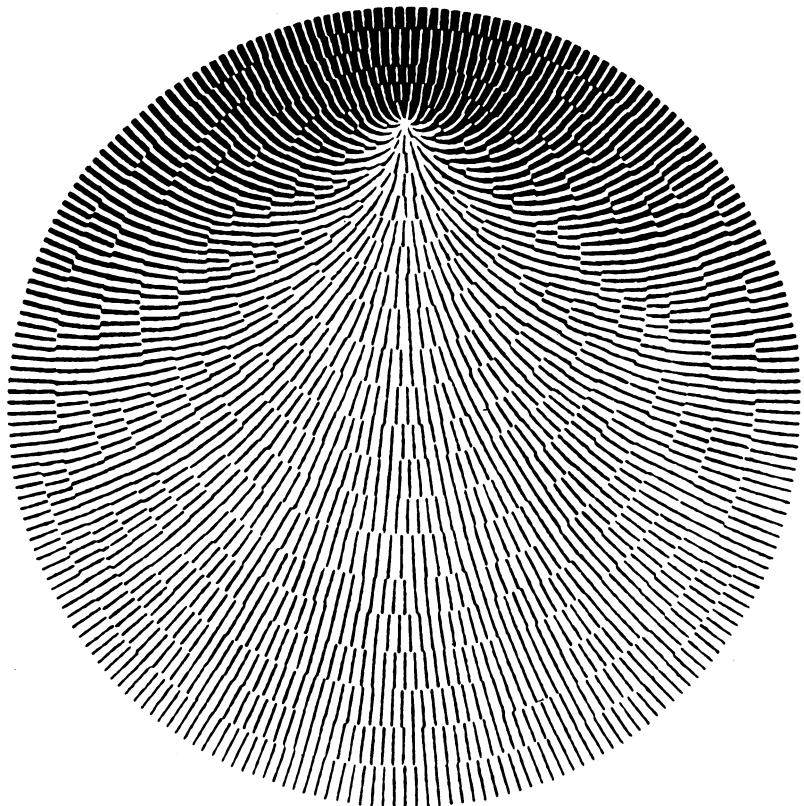


Ф.И.Перегудов, Ф.П.Тарасенко
ОСНОВЫ СИСТЕМНОГО АНАЛИЗА



Ф.И.Перегудов, Ф.П.Тарасенко

ОСНОВЫ
СИСТЕМНОГО
АНАЛИЗА

Томск-1997

ББК 22.161

П27

УДК 517

Р е ц е н з е н т: д-р техн. наук, проф. А.А.Золотарев (Московский институт инженеров гражданской авиации)

Перегудов Ф.И., Тарасенко Ф.П.

П27 Основы системного анализа: Учеб. 2-е изд., доп.–Томск:

Изд-во НТЛ, 1997.–396 с.: ил.

ISBN 5-89503-004-1

Перед специалистами любого профиля часто возникают следующие вопросы: как решить реальную проблему и не создать при этом новых проблем? как уменьшить сложность возникшей ситуации? как правильно организовать исследование существующей системы или проектирование новой? На эти вопросы отвечает современный прикладной системный анализ. Знакомство с методами системного анализа необходимо практически всем специалистам, и его элементы вошли в учебные планы многих вузов. В данной книге излагаются основные понятия и методы системного анализа.

Во второе издание (1-е – 1989 г.) внесены некоторые дополнения.

ББК 22.161

517.2

ISBN 5-89503-004-1

© Ф.И.Перегудов,
Ф.П.Тарасенко, 1989
© Ф.И.Перегудов,
Ф.П.Тарасенко, 1997,
с изменениями

Подписано к печати 25.10.96. Формат бумаги 60×84/16. Бумага книжно-журнальная. Печать офсетная. Усл. печ. л. 23,02. Уч.-изд. л. 20,63. Тираж 1000 экз.

Отпечатано с оригинал-макета, подготовленного АО “Издательство НТЛ”.

Лицензия ЛР № 064401 от 22.01.96.

634050, Томск, пл.Революции 1, т. (382-2) 23-33-35

Типография издательства “Красное Знамя”. 634050, г.Томск, пр.Фрунзе, 103/1

ОГЛАВЛЕНИЕ

Предисловие **3**

Введение **5**

ВОЗНИКНОВЕНИЕ И РАЗВИТИЕ СИСТЕМНЫХ ПРЕДСТАВЛЕНИЙ **7**

- 1.1. Предварительные замечания **7**
 - 1.2. Роль системных представлений в практической деятельности **8**
 - 1.3. Внутренняя системность познавательных процессов **14**
 - 1.4. Системность как всеобщее свойство материи **17**
 - 1.5. Краткий очерк истории развития системных представлений **21**
 - Заключение **32**
 - Литература **33**
 - Упражнения **33**
 - Вопросы для самопроверки **34**
-

МОДЕЛИ И МОДЕЛИРОВАНИЕ **35**

- 2.1. Широкое толкование понятия модели **35**
 - 2.2. Моделирование – неотъемлемый этап всякой целенаправленной деятельности **37**
 - 2.3. Способы воплощения моделей **41**
 - 2.4. Условия реализации свойств моделей **47**
 - 2.5. Соответствие между моделью и действительностью: различия **49**
 - 2.6. Соответствие между моделью и действительностью: сходство **54**
 - 2.7. О динамике моделей **57**
 - Заключение **62**
 - Литература **64**
 - Упражнения **65**
 - Вопросы для самопроверки **68**
-

СИСТЕМЫ. МОДЕЛИ СИСТЕМ **69**

- 3.1. Множественность моделей систем **69**
 - 3.2. Первое определение системы **69**
 - 3.3. Модель “черного ящика” **72**
 - 3.4. Модель состава системы **78**
 - 3.5. Модель структуры системы **81**
 - 3.6. Второе определение системы. Структурная схема системы **84**
 - 3.7. Динамические модели систем **87**
 - Заключение **94**
 - Литература **96**
 - Упражнения **97**
 - Вопросы для самопроверки **98**
-

ИСКУССТВЕННЫЕ И ЕСТЕСТВЕННЫЕ СИСТЕМЫ 99

- 4.1. Искусственные системы и естественные объекты **99**
 - 4.2. Обобщение понятия системы. Искусственные и естественные системы **101**
 - 4.3. Различные классификации систем **104**
 - 4.4. О больших и сложных системах **112**
 - Заключение **118**
 - Литература **120**
 - Упражнения **120**
 - Вопросы для самопроверки **122**
-

ИНФОРМАЦИОННЫЕ АСПЕКТЫ ИЗУЧЕНИЯ СИСТЕМ 123

- 5.1. Информация как свойство материи **123**
 - 5.2. Сигналы в системах **124**
 - 5.3. Случайный процесс – математическая модель сигналов **127**
 - 5.4. Математические модели реализаций случайных процессов **130**
 - 5.5. О некоторых свойствах непрерывных сигналов **137**
 - 5.6. Энтропия **143**
 - 5.7. Количество информации **149**
 - 5.8. Об основных результатах теории информации **154**
 - Заключение **163**
 - Литература **165**
 - Упражнения **165**
 - Вопросы для самопроверки **168**
-

РОЛЬ ИЗМЕРЕНИЙ В СОЗДАНИИ МОДЕЛЕЙ СИСТЕМ 169

- 6.1. Эксперимент и модель **169**
 - 6.2. Измерительные шкалы **172**
 - 6.3. Расплывчатое описание ситуаций **191**
 - 6.4. Вероятностное описание ситуаций. Статистические измерения **196**
 - 6.5. Регистрация экспериментальных данных и ее связь с последующей их обработкой **200**
 - Заключение **206**
 - Литература **206**
 - Упражнения **207**
 - Вопросы для самопроверки **208**
-

ВЫБОР (ПРИНЯТИЕ РЕШЕНИЙ) 209

- 7.1. Многообразие задач выбора **209**
7.2. Критериальный язык описания выбора **211**
7.3. Описание выбора на языке бинарных отношений **221**
7.4. Язык функций выбора **228**
7.5. Групповой выбор **231**
7.6. Выбор в условиях неопределенности **237**
7.7. О выборе в условиях статистической неопределенности **241**
7.8. Выбор при расплывчатой неопределенности **251**
7.9. Достоинства и недостатки идеи оптимальности **255**
7.10. Экспертные методы выбора **259**
7.11. Человеко–машины системы и выбор **263**
7.12. Выбор и отбор **267**
 Заключение **272**
 Литература **273**
 Упражнения **275**
 Вопросы для самопроверки **276**
-

ДЕКОМПОЗИЦИЯ И АГРЕГИРОВАНИЕ КАК ПРОЦЕДУРЫ СИСТЕМНОГО АНАЛИЗА 277

- 8.1. Анализ и синтез в системных исследованиях **277**
8.2. Модели систем как основания декомпозиции **281**
8.3. Алгоритмизация процесса декомпозиции **289**
8.4. Агрегирование, эмерджентность, внутренняя целостность систем **298**
8.5. Виды агрегирования **301**
 Заключение **312**
 Литература **314**
 Упражнения **315**
 Вопросы для самопроверки **316**
-

О НЕФОРМАЛИЗУЕМЫХ ЭТАПАХ СИСТЕМНОГО АНАЛИЗА 317

- 9.1. Что такое системный анализ **317**
9.2. Формулирование проблемы **319**
9.3. Выявление целей **324**
9.4. Формирование критериев **328**
9.5. Генерирование альтернатив **334**
9.6. Алгоритмы проведения системного анализа **342**
9.7. Претворение в жизнь результатов системных исследований **345**
9.8. О специфике социальных систем **354**
 Заключение **366**
 Литература **368**
 Упражнения **369**
 Вопросы для самопроверки **371**
-

Краткий словарь специальных терминов **373**

ПРЕДИСЛОВИЕ

В последние годы резко усилились тенденции к фундаментализации обучения в высшей школе, уходу от чрезмерно суженной специализации, расширению кругозора специалистов, развитию междисциплинарных связей, воспитанию системного мышления. Наряду с реализацией этих тенденций в преподавании традиционных предметов, в учебных планах многих вузов появились и специфические “системные” курсы, излагающие методологию исследования и проектирования сложных систем. Настоящая книга отвечает потребности в учебной литературе по таким курсам.

Особенность данного учебника состоит в том, что его можно использовать при обучении студентов разных специальностей, опуская или, наоборот, развивая определенные темы, изложенные в нем. Излагаемый в книге подход скорее всего ближе специалистам, имеющим дело с разработкой математических моделей реальных явлений. Однако авторы на практике убедились, что данный учебник может быть полезен не только для инженеров-системотехников и математиков-прикладников, но и для студентов-гуманитариев. Например, в Томском университете многие разделы книги используются в курсе правовой кибернетики для юристов, в спецкурсах по методам исследования социальных систем на философском и историческом факультетах.

На наш взгляд, в ближайшее время прикладной системный анализ станет объектом профессиональной деятельности. Ряд университетов США уже развернул подготовку специалистов в этой области. Необходимость в этом ощущается и в нашей стране. Это подтверждает опыт изложения идей системного анализа слушателям курсов повышения квалификации, т.е. людям, сталкивавшимся с необходимостью решать проблемы реальной жизни и на собственном опыте прочувствовавшим, насколько это не просто.

В девяти главах книги излагаются основные понятия и методы прикладного системного анализа. Сюда относятся и общие фундаментальные понятия (система, модель, информация и др.), и важные конкретные понятия (сигналы, измерительные шкалы, “большие” и “сложные” системы, эмерджентность и пр.), понятия, специфичные для системного анализа (декомпозиция, агрегирование, конфигуратор, проблематика, “заинтересованные стороны” и др.) К арсеналу используемых в системном анализе методов относятся и строго формализованные (оптимизация, принятие решений, кодирование информации), и направленные на формализацию (экспериментальные исследования, построение моделей), и слабо формализованные (экспертные оценки, коллективный выбор), и в принципе неформализуемые операции (формулирование проблем, выявление целей, определение критериев, генерация альтернатив). Все это объединяется общей методологией, вытекающей из диалектического метода, что и позволяет определить системный анализ и как прикладную диалектику.

Книга содержит обширный и разнообразный вспомогательный материал. На полях приведены некоторые важные выводы, а также наиболее употребительные термины и их английские эквиваленты. Каждый параграф завершается кратким итогом (на русском и английском языках). В конце каждой главы приводятся: заключение (в котором отражены основные результаты данной главы), литература, упражнения и вопросы для самопроверки. В конце книги помещен краткий словарь специальных терминов. Все это, по мнению авторов, должно способствовать лучшему пониманию и усвоению изучаемого материала.

К несчастью, Феликс Иванович Перегудов умер, не увидев второго издания нашей книги. Многие из ее достоинств – его заслуга.

Выражаю искреннюю признательность рецензентам первого издания профессорам Т.Вашко (Лаксенбург, Австрия), А.И.Половинкину (Волгоград), Л.А.Растигину (Рига) и рецензенту второго издания проф. А.А.Золотареву (Москва) за полезные замечания и советы, способствовавшие улучшению книги.

Особую благодарность хотелось бы выразить главному редактору издательства “Высшая школа” А.Д.Суходскому, который во многом улучшил форму представления материала первого издания и существенно помог в подготовке второго издания книги.

Ф.П.Тарасенко

ВВЕДЕНИЕ

Научно-техническая революция привела к возникновению таких понятий, как большие и сложные системы, обладающие специфическими для них проблемами. Необходимость решения этих проблем вызвала к жизни множество приемов, методов, подходов, которые постепенно накапливались, развивались, обобщались, образуя в конце концов определенную технологию преодоления количественных и качественных сложностей. В разных сферах практической деятельности возникали такие ситуации, а соответствующие технологии вместе с их теоретическими основами получали разные названия: в инженерной деятельности – “методы проектирования”, “методы инженерного творчества”, “системотехника”; в военных и экономических вопросах – “исследование операций”; в административном и политическом управлении – “системный подход”, “политология”, “футурология”; в прикладных научных исследованиях – “имитационное моделирование”, “методология эксперимента” и т.д.

С другой стороны, теоретическая мысль на разных уровнях абстракции отражала системность мира вообще и системность человеческого познания и практики: на философском уровне – диалектический материализм; на общенаучном – системология, общая теория систем, теория организации; на естественно-научном – кибернетика; с развитием вычислительной техники возникли информатика и искусственный интеллект.

В начале 80-х годов уже стало очевидным, что все эти теоретические и прикладные дисциплины образуют как бы единый поток, “системное движение”. Системность стала не только теоретической категорией, но и осознанным аспектом практической деятельности. Поскольку большие и сложные системы по необходимости стали предметом изучения, управления и проектирования, потребовалось обобщение методов исследования систем и методов воздействия на них. Должна была возникнуть некая прикладная наука, являющаяся “мостом” между абстрактными теориями системности и живой системной практикой. Она и возникла – сначала, как мы видели, в разных областях и под разными названиями, но в последние годы оформилась в науку, которая получила название “системный анализ”. Хотя системный анализ находится в развитии, сегодня он выступает уже как самостоятельная дисциплина, имеющая свой объект деятельности, накопившая достаточно мощный арсенал средств и обладающая значительным практическим опытом.

Особенности современного системного анализа вытекают из самой природы сложных систем. Имея в качестве цели ликвидацию проблемы или, как минимум, выяснение ее причин, системный анализ привлекает для этого широкий спектр средств, использует возможности различных наук и практических сфер деятельности. Являясь по существу прикладной диалектикой, системный анализ придает большое значение методологическим аспектам любого системного исследования. С другой стороны, прикладная направленность системного анализа приводит к использованию всех современных средств научных исследований – математики, вычислительной техники, моделирования, натурных наблюдений и экспериментов.

В ходе исследования реальной системы обычно приходится сталкиваться с самыми разнообразными проблемами; быть профессионалом в каждой из них невозможно одному человеку. Выход видится в том, чтобы тот, кто берется осуществлять системный анализ, имел образование и опыт, необходимые для опознания и классификации конкретных проблем, для определения того, к каким специалистам следует обратиться для продолжения анализа. Это предъявляет особые требования к специалистам-системщикам: они должны обладать широ-

кой эрудицией, раскованностью мышления, умением привлекать людей к работе, организовывать коллективную деятельность.

Прочтя данную книгу и даже дополнительно изучив нужные разделы глубже, чем это дано в ней, нельзя стать специалистом по системному анализу. Как выразился У. Шекспир: “Если бы делать было бы столь же легко, как знать, что надо делать, – часовни были бы соборами, хижины – дворцами”. Профессионализм приобретается в практике.

Полезно было бы дополнить данную книгу сборником реальных примеров системного анализа. Но и это будут все-таки уже проработанные кем-то ранее ситуации. Самая же сложная и интересная часть системного анализа – это “вытаскивание” проблем из реальной жизни, отделение важного от несущественного, поиск правильной формулировки для каждой из возникающих задач. Авторы надеются, что данная книга будет полезной как введение в системный анализ.

ВОЗНИКОВЕНИЕ И РАЗВИТИЕ СИСТЕМНЫХ ПРЕДСТАВЛЕНИЙ

ГЛАВА ПЕРВАЯ

§ 1.1. ПРЕДВАРИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

В современном обществе системные представления уже достигли такого уровня, что мысль о полезности и важности системного подхода к решению возникающих в практике проблем вышла за рамки специальных научных истин и стала привычной, общепринятой. Уже не только ученые, но и инженеры, педагоги, организаторы производства, деятели культуры обнаружили системность собственной деятельности и стараются осуществлять свою работу осознанно системно. Широко распространилось понимание того, что наши успехи связаны с тем, насколько системно мы подходим к решению проблем, а наши неудачи вызваны отступлениями от системности.

Было бы неверным считать, что “мышление стало системным” только во второй половине XX века. Мышление системно всегда и другим быть не может. Системность – это не такое качество, которым можно обладать или не обладать. Однако системность имеет разные уровни. Сигналом о недостаточной системности существующей деятельности является появление проблемы; разрешение возникшей проблемы осуществляется путем перехода на новый, более высокий уровень системности в нашей деятельности. Поэтому системность не столько состояние, сколько процесс.

Иллюстрацией этого может служить состояние знаний тех, кому только что сказанное представляется расплывчатым, не очень ясным: что означает само слово “система”, что означает “действовать системно”, почему “не системного” знания не бывает? Налицо проблема понимания сказанного (кстати, совсем не уникальная, а типичная для обучения). Эту проблему мы будем решать, постепенно повышая уровень системности знаний, в чем и состоит цель данной книги. Пока же нам достаточно тех ассоциаций, которые возникают, когда мы употребляем в обыденной речи слово “система” в сочетании со словами “общественно-политическая”, “Солнечная”, “нервная”, “отопительная” или “уравнений”, “показателей”, “взглядов и убеждений”, – ведь эти словосочетания обозначают (наряду с различным) и нечто общее: системность.

В данной главе мы покажем, что хотя осознание системности мира пришло с трудом и не сразу, оно не могло не прийти: системные представления возникли по объективным причинам и развиваются под действием объективных факторов.

§ 1.2. РОЛЬ СИСТЕМНЫХ ПРЕДСТАВЛЕНИЙ В ПРАКТИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ

Человек – активная часть природы. Добиваясь своих целей, человек использует природу, воздействует на нее, преобразует ее и себя. Без преувеличения можно сказать, что самым важным и самым интересным для человечества кругом вопросов являются вопросы о возможностях человека в его отношениях с природой, о способах реализации этих возможностей, о факторах, способствующих и препятствующих расширению этих возможностей. К этому же кругу вопросов принадлежит и философская проблема соотношения материи и сознания.

Начнем с рассмотрения практической деятельности человека, т.е. его активного и целенаправленного воздействия на окружающую среду. Наша первая задача – показать, что человеческая практика системна. Впоследствии мы будем подробно и всесторонне рассматривать признаки системности, а сейчас отметим только самые очевидные и обязательные из них: *структурированность системы, взаимосвязанность составляющих ее частей, подчиненность организации всей системы определенной цели*. По отношению к человеческой деятельности эти признаки и в самом деле очевидны, поскольку каждый из нас легко обнаружит их в своем собственном практическом опыте. Всякое наше осознанное действие преследует определенную цель (пока оставим в стороне неосознанные действия). Во всяком действии легко увидеть его составные части, более мелкие действия. При этом легко убедиться, что эти составные части должны выполняться не в произвольном порядке, а в определенной их

AUTOMATION	последовательности. Это и есть та самая определенная, подчиненная
<i>автоматизация</i>	цели взаимосвязанность составных частей, которая и является признаком системности.
ALGORITHM алгоритм	СИСТЕМНОСТЬ И АЛГОРИТМИЧНОСТЬ
MECHANIZATION механизация	Другое название для такого построения деятельности –
PRACTICE практика	алгоритмичность . Понятие алгоритма возникло сначала в математике
PRODUCTIVITY производительность	и означало задание точно определенной последовательности однозначно понимаемых операций над числами или другими математическими объектами. В последние годы стала осознаваться алгоритмичность любой деятельности, и уже всерьез говорят не только об алгоритмах принятия управлеченческих решений, об алгоритмах обучения, алгоритмах игры в шахматы, но и об алгоритмах изобретательства [2], алгоритмах композиции музыки [8]. Подчеркнем, что при этом делается отход от математического понимания алгоритма: сохраняя логическую принудительность последовательности действий, мы допускаем, что в алгоритме данной деятельности могут присутствовать и такие действия, которые не формализованы; важно лишь, чтобы этот этап алгоритма успешно выполнялся человеком, хотя и не осознанно. Например, Р.Х. Зарипов отмечает:

“...подавляющее большинство элементов творческой деятельности, реализуемых человеком “легко и просто”, “не думая”, “по интуиции”, на самом деле являются неосознанной реализацией определенных алгоритмизируемых закономерностей, реализацией неосознаваемых, но объективно существующих и формализуемых критериев красоты и вкуса” [8, с.12].

Здесь важными являются следующие моменты. Во-первых, всякая деятельность алгоритмична. Во-вторых, не всегда алгоритм реальной деятельности осознается (композитор сочиняет музыку, шофер мгновенно реагирует на изменения дорожной обстановки, вратарь ловит в броске мяч – “не думая”). В-третьих, в случае неудовлетворенности результатом деятельности возможную причину неудачи следует искать в несовершенстве алгоритма. Это означает – пытаться выявить алгоритм, исследовать его, искать “слабые места”, устранять их, т.е. совершенствовать алгоритм и, следовательно, повышать системность деятельности.

Успешность всякой деятельности тем более вероятна, чем выше уровень ее системности; неудачи вызваны недостаточной системностью.

Можно выделить три уровня системности:

Таким образом, явная алгоритмизация любой практической деятельности является важным средством ее развития.

Перейдем теперь к другой задаче – показать, что роль системных представлений в практике постоянно увеличивается, что растет сама системность человеческой практики. Последнее можно проиллюстрировать многими примерами, но поучительно сделать это на несколько схематизированном примере проблемы повышения производительности труда.

Одна из важнейших особенностей общественного производства состоит в непрерывном росте его эффективности, и прежде всего в повышении производительности труда. Подчеркнем, что этот процесс роста очень сложен, многогранен, но его итог выражается, овеществляется в развитии средств труда и методов его организации.

МЕХАНИЗАЦИЯ

Простейший и исторически первый способ повышения эффективности труда – *механизация*. Человек вооружается **механизмами** – от простейших орудий и приспособлений, приводимых в действие мускульной силой, до сложнейших машин со встроенными в них двигателями. С помощью механизмов и машин один человек выполняет физическую работу, которую без них пришлось бы выполнять многим людям.

Механизация позволяет решать многие проблемы. Например, по подсчетам акад. А.И. Берга, если бы механизация строительных работ у нас оставалась на уровне строительства ДнепроГЭСа, то для сооружения только электростанций уже в прошлом десятилетии потребовалось бы все трудоспособное население Советского Союза. В строительстве электростанций этого не произошло – благодаря механизации. Но в целом в народном хозяйстве страны еще весьма высок процент работ, выполняемых вручную: в промышленности – несколько меньше половины, в сельском хозяйстве и в сфере обслуживания – значительно выше. Таким образом, возможности механизации еще далеко не исчерпаны.

Однако механизация имеет естественный предел: работой механизмов управляет человек, а его возможности ограничены физиологически. Нельзя делать лопату слишком широкой – поднимать ее придется человеку. Машина не должна иметь слишком много приборов-индикаторов и рычагов управления: у человека всего два глаза и две руки. Скорость реакции человека ограничена, поэтому механизация очень быстрых процессов бессмысленна. Короче говоря, сам человек является “узким местом” механизации.

АВТОМАТИЗАЦИЯ

Решение проблемы состоит в том, чтобы вообще исключить участие человека из конкретного производственного процесса, т.е. возложить на машины не только выполнение самой работы, но и выполнение операций по регулированию хода, течения процесса работы. Технические устройства, объединившие эти две функции, называются **автоматами**.

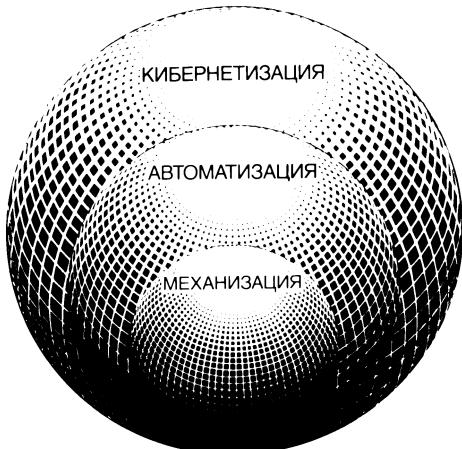
В соответствии с этим второй способ повышения производительности труда (он же второй этап по времени и второй уровень системности общественного производства) получил название *автоматизации*.

Автоматы полностью освобождают человека от выполнения данной работы*. Они могут иметь разную сложность и выполнять разнообразные работы. В быт вошли торговые и игровые автоматы, автоматическая телефонная связь, в промышленности уже существуют целые автоматические линии, цехи и заводы, развивается промышленная и транспортная робототехника. Расширенные возможности представляют перестраиваемые, многофункциональные автоматы, среди которых особое место занимают ЭВМ. Автоматизации поддаются все более сложные работы, в том числе такие, которые прежде выполнялись только в виде мыслительной деятельности. Это опять-таки вызывается объективной необходимостью. А.И. Берг приводил данные, что если бы плановые, экономические и финансовые органы нашей страны обрабатывали всю информацию по-старому, на счетах и арифмометрах, то сейчас все трудоспособное население страны должно было бы работать “в бухгалтериях”. Этого не произошло – благодаря автоматизации с помощью ЭВМ.

Однако очень важно понять, что автоматизировать, т.е. полностью возложить на машину, можно только те работы, которые детально изучены, подробно и полно описаны, в которых точно известно, что, в каком порядке и как надо делать в каждом случае, и точно известны все возможные случаи и обстоятельства, в которых может оказаться автомат. Только при таких условиях можно сконструировать соответствующий автомат, и только в этих условиях он может успешно выполнять работу, для которой предназначен. Пользуясь уже знакомой нам терминологией, можно сказать, что автомат реализует некоторый алгоритм (в математическом смысле этого слова), и если алгоритм в какой-то своей части неправилен или неточен либо встретилась ситуация, не предусмотренная алгоритмом, то поведение автомата не может соответствовать целям его создания.

Итак, автоматизация является мощным средством повышения производительности труда: по мере совершенствования наших знаний о тех или иных производственных процессах последние могут быть автоматизированы во все большей степени. Однако и у автоматизации в свою очередь существует естественный предел: в реальной жизни часто приходится сталки-

*Правда при этом человеку приходится выполнять другую работу – по созданию и обслуживанию автоматов, но это труд совсем другого уровня и квалификации, и, кстати, также в принципе автоматизируемый.



1.1 ————— Схема этапов повышения производительности труда

вательском) в жизнедеятельность человеческого организма, при воздействии человека на природу, т.е. в тех случаях, когда приходится взаимодействовать со сложными системами*. Повышение эффективности такого взаимодействия является как объективной, так и субъективной необходимостью, и, естественно, человечество вырабатывает способы решения возникающих при этом проблем.

Совокупность таких способов представляет собой третий уровень системности практической деятельности человека. Этот уровень можно назвать *кибернетизацией*, поскольку кибернетика первой среди других подходов стала претендовать на научное решение проблем управления сложными системами (хотя она оправдала не все ожидания). Соотношение между тремя рассмотренными уровнями организации труда иллюстрирует рис.1.1.

Основная идея разрешения проблем, связанных со сложными системами, состоит в том, чтобы в тех случаях, когда автоматизация (т.е. формальная алгоритмизация) невозможна, использовать ту человеческую способность, которая именно в таких случаях проявляется и которая называется *интеллектом*: способность ориентироваться в незнакомых условиях и находить решение слабо формализованных задач (сказочный герой мог решать даже такую задачу: “Пойди туда, не знаю куда, принеси то, не знаю что”). При этом человек выполняет именно те операции в общем алгоритме, которые не поддаются формализации (например, экспертная оценка или сравнение многомерных и неколичественных вариантов, принятие управлеченческих решений, взятие на себя ответственности). Именно на этом принципе строятся автоматизированные (в отличие от автоматических) системы управления, в которых формализованные операции выполняют автоматы и ЭВМ, а неформализованные (и, возможно, неформализуемые) операции – человек. Этот путь, следовательно, состоит в разумном использовании естественного человеческого интеллекта.

Однако на этом возможности кибернетизации не кончаются, а, скорее, лишь начинаются. Вполне логично, в духе лучших научных традиций, возникает вопрос: нельзя ли смоделировать интеллектуальные возможности человека – хотя бы в той части, которая необходима для выполнения конкретных, пусть частных, интеллектуальных операций? Здесь опыт науки и техники подсказывает два пути: “подглядеть” у природы алгоритмы интеллектуальной деятельности (т.е. изучать естественный интеллект) либо “изобрести” эвристически алгоритм

ваться с непредвиденными условиями и с невозможностью полной формализации многих практических действий.

КИБЕРНЕТИЗАЦИЯ

Наиболее остро такие проблемы возникают в процессе руководства человеческими коллективами, при управлении производственными системами, при проектировании и эксплуатации крупных технических комплексов, при вмешательстве (например, медицинском или исследо-

* В гл. 4 термину “сложная система” будет придан более точный смысл, а сейчас нам достаточно того (пусть расплывчатого) общеупотребительного смысла, который это понятие имеет в естественном языке. Впрочем, одну важную сторону сложности мы уже отметили: неформализуемость ряда процессов, происходящих в системе и непредвиденность некоторых внешних условий.

предположительно с интеллектуальными свойствами и исследовать, что это даст (т.е. конструировать интеллект искусственно).

По ходу изложения мы еще не раз будем возвращаться к проблемам и результатам кибернетизации. Пока же отметим, что техническими средствами для этого служат автоматы и ЭВМ. Их совершенствование и использование для самых разнообразных нужд, создание алгоритмов и программных систем для ЭВМ приобрели исключительно важное значение. Настолько важное, что все связанное с использованием ЭВМ стали выделять в отдельную область науки и техники, названную *информатикой*.

Но не менее важная сторона процесса кибернетизации связана с изучением самих систем, созданием методов организации и управления сложными системами, развитием системных представлений, системных теорий.

Подведем итог

Summary

Природная системность человеческой практики является одним из объективных факторов возникновения и развития системных понятий и теорий. Естественный, спонтанный рост системности человеческой деятельности сопровождается, дополняется осознанием этой системы, а затем ускоряется сознательным повышением. Роль знания и соблюдения принципов системности в практике возрастает. Алгоритмизация любого вида деятельности – важный способ повышения ее системности.

The natural and spontaneous increase in the systematicity of human activity, combined with and supported by the recognition of this systematicity, has been an objective factor in the appearance and development of systems principles in human endeavor is increasing. Algorithmization is an important means of raising the level of systematicity of any given activity.

§ 1.3. ВНУТРЕННЯЯ СИСТЕМНОСТЬ ПОЗНАВАТЕЛЬНЫХ ПРОЦЕССОВ

Рассматривая объективные причины возникновения и факторы развития системных представлений и соответствующих методов и теорий, мы должны отметить объективные особенности человеческого мышления. Покажем, что сам процесс познания системен и что знания, добытые человечеством, также системны.

Современные представления об окружающем нас мире позволяют говорить о его бесконечности в пространстве и времени, о возможности неограниченного расширения и углубления наших знаний о любом объекте, сколь бы велик или мал он ни был. Человек существует конечное время и располагает конечными материальными, энергетическими и информационными ресурсами. И все же человечеству удается познавать мир и, как показывает его практика, познавать верно. А. Эйнштейн отмечал, что самое удивительное в природе то, что она познаваема.

Противоречия между неограниченностью желаний человека познать мир и ограниченностью существующих возможностей сделать это, между бесконечностью природы и конечностью ресурсов человечества имеют много важных последствий, в том числе и в самом процессе познания человеком окружающего мира.

АНАЛИЗ И СИНТЕЗ В ПОЗНАНИИ

Одна из таких особенностей познания, которые позволяют постепенно, поэтапно разрешать эти противоречия, – наличие аналитического и синтетического образов мышления. Суть анализа состоит в разделении целого на части, в представлении сложного в виде сово-

ANALYSIS анализ
MATTER материя
THINKING,
THOUGHT мышление
COGNITION познание
SYNTHESIS синтез

купности более простых компонент. Но чтобы познать целое, сложное, необходим и обратный процесс – *синтез*. Это относится не только к индивидуальному мышлению, но и к общечеловеческому знанию.

Аналитичность человеческого знания находит свое отражение в существовании различных наук, в продолжающейся дифференциации наук, во все более глубоком изучении все более узких вопросов, каждый из которых сам по себе тем не менее интересен, важен и необходим. Вместе с тем столь же необходим и обратный процесс синтеза знаний. Так возникают “пограничные” науки типа биохимии, физикохимии, биофизики или бионики. Однако это лишь одна из форм синтеза. Другая, более высокая форма синтетических знаний реализуется в виде наук о самых общих свойствах природы. Философия выявляет и отображает все (любые) общие свойства всех форм материи; математика изучает некоторые, но также всеобщие, отношения. К числу синтетических относятся и системные науки: кибернетика, теория систем, теория организации и др. В них необходимым образом соединяются технические, естественные и гуманитарные знания.

Итак, расчлененность мышления (на анализ и синтез) и взаимосвязанность этих частей являются очевидными признаками системности познания.

Системность мышления связана с системностью мира. В то же время мышление обладает самостоятельностью и большей свободой: в воображении можно построить образы нереализуемых в практике конструкций.

Для последующего этапа – метафизического способа мышления – характерно преобладание анализа: “Разложение природы на ее отдельные части, разделение различных процессов и предметов природы на определенные классы, исследование внутреннего строения органических тел по их многообразным анатомическим формам – все это было основным условием тех исполнительских успехов, которые были достигнуты в области познания природы за последние четыреста лет. Но тот же способ оставил нам вместе с тем и привычку рассматривать вещи и процессы природы в их обособленности, вне их великой связи, и в силу этого – не в движении, а в неподвижном состоянии, не как существенно изменчивые, а как вечно неизменные, не живыми, а мертвыми” [11, с. 20-21].

Сами метафизики ощущали незавершенность аналитического знания, видели возможность и даже необходимость синтеза, но считали синтез произвольной деятельностью ума, не имеющей отношения к природе. Так, Ф. Бэкон писал:

“Человеческий разум в силу своей склонности легко предполагает в вещах больше порядка и единства, чем их находит. И в то время, как многое в природе единично и совершенно не имеет себе подобия, он придумывает параллели, соответствия и отношения, которых нет” [4, с. 20].

Р. Декарт утверждал, что разум должен предполагать “порядок даже и там, где объекты мышления вовсе не даны в их естественной связи” [7, с. 272].

Новый, более высокий уровень системности познания представляет собой диалектический способ мышления.

ЭВОЛЮЦИЯ ВЗГЛЯДОВ НА СИСТЕМНОСТЬ МЫШЛЕНИЯ

Интересно отметить, что осознание диалектического единства анализа и синтеза пришло не сразу, и в разные исторические эпохи система мышления имела различный характер [1]. Ф. Энгельс отмечал, что в древней Греции преобладало нерасчлененное знание: “...природа еще рассматривается в общем, как одно целое. Всеобщая связь явлений природы не доказывается в подробностях: она является для греков результатом непосредственного созерцания” [11, с. 369].

Для последующего этапа – метафизического способа мышления – характерно преобладание анализа: “Разложение природы на ее отдельные части, разделение различных процессов и предметов природы на определенные классы, исследование внутреннего строения органических тел по их многообразным анатомическим формам – все это было основным условием тех исполнительских

успехов, которые были достигнуты в области познания природы за последние четыреста лет. Но тот же способ оставил нам вместе с тем и привычку рассматривать вещи и процессы природы в их обособленности, вне их великой связи, и в силу этого – не в движении, а в неподвижном состоянии, не как существенно изменчивые, а как вечно неизменные, не живыми, а мертвыми” [11, с. 20-21].

Сами метафизики ощущали незавершенность аналитического знания, видели возможность и даже необходимость синтеза, но считали синтез произвольной деятельностью ума, не имеющей отношения к природе. Так, Ф. Бэкон писал:

“Человеческий разум в силу своей склонности легко предполагает в вещах больше порядка и единства, чем их находит. И в то время, как многое в природе единично и совершенно не имеет себе подобия, он придумывает параллели, соответствия и отношения, которых нет” [4, с. 20].

Р. Декарт утверждал, что разум должен предполагать “порядок даже и там, где объекты мышления вовсе не даны в их естественной связи” [7, с. 272].

Новый, более высокий уровень системности познания представляет собой диалектический способ мышления.

В развитие диалектики значительный вклад внесли представители немецкой классической философии – И.Кант, И.Фихте, Ф. Шеллинг.

У Канта имеются прямые суждения о системности самого мышления и его развития: “Достигаемое разумом единство есть единство системы” [9, с. 577].

“Системы [знаний] кажутся, подобно червям, возникающими путем generatio aequovoca* из простого скопления собранных вместе понятий, сначала в изуродованной, но с течением времени в совершенно развитой форме” [9, с. 681 – 682].

Но Кант считал систему только идеальной:

“Под системой же я разумею единство многообразных знаний, объединенных одной идеей” [9, с. 680].

Исследование процессов развития познания, соотношения познания и реальности привело Гегеля к окончательному формированию диалектики. Материалистическая интерпретация диалектики, построенная Марксом, дополнила общую картину до современного понимания системности мышления. При этом диалектический подход состоит в том, чтобы не только противопоставлять материалистическую диалектику идеалистической, но и видеть их взаимодействие. Если рассматривать диалектику как методологию согласования системности мышления с системностью природы, то такое согласование возможно вести в двух направлениях, взяв за начальную точку либо материю, либо сознание. Ясно, что полную картину мы будем иметь, лишь рассмотрев весь замкнутый цикл такого согласования. Истина состоит не только в том, что бытие определяет сознание, но и в том, что сознание определяет бытие.

Подведем итог

Одной из важнейших объективных причин возникновения системных наук является системность самого человеческого мышления*. В данном параграфе основное внимание было уделено системности процесса логического познания, что прежде всего проявляется в его структурированности, в частности в выделении подпроцессов анализа и синтеза. Системен и сам результат познания, т.е. сами полученные знания, само их представление; однако этот вопрос столь важен для данного курса, что ему будет посвящена отдельная глава (см. гл. 2).

§ 1.4. СИСТЕМНОСТЬ КАК ВСЕОБЩЕЕ СВОЙСТВО МАТЕРИИ

Обсуждая объективные причины развития системных представлений, мы убедились, что к важнейшим из этих причин относятся системность человеческой практической деятельности (см. § 1.2) и внутренняя системность человеческого мышления (см. § 1.3). Возникает вопрос не является ли системность специфическим свойством человека, своего рода приспособлением, выработанным им для собственного удобства, облегчения, упрощения своей деятельности.

Summary

One of the most significant objective factors in the development of systems sciences is the systematicity of human thought itself.* This section focuses on the systems nature of logical reasoning, which is apparent in its organizational structure, in particular, its division into component processes of analysis and synthesis. The products of cognition – i.e., knowledge and its presentation – are also systemic in nature, but this issue is sufficiently important that it will be discussed separately in Chapter Two.

* Отметим, что здесь мы делаем акцент на логическом, сознательном мышлении и не касаемся других важных аспектов интеллекта – роли социальных, эмоциональных, подсознательных, физиологических и других компонент в человеческом мышлении.

сти в окружающем мире, а сам мир не только безразличен к тому, кто и как его познает, и познает ли вообще, но может и не иметь ничего общего с нашими представлениями о нем?

Поставленный вопрос порождает множество переплетенных между собой проблем. Можно (как это было сделано в первом издании данной книги) рассмотреть философские аспекты познавательного процесса, сопоставить мнения философов различных направлений. Несмотря на различия между взглядами разных авторов на процессы познания и практики, почти все они признавали существование связи между мышлением и реальностью (мышление как отражение реальности, мышление как один из природных процессов),
“ввиду чего мы должны предполагать систематическое единство природы как объективно значимое и необходимое” [9, с 588].

К выводам, аналогичным приведенному высказыванию И. Канта, приходили многие философы. Системность природы не только логически выводится в рамках теоретических построений, но и практически проявляется в реально наблюдаемых явлениях, как с участием человека, так и без него.

