов систем / Ю.А. Дубов, С.И. Травкин, В.Н. Якимец. – М. : Наука, 1986. – 296 с.
12. Добкин, В.М. Системный анализ в управлении / В.М. Добкин. – М. : Химия, 1984. – 224 с.
13. Поспелов, Д.А. Ситуационное управление: теория и практика / Д.А. Поспелов. – М. : Наука, 1986. – 288 с.
14. Губанов, В.А. Введение в системный анализ / В.А. Губанов, В.В. Захаров, А.Н. Коваленко. – Л. : Изд-во Ленинградского ун-та, 1988. – 232 с.
15. Директор, С. Введение в теорию систем / С. Директор, Р. Рорер. – М. : Мир, 1974. – 464 с.
16. Моисеев, Н.Н. Элементы теории оптимальных систем / Н.Н. Моисеев. – М. : Наука, 1975. – 526 с.
17. Алексеев, В.М. Оптимальное управление / В.М. Алексеев, В.М. Тихомиров, С.В. Фомин. – М. : Наука, 1979. – 429 с.
18. Иоффе, А.Д. Теория экстремальных задач / А.Д. Иоффе, В.М. Тихомиров. – М. : Наука, 1974. – 479 с.
19. Красовский, Н.Н. Управление динамической системой / Н.Н. Красовский. – М. : Наука, 1985. – 520 с.
20. Мулен, Э. Кооперативное принятие решений: аксиомы и модели / Э. Мулен. – М. : Мир, 1991. – 464 с.
21. Куржанский, А.Б. Управление и наблюдение в условиях неопределенности / А.Б. Куржанский. – М. : Наука, 1977. – 392 с.

ОГЛАВЛЕНИЕ