ВСЯ ПРИРОДА СИСТЕМНА

Ныне системность понимается не только как свойство человеческой практики (включающей и внешнюю активную деятельность, и мышление, и даже пассивное созерцание), но и как свойство всей материи. Системность нашего мышления вытекает из системности мира. Современные представления позволяют говорить о мире как о бесконечной иерархической системе систем, находящихся в развитии, на разных стадиях развития, на разных уровнях системной иерархии, взаимодействующих друг с другом. Все новые и новые естественнонаучные данные подкрепляют такие представления. Например, И. Пригожин недавно построил теорию, успешно объясняющую процессы самоорганизации в природе. Оказалось, что неоспоримые факты спонтанного возникновения организованных систем из хаоса вполне объяснимы внутренними свойствами самой материи [13].

СИСТЕМЫ КАК АБСТРАКЦИЯ

Присоединяясь к точке зрения на системность как на всеобщее, универсальное свойство всей природы, должны отметить, что существует и другая точка зрения. Некоторые авторы считают, что возможность по-разному выделять систему из цельного мира означает, что системы – лишь наши представления о мире. Например, в монографии Дж. Клира приводится такое определение системы:

“Система – это то, что различается как система... Системой является все то, что мы хотим рассматривать как систему” [10, с. 24].

На первый взгляд, такой подход снимает многие вопросы, которые нам приходится подробно обсуждать в данной главе. Во-первых, вместе с понятием естественных систем исчезают трудности объяснения целесообразности и самоорганизации в природе. Во-вторых, исследование систем сводится к исследованию построенных моделей, и единственным экспериментальным методом системологии становятся машинные (вычислительные) эксперименты.

Допустимый для специалистов, занимающихся разработкой высокосложных компьютерных технологий, такой подход совершенно неприемлем для специалистов по системному анализу, объектом исследований которых являются реальные жизненные ситуации, требующие практического (и улучшающего!) вмешательства. В системном анализе на первый план выступает вопрос об адекватности моделей, об их соответствии реальности, а это неизбежно ведет от представлений о системности наших знаний о мире к представлениям о системности самого мира.

И все-таки в приведенных выше высказываниях Дж. Клира есть рациональное зерно. Возможность по-разному расчленять реальный

* Самозарождения (лат). В те времена считалось, что черви зарождаются в питательной среде самопроизвольно, без участия других живых клеток.

мир на относительно отдельные части, каждую из которых можно рассматривать как систему, и создает впечатление, будто системой является все то, что мы хотим считать системой. Еще дальше идущей абстракцией является понятие множества как произвольно формулируемой совокупности.

Примененное к реальным объектам, это понятие совпадает с понятием системы (система всех рек, система всех элементарных частиц и т.п.). Однако из-за реальной всеобщей взаимосвязанности не бессмысленно говорить о свойствах множества всех рек региона, континента, Земли и практически изучать эти свойства. Таким образом, дело оказывается не в том, что мы хотим рассматривать некоторое множество реальных объектов как систему, а в том, что оно действительно образует систему со своими, специфическими для нее свойствами.

СВОЙСТВА ЛЮБЫХ СИСТЕМ

Однако, как бы ни понималась системность (как свойство разума, позволяющее познать мир, или как свойство самой природы), все исследователи этого феномена сходятся во мнениях относительно характеристических признаков любой системы независимо от ее происхождения. Не претендуя на законченность списка таких особенностей, приведем перечень важнейших, необходимых признаков того, что данный объект, явление, процесс или их модель имеют системный характер.

1. Всякая система обладает целостностью, обособленностью от окружающей ее среды, выступает как нечто отдельное, единое. (Примеры: рыба в воде, море и окружающая его суши, наука в культуре, солнечная система в галактике, геометрия в математике и т.д.).

2. Обособленность, выделенность системы в среде не означает ее изолированности от среды: система связана со средой, существует в ней, взаимодействует с ней, обменивается со средой энергией, материей, информацией (в разных пропорциях, в зависимости от природы системы). Иными словами – все системы открыты; замкнутых (т.е. изолированных от среды) систем не бывает. Можно вообразить замкнутую систему, но проверить, доказать ее реальность невозможно – ведь с ней нельзя взаимодействовать, т.е. нет опыта, в котором проявилось бы ее существование.

3. Цельность системы не означает ее однородности и неделимости: наоборот, в системе можно различать определенные составные части.

4. Разделимость системы на части не означает, что эти части полностью изолированы друг от друга. Наоборот, части образуют целое благодаря связям между ними. Открытость системы означает, что ее части связаны и с внешней средой, но цельность системы основана на том, что внутренние связи частей, образующие структуру системы, в каком-то отношении сильнее, существеннее, важнее, чем их внешние связи.

5. Целостность системы обусловлена тем, что система как целое обладает такими свойствами, которых нет и не может быть у составляющих ее частей. Свойства системы не сводятся к свойствам ее частей, не являются простой совокупностью этих свойств. Система и существует, и выделяется, и описывается как носитель этих качественно новых свойств. (Возникновение принципиально нового качества, не существующего без объединения частей в систему, называется *эмерджентностью*.) Понятие эмерджентности проясняет разницу между внешними и внутренними связями системы: свойство системы как целого проявляется в ее взаимодействии с окружающей средой (т.е. реализуется через внешние связи как *функция системы*), но само это свойство возникает и может существовать лишь благодаря взаимодействию частей (т.е. благодаря внутренним связям, т.е. благодаря *структуре системы*).

6. Понятие эмерджентности позволяет подчеркнуть еще один аспект внутренней целостности системы. Изъятие части из системы приводит к тому, что система при этом теряет какие-то существенные свойства, т.е. становится другой системой. Более того, часть, изъятая из системы, также теряет свои существенные свойства, которые могли реализовываться лишь до тех пор, пока эта часть находилась в системе. Поэтому основа холистического (целостно-

го) подхода состоит в недопустимости рассмотрения частей системы по отдельности, вне их взаимодействия с другими частями.

7. Открытость системы, ее связанность со средой означает, что она (система) в свою очередь входит в какую-то большую систему, является частью в этой большей системе. В результате мир выглядит (существует!) как иерархическая система вложенных друг в друга, перекрывающихся частично или полностью, или разделенных, но взаимодействующих систем.

8. Внутренняя и внешняя целостность систем обобщаются, объединяются, синтезируются в понятии *цели*, которая как бы диктует и структуру, и функцию системы. Функция системы интерпретируется как проявление целесустримленности системы; структура системы выступает при этом как вариант реализации цели. В связи с этим рассмотрение целей системы становится одной из центральных проблем системологии. В частности, проводится различие между субъективными и объективными целями (и, соответственно, между искусственными и естественными системами).

9. Системы не являются застывшими, неизменными образованиями. Наоборот, в результате внешних и внутренних взаимодействий, все системы находятся в динамике, подвержены постоянным изменениям, происходящим с разной интенсивностью. Многообразие процессов, происходящих с системами, велико; их классификация проводится по разным основаниям (развитие – рост – равновесие – убыль – деградация; цикличность – непериодичность; детерминированность – случайность; рождение – жизнь – смерть; и т.д.). Многие явления в системах невозможно понять без учета их динамики.

В последующих главах мы подробно детализируем и обоснуем эти тезисы, а также постараемся показать, как на их основе строятся методология исследования и проектирования систем, методы целевого (успешного, правильного, системного) воздействия на системы.

Подведем итог

Системы не только человеческая практика и мышление, но и сама природа, вся Вселенная. Системность является настолько присущим и всеобщим свойством материи, что его можно назвать формой существования материи. Известные формы существования материи – время, пространство, движение, структурированность – представляют собой частные проявления, аспекты системности мира.

Summary

Human activity and thought display properties of systems, but so do Nature and the entire Universe. Systematicity is so inherent and general a property of matter that it can be called a form of existence of matter. The other forms of material existence known to us – time, space, movement, and organization – are only particular phenomena or aspects of world systematicity.

§ 1.5. КРАТКИЙ ОЧЕРК ИСТОРИИ РАЗВИТИЯ СИСТЕМНЫХ ПРЕДСТАВЛЕНИЙ

В предыдущих параграфах мы отмечали, что для возникновения системных понятий и теорий имеются непреходящие объективные причины, и поэтому такие теории не могли не возникнуть, а возникнув – не могут не развиваться, и существующее состояние является лишь очередным этапом этого развития, которое будет продолжаться. Однако этот процесс происходит негладко, в нем возможны не только рывки вперед, но и остановки и даже отступления назад. Осознание системности мира и модельности мышления всегда отставало от эмпирической системности человеческой практики. Хотя всегда существует возможность придать слишком большое значение событиям настоящего времени, но, пожалуй, не будет преувеличением сказать, что сейчас в процессе развития системных представлений происходит нечто

качественно новое: системное мышление и системная методология становятся массовыми и в этом качестве обращаются в “материальную силу”, повышая уровень человеческой практики.

История развития системных представлений первоначально шла по нескольким отдельным направлениям. С разных исходных позиций приближались к современному пониманию системности философская мысль и конкретно-практическая научная и техническая методология. В своем движении к единой, объективной истине они неминуемо должны были сойтись, сопоставить результаты, понять общность и различия; свидетелями и участниками этого этапа синтеза научного знания мы и являемся.

Интересно отметить, что философия примерно на сто лет раньше вышла “в район встречи”, на высшую позицию в осознании системности материи, сознания и их отношения. Итогом обобщений, развития и борьбы в философии стала материалистическая диалектика. Результаты философии относятся к множеству всех существующих и мыслимых систем, носят всеобщий характер. Их применение к рассмотрению конкретных ситуаций означает переход к суженному множеству систем, учет его особенностей в проявлении общих закономерностей (дедуктивный метод).

СИСТЕМНОСТЬ КАК ОБЪЕКТ ИССЛЕДОВАНИЯ

Конкретные науки большей частью придерживаются противоположного, индуктивного метода – от исследования реальных систем к установлению общих закономерностей. Это относится и к собственно системным закономерностям. Например, Эддингтон подчеркивал, что “первая” физика, изучающая свойства отдельных частиц и тел, это лишь первые шаги в изучении природы, что принципиально новыми свойствами обладают взаимодействующие совокупности, ансамбли атомов и молекул, являющиеся объектами изучения “второй” физики. Законы системы, отмечал Эддингтон, невыводимы только из первичных закономерностей.

Интересно и полезно проследить, как именно шло развитие системных представлений в той или иной конкретной науке; конечно, это нашло (и не могло не найти) отражение в исследованиях по истории каждой естественной и гуманитарной науки. Несмотря на это, специальное изучение этапов повышения системности, с намеренным интересом именно к системным проблемам данной науки, остается актуальной задачей.

В свете современных представлений системность всегда, осознанно или неосознанно, была методом любой науки; любой ученый прошлого, и не помышлявший о системах и моделях, именно с ними и имел дело. Как уже отмечалось, быстрее всего была осознана системность самого человеческого познания. Философия, логика, основания математики – области, в которых споры по системным проблемам уходят в глубь веков. Однако для нас особый интерес представляют те моменты в истории, когда системность привлекала внимание как объект исследования для естественных и технических наук.

ПЕРВЫЕ ШАГИ КИБЕРНЕТИКИ

Первым в явной форме вопрос о научном подходе к управлению сложными системами поставил М.-А. Ампер. При построении классификации всевозможных, в том числе и несуществовавших тогда, наук (“Опыт о философии наук, или аналитическое изложение классификации всех человеческих знаний”, ч. I – 1834 г., ч. II – 1843 г.) он выделил специальную науку об управлении государством и назвал ее кибернетикой. При этом он не только обозначил необходимое место для кибернетики в ряду других наук, но и подчеркнул основные ее системные особенности:

“Беспрестанно правительству приходится выбирать среди различных мер ту, которая более всего пригодна к достижению цели (...) и лишь благодаря углубленному и сравнительному изучению различных элементов, доставляемых ему для этого выбора, знанием всего того, что касается управляемого им народа, – характера, воззрений, истории, религии, средств существования и процветания, организаций и законов, – может оно составить себе общие правила поведения, руководящие им в каждом конкретном случае. Эту науку я называю кибернети-

кой от слова κυβερνητικη, обозначавшего сперва, в узком смысле, искусство управления кораблем, а затем постепенно получившего у самих греков гораздо более широкое значение искусства управления вообще”.

Ампер только еще пришел к выводу о необходимости кибернетики, а Б. Трентовский, польский философ-гегельянец, уже читал во Фрейбургском университете курс лекций, содержание которого опубликовал на польском языке в 1843 г. Его книга называлась “Отношение философии к кибернетике как искусству управления народом”*. Трентовскийставил целью построение научных основ практической деятельности руководителя (“кибернета”): “Применение искусства управления без сколько-нибудь серьезного изучения соответствующей теории подобно врачеванию без сколько-нибудь глубокого понимания медицинской науки”.

Он подчеркивал, что действительно эффективное управление должно учитывать все важнейшие внешние и внутренние факторы, влияющие на объект управления: “При одной и той же политической идеологии кибернет должен управлять различно в Австрии, России или Пруссии. Точно так же и в одной и той же стране он должен управлять завтра иначе, чем сегодня”.

Главная сложность в управлении, по Трентовскому, связана со сложностью поведения людей:

“Люди не математические символы и не логические категории, и процесс управления – это не шахматная партия. Недостаточное знание целей и стремлений людей может опрокинуть любое логическое построение. Людьми очень трудно командовать и предписывать им наперед заданные действия. Приказ, если кибернет вынужден его отдавать, всегда должен четко формулироваться. Исполняющему всегда должен быть понятен смысл приказа, его цели, результат, который будет достигнут, и кара, которая может последовать за его невыполнением, – последнее обязательно”.

С позиций диалектики Трентовский понимал, что общество, коллектив, да и сам человек – это система, единство противоречий, разрешение которых и есть развитие. Поэтому кибернет должен уметь, исходя из общего блага, одни противоречия примирять, другие – обострять, направляя развитие событий к нужной цели:

“Короче говоря, кибернет не проектирует будущее, как старается сделать некий радикальный философ, – он позволяет будущему рождаться своим собственным независимым способом. Он оказывает будущему помочь как опытный и квалифицированный политический акушер”.

Даже по этим нескольким кратким высказываниям из книги Трентовского можно видеть, насколько далеко ему удалось продвинуться в понимании необходимости алгоритмизации человеческой деятельности в осознании системности человеческих коллективов, групп, формальных и неформальных образований, в понимании сложности управления людьми. Можно согласиться с оценкой Н.Н. Моисеева:

“Я думаю, что его книга – одно из удачных изложений методологических принципов управления в домарксистский период. Это веха, показывающая становление кибернетики как общей науки об управлении, о каркасе, как говорил Б. Трентовский, через который отдельные науки могут соединиться и взаимодействовать для достижения общих целей” [11, с. 20].

И все же общество середины прошлого века оказалось не готовым воспринять идеи кибернетики. Практика управления еще могла обходиться без науки управления. Кибернетика родилась слишком рано и была позабыта.

Прошло около полувека, и системная проблематика снова появилась в поле зрения науки. На этот раз внимание было сосредоточено на вопросах структуры и организации систем. Поразительным, например, оказалось открытие, опубликованное в 1891 г. акад. Е.С. Федоровым, что может существовать только 230 разных типов кристаллической решетки, хотя любое вещество при определенных условиях может кристаллизоваться. Конечно, это открытие

* Мы будем цитировать Б. Трентовского по книге Н.Н. Моисеева [11].

CYBERNETICS	
<i>кибернетика</i>	
ORGANIZATION	
<i>организация</i>	
REALIZING, UNDERSTANDING	
<i>осознание</i>	
DEVELOPMENT	
<i>развитие</i>	
CONSCIOUSNESS	
<i>сознание</i>	

прежде всего в области минералогии и кристаллографии, но его более общий смысл и значение отметил сам Федоров. Важным было осознать, что все невообразимое разнообразие природных тел реализуется из ограниченного и небольшого числа исходных форм. Это оказывается верным и для языковых устных и письменных построений, архитектурных конструкций, строения вещества на атомном уровне, музыкальных произведений, других систем. Но Федоров пошел дальше. Развивая системные представления, он установил и некоторые закономерности развития систем. Ему принадлежит наблюдение, что главным средством жизнеспособности и прогресса систем является не их приспособленность, а способность к приспособлению (“жизненная подвижность”), не стройность, а способность к повышению стройности. Е.С. Федорова можно заслуженно отнести к числу основоположников теории систем.

ТЕКТОЛОГИЯ БОГДАНОВА

Следующая ступень в изучении системности как самостоятельного предмета связана с именем А.А. Богданова*. В 1911 г. вышел в свет первый том, а в 1925 г. – третий том его книги “Всеобщая организационная наука (тектология)” [3]. Большая общность тектологии связана с идеей Богданова о том, что все существующие объекты и процессы имеют определенную степень, уровень организованности. В отличие от конкретных естественных наук, изучающих специфические особенности организации конкретных явлений, тектология должна изучать общие закономерности организации для всех уровней организованности. Все явления рассматриваются как непрерывные процессы организации и дезорганизации. Богданов не дает строгого определения понятия организации, но отмечает, что уровень организации тем выше, чем сильнее свойства целого отличаются от простой суммы свойств его частей. Пожалуй, самой важной особенностью тектологии является то, что основное внимание уделяется закономерностям развития организации, рассмотрению соотношений устойчивого и изменчивого, значению обратных связей, учету собственных целей организации (которые могут как содействовать целям высшего уровня организации, так и противоречить им), роли открытых систем. Богданов довел динамические аспекты тектологии до рассмотрения проблемы кризисов, т.е. таких моментов в истории любой системы, когда неизбежна коренная, “взрывная” перестройка ее структуры. Он подчеркивал роль моделирования и математики как потенциальных методов решения задач тектологии.

Обратите внимание на специальности классиков системности: Ампер – физик, Трентовский – философ, Федоров – геолог, Богданов – медик, Винер – математик, Берталанфи – биолог, Пригожин – физик. Можно ли

Даже из столь беглого обзора основных идей тектологии видно, что Богданов предвосхитил, а кое в чем и превзошел многие положения современных кибернетических и системных теорий. Тот факт, что к тектологии стали обращаться лишь в самые последние годы, объясняется в большой степени противоречивостью, сложностью личности и судьбы Богданова. Талантливый и самонадеянный, увлекающийся и теряющий чувство меры, по профессии медик, он всерьез заинтересовался философией, но от материализма перешел на позиции марксизма, создав собственную философию – эмпириомонизм. В.И. Ленин в “Материализме и эмпириокритицизме” подверг его жесткой критике, после чего Богданов отошел вообще от всякой философии, включая собственную. Он активно участвовал в политической деятельности, но когда русская социал-демократия отвергла его вместе с “богоискателями”, вообще прекратил работу в партии. Тем не менее после революции он вошел в состав Коммунисти-

* Богданов – это псевдоним, его настоящая фамилия – Малиновский.

ческой Академии; написал не потерявший значение и поныне “Краткий курс политической экономии”. В.И. Ленин поддержал также его идею создать первый в мире Институт переливания крови, и Богданов, став его директором, начал проверять некоторые выводы тектологии, которую разрабатывал последние 20 лет своей жизни, на примере действительно сложной системы – кровеносной. Рискованные опыты он проводил на себе, и его абсолютная вера в безошибочность всех своих гипотез привела к трагедии: один из таких опытов окончился его гибелью. Преувеличение значимости своих идей Богдановым проявилось и в том, как он оценивал соотношение тектологии с философией:

“По мере своего развития тектология должна сделать излишней философию, и уже с самого начала стоит над нею, соединяя с ее универсальностью научный и практический характер” [3, с. 209].

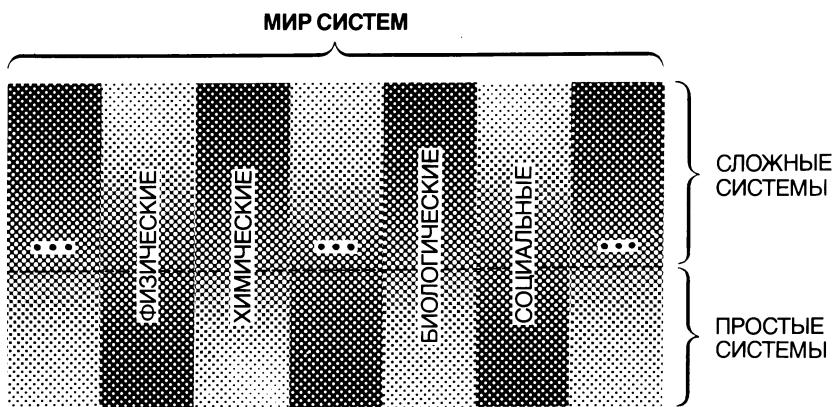
Все это в совокупности с новизной предмета тектологии и первоначальным непониманием ее задач и привело к тому, что о ней вспомнили только тогда, когда другие начали приходить к тем же результатам. Теперь ясно, что приоритет и вклад А.А. Богданова в развитие системных представлений нельзя недооценивать.

КИБЕРНЕТИКА ВИНЕРА

По-настоящему явное и массовое усвоение системных понятий, общественное осознание системности мира, общества и человеческой деятельности началось с 1948 г., когда американский математик Н. Винер опубликовал книгу под названием “Кибернетика” [5]. Первоначально он определил кибернетику как “науку об управлении и связи в животных и машинах”. Однако очень быстро стало ясно, что такое определение, сформировавшееся благодаря особому интересу Винера к аналогиям процессов в живых организмах и машинах, неоправданно сужает сферу приложения кибернетики. Уже в следующей книге [6] Н. Винер анализирует с позиций кибернетики процессы, происходящие в обществе.

Сначала кибернетика привела многих ученых в замешательство: оказалось, что кибернетики берутся за рассмотрение и технических, и биологических, и экономических, и социальных объектов и процессов. Возник даже спор – имеет ли кибернетика свой предмет исследования. Первый международный конгресс по кибернетике (Париж, 1956) даже принял предложение считать кибернетику не наукой, а “искусством эффективного действия”. В нашей стране кибернетика была встречена настороженно и даже враждебно. Рекламные заявления американских кибернетиков о работе над созданием “мыслящих машин” некоторыми философами были восприняты буквально, и кибернетика была объявлена ими идеалистической лжен наукой.

По мере развития кибернетики, уточнения ее понятий, разработки ее собственных методов, получения конкретных результатов в разных областях стало очевидным, что кибернетика – это самостоятельная наука, со своим, характерным только для нее предметом изучения, со своими специфическими методами исследования. В становление кибернетики внесли вклад и советские ученые. Важную роль сыграли определения, сформулированные в период горячих дискуссий о сути кибернетики: *кибернетика – это наука об оптимальном управлении сложными динамическими системами* (А.И. Берг); *кибернетика – это наука о системах, воспринимающих, хранящих, перерабатывающих и использующих информацию* (А.Н. Колмогоров). Эти определения признаны весьма общими и полными. Хотя в них необходимо разъяснить практически каждое слово (что и будет сделано в следующих главах), из них видно, что предметом кибернетики является исследование систем. Важно подчеркнуть, что, хотя при изучении системы на каком-то этапе потребуется учет ее конкретных свойств, для кибернетики в принципе несущественно, какова природа этой системы, т.е. является ли она физической, биологической, экономической, организационной или даже воображаемой, нереальной системой. Это делает понятным, почему кибернетика “вторгается” в совершенно разнородные сферы. Если (в порядке шутки) представить себе весь мир как “булку хлеба”, от которой каждая наука “отрезает себе ломоть”, то, образно говоря, все



1.2 ————— Отношение между кибернетикой и другими науками

науки разрезают эту булку “поперек”, а кибернетика – “вдоль”: в кибернетический “ломтик” попадают объекты любой природы, как только они оказываются сложными системами (рис. 1.2). То, что кибернетические методы могут применяться к исследованию объектов, традиционно “закрепленных” за той или иной наукой, должно рассматриваться не как “постороннее вмешательство неспециалистов”, а как рассмотрение этих объектов с другой точки зрения. Более того, при этом происходит взаимное обогащение кибернетики и других наук: с одной стороны, кибернетика получает возможность развивать и совершенствовать свои модели и методы, с другой – кибернетический подход к системе определенной природы может прояснить некоторые проблемы данной науки или даже выдвинуть перед ней новые проблемы, а главное – содействовать повышению ее системности.

С кибернетикой Винера связаны такие продвижения в развитии системных представлений, как типизация моделей систем, выявление особого значения обратных связей в системе, подчеркивание принципа оптимальности в управлении и синтезе систем, осознание информации как всеобщего свойства материи и возможности ее количественного описания, развитие методологии моделирования вообще и в особенности идеи математического эксперимента с помощью ЭВМ. Все это, без преувеличения, сыграло революционизирующую роль в развитии общественного сознания, человеческой практики и культуры, подготовило почву для того невиданного ранее размаха компьютеризации, которая происходит на наших глазах.

Однако необходимо воздержаться от преувеличенных оценок винеровской кибернетики. Некоторых раздражала рекламная шумиха, созданная вокруг кибернетики не только дельцами, но и некоторыми учеными. Впрочем, дело не только в субъективных оценках. Простое сравнение идей Винера с идеями Трентовского и Богданова показывает, что кибернетика не смогла дойти до рассмотрения действительно сложных систем, что винеровской кибернетике свойствен определенный техницизм, современная разновидность механицизма. В рассмотрении информационных процессов качественная сторона информации принесена в жертву количественной; принцип оптимальности реализуется только в полностью формализованных задачах; при моделировании интеллекта учитывается только логическая компонента мышления. Это действительно так, но все же стремление некоторых специалистов по информатике отмежеваться от винеровской кибернетики выглядит как сверхреакция на ее недостатки. Справедливее рассматривать кибернетику Винера как важный этап в развитии системных представлений, давший ценные идеи и результаты, этап, на котором встретились непреодоленные трудности и обнаружились недостатки самой теории.

ПОПЫТКИ ПОСТРОЕНИЯ ОБЩЕЙ ТЕОРИИ СИСТЕМ

Параллельно и как бы независимо от кибернетики прокладывается еще один подход к науке о системах – *общая теория систем*. Идея построения теории, приложимой к системам любой природы, была выдвинута австрийским биологом Л. Берталанфи [14]. Один из путей реализации этой идеи Берталанфи видел в том, чтобы отыскивать структурное сходство законов, установленных в различных дисциплинах, и, обобщая их, выводить общесистемные закономерности. Пожалуй, самым важным достижением Берталанфи является введение понятия *открытой системы*. В отличие от винеровского подхода, где изучаются внутрисистемные обратные связи, а функционирование систем рассматривается просто как отклик на внешние воздействия, Берталанфи подчеркивает особое значение обмена системы веществом, энергией и информацией (негэнтропией) с окружающей средой. В открытой системе устанавливается динамическое равновесие, которое может быть направлено в сторону усложнения организации (вопреки второму закону термодинамики, благодаря вводу негэнтропии извне), и функционирование является не просто откликом на изменение внешних условий, а сохранением старого или установлением нового подвижного внутреннего равновесия системы. Здесь усматриваются как кибернетические идеи гомеостазиса, так и новые моменты, имеющие свои истоки в биологии. Берталанфи и его последователи работают над тем, чтобы придать общей теории систем формальный характер. Однако заманчивый замысел построить общую теорию систем как новую логико-математическую дисциплину не реализован полностью до сих пор. Не исключено, что наибольшую ценность общей теории систем представит не столько ее математическое оформление, сколько разработка целей и задач системных исследований, развитие методологии анализа систем, установление общесистемных закономерностей.

Современный “прорыв в незнаное” в исследовании систем совершен бельгийской школой во главе с И. Пригожиным. Развивая термодинамику неравновесных физических систем (за результаты этих исследований Пригожин был удостоен Нобелевской премии 1977 г.), он вскоре понял, что обнаруженные им закономерности относятся к системам любой природы. Наряду с переоткрытием уже известных положений (иерархичность уровней организации систем; несводимость друг к другу и невыводимость друг за друга закономерностей разных уровней организации; наличие наряду с детерминированными случайных процессов на каждом уровне организации и др.) Пригожин предложил новую, оригинальную теорию системодинамики. Отметим, что наибольший интерес и внимание привлекли те ее моменты, которые раскрывают механизм самоорганизации систем. Согласно теории Пригожина [12], материя не является пассивной субстанцией; ей присуща спонтанная активность, вызванная неустойчивостью неравновесных состояний, в которые рано или поздно приходит любая система в результате взаимодействия с окружающей средой. Важно, что в такие переломные моменты (называемые “особыми точками” или “точками бифуркации”) принципиально невозможно предсказать, станет ли система менее организованной или более организованной (“дисциплинарной”, по терминологии Пригожина).

Заметим, что в данной главе история системного мышления была прослежена лишь до периода возникновения системного анализа. Более поздние события будут описаны в последующих главах, особенно в гл. 9.

Подведем итог

Наращивание системности знаний – естественный про-

цесс, происходящий во всех областях человеческой деятельности стихийно (как результат обратной связи через практику, как форма развития). Осознание же системности нашего познания и окружающего мира – это более высокий уровень системности знаний, и оно происходит труднее, медленнее, с отставанием, эадержками и петлянием, свойственными процессам блуждания и поиска. Это не бесцельное, хаотическое блуждание, а процесс поиска истины, в котором возможны задержки и ошибки, но его содержание и смысл не в них, а в продвижении к истине.

Summary

The increasing systematicity of knowledge is a natural process occurs spontaneously in all areas of human activity as a result of feedback via praxis, or as a form of development. But the recognition and conceptualization of the systematic nature of both human cognition and the surrounding world represents a higher level of understanding, and it proceeds more slowly, with the difficulties, setbacks, and repetitiveness characteristic of the investigative process. This is not an aimless, chaotic activity but the search for truth, and while delays and errors are possible, the substance of this search lies not in the errors but in its approach to the truth.

ЗАКЛЮЧЕНИЕ

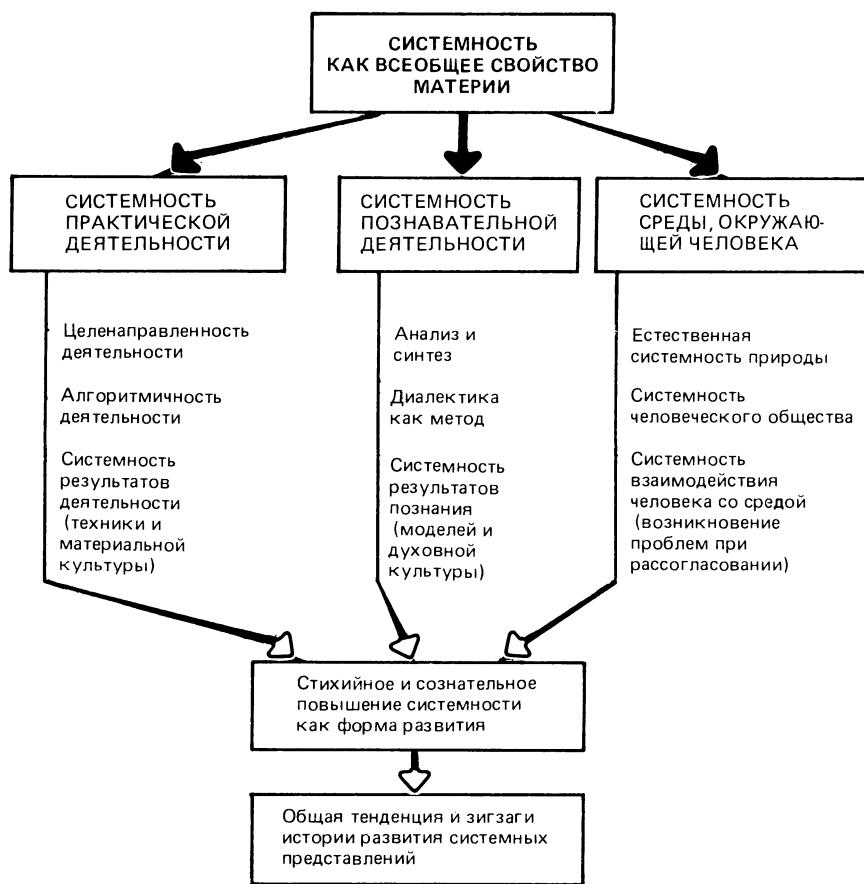
Для облегчения запоминания и в качестве “информации к размышлению” приведем схему, отображающую связь тем, рассмотренных в данной главе (рис.1.3).

ЛИТЕРАТУРА

1. Аверьянов А.Н. Системное познание мира. – М.: Политиздат, 1985.
2. Альтшулер Г.С. Алгоритм изобретения. – М.: Московский рабочий, 1973.
3. Богданов А.А. Всеобщая организационная наука (текнология). В 3 т. М., 1905–1924. Т. 3.
4. Бэкон Ф. Соч. В 2 т. М., 1978. Т. 2.
5. Винер Н. Кибернетика. – М.: Сов. радио, 1968.
6. Винер Н. Кибернетика и общество. – М.: ИЛ, 1958.
7. Декарт Р. Избр. произв.– М., 1950.
8. Зарипов Р.Х. Машинный поиск вариантов при моделировании творческого процесса. – М.: Наука, 1983.
9. Кант И. Соч. В 6 т. М. Т. 3.
10. Клир Дж. Системология.–М.: Радио и связь, 1990.
11. Маркс К., Энгельс Ф. Соч. 2-е изд. Т. 20.
12. Мусеев Н.Н. Люди и кибернетика. – М.: Молодая гвардия, 1984.
13. Пригожин И., Стенгерс И. Порядок из хаоса. – М.: Прогресс, 1986.
14. Bertalanffy L. An Outline of General Systems Theory. – “British J. for Phil. of Sci”. 1950. Vol. 1. N 2. 134 – 165.

УПРАЖНЕНИЯ

§ 1.2



1.3 ————— Взаимосвязь тем, рассмотренных в гл.1

- Приведите известные вам примеры современной механизации физических работ. Приблизительно подсчитайте, насколько механизация увеличила производительность труда при выполнении этих работ.
- Приведите примеры автоматизации труда.
- Приведите пример, доказывающий, что автомат может успешно работать только в тех условиях, на которые он рассчитан.
- Приведите пример деятельности, которая, по вашему мнению, не может быть автоматизирована. Обоснуйте это мнение.
- Приведите пример деятельности, которая в прошлом считалась чисто эвристической, а сейчас алгоритмизована.

§ 1.3

- Приведите известные вам примеры анализа и синтеза в познавательной деятельности.
- Обсудите с разных сторон изменения в системности наших знаний о природе после открытия Д.И. Менделеевым периодической системы элементов.
- Попробуйте вообразить себе отсутствие какого-либо признака системности в познавательной деятельности.

§ 1.4

- Чем, по-вашему, ограничена свобода воображения? Например, насколько разум может оторваться от условий реализуемости воображаемых вещей?
- Представьте себе, что сила тяжести на Земле уменьшилась вдвое. Какие изменения в конструкции жилых зданий следовало бы внести?

Вопросы для самопроверки

1. Может ли какой-нибудь объект или явление быть несистемным? Обоснуйте свой ответ.
2. Что такое проблемная ситуация?
3. Что называется алгоритмом?
4. В чем различие между полностью формализованным и полностью определенным алгоритмом?
5. Каковы три способа повышения производительности труда?
6. Чем ограничены возможности механизации?
7. Каково главное условие автоматизации?
8. Какие особенности мышления позволяют утверждать, что оно системно?
9. Каковы аргументы в пользу системности всей материи?
10. Какие правила мышления позволяют согласовать системность мышления с системностью окружающего мира? Когда возникает необходимость такого согласования?
11. Каковы основные события в развитии системных представлений в течение последних 150 лет?

§ 2.1. ШИРОКОЕ ТОЛКОВАНИЕ ПОНЯТИЯ МОДЕЛИ

Мы уже отмечали, что осознание объективной системности любой деятельности происходит с большой задержкой, на более поздних этапах стихийного повышения системности. Не является исключением и такая форма человеческой деятельности, как *моделирование*, т.е. построение, использование и совершенствование моделей. Интересно, например, проследить, как развивалось само понятие модели [11].

Первоначально *моделью* называли *некое вспомогательное средство, объект, который в определенной ситуации заменял другой объект*. При этом далеко не сразу была понята универсальность законов природы, всеобщность моделирования, т.е. не просто возможность, но и необходимость представлять любые наши знания в виде моделей. Например, древние философы считали невозможным моделирование естественных процессов, так как, по их представлениям, природные и искусственные процессы подчинялись различным закономерностям. Они полагали, что отобразить природу можно только с помощью логики, методов рассуждений, споров, т.е., по современной терминологии, языковых моделей. Через несколько столетий девизом английского Королевского научного общества стал лозунг “Ничего словами!”, который явился кратчайшим изложением принципов естествознания: признавались только выводы, подкрепленные экспериментально или математическими выкладками. В английском языке до сих пор в понятие “наука” не входят области знания, которым в русском языке соответствует термин “гуманитарные науки”, – они отнесены к категории “искусств”. В результате очень долго понятие “модель” относилось только к материальным объектам специального типа, например манекен (модель человеческой фигуры), гидродинамическая уменьшенная модель плотины, модели судов и самолетов, чучела (модели животных) и т.п.

РАЗВИТИЕ ПОНЯТИЯ МОДЕЛИ

Осмысливание основных особенностей таких моделей привело к разработке многочисленных определений, типичным примером которых служит следующее: *моделью* называется некий *объект-заместитель, который в определенных условиях может заменять объект-оригинал, воспроизводя интересующие нас свойства и характеристики оригинала, причем имеет существенные преимущества удобства* (наглядность, обозримость, доступность испытаний, легкость оперирования с ним и пр.). Затем были осознаны модельные свойства чертежей, рисунков, карт – реальных объектов искусственного происхождения, воплощающих абстракцию довольно высокого уровня. Следующий шаг заключался в признании того, что моделями могут служить не только реальные объекты, но и абстрактные, идеальные построения. Типичным примером служат математические модели. В результате деятельности математиков, логиков и философов, занимавшихся исследованием оснований математики, была создана *теория моделей*. В ней модель определяется как результат отображения одной абстрактной математической структуры на другую, тоже абстрактную, либо как результат интерпретации структуры в терминах и образах второй.

В XX в. понятие модели становится все более общим, охватывающим и реальные, и идеальные модели. При этом понятие абстрактной модели вышло за пределы математических моделей, стало относиться к любым знаниям и представлениям о мире. Следует отметить, что споры вокруг такого широкого толкования понятия модели продолжаются и поныне. Рассмотрим аргументы, фигурирующие в таких спорах.

Стоит ли понятие абстрактной модели распространять на такие формы научных знаний, как законы, гипотезы, теории? Сторонники положительного ответа на этот вопрос отмечают, что психологический барьер неприятия объясняется тем, что понятия гипотезы, закономерности, теории сформировались и установились в языке науки и философии значительно раньше, чем понятие модели. Эти понятия, будучи исторически первыми, воспринимаются и как логически первичные, причем в этой схеме модели отводится роль лишь вспомогатель-

ного средства. Однако при этом содержание понятия модели обедняется, неоправданно сужается. Дело в том, что классифицировать гипотезу или теорию как модель вовсе не означает подмену одного понятия другим или отождествление этих, безусловно, разных понятий. Модели могут быть качественно различными, они образуют *иерархию*, в которой модель более высокого уровня (например, теория) содержит модели нижних уровней (скажем, гипотезы) как свои части, элементы. Важно также, что признание идеальных представлений, научных построений, законов в качестве моделей подчеркивает их относительную истинность.

МОДЕЛЬ КАК ФИЛОСОФСКАЯ КАТЕГОРИЯ

Другой вопрос, часто возникающий в спорах: не означает ли такое широкое толкование модели, что это понятие становится применимым ко всему и, следовательно, логически пустым? Этот вопрос дает возможность обсудить некоторые особенности моделей. Во-первых, еще раз отметим иерархичность моделей, поэтому применительно к разным объектам понятие модели может иметь разное содержание. Во-вторых, тот факт, что любой объект может быть использован как модель, вовсе не означает, что он не может быть ничем иным. Например, ботинок также может являться моделью его владельца (скажем, по запаху ботинка сыскная собака отыщет преследуемого; по состоянию ботинка можно судить о некоторых особенностях сложения и даже чертах характера его хозяина), но это не лишает смысла ни понятие “обувь”, ни понятие “модель”. В-третьих, самые общие понятия совсем не являются логически пустыми: материя, движение, энергия, организация, система, ..., модель.

Подведем итог

Сначала в сфере научных дисциплин информационного, кибернетического, системного направления, а затем и в других областях науки модель стала осознаваться как нечто универсальное, хотя и реализуемое различными способами. Модель есть способ существования знаний.

§ 2.2. МОДЕЛИРОВАНИЕ – НЕОТЪЕМЛЕМЫЙ ЭТАП ВСЯКОЙ ЦЕЛЕНАПРАВЛЕННОЙ ДЕЯТЕЛЬНОСТИ

Всякий процесс труда есть деятельность, направленная на достижение определенной цели. Рабочий обтачивает заготовку, которая должна превратиться в деталь. Скульптор обтесывает мраморную глыбу, оставляя в ней то, что станет скульптурой. Земледелец пашет, сеет, собирает урожай, чтобы произвести продукты питания. Спортсмен настойчиво тренируется, чтобы победить в соревновании. Студент учится, чтобы иметь профессию.

Эта идея в обобщенном виде была выражена К. Марксом в “Капитале”: “В конце процесса труда получается результат, который уже в начале этого процесса имелся в представлении человека, т.е. идеально. Человек не только изменяет форму того, что дано природой; в том, что дано природой, он осуществляет вместе с тем и свою сознательную цель, которая как закон определяет способ и характер его действий” [10, с. 189].

Целевой характер имеет не только трудовая деятельность. Отдых, развлечения, прогулки, игры, физзарядка, чтение, коллекционирование и т.п. обычно не рассматриваются как труд, но их целевой характер очевиден. Поэтому следует говорить о различных видах целесообразной деятельности человека.

Важнейшим, организующим элементом такой деятельности является цель – образ желаемого будущего, т.е. модель состояния, на реализацию которого и направлена деятельность. Однако роль моделирования этим не ограничивается. Системность деятельности проявляется

Summary

First in the information, cybernetics and systems sciences and later in other areas of science, a model has been conceptualized as something universal, though able to vary in its material substance. A model is a mode of existence of knowledge.

в том, что она осуществляется по определенному плану, или, как мы уже отмечали, по определенному алгоритму. Следовательно, алгоритм – образ будущей деятельности, ее модель.

Как правило, деятельность редко осуществляется по жесткой программе, без учета (точнее – только с априорным учетом) того, что происходит на промежуточных этапах. Чаще приходится оценивать текущий результат предыдущих действий и выбирать следующий шаг из числа возможных. Это означает, что необходимо сравнивать последствия всех возможных шагов, не выполняя их реально, т.е. “проиграть” их на модели.

Таким образом, моделирование является обязательным, неизбежным действием во всякой целесообразной деятельности, пронизывает и организует ее, представляет собой не часть, а

ACTIVITY <i>деятельность</i>
MODEL <i>модель</i>
FACILITIES, CONVENiences <i>удобства</i>
PURPOSEFUL, TELIC <i>целенаправленный</i>
PURPOSE, OB- JECTIVE, AIM, GOAL <i>цель</i>

аспект этой деятельности.

ЦЕЛЬ КАК МОДЕЛЬ

Из предыдущего обсуждения также видно, что модель является не просто образом – заменителем оригинала, не вообще каким-то отображением, а отображением целевым.

Чтобы подчеркнуть это, представим, какие модели одного и того же бревна используют в своей деятельности разные члены туристской группы, пришедшей к месту новой стоянки: одному поручено оборудовать лагерь, и он прикидывает, использовать ли это бревно для стола или как сиденье; другой отвечает за кострище, а для дров от бревна требуются не геометрические, а совсем другие качества; третьего интересует возраст дерева, и он обследует спил бревна; художник ищет у бревна сук с замысловатым изгибом... Короче говоря, модель отображает не сам по себе объект-оригинал, а то, что в нем нас интересует, т.е. то, что соответствует поставленной цели.

Из того, что модель является целевым отображением, с очевидностью следует множественность моделей одного и того же объекта: для разных целей обычно требуется разные модели. Сама целевая предназначенностю моделей позволяет все разнообразное множество моделей разделить на основные типы – по типам целей [17].

ПОЗНАВАТЕЛЬНЫЕ И ПРАГМАТИЧЕСКИЕ МОДЕЛИ

Поскольку модели играют чрезвычайно важную роль в организации любой деятельности человека, все виды деятельности удобно разделить по направленности основных потоков информации, циркулирующих между субъектом и окружающим его миром. Разделим модели на **познавательные** и **прагматические**, что соответствует делению целей на теоретические и практические. Хотя это деление (как, впрочем, и всякое другое) относительно

Модель понимается широко: модели бывают не только материальные, но и идеальные. Например, все наши знания представлены моделями. Модели – это специальные систе-

и легко привести примеры, когда конкретную модель нелегко однозначно отнести к одному из классов, оно все же не целиком условно и отображает реальные различия.

Проявления этих различий разнообразны, но, пожалуй, наиболее наглядно разница между познавательными и прагматическими моделями проявляется в их отношении к оригиналам в процессе деятельности.

Познавательные модели являются формой организации и представления знаний, средством соединения новых знаний с имеющимися. Поэтому при обнаружении расхождения

между моделью и реальностью встает задача устранения этого расхождения с помощью изменения модели.

О способах изменения моделей мы будем говорить позже, а сейчас важно подчеркнуть, что познавательная деятельность ориентирована в основном на *приближение модели к реальности*, которую модель отображает (рис. 2.1, а).

Прагматические модели являются средством управления, средством организации практических действий, способом представления образцово правильных действий или их результата, т.е. являются рабочим представлением целей. Поэтому использование прагматических моделей состоит в том, чтобы при обнаружении расхождений между моделью и реаль-



2.1 ————— Различие между познавательной и прагматической моделью: а) познавательная модель (модель подготавливается под реальность); б) прагматическая модель (реальность направляет усилия на изменение реальности так, чтобы приблизить реальность к модели)

нностью направить усилия на изменение реальности так, чтобы *приблизить реальность к модели* (рис. 2.1, б). Таким образом, прагматические модели носят нормативный характер, играют роль стандарта, образца, под которые “подгоняются” как сама деятельность, так и ее результат. Примерами прагматических моделей могут служить планы и программы действий, уставы организаций, кодексы законов, алгоритмы, рабочие чертежи и шаблоны, параметры отбора, технологические допуски, экзаменационные требования и т.д.

Другими словами, основное различие между познавательными и прагматическими моделями можно выразить так: *познавательные модели отражают существующее, а прагматические – не существующее, но желаемое и (возможно) осуществимое*.

СТАТИЧЕСКИЕ И ДИНАМИЧЕСКИЕ МОДЕЛИ

Не следует понимать сказанное как абсолют. Например, возникают сложности с отнесением эстетических моделей (произведений искусства) или моделей медицины либо, например, детских игрушек или карт местности к одному из указанных классов. С другой стороны, прагматические модели так же изменяются, развиваются, как и познавательные, и в их динамике так много общего, что в эти моменты их различие начинает размываться, иногда исчезать.

Другим принципом классификации целей моделирования, другим ответом на вопрос о том, что отображается в модели, может служить деление моделей на статические и динамические. Для одних целей нам может понадобиться модель конкретного состояния объекта, своего рода “моментальная фотография” интересующего нас объекта. Такие модели называются **статическими**. Примером являются структурные модели систем. В тех же случаях, когда наши цели связаны не с одним состоянием, а с различием между состояниями, возникает необходимость в отображении процесса изменений состояния. Такие модели называются **динамическими**; примером их служат функциональные модели систем.

Подведем итог

Модели имеют целевой ха-

Summary

тер: один из принципов классификации моделей состоит в соотнесении типов моделей с типами целей; примером такой классификации служит деление моделей на познавательные и

Models are purposeful. One of the principles by which we classify models is based on a difference between purposes. For example, we may classify models as cognitive or pragmatic, or as static or dynamic.

pragmaticальные и, кроме того, на статические и динамические.

§ 2.3. СПОСОБЫ ВОПЛОЩЕНИЯ МОДЕЛЕЙ

Отображение, которым является модель, есть отношение между отображаемым и отображающим объектами. В предыдущем параграфе мы обсудили, какие аспекты объекта-оригинала отображаются в модели. Выяснилось, что на отношение отображения весьма существенным образом влияет (точнее, в этом отношении непосредственно участвует) все, что связано с целью, под которую создается модель. Перейдем теперь к рассмотрению того, на чем осуществляется отображение, т.е. из чего строятся модели.

Так как мы обсуждаем (пока) только модели, сознательно создаваемые человеком, то в его распоряжении имеется два типа материалов для построения моделей – средства самого сознания и средства окружающего материального мира. Соответственно этому модели делятся на **абстрактные (идеальные) и материальные (реальные, вещественные)**.

АБСТРАКТНЫЕ МОДЕЛИ И РОЛЬ ЯЗЫКОВ

Абстрактные модели являются идеальными конструкциями, построенными средствами мышления, сознания. Очевидно, что к абстрактным моделям относятся языковые конструкции; однако современные представления о мышлении и сознании утверждают, что языковые модели (т.е. модели, построенные средствами естественного языка) являются своего рода конечной продукцией мышления, уже готовой или почти готовой для передачи другим носителям языка. Установлено, что на более ранних стадиях работы человеческого мозга важную, хотя сейчас во многом неясную, роль играют и неязыковые формы мышления, которые обозначаются терминами “эмоции”, “бессознательное”, “интуиция”, “озарение”, “образное мышление”, “подсознание”, “эвристика” и т.п. Все это, однако, касается механизма мышления, т.е. внутренних моделей нашего мозга. Нас же сейчас интересуют абстрактные модели, предназначенные для общения между людьми. Оставив в стороне иррациональные формы общения (искусство, гипноз, телепатию, сопереживание, мимику и жестикуляцию и пр.), со средоточим внимание на моделях, создаваемых средствами языка.

На естественном языке мы можем говорить обо всем, он является универсальным средством построения любых абстрактных моделей. Эта универсальность обеспечивается не только возможностью введения в язык новых слов, но и возможностью иерархического построения все более развитых языковых моделей (слово – предложение – текст; понятия – отношения – определения – конструкции...). Универсальность языка достигается, кроме прочего, еще и тем, что языковые модели обладают неоднозначностью, расплывчатостью, размытостью. Это свойство проявляется уже на уровне слов. Многозначность почти каждого слова (см. толковый словарь любого языка) или неопределенность слов (например, “много”, “несколько”) вместе с многовариантностью их возможных соединений во фразы позволяет любую ситуацию отобразить с достаточной для обычных практических целей точностью. Эта приблизительность – неотъемлемое свойство языковых моделей. Человек преодолевает в практике их расплывчатость с помощью “понимания”, “интерпретации” [4]. Иногда эта расплывчатость сознательно используется – в юморе, дипломатии, поэзии; она придает прелесть

человеческому общению; в других случаях она мешает выразиться так точно, как хотелось бы (“у меня не хватает слов”, или тютчевское “мысль изреченная есть ложь”).

Рано или поздно практика сталкивает нас с ситуациями, когда приблизительность естественного языка оборачивается недостатком, который необходимо преодолеть на постоянной основе. Такую основу предоставляет выработка “профессионального” языка людьми, связанными общей для них, но частной для всех остальных деятельностью. Например, у северных народов имеется несколько десятков разных слов, обозначающих различные состояния снега; у африканского скотоводческого племени масаев столько разных слов, выражающих различия между коровами, что масай по одному слову может выделить одно животное из огромного стада.

Наиболее ярко это видно на примере языков конкретных наук. Дифференциация наук объективно потребовала создания специализированных языков, более четких и точных, чем естественный. Модели специальных наук более точны, более конкретны, они содержат больше информации. Новые знания аккумулируются в новых моделях, и если старых языковых средств для их построения не хватает, то возникают еще более специализированные языки. В результате приходим к иерархии языков и соответствующей иерархии типов моделей. На верхнем уровне этого спектра находятся модели, создаваемые средствами естественного языка, и так вплоть до моделей, имеющих максимально достижимую определенность и точность для сегодняшнего состояния данной отрасли знаний. Видимо, так и следует понимать известные высказывания Канта и Маркса о том, что любая отрасль знания может тем с большим основанием именоваться наукой, чем в большей степени в ней используется математика. Математические модели обладают абсолютной точностью, но чтобы дойти до их использования в данной области, необходимо получить достаточно для этого количество знаний. Нематематизированность какой-то науки не означает ее “ненаучность”, а есть следствие сложности, недостаточной познанности ее предмета, есть временное явление.

МАТЕРИАЛЬНЫЕ МОДЕЛИ И ВИДЫ ПОДОБИЯ

Перейдем теперь к рассмотрению *материальных (реальных, вещественных)* моделей. Чтобы некоторая материальная конструкция могла быть отображением, т.е. замещала в каком-то отношении оригинал, между оригиналом и моделью должно быть установлено отношение похожести подобия. Существуют разные способы установления такого подобия, что придает моделям особенности, специфичные для каждого способа.

Прежде всего это подобие, устанавливаемое в результате физического взаимодействия (или цепочки физических взаимодействий) в процессе создания модели. Примерами таких отображений являются фотографии, масштабированные модели самолетов, кораблей или гидротехнических сооружений, макеты зданий, куклы, протезы, шаблоны, выкройки и т.п. Назовем такое подобие **прямым**. И прямое подобие может быть иногда лишь отдаленным сходством, но только при прямом подобии возможна трудно обнаружимая взаимозаменяемость модели и оригинала (например, копии произведений искусства, голограммические изображения предметов и т.п.) и даже фактическая перемена их местами (натурщик является моделью в работе художника, манекенщица моделирует будущих потребителей одежды, актер является моделью персонажа пьесы). С другой стороны, как бы хороша ни была модель, она все-таки лишь заменитель оригинала, выполняющий эту роль только в определенном отношении. Даже тогда, когда модель прямого подобия осуществлена из того же материала, что и оригинал, т.е. подобна субстратно, возникают проблемы переноса результатов моделирования на оригинал. Рассмотрим, например, испытания уменьшенной модели корабля на гидродинамические качества. Часть условий эксперимента можно привести в соответствие масштабам модели (скорость течения), другая же часть условий (вязкость и плотность воды, сила тяготения, определяющие свойства волн, и т. д.) не может быть масштабирована. Задача пересчета данных модельного эксперимента на реальные условия становится нетривиальной;

возникла разветвленная, содержательная теория подобия [19], относящаяся именно к моделям прямого подобия.

Второй тип подобия в отличие от прямого назовем **косвенным**. Косвенное подобие между оригиналом и моделью устанавливается не в результате их физического взаимодействия, а объективно существует в природе, обнаруживается в виде совпадения или достаточной близости их абстрактных моделей и после этого используется в практике реального моделирования. Наиболее известным примером этого является электромеханическая аналогия. Оказалось, что некоторые закономерности электрических и механических процессов описываются одинаковыми уравнениями; различие состоит лишь в разной физической интерпретации пе-

**SIGN, SYMBOL,
MARKзнак
HIERARCHY
иерархия
INSIGHT озарение
CONDITIONAL,
CONVENTIONAL
условный
LANGUAGE язык**

Цель моделирования диктует, какие стороны оригинала должны быть отражены в модели. Различным целям соответствуют разные модели одного и того же объекта.

Модели могут строиться средствами мышления (абстрактные модели) либо средствами материального мира (реальные модели). Особое место среди абстрактных моделей занимают

ременных, входящих в эти уравнения. В результате оказывается возможным не только заменить неудобное и громоздкое экспериментирование с механической конструкцией на простые опыты с электрической схемой, перепробовать множество вариантов, не переделывая конструкцию, но и “проиграть” на модели варианты, в механике пока неосуществимые (например, с произвольным и непрерывным изменением масс, длин и т.д.). Роль моделей, обладающих косвенным подобием оригиналу, очень велика. Часы – аналог времени; подопытные животные у медиков – аналоги человеческого организма; автопилот – аналог летчика; электрический ток в подходящих цепях может моделировать транспортные потоки информации в сетях связи, течение воды в городской водопроводной сети; аналоговые вычислительные машины позволяют найти решение почти всякого дифференциального уравнения, представляя собой таким образом модель, аналог процесса, описываемого этим уравнением.

Особенности аналогий, ограничения на перенос результатов, полученных моделированием, на сам оригинал оживленно обсуждались в период дискуссий о кибернетике (см., например, [7; 23]). Мы еще вернемся к вопросу о силе аналогий, а пока отметим, что роль аналогий (моделей косвенного подобия) в науке, технике, практике вряд ли можно переоценить: они просто незаменимы, альтернативой модели может быть только другая модель*.

Третий, особый класс реальных моделей образуют модели, подобие которых оригиналу не является ни прямым, ни косвенным, а устанавливается в результате соглашения. Назовем такое подобие **условным**. Примерами условного подобия служат деньги (модель стоимости), удостоверения личности (официальная модель владельца), всевозможные и разнообразные сигналы (модели сообщений), рабочие чертежи (модели будущей продукции), карты (модели местности) и т.д.

С моделями условного подобия приходится иметь дело очень часто, поскольку они являются способом материального воплощения абстрактных моделей, вещественной формой, в которой абстрактные модели могут передаваться от одного человека к другому, храниться до (иногда очень отдаленного) момента их использования, т.е. отчуждаться от сознания и все-таки сохранять возможность возвращения в абстрактную форму. Это достигается с помощью

* Роль аналогий между абстрактными моделями подробно рассматривается в [13].

соглашения о том, какое состояние реального объекта ставится в соответствие данному элементу абстрактной модели. Такое соглашение принимает вид совокупности правил построения моделей условного подобия и правил пользования ими.

ЗНАКОВЫЕ МОДЕЛИ И СИГНАЛЫ

Эта общая схема конкретизируется и углубляется в ряде конкретных наук, в которых используются или непосредственно изучаются модели условного подобия. Например, теория связи, теория информации, радиотехника, теория управления и ряд других наук имеют дело со специфическими моделями условного подобия, которые применяются в технических устройствах без участия человека; они получили название сигналов. В этих науках правила построения и способы использования сигналов, названные *кодом, кодированием и декодированием*, сами стали предметом углубленных исследований (например, возникла очень развитая теория кодирования).

С иных позиций рассматриваются модели условного подобия в науках, изучающих создание и использование этих моделей самим человеком. У пред назначенных для этого моделей имеется своя специфика, позволяющая дать им специальное название – **знаки** – и требующая специальных методов для ее исследования. Возникшая в связи с этим область знаний получила название *семиотики* (от греч. “знак”). Семиотика изучает знаки не в отдельности, а как входящие в знаковые системы, в которых выделено три основных группы отношений:

синтаксис (греч. “построение, порядок”), т.е. отношения между различными знаками, позволяющие отличать их и строить из них знаковые конструкции все более высокой сложности;

семантика (греч. “обозначение”), т. е. отношения между знаками и тем, что они обозначают, или вложенный, изначальный смысл знаков;

прагматика (греч. “дело, действие”), т. е. отношения между знаками и теми, кто их использует в своей деятельности, или понятый, воспринятый смысл знаков.

Существуют и другие многочисленные аспекты исследования моделей условного подобия (языкознание, картография, криптография, графология, техническое черчение, нумизматика, информатика, литературоведение и т.д.).

Не вдаваясь в другие подробности, отметим в заключение, что хотя условное подобие в принципе не требует фактического сходства, но оно должно строиться с учетом особенностей человека – создателя и потребителя моделей условного подобия. В разных множествах реальных объектов мы будем искать “сырье” для знаковых моделей, предназначенных для слепых, глухонемых и обычных людей. Выбор символики для обозначения цифр только на первый взгляд кажется произвольным: в практике вычислений арабская символика в конце концов вытеснила римскую из-за существенного различия в удобстве ручного выполнения операций над знаками чисел. На ЭВМ двоичная символика вытеснила арабскую по подобным соображениям. Такие примеры можно множить.

Подведем итог

Summary

Можно строить модели в воображении, идеально, средствами мышления – такие модели называются абстрактными, а abstract. One can build models using real objects and processes. реальных объектов и процессов Such models are called real, material, substantial, and sometimes – такие модели называют вещественными, ино- physical. Sign models which are гда физическими. В некотором смысле промежуточное, по- material but have an abstract meaning, have, in a sense, an intermediate nature.

мают реальные модели, имеющие абстрактное содержание, – знаковые модели.

§ 2.4. УСЛОВИЯ РЕАЛИЗАЦИИ СВОЙСТВ МОДЕЛЕЙ

Для того чтобы модель отвечала своему назначению, недостаточно взять готовую модель или создать новую; необходимо, чтобы существовали условия, обеспечивающие ее функционирование. Отсутствие (или недостаточность) таких условий лишает модель ее модельных свойств, т.е. переводит модель в качественно иное состояние нераскрытия ее потенциальных возможностей.

Приведем примеры, иллюстрирующие эту особенность моделей. Бумажные денежные знаки могут играть роль модели стоимости только до тех пор, пока в среде их обращения существуют правовые нормы и финансовые учреждения, поддерживающие функционирование денег: царские ассигнации и “керенки”, может быть, и имеют сейчас некоторую историческую ценность, но уже не как деньги. Древнеегипетские иероглифические надписи не могли быть прочитаны, пока не был найден знаменитый розеттский камень с одинаковым текстом на забытом древнеегипетском языке и на сохранившемся в памяти специалистов древнегреческом языке: можно расшифровать сигнал только зная его код. В истории известны прекрасные идеи, “обогнавшие свое время”, т.е. не соответствовавшие общественному уровню знания и технологии, своего времени и потому не воспринятые обществом: вспомним вертолет Леонардо да Винчи (XV в.), кибернетику Трентовского (1848), универсальную вычислительную машину Бэббиджа (1883) или шутливую историю янки при дворе короля Артура, придуманную Марком Твеном. Зато идеи (т.е. абстрактные модели), попадающие в благоприятную общественную среду, “овладевшие массами, – по выражению К. Маркса, – становятся материальной силой”. Еще один яркий пример отмеченной особенности дает программа для ЭВМ – машинная модель алгоритма: программисты знают, что малейшее рассогласование в ней с языком машины полностью обесценивает программу.

Итак, для реализации своих модельных функций необходимо, чтобы модель была согласована с культурной средой, в которой ей предстоит функционировать, входила в эту среду не как чуждый ей элемент, а как ее естественная часть.

Такое свойство *согласованности с культурной средой* столь важно и характерно для модели, что стоило бы связать с ним специальный термин. Может быть, приживется термин *ингерентность*: во-первых, он происходит от английского *inherent* – внутренний, собственный (чему-то), накрепко связанный (с чем-то), существующий как неотъемлемая часть (чего-то), что хорошо отвечает смыслу того, что мы хотим обозначить этим термином; во-вторых, он перекликается и по происхождению, и по звучанию, и по смыслу с уже прижившимся в русском языке термином “*когерентность*”. Чтобы прижиться, новый термин должен стать ингерентным языку; будущее покажет, произойдет ли это.

Это очень общее свойство моделей можно при необходимости рассматривать и с более конкретных, частных позиций, выделяя конкретные аспекты согласованности модели со средой. Например, чрезвычайно важным аспектом такой согласованности (ингерентности) является обеспеченность функционирования модели ресурсами. Рассмотрение этого аспекта дает возможность четкого различия между теми системами, которые можно назвать “большими”, и теми, которые можно назвать “сложными”. Отложим обсуждение классификации систем (см. § 4.3): мы пока еще не дали самого определения системы, но все же обратим особое внимание на то, что для функционирования модели нужны ресурсы, в том числе и материальные, даже если модель абстрактна. Особый аспект необходимости согласованности модели со средой состоит в обеспечении операционности модели, т.е. ее работы, реализации самого процесса моделирования. Это означает, что в модели должны быть предусмотрены не только “стыковочные узлы” (интерфейсы) со средой, но и в самой среде должны быть реализованы подсистемы, другие модели и алгоритмы, обеспечивающие, поддерживающие функционирование модели, использующие результаты ее функционирования, управляющие ходом процесса моделирования: не только модель должна приспосабливаться к среде, но и среда к модели.

Подведем итог

Понятие модели снова расширилось; как выражаются математики, модель оказывается “многоместным отношением”. Это означает, что понятие модели относится не просто к объекту, который мы так называем; оно включает (в определенных отношениях): субъекта, организующего моделирование, и задачу, ради которой проводится моделирование (обоих – через цель); объект-оригинал, который моделируется; среды, из которых создается модель; а теперь и среду, в которой модель должна функционировать.

§ 2.5. СООТВЕТСТВИЕ МЕЖДУ МОДЕЛЬЮ И ДЕЙСТВИТЕЛЬНОСТЬЮ: РАЗЛИЧИЯ

В предыдущих параграфах мы рассмотрели вопросы о том, что отображает модель, из чего и как она может быть построена, каковы внешние условия актуализации (осуществления функций) модели. Теперь необходимо рассмотреть те качества моделей, которые определяют ценность самого моделирования, т.е. отношение моделей с отображаемой ими реальностью: чем отличаются модели и моделируемые объекты или явления, в каком смысле и до какой степени можно отождествлять модель с оригиналом. Обсудим главные различия между моделью и действительностью: конечность, упрощенность и приближенность модели.

КОНЕЧНОСТЬ МОДЕЛЕЙ

Начнем обсуждение с *конечности моделей*. Мир, частью которого мы являемся, бесконечен, как бесконечен и любой объект, не только в пространстве и времени, но и в своих связях с другими объектами, и в том, что к любому числу отношений, в которых мы рассматривали данный объект, всегда можно добавить еще одно. И мы сами, как и все природные объекты,

Summary

In this section we have broadened the concept of a model: as mathematicians say, a model turns out to be “a many – placed relation”. This means that the concept of a model contains not only an object called “a model”: it also encompasses (in certain relations) a subject who organizes the modeling; a problem for which modeling is performed; a modeled object; the means by which modeling is effected; and, finally, an environment in which the model will function.

также бесконечны. Однако если иметь в виду не любые наши качества (скажем, химический состав тканей, строение кровяных телец, структуру глаза, волоса или зуба и т.д.), а лишь те, которые отличают нас от других объектов, в том числе и живых (а к таким качествам прежде всего относятся мышление и труд), то, к сожалению, здесь возможности природы ограничены, конечны. Прежде всего ограничены наши собственные ресурсы – число нервных клеток мозга, число действий, которые мы можем выполнить в единицу времени, да и само время, которое мы можем затратить для решения какой-то задачи, – максимум вся сознательная жизнь. Ограничены и внешние ресурсы, которые мы можем вовлечь в конкретный процесс практической или теоретической деятельности (даже физики, строящие грандиозные радио-

ADEQUACY*адекватность*

INHERENCE*ингерентность*

FINITENESS*конечность*

ENVIRONMENT*среда (обитания)*

SIMPLIFICATION*упрощение*

телескопы или синхрофазотроны, жалуются на недостаточность отпускаемых им ресурсов).

Возникает явное противоречие: необходимо познавать бесконечный мир конечными средствами. Как ни странно, но это оказывается возможным – такова человеческая практика. Способ преодоления этого противоречия состоит в построении моделей. Так, Розенблют и Винер отмечали:

“Частные модели, при всех их несовершенствах, – единственное средство, выработанное наукой дня понимания мира. Из этого положения не вытекает пораженческой установки. В нем признается только, что главное орудие науки – человеческий разум, а этот разум конечен” [18].

Прежде чем пытаться ответить на интригующий вопрос, каким образом бесконечное удается отображать конечными моделями, завершим обсуждение самой конечности моделей. Если конечность абстрактных моделей бесспорна (они сразу наделяются строго фиксированным числом свойств), то реальные модели – это ведь некоторые вещественные объекты, а как всякие объекты, они бесконечны. Здесь-то и оказывается различие между самим объектом и тем же объектом, используемым в качестве модели другого объекта. Из необозримого множества свойств объекта-модели выбираются и используются лишь некоторые свойства, подобные интересующим нас свойствам объекта-оригинала. Особенно наглядна конечность (в этом смысле) знаковых моделей: цветок в окне явочной квартиры Штирлица означал провал явки; ясно, что многочисленные свойства цветка, изучаемые ботаникой, физиологией растений, цветоводством, искусством и икебаной, не имели прямого отношения к знаковой функции цветка. Модель подобна оригиналу в конечном числе отношений – это один аспект конечности реальных моделей.

Чтобы осуществить то, для чего модель была создана, недостаточно только наличия модели. Модель должна быть в достаточной степени согласована со средой (культурой). Это свойство модели называется ингирентностью.

Модель конечна, так как:

1) она отображает оригинал лишь в конечном числе отношений;

2) ресурсы моделирования конечны. Модель всегда упрощенно отображает оригинал. Причины упрощенности:

1) конечность модели;

2) модель отображает только главные, наиболее

Другой аспект возникает в связи с реальными моделями, обладающими свойством непрерывности: ведь непрерывность – одно из проявлений бесконечности. Видимо, можно говорить о конечности моделей и в этом случае, если считать, что непрерывность – это свойство природы, а не модели. В абстракции же понятие непрерывности получается в результате рассмотрения дискретной совокупности последовательных состояний. Кроме того, после открытия атомарности вещества, пространства, действия (а теперь исследуется вопрос о возможной дискретности и времени) можно усомниться в реальности непрерывности: не является ли

она только удобной, экономной абстракцией. Например, для непрерывных сред мы ввели понятие плотности, но это понятие теряет смысл, если объем взять очень малым (охватывающим несколько молекул) или очень большим (скажем, размера Солнечной системы): в реальности нет непрерывных сред, это просто удобная абстрактная модель.

Однако не следует понимать сказанное так, будто понятию непрерывности или плотности на самом деле ничто не соответствует в реальности. Эти понятия отображают определенные свойства окружающей нас реальности, но проявляться (и вообще существовать) эти свойства могут лишь при определенных условиях. С другой стороны, не всякое понятие отображает нечто осозаемое, непосредственно существующее; абстракция может быть “многоэтажной”, иерархической; можно говорить не только о модели чего-то реального, но и о модели моделей, и число таких ступеней ограничено, по-видимому, только практической надобностью.

Перейдем теперь к рассмотрению тех факторов, которые позволяют с помощью конечных моделей отображать бесконечную действительность, и не просто отображать, а отображать эффективно, т.е. достаточно правильно.

УПРОЩЕННОСТЬ МОДЕЛЕЙ

Первый фактор – это *упрощенность моделей*. Прежде всего отметим, что *сама конечность моделей делает их упрощенность неизбежной*, но, как мы впоследствии увидим, это ограничение не настолько сильно, как этого можно было ожидать (иерархичность моделей обладает потенциальной практической неограниченностью). Далее, гораздо более важным является то, что в человеческой практике упрощенность моделей является допустимой. К счастью, для любой цели оказывается вполне достаточным неполное, упрощенное отображение действительности. Более того, для конкретных целей такое упрощение является даже необходимым (а не только достаточным).

Вот как описывал В.И. Ленин эту необходимость на примере формирования модели движения:

“Мы не можем представить, выразить, смерить, изобразить движения, не прервав непрерывного, не упростив, угрубив, не разделив, не омертвив живого. Изображение движения мыслью есть всегда огрубление, омертвление, – и не только мыслью, но и ощущением, и не только движения, но и **всякого** понятия” [9, с. 233].

Еще один наглядный пример этого приводит Б. Рассел:

“Допустим, что вашего друга зовут м-р Джонс. Его очертания с физической точки зрения довольно неопределенны как потому, что он непрерывно теряет и приобретает электроны, так и потому, что каждый электрон (...) не имеет резкой границы (...). Поэтому внешние очертания м-ра Джонса имеют в себе нечто прозрачно-неосознанное, что никак не ассоциируется с видимой плотностью вашего друга. Нет никакой необходимости вдаваться в тонкости теоретической физики для того, чтобы показать, что м-р Джонс есть некая печальная неопределенность. Когда он стрижет ногти, то в этом процессе есть определенный, хотя и короткий, период времени, когда нельзя сказать продолжают ли еще обрезки его ногтей быть частью его самого или уже нет. Когда он ест баранью котлету, то можно ли с точностью установить момент, когда котлета становится частью его самого? Когда он выдыхает углекислый газ, то является ли этот последний частью его самого, пока не выйдет из его ноздрей? (...) Так или иначе, все же становится неясным, что еще является частью м-ра Джонса, а что не является” [16, с. 95].

Что именно из свойств такого объекта, как мистер Джонс, включать в его модель, а что – нет, зависит от целей моделирования; однако выбор цели определит, что можно и что нужно отбросить, в каком направлении упрощать модель по сравнению с отображаемым оригиналом. Упрощение является сильным средством для выявления главных эффектов в исследуемом явлении: это видно на примере таких моделей физики, как идеальный газ, непоглощающее зеркало, абсолютно черное тело, математический маятник, пружина без массы, конденсатор без утечки, абсолютно твердый рычаг и т.д.

Следующая причина вынужденного упрощения модели связана с *необходимостью опирания с ней*. За неимением методов решения нелинейного уравнения мы его линеаризуем; в других случаях искусственно уменьшаем размерность, заменяя переменные величины постоянными, случайные – детерминированными и т.д. Ресурсное, навязанное происхождение таких упрощений иллюстрируется тем, что по мере распространения ЭВМ и с развитием численных методов эти упрощения ликвидируются, что дает существенное продвижение в исследовании явлений и даже приводит к открытиям.

Есть и еще один, довольно загадочный, аспект упрощенности моделей. Почему-то оказывается так, что из двух моделей, одинаково хорошо описывающих данное явление, та, которая проще, оказывается ближе к истинной природе отображаемого явления. В истории науки имеется немало замечательных примеров этого, и, возможно, самый яркий из них – переход от геоцентрической модели Птолемея к гелиоцентрической модели Коперника. Ведь и геоцентрическая модель позволяла с нужной точностью рассчитать движения планет, предсказать затмения Солнца – хотя и по очень громоздким формулам, с переплетением многочисленных “циклов”. Этот пример выявляет разницу между эффективностью и правильностью (истинностью) моделей. У физиков имеется неформальный, эвристический критерий: если уравнение “красивое”, то оно, скорее всего, правильное – эстетическая оценка простоты и истинности модели. Можно предположить, что простота правильных моделей отражает некое глубинное свойство природы, и, видимо, именно это имел в виду И. Ньютон, говоря, что природа проста и не излишествует причинами вещей, или древние схоласты, которые подметили эвристически, что простота – печать истины.

Итак, упрощенность моделей основана как на свойствах мышления, ресурсов моделирования, так и на свойствах самой природы.

ПРИБЛИЖЕННОСТЬ МОДЕЛЕЙ

Второй фактор, позволяющий преодолевать бесконечность мира в конечном познании, – это *приближенность (приблизительность) отображения действительности с помощью моделей*. Конечность и упрощенность моделей также можно интерпретировать как приближенность (пример расплывчатости терминов естественного языка), но мы хотим разделить качественные различия между оригиналом и моделью (их и будем связывать с конечностью и упрощением) и такие их различия, которые допускают количественное (“больше – меньше”) или хотя бы ранговое (“лучше – хуже”) сравнение: этот аспект и свяжем с термином “*приближенность*”.

Приближенность модели может быть очень высокой (так, некоторые подделки произведений искусства даже эксперты не могут отличить от оригинала; сильное впечатление производят голограммические фотографии предметов; у английского фельдмаршала Монтгомери во время войны был двойник, появление которого на разных участках фронта намеренно дезинформировало разведку немцев, и т.д.), в других случаях приближенность модели видна сразу и может варьироваться (например, карты местности в разных масштабах); но во всех случаях модель – это другой объект, и различия неизбежны (единственной совершенно точной картой страны является сама эта страна).

Различие само по себе не может быть ни большим, ни малым: само по себе оно либо есть, либо его нет. Величину, меру, степень приемлемости различия мы можем ввести только сопоставляя его с целью моделирования. Скажем, точность наручных часов, вполне достаточная для бытовых целей, совершенно недостаточна при регистрации спортивных рекордов или для целей астрономии.

АДЕКВАТНОСТЬ МОДЕЛЕЙ

Модель, с помощью которой успешно достигается поставленная цель, будем называть *адекватной* этой цели. Подчеркнем, что введенное таким образом понимание адекватности не полностью совпадает с требованиями полноты, точности и правильности (истинности): *адекватность означает, что эти требования выполнены не вообще* (так сказать, безмерно),

а лишь в той мере, которая достаточна для достижения цели. Как уже отмечалось, геоцентрическая модель Птолемея была неправильной, но адекватной в смысле точности описания движения планет. Следует добавить, что эта модель не была лишена истинности вовсе: ведь и Солнце, и планеты действительно движутся относительно Земли. В противоположность этому христианская модель поведения “Десять заповедей” адекватна по отношению к целям морали (чем и можно объяснить ее живучесть), но не обладает истинностью своего якобы божественного происхождения. Шаман, объясняющий свое успешное врачевание силами духов, также предлагает адекватную, но ложную модель.

В ряде случаев удается ввести некоторую меру адекватности модели, т.е. указать способ сравнения двух моделей по степени успешности достижения цели с их помощью. Если к тому же такой способ приводит к количественно выражаемой мере адекватности, то задача улучшения модели существенно облегчается. Именно в таких случаях можно количественно ставить вопросы об идентификации модели (т.е. о нахождении в заданном классе моделей наиболее адекватной), об исследовании чувствительности и устойчивости моделей (т.е. о зависимости меры адекватности модели от ее точности), об адаптации моделей (т.е. подстройке параметров модели с целью повышения адекватности) и т.п.

Подведем итог

Summary

Различие между моделью и оригиналом вызвано тем, что мы можем отображать реальность лишь в конечном числе отношений, конечными средствами; в результате упрощение и приближенность модели необходимы, неизбежны; но замечательное свойство мира и нас самих состоит в том, что этого достаточно для человеческой практики.

The difference between a model and its origin stems from the fact that we are able to represent reality only in a finite number of relations and with finite resources. As a result, the simplification and approximate ness of models are necessary and unavoidable; but an amazing property of the world and ourselves is that this is enough for human activity.

§ 2.6. СООТВЕТСТВИЕ МЕЖДУ МОДЕЛЬЮ И ДЕЙСТВИТЕЛЬНОСТЬЮ: СХОДСТВО

В предыдущем параграфе мы обсуждали соответствие между моделью и оригиналом с позиций различия между ними. Рассмотрим теперь более трудный вопрос – о сходстве между ними. Поскольку различия между моделью и реальностью принципиально неизбежны и неустранимы, существует предел истинности, правильности наших знаний, сконцентрированных в моделях. Является ли этот предел вечным, принципиально неотодвигаемым или имеется возможность неограниченно увеличивать сходство наших моделей с реальностью? Или, говоря философским языком, доступна ли объективная истина субъективному познанию?

Различные философские учения отвечали на этот вопрос по-разному. Субъективные идеалисты вообще отрицали наличие какой-либо реальности вне сознания субъекта либо считали ее продуктом сознания. Античные софисты считали и саму реальность, и всякое знание о внешней реальности относительными, субъективными и индивидуальными; разные мнения об одном, принадлежащие разным субъектам, считались одинаково истинными (“что кому как является, так оно и есть для того, кому является”). Агностики признавали ощущения единственной реальностью и потому считали, что познанию доступны лишь явления, а сущности и закономерности реального мира принципиально непознаваемы. Диалектический материализм освещает этот вопрос в плане отношений между относительной истиной, т.е. объективным, но не полным содержанием сегодняшнего человеческого познания, и истиной абсолютной: истина относительная может пополняться, развиваться, приближаясь к истине абсолютной, и принципиального предела этому приближению не существует. Степень пра-

**TRUE, CORRECT,
PROPER**истинный
**FALSE, ERRONE-
OUS**ложный
CAUSE, REASON,
MOTIVE причина
**CORRESPOND-
ENCE, ACCORD-
ANCE**соответствие
EVOLUTIONэволю-
ция

Реальные модели
**служат заместите-
лями оригинала**
благодаря своему
**подобию с ним. Воз-
можны три вида**
подобия: прямое
(устанавливаемое
через физическое
взаимодействие),
**косвенное (устанав-
ливаемое через ана-
логию, т.е. через об-
щую абстрактную**
модель) и условное
(устанавливаемое
через соглашение). В
каждой модели есть
доля истины, т.е. в
чем-то любая мо-
дель правильно от-
ражает оригинал.

Степень истинности
модели проявляется
при правильном
полученным результатам

вильности относительных истин, наличие в моделях элементов, не обладающих истинностью, обнаруживаются в практике человеческого общества.

ИСТИННОСТЬ МОДЕЛЕЙ

Весь исторический опыт науки, техники, общественной практики подтверждает верность диалектико-материалистической теории познания. Например, в современной теории тепловых процессов многое сохранилось со времен С. Карно, но от модели теплорода как особой субстанции – носителя тепла пришлось отказаться; в основе современной электродинамики лежат уравнения Максвелла, но от механической максвелловской модели эфира до современных представлений о физическом вакууме – огромная дистанция [12].

Рассмотрение проблемы истинности знаний с философских позиций требует глубокого и всестороннего анализа многих аспектов даже на примере одной науки (так, о физике см. [21]). Наша задача гораздо конкретнее: обратить внимание на сочетание истинного и предполагаемого (могущего быть как верным, так и неправильным) во всех моделях.

Еще раз подчеркнем, что об истинности, правильности или ложности модели самой по себе говорить бессмысленно: только в практическом соотнесении модели с отображаемой еюатурой выявляется степень истинности. При этом изменение условий, в которых ведется сравнение, весьма существенно влияет на его результат: именно из-за этого возможно существование двух противоречивых, но “одинаково” истинных моделей одного объекта. Яркий пример этого – волновая и корпускулярная модели света или электрона; эти модели различны, противоположны и истинны, каждая в своих условиях. Важно отметить, что каждая модель явно или неявно содержит условия своей истинности, и одна из опасностей практики моделирования состоит в применении модели без проверки выполнения этих условий. В инженерной практике такая ситуация встречается чаще, чем принято думать. Например, для определения пропускной способности различных каналов связи нередко используют формулу Шэннона – Таллера, хотя она верна только для конкретных (гауссовых) каналов. Для обработки экспериментальных данных часто употребляют статистические процедуры, не проверяя условий их применимости (скажем, нормальности или независимости). Иногда это делается вынужденно (не всякое условие бывает возможным проверить), но тогда и к полученным результатам должно быть осторожное, условное отношение, что, к сожалению, не всегда имеет место. Такие ситуации выдвинули перед исследователями специальную проблему – создание моделей, применимость которых сохраняется в некотором диапазоне условий; например, в математической статистике этому соответствуют непараметрические и robustные процедуры обработки данных в теории управления и прикладной математике – исследование устойчивости моделей и регуляризация алгоритмов и т. п.

СОЧЕТАНИЕ ИСТИННОГО И ЛОЖНОГО В МОДЕЛИ

Еще один важный аспект соотношения истинного (т.е. определенно известного и правильного) с предполагаемым (т.е. возможным, но не обязательно действительным) при построении моделей состоит в том, что ошибки в предположениях имеют разные последствия для прагматических и познавательных целей. Если ошибки в предположениях вредны и даже губительны при использовании прагматических моделей, то при создании познавательных

моделей поисковые предположения, истинность которых еще предстоит проверить, – единственный способ оторваться от фактов. Роль гипотез в науке настолько важна, что образно можно сказать, что вся научная работа состоит в выдвижении и проверке гипотез. А. Эйнштейн писал:

“Воображение важнее знания, ибо знание ограничено. Воображение же охватывает все на свете, стимулирует прогресс и является источником его эволюции”.

Среди других аспектов проблемы правильности моделей, истинности знаний для целей системного анализа важен аспект явной рефлексии уровня истинности: что известно точно, достоверно; что – с оцениваемой степенью неопределенности (например, с известной вероятностью для стохастических моделей или с известной функцией принадлежности для расплывчатого описания); что – с неопределенностью, не поддающейся оценке; что может считаться достоверным только при выполнении определенных условий; наконец – что известно о том, что неизвестно.

Подведем итог

Summary

Главная ценность моделей как формы знаний состоит в том, что они содержат объективную истину, т.е. в чем-то правильно отображают моделируемое. Однако кроме безусловно истинного содержания в модели имеется и условно-истинное (т.е. верное лишь при определенных условиях), и предположительно-истинное (т.е. условно-истинное при неизвестных условиях), а следовательно, и ложное. При этом в каждой конкретных условиях неизвестно точно, каково же фактическое соотношение истинного и ложного в данной модели. Ответ на этот вопрос

The main virtue of models as a representation of knowledge is that they contain an objective truth; i.e., they correctly represent modeled objects in certain respects. But besides a definitely correct content, each model contains a conditionally correct (i.e., true only under certain conditions) and a wrong content. In any given circumstance the real proportions between correct and erroneous content in a given model are not fully known. Only practice can give an answer to this. But in any case a model is principally poorer than the original; this is its fundamental property.

дает только практика. Однако в любом случае модель принципиально беднее оригинала, это ее фундаментальное свойство.

§ 2.7. О ДИНАМИКЕ МОДЕЛЕЙ

Как и все в мире, модели проходят свой жизненный цикл: они возникают, развиваются, сотрудничают или соперничают с другими моделями, уступают место более совершенным. Одни модели живут дольше отдельных людей, и тогда этапы жизненного цикла моделей изучаются в виде истории той или иной отрасли знаний или деятельности (например, истории физики, истории авиации, истории гончарного искусства и т.д.). Жизненный цикл других моделей должен быть обязательно завершен в обозримый срок, и тогда перевод модели от этапа к этапу становится технологическим действием, должен выполняться как можно эффективнее (быстрее, лучше, дешевле). Как мы уже знаем, это невозможно без моделирования самого процесса моделирования, т.е. алгоритмизации моделирования.

Наиболее полно необходимость алгоритмизации моделирования осознана там, где проблема эффективности действия стоит особенно остро: в проектной деятельности [5], в исследо-

довании операций [3], в изобретательском поиске [2], в создании АСУ [14], в имитационном моделировании [22].

Остановимся на важнейших причинах и закономерностях динамики моделей. Прежде всего очевидно, что процесс моделирования структурирован, организован, состоит из последовательности этапов. Этапы отличаются качественно, конкретными целями и средствами и должны выполняться в определенной очередности. Например, при конструировании новой технической системы ее модель развивается от воплощения в виде результатов предыдущей научно-исследовательской работы по стадиям технического задания, технического проекта, рабочего проекта, опытного образца, мелкой серии до модели, предназначенной для промышленного выпуска. Другой пример дают рекомендации по последовательности этапов имитационного моделирования: формирование целей моделирования – построение абстрактной модели – создание имитационной реальной модели – ее исследование – обработка и интерпретация результатов.

СЛОЖНОСТИ АЛГОРИТМИЗАЦИИ МОДЕЛИРОВАНИЯ

Стремление к совершенству требует алгоритмизации, и для многих исследователей исходным стремлением было найти наиболее эффективную последовательность этапов в работе с моделями. Однако здесь обнаружилось, что в практике моделирования чаще всего не удается строго выдержать рекомендуемую последовательность действий. Более того, стало очевидным, что вообще не существует какого-то единого, пригодного для всех случаев алгоритма работы с моделями. Это вызвано разными причинами.

Во-первых, как мы знаем (см. § 2.4), модель функционирует в культурной среде, и конкретное окружение каждой модели может настолько отличаться, что опыт работы с одной моделью не может без изменений переноситься на другую.

Во-вторых, требования, предъявляемые к модели, противоречивы: полнота модели противоречит ее простоте, точность модели – ее размерности, эффективность – затратам на реализацию. Многое в истории данной модели зависит от того, какой именно компромисс выбран между этими противоречащими критериями.

В-третьих, с самого начала невозможно предусмотреть все детали того, что произойдет в будущем с любой моделью. Моделирование призвано устраниć неопределенность, но существует неопределенность и в том, что именно надо устранять. Начальные цели впоследствии могут оказываться неполными. Например, по результатам испытаний опытного образца часто приходится вносить изменения в техническое задание и снова возвращаться к этапам проектирования образца. Приведем другой пример: после интерпретации результатов имитационного моделирования цели уточняются, в модель вносятся изменения и моделирование повторяется. Недостатки модели проще и легче обнаружить и исправить в ходе моделирования, чем предусмотреть их заранее. Это еще одна причина динамичности моделей.

Однако среди всех причин невозможности полной алгоритмизации процесса моделирования особо выделяются еще две, последние по счету, но не по важности. Они различаются тем, насколько большой вес придается роли отдельного человека в развитии модели.

Желая выявить алгоритм моделирования, естественно обратиться к инженерной и научной практике: ведь именно в этих видах деятельности наиболее сильно выражено стремление к формализации, точности, определенности. Появляется надежда на обнаружение закономерностей моделирования при наблюдении работы инженеров и ученых над созданием моделей. Оказалось, что даже в этой области

“любой набор правил для разработки моделей в лучшем случае имеет ограниченную полезность и может служить лишь предположительно в качестве каркаса будущей модели или отправного пункта в ее построении” [22, с. 33].

Более того, даже в чистой математике разработка моделей невозможна без эвристического, творческого, неформального начала [1]. К. Жаблон и К. Симон анализируют эту ситуацию следующим образом:

“Как же тогда можно убедиться, что та или иная реальная задача может быть простым способом смоделирована? Только если найдена простая модель! Этот ответ не шутка. Теория разрешимости Геделя и Клини дает такой результат: не существует универсальной программы, которая при рассмотрении реальной задачи могла бы решить, существует ли подходящая модель, и которая могла бы ее построить” [6].

Обсуждение вопроса о соотношении формальных и эвристических приемов в процессе построения моделей завершим словами Р. Шэннона:

“Искусством моделирования могут овладеть те, кто обладает оригинальным мышлением, изобретательностью и находчивостью, равно как и глубокими знаниями систем и физических явлений, которые необходимо моделировать. Не существует твердых и эффективных правил относительно того, как надо формулировать задачу в самом начале процесса моделирования, т.е. сразу же после первого знакомства с ней. Не существует и магических формул для решения при построении модели таких вопросов, как выбор переменных, и параметров, соотношений, описывающих поведение системы, и ограничений, а также критериев оценки эффективности модели. Помните, что никто не решает задачу в чистом виде, каждый оперирует с моделью, которую он построил, исходя из поставленной задачи. Все эти соображения должны помочь читателю правильно разобраться в особенностях моделей и в некоторых вопросах искусства моделирования” [22, с. 35].

Итак, говоря о том, как отдельные люди осуществляют построение моделей, мы должны признать, что в этом процессе кроме осознанных формализованных, технических и научных приемов огромную, решающую роль играет то, что мы называем творчеством, интуитивным искусством. В этом одна из главных причин невозможности полной формализации процесса моделирования.

ЕСТЕСТВЕННАЯ ЭВОЛЮЦИЯ МОДЕЛЕЙ

Еще с одним аспектом этой невозможности мы сталкиваемся тогда, когда моделирование осуществляется не отдельным индивидуумом, а коллективно, не осознанно, а как бы стихийно, естественным эволюционным путем. Проявлением такого способа моделирования является фольклор. Другие примеры дают мозговой штурм и синектика (о которых мы еще будем говорить в § 9.6 как о методах коллективного системного анализа). Естественное развитие моделей происходит не только у людей. Интересный пример этого наблюдали орнитологи. В некоторых европейских странах принято доставлять молоко заказчикам, расставляя по утрам бутылки прямо у дверей или калиток домов. Было замечено, что в одном районе большого города синицы научились пробивать клювом сделанные из фольги крылечки бутылок и отпивать молоко. Вскоре это умение синиц распространилось на весь город и за его пределы.

Это явление развития модели “естественным путем”, как бы “самостоятельной жизни” модели в подходящей культурной среде, наблюдается не только в действительно естественных условиях (скажем, в языковой практике – слухи и анекдоты), но и в более специфических условиях инженерной практики. Например, в имитационном моделировании все чаще начинают с реализации самой простой модели самой сложной ситуации (минимаксный подход), а затем ее усложняют и отрабатывают. В проектировании технических систем все чаще отступают от традиционного надежного, но длительного пути “НИР – ОКР – производство” и прибегают к созданию реальной, пусть плохой, но действующей системы-макета, а затем ее “доводят”, постепенно улучшая. Пожалуй, наиболее ярко это проявляется сейчас в разработке систем программного обеспечения ЭВМ. Однако эволюционное развитие научно-технических моделей происходит не только тогда, когда мы “организуем” его, но и независимо от нашего желания. Эту тенденцию на примере истории техники очень последовательно проследил С. Лем [8].

Итак, еще один путь неформализуемого развития модели – это ее эволюционная динамика в среде, которой она ингерентна (см. § 2.4).

Подведем итог

Основные положения данного параграфа заключаются в следующем:

как и любые реальности, модели также претерпевают изменения, проходят свой “жизненный цикл”;

пользование готовыми моделями обычно не требует знания их истории, но знание закономерностей развития моделей полезно и необходимо разработчикам;

на этапе построения моделей решающую роль играют неформализуемые эвристические способности человеческого интеллекта;

существует эволюционный тип развития моделей, при котором вклад отдельных личностей мал, случаен, но в целом приводит к прогрессу модели;

неформализуемость основных этапов построения модели, превращающая этот вид деятельности в искусство, не означает, что наука вообще и алгоритмизация в частности непричастны к этому, что нельзя дать никаких полезных и конкретных рекомендаций, повышающих эффективность моделирования;

наука моделирования состоит в разделении процесса моделирования на этапы (см., например, [20]), детальном изучении каждого этапа и его описания с максимально возможной степенью формализации, рассмотрения и оценивания вариантов;

в целом моделирование является неразделимым сочетанием, сплавом науки и искусства.

Summary

The main principles set forth in this section are as follows:

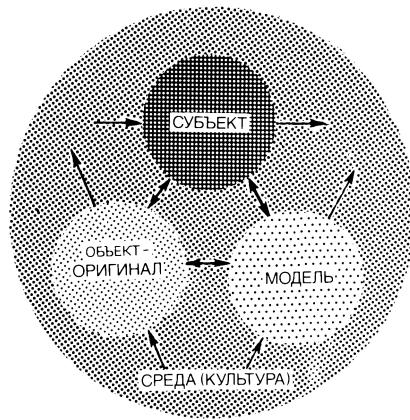
- Like all of reality, models are subject to change; they possess a life cycle.
- One does not need to know the history of a model in order to use it; but a knowledge of the general laws governing model development is useful and even necessary for designers of models.
- In model design a decisive role is played by the nonformalizable heuristic capacities of human intelligence.
- There is also an evolutionary mode of model development, in which progress is made by small and random contributions of many people involved.
- The fact that many important stages of modeling are nonformalizable makes modeling activity an art; but this does not mean that science and algorithmization have nothing to do with modeling, that useful, concrete recommendations and methods facilitating modeling cannot be given.
- The science of modeling encompasses the division of modeling into stages (cf., for example [20]), the detailed study of each stage with as formalized a description as possible, and the consideration and assessment of alternative models.
- The process of modeling is an inseparable fusion of science, craft and art.

ЗАКЛЮЧЕНИЕ

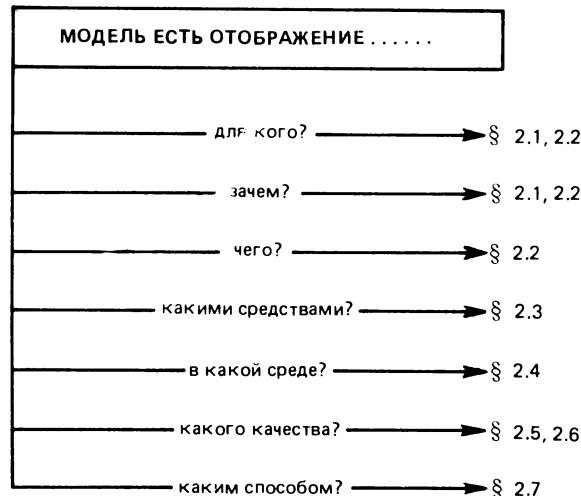
Обратим внимание на то, что понятие модели оказалось невозможным ограничить только тем самым, что мы непосредственно называем моделью. Если мы хотим включить в это понятие все существенное, т.е. учесть наиболее общие, важные свойства модели и основные условия проявления этих свойств, то мы вынуждены рассматривать все взаимодействия между всеми “сторонами-участниками” моделирования. Таких “участников”, в самом укрупненном представлении, четыре (рис. 2.2): “субъект” – инициатор моделирования и/или пользователь его результатов; “объект-оригинал” – предмет моделирования; “модель” – отображение объекта; “культура” – среда, в которой находятся и с которой взаимодействуют все остальные “участники”.

Все сказанное в предыдущем абзаце и изображенное на рис. 2.2 является моделью модели. Для удобства приведем схему построения рассуждений главы в целом. Если в центр всех рассуждений поставить саму модель – концентратор нужной имеющейся информации об объекте-оригинале (а ведь для каких-то целей можно сделать иначе – мы теперь знаем, что всякая модель, в том числе и модель модели, является целевой), то общая схема данной главы может быть проиллюстрирована рис. 2.3.

Если же акцент переместить с того, что такая модель, на то, как она функционирует, то следует отметить один из многих аспектов этого вопроса, а именно: иерархическую организованность моделей, их разноуровневость, включенность одних в другие. Такая организация мира моделей имеет множество важных последствий, которые мы упоминали в разной степени; наиболее важные из них – соответствие модельных построений и отображаемых ими структур природы (см. гл. 1) и ингерентность моделей (см. § 2.4). Сейчас же стоит обратить внимание на то, что иерархический способ построения модели позволяет ей успешно функционировать при различной подробности, детализации, продвинутости, формализованности ее составляющих частей. Оказывается возможным в единой структуре направить на общую цель самые разнообразные модели – от языковых, содержательных, до математических, формальных, и даже модели, отличающиеся не только степенью формализации, но и самими языками описания. Например, в территориальных АСУ основными языками описания любой подсистемы (до уровня производственного объекта) являются языки: финансовый, производственный и социально-ценностный [15]. Эта же особенность иерархического построения модели проявляется на всех уровнях, вплоть до



2.2 Схема, отображающая модель как многоместное отношение



2.3 Последовательность тем, рассмотренных в гл.2

уровня математических моделей. Так, Розенблют и Винер, говоря о математических моделях, подчеркивают, что *формальная модель может быть разнородной совокупностью элементов, часть из которых изучена детально, т.е. конкретно и структурно, а другая часть – только исходя из соображений ее общей работоспособности, т.е. обобщенно и функционально* [18].

Известно много определений модели. Это объясняется тем, что определение также модель; разным целям требуется сопоставить разные аспекты моделей, поэтому различие в определениях модели неизбежно. Еще одна причина различия – ингерентность. Даже одинаковую мысль на разных языках приходится выражать по-разному.

Если попытаться подытожить все сказанное в данной главе в виде еще одного определения модели, то оно может выглядеть так:

Модель есть отображение: целевое; абстрактное или реальное, статическое или динамическое; ингерентное; конечное, упрощенное, приближенное; имеющее наряду с безусловно-истинным условно-истинное и ложное содержание; проявляющееся и развивающееся в процессе его создания и практического использования.

Если же это определение покажется чересчур громоздким, то его вполне можно заменить кратким эквивалентом: *модель есть системное отображение оригинала*. Это сокращение мы приводим не столько для определения модели, сколько для иллюстрации того, насколько богато понятие системности.

ЛИТЕРАТУРА

1. Адамар Ж. Исследование процесса изобретения в области математики. – М.: Сов. радио, 1970.
2. Альтшулер Г.С. Найти идею. – Новосибирск: Наука, 1986.
3. Дегтярев Ю.И. Исследование операций. – М.: Высш. шк., 1986.
4. Демьяненков В.З. Проблема понимания как предмет вычислительной лингвистики. – В кн.: Лингвистическое обеспечение информационных систем, – М., 1987.
5. Джонс Дж. К. Методы проектирования. – М.: Мир, 1986.
6. Жаблон К., Симон Ж.-К. Применение ЭВМ для численного моделирования в физике. – М.: Наука, 1983.
7. Кемени Дж., Снелл Дж. Кибернетическое моделирование. Некоторые приложения. – М.: Сов. радио, 1972.
8. Лем С. Сумма технологий, – М.: Мир, 1968.
9. Ленин В.И. Полн. собр. соч. Т. 29.
10. Маркс К., Энгельс Ф. Соч. 2-е изд. Т. 23.
11. Неуймин Я.Г. Модели в науке и технике. – Л.: Наука, 1984.
12. Подольный Р. Нечто по имени нечто. – М.: Знание, 1983.
13. Пойа Д. Математика и правдоподобные рассуждения. – М.: ИЛ, 1957.
14. Перегудов Ф.И. Основы системного проектирования АСУ организационными комплексами. – Томск: ТГУ, 1984.
15. Перегудов Ф.И. и др. Системное проектирование АСУ хозяйством области – М.: Статистика, 1977.
16. Рассел Б. Человеческое познание. М., 1957.
17. Растрогин Л.А. Кибернетика и познание. – Рига: Зинатне, 1978.
18. Розенблют А., Винер Н. Роль моделей в науке. – В кн.: Неуймин Я.Г. Модели в науке и технике. – Л.: Наука, 1984.
19. Седов Л.И. Теория подобия и размерности в механике. – М.: ГИТТЛ, 1954.
20. Хагер Н. Этапы формирования моделей. – В сб.: Эксперимент. Модель. Теория. – М. – Берлин: Наука, 1982.
21. Чешев В.В. Проблема реальности в классической и современной физике. – Томск: ТГУ, 1984.

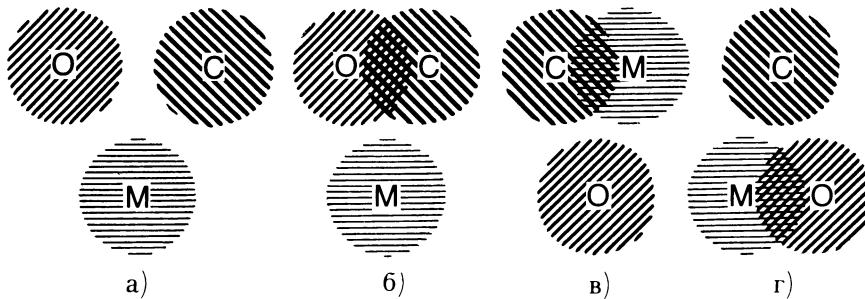
22. Шэннон Р. Имитационное моделирование систем – искусство и наука. – М.: Мир; 1978.

23. Штрафф В.А. Моделирование и философия. – М. – Л.: Наука, 1966.

УПРАЖНЕНИЯ

§ 2.2

- Обсудите различия в моделях лошади с позиций крестьянина, жокея, кавалериста, скульптора, коневода, повара. Задача обсуждения – иллюстрация целевого характера моделей.
- В каких обстоятельствах карта местности является познавательной, а в каких – прагматической?



2.4 ————— Различные варианты физического совмещения отдельных компонент процесса моделирования

ческой моделью?

§ 2.3

- Попробуйте рассмотреть ваше любимое стихотворение как модель. Обсудите реальные и абстрактные аспекты дорожных знаков.
- Если условное подобие определяется соглашением, то чем ограничена свобода выбора моделей условного подобия?

§ 2.4

- На рис. 2.2 субъект (С), объект моделирования (О) и модель (М) изображены отдельно, что можно интерпретировать как их раздельное физическое существование. Однако могут иметь место ситуации, когда они полностью или частично совмещены (например, человек изучает свой собственный организм, скажем, измеряет себе температуру). На рис. 2.4 приведено несколько вариантов такого совмещения. Придумайте реальную ситуацию, отвечающую каждому варианту схемы моделирования.

§ 2.5

- “Я знаю, что ты знаешь, что я знаю”. Сколько ветвей и уровней в этой иерархии моделей? Попробуйте изобразить это графом.
- Экстрасенс, делая пассы руками, снимает головную боль у пациента и объясняет это взаимодействием своего биополя с биополем пациента. Обсудите соотношение адекватности и истинности модели, предложенной экстрасенсом.
- Приведите примеры конечности, упрощенности и приближенности моделей.

§ 2.6

- (Для знающих физику.) Обсудите, в каких условиях проявляются волновые свойства элементарных частиц и чем эти условия отличаются от тех, когда корпускулярная модель более адекватна.
- Попробуйте в знакомой вам модели (например, в законе Ома или в описании купленной вами вещи) отыскать безусловно-истинное, условно-истинное и ложное содержание.

- Тема для обсуждения: насколько хорошо может простая модель отразить сложное явление? Аргументы в дискуссии подкрепляйте примерами. В качестве одного из таких примеров можете рассмотреть следующий (не претендующий на полную серьезность).
 - “Теория счастья”. Для построения содержательной теории некоторого явления совсем не обязательно его определять; достаточно ввести его количественную меру. Например, в теории информации не рассматривается, что такая информация, а все содержание теории основано на количестве информации. В теоретической механике не обсуждается вопрос о том, что такое движение; в электродинамике – что такое электричество и магнетизм; этим теориям достаточно количественных характеристик изучаемых явлений. То же имеет место и для предлагаемой “теории счастья”. В качестве аксиом предлагаются следующие положения:
 - 1°. Каждый индивид наблюдает окружающую его реальность (P) и имеет свое представление о желательном состоянии реальности, т.е. идеал (I). (Это может быть идеал любого уровня – идеал вкусного обеда, любимой работы, человеческого общения, настоящего друга, справедливости и т.п.)
 - 2°. Мерой счастливости индивида можно считать функцию, монотонно зависящую от степени совпадения P и I : чем сильнее они различаются, тем менее счастлив индивид.
 - 3°. Каждому индивиду присущее природное стремление к повышению своей счастливости. Этих аксиом достаточно для вывода следующих суждений. Существуют разные способы повышения счастливости, и все из них имеют место в жизни:
 - изменяя реальность, приближать ее к идеалу ($P \rightarrow I$). В зависимости от идеалов имеем спектр индивидов, идущих по этому пути – от революционеров, реформаторов, просветителей, через “ рядовых ” тружеников, до (к сожалению) нарушителей законов;
 - сближать реальность и идеал с помощью изменения идеала ($I \rightarrow P$). Снова имеем спектр лиц, следующих такой возможности: от конформистов до циников, но в этом ряду есть и обычные люди, отказывающиеся от несбыточных идеалов;
 - подменять реальность ее иллюзией, уменьшая тем самым расхождение реальности и идеала. Здесь очевидны роль алкоголя, наркотиков; но не стоит ли в этом же ряду “ массовая культура ”? А “ ложь во спасение ”, например в отношениях между врачами и смертельно больными пациентами?
 - сохранять баланс при привитии идеалов: идеал должен быть потенциально достижимым.
- § 2.7
- Тема для обсуждения: культура как “вторая природа” (возможные аспекты дискуссии: объективные процессы в развитии культуры; роль моделей в культуре; возможны ли резкие ускорения в развитии культуры; формы взаимодействия модели с культурой и т.д.).
 - В § 2.7 перечислены причины, способствующие динамике, изменчивости моделей. Приведите примеры, иллюстрирующие каждую из этих причин.

Вопросы для самопроверки

1. Что заставляет нас пользоваться моделями вместо самих моделируемых объектов?
2. Какие функции выполняют модели во всякой целесообразной деятельности? Можно ли осуществлять такую деятельность без моделирования?
3. В каком смысле модель можно назвать “окном”, сквозь которое мышление “видит” мир?
4. Каково главное отличие между познавательной и прагматической моделями?
5. Какими средствами располагает человек для построения моделей?
6. Что необходимо для перехода от моделей в терминах естественного языка к математическим моделям?
7. Что общего между моделью и оригиналом при косвенном подобии?
8. Почему знаки можно назвать материальными по форме и абстрактными по существу моделями?
9. Что такое ингерентность модели?
10. В каком смысле можно говорить о конечности моделей?
11. В чем различие между адекватностью и истинностью модели?
12. Каковы причины того, что любая модель со временем изменяется?

INPUTS	
<i>входы (системы)</i>	
OUTPUTS	
<i>выходы (системы)</i>	
PROBLEM	
<i>проблема</i>	
COMPLEXITY	
<i>сложность</i>	
BLACK-BOX	
<i>черный ящик</i>	

СИСТЕМЫ МОДЕЛИ СИСТЕМ

ГЛАВА ТРЕТЬЯ

§ 3.1. МНОЖЕСТВЕННОСТЬ МОДЕЛЕЙ СИСТЕМ

Центральной концепцией теории систем, кибернетики, системного подхода, всей системологии является понятие системы. Поэтому очень многие авторы анализировали это понятие, развивали определения системы до различной степени формализации. Например, в [8] собрано 35 (!) различных определений системы. Изложенное в предыдущей главе понимание модели позволяет легко объяснить эту множественность: определение – это языковая модель системы, и, следовательно, различия целей и требований к модели приводят к разным определениям. Кроме того, разная языковая среда, в силу ингерентности модели, также обуславливает видоизменение определений (уже только поэтому отличаются, например, философское и математическое определения системы).

В данной главе мы изложим понятие системы в его развитии. На примере этого понятия проиллюстрируем идею поэтапного уточнения, конкретизации модели, ее постепенного “движения” от уровня естественно-языковой формы до математической (об этом было сказано в § 2.3). Поэтому формулировки определений системы выглядят просто как фиксация того или иного этапа процесса развития понятия системы. К тому же и усвоение такого непростого понятия, как система, можно облегчить, если начинать с его простейших версий, постепенно расширяя и обобщая его смысл.

§ 3.2. ПЕРВОЕ ОПРЕДЕЛЕНИЕ СИСТЕМЫ

Начнем с рассмотрения искусственных, т.е. создаваемых человеком, систем. Как уже отмечалось, любая деятельность человека носит целенаправленный характер. Наиболее четко это прослеживается на примере трудовой деятельности.

ПРОБЛЕМЫ И СИСТЕМЫ

Цели, которые ставит перед собой человек, редко достижимы только за счет его собственных возможностей или внешних средств, имеющихся у него в данный момент. Такое стечие обстоятельств называется **проблемной ситуацией**. Примером такой ситуации, требующей создания АСУ, является случай, когда обычные способы сбора и переработки информации не обеспечивают необходимой полноты и быстроты ее обработки, что значительно снижает качество принимаемых управлений решений.

Проблемность существующего положения осознается в несколько “стадий”: от смутного ощущения, что “что-то не так”, к осознанию потребности, затем к выявлению проблемы, наконец, к формулировке цели. Цель – это субъективный образ (абстрактная модель) несуществующего, но желаемого состояния среды, которое решило бы возникшую проблему.

Система (искусственная) есть средство достижения цели. Однако соответствие цели и системы неоднозначно: в чем-то разные системы могут быть ориентированы на одну цель; одна система может иметь (и, как правило, имеет) несколько разных целей.

Модель типа “черный ящик” отображает только связи системы со средой, в виде перечня “входов” и “выходов”.

Трудность построения модели “черного ящика” по требости, затем к выявлению проблемы, наконец, к формулировке цели. Цель – это субъективный образ (абстрактная модель) несуществующего, но желаемого состояния среды, которое решило бы возникшую проблему.

Вся последующая деятельность, способствующая решению этой проблемы, направлена на достижение поставленной цели. В дальнейших главах мы будем подробно рассматривать этапы такой деятельности, а сейчас лишь укрупненно в общих чертах опишем ее как действия, направленные на отбор из окружающей среды объектов, свойства которых можно использовать для достижения цели, и на объединение этих объектов надлежащим образом, т. е.

как работу по созданию того, что мы и будем называть системой. Другими словами, *система есть средство достижения цели*. Это и есть первое определение системы.

Приведем несколько упрощенных примеров систем, предназначенных для реализации определенных целей.

№	Цель	Система
1.	В произвольный момент указать время	Часы
2.	Обеспечить выпечку хлеба в заданном ассортименте для значительного количества людей	Пекарня
3.	Передать зрительную информацию в звуковом сопровождении на большие расстояния практически мгновенно	Телевидение
4.	Обеспечить быстрое перемещение большого числа людей по их желанию в пределах города	Городской транспорт

Полезно разобраться, в чем именно состоит упрощенность этих примеров. Прежде всего обратим внимание на то, что здесь речь идет об уже существующих и общеизвестных системах. Тем не менее далеко не просто сформулировать их цели так, чтобы имелось действительно очевидное соответствие между целями и системами. Например, только слова “практически мгновенно” в примере 3 отличают цель телевидения от цели кино или пересылки видеокассет.

СЛОЖНОСТИ ВЫЯВЛЕНИЯ ЦЕЛЕЙ

В примере 4 цель городского транспорта сформулирована весьма грубо: придется сделать много дополнений, чтобы отразить маршрутность транспорта или чтобы в городской транспорт вписалось такси. Одна из причин подобных трудностей состоит в том, что между целью (абстрактной и конечной моделью) и реальной системой нет и не может быть однозначного соответствия: для достижения заданной цели могут быть избраны разные средства – системы (к этому непростому моменту мы еще не раз будем возвращаться); с другой стороны, заданную реальную систему можно использовать и для других целей, прямо не предусмотренных при ее создании (например, действия армии или городского транспорта в момент стихийного бедствия или использование королевской печати для колки орехов, как рассказал об этом М. Твен в “Принце и нищем”).

Если так не просто обстоит дело с формулировкой целей уже существующих систем, то тем более осторожно, тщательно, продуманно, условно (осознанно системно) следует подходить к формулировке целей еще не существующих, а только проектируемых систем. Известны случаи, когда созданная система полностью отвечала поставленным целям, но совершенно не удовлетворяла тех, кто эти цели формулировал. Такие случаи дали повод для серии анекдотов из инженерного фольклора о неосмотрительно сформулированном техническом задании.

В инженерной практике момент постановки целей (формулировки технического задания) – один из важнейших этапов создания систем. Специалисты отдают себе отчет в сложности этого этапа и на деле сугубо формально трактуют его результат только тогда, когда наступает юридическая ответственность за неудачу. Обычно же цели уточняются итеративно, с многократными изменениями и дополнениями. Мы рассмотрим этот момент более подробно при обсуждении техники системного анализа (см. § 9.3).

Все отмеченное не изменяет и тем более не противоречит смыслу первого определения системы. Система есть средство достижения цели, средство решения проблемы. Акцентируем внимание на этом моменте, используя образное выражение “без проблемы нет системы”.

Другой аспект первого определения системы состоит в его конструктивности. Оно не только отвечает на вопрос “зачем нужна система?”, но и ориентирует при решении вопроса, следует или не следует включать данный объект из окружающей среды в состав системы: да, если его свойства могут быть использованы для достижения цели. Поэтому можно сформулировать еще одно образное выражение, подчеркивающее этот аспект: система есть *тень цели на среде* [6].

Подведем итог

Первое определение (искусственной) системы (“средство достижения цели”) system (as “a means of achieving a purpose”) emphasizes the subordination of all aspects of organization of the system to its purpose. However, even in the simplest examples, complexities arise: the correspondence between purposes and systems is not always one-to-one (one system may be related to several purposes; different systems may serve one purpose), and it is not always evident (to clarify the true purposes of existing system is not easy task). Nevertheless, the purposeful nature of a system is its primary, most important property.

Summary

§ 3.3. МОДЕЛЬ “ЧЕРНОГО ЯЩИКА”

В определении системы, приведенном в предыдущем параграфе, сделан акцент на назначении системы, а об ее устройстве говорится лишь косвенно. Для более определенной и точной характеристики конструкции системы следует развивать ее модель, преобразуя имеющиеся сведения так, чтобы в результате получить более удобную форму модели, включая в модель по мере необходимости дополнительные сведения.

КОМПОНЕНТЫ “ЧЕРНОГО ЯЩИКА”

Важную для человека роль играют наглядные, образные, визуальные модели. Перейдем от первого определения системы к его визуальному эквиваленту.

Во-первых, приведенное определение ничего не говорит о внутреннем устройстве системы. Поэтому ее можно изобразить в виде непрозрачного “ящика”, выделенного из окружающей среды. Подчеркнем, что уже эта, максимально простая, модель по-своему отражает два следующих важных свойства системы: *целостность и обособленность от среды*.

Во-вторых, в определении системы косвенно говорится о том, что хотя “ящик” и обособлен, выделен из среды, но не является полностью от нее изолированным.

В самом деле, ведь достигнутая цель – это запланированные заранее изменения в окружающей среде, какие-то продукты работы системы, предназначенные для потребления вне ее. Иначе говоря, система связана со средой и с помощью этих связей воздействует на среду. Изобразим связи в виде стрелок, направленных от системы в среду. Эти связи называются **выходами** системы. Подчеркнем еще раз, что выходы системы в данной графической модели соответствуют слову “цель” в словесной модели (первом определении) системы.

Кроме того, в определении имеется указание и на наличие связей другого типа: система является средством, поэтому должны существовать и возможности ее использования, воздействия на нее, т.е. и такие связи со средой, которые направлены извне в систему. Изобразим эти связи также в виде соответствующих стрелок, направленных от среды в систему, и назовем их **входами** системы.

В результате мы построили модель системы, которая получила название **черного ящика** (рис. 3.1). Это название образно подчеркивает полное отсутствие сведений о внутреннем содержании “ящика”: в этой модели задаются, фиксируются, перечисляются только входные и выходные связи системы со средой (даже “стенки ящика”, т.е. границы между системой и средой, в этой модели обычно не описываются, а лишь подразумеваются, признаются существующими). Такая модель, несмотря на внешнюю простоту и на отсутствие сведений о внутренности системы, часто оказывается полезной.

Во многих случаях достаточно содержательного словесного описания входов и выходов; тогда модель “черного ящика” является просто их списком. Например, бытовая модель теле-



3.1

Модель “черного ящика”

визора такова: входы – шнур электропитания, антенна, ручки управления и настройки; выходы – экран кинескопа и звукодинамики. В других случаях требуется количественное описание некоторых или всех входов и выходов. Пытаясь максимально формализовать модель “черного ящика”, мы приходим к заданию двух множеств X и Y входных и выходных переменных, но никаких других отношений между этими множествами фиксировать нельзя (иначе это уже будет не “черный”, а прозрачный ящик).

СЛОЖНОСТИ ПОСТРОЕНИЯ МОДЕЛИ “ЧЕРНОГО ЯЩИКА”

Теперь рассмотрим принципиально важный вопрос об обманчивой простоте модели “черного ящика”. Казалось бы, так просто: перечислить входы и выходы системы – и модель готова. Но как только это потребуется сделать для конкретной реальной системы, мы сталкиваемся с трудностями. Проиллюстрируем это сначала на хорошо знакомых примерах.

Пример 1. Опишем выходы системы “наручные часы”. Учитывая, что выходы соответствуют конкретизации цели, фиксируем в качестве выхода *показание времени в произвольный момент*. Затем принимаем во внимание, что сформулированная таким образом цель относится ко всем часам, а не только к *нашим наручным часам*. Чтобы различить их, вносим следующее добавление (выход): *удобство ношения часов на запястье*; тогда появляется обязательность ремешка или браслета, а с ним и еще один выход: *удовлетворение требований санитарии и гигиены*, так как не любое крепление часов на руке допустимо с этой точки зрения.

Далее, представив себе условия эксплуатации часов, можно добавить *достаточную в бытовых условиях прочность; пылевлагонепроницаемость*.

Затем, расширив понятие “условия эксплуатации часов”, добавим еще два выхода: *достаточную для бытовых нужд точность; легкость прочтения показаний часов при беглом взгляде на циферблат*.

Можно еще более расширить круг учитываемых требований к часам, что позволит добавить несколько выходов: *соответствие моде и понятию красоты; соответствие цены ча-*

сов покупательной способности потребителя. Очевидно, что список желаемых, т.е. включаемых в модель, выходов можно продолжать. Например, можно потребовать, чтобы имелась возможность прочтения показаний часов в полной темноте, и реализация этого выхода приведет к существенному изменению конструкции часов, в которой могут быть различные варианты самосвещения, подсветки, считывания на ощупь или подачи звуковых сигналов. А ведь мы в явной форме еще не говорили о габаритах, весе, многих других физических, химических, экономических и социальных аспектах использования наручных часов...

Пример 2. Попробуем перечислить входы системы “легковой автомобиль”. Исходя из определения системы как средства достижения цели, мы связали понятие входа с управляющим воздействием на систему, воздействием, “подталкивающим” систему к цели. Поэтому сразу же выделим в автомобиле в качестве входов те его элементы, которые предназначены для управления во время движения: *руль, педали сцепления, газа и тормоза, рычаг переключения коробки передач, переключатели сигнализации и освещения, ручка аварийного и стояночного тормоза.*

Затем, учитывая, что регулирующие воздействия приходится осуществлять не только на ходу, в список входов автомобиля вносим *регулировочные винты, гайки, эксцентрики*.

Смазка и заправка – это также регулирующее и управляющее воздействия. Поэтому *точки смазки и заправочные отверстия являются входами*.

Нельзя не учитывать входы в буквальном смысле. Поэтому добавляем *двери салона и (задние) крышки багажника и капота*.

И тут мы начинаем понимать, что входное воздействие на автомобиль оказывает не только водитель, но и пассажиры, а также окружающая среда. Записываем в перечень входов *окна и зеркала*, с помощью которых поступает информация к водителю и пассажирам. Но тогда можно отметить, что свойства поверхности, по которой движется автомобиль, также оказывают входное воздействие: по-разному приходится действовать водителю при езде по асфальту, песку, гравию, в случае гололеда, грязи... Добавляем к списку входов *механическое воздействие грунта на колеса*.

Однако различие между песком и асфальтом для автомобиля существенно лишь потому, что существует *поле тяготения Земли*.

Вместе с тем мы еще не упомянули многие реально существующие способы воздействия среды на данную систему: *руочки стеклоподъемников, аэродинамическое сопротивление воздуха, кнопки радиоприемника или кондиционера*, а в последних моделях – *входы вычислительных устройств*. А разве не влияют на автомобиль и его пассажиров *электрические и магнитные поля*? Не зря же рекомендуют прикреплять к автомобилю проводящий ремень, который отводит накапливающиеся на кузове электрические заряды. Далее, стали обязательными пристегенные ремни, так как нельзя пренебречь тем, что существует еще один вход – *силы инерции*, которые при авариях достигают опасных для здоровья и жизни величин. Очевидно, что список входов может быть еще продолжен.

Рассмотренные примеры свидетельствуют, что построение модели “черного ящика” не является тривиальной задачей, так как на вопрос о том, сколько и какие именно входы и выходы следует включать в модель, ответ не прост и не всегда однозначен. Установим причины этого факта.

МНОЖЕСТВЕННОСТЬ ВХОДОВ И ВЫХОДОВ

Главной причиной множественности входов и выходов в модели “черного ящика” является то, что всякая реальная система, как и любой объект, взаимодействует с объектами окружающей среды неограниченным числом способов. Строя модель системы, мы из этого бесчисленного множества связей отбираем конечное их число для включения в список входов и выходов. Критерием отбора при этом является целевое назначение модели, существенность той или иной связи по отношению к этой цели. То, что существенно, важно, включается в модель, то, что несущественно, неважно, – не включается. Именно здесь возможны

ошибки. Тот факт, что мы не учитываем в модели, исключаем из рассмотрения остальные связи, не лишает их реальности, они все равно действуют независимо от нас. И нередко оказывается, что казавшееся несущественным или неизвестным для нас на самом деле является важным и должно быть учтено.

Особое значение этот момент имеет при задании цели системы, т.е. при определении ее выходов главную цель приходится сопровождать заданием дополнительных целей. Важно подчеркнуть, что выполнения только основной цели недостаточно, что невыполнение дополнительных целей может сделать ненужным или даже вредным и опасным достижение основной цели. Этот момент заслуживает особого внимания, так как на практике часто обнаруживается незнание, непонимание или недооценка важности указанного положения. Между тем оно является одним из центральных во всей системологии.

Пример 3. Лет тридцать назад свечение цифр и стрелок наручных часов было достигнуто применением фосфоресцирующей краски. Впоследствии оказалось, что кроме полезного эффекта возникали вредные для здоровья излучения, и выпуск таких часов пришлось прекратить. Теперь найдены нерадиоактивные светящиеся материалы, и светящиеся часы вновь появились в продаже.

Модель “черного ящика” часто оказывается не только очень полезной, но в ряде случаев единственно применимой при изучении систем. Например, при исследовании психики человека или влияния лекарства на живой организм мы лишены возможности вмешательства в систему иначе как только через ее входы, и выводы делаем только на основании наблюдения за ее выходами. Это вообще относится к таким исследованиям, в результате проведения которых нужно получить данные о системе в обычной для нее обстановке, где следует специально заботиться о том, чтобы измерения как можно меньше влияли на саму систему. Другая причина того, что приходится ограничиваться только моделью “черного ящика”, – действительное отсутствие данных о внутреннем устройстве системы. Например, мы не знаем, как “устроен” электрон, но знаем, как он взаимодействует с электрическими и магнитными полями, с гравитационным полем. Это и есть описание электрона на уровне модели “черного ящика”.

Сделаем еще одно замечание. Модель “черного ящика” – это уже структурированная модель: в ней про каждую связь со средой известно, относится ли она к числу входов или является выходом. Однако на ранних стадиях исследования системы такая информация может отсутствовать: мы можем выделить некоторую связь системы со средой, можем даже наблюдать или измерять параметр, характеризующий эту связь, но не иметь оснований безусловно говорить о направленности этой связи. В таких случаях иногда полезно рассмотреть две конкурирующих модели “черного ящика”, в одной из которых эта связь причислена ко входам, а в другой – к выходам. Примером является исследование связи между двумя процессами, когда неизвестно, какой из них – причина, а какой – следствие, или даже вообще является ли их связь причинно-следственной (т.е. возможно, что они оба – следствия какой-то ненаблюдаемой причины).

Подведем итог

Простота модели “черного ящика” обманчива. Всегда существует опасность неполноты составления перечня входов и выходов как вследствие того, что важные из них могут быть сочтены несущественными, так и в силу неизвестности некоторых из них в момент построения модели.

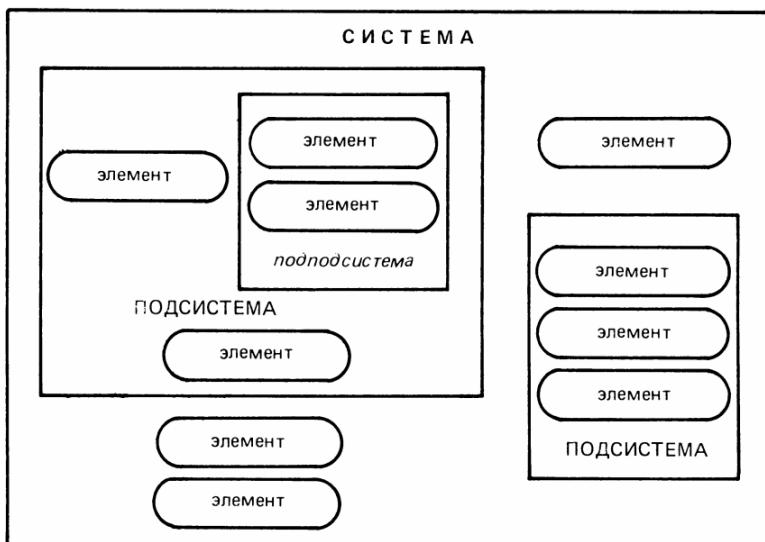
The simplicity of the “black-box” model is deceptive. There is always the danger that a compiled list of inputs and outputs might be incomplete, – either because some important ones were omitted as irrelevant, or because some of them are unknown at the time the model is created.

§ 3.4. МОДЕЛЬ СОСТАВА СИСТЕМЫ

Очевидно, что вопросы, касающиеся внутреннего устройства системы, невозможно решить только с помощью модели “черного ящика”. Для этого необходимы более развитые, более детальные модели.

КОМПОНЕНТЫ МОДЕЛИ СОСТАВА

При рассмотрении любой системы прежде всего обнаруживается то, что ее целостность и обособленность (отображенные в модели черного ящика) выступают как внешние свойства. Внутренность же “ящика” оказывается неоднородной, что позволяет различать составные части самой системы. При более детальном рассмотрении некоторые части системы могут быть, в свою очередь, разбиты на составные части и т.д. Те части системы, которые мы рассматриваем как неделимые, будем называть **элементами**. Части системы, состоящие более чем из одного элемента, назовем **подсистемами**. При необходимости можно ввести обозна-



3.2

Модель состава системы

чения или термины, указывающие на иерархию частей (например, “подподсистемы”, или “подсистемы такого-то уровня”).

В результате получается **модель состава системы**, описывающая, из каких подсистем и элементов она состоит (рис. 3.2).

Рассмотрим упрощенные примеры моделей состава для некоторых систем.

№	Система	Подсистемы	Элементы
---	---------	------------	----------

1.	Система телевидения “Орбита”	Подсистема передачи	Центральная телестудия Антенно-передающий центр
	Канал связи	Среда распространения радиоволн Спутники-ретрансляторы	
	Приемная подсистема	Местные телекомпании Телевизоры потребителей	
2.	Семья	Члены семьи	Муж Жена Предки Потомки <u>Другие родственники</u>
		Имущество семьи	Общее жилье и хозяйство Личная собственность членов семьи
3.	Отопительная система жилого дома	Источники тепла	Котельная или отвод от центральной теплотрасцы
		Подсистема распределения и доставки тепла	Трубы Калориферы Вентили
		Подсистема эксплуатации	Службы эксплуатации и ремонта Персонал

СЛОЖНОСТИ ПОСТРОЕНИЯ МОДЕЛИ СОСТАВА

Построение модели состава системы только на первый взгляд кажется простым делом. Если дать разным экспертам задание определить состав одной и той же системы, то результаты их работы будут различаться, и иногда довольно значительно. Причины этого состоят не только в том, что у них может быть различная степень знания системы: один и тот же эксперт при разных условиях также может дать разные модели. Существуют по крайней мере еще три важные причины этого факта.

Во-первых, разные модели состава получаются вследствие того, что понятие элементарности можно определить по-разному. То, что с одной точки зрения является элементом, с другой – оказывается подсистемой, подлежащей дальнейшему разделению.

Во-вторых, как и любые модели, модель состава является целевой, и для различных целей один и тот же объект потребуется разбить на разные части. Например, один и тот же завод для директора, главного бухгалтера, начальника пожарной охраны состоит из совершенно различных подсистем. Точно так же модели состава самолета с точек зрения летчика, стюардессы, пассажира и аэродромного диспетчера окажутся различными. То, что для одного обязательно войдет в модель, может совершенно не интересовать другого.

В-третьих, модели состава различаются потому, что всякое разделение целого на части, всякое деление системы на подсистемы является относительным, в определенной степени условным. Например, тормозную систему автомобиля можно отнести либо к ходовой части, либо к подсистеме управления. Другими словами, границы между подсистемами условны, относительны, модельны.

Это относится и к границам между самой системой и окружающей средой; поэтому остановимся на этом моменте подробнее. В качестве примера рассмотрим систему “часы”. Какую бы природу ни имели устройства, которые мы называем часами, в них можно выделить две подсистемы: *датчик времени*, т.е. процесс, ход которого изображает течение времени (это

BOUNDARY

граница

PROPERTY

свойство

PARTITION

состав

STRUCTURE, LINKS, CONNECTIONS

структура

может быть равномерное раскручивание пружины, электрический ток с некоторым постоянным параметром, равномерное течение струйки песка, вращение Земли вокруг своей оси, колебания некоторой молекулы и т.д.); *индикатор времени*, т.е. устройство, преобразующее, отображающее состояние датчика в сигнал времени для пользователя. Модель состава часов можно считать полностью исчерпанной (если снова не разбивать эти две подсистемы). Однако поскольку фактически каждые часы показывают состояние своего датчика, рано или поздно их показания разойдутся между собой. Выход из этого положения состоит в синхронизации всех часов с неким общим для всех *эталоном времени*, например с помощью сигналов “точного времени”, передаваемых по радио. Здесь и возникает вопрос: включать ли этalon времени в состав часов как системы или рассматривать часы как подсистему в общей системе указания времени?

Можно привести много подобных примеров относительно системных границ. Вспомните пример с мистером Джонсоном, поглощающим котлету (см. § 2.5); солдат в отпуске или студент на каникулах в определенной степени остаются элементами соответствующих систем; возвращающийся с работы человек, упав и получив травму, имеет разные права в зависимости от того, произошло ли это на крыльце заводской проходной или его собственного дома (травма считается производственной или бытовой, и юри-

Модель состава системы отображает из каких частей (подсистем и элементов) состоит система. Главная трудность в построении модели состава заключается в том, что разделение целостной системы на части является относительным, условным, зависящим от целей моделирования (это относится не только к границам между частями системы, но и к границам самой системы). Кроме того, относительным является и определение самой малой части – элемента.

Модель структуры системы

дические последствия различны).

Подведем итог

Модель состава ограничивается снизу тем, что считается “from below” by what is considered an element, and “from above” by a system boundary. Так и границы разбиения на подсистемы определяются целями построения модели и, следовательно, не имеют абсолютного характера. Это не означает, что сама система или ее состав нереальны. Мы имеем дело не с разными системами, а с разными моделями системы.

These boundaries, as well as boundaries between subsystems, depend on purpose of the model and hence are relative in nature. This does not mean that the system itself or its parts are unreal. We deal not with different systems but with different models of the same system.

§ 3.5. МОДЕЛЬ СТРУКТУРЫ СИСТЕМЫ

Еще раз подчеркнем, что для достижения ряда практических целей достаточно модели “черного ящика” или модели состава. Однако очевидно, что есть вопросы, решить которые с помощью этих моделей нельзя. Чтобы получить велосипед недостаточно иметь “ящик” со всеми отдельными его деталями (состав налицо). Необходимо еще правильно соединить все детали между собой, или, говоря более общо, установить между элементами определенные связи – отношения. Совокупность необходимых и достаточных для достижения цели отношений между элементами называется структурой системы.

ОТНОШЕНИЯ И СТРУКТУРЫ

Перечень связей между элементами (т.е. структура системы) является отвлеченной, абстрактной моделью: установлены только отношения между элементами, но не рассмотрены сами элементы. Хотя на практике безотносительно к элементам говорить о связях можно лишь после того, как отдельно рассмотрены сами элементы (т.е. рассмотрена модель состава), теоретически модель структуры можно изучать отдельно.

Бесконечность природы проявляется и в том, что между реальными объектами, вовлеченными в систему, имеется невообразимое (может быть, бесчисленное) количество отношений. Однако когда мы рассматриваем некоторую совокупность объектов как систему, то из всех отношений важными, т.е. существенными для достижения цели, являются лишь некоторые. Точнее, в модель структуры (т.е. в список отношений) мы включаем только конечное число связей, которые, по нашему мнению, существенны по отношению к рассматриваемой цели.

Пример 1. При расчете механизма не учитываются силы взаимного притяжения его деталей, хотя, согласно закону всемирного тяготения, такие силы объективно существуют. Зато вес деталей (т.е. сила их притяжения к Земле) учитывается обязательно.

Пример 2. Между деталями молотка наряду с отношениями, существенными для его использования, имеются и несущественные. Например, если соприкасающиеся детали изготовлены из разных металлов, то между ними есть контактная разность потенциалов.

Пример 3. Рассмотрим систему “часы вообще”. Считаем, что в состав такой системы входят три элемента: датчик, индикатор и этalon времени (см. § 3.4). Структура часов определяется следующими отношениями между парами элементов:

Пара элементов	Связь между ними
Датчик и индикатор	Однозначное соответствие
Эталон и датчик	Приблизительное соответствие
Индикатор и эталон	Периодическое сравнение и устранение расхождения

Отношения между элементами могут быть самыми разнообразными. Однако можно попытаться их классифицировать и по возможности перечислить. Трудность состоит в том, что мы знаем не все реально существующие отношения и вообще неизвестно, является ли конечным их число. Интересное исследование было проведено с естественными языками. Выделение языковых конструкций, выражающих отношения (типа находиться на (под, около,...), быть причиной, быть подобным, быть одновременно, состоять из, двигаться к (от, вокруг,...) и т.п.), привело к выводу, что в английском, итальянском и русском языках число выражаемых отношений примерно одинаково и немного превышает 200. Этот результат не может служить доказательством конечности числа отношений, но сам факт дает повод для размышлений.

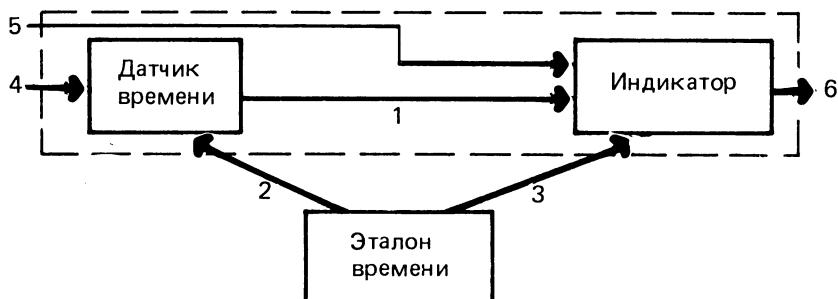
СВОЙСТВО И ОТНОШЕНИЕ

В заключение данного параграфа рассмотрим связь между понятиями “отношение” и “свойство”. В отношении участвует не менее двух объектов, а свойством мы называем некоторый атрибут одного объекта. Это различие отражается и при их математическом описании.

Пусть E – множество. Любое свойство, которым может обладать элемент $x \in E$, задает в E подмножество $A \subseteq E$ всех элементов, обладающих этим свойством. Пусть задано некоторое отношение R , в котором могут находиться элементы x и y множества E , записанные в указанном порядке. Если они находятся в заданном отношении, то используется запись xRy , если нет – запись $x\bar{R}y$. Множество всех упорядоченных пар (x, y) (т.е. (x, y) и (y, x) – разные пары при $x \neq y$) называется произведением $E \times E$. Рассмотрим подмножество $R \subseteq E \times E$ всех пар, для которых xRy . Задание этого подмножества является заданием отношения. Если теперь ввести понятие многоместного (а не только двуместного, бинарного) отношения, то свойство оказывается *одноместным (унарным) отношением*.

Однако нас интересует не только формальное доказательство того, что свойство есть частный случай отношения: это скорее следствие теоретико-множественного определения отношения как подмножества. Важнее проследить содержательную связь свойства и отношения. Во-первых, любое свойство, даже если его понимать как потенциальную способность обладать определенным качеством, выявляется в процессе взаимодействия объекта (носителя свойства) с другими объектами, т.е. в результате установления некоторого отношения. Чтобы убедиться в том, что мяч красный, мало иметь мяч, нужны еще источник белого света и анализатор света, отраженного от мяча (еще и не всякий анализатор пригоден; например, глаз дальтоника не может установить цвет мяча). Во-вторых, можно сделать дальнейшее обобщение и выдвинуть следующее предположение: свойство – это не атрибут объекта, а лишь определенная абстракция отношения, экономящая мышление. Мы “коротко и ясно” говорим, что стекло прозрачно, вместо того чтобы каждый раз говорить об отношении между лучом света, падающим на поверхность стекла, самим листом стекла и приемником света, находящимся по другую сторону этого листа. Другими словами, можно утверждать, что свойство – это свернутое отношение (а если использовать понятие модели так, как это предложено в гл. 2, то свойство есть модель отношения).

Остается проверить, является ли это утверждение справедливым для всех свойств или все же имеются некоторые абсолютные атрибуты объектов. Мы склоняемся к первому утверждению. Даже свойства элементарных частиц определяются через их взаимодействия. Ведь не зря же В.И. Ленин даже такое свойство материи, как ее объективное существование, характеризовал через его данность нам в ощущениях, т.е. опять-таки как некоторое отношение.



3.3

Структурная схема часов

Подведем итог

Очередным шагом в развитии модели системы является модель структуры, которая описывает существенные связи между элементами (компонентами модели состава). Говоря, что свойства какого-то объекта можно использовать в системе, мы имеем в виду установление некоторых отношений между некоторыми отношениями между данным объектом и другими частями системы, т.е. включение этих отношений в структуру системы.

Summary

The next step in developing a model of a system is the structure model which describes the essential links between elements (components of the system). Saying that certain properties of an object could be useful for the system, we mean that definite relations could be established between this object and other parts of the system, i.e. that these relations should be included in the system structure.

§ 3.6. ВТОРОЕ ОПРЕДЕЛЕНИЕ СИСТЕМЫ.

СТРУКТУРНАЯ СХЕМА СИСТЕМЫ

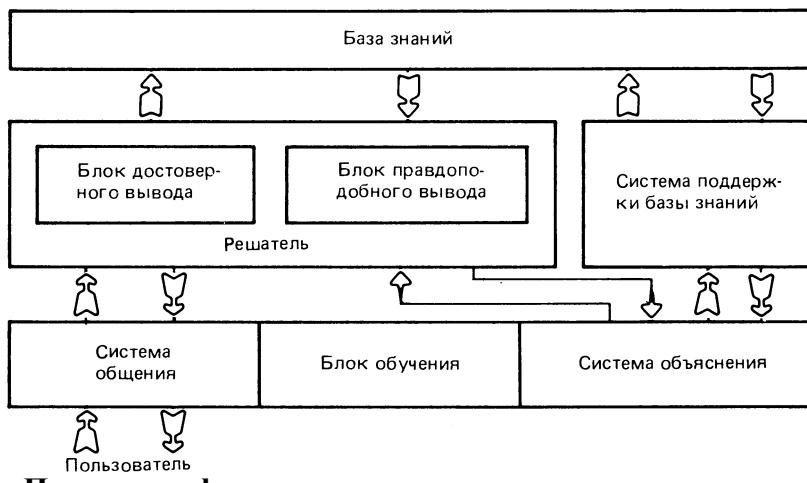
Объединяя все изложенное в предыдущих параграфах, можно сформулировать второе определение системы: система есть совокупность взаимосвязанных элементов, обособленная от среды и взаимодействующая с ней как целое.

СТРУКТУРНАЯ СХЕМА КАК СОЕДИНЕНИЕ МОДЕЛЕЙ

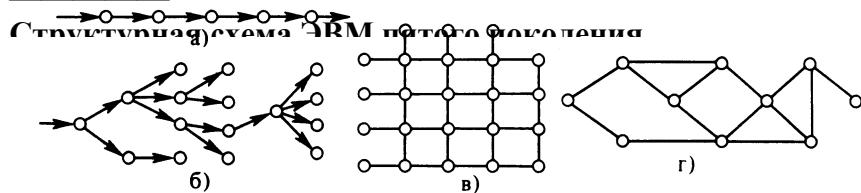
Очевидно, что это определение охватывает модели “черного ящика”, состава и структуры. Все вместе они образуют еще одну модель, которую будем называть **структурной схемой системы**; в литературе встречаются также термины “белый ящик”, “прозрачный ящик”, подчеркивающие ее отличие от модели “черного ящика”, а также термин “конструкция системы”, который мы будем использовать для обозначения материальной реализации структурной схемы системы. В структурной схеме указываются все элементы системы, все связи между элементами внутри системы и связи определенных элементов с окружающей средой (входы и выходы системы).

Пример 1. Структурная схема системы “синхронизируемые часы” приведена на рис. 3.3. Элементы системы изображены в виде прямоугольников; связи 1 – 3 между элементами описаны в примере 3 § 3.5; вход 4 изображает поступление энергии извне; выход 5 соответствует регулировке индикатора; выход 6 – показание часов.

Все структурные схемы имеют нечто общее, и это побудило математиков рассматривать их как особый объект математических исследований. Для этого пришлось абстрагироваться от содержательной стороны структурных схем, оставив в рассматриваемой модели только общее для каждой схемы. В результате получилась схема, в которой обозначается только наличие элементов и связей между ними, а также (в случае необходимости) разница между элементами и между связями. Такая схема называется **графом**. Следовательно, граф состоит



3. Пример графа



3.5

из обозначений элементов произвольной природы, называемых *вершинами*, и обозначений связей между ними, называемых *ребрами* (иногда дугами). На рис. 3.4 изображен график: вершины обозначены в виде кружков, ребра – в виде линий. Часто бывает необходимо отразить несимметричность некоторых связей; в таких случаях линию, изображающую ребро, снабжают стрелкой. Если направления связей не обозначаются, то график называется *неориентированным*, при наличии стрелок – *ориентированным* (полностью или частично). Данная пара вершин может быть соединена любым количеством ребер; вершина может быть соединена сама с собой (тогда ребро называется *петлей*). Если в графике требуется отразить другие различия между элементами или связями, то либо приписывают разным ребрам различные веса (*взвешенные графы*), либо раскрашивают вершины или ребра (*раскрашенные графы*).

Оказалось, что для графов может быть построена интересная и содержательная теория, имеющая многочисленные приложения. Разнообразные задачи этой теории связаны с различными преобразованиями графов, а также с возможностью рассмотрения различных отношений на графах: весов, рангов, цветов, вероятностных характеристик (*стохастические графы*) и т.д. В связи с тем что множества вершин и ребер формально можно поменять местами, получается два разных представления системы в виде *вершинного* или в виде *реберного* графа. Оказывается, что в одних задачах удобнее использовать вершинный, а в других – реберный график [6].

Графы могут изображать любые структуры, если не накладывать ограничений на пересекаемость ребер. Некоторые типы структур имеют особенности, важные для практики, они выделены из других и получили специальные названия. Так, в организационных системах часто встречаются *линейные*, *древовидные* (иерархические) и *матричные структуры*; в технических системах чаще встречаются *сетевые структуры* (рис. 3.5); особое место в теории систем занимают *структуры с обратными связями*, которые соответствуют кольцевым путям в ориентированных графах (такие структуры будут рассмотрены отдельно).

Пример 2. Структурная схема ЭВМ пятого поколения, с помощью которой пользователь, не умеющий программировать, может решать достаточно сложные задачи [7], приведена на рис. 3.6. Отметим, что в этой схеме имеются и иерархические, и линейные, и обратные связи.

Одной структурной информации, которая содержится в графах, для ряда исследований недостаточно. В таких случаях методы теории графов становятся вспомогательными, а главным является рассмотрение конкретных функциональных связей между входными, внутренними и выходными переменными системы.

Подведем итог

Summary

Структурная схема системы The structural scheme of a system is the most detailed and exhaustive model of any system at стемы на данном этапе нашего познания. При этом остается актуальным вопрос об адекватности этой модели, разрешаемый только на практике.

The structural scheme of a system is the most detailed and exhaustive model of any system at a given stage of our knowledge. The problem of the adequacy of the model still exists, and can be solved only in practice.

§ 3.7. ДИНАМИЧЕСКИЕ МОДЕЛИ СИСТЕМ

До сих пор основное внимание было уделено понятию системы, ее составу и устройству. Были построены модели, которые являются как бы “фотографиями” системы, отображают ее в некоторый момент времени. В этом смысле рассмотренные варианты моделей “черного ящика”, состава, структурной схемы системы могут быть названы **статическими моделями**, что подчеркивает их неподвижный, как бы застывший характер.

ОТОБРАЖЕНИЕ ДИНАМИКИ СИСТЕМЫ

Следующий шаг в исследовании систем состоит в том, чтобы понять и описать, как система “работает”, что происходит с ней самой и с окружающей средой в ходе реализации поставленной цели. Очевидно, и подход к описанию, и степень подробности описания происходящих процессов могут быть различными. Однако общим при этом является то, что разрабатываемые модели должны отражать поведение систем, описывать происходящие с течением времени изменения, последовательность каких-то этапов, операций, действий, причинно-следственные связи.

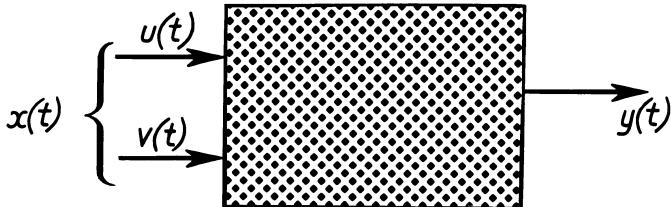
Системы, в которых происходят какие бы то ни было изменения со временем, будем называть **динамическими**, а модели, отображающие эти изменения, – **динамическими моделями** систем. Заметим, что термин “динамический” в русском языке неоднозначен; здесь он будет использован в самом широком смысле как обозначение любых изменений во времени.

Для разных объектов и систем разработано большое количество динамических моделей, описывающих процессы с различной степенью детальности: от самого общего понятия динамики, движения вообще, до формальных математических моделей конкретных процессов типа уравнений движения в механике или волновых уравнений в теории поля. Развитие моделей происходит примерно в той последовательности, как это было изложено: от “черного ящика” к “белому”. Однако этот путь конкретизации моделей непрост и нелегок и для многих систем еще не закончен из-за недостаточности имеющихся знаний.

ФУНКЦИОНИРОВАНИЕ И РАЗВИТИЕ

Уже на этапе “черного ящика” различают два типа динамики системы: ее функционирование и развитие. Под **функционированием** подразумеваются процессы, которые происходят в системе (и окружающей ее среде), стабильно реализующей фиксированную цель (функционируют, например, часы, городской транспорт, кинотеатр, канцелярия, радиоприемник, станок, школа и т.д.). **Развитием** называют то, что происходит с системой при изменении ее целей. Характерной чертой развития является тот факт, что существующая структура перестает соответствовать новой цели, и для обеспечения новой функции приходится изменять структуру, а иногда и состав системы, перестраивать всю систему*.

* Понятие “развитие” будет уточнено в § 9.7.



3.7

Динамическая модель “черного ящика”: задание

Не следует считать, что система всегда находится либо в фазе развития, либо в состоянии функционирования. При реконструкции одного цеха остальные функционируют, завод в целом развивается. Даже при коренной перестройке системы какие-то элементы и даже подсистемы старой структуры могут продолжать функционировать в новой по-прежнему. Возможны и такие системы, для функционирования которых какие-то ее подсистемы должны быть постоянно в развитии.

Следующий шаг в построении динамических моделей состоит в том, чтобы конкретнее отобразить происходящие изменения. Это означает, что следует различать части, этапы происходящего процесса, рассматривать их взаимосвязи. Иными словами, типы динамических моделей такие же, как и статических, только элементы этих моделей имеют временной характер. Например, динамический вариант “черного ящика” – указание начального (“вход”) и конечного (“выход”) состояний системы (например, как в пятилетнем плане). Модели состава соответствует перечень этапов в некоторой упорядоченной последовательности действий. Например, доказано, что любой алгоритм можно построить, используя всего три оператора: “выполнять последовательно”, “если... то...” и “выполнять, пока не удовлетворится условие”. Эти операторы можно рассматривать как модель минимального состава алгоритма, хотя не обязательно составлять алгоритм только из этих операторов. Динамический вариант “белого ящика” – это подробное описание происходящего или планируемого процесса. Например, на производстве широко используют так называемые сетевые графики – графы, имеющие сетевую структуру; их вершинами служат выполняемые производственные операции, а ребра указывают, какие операции не могут начаться, пока не окончатся предыдущие. Здесь же некоторым образом (например, с помощью задания длин или весов ребер) изображается длительность выполнения операций, что и позволяет находить на графике “критические” пути, т.е. последовательности операций, от которых главным образом зависит ритмичность всей работы.

ТИПЫ ДИНАМИЧЕСКИХ МОДЕЛЕЙ

Те же типы моделей прослеживаются и при более глубокой формализации динамических моделей [4]. При математическом моделировании некоторого процесса его конкретная реализация описывается в виде соответствия между элементами множества X возможных “значений” x и элементов упорядоченного множества T “моментов времени” t , т.е. в виде отображения $T \rightarrow X: x(t) \in X^T, t \in T$. С помощью этих понятий можно строить математические модели систем.

Рассматривая выход $y(t)$ системы (это может быть вектор) как ее реакцию на управляемые $u(t)$ и неуправляемые $v(t)$ входы $x(t) = \{u(t), v(t)\}$ (рис. 3.7), можно модель “черного ящика” выразить как совокупность двух процессов: $X^T = \{x(t)\}$ и $Y^T = \{y(t)\}, t \in T$. Если даже считать $y(t)$ результатом некоторого преобразования Φ процесса $x(t)$, т.е. $y(t) = \Phi(x(t))$, то модель “черного ящика” предполагает, что это преобразование неизвестно. В том же случае, когда мы имеем дело с “белым ящиком”, соответствие между входом и выходом можно описать тем или иным способом. Какой именно способ – зависит от того, что нам известно, и в какой форме можно использовать эти знания.

Например, иногда бывает известно, что система мгновенно преобразует вход в выход, т.е. что $y(t)$ является функцией только $x(t)$ в тот же момент времени. Остается задать или найти эту функцию. На практике чаще всего известна лишь безынерционность системы и требуется, наблюдая входы и выходы, восстановить неизвестную функцию $y = \Phi(x)$. По существу, это задача о переходе от модели “черного ящика” к модели “белого ящика” по наблюдениям входов и выходов при наличии информации о безынерционности системы. Даже в такой достаточно простой постановке задача имеет совсем не простые варианты, которые зависят от того, что известно о функции Φ (в параметризованном случае Φ принадлежит семейству функций, известных с точностью до параметров; в непараметризованном – вид функции Φ неизвестен), и от наличия или отсутствия некоторых общих сведений о ее свойствах (непрерывности, гладкости, монотонности, симметричности и т.п.). Дополнительные варианты (и дополнительные трудности) возникают, если входы или выходы наблюдаются с помехами или искажениями. При этом разные предположения о природе этих помех приводят к принципиально отличающимся решениям задачи (например, в случае, когда распределение помехи известно точно, известно с точностью до числовых параметров или неизвестно совсем).

Однако класс систем, которые можно считать безынерционными, весьма узок. Необходимо строить математические модели систем, выход которых определяется не только значением входа в данный момент времени, но и теми значениями, которые были на входе в предыдущие моменты. Более того, в самой системе с течением времени как под влиянием входных воздействий, так и независимо от них могут происходить изменения, что также следует отразить в модели.

ОБЩАЯ МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ДИНАМИКИ

В наиболее общей модели [2; 4] это достигается введением понятия **состояния системы** как некоторой (внутренней) характеристики системы, значение которой в настоящий момент времени определяет текущее значение выходной величины. Состояние можно рассматривать как своего рода хранилище информации, необходимой для предсказания влияния настоящего на будущее [1, с. 12 – 13]. Обозначим это состояние через $z(t)$. Все сказанное выше означает существование такого отображения $\eta: Z \times T \rightarrow Y$, что

$$y(t) = \eta(t, z(t)), t \in T. \quad (1)$$

Явная зависимость η от t введена для учета возможности изменения зависимости выхода от состояния с течением времени. Это отображение называется **отображением выхода**.

Для завершения построения модели нужно описать связь между входом и состоянием, т.е. ввести параметрическое семейство отображений $\mu_{\tau}t: Z \times X(\cdot) \rightarrow Z$, заданных для всех значений параметров $t \in T$, $\tau \in T$ и $\tau \leq t$. Это означает принятие аксиомы о том, что состояние в любой момент $t > \tau$ однозначно определяется состоянием z_{τ} в момент τ и отрезком реализации входа $x(\cdot)$ от τ до t :

$$z(t) = \mu_{\tau}t(z_{\tau}, x(\cdot)) = ??(t; ?, z_{\tau}, x(\cdot)). \quad (2)$$

Такое отображение называют **переходным отображением**.

Итак, математическая модель системы, соответствующая уровню “белого ящика”, – это задание множеств входов, состояний и выходов, и связей между ними:

$$X \xrightarrow{\sigma} Z \xrightarrow{\eta} Y. \quad (3)$$

Конкретизируя множества X , Z и Y и отображения σ и η , можно перейти к моделям различных систем. Так, говорят о *дискретных и непрерывных по времени* системах в зависимости от того, дискретно или непрерывно множество T . Далее, если множества X , Z и Y дискретной по времени системы имеют конечное число элементов, то такую систему называют *конечным автоматом*. Это довольно простой класс систем в том смысле, что для исследования конечных автоматов необходимы лишь методы логики и алгебры; в то же время это широкий и практически важный класс, так как в него входят все дискретные (цифровые) измерительные, управляющие и вычислительные устройства, в том числе и ЭВМ.

VERTEX
 вершина (графа)
ARC
 дуга (графа)
LINEAR
 линейный
STATE
 состояние
FUNCTION
 функция

Объединив модели “черного ящика”, состава и структуры системы, мы получим самую полную (для наших целей), самую подробную (для нашего уровня знаний) модель системы – ее структурную схему.
Модели, отображающие процессы, происходящие в системе с течением времени, называются динамическими моделями. Динамические модели включают те же типы, что и статические, но с явным выделением, подчеркиванием роли времени.
Всякая реальная динамическая система подчине-

Если X , Z и Y – линейные пространства, а σ и η – линейные операторы, то и система называется *линейной**. Если к линейной системе дополнительно предъявить требования, состоящие в том, чтобы пространства имели топологическую структуру**, а σ и η были бы непрерывны в этой топологии, то мы приходим к *гладким* системам. Этот класс систем имеет большое значение, так как оказалось [2], что для гладких систем переходное отображение σ является общим решением дифференциального уравнения

$$\frac{dz}{dt} = f(t, z, x) , \quad (4)$$

а для дискретных систем – общим решением уравнения

$$z(t_{k+1}) = f(t_k, z, x) = \sigma(t_{k+1}; t_k, z, x(\cdot)) , \quad (5)$$

где $x(\cdot)$ – траектория для моментов времени $t \leq t_k$.

Интенсивно исследовались *стационарные системы*, т.е. такие системы, свойства которых со временем не изменяются. Стационарность означает независимость от t функции η и инвариантность функции σ к сдвигу во времени:

$$\eta(t, z(t)) = \eta(z(t)), \sigma(t; t_0, z, x(\cdot)) = \sigma(t + \tau; t_0 + \tau, z, \bar{x}(\cdot)) ,$$

где $\bar{x}(\cdot)$ есть $x(\cdot)$, сдвинутое на время τ .

Конкретизация моделей динамических систем на этом, конечно, не заканчивается; приведенные модели скорее всего являются просто примерами, которые можно рассматривать отдельно. Но на одном свойстве реальных динамических систем следует остановиться. Речь идет о подчиненности реальных систем **принципу причинности**. Согласно этому принципу, *отклик системы на некоторое воздействие не может начаться раньше самого воздействия*. Это условие, очевидное для реальных систем, совсем не автоматически выполняется в рамках их математических моделей. При этом модель, в которой нарушается принцип причинности, не обязательно является “плохой”, бесполезной. Примером служит модель фильтра с конечной полосой пропускания: отклик такой системы на короткий импульс имеет вид $\sin\omega_0 t / (\omega_0 t)$, т.е. начинается в минус бесконечности. Несмотря на явное нарушение принципа причинности, такую модель широко используют в радиотехнике. Однако, как только возникает вопрос о практической реализации такого фильтра, становится ясно, что она невозможна в точном смысле, хотя допустимы различные приближения. В связи с этим одна из проблем теории динамических систем состоит в выяснении *условий физической реализуемости* теоретических моделей, т.е. конкретных ограничений, которые приходится накладывать на модель при соблюдении принципа причинности.

* Основным свойством линейных систем является выполнение принципа суперпозиции, т. е. условия $[x(t)=x_1(t) + x_2(t)] \rightarrow [y(t)=y_1(t)+y_2(t)]$, где $x_1(t)$ и $x_2(t)$ – некоторые входные воздействия, а $y_1(t)$ и $y_2(t)$ – выходные отклики на каждый из них в отдельности.

** Не вдаваясь в математические подробности, отметим, что задание топологической структуры множества позволяет строго определить основные понятия анализа на этом множестве, например *сходимость* последовательностей на нем, а также *вводить метрику* (меру близости между элементами пространства).

Подведем итог

Отображение процессов, происходящих в системе и в окружающей ее среде, осуществляется с помощью динамических моделей. Все, что говорилось о моделях вообще (см. гл. 2), конечно, относится и к динамическим моделям. В частности, динамические модели систем делятся на те же типы, что и статические модели (см. § 3.3-3.5). Интересным и важным примером различия между моделью и оригиналом служит то, что динамическая модель может не удовлетворять принципу причинности, но тем не менее обладать полезными свойствами.

Summary

The processes taking place inside a system and in its environs are represented by dynamic models. Everything that has been said about models (cf. Chapter 2) is, of course, valid for dynamic models as well. In particular, dynamic models are of the same three types as the static ones (cf. Sections 3.3 – 3.5). The fact that a dynamic model cannot satisfy the causality principle but at the same time can possess some useful properties is an interesting and important example of the difference between a model and its origin.

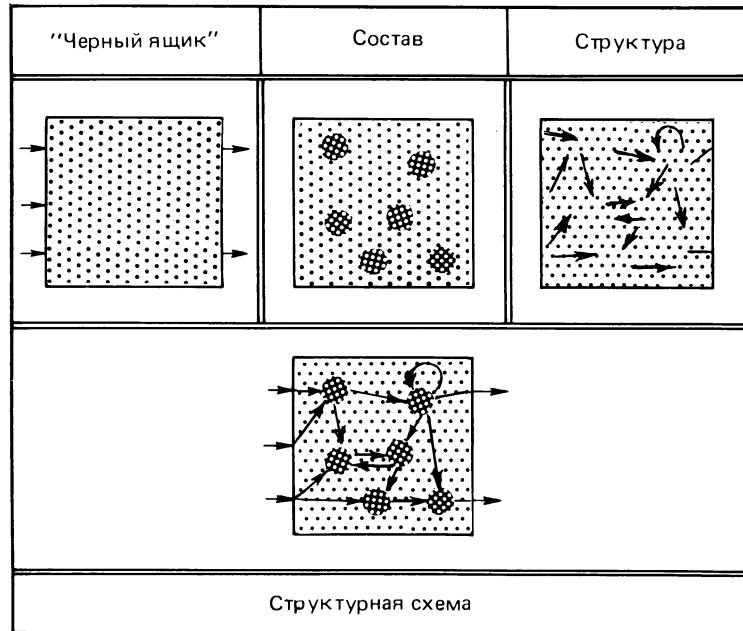
ЗАКЛЮЧЕНИЕ

Оказывается, что при всем невообразимом многообразии реальных систем **п р и н ц и п и - а льно разлиных типов** моделей систем очень немного: модель типа “черный ящик”, модель состава, модель структуры, а также их разумные сочетания и прежде всего объединение всех трех моделей, т.е. структурная схема системы. Это относится как к статическим моделям, отображающим фиксированное состояние системы (рис. 3.8, а), так и к динамическим моделям, отображающим характер временных процессов, которые происходят с системой (рис. 3.8, б). Рис. 3.8 иллюстрирует общность типов статических и динамических моделей. Можно сказать, что структурная схема (“белый ящик”) получается как результат “суммирования” моделей “черного ящика”, состава и структуры системы.

Все указанные типы моделей являются формальными, относящимися к любым системам и, следовательно, не относящимися ни к одной конкретной системе. Чтобы получить модель заданной системы, нужно придать формальной модели конкретное содержание, т.е. решить, какие аспекты реальной системы включать как элементы модели избранного типа, а какие – нет, считая их несущественными. Этот процесс, как было показано, обычно неформализуем, поскольку признаки существенности или несущественности в очень редком случае удается формализовать (к таким случаям относится, например, возможность принять в качестве признака существенности частоту встречаемости данного элемента в различных подобных, т.е. одинаково классифицируемых, системах). Столь же слабо формализованными являются признаки элементарности и признаки разграничения между подсистемами.

В силу указанных причин, процесс построения содержательных моделей является процессом интеллектуальным, творческим. Тем не менее интуиции эксперта, разрабатывающего содержательную модель, немало помогают формальная модель и рекомендации по ее наполнению конкретным содержанием. Формальная модель является “окном”, через которое эксперт смотрит на реальную систему, строя содержательную модель.

В процессе построения содержательных моделей систем отчетливо прослеживается необходимость использования диалектики. В этом процессе главной является задача создания *полной модели*. Общие рекомендации по достижению полноты вытекают из основных положений диалектики:



a)



б)

3.8

Типы моделей: а) статический вариант; б) динамический вариант

необходимо стремиться учесть все существенные факторы, влияющие на рассматриваемое явление; поскольку такая существенность не всегда очевидна, лучше включить в модель несущественный элемент, чем не включить существенный;

одним из необходимых признаков полноты модели является наличие в ней противоречивых элементов; следует уделить специальное внимание этому моменту: например, при перечислении выходов надо включать в перечень не только желательные целевые выходы (связи, продукцию и т.п.), но и нежелательные (отходы, брак, и т.п.);

как бы ни были обширны наши знания о данном явлении, реальность богаче моделей – в ней всегда есть неизвестные факторы; чтобы не упустить из виду возможность чего-то существенного, но пока неизвестного, рекомендуется включать в модель неявные “запасные”, неконкретизированные элементы (типа “все остальное”, “что-то еще”) и на различных стадиях системного анализа обращаться к этим элементам, как бы ставя перед собой вопрос: не пора ли дополнить модель еще одним явным элементом?

Эти рекомендации, конечно, не исчерпывают всех возможностей: в арсенал искусства моделирования входит много научно обоснованных методов и эмпирических эвристик (см., например, [1; 3; 9]).

ЛИТЕРАТУРА

1. Блехман И.И., Мышикис А.Д., Пановко Я.Г. Прикладная математика: предмет, логика, особенности подходов. – Киев: Наукова думка, 1976.
2. Калман Р., Фалб П., Арбіб М. Очерки по математической теории систем. – М.: Мир, 1971.
3. Кемени Дж., Снелл Дж. Кибернетическое моделирование. Некоторые приложения. – М.: Сов. радио, 1972.
4. Мороз А.И. Курс теории систем. – М.: Высш. шк., 1987.
5. Нечипоренко В.И. Структурный анализ систем. – М.: Сов. радио, 1977.
6. Перегудов Ф.И и др. Основы системного подхода. – Томск: ТГУ, 1976.
7. Поспелов Д.А. Инженерия знаний // Наука и жизнь. 1987. № 6. С. 11 – 18.
8. Уемов А.И. Системный подход и общая теория систем. – М.: Мысль, 1978.
9. Шэннон Р. Имитационное моделирование систем – искусство и наука. – М.: Мир, 1978.

УПРАЖНЕНИЯ

3.2

- Приведите примеры:
 - а) системы, которая предназначена для выполнения определенной цели, но которую можно использовать и для других целей;
 - б) системы, спроектированной специально для реализации одновременно нескольких различных целей;
 - в) разных систем, предназначенных для одной и той же цели.
- Сформулируйте цель работы вашего факультета так, чтобы она не была общей для других факультетов, в том числе для родственных факультетов других вузов.

§ 3.3

- Обсудите проблему множественности входов и выходов на примере знакомой вам системы (радиоприемника, столовой, велосипеда и т.п.). Перечислите при этом нежелательные входы и выходы. Установите, как можно устраниТЬ недостатки системы (нежелательные связи со средой).

- Выделите в примере из предыдущего упражнения главную цель системы, дополнительные цели и ограничения.
- Приведите случай, когда другие модели, кроме модели “черного ящика”, не нужны или недопустимы.

§ 3.4

- Приведите пример, когда различным целям соответствуют разные модели состава: а) различающиеся субстратно, т.е. содержащие разные элементы; б) различающиеся делением на подсистемы, но содержащие одни и те же элементы; в) различающиеся из-за отличий в определении элементарности.

§ 3.5

- Попробуйте привести пример свойства, которое хотя и проявляется во взаимодействии, но существует само по себе как атрибут объекта.

§ 3.6

- Сравните формальную структурную схему какой-нибудь известной вам организации с ее реальной структурой. Обсудите расхождения.

§ 3.7

- Рассмотрите процессы функционирования и развития на примере конкретной системы. Обсудите общее и различное в росте и развитии системы.

Вопросы для самопроверки

1. Чем объясняется существование различных определений системы? Как совместить справедливость каждого из них с тем, что они различны?
2. Соответствие между конструкцией системы и ее целью не однозначно, но и не произвольно. Что же их связывает?
3. От чего зависит количество входов и выходов модели “черного ящика” для данной системы?
4. Какими признаками должна обладать часть системы, чтобы ее можно было считать элементом?
5. Что общего и в чем различие между понятием элемента и его моделью “черного ящика”?
6. Какова связь между вторым определением системы и ее структурной схемой?
7. Какие особенности системы отражены в ее графе и какие свойства системы не отображаются этой моделью?
8. В чем различие между функционированием и развитием?
9. Каким способом удается компактно описать связь между входом и выходом системы, если значение выхода в данный момент зависит от всей предыстории входа?
10. В чем состоит условие физической реализуемости динамической модели?
11. Какие приемы могут помочь повысить степень полноты содержательных моделей систем?

ИСКУССТВЕННЫЕ И ЕСТЕСТВЕННЫЕ СИСТЕМЫ

§ 4.1. ИСКУССТВЕННЫЕ СИСТЕМЫ И ЕСТЕСТВЕННЫЕ ОБЪЕКТЫ

Один из основных признаков системы состоит в ее структурированности, в целесообразности связей между ее элементами. Понятное и очевидное, если речь идет о системах, созданных человеком, такое описание (определение) системы приводит к сложным вопросам, когда приходится сталкиваться с естественной структурированностью реальных природных объектов. Как красиво и правильно растут кристаллы! Как стройна наша Солнечная система! Как целесообразно устроены живые организмы! Явно налицо необходимые признаки систем. Но в таком случае мы должны вернуться к первому определению системы и поставить перед собой вопрос: на достижение каких целей направлено функционирование этих систем, и если такие цели существуют, то кто их поставил?

На ранней стадии развития общественного сознания стройность и целесообразность природы пытались объяснить сравнительно просто – допустив существование некоторого “внечеловеческого” разума. Это и явилось одной из причин возникновения различных религий, отличие между которыми в основном состоит в том, где находится этот “всезнающий и всемогущий” разум и какие намерения ему присущи. Такая гипотеза нас не устраивает не только потому, что ее принятие сопровождается рядом отрицательных* общественных последствий (побуждением людей к пассивному подчинению “воле всевышнего”, использованием религии реакционными классами, отрицанием или ограничением возможностей науки в познании мира и т.д.), но и потому, что эта гипотеза по существу не отвечает на поставленный вопрос. Ведь предполагаемый “сверхъестественный” разум также система, но тогда кто, в свою очередь, “создал” эту систему и откуда возникают ее цели?

Наука не нуждается в гипотезе о существовании бога. Противоречие, которое имеет место, возникло из-за того, что мы сначала абсолютизировали второе определение системы, а затем вспомнили, что всякая система есть средство достижения цели. Это логическая ошибка, такая же, как допущенная в следующем рассуждении: “Всякий человек имеет четыре конечности, туловище и голову. Следовательно, всякое существо с четырьмя конечностями, туловищем и головой есть человек”.

Попробуем последовательно придерживаться принятой точки зрения. Мы признаем, во-первых, первое и второе определения системы, а во-вторых, что окружающий мир состоит из структурированных объектов, имеющих связанные между собой части. Следовательно, всякая система есть объект, но не всякий объект есть система.

Пример 1. Лес – это объект, но лишь до тех пор, пока его свойства не используются для конкретных целей. Например, при постройке нового города на лесистой территории часть леса в черте города сохранена в целях, поставленных проектировщиками; теперь эта часть леса входит в систему “город”. Или: леспромхозу отведен новый участок леса для разработки. С этого момента такой участок становится элементом системы “леспромхоз”.

У леса нет “целей” украсить город или способствовать сохранению его хорошего микроклимата, предоставить леспромхозу возможность выполнить план или снабдить вас грибами. Лес имеет только свойства, знание и использование которых дает возможность всем этим системам достичь своих целей. У леса есть еще очень много (бесконечно много!) свойств, которые, будучи познанными, потенциально могут быть использованы для создания систем. Подобное рассуждение применимо к любому объекту.

Такой подход обладает следующим достоинством: он подчеркивает, что если человек и может добиться любой цели, то не любым образом. Законы природы (обобщенное выраже-

ГЛАВА ЧЕТВЕРТАЯ

* Очевидно, у религий имеются многие положительные (например, моральные и культурные) аспекты, чем и объясняется их “живучесть”, но нам важны отношения религии и науки.

ние познанных свойств объектов) нельзя нарушать, их можно только использовать. Мы намечаем конечный пункт (цель), но пути его достижения определяет природа. Реализовать поставленную цель можно только соподчиняя, организуя, используя естественные свойства объектов. Недостижимы лишь цели, противоречащие объективным природным закономерностям. Утверждая это, однако, не следует путать действительно объективные свойства природы и конкретные знания о них на современном этапе.

Пример 2. Известно, что из закона сохранения энергии следует невозможность создания вечного двигателя, т.е. системы, дающей на выходе энергию и ничего не имеющей на входе. Поэтому цель создать вечный двигатель недостижима.

Пример 3. Достижима ли цель создания мыслящей машины? Законы природы, которые запрещали бы это, неизвестны. Конечно, отсюда еще не следует, что такая цель достижима, но поиски способов ее достижения не выглядят напрасными; на современном уровне знаний достижение этой цели не исключено, но и не гарантировано.

Подведем итог

Summary

В данном параграфе сделана попытка сохранить понятие системы только для искусственных систем. При таком подходе реальные объекты, несмотря на свою структурированность, системами не признаются, поскольку отсутствуют цели их возникновения и

In this section we have attempted to leave the notion of system to man-made systems. According to this approach, natural objects, in spite of their having a certain structure, are considered not to be systems because of the lack of purposes for their existence.

развития.

§ 4.2. ОБОБЩЕНИЕ ПОНЯТИЯ СИСТЕМЫ.

ИСКУССТВЕННЫЕ И ЕСТЕСТВЕННЫЕ СИСТЕМЫ

Из рассмотренного в предыдущем параграфе следует, что можно рассматривать структурированные объекты, созданные с известной целью, как системы, а все остальные объекты – как “не системы”, как объекты с известными и/или неизвестными свойствами. Однако отсюда еще не следует, что только так и нужно всегда делать. Очевидно, что если неизвестна цель, достижению которой служит система, то это еще не означает, что данный объект перестает быть системой (представьте себе, например, некоторый прибор, назначение которого вам неизвестно).

СТРУКТУРИРОВАННОСТЬ ЕСТЕСТВЕННЫХ ОБЪЕКТОВ

Таким образом, сделаем первый шаг к обобщению: признаем, что могут существовать системы с неизвестными нам целями и/или устройством. Тогда возникает задача изучения этих систем, и понятно, что сам подход и некоторые методы такого изучения существенно отличаются от тех, которые применяются при рассмотрении обычных объектов: ведь нужно прежде всего познать данный объект как систему, выявить ее назначение (цель) или способ действия.

Создание специальных методов для исследования систем – одна из важнейших задач всей системологии. Ряд таких методов уже создан (в дальнейшем мы познакомимся с ними), другие еще только разрабатываются и совершенствуются, третьи пока еще ждут своих создателей (может быть, это будет именно вы). Поскольку такие методы уже имеются, почему бы не применить их к исследованию сложных объектов, существующих независимо от человека, не созданных людьми? Почему бы, например, не изучать растения как системы, не считать “настоящими” системами нервную или сердечно-сосудистую системы человека, муравейник или лес вместе со всеми животными в данном лесу? Более того, уже сейчас ясно, что, например, не считая мозг сложной системой, нельзя познать то, как все же мы мыслим. Если такой подход практически полезен (а так оно и есть), то отказываться от него вредно. Итак, мы

NATURAL
естественный
LAW, RULE,
PRINCIPLE
закон
REGULARITY,
CONFORMITY
закономерность
ARTIFICIAL,
MAN-MADE
искусственный

Попытка сохра-
нить термин “си-
стема” толь-
ко для искус-
ственных кон-
струкций приво-
дит к трудностям.
Например: 1) на
каком уровне чле-
нения системы ее
часть оказывается
уже не искус-
ственной систе-
мой, а естествен-
ным объектом? 2)
кем (системой или
объектом) нам
считать самих се-

пришли к тому, чтобы сделать второй шаг – признать, что существуют **искусственные** (созданные человеком) и **естественные** (возникшие в природе без участия человека) системы. Полезность этого шага видна хотя бы из возникновения таких наук, как экология, бионика, биокибернетика, нейрокибернетика, исследующих естественные системы в различных аспектах.

**//(биосоциальные системы являются искусственными или есте-
ственными, согласно этой классификации. Пример: прайд, стая,
пара особей. Взаимодействие по схожим законам с современными
созданными человеком соц системами)//**

Признав существование естественных систем, не следует противоречить себе в вопросе о целях систем. Теперь мы должны расширить понятие цели так, чтобы оно охватывало предыдущее понятие цели искусственной системы, определяло цель естественной системы и указывало, в чем состоит их общность и различие.

СУБЬЕКТИВНЫЕ И ОБЪЕКТИВНЫЕ ЦЕЛИ

Цель искусственной системы была определена как идеальный образ желаемого результата ее деятельности, т.е. как образ того, что должно бы быть, должно бы осуществиться (частица “бы” подчеркивает тот факт, что цели бывают как осуществимые, так и неосуществимые; см. § 4.1). Такой идеальный образ будущего состояния системы и окружающей ее среды естественно назвать **субъективной целью**. Теперь представим себе, что прошел срок, который был отведен для реализации субъективной цели; система и непосредственно окружающая ее среда оказались в некотором реализовавшемся состоянии. В это состояние система пришла объективно, т.е. в результате реализации объективных закономерностей. По отношению к прошлому моменту, когда результат еще только планировался, это состояние можно назвать **объективной целью** системы. Другими словами, будущее реальное состояние системы мы представляем себе как ее объективную цель. В этом и состоит обобщение: теперь мы можем рассматривать субъективные и объективные цели системы, т.е. такие, которые ставит человек, и такие, которые реализует природа.

Теперь совсем просто сделать шаг к тому, чтобы признать, что у естественных неодушевленных систем есть только объективные цели. Правда, при этом возникает затруднение в объяснении поведения некоторых животных, но оно носит не принципиальный характер, а, по-видимому, связано с недостаточностью знаний о степени абстракции, которой обладают животные, и о нашем собственном мышлении: пока еще далеко не всегда мы можем определить, где кончается поведение живых систем на уровне рефлексов, инстинктов и генетически определенных программ, а где начинается поведение на основе абстрактных моделей, которые уже можно назвать идеальными.

Итак, мы пришли к пониманию того, что любой объект можно рассматривать как систему, но это означает, что этот объект рассматривается под определенным углом зрения, т.е. только в определенном отношении, а именно в отношении к цели.

Отметим, что интерпретация будущих состояний естественных объектов как “целевых” состояний имеет длинную историю [10]. Еще Аристотель развивал телеологию, согласно которой “природа” вещи предопределяет ее естественное положение и движение объясняется предполагаемым стремлением ее достичь этого положения. Галилей и его последователи, отвергая такой подход, стремились изучать не конечное состояние, а его причины, т.е. комбинации сил, действующих в данный момент. Любопытно подчеркнуть, что дальнейшее раз-

вление физики привело к установлению того факта, что реальное движение и состояния физических тел подчиняются определенным принципам экстремальности: минимума действия, минимума потенциальной энергии и т.п. Далее оказалось, что физика “чисто белого ящика” испытывает затруднения уже при рассмотрении $N \geq 3$ взаимодействующих тел: такие задачи не получили аналитического решения. Прогресс был достигнут “физикой черного ящика” – статистической физикой. Примечательно, что и тут был обнаружен экстремальный принцип – принцип максимума энтропии. Открытие принципов экстремальности природных процессов следует, видимо, рассматривать как проявление объективной системности природы. Наличие субъективных и объективных целей можно рассматривать как одно из проявлений общности и различия между системностью мышления и системностью природы.

Подведем итог

Summary

В гл. 1 мы с философских позиций уже пришли к выводу о всеобщей системности природы. Теперь мы приходим к тому же выводу с позиций практических. Дополнив понятие субъективных целей понятием целей объективных (как осуществимых будущих состояний), мы получаем возможность рассматривать не только искусственные, но и естественные системы.

In Chapter 1 we have reached certain conclusions about the general systematicity of nature – from the point of view of philosophy. Now we arrive at the same of view. To the notion of subjective purposes we add the notion of objective purposes (as future actual states) and thus obtain the possibility of considering not only artificial but – natural systems as well.

§ 4.3. РАЗЛИЧНЫЕ КЛАССИФИКАЦИИ СИСТЕМ

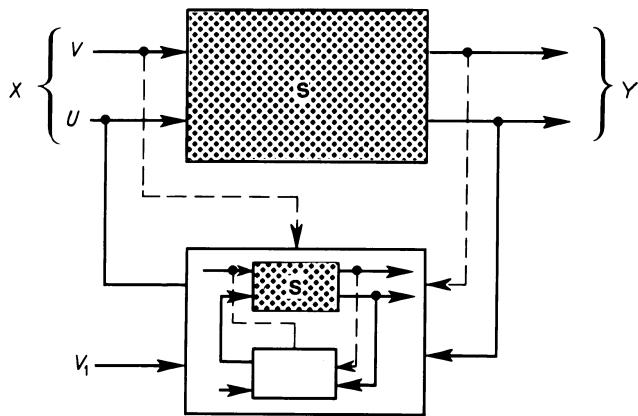
Начиная сравнивать и различать системы, считать одни из них одинаковыми, другие – различными, мы тем самым вводим и осуществляем их *классификацию*. Например, в предыдущем параграфе были введены классы искусственных и естественных систем.

Важно понять, что классификация – это только модель реальности. (Можно было бы даже сказать – первичная, простейшая модель, если бы не тот факт, что вследствие многоуровневой, кратной, вложенной классификации полученная модель может и не отвечать понятию “простейший”.) Поэтому классификацию не следует абсолютизировать: реальность всегда сложнее любой модели. Например, исчезают ли введенные нами классы все возможные типы систем? Так, можно ли считать, что в систему “рыболовецкий совхоз” не входит промысловая акватория, в “леспромхоз” не входят лесные делянки, а “всадник” мыслим без лошади? Естественно, напрашивается введение еще одного класса **смешанных систем**, объединяющих искусственные и естественные подсистемы. Теперь такая классификация уже может претендовать на полноту*. Вообще, *полнота классификации* является предметом особого внимания при ее построении. Иногда есть уверенность в полноте вводимой классификации, иногда нет (и тогда имеет смысл вводить класс “все остальное”), а иногда наша уверенность оказывается самонадеянным незнанием (например, “очевидное” разделение всех людей на мужчин и женщин “не срабатывает” не только для гермафродитов, но и на генетическом уровне: в начале 80-х годов в международном спорте был введен генетический тест в некоторых женских олимпийских видах). Это еще одна иллюстрация модельности, условности всякой классификации.

КЛАССИФИКАЦИЯ СИСТЕМ ПО ИХ ПРОИСХОЖДЕНИЮ

Часто оказывается необходимым провести разграничение внутри одного класса, не отказываясь тем не менее от общности в его рамках. Так появляются подклассы, что приводит к

* Правда, следовало бы расширить понятие “искусственное”, чтобы классификация не нарушилась, если когда-нибудь обнаружится внеземная цивилизация...



4.2

Схема функционирования управляемой системы

приведенным в § 1.2; неполнота на этом уровне связана, например, с еще незавершенным развитием систем искусственного интеллекта. В качестве примеров подклассов смешанных систем можно привести эргономические системы (комплексы машина – человек-оператор), биотехнические (системы, в которые входят живые организмы и технические устройства) и организационные системы (состоящие из людских коллективов, которые оснащены необходимыми средствами). Классификация естественных систем ясна из рис. 4.1; ее неполнота очевидна. Например, не решен окончательно вопрос о том, куда следует отнести вирусы: к живым или неживым системам. Или, скажем, идея В.И. Вернадского о ноосфере шире рамок экологических и социальных систем.

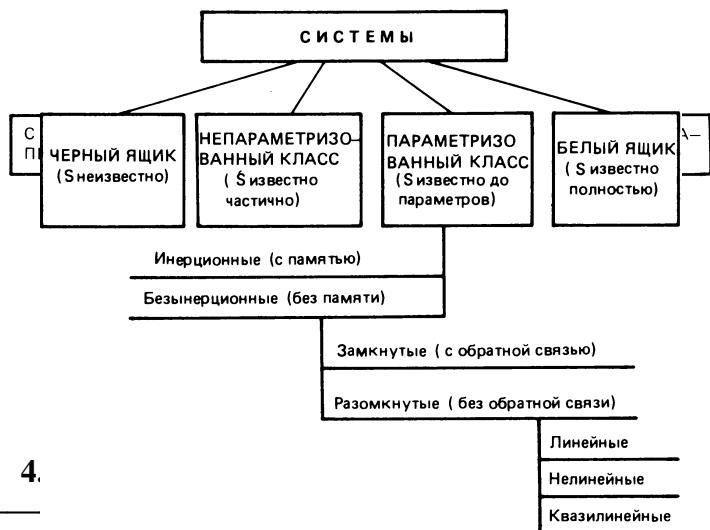
В гл. 2 был подчеркнут целевой характер моделей. Не является исключением и классификация: новые цели, учет новых различий между системами порождают и новые классификации. Чтобы как-то упорядочить подходы к классификации систем, воспользуемся общей схемой функционирования системы (рис. 4.2), выделив отдельно систему S , подлежащую управлению U , и управляющую систему, которая это управление вырабатывает. Подчеркнем, что для выработки управления U требуется предсказание его последствий, т.е. нужна модель всей ситуации; с помощью этой модели управляющая система и определяет, какое управление подать на управляемый вход системы. Это иллюстрирует рис. 4.2, где схема изображена еще раз внутри управляющего блока.

И методы нахождения управления U и способы его осуществления, и сам результат управления в немалой степени определяются тем, что известно о системе и что учитывается при выработке управления, т.е. тем, какова модель управляемой системы, и тем, в какой сте-

многоуровневой, иерархической классификации. При необходимости такая классификация может быть продолжена без изменения ее верхних уровней. Например, двухуровневая классификация систем по происхождению изображена на рис. 4.1. Если полнота классификации первого уровня логически ясна, то второй уровень на полноту не претендует. Разбиение искусственных систем соответствует рассуждениям,



4.1



4.

Фрагмент классификации систем по описанию переменных 4.4

Фрагмент классификации систем по типу их операторов

На рис. 4.3 приведена трехуровневая классификация систем по типу входных (X), выходных (Y) и внутренних (Z) (если описание ведется не на уровне “черного ящика”) переменных. Принципиально разных подходов требуют переменные, описываемые качественно и количественно, что и дает основание для первого уровня классификации. Для полноты введен третий класс, к нему отнесены системы, у которых часть переменных носит качественный характер, а остальные являются количественными. На следующем уровне классификации систем с качественными переменными различаются случаи, когда описание ведется средствами естественного языка, и случаи, допускающие более глубокую формализацию. Второй уровень классификации систем с количественными переменными вызван различиями в методах дискретной и непрерывной математики, что и отражено в названиях вводимых классов; предусмотрен и случай, когда система имеет как непрерывные, так и дискретные переменные. Для систем со смешанным количественно-качественным описанием переменных второй уровень является объединением классов первых двух ветвей и на рисунке не приводится. Третий уровень классификации одинаков для всех классов второго уровня и изображен только для одного из них. Различия между классами третьего уровня будут рассмотрены в гл. 6.

ТИПЫ ОПЕРАТОРОВ СИСТЕМЫ

Следующая классификация (рис. 4.4) – по особенностям оператора S системы, т.е. классификация типов связей между входными и выходными переменными. Материал для этого типа классификации дает § 3.7. На первом уровне расположены классы систем, отличающиеся степенью известности оператора S . Ветвь “черного ящика” на этом уровне кончается: S считается вообще неизвестным. Чем больше сведений об S мы имеем, тем больше различий можно рассмотреть и тем более развитой окажется классификация. Например, информация об S может носить настолько общий характер, что модель нельзя привести к параметризованной функциональной форме. Так, может быть известно, что в соотношении $Y = S(X)$ функция S непрерывна, монотонна или симметрична; отсюда не следует никаких конкретных выводов о функциональном виде этой зависимости.

Непараметризованный класс операторов системы (второй блок первого уровня) и соответствует подобным ситуациям с очень скучной априорной информацией об S .

Наши знания об S могут соответствовать уровню, который позволяет предложить параметрическую модель этого оператора, т.е. записать зависимость $y(t)$ от $x(t)$ в явной форме с

пени эта модель соответствует реальной системе. Рассматривая разные аспекты этого соответствия, можно строить разные классификации систем. Например, представляют интерес следующие классификации: по описанию входных и выходных процессов; по описанию оператора S системы; по типу управления; по обеспеченности управления ресурсами.

ТИПЫ ПЕРЕМЕННЫХ СИСТЕМЫ

Рассмотрим их подроб-



4.5

Разделение распределенных систем на классы по способу управления

Классификация множества объектов на классы является первым, простейшим актом моделирования этого множества. Как и всякая модель, классификация, отражая объективные различия, в то же время носит целевой характер и является относительной, услов-

непараметризованных операторов связана с типом имеющейся информации). Второй, третий и четвертый уровни ясны из самого рисунка. Конечно, классификация может быть продолжена (например, линейные операторы принято делить на дифференциальные, интегральные и суммарно-разностные), но мы на этом остановимся.

ТИПЫ СПОСОБОВ УПРАВЛЕНИЯ

Следующая классификация систем – по способам управления – приведена на рис. 4.5. Первый уровень классификации определяется тем, входит ли управляющий блок в систему или является внешним по отношению к ней; выделен также класс систем, управление которыми разделено и частично осуществляется извне, а частично – внутри самой системы.

Независимо от того, включен ли в систему и вынесен ли из нее управляющий блок, можно выделить четыре основных типа (способы) управления, что и отражено на втором уровне классификации (рис. 4.5). Эти способы различаются в зависимости от степени известности траектории*, приводящей систему к цели, и возможности управляющей системы удерживать управляемую систему на этой траектории.

Первый (простейший) случай имеет место тогда, когда нужная траектория известна точно, а следовательно, априори известно и правильное программное управление $u_0(t)$. В таком случае это управление можно осуществлять, не обращая внимание на развитие событий; ведь и так известно, как они должны (и будут) развиваться. Стрельба из ружья, работа ЭВМ по программе, рост зародыша живого организма, пользование телефоном-автоматом являются примерами такой ситуации.

Однако случаи, когда управление $u_0(t)$ без обратной связи, только по априорной информации, приводят к достижению цели, возможны лишь при том условии, что все будет происходить именно так, как предписывает заданная траектория.

$$\langle x_0(t), y_0(t) \rangle = \langle x_0(t) = \langle u_0(t), v_0(t) \rangle, y_0(t) \rangle.$$

Чаще оказывается, что процессы на неуправляемых входах $v_0(t)$ отличаются от ранее предполагаемых, либо существенным оказывается действие неучитываемых входов и система “сходит с нужной траектории”. Пусть имеется возможность наблюдать текущую траекто-

точностью до конечного числа параметров $\theta = (\theta_1, \dots, \theta_k)$: $y(t) = S(x(\cdot), \theta)$. Этому соответствует третий блок первого уровня классификации. Наконец, если эти параметры также заданы точно, то всякая неопределенность исчезает и мы имеем системы с полностью определенным оператором, т.е. “белый ящик”.

Дальнейшие уровни классификации на рис. 4.4 приведены только для последующих двух ветвей (“черный ящик” не подлежит дальнейшей классификации, а классификация

рию $y(t)$, находить разность $y(t) - y_0(t)$ и определять дополнительное к программному управление, которое в ближайшем будущем возвратит выходы системы на нужную траекторию $y_0(\cdot)$. Такой способ управления называется **регулированием**, а соответствующие системы выделены во второй класс второго уровня классификации (рис. 4.5). Например, этому классу принадлежит управление, которое осуществляется операторами-станочниками, регулятором Уатта, автопилотом, судовым авторулевым, в рефлекторных реакциях животных и т.п.

Следующие способы управления и соответствующие им типы систем возникают в связи с необходимостью управления в условиях, когда либо невозможно задать опорную программную траекторию на весь период времени, либо уклонение от нее столь велико, что невозможно вернуться на нее (регулирование обычно осуществляется при “малых” в известном смысле уклонениях* $y(t) - y_0(t)$). Теперь нам необходимо спрогнозировать текущую траекторию $y(t)$ на будущее и определить, пересечет ли она целевую область Y^* . Управление состоит в *подстройке параметров* системы до тех пор, пока такое пересечение не будет обеспечено. Этому и соответствует третий класс систем. Примерами такого управления являются процессы адаптации живых организмов к изменяющимся условиям жизни, работа пилотов и шоферов, адаптивные и автоматизированные системы управления и т.п.

Иногда может оказаться, что среди всех возможных комбинаций значений управляемых параметров системы не найдется такой, при которой ее траектория пересечет целевую область. Это означает, что цель для данной системы недостижима. Но, может быть, она достижима для другой системы? Сказанное дает еще один способ управления: изменять структуру системы в поисках такой, при которой возможно попадание в целевую область. По существу, имеет место перебор различных систем, но это системы с одинаковыми выходами Y , создаваемыми не произвольно, а в соответствии с наличными средствами. Такое управление, называемое **структурной адаптацией** [11], выделим в четвертый класс классификации второго уровня (рис. 4.5). Примерами реализации указанного управления являются гибкие автоматизированные производства, вычислительные сети, сельскохозяйственные машины со сменными навесными и прицепными устройствами, **мутации организмов в процессе естественного отбора, организационные изменения в государственном аппарате** и т.д.

На этом закончим классификацию по типам управления, хотя ее можно развивать и дальше, не только “вглубь”, но и “вширь”. Например, в [11] предлагается рассматривать и тот случай, когда всевозможные преобразования структуры также не приводят к реализации цели. Это означает принципиальную недостижимость цели при имеющихся ресурсах, и предлагается отказ от старой цели и задание новой рассматривать как “управление (адаптацию) по целям”. Мы воздержимся от такого обобщения, хотя примерно так действует руководитель коллектива, ставя перед подчиненными посильные для них цели. Условимся этап выявления цели считать не входящим в понятие управления, предшествующим ему во времени.

Наиболее интересной с нашей точки зрения является классификация систем, рассмотренная в следующем параграфе, поскольку она придает достаточно конкретный смысл терминам “большая система” и “сложная система”. Это очень важные понятия во всей системологии, поэтому этой классификации и посвящен отдельный параграф.

* Имеется в виду траектория системы в “фазовом пространстве”, т.е. $\langle x(t), y(t) \rangle$; траектория $\langle x_0(t), y_0(t) \rangle$, приводящая к цели, пересекает целевую область Y^* : соответствующее этому управление $u_0(t)$

назовем правильным.

* Быть может, разница между этой и предыдущей группами способов управления станет очевидной, если провести аналогию с двумя способами турпохода по неизвестной территории; можно наметить маршрут заранее и в походе придерживаться выбранного пути, а можно идти без заранее составленного плана, ориентируясь лишь по компасу и учитывая особенности рельефа, стремясь достичь заданного пункта.

Подведем итог

Основным результатом данного параграфа следует считать представленные на рис. 4.1 – 4.5 классификации систем. Попутным, но важным результатом является также понимание всякой классификации как простейшей модели рассматриваемого множества объектов.

§ 4.4. О БОЛЬШИХ И СЛОЖНЫХ СИСТЕМАХ

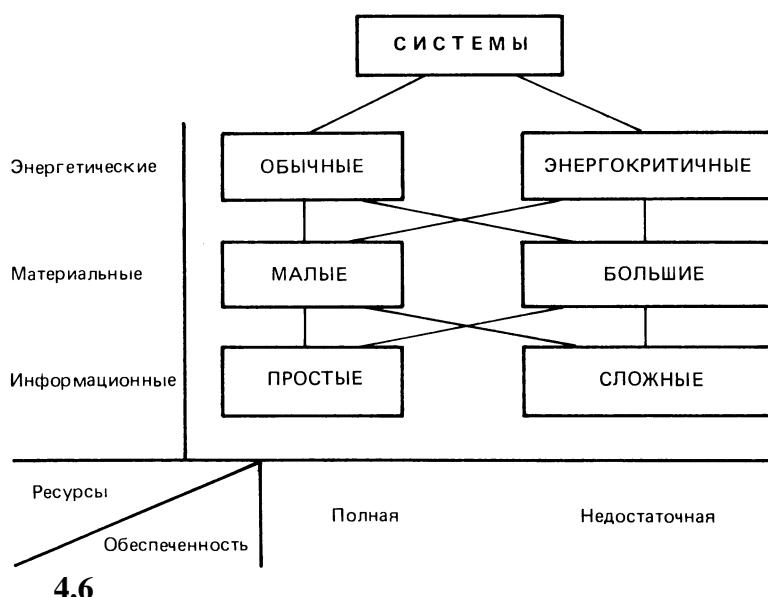
Как подчеркивалось в § 4.3, нужное управление $u_0(t)$ отыскивается с помощью отбора среди возможных управлений, сравнения по каким-то критериям последствий каждого из них. Определить эти последствия и сравнить их можно лишь в том случае, если в управляющей системе имеется модель управляемой системы, на которой и испытываются разные управления (см. рис. 4.2).

Для того чтобы модель “заработала”, или, как говорят, была актуализирована, необходимы затраты ресурсов: модель нужно не только воплотить в каком-то реальном виде, но и обеспечить, чтобы она позволяла получать решение нужного качества и к нужному моменту времени (ведь даже самое лучшее решение становится ненужным, если оно появляется позже, чем это необходимо).

РЕСУРСЫ УПРАВЛЕНИЯ И КАЧЕСТВО СИСТЕМЫ

В реальности же оказывается, что имеющиеся ресурсы не всегда позволяют обеспечить полное выполнение этих условий. Поэтому имеют место принципиально разные ситуации в зависимости от того, в какой степени обеспечено ресурсами управление; системы при этом выступают как качественно различные объекты управления. Это и отражено в классификации систем, приведенной на рис. 4.6.

Рассмотрим, например, энергетические затраты на актуализацию модели и выработку управления. Обычно они настолько малы по сравнению с количеством энергии, потребляемой или производимой в управляющей системе, что их просто не принимают во внимание. Однако представим себе случай, когда, во-первых, управляющая и управляемая системы пытаются от одного ограниченного источника энергии, и, во-вторых, энергопотребление обеих систем имеет одинаковый порядок: возникает интересный и нетривиальный класс задач о



Summary

The main accomplishment of this section is the design of classifications of the systems presented in Figs. 4.1 to 4.5. An auxiliary but also important result is an understanding of any classification as the simplest model of the set of objects under consideration.

наилучшем распределении энергии между ними. С подобными задачами приходится иметь дело не часто, но в ответственных случаях: выполнение энергоемких задач автономными системами (например, космическими аппаратами или исследовательскими роботами), некоторые эксперименты в физике частиц высоких энергий и т.п.

Следующее деление систем связано с материальными ресурсами, затрачи-

ваемыми на актуализацию модели. В случае моделирования на ЭВМ это объем памяти и машинное время; такие ресурсы лимитируют возможности решения задач большой размерности в реальном масштабе времени. К подобным задачам приводится моделирование ряда экономических, метеорологических, организационно-управленческих, нейрофизиологических и других систем. *Системы, моделирование которых затруднительно вследствие их размерности*, будем называть **большими**. Существует два способа перевода больших систем в разряд малых [5]: разрабатывать более мощные вычислительные средства либо осуществлять декомпозицию многомерной задачи на совокупность связанных задач меньшей размерности (если природа системы это позволяет).

LARGE-SCALE

большие (системы)

DISTINCTION,

DISCRIMINATION

различие

DIMENSION

размерность

RESOURCES

ресурсы

COMPLEX

сложный

Для выбора управляющего воздействия на систему необходимо сравнить последствия разных возможных воздействий, не выполняя их реально. Такую возможность мы получаем, используя модель всей ситуации. Осуществление моделирования требует затраты ресурсов, наличие которых всегда ограничено. В зависимости от того, полностью или лишь частично удовлетворяется потребность моделирования в ресурсах, управляемая система выступает в принципиально различном качестве по отношению к управлению. Эти различия являются основой

Второй способ – сменить цель, что в технических системах обычно неэффективно (подобно забиванию гвоздей магнетроном), но в отношениях между людьми это часто единственный выход.

Как и все предыдущие, такая классификация систем при необходимости может быть развита – либо благодаря более подробному рассмотрению видов ресурсов, либо в результате введения большего числа градаций степени обеспеченности ими. Например, С. Бир [1] пред-

Наконец, третий тип ресурсов – информация – дает основание для еще одной классификации систем. Имеющаяся о системе информация, сколько бы ее ни было, представлена в концентрированном виде как та самая модель, об использовании которой идет речь. Признаком простоты системы, т.е. достаточности информации для управления, является успешность управления. Однако если полученное с помощью модели управление приводит к неожиданным, непредвиденным или нежелательным результатам, т.е. отличающимся от предсказанных моделью, это интерпретируется как сложность системы, а объясняется как недостаточность информации для управления. Поэтому **сложной системой** мы будем называть *систему, в модели которой не хватает информации для эффективного управления*.

Таким образом, свойство простоты или сложности управляемой системы является свернутым отношением между нею и управляющей системой, точнее, между системой и ее моделью. Это отношение носит объективный характер: например, кодовый замок действительно качественно различен для того, кто знает код и кто его не знает; каждому человеку родной язык кажется проще иностранного; люди, умеющие и не умеющие водить автомобиль, имеют объективно разные возможности обращения с ним. Имеется два способа перевода системы из разряда сложной в разряд простой. Первый состоит в выяснении конкретной причины сложности, получении недостающей информации и включении ее в модель; это и является основной задачей науки, познания вообще и системного анализа в частности. Так, У. Эшби пишет: “Не подлежит сомнению, что наука упрощения обладает своими методами и тонкостями. Я убежден, что в будущем теоретик систем должен стать экспертом по упрощению” [16].

лагает выделить в отдельный класс очень сложные системы (к этому классу он относит мозг, экономику, фирму).

РАЗЛИЧЕНИЕ БОЛЬШИХ И СЛОЖНЫХ СИСТЕМ

Чтобы не возникло недоразумений при чтении специальной литературы, необходимо отметить, что термины “большая” и “сложная” система в системологии окончательно еще не установились и при всей, по нашему мнению, естественности придаваемого им здесь смысла имеются разные варианты их употребления. Некоторые авторы вообще не используют эти понятия [4], другие используют их как синонимы [2; 3; 8], третьи считают разницу между ними чисто количественной [6], четвертые связывают сложность с некоторыми особенностями самих систем [11; 14]. Попытки “примирить” эти подходы, объединить их приводят к путанице. Например, в [15] читаем:

“Сложные системы управления – собирательное название систем, состоящих из большого числа взаимосвязанных элементов (...), которые нельзя описать корректно математически (...), для изучения которых необходимо было бы решать задачи с непомерно большим объемом вычислений. (...) Понятия “сложная система” и “большая система” не являются тождественными, так как последний термин характеризует только одну черту “сложности” – размерность системы”.

Смысл понятий “большой” и “сложный”, вытекающий из предложенной классификации, не является совершенно новым. Многие авторы из других соображений так или почти так и использовали эти термины.

У. Эшби: “Наблюдатель говорит, что система “очень большая”, если она в чем-либо побивает его своим богатством” [16].

Дж. Б. Форрестер: “Я сделал попытку показать, как сложные системы приводят нас к ошибкам из-за того, что наша интуиция и оценки заставляют ожидать поведения, отличного от того, которым они в действительности обладают” [13].

И. Пригожин: “Очень часто отклик системы на возмущение оказывается противоположным тому, что подсказывает нам интуиция. Наше состояние обманутых ожиданий в этой ситуации хорошо отражает термин “антиинтуитивный”: “Эта проклятая штука ведет себя не так, как должна бы вести!”. И далее. “Единственной специфической особенностью сложных систем является то, что наше знание о них ограничено и неопределенность со временем возрастает” [9].

Чтобы подробнее пояснить вытекающую из сделанной классификации разницу между большими и сложными системами, отметим, что возможны все четыре комбинации; существуют системы: 1) “малые простые”; 2) “малые сложные”; 3) “большие простые”; 4) “большие сложные”. Приведем примеры систем всех видов (в том же порядке):

1) исправные бытовые приборы (утюг, часы, холодильник, телевизор и т.д.) – для пользователя; неисправные – для мастера; шифрзамок – для хозяина сейфа;

2) неисправный бытовой прибор – для пользователя;

3) шифрзамок для похитителя – система простая (требуется лишь перебор вариантов) и большая (имеющаяся на вскрытие сейфа времени может не хватить на перебор вариантов); точный прогноз погоды; полный межотраслевой баланс (определяется с задержкой на годы);

4) мозг; экономика; живой организм.

ДРУГИЕ ПОДХОДЫ К ПОНЯТИЮ СЛОЖНОСТИ

Обсуждение проблемы сложности было бы неполным, если не упомянуть и о другом подходе к ней. В науке нередко бывает так, что вместо содержательного определения чего-то вводится его количественная мера, что позволяет продвинуться вперед в изучении, отвечая на вопрос “как?” и не затрагивая вопрос “что?”. Примерами являются теоретическая механика (использующая не понятие движения, а его количественные характеристики), теория информации (не требующая определения самой информации, так как вполне достаточно понятия количества информации) и т.д. Аналогичные попытки сделаны и по введению количе-

ственной меры сложности. Пока такой единой меры построить не удалось. Одна из причин (но не единственная) состоит в том, что термин “сложный” употребляется и как синоним “большой” (например, теоретико-множественное понятие сложности системы связывается с мощностью множества ее элементов). Наиболее известные концепции, в которых сложность выражается некоторым числом, таковы:

логическая концепция, основанная на анализе свойств предикатов, характеризующих систему;

теоретико-информационная концепция, связывающая сложность системы с ее энтропией;

статистическая концепция, характеризующая сложность через меру различимости распределений вероятностей;

алгоритмическая концепция, определяющая сложность как длину алгоритма воссоздания системы;

теоретико-множественная концепция, отождествляющая сложность системы с числом ее элементов;

вычислительная концепция, “привязывающая” алгоритмическую сложность к средствам вычислений.

Подробнее эти концепции рассмотрены в [12; 18].

Подведем итог

Главным результатом является определение сложности системы как следствия недостаточности информации для желаемого качества управления системой. Это не только упорядочивает терминологию (благодаря введению четкого различия между терминами “большая система” и “сложная система”), но и указывает пути преодоления сложности.

Summary

The main accomplishment of this section is the definition of system complexity as a consequence of the lack of information, which prevents the desired quality of system control from being reached. This not only makes terminology more exact (due to the clear distinction between the terms “large systems” and “complex systems”), but also points out ways of lessening this complexity.

ЗАКЛЮЧЕНИЕ

Если рассматривать системы только как создаваемые человеком средства достижения поставленных им целей, то все остальные природные предметы и их совокупности являются естественными объектами, познанные свойства которых человек может использовать, включая их в искусственные системы (см. § 4.1).

Такое ограничительное толкование понятия “система” наталкивается на ряд трудностей: наличие у естественных объектов структурированности и упорядоченной взаимосвязанности их частей – характерных признаков систем; необходимость отрицать системность искусственного сооружения, как только неизвестна цель, ради которой оно создано; невозможность признания системности самого человека и всей природы.

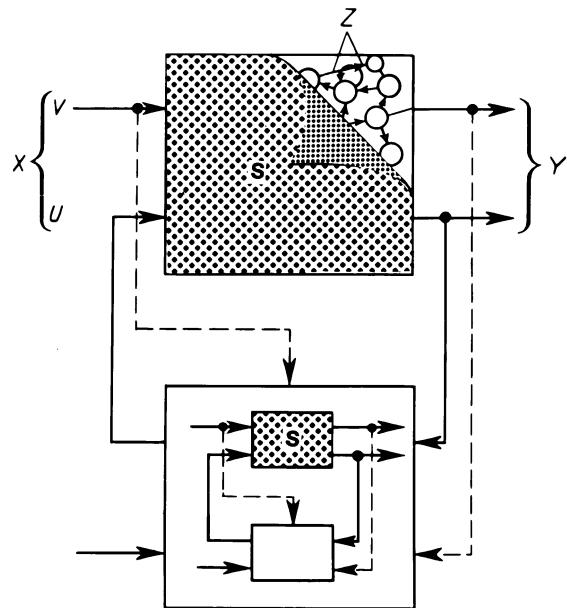
Эти сложности исчезают, если признать не только то, что искусственная система остается системой, даже если ее цель неизвестна, но и то, что вся природа объективно системна, т.е. что наряду с искусственными существуют естественные системы. Эта идея требует обобщения понятия цели – наряду с целями субъективными (желательными состояниями) вводится понятие объективной цели как будущего реального состояния, в которое придет объект через какое-то время (см. § 4.2).

Споры о существовании объективных целей идут в науке давно. Например, в биологии сложилось положение, при котором стремление объяснить поведение живых организмов через цель (телеология) пришло в противоречие с понятием о сугубой субъективности целей. Как выразился М. Месарович, “телеология – это дама, без которой не может прожить ни один биолог, но с которой он стыдится показаться в обществе” [7]. Введение понятия объективной цели как будущего состояния, по-видимому, снимает это противоречие, позволяя говорить не только об искусственных, но и о естественных системах.

Означает ли это, что “не систем” не существует? Чтобы акцентировать внимание на особенностях системного подхода, скажем без оговорок: да, означает. Другое дело, что мы можем (или по незнанию иначе не можем) рассматривать некий объект и обращаться с ним, не полностью считаясь или совсем не считаясь с его системностью. Рано или поздно недостаточная системность нашего подхода выльется в появление проблем.

После окончательного признания всеобщей системности мира целесообразно обсудить характерные различия между всевозможными системами. Это делается с помощью классификации систем. Отметив, что можно ввести много классификаций, мы рассмотрели несколько из них, важных для целей данного курса (см. § 4.3). Классификацию, в которой появляется возможность различения “больших” и “сложных” систем, мы обсудили подробнее в § 4.4. Конкретизация понятия сложности проясняет и происхождение проблем: осознание сложности и приводит к формулировке проблемы. Решение проблемы есть упрощение ситуации.

Для облегчения восприятия и запоминания приведенных классификаций обратимся еще раз к схеме управляемой системы, слегка дополнив ее (рис. 4.7). Эта схема позволяет выделить следующие аспекты, характеризующие управление системой: описание природы (происхождения) системы S ; задание типов переменных X, Y, Z ; конкретизация типа оператора S ; описание способа управления (получения U); задание условий получения U (обеспеченности управления ресурсами). Каждый из перечисленных аспектов дает основания для построения конкретной классификации систем. Выявление других аспектов описания может привести к новым полезным классификациям систем.



4.7

Схема взаимодействия управляемой и управляемой

ЛИТЕРАТУРА

1. Бир С. Кибернетика и управление производством. – М.: Наука, 1965.
2. Гуд Г.Х., Макол Р.Э. Системотехника. Введение в проектирование больших систем. – М.: Сов. радио, 1962.
3. Калашников В.В. Сложные системы и методы их анализа. – М.: Знание, 1980.
4. Квейд Э. Анализ сложных систем. – М.: Сов. радио, 1969.
5. Лэсдон Л. С. Оптимизация больших систем. – М.: Наука, 1975.
6. Меерович Г.А. Эффект больших систем. – М.: Знание, 1985.
7. Месарович М. Теория систем и биология. Точка зрения теоретика.- В сб.: Теория систем и биология – М.: Мир, 1971.
8. Нечипоренко В.И. Структурный анализ систем. – М.: Сов. радио, 1977.
9. Пригожин И., Стенгерс И. Порядок из хаоса. – М.: Прогресс, 1986.
10. Рапопорт А. Математические аспекты абстрактного анализа систем. – В сб.: Исследования по общей теории систем. – М.: Мир, 1969.
11. Растиригин Л.А. Адаптация сложных систем. – Рига: Зинатне, 1981.
12. Уемов А.И. Системный подход и общая теория систем. – М.: Мысль, 1978.
13. Форрестер Дж. Б. Антиинтуитивное поведение сложных систем. – В сб.: Современные проблемы кибернетики. – М.: Знание, 1977.
14. Шрейдер Ю.А., Шаров А.А. Системы и модели. – М.: Радио и связь, 1982.

15. Энциклопедия кибернетики. – Киев: Укр. энц., 1975. Т. 11.
16. Эшби У.Р. Введение в кибернетику. – М.: ИЛ, 1959.
17. Эшби У..Р. Несколько замечаний. – В сб.: Общая теория систем. – М.: Мир, 1966.
18. Юдин Д.Б., Юдин А.Д. Число и мысль. – М.: Знание, 1985. Вып. 8.

УПРАЖНЕНИЯ

§ 4.1

- Приведите несколько примеров, иллюстрирующих использование свойств естественных объектов в искусственных системах.
- В фантастическом романе Стругацких “Пикник на обочине” описываются вещи, оставленные космическими пришельцами на месте их временной стоянки на Земле. Например, предмет, получивший земное название “пустышка”, состоял из двух металлических дисков, разделенных небольшим расстоянием и ничем видимым не скрепленных, но которые нельзя было ни сблизить, ни разобщить. Приведите аргументы за то, что этот предмет следует считать искусственной системой.

§ 4.2

- Как показывает история, в обществе многие системы сначала возникают естественным образом и лишь потом осознается факт их существования, выявляются и формулируются их цели. Обсудите в качестве примеров этого факта появление социальных классов, мировой системы связи, неформальных общественных групп. Приведите свои примеры.
- Обсудите естественную системность дерева, Солнечной системы, озера, других природных объектов.
- Приведите и обсудите свои примеры достижимых и недостижимых субъективных целей.

§ 4.3

- Попробуйте перечислить цели, в которых может использоваться каждая из приведенных классификаций систем.
- Подкрепите известными вам примерами классификацию, приведенную на рис. 4.5.
- Попытайтесь продолжить любую из ветвей классификации, которая вызвала у вас интерес. Вернитесь к этому упражнению после прочтения гл. 8.

§ 4.4

- Обсудите на новых примерах разницу между большими и сложными системами.

Вопросы для самопроверки

1. Почему целевой характер искусственных систем не позволяет без оговорок перенести понятие системы на естественные объекты?
2. Как обобщить понятие цели, чтобы в него входило не только понятие субъективной цели, но и объективная тенденция процессов, происходящих с любым естественным объектом?
3. Что, кроме возможности ввести понятие естественных систем, дает такое обобщение?
4. Какие особенности управления дают основания для различия программного управления, регулирования, параметрической адаптации и структурной адаптации?
5. На что расходуются ресурсы в процессе выработки управляющего воздействия? Почему степень обеспеченности управления ресурсами определяет качественное состояние управляемой системы?
6. Чем отличается “большая” система от “сложной”?
7. Не смущает ли вас то, что сложность системы оказывается не атрибутом системы, а отношением между системой и ее моделью, используемой в управлении? Помните ли вы, как рассматривалась связь между свойством и отношением в § 3.5?

§ 5.1. ИНФОРМАЦИЯ КАК СВОЙСТВО МАТЕРИИ

Современное понимание того, что такая информация и какую роль она играет в искусственных и естественных системах, сложилось не сразу; оно представляет собой совокупность знаний, полученных разными науками: физикой, биологией, философией, теорией связи и т.д.

Хотя физика старается изучать явления природы в максимально объективированной (т.е. не связанной с человеком и его воздействием на окружающий мир) форме, ей не удалось полностью исключить “человеческий фактор”. Во-первых, при экспериментальном исследовании физических явлений невозможно обойтись без измерения ряда величин, и наличие этих величин в теоретических моделях рано или поздно требует специального рассмотрения того, как именно проводятся измерения – а без влияния человека не проводятся даже автоматические эксперименты. Во-вторых, физика не могла не заняться изучением технических (т.е. созданных человеком) устройств и тут-то и “попала в ловушку”: полностью объективистский подход к машинам все равно обнаружил в их поведении следы человеческой деятельности. Впервые это произошло в термодинамике – науке, изучающей процессы в тепловых машинах. Оказалось, что без введения специального понятия энтропии невозможно дать исчерпывающего описания их действия. Скачок в понимании природы этой величины произошел, когда Л.Больцман дал ей статистическую интерпретацию (1877); уже сам Больцман обронил фразу о том, что *энтропия характеризует недостающую информацию*, но тогда этой фразы никто не понял.

После построения К.Шэнноном теории информации (1948), когда оказалось, что формула Шэннона для информационной энтропии и формула Больцмана для термодинамической энтропии оказались тождественными, разгорелись споры о том, является ли это совпадение чисто формальным или оно выражает глубокую связь термодинамики и теории информации. Дискуссии (в особенности благодаря работам Бриллюэна [1, 2]) привели к современному пониманию этой неразрывной связи.

Совсем с другой стороны к этой проблеме пришла философская теория познания. Изначальный смысл слова “информация” как “знания, сведения, сообщения, уведомления, известия, ведомости”, т.е. нечто присущее только человеческому сознанию и общению, начал расширяться и обобщаться. Признав, что наше знание есть отражение реального мира, материалистическая теория познания установила, что отражение является всеобщим свойством материи. Сознание человека является высшей, специфической формой отражения, но существуют и другие формы – психическая (присущая не только человеку, но и животным), раздражимость (охватывающая, кроме того, растения и простейшие организмы) и, наконец, самая элементарная форма – запечатление взаимодействия (присущая и неорганической природе, и элементарным частицам, т.е. всей материи вообще). И теперь, как только состояния одного объекта находятся в соответствии с состояниями другого объекта (будь то соответствие между положением стрелки вольтметра и напряжением на его клеммах или соответствие между нашим ощущением и реальностью), мы говорим, что один объект *отражает другой, содержит информацию о другом*.

Так вновь сомкнулись результаты философского и естественнонаучного исследования природы. Иначе и быть не могло – ведь предмет изучения един, хотя это изучение ведется с разных позиций и разными методами. *В настоящее время информация рассматривается как фундаментальное свойство материи.*

Таким образом, ясно, что роль информации в самом существовании систем, искусственных и естественных, огромна. Понятие информации, обладая всеобщностью, приобрело смысл философской категории. Для кибернетики, да и всей системологии, понятие информации столь же фундаментально, как понятие энергии для физики. Недаром признается не

менее общим, чем определение А.И.Берга (см. гл.1), определение кибернетики, предложенное А.Н.Колмогоровым: *кибернетика – это наука, которая занимается изучением систем любой природы, способных воспринимать, хранить и перерабатывать информацию и использовать ее для управления и регулирования.*

Подведем итог

Summary

Исследуя информацию, кибернетика не открыла нового свойства материи: оно известно в философии под названием свойства отражения. Новое свойство состоит в том, что информацию можно исследовать количественно.

In studying information, cybernetics has not discovered a new property of matter: it is known in philosophy as the reflection property. What is new is its quantitative study of information.

§ 5.2. СИГНАЛЫ В СИСТЕМАХ

Как мы уже отмечали, для того чтобы два объекта содержали информацию друг о друге, необходимо, чтобы между их состояниями существовало соответствие: только при этом условии по состоянию одного объекта можно судить о состоянии другого. Такое соответствие может установиться только в результате физического взаимодействия между этими объектами. Редко бывает, чтобы два объекта взаимодействовали “непосредственно”: даже между печатью и документом нужен слой мастики; сквозь кинопленку до экрана проносится пучок света; от речевого аппарата оратора до ушей слушателей звук переносится колебаниями воздуха. Другими словами, соответствие между состояниями двух объектов может устанавливаться и с помощью взаимодействия с промежуточными объектами, часто даже целой совокупностью промежуточных объектов.

ПОНЯТИЕ СИГНАЛА

Итак, *сигнал есть материальный носитель информации, средство перенесения информации в пространстве и времени.*

Утверждая что объекты выступают в качестве сигналов, мы должны сделать уточнение. Один и тот же объект может “выступать в качестве” разных сигналов: колебания воздуха могут нести звуки музыки, речь лектора, пение птиц или шум самолета; с магнитной ленты можно стереть одну запись и сделать другую и т.д. Следовательно, *в качестве сигналов используются не сами по себе объекты, а их состояния.*

Далее, не всякое состояние имеет сигнальные свойства. Точнее говоря, данный объект взаимодействует не только с тем объектом, информацию о котором мы хотели бы получить, но и с другими, не интересующими нас объектами. В результате соответствие состояний ослабевает, разрушается. Условия, обеспечивающие установление и способствующие сохранению сигнального соответствия состояний, называются **кодом***. Посторонние воздействия, нарушающие это соответствие, называются **помехами** или **шумами**. Нарушение соответствия может происходить не только вследствие помех, но и из-за рассогласования кодов взаимодействующих объектов. В искусственных системах, где такое согласование организуется специально, это явно видно на примере криптографии, основанной на засекречивании кодов. В природных системах согласование кодов происходит в самой структуре систем через естественный отбор различных вариантов.

ТИПЫ СИГНАЛОВ

Поскольку сигналы служат для переноса информации в пространстве и времени, для образования сигналов могут использоваться только объекты, состояния которых достаточно

* В искусственных системах кодом называют комплекс правил образования сигнала.

При более подробном рассмотрении этого процесса в технических системах кодом называют условные, варьируемые правила, а диктуемые техникой условия называют модуляцией. Мы же будем пока использовать самое общее употребление слова “код”.

устойчивы по отношению к течению времени или к изменению положения в пространстве. С этой точки зрения сигналы делятся на два типа.

К первому типу относятся сигналы, являющиеся стабильными состояниями физических объектов (например, книга, фотография, магнитофонная запись, состояние памяти ЭВМ, положение триангуляционной вышки и т.д.). Такие сигналы называются статическими.

Ко второму типу относятся сигналы, в качестве которых используются динамические состояния силовых полей. Такие поля характеризуются тем, что изменение их состояния не может быть локализовано в (неизолированной) части поля и приводит к распространению возмущения. Конфигурация этого возмущения во время распространения обладает определен-

INFORMATION

информация

BEARER HOLDER

носитель

INTERFERENCE

помехи

SIGNAL

сигнал

**Информация есть
свойство материи,
состоящее в том, что
в результате взаимодействия
объектов
между их состояниями уста-
навливается определенное
соответствие. Чем сильнее
выражено это соответствие,
тем полнее состояние одного
объекта отражает состояние
другого объекта, тем больше
информации один объект со-
держит
о другом. Сигнал есть мате-
риальный носитель инфор-**

ной устойчивостью, что обеспечивает сохранение сигнальных свойств. Примерами таких сигналов могут служить звуки (изменение состояния поля сил упругости в газе, жидкости или твердом теле), световые и радиосигналы (изменения состояния электромагнитного поля). Сигналы указанного типа называются *динамическими* *.

Понятно, что динамические сигналы используются преимущественно для передачи, а статические – для хранения информации, но можно найти и противоположные примеры (динамические запоминающие устройства, письма, газеты).

Сигналы играют в системах особую, очень важную роль. Если энергетические и вещественные потоки, об разно говоря, питают систему, то потоки информации, переносимые сигналами, организуют все ее функционирование, управляют ею. Н. Винер, например, подчеркивал, что общество простирается до тех пределов, до каких распространяется информация. Пожалуй, это следует отнести к любой системе.

* Обратим еще раз внимание на относи-
тельность вся-
кой классификации. К какому из введен-
ных классов
вы отнесете дымовые сигналы, или запах,
или голо-
графическое изображение?

Подведем итог

Первое и, быть может, главное отличие подхода к изучению любого объекта как системы, а не как просто объекта, состоит в том, что мы ограничиваемся не только рассмотрением и описанием вещественной и энергетической его сторон, но

The first and perhaps the main distinction of the approach to studying an object as a system, and not just as a mere object, is that we do not restrict ourselves to studying and describing it only in terms of substance and energy, but also (and preferably) from the

Summary

RANDOM
случайный
NOISE
шум
OSCILLATION
колебание
ENVELOPE
огибающая

Случайный процесс хорошо отображает главное свойство реальных сигналов – их неизвестность до момента приема.
Однако некоторые модели облашают и “лишними”, не присущими реальным сигналам свойствами.
Например, описание ре-

и (прежде всего) проводим standpoint of information: its исследование его информационных аспектов: целей, сигналов, information purposes, signals, information flows, control, organization, etc. лов, информационных потоков, управления, организации и т.д.

§ 5.3. СЛУЧАЙНЫЙ ПРОЦЕСС – МАТЕМАТИЧЕСКАЯ МОДЕЛЬ СИГНАЛОВ

Казалось бы, после того как мы установили, что сигналами служат состояния физических объектов, никаких проблем с их математическим описанием не должно быть: ведь физика имеет богатый опыт построения математических моделей физических процессов и объектов. Например, можно зафиксировать звуковые колебания, соответствующие конкретному сигналу, в виде зависимости давления x от времени t и изобразить этот сигнал функцией $x(t)$; такой же функцией можно изобразить и статический сигнал, например запись этого звука на магнитной ленте или на грампластинке, поставив параметру t в соответствие протяженность (длину) записи.

НЕПРЕДСКАЗУЕМОСТЬ – ОСНОВНОЕ СВОЙСТВО СИГНАЛОВ

Однако имеется существенное различие между просто состоянием $x(t)$ объекта и сигналом $x(t)$. Оно состоит в том, что единственная функция $x(t)$ не исчерпывает всех важных свойств сигналов. Ведь понятие функции предполагает, что нам известно значение x (либо правило его вычисления) для каждого t . Если же это известно получателю сигнала, то отпадает необходимость в передаче: функция $x(t)$ может быть и без этого воспроизведена на приемном конце.

Следовательно, единственная однозначная функция вещественного аргумента не может служить моделью сигнала. Такая функция приобретает сигнальные свойства только тогда, когда она является *одной из возможных функций*. Другими словами, моделью сигнала может быть набор (или, как еще говорят, *ансамбль*) функций параметра t , причем до передачи неизвестно, какая из них будет отправлена; это становится известным получателю только после передачи. Каждая такая конкретная функция называется *реализацией*. Если теперь еще ввести *вероятностную меру* на множество реализаций, то мы получим математическую модель, называемую **случайным процессом***.

Прежде чем перейти к некоторым подробностям этой модели, отметим, что у нее все-таки имеются качества, которых нет у реальных сигналов (еще раз обратим внимание на нетождественность модели и оригинала). Дело в том, что реальные системы всегда оперируют только с конечным объемом данных, а понятие аналитической функции предполагает ее точное значение для каждого значения аргумента, т.е. такое, которое может быть представлено только с помощью бесконечного ряда цифр. Иначе говоря, в “более правильной” модели сигнала любая реализация не должна определяться с бесконечной точностью. Но пока не существует подходящего математического аппарата, столь же удобного, как математический анализ, и мы вынуждены пользоваться традиционным понятием функции. Это сопряжено с тем, что в выводах теории могут появиться (и на самом деле появляются [6]) парадоксы, природа которых связана не с реальными сигналами, а с их моделью. Во всяком случае, там, где конечная точность реальных сигналов существенна для самой постановки задачи, ее вводят в модель либо как добавочный “шум”, либо как “квантование” непрерывного сигнала (наподобие шкалы с делениями у измерительных приборов).

* Изложение будет вестись в предположении, что читателю знакомы элементы теории вероятностей.

Вернемся к рассмотрению случайных процессов как моделей сигналов. Имеется несколько различных подходов к тому, как вводить вероятностную меру на множестве реализаций. Для инженерных приложений оказывается удобным определение случайного процесса как такой функции времени $x(t)$, значение которой в каждый данный момент является случайной величиной. Случайная величина полностью характеризуется распределением вероятностей, например плотностью $P_1(x_1 | t_1)$; однако, чтобы охарактеризовать случайный процесс, нужно описать, связаны ли (и если да, то как) значения реализации, разделенные некоторыми интервалами времени. Так как связь только двух таких значений, описываемая распределением второго порядка $P_2(x_1, x_2 | t_1, t_2)$, может неполно характеризовать процесс в целом, вводят распределения третьего, четвертого, ..., n -го порядков: $P_n(x_1, \dots, x_n | t_1, \dots, t_n)$. В конкретных задачах обычно ясно, до какого порядка следует доходить в описании процесса.

КЛАССЫ СЛУЧАЙНЫХ ПРОЦЕССОВ

Необходимость моделирования самых разнообразных сигналов приводит к построению частных моделей случайных процессов, т.е. наложению дополнительных ограничений на параметры распределений и на сами распределения. Перечислим наиболее важные классы случайных процессов.

Непрерывные и дискретные по времени процессы. Случайный процесс с непрерывным временем характеризуется тем, что его реализации определяются для всех моментов из некоторого (конечного или бесконечного) интервала T параметра t . Дискретный по времени процесс* задается на дискретном ряде точек временной оси (обычно равнотстоящих).

Непрерывные и дискретные по информативному параметру процессы. Эти процессы различаются в зависимости от того, из какого (непрерывного или дискретного) множества принимает значение реализация x случайной величины X .

Стационарные и нестационарные процессы. Так называются процессы в зависимости от постоянства или изменчивости их статистических характеристик. Случайный процесс называется *стационарным в узком смысле*, если для любого n конечномерные распределения вероятностей не изменяются со временем, т.е. при любом τ выполняется условие

$$P_n(x_1, \dots, x_n | t_1 + \tau, t_2 + \tau, \dots, t_n + \tau) = P_n(x_1, \dots, x_n | t_1, t_2, \dots, t_n).$$

Если же условие независимости от времени выполняется только для первых двух моментов (среднего и функции автокорреляции), то процесс называется *стационарным в широком смысле* (или в смысле Хинчина).

Эргодические и неэргодические процессы. На практике при описании случайных величин вместо рассмотрения их распределений часто ограничиваются только их числовыми характеристиками, обычно моментами. В тех случаях, когда распределение неизвестно, моменты (и другие числовые характеристики) можно оценить статистически.

Перенос такой практики на произвольные случайные процессы требует не только учета зависимости отстоящих друг от друга (“разнесенных”) во времени значений, но и наложения дополнительных требований. Требование совпадения величин, получающихся при усреднении по ансамблю (т.е. при фиксированном времени) и при усреднении по времени (точнее, по одной реализации), и называется условием эргодичности. Это требование можно толковать и как совпадение результатов *усреднения по любой реализации**. Как и для стационарности, можно различать эргодичность в узком и широком смысле.

Можно продолжать классификацию случайных процессов и дальше, но мы будем делать это при рассмотрении конкретных вопросов.

Подведем итог

Summary

Основной результат данного параграфа состоит в том, что conclusion that a random process

* Дискретные по времени случайные процессы иногда называются *случайными последовательностями*, так как само слово “процесс” имеет оттенок непрерывности.

*Точнее, эргодический ансамбль может содержать реализации, не удовлетворяющие этому условию, но суммарная вероятность таких реализаций должна быть сколь угодно близкой к нулю.

случайный процесс может can be used as a mathematical служить математической моделью model of a signal. But it is necessary to pay special attention to только следить за тем, чтобы the correctness of mapping конкретные особенности изучаемых сигналов были корректно отображены в свойствах случайного процесса.

§ 5.4. МАТЕМАТИЧЕСКИЕ МОДЕЛИ РЕАЛИЗАЦИЙ СЛУЧАЙНЫХ ПРОЦЕССОВ

Для рассмотрения конкретных свойств систем бывает необходимо учесть особенности сигналов, циркулирующих по каналам связи этих систем. Такие особенности можно описать по-разному: просто перечислить возможные реализации (если число их конечно) либо задать в той или иной форме общие свойства реализаций, входящих в ансамбль. О дискретных процессах мы будем говорить отдельно (см. § 5.8), а сейчас рассмотрим математические модели реализаций непрерывных сигналов.

Приведем примеры, с которыми часто имеют дело в теории сигналов [7].

МОДЕЛИРОВАНИЕ КОНКРЕТНЫХ РЕАЛИЗАЦИЙ

Гармонические сигналы. Обозначим через S_c множество всех синусоидальных сигналов:

$$S_c = \{x: x(t) = S \cos(\omega t + \varphi)\}; \quad -\infty < t < \infty; S, \omega \in \mathbb{R}^+, \varphi \in [0, 2\pi]. \quad (1)$$

Здесь \mathbb{R}^+ – множество всех положительных действительных чисел; S называется *амплитудой*, ω – *круговой частотой* ($\omega = 2\pi f$, f – *частота*), φ – *фазой* гармонического колебания.

Модулированные сигналы. В технических системах полезная информация может передаваться каким-нибудь одним параметром “гармонического” колебания. Конечно, колебание при изменении этого параметра во времени перестает быть гармоническим. Процесс изменения параметра синусоиды называется *модуляцией*, а выделение этого изменения в чистом виде, как бы “снятия” модуляции, называется *демодуляцией*. Само колебание называется *несущим*. Различают *амплитудную*, *частотную* и *фазовую модуляцию* в зависимости от того, на какой из параметров несущего колебания “накладывают” полезную информацию (параметры с индексом 0 считаются постоянными):

$$S_{\text{ам}} = \{x: x(t) = S(t) \cos(\omega_0 t + \varphi_0)\}, \quad (2)$$

$$S_{\text{чм}} = \{x: x(t) = S_0 \cos(\omega(t) \cdot t + \varphi_0)\}, \quad (3)$$

$$S_{\text{фм}} = \{x: x(t) = S_0 \cos(\omega_0 t + \varphi(t))\}. \quad (4)$$

Необходимо отметить, что “физический смысл” модуляции сохраняется лишь в том случае, когда модулирующий сигнал является “медленно меняющимся” по сравнению с немодулированным несущим колебанием: только при этом условии имеется возможность нестрогого, “по-инженерному”, но все-таки не без смысла говорить о “синусоиде с переменной амплитудой (фазой, частотой)”. Однако удобнее все-таки говорить не о “переменной амплитуде”, а об “огибающей”, как это принято в радиотехнике.

Периодические сигналы. Сигналы называются *периодическими*, а временному интервалу τ – *периодом*, если

$$S_\tau = \{x: x(t) = x(t+\tau); -\infty < t < \infty\}. \quad (5)$$

Сигналы с ограниченной энергией.

О сигналах из множества

$$S_3 = \{x: \int_{-\infty}^{\infty} x^2(t) dt \leq K < \infty\} \quad (6)$$

говорят, что их *энергия ограничена величиной K* . Происхождение этого названия связано с тем, что если $x(t)$ есть напряжение, то интеграл в формуле (6) представляет собой энергию,

выделенную сигналом $x(t)$ на единичном сопротивлении. Конечно, если $x(t)$ есть, например, глубина бороздки на грампластинке, то интеграл связан с энергией лишь косвенно.

Сигналы ограниченной длительности. Интервал T называется *длительностью сигнала* $x(t)$, если

$$S_T = \{x: x(t) \equiv 0, t \notin T\}; \quad (7)$$

конечно, предполагается, что внутри этого интервала $x(t)$ не везде обращается в нуль.

Особую роль среди сигналов с ограниченной длительностью играют *импульсные сигналы*, их отличие состоит в “кратковременности” T , которую трудно формализовать, но которая проявляется в практике: звуки типа “щелчок”, “взрыв”, “хлопок”; световые “вспышки”; тактильные сигналы “укол”, “щипок”, “удар” и т.п. В таких случаях $x(t)$ обычно называется “формой импульса”. На практике очень распространены периодические последовательности импульсов (радиолокация, электрокардиография, ультразвуковая гидролокация и т.п.); не менее важны непериодические импульсные последовательности (телеграфия, цифровая telemetry); в реальности все импульсные последовательности, в свою очередь, имеют ограниченную длительность.

Сигналы с ограниченной полосой частот. Фурье-преобразование $X(f)$ временной реализации $x(t)$ сигнала называется ее *спектром*:

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{i2\pi ft} dt. \quad (8)$$

Физический смысл спектра состоит в том, что колебание $x(t)$ представляется в виде суммы (в общем случае в виде интеграла) составляющих его гармонических колебаний с определенными амплитудами $|x(f)|$, частотами f и соответствующими фазами. Между $x(t)$ и $X(f)$ имеется взаимно однозначное соответствие, так как

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{-i2\pi ft} dt. \quad (9)$$

Условием существования и обратимости Фурье-преобразования является ограниченность энергии сигнала (интегрируемость в квадрате функций $x(t)$ и $X(f)$) и его непрерывность.

Если функция $X(f)$ на оси f имеет ограниченную “длительность” F (в смысле соотношения (7), но в частотной области), то говорят, что сигнал $x(t)$ имеет ограниченную полосу частот шириной F :

$$S_B = \{x: X(f) = \int_{-\infty}^{\infty} x(t) e^{i2\pi ft} dt = 0 \text{ для всех } f > |F|\}. \quad (10)$$

НЕКОТОРЫЕ МОДЕЛИ АНСАМБЛЯ РЕАЛИЗАЦИЙ

Нормальный шум. Удобной моделью помех и некоторых полезных сигналов является стационарный нормальный случайный процесс. Такое название вызвано тем, что случайные мгновенные значения величины $x(t)$ предполагаются подчиненными нормальному закону (с нулевым средним), т.е. плотность распределения первого порядка выражается формулой

$$p_1(x_1 | t_1) = \frac{1}{\sqrt{2\pi N}} \exp\left\{-\frac{1}{2N} x_1^2\right\}; \quad (11)$$

для плотности распределения второго порядка справедлива формула

$$\begin{aligned} p_2(x_1, x_2 | t_1, t_2) &= p_2(x_1, x_2 | \tau = t_2 - t_1) = \\ &= \frac{1}{2\pi N \sqrt{1-\rho^2(\tau)}} \exp\left\{-\frac{1}{2(1-\rho^2(\tau))N} [x_1^2 - 2\rho(\tau)x_1 x_2 + x_2^2]\right\}, \end{aligned} \quad (12)$$

а для n -мерной плотности – формула

$$p_1(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2} |C|^{1/2}} \exp\left\{-\frac{1}{2|C|} \sum_{i,j=1}^n |c_{ij}| x_i x_j\right\}. \quad (13)$$

Использованные здесь обозначения таковы:

$N = E(X^2)$ – мощность шума (дисперсия мгновенных значений);

$\rho(\tau) = N^{-1}E(X_t X_{t+\tau})$ – нормированная функция автокорреляции;

$c_{ij} = E(X_i X_j)$, $c_{ii} = N$ – коэффициенты ковариации; $C = \|c_{ij}\|$,

$|C|$ – определитель матрицы C ;

$|c_{ij}|$ – алгебраическое дополнение элемента c_{ij} , т.е. умноженный на $(-1)^{i+j}$ определитель матрицы $(n-1)$ -го порядка, полученной из C вычеркиванием i -й строки и j -го столбца.

Замечательной особенностью нормального процесса является то, что его распределение любого порядка полностью определяется только первыми двумя моментными функциями:

средним (которое в формулах (11) – (13) принято равным нулю) и ковариационной функцией $\sqrt{N\rho(\tau)}$, так как $c_{ij} = N\rho(t_i - t_j)$.

Сигнал как колебание со случайными огибающей и фазой. Понятия амплитуды и фазы, введенные первоначально для гармонических сигналов, с помощью модуляции были обобщены на сигналы, которые уже не являются гармоническими. Легко обобщить их на произвольные сигналы: пока чисто формально можно задать такие функции $R(t)$ и $\phi(t)$, чтобы для заданной функции $x(t)$ было выполнено равенство

$$x(t) = R(t)\cos(\omega_0 t + \phi(t)), \quad (14)$$

и, сравнив его с соотношениями (1) – (4), можно трактовать $R(t)$ и $\phi(t)$ как “огибающую” и “фазу” колебания с частотой ω_0 .

Возникает вопрос: соответствуют ли эти функции чему-нибудь реальному? Оказывается, свобода выбора в задании функций R и ϕ при определенных условиях весьма ограничена. Эти условия мы уже упоминали как “медленность” $R(t)$ по сравнению с $\cos\omega_0 t$ и $\phi(t)$ по сравнению с $\omega_0 t$, а в целом этот комплекс условий в силу причин (которые мы не будем обсуждать) получил название *узкополосности* сигнала $x(t)$.

Очень наглядным является векторный вариант модели (14): R и ϕ можно рассматривать как полярные координаты некоторого вектора. Тогда всякое гармоническое колебание $x(t) = S\cos(\omega_0 t + \phi)$, имеющее частоту ω_0 , изображается как постоянный вектор с амплитудой S и углом ϕ к направлению, принятому за ось Ox .

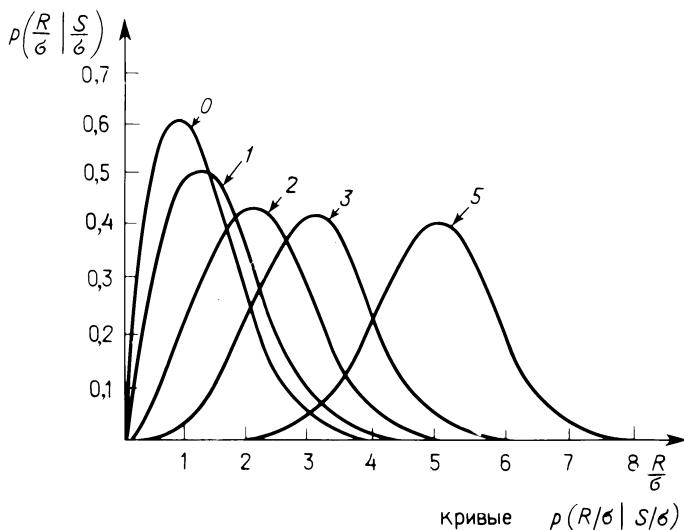
Если гармонический сигнал имеет частоту, отличную от ω_0 (например, ω_1), то в рассматриваемой системе координат соответствующий ему вектор будет вращаться с частотой

$$\omega_1 - \dots - \omega_0, \text{ поскольку } \cos\omega_1 t = \cos(\omega_0 t + \phi(t)),$$

$\phi(t) = (\omega_1 - \omega_0)t$. Негармонический сигнал изобразится как вектор, совершающий вращения или колебания (изменяется угол) и изменяющий свою длину R .

При всей простоте такая модель позволяет легко описывать различные преобразования сигналов. Например, сумма двух колебаний с частотой ω_0 изобразится как сумма двух соответствующих векторов (при этом автоматически учитываются фазовые соотношения).

Пусть, например, принимаемый сигнал является суммой полезного сигнала $s(t)$ и шума $n(t)$:



(15)

$$x(t) = s(t) + n(t).$$

Рассмотрим случай, когда полезный сигнал есть гармоническое колебание

$$s(t) = S \cdot \cos \omega_0 t,$$

а шум является нормальным с дисперсией σ^2 и нулевым средним. Будем считать, что ω_0 значительно превышает ширину полосы частот, занимаемую сигналом $x(t)$, так что такая узкополосность обеспечивает физический смысл понятиям его огибающей и фазы (см.

формулу (14)). Направив ось $0x$ вдоль вектора сигнала $s(t)$, находим его координаты $x = S$ и $y = 0$. Вектор шума случаен, обе его компоненты нормальны с нулевым средним и дисперсией σ^2 . Сложение векторов сигнала и шума приводит к вектору с нормальными компонентами, у одной из которых среднее равно S , а у другой – нулю:

$$p(x, y) dx dy = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2} \left[(x - S)^2 + y^2 \right]\right\} dx dy. \quad (17)$$

Поскольку существуют радиотехнические устройства, чувствительные только к огибающей или только к фазе принимаемого сигнала, имеет смысл рассмотреть их статистические свойства. Переходим от декартовых координат (x, y) к полярным (R, φ) . Используя очевидные соотношения $x = R \cos \varphi$, $y = R \sin \varphi$, $dx dy = R dR d\varphi$, из равенства (17) сразу получаем распределение

$$p(R, \varphi) dR d\varphi = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2} \left(R^2 - 2RS \cos \varphi + S^2 \right)\right\} dR d\varphi. \quad (18)$$

Легко установить, что если $S = 0$ (т.е. имеется только чистый шум), то

$$p(R, \varphi) dR d\varphi = \frac{R}{\sigma^2} \exp\left(-\frac{R^2}{2\sigma^2}\right) dR \cdot \frac{1}{2\pi} d\varphi, \quad (19)$$

т.е. огибающая распределена по закону Рэлея, фаза – равномерно в интервале $[0, 2\pi]$, R и φ независимы (так как их совместное распределение факторизовалось).

Распределения вероятностей отдельно для огибающей R и фазы φ получаются при интегрировании совместного распределения (18) по переменной, которую нужно исключить. Например, распределение огибающей имеет вид

$$\begin{aligned} p(R) dR &= dR \int_0^{2\pi} p(R, \varphi) d\varphi = \\ &= \frac{R dR}{\sigma^2} \exp\left(-\frac{R^2 + S^2}{2\sigma^2}\right) \frac{1}{2\pi} \int_0^{2\pi} \exp\left(\frac{RS}{\sigma^2} \cos \varphi\right) d\varphi = \\ &= \frac{R}{\sigma^2} \exp\left(-\frac{R^2 + S^2}{2\sigma^2}\right) I_0\left(\frac{RS}{\sigma^2}\right) dR, \end{aligned} \quad (20)$$

5.1

График плотностей распределения вероятностей огибающей смеси синусоидального сигнала с гауссовым шумом

где через I_0 обозначена модифицированная функция Бесселя первого рода нулевого порядка (так называется трансцендентная функция, выражаемая полученным интегралом). Это распределение называется **распределением Рэля–Райса** или **обобщенным распределением Рэлея**, поскольку при $S = 0$ оно обращается в обычное распределение Рэлея. Семейство кривых $p(R)$ для различных S и σ изображено на рис. 5.1.

Распределение фазы выражается формулой

$$p(\varphi) d\varphi = d\varphi \int_0^\infty p(R, \varphi) dR. \quad (21)$$

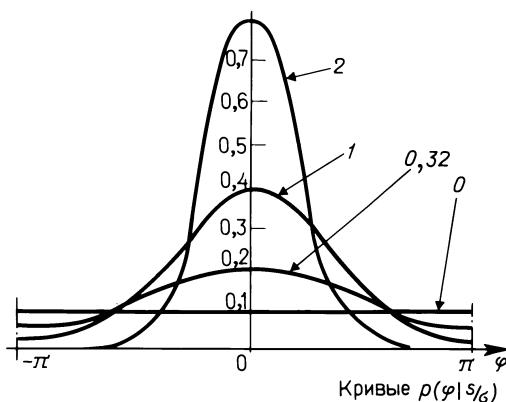
Не приводя аналитических выражений для $p(\varphi)$ из-за их громоздкости (при необходимости см. [5]), изобразим графики этого распределения при разных S/σ на рис. 5.2.

Подведем итог

В данном параграфе приведено несколько наиболее употребительных моделей сигналов. signals have been presented in this section. It is necessary to Следует помнить, что в ряде случаев эти модели достаточно хорошо отображают реальные сигналы, но абсолютно точными они не бывают никогда.

Summary

§ 5.5. О НЕКОТОРЫХ СВОЙСТВАХ НЕПРЕРЫВНЫХ СИГНАЛОВ



5.2

График плотностей распределения вероятностей для фазы колебания, являющегося аддитивной смесью смеси синусоидального сигнала с шумом

Хотя мы излагаем лишь элементы теории сигналов, преследуя только ознакомительные цели, представляется интересным рассмотреть два основных ее аспекта, относящихся к свойствам непрерывных сигналов.

ЧАСТОТНО-ВРЕМЕННОЕ ПРЕДСТАВЛЕНИЕ СИГНАЛОВ

Первый – это частотно-временная неопределенность сигналов.

Известно, что некоторая функция $x(t)$ и ее спектр $X(f)$ однозначно выражаются друг через друга (см. формулы (8) и (9) § 5.4). Следовательно, сигнал можно рассматривать в любом из этих эквивалентных представлений – временном или частотном. При этом масштабные параметры этих представлений связаны обратно пропорциональной зависимостью. Пусть $x(t)$ имеет спектр $X(f)$. Изменим масштаб по оси времени в a раз (например, воспроизведем запись $x(t)$ с другой скоростью) и найдем спектр функции $x(at)$:

$$X_a(f) = \int_{-\infty}^{\infty} x(at)e^{j2\pi ft} dt = \frac{1}{a} X\left(\frac{f}{a}\right). \quad (1)$$

Как видим, масштаб по частотной оси изменился в $1/a$ раз. Более того, из свойств преобразования Фурье следует, что *сигналы с ограниченной длительностью имеют спектры неограниченной ширины, а сигналы с ограниченной полосой частот делятся бесконечно долго*.

Этот математический результат находится в противоречии с практикой: в реальности все сигналы конечны по длительности, а все чувствительные к сигналам устройства не могут воспринимать и воспроизводить абсолютно все частоты. Например, диапазон частот, к которым чувствителен слух человека, простирается от нескольких герц до 20 – 30 кГц, а все различные звуки человеческой речи делятся доли секунды.

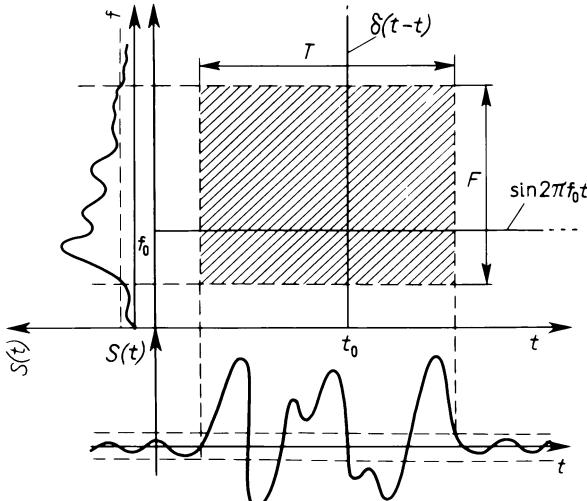
Тот факт, что аналитическая функция времени не может быть одновременно ограниченной и по длительности, и по ширине спектра, является, как видим, не свойством реальных сигналов, а свойством данной модели сигналов. В § 5.3 мы уже отмечали, что если не отказываться от достоинств аппарата аналитических функций, то выход состоит в том, чтобы как-то иначе ввести в рассмотрение конечную точность реализаций функций времени. Правда, пока эта конечная точность не будет свойством самих реализаций, ее искусственное введение в модель можно проводить на разных этапах, что придает результатам некоторую относительность.

Например, говорить об одновременной ограниченности сигналов и по времени, и по спектру оказывается возможным при использовании энергетического критерия точности: сигнал считается имеющим конечную длительность ΔT , если в этом интервале времени со средоточена основная часть всей энергии функции $x(t)$; в то же время и ширина спектра ΔF сигнала определяется как область частот, содержащая эту же часть всей энергии спектра $X(f)$:

$$\int_{\Delta T} x^2(t) dt = \int_{\Delta F} X^2(f) df = ? \int_{-\infty}^{\infty} x^2(t) dt = ? \int_{-\infty}^{\infty} X^2(f) df, \quad (2)$$

Здесь величина $?$ меньше единицы, но достаточно близка к ней, а величина $1 - ?$ характеризует косвенным образом точность, о которой шла речь.

Теперь можно говорить о том, какую “площадь” на плоскости “частота – время” занимает тот или иной сигнал. Если строго следовать теории Фурье-преобразований, то получим, что эта площадь для всех сигналов бесконечна, но для большинства из них энергетический критерий позволит ограничить ее естественным образом (рис. 5.3). Меняя форму сигнала $s(t)$, можно менять и занимаемую им площадь. Оказывается [4], что уменьшать эту площадь можно лишь до некоторого предела. Этот предел достигается на кривой, являющейся гармоническим колебанием, которое модулировано по амплитуде гауссовым импульсом; интересно, что спектр этой кривой имеет такую же форму:



5.3

Иллюстрация частотно-временной неопределенности

ДИСКРЕТНОЕ ПРЕДСТАВЛЕНИЕ СИГНАЛОВ

UNCERTAINTY
неопределенность
CONTINUOUS
непрерывный
LIMIT
предел
FREQUENCY
частота

Все реальные сигналы длятся конечное время, и все реальные системы не могут реагировать на абсолютно все частоты. Следовательно, как длительность, так и ширина спектра сигнала ограничены. Аналитическая модель сигнала позволяет говорить об этих характеристиках приближенно. В этом мы видим конкретный пример проявления расхождений между моделью и оригиналом.

Практически всякий сигнал можно представить состоящим из более простых компонент. В математике этой идеи соответствует представление функций их разложениями в ряды и интегралы. В теории сигналов специальное внимание уделяется

функциям $\{\psi_k(t)\}$ удовлетворяют условиям

$$s(t) = e^{-a(t-t_0)^2} e^{i2\pi f_0 t}$$

$$S(f) = e^{-\left(\frac{\pi}{a}\right)^2 (f-f_0)^2} e^{-i2\pi f_0 t}.$$

Существование предела, ниже которого нельзя сжать площадь сигнала, занимаемую им на плоскости "частота – время", и называется (по аналогии с принципом неопределенности в квантовой механике) **принципом частотно-временной неопределенности сигналов:**

$$F \cdot T \geq \text{const} > 0.$$

Вторым важным аспектом теории сигналов является **проблема дискретного представления непрерывных сигналов**. Вопрос формулируется так: существуют ли условия (и если да, то каковы они), при которых любой непрерывной функции $x(t)$ можно поставить во взаимно однозначное соответствие дискретное множество чисел $\{C_k(x)\}$, $k = \dots -2, -1, 0, 1, 2, \dots$? Положительный ответ на этот вопрос имел бы как теоретическое, так и практическое значение. Во-первых, рассмотрение случайных величин вместо реализаций непрерывных случайных процессов существенно упрощает решение многих задач, вся теория становится проще и может быть продвинута дальше. Во-вторых, соответствие $x(t) \leftrightarrow \{C_k(x)\}$ можно использовать и в технических устройствах, работающих с непрерывными сигналами (например, иногда технически проще хранить или передавать $\{C_k(x)\}$ вместо $x(t)$).

Ограничимся более конкретной формулировкой поставленной задачи и рассмотрим условия выполнения равенства

$$x(t) = \sum_k C_k(x) \psi_k(t). \quad (6)$$

Функции $\{\psi_k(t)\}$ называются *координатными функциями*, они не должны зависеть от $x(t)$, более того, они заранее известны. Ряд в правой части равенства называется *разложением* $x(t)$ по координатным функциям. Числовые коэффициенты $\{C_k(x)\}$ содержат всю информацию об $x(t)$, необходимую для восстановления этой функции по формуле (6); следовательно, $\{C_k(x)\}$ являются функционалами* от функции $x(t)$.

Наиболее известны разложения по системе ортогональных и нормированных функций. Это означает, что

* *Функционалом* называется отображение множества функций в множество чисел

$$\int j_i(t)j_k(t)dt = d_{ik} = \begin{cases} 1 & \text{при } i = k, \\ 0 & \text{при } i \neq k. \end{cases} \quad (7)$$

Умножим обе части равенства (6) на $\varphi_i(t)$ и проинтегрируем (опуская тонкости, будем считать, что все операции обоснованы):

$$\int \varphi_i(t)x(t)dt = \sum_k C_k(x) \int \varphi_i(t)\varphi_k(t)dt = C_i. \quad (8)$$

Такое представление называют *рядом Фурье*, а $C_k(x)$ – *коэффициентами Фурье*. Условия сходимости ряда Фурье к функции $x(t)$ подробно исследованы и, кратко говоря, сводятся к тому, чтобы были оправданы все необходимые математические операции, а коэффициенты Фурье убывали достаточно быстро (точнее, $\sum_k C_k^2 \log^2 k < \infty$). Это не очень жесткое ограничение, но все же оно связывает свойства системы координатных функций и самих функций $x(t)$. Например, если $\{\varphi_k(t)\}$ – гармонические функции кратных частот, то $x(t)$ должна быть периодической функцией с периодом T , равным периоду самой низкочастотной гармоники:

$$x(t) = \sum_{k=-\infty}^{\infty} C_k e^{\frac{i \pi k t}{T}}, \quad (9)$$

$$C_k = \frac{1}{2F} \int_{-F}^F x(t) e^{-i \frac{2\pi k t}{T}} dt. \quad (10)$$

Значительный интерес привлекли разложения реализаций случайного процесса с ограниченной полосой частот. Для таких сигналов В.А. Котельников доказал (1946) следующую теорему (**теорему отсчетов**):

любая функция со спектром, находящимся в интервале $[0, F]$, полностью определяется последовательностью ее значений в точках, отстоящих друг от друга на $1/(2F)$ единиц времени.

Пусть $x(t)$ имеет спектр $X(f)$; эти функции связаны соотношениями (8) и (9) § 5.4, причем $X(f)$ отлично от нуля только в интервале $|f| \leq F$.

В этом интервале применимо разложение (9):

$$X(f) = \sum_{k=-\infty}^{\infty} C_k(X) e^{i \frac{\pi k f}{F}} = \sum_{k=-\infty}^{\infty} C_k(X) e^{i \frac{2\pi k f}{2F}},$$

а коэффициенты Фурье этого разложения таковы:

$$C_k(X) = \frac{1}{2F} \int_{-F}^F X(f) e^{i \frac{2\pi k f}{2F}} df = \frac{1}{2F} x\left(\frac{k}{2F}\right).$$

Следовательно,

$$X(f) = \frac{1}{2F} \sum_{k=-\infty}^{\infty} x\left(\frac{k}{2F}\right) e^{i \frac{2\pi k f}{2F}}, \quad (11)$$

и это соотношение уже доказывает теорему отсчетов в силу однозначной связи $X(f)$ с $x(t)$.

Чтобы показать в явном виде, как восстанавливать $x(t)$ для значений t между точками отсчетов, воспользуемся формулой (9) § 5.4:

$$\begin{aligned} x(t) &= \int_{-F}^F X(f) e^{i 2\pi f t} df = \frac{1}{2F} \sum_{k=-\infty}^{\infty} x\left(\frac{k}{2F}\right) \int_{-F}^F e^{i 2\pi f \left(\frac{k}{2F} - t\right)} df = \\ &= \sum_{k=-\infty}^{\infty} x\left(\frac{k}{2F}\right) \frac{\sin(2\pi F t - k\pi)}{2\pi F t - k\pi}. \end{aligned} \quad (12)$$

Итак, мы получили разложение реализации, координатными функциями которого являются функции вида $(\sin u)/u$, сдвинутые друг относительно друга на интервалы времени $1/(2F)$, а коэффициентами – значения (“отсчеты”) самой реализации, взятые в моменты $k/(2F)$.

Иногда говорят, что эта теорема является теоретическим обоснованием возможности на практике восстанавливать $x(t)$ по отсчетам $x\left(\frac{k}{2F}\right)$. Однако дело в том, что координатные функции имеют неограниченную длительность и, следовательно, физически нереализуемы. Кроме того, ряд (12) имеет неограниченное число членов. Все это снова возвращает нас к проблеме точности фиксации сигналов, – проблеме, пока не получившей полного освещения*.

Подведем итог

Из многочисленных результатов теории сигналов мы выделяем два, как существенно поясняющие природу непрерывных сигналов. Первый состоит в том, что сигналы обнаруживают своеобразную “упругость” занимаемой ими площади на плоскости “частота-время”. Это явление называется частотно-временной неопределенностью сигналов. Второй результат заключается в том, что определенный класс непрерывных сигналов допускает взаимно однозначное соответствие между любой реализацией из этого класса и дискретным набором отсчетов данной реализации.

Summary

From the numerous results of signal theory we single out two that substantially clarify the nature of continuous signals. The first is that the simultaneous time-and-frequency representation of signals exposes a sort of “elasticity” of an area occupied by a signal on a time-frequency plane. This phenomenon is called the time-and-frequency uncertainty of signals. The second result is that a certain class of continuous signals allows a one-to-one correspondence between any function from this class and a set of the values of this function at discrete moments in time.

§ 5.6. ЭНТРОПИЯ

Установив, что случайные процессы являются адекватной моделью сигналов (см. § 5.3), мы получаем возможность воспользоваться результатами и мощным аппаратом теории случайных процессов. Кроме того, обнаружив, что некоторые типы непрерывных сигналов допускают дискретное представление (см. § 5.5), мы упрощаем задачу, сводя все к рассмотрению случайных величин.

Это не означает, что теория вероятностей и теория случайных процессов дают готовые ответы на все вопросы о сигналах: подход с новых позиций выдвигает такие вопросы, которые раньше просто не возникали. Так и родилась *теория информации* [9], специально рассматривающая сигнальную специфику случайных процессов. При этом были построены принципиально новые понятия (которые мы рассмотрим в данном и следующем параграфах) и получены новые, неожиданные результаты, имеющие характер научных открытий (наиболее важные из них мы обсудим в § 5.8).

ПОНЯТИЕ НЕОПРЕДЕЛЕННОСТИ

Первым специфическим понятием теории информации является понятие *неопределенности* случайного объекта, для которой удалось ввести количественную меру, названную энтропией. Начнем с простейшего варианта – со случайного события. Пусть, например, неко-

* Некоторые подробности см. в [5].

торое событие может произойти с вероятностью 0,99 и не произойти с вероятностью 0,01, а другое событие имеет вероятности соответственно 0,5 и 0,5. Очевидно, что в первом случае результатом опыта “почти наверняка” является наступление события, во втором же случае неопределенность исхода так велика, что от прогноза разумнее воздержаться.

Для характеристики размытости распределений широко используется второй центральный момент (дисперсия) или доверительный интервал. Однако эти величины имеют смысл лишь для случайных числовых величин и не могут применяться к случайным объектам, состояния которых различаются качественно, хотя и в этом случае можно говорить о большей или меньшей неопределенности исхода опыта. Следовательно, мера неопределенности, связанная с распределением, должна быть некоторой его числовой характеристикой, функционалом от распределения, никак не связанным с тем, в какой шкале измеряются реализации случайного объекта.

ЭНТРОПИЯ И ЕЕ СВОЙСТВА

Примем (пока без обоснования) в качестве меры неопределенности случайного объекта A с конечным множеством возможных состояний

A_1, \dots, A_n с соответствующими вероятностями p_1, \dots, p_n величину

$$H(p_i) = -\sum_{k=1}^n p_i \log p_i , \quad (1)$$

которую и называют **энтропией** случайного объекта A (или распределения $\{p_i\}$). Убедимся, что этот функционал обладает свойствами, которые вполне естественны для меры неопределенности.

1⁰. $H(p_1, \dots, p_n) = 0$ в том и только в том случае, когда какое-нибудь одно из $\{p_i\}$ равно единице (а остальные – нули). Это соответствует случаю, когда исход опыта может быть предсказан с полной достоверностью, т.е. когда отсутствует всякая неопределенность. Во всех других случаях энтропия положительна. Это свойство проверяется непосредственно.

2⁰. $H(p_1, \dots, p_n)$ достигает наибольшего значения при $p_1 = p_2 = \dots = p_n = 1/n$, т.е. в случае максимальной неопределенности.

Действительно, вариация H по p_i при условии $\sum p_i = 1$ дает $p_i = \text{const} = \frac{1}{n}$.

3⁰. Если A и B – независимые случайные объекты, то

$$H(A \cap B) = H(\{?_{ik} = p_i q_k\}) = H(\{p_i\}) + H(\{q_k\}) = H(A) + H(B).$$

Это свойство проверяется непосредственно.

4⁰. Если A и B – зависимые случайные объекты, то

$$H(A \cap B) = H(A) + H(B | A) = H(B) + H(A | B), \quad (2)$$

где условная энтропия $H(B | A)$ определяется как математическое ожидание энтропии условного распределения.

Это свойство проверяется непосредственно.

5⁰. Имеет место неравенство $H(A) \geq H(A | B)$, что согласуется с интуитивным представлением о том, что знание состояния объекта B может только уменьшить неопределенность объекта A , а если они независимы, то оставит ее неизменной.

Это свойство доказывается с помощью тождественного неравенства

$$\sum_k \lambda_k f(\sum_k \lambda_k x_k) \leq f(\sum_k \lambda_k x_k) ,$$

справедливого для любой выпуклой функции $f(x)$, если в этом неравенстве положить $f(x) = x \log x$, $\lambda_k = p_k$, $x_k = q_k | l$.

Как видим, свойства функционала H позволяют использовать его в качестве меры неопределенности. Интересно отметить, что если пойти в обратном направлении, т.е. задать желаемые свойства меры неопределенности и искать обладающий указанными свойствами функционал, то уже только условия 2⁰ и 4⁰ позволяют найти этот функционал, и притом единственный образом (с точностью до постоянного множителя).

ДИФФЕРЕНЦИАЛЬНАЯ ЭНТРОПИЯ

Обобщение столь полезной меры неопределенности на непрерывные случайные величины наталкивается на ряд сложностей. Можно по-разному преодолеть эти сложности; выберем кратчайший путь. Прямая аналогия

$$-\sum_k p_k \log p_k \rightarrow \int_X p(x) \log p(x) dx$$

не приводит к нужному результату; плотность $p(x)$ является размерной величиной*, а логарифм размерной величины не имеет смысла. Однако положение можно исправить, умножив $p(x)$ под знаком логарифма на величину E , имеющую ту же размерность, что и x :

$$-\sum_k p_k \log p_k \rightarrow \int_X p(x) \log [E \cdot p(x)] dx .$$

Теперь величину E можно принять равной единице измерения x , что приводит к функционалу

$$h(X) = -\int_X p(x) \log [p(x)] dx , \quad (3)$$

* Размерность плотности $p(x)$ обратна размерности x , так как элемент вероятности $dP(x) = p(x) dx$ безразмерен.

PROBABILITY
 вероятность
INDEPENDENT
 независимый
SYMBOL
 символ
ENTROPY
 энтропия
ERGODICITY
 эргодичность

Важным шагом в построении теории информации является введение количественной меры неопределенности – энтропии.
Оказывается, что функционал (1) обладает качествами, которые логично ожидать от меры неопределенности, и, наоборот, единственным функционалом с такими свойствами является именно функционал энтропии.
Обобщение понятия энтропии на непрерывные случайные величины приводит к выводу, что такое обобщение – дифференциальная энтропия – возможно лишь

множества возможных состояний – алфавитом, их число m – объемом алфавита. Число всевозможных последовательностей длины n , очевидно, равно m^n . Появление конкретной последовательности можно рассматривать как реализацию одного из m^n возможных событий. Зная вероятности символов и условные вероятности появления следующего символа, если известен предыдущий (в случае их зависимости), можно вычислить вероятность $P(C)$ для каждой последовательности C . Тогда энтропия множества $\{C\}$, по определению, равна

$$H_n = - \sum_{C \in \Omega} P(C) \log P(C) \quad (4)$$

Определим энтропию процесса H (среднюю неопределенность, приходящуюся на один символ) следующим образом*:

$$H = \lim_{n \rightarrow \infty} \frac{H_n}{n}. \quad (5)$$

На множестве $\{C\}$ можно задать любую числовую функцию $f_n(C)$, которая, очевидно, является случайной величиной. Определим $f_n(C)$ с помощью соотношения

$$f_n(C) = -\frac{1}{n} \log P(C).$$

Математическое ожидание этой функции

$$Mf_n(C) = -\frac{1}{n} \sum_C P(C) \log P(C)$$

который получил название **дифференциальной энтропии**. Это аналог энтропии дискретной величины, но аналог условный, относительный: ведь единица измерения произвольна. (Здесь $[p(x)]$ есть безразмерное представление плотности.) Запись (3) означает, что мы как бы сравниваем неопределенность случайной величины, имеющей плотность $p(x)$, с неопределенностью случайной величины, равномерно распределенной в единичном интервале. Поэтому величина $h(X)$ в отличие от $H(X)$ может быть не только положительной. Кроме того, $h(X)$ изменяется при нелинейных преобразованиях шкалы x , что в дискретном случае не играет роли. Остальные свойства $h(X)$ аналогичны свойствам $H(X)$, что делает дифференциальную энтропию очень полезной мерой.

Пусть, например, задача состоит в том, чтобы, зная лишь некоторые ограничения на случайную величину (типа моментов, пределов сверху и снизу области возможных значений и т.п.), задать для дальнейшего (каких-то расчетов или моделирования) конкретное распределение. Одним из подходов к решению этой задачи дает **принцип максимума энтропии**: из всех распределений, отвечающих данным ограничениям, следует выбирать то, которое обладает максимальной дифференциальной энтропией. Смысл этого критерия состоит в том, что, выбирая экстремальное по энтропии распределение, мы гарантируем наибольшую неопределенность, связанную с ним, т.е. имеем дело с наихудшим случаем при данных условиях.

ФУНДАМЕНТАЛЬНОЕ СВОЙСТВО ЭНТРОПИИ СЛУЧАЙНОГО ПРОЦЕССА

Особое значение энтропия приобретает в связи с тем, что она связана с очень глубокими, фундаментальными свойствами случайных процессов. Покажем это на примере процесса с дискретным временем и дискретным конечным множеством возможных состояний.

Назовем каждое такое состояние **символом**, множество возможных состояний – **алфавитом**, их число m – **объемом алфавита**. Число всевозможных последовательностей длины n , очевидно, равно m^n . Появление конкретной последовательности можно рассматривать как реализацию одного из m^n возможных событий. Зная вероятности символов и условные вероятности появления следующего символа, если известен предыдущий (в случае их зависимости), можно вычислить вероятность $P(C)$ для каждой последовательности C . Тогда энтропия множества $\{C\}$, по определению, равна

* Существование такого предела для любого стационарного процесса можно строго доказать.

откуда следует, что

$$M - \frac{1}{n} \log P \text{log} = \frac{H_n}{n}, \text{ и}$$

$$\lim_{n \rightarrow \infty} M - \frac{1}{n} \log P \text{log} = H. \quad (6)$$

Это соотношение, весьма интересное уже само по себе, является, однако, лишь одним из проявлений гораздо более общего свойства дискретных эргодических процессов. Оказывается, что не только математическое ожидание величины $f_n(C)$ при $n \rightarrow \infty$ имеет своим пределом H , но сама эта величина $f_n(C)$ стремится к H при $n \rightarrow \infty$. Другими словами, как бы малы ни были $\epsilon > 0$ и $\delta > 0$, при достаточно большом n справедливо неравенство [9]

$$P \left| \frac{1}{n} \log P \text{log} H \right| > \epsilon < \delta, \quad (7)$$

т.е. близость $f_n(C)$ к H при больших n является почти достоверным событием.

Для большей наглядности сформулированное фундаментальное свойство случайных процессов обычно излагают следующим образом. Для любых заданных $\epsilon > 0$ и $\delta > 0$ можно найти такое n_0 , что реализации любой длины $n > n_0$ распадаются на два класса:

группа реализаций, вероятности $P(C)$ которых удовлетворяют неравенству

$$\left| \frac{1}{n} \log P \text{log} H \right| < \epsilon; \quad (8)$$

группа реализаций, вероятности которых этому неравенству не удовлетворяют.

Так как согласно неравенству (7) суммарные вероятности этих групп равны соответственно $1 - \epsilon$ и ϵ , то первая группа называется *высоковероятной*, а вторая – *маловероятной*.

Это свойство эргодических процессов приводит к ряду важных следствий, из которых три заслуживают особого внимания.

1⁰. *Независимо от того, каковы вероятности символов и каковы статистические связи между ними, все реализации высоковероятной группы приблизительно равновероятны* (см. формулу (8)).

В связи с этим фундаментальное свойство иногда называют “свойством асимптотической равнораспределенности”. Это следствие, в частности, означает, что по известной вероятности $P(C)$ одной из реализаций высоковероятной группы можно оценить число N_1 реализаций в этой группе:

$$N_1 = \frac{1}{P \text{log}}.$$

2⁰. *Энтропия H_n с высокой точностью равна логарифму числа реализаций в высоковероятной группе:*

$$H_n = nH = \log N_1. \quad (9)$$

3⁰. *При больших n высоковероятная группа обычно охватывает лишь ничтожную долю всех возможных реализаций* (за исключением случая равновероятных и независимых символов, когда все реализации равновероятны и $H = \log m$).

Действительно, из соотношения (9) имеем $N_1 = a^{nH}$, где a – основание логарифма. Число N всех возможных реализаций есть $N = m^n = a^{n \log m}$. Доля реализаций высоковероятной группы в общем числе реализаций выражается формулой

$$N_1/N = a^{n(\log m - H)}, \quad (10)$$

и при $H < \log m$ эта доля неограниченно убывает с ростом n . Например, если $a = 2$, $n = 100$, $H = 2,75$, $m = 8$, то $N_1/N = 2^{-25} \sim (3 \cdot 10^7)^{-1}$, т.е. к высоковероятной группе относится лишь одна тридцатимиллионная доля всех реализаций!

Строгое доказательство фундаментального свойства эргодических процессов сложно и здесь не приводится. Однако следует отметить, что в простейшем случае независимости символов это свойство является следствием закона больших чисел. Действительно, закон больших чисел утверждает, что с вероятностью, близкой к 1, в длинной реализации i -й символ, имеющий вероятность p_i , встретится примерно $n p_i$ раз. Следовательно, вероятность реализа-

ции высоковероятной группы есть $P \log \prod_{i=1}^m p_i^{n p_i}$, откуда $-\log P(C) = -n \sum_{i=1}^m p_i \log p_i = nH$, что и доказывает справедливость фундаментального свойства в этом случае.

Подведем итог

Связав понятие неопределенности дискретной величины с распределением вероятности по возможным состояниям и потребовав некоторых естественных свойств от количественной меры неопределенности, мы приходим к выводу, что такой мерой может служить только функционал (1), названный энтропией. С некоторыми трудностями энтропийный подход удалось обобщить на непрерывные случайные величины (введением (3)) и на дискретные случайные процессы.

Summary

Linking the concept of the uncertainty of a discrete random variable and the form of its probability distribution, and demanding certain reasonable properties from the quantitative measure of uncertainty, we arrive at the conclusion that such a measure may only be the functional (1), which is called entropy. The entropy approach may be extended (with some difficulty) to continuous random variables – by the introduction of differential entropy as well as to random processes (we have considered here (3)) – only discrete processes).

§ 5.7. КОЛИЧЕСТВО ИНФОРМАЦИИ

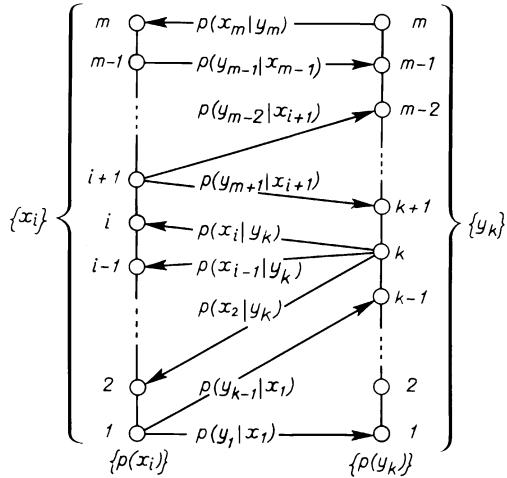
В основе всей теории информации лежит открытие, что *информация допускает количественную оценку*. В простейшей форме эта идея была выдвинута еще в 1928 г. Хартли, но завершенный и общий вид придал ей Шэннон в 1948 г. [9]. Не останавливаясь на том, как развивалось и обобщалось понятие количества информации, дадим сразу его современное толкование.

КОЛИЧЕСТВО ИНФОРМАЦИИ

КАК МЕРА СНЯТОЙ НЕОПРЕДЕЛЕННОСТИ

Процесс получения информации можно интерпретировать как изменение неопределенности в результате приема сигнала. Проиллюстрируем эту идею на примере достаточно простого случая, когда передача сигнала происходит при следующих условиях: 1) полезный (отправляемый) сигнал является последовательностью статистически независимых символов с вероятностями $p(x_i)$, $i = \overline{1, m}$; 2) принимаемый сигнал является последовательностью символов y_k того же алфавита; 3) если шумы (искажения) отсутствуют, то принимаемый сигнал совпадает с отправляемым $y_k = x_i$; 4) если шум имеется, то его действие приводит к тому, что данный символ может либо остаться прежним (i -м), либо быть подмененным любым другим (k -м) символом, вероятность этого равна $p(y_k | x_i)$; 5) искажение очередного символа является событием, статистически независимым от того, что произошло с предыдущими символами. Конечно, можно рассматривать ситуацию и со стороны передатчика, используя вероятности $p(x_i | y_k)$ (рис. 5.4). В этих условиях энтропия процесса есть энтропия одного символа, и все сводится к рассмотрению посимвольного приема.

Итак, до получения очередного символа ситуация характеризуется неопределенностью того, какой символ будет отправлен, т.е. априорной энтропией $H(X)$. После получения символа y_k неопределенность относительно того, какой символ был отправлен, меняется: в случае отсутствия шума она вообще исчезает (апостериорная энтропия равна нулю, поскольку точно известно, что был передан символ $x_k = y_k$), а при наличии шума мы не можем быть уверены, что полученный нами символ и есть отправленный, и возникает неопределенность,



5.4

Схема “вееров неопределенности” при наличии шума в канале

характеризуемая апостериорной энтропией $H(X | y_k) = H(\{p(x_i | y_k)\}) > 0$. В среднем после получения очередного символа энтропия $H(X | Y) = M_y H(X | y_k)$.

Определим теперь **количество информации** как меру снятой неопределенности: *числовое значение количества информации о некотором объекте равно разности априорной и апостериорной энтропий этого объекта*, т.е.

$$I(X, Y) = H(X) - H(X | Y). \quad (1)$$

Используя равенство (2) § 5.6, легко получить, что

$$I(X, Y) = H(Y) - H(Y | X). \quad (2)$$

В явной форме равенство (1) запишется так:

$$\begin{aligned} I(X, Y) &= H(X) - H(X | Y) = -\sum_{i=1}^m p(x_i) \log p(x_i) \\ &+ \sum_{k=1}^m p(y_k) \sum_{i=1}^m p(x_i | y_k) \log p(x_i | y_k) = \\ &= -\sum_{k=1}^m \sum_{i=1}^m p(x_i, y_k) \log p(x_i, y_k) + \sum_{i=1}^m \sum_{k=1}^m p(x_i, y_k) \log p(x_i | y_k) = \\ &= \sum_{i=1}^m \sum_{k=1}^m p(x_i, y_k) \log \frac{p(x_i, y_k)}{p(x_i)}, \end{aligned} \quad (3)$$

а для равенства (2) имеем

$$I(X, Y) = \sum_{i=1}^m \sum_{k=1}^m p(x_i, y_k) \log \frac{p(y_k | x_i)}{p(y_k)}. \quad (4)$$

КОЛИЧЕСТВО ИНФОРМАЦИИ КАК МЕРА СООТВЕТСТВИЯ СЛУЧАЙНЫХ ОБЪЕКТОВ

Этим формулам легко придать полную симметричность: умножив и разделив логарифмируемое выражение в (3) на $p(y_k)$, а в (4) на $p(x_i)$, сразу получим, что

$$I(X, Y) = \sum_{i=1}^m \sum_{k=1}^m p(x_i, y_k) \log \frac{p(x_i, y_k)}{p(x_i)p(y_k)} \quad (5)$$

Эту симметрию можно интерпретировать так: количество информации в объекте X об объекте Y равно количеству информации в объекте Y об объекте X . Таким образом, количество информации является не характеристикой одного из объектов, а характеристикой их связи, соответствия между их состояниями. Подчеркивая это, можно сформулировать еще одно определение: *среднее количество информации, вычисляемое по формуле (5), есть мера соответствия двух случайных объектов*.

INTERACTION
<i>взаимодействие</i>
DISTORTION
<i>искажение</i>
QUANTITY
<i>количество</i>
DENSITY
<i>плотность</i>
AVERAGE
<i>средний</i>

Количество информации можно определить как меру уменьшения неопределенности в результате получения сигнала. Это соответствует разности энтропий до и после приема сигнала. Среди свойств количества информации выделяются следующие: 1) количество информации (в отличие от эн-

$$I(X, X) = H(X).$$

4⁰. Преобразование ?(·) одной случайной величины не может увеличить содержание в ней информации о другой, связанной с ней, величине:

$$I(?(X), Y) \leq I(X, Y) \quad (9)$$

5⁰. Для независимых пар величин количество информации аддитивно:

$$I(\{X_i, Y_i\}) = \sum_i I(X_i, Y_i). \quad (10)$$

ЕДИНИЦЫ ИЗМЕРЕНИЯ ЭНТРОПИИ И КОЛИЧЕСТВА ИНФОРМАЦИИ

Рассмотрим теперь вопрос о единицах измерения количества информации и энтропии. Из определений I и H следует их безразмерность, а из линейности их связи – одинаковость их единиц. Поэтому будем для определенности говорить об энтропии. Начнем с дискретного случая. За единицу энтропии примем неопределенность случайного объекта, такого, что

$$H(Y) = -\sum_{i=1}^m p_i \log p_i = 1. \quad (11)$$

Легко установить, что для однозначного определения единицы измерения энтропии необходимо конкретизировать число m состояний объекта и основание логарифма. Возьмем для

Это определение позволяет прояснить связь понятий информации и количества информации. Информация есть отражение одного объекта другим, проявляющееся в соответствии их состояний. Один объект может быть отражен с помощью нескольких других, часто какими-то лучше, чем остальными. Среднее количество информации и есть числовая характеристика степени отражения, степени соответствия. Подчеркнем, что при таком описании как отражаемый, так и отражающий объекты выступают совершенно равноправно. С одной стороны, это подчеркивает обоюдность отражения: каждый из них содержит информацию друг о друге. Это представляется естественным, поскольку отражение есть результат взаимодействия, т.е. взаимного, обоюдного изменения состояний. С другой стороны, фактически одно явление (или объект) всегда выступает как причина, другой – как следствие; это никак не учитывается при введенном количественном описании информации.

Формула (5) обобщается на непрерывные случайные величины, если в соотношения (1) и (2) вместо H подставить дифференциальную энтропию h ; при этом исчезает зависимость от стандарта Е и, значит, количество информации в непрерывном случае является столь же безотносительным к единицам измерения, как и в дискретном:

$$I(X, Y) = \int_{XY} b \log \frac{p(x, y)}{p(x)p(y)} dx dy, \quad (6)$$

где $p(x)$, $p(y)$ и $p(x, y)$ – соответствующие плотности вероятностей.

СВОЙСТВА КОЛИЧЕСТВА ИНФОРМАЦИИ

Отметим некоторые важные свойства количества информации.

1⁰. Количество информации в случайном объекте X относительно объекта Y равно количеству информации в Y относительно X :

$$I(X, Y) = I(Y, X). \quad (7)$$

2⁰. Количество информации неотрицательно:

$$I(X, Y) \geq 0. \quad (8)$$

Это можно доказать по-разному. Например, варьированием $p(x, y)$ при фиксированных $p(x)$ и $p(y)$ можно показать, что минимум I , равный нулю, достигается при $p(x, y) = p(x)p(y)$.

3⁰. Для дискретных X справедливо равенство

$$I(X, X) = H(X).$$

4⁰. Преобразование ?(·) одной случайной величины не может увеличить содержание в ней информации о другой, связанной с ней, величине:

$$I(?(X), Y) \leq I(X, Y) \quad (9)$$

5⁰. Для независимых пар величин количество информации аддитивно:

$$I(\{X_i, Y_i\}) = \sum_i I(X_i, Y_i). \quad (10)$$

определенности наименьшее число возможных состояний, при котором объект еще остается случайным, т.е. $m = 2$, и в качестве основания логарифма также возьмем число 2. Тогда из равенства $-p_1 \log_2 p_1 - p_2 \log_2 p_2 = 1$ вытекает, что $p_1 = p_2 = 1/2$. Следовательно, единицей неопределенности служит энтропия объекта с двумя равновероятными состояниями. Эта единица получила название “бит”. Бросание монеты дает количество информации в один бит. Другая единица (“нит”) получается, если использовать натуральные логарифмы, обычно она употребляется для непрерывных величин.

КОЛИЧЕСТВО ИНФОРМАЦИИ В ИНДИВИДУАЛЬНЫХ СОБЫТИЯХ

Остановимся еще на одном важном моменте. До сих пор речь шла о среднем количестве информации, приходящемся на любую пару состояний (x_i, y_k) объектов X и Y . Эта характеристика естественна для рассмотрения особенностей стационарно функционирующих систем, когда в процессе функционирования (рано или поздно, реже или чаще) принимают участие все возможные пары (x_i, y_k) . Однако в ряде практических случаев оказывается необходимым рассмотреть информационное описание конкретной пары состояний, оценить содержание информации в конкретной реализации сигнала. Тот факт, что некоторые сигналы несут информации намного больше, чем другие, виден на примере того, как отбираются новости средствами массовой информации (скажем, все радиостанции и газеты сообщают о рождении шестерых близнецов где-то в Южной Америке, но о рождении двойни обычно не пишут).

Допуская существование количественной меры информации $i(x_i, y_k)$ в конкретной паре (x_i, y_k) , естественно потребовать, чтобы индивидуальное и среднее количества информации удовлетворяли соотношению

$$i(x_i, y_k) \leq \sum_{i,k} p(x_i, y_k) i(x_i, y_k). \quad (12)$$

Хотя равенство сумм имеет место не только при равенстве всех слагаемых, сравнение формул (12) и, например, (4) наталкивает на мысль, что мерой индивидуальной информации в дискретном случае может служить величина

$$b_{x_i, y_k} = \log \frac{p(x_i, y_k)}{p(x_i) p(y_k)} = \log \frac{p(x_i | y_k)}{p(y_k)} = \log \frac{p(x_i, y_k)}{p(x_i) p(y_k)}, \quad (13)$$

а в непрерывном – величина

$$b_{x_i, y_k} = \ln \frac{p(x_i | y_k)}{p(x_i)} = \ln \frac{p(x_i | y_k)}{p(y_k)} = \ln \frac{p(x_i, y_k)}{p(x_i) p(y_k)}, \quad (14)$$

называемая *информационной плотностью*. Свойства этих величин согласуются с интуитивными представлениями (в том числе и возможная отрицательность при положительности в среднем) и, кроме того, доказана единственность меры, обладающей указанными свойствами. Полезность введения понятия индивидуального количества информации проиллюстрируем на следующем примере.

Пусть по выборке (т.е. совокупности наблюдений) $x = x_1, \dots, x_N$ требуется отдать предпочтение одной из конкурирующих гипотез (H_1 или H_0), если известны распределения наблюдений при каждой из них, т.е. $p(x | H_0)$ и $p(x | H_1)$. Как обработать выборку? Из теории известно, что никакая обработка не может увеличить количества информации, содержащегося в выборке x (см. формулу (9)). Следовательно, выборке x нужно поставить в соответствие число, содержащее всю полезную информацию, т.е. обработать выборку без потерь. Возникает мысль о том, чтобы вычислить индивидуальные количества информации в выборке x о каждой из гипотез и сравнить их:

$$\Delta i = i(H_1) - i(H_0) = \ln \frac{p(H_1)}{p(H_0)} - \ln \frac{p(H_0)}{p(H_1)} = \ln \frac{p(H_1)}{p(H_0)}. \quad (15)$$

Какой из гипотез отдать предпочтение, зависит теперь от величины Δi и от того, какой порог сравнения мы назначим. Оказывается, что мы получили статистическую процедуру, оптимальность которой специально доказывается в математической статистике, – именно к этому

сводится содержание фундаментальной леммы Неймана – Пирсона. Данный пример иллюстрирует эвристическую силу теоретико-информационных представлений.

Подведем итог

Summary

Главным результатом данного параграфа является открытие К.Шэнноном возможности количественного описания информационных процессов в системах и получение им формулы для количества информации.

The gist of this section is the possibility, discovered by C.Shannon, of quantitatively describing information processes in systems and finding the formula that expresses the quantity of information.

§ 5.8. ОБ ОСНОВНЫХ РЕЗУЛЬТАТАХ ТЕОРИИ ИНФОРМАЦИИ

Теория информации приводит к ряду новых понятий, описывающих информационные процессы, происходящие в любых системах, к введению новых количественных параметров, позволяющих проводить измерения и расчеты. Часть этих понятий и величин мы рассмотрели в предыдущих параграфах, некоторые другие опишем ниже. Однако главная ценность теории информации заключается в полученных ею новых результатах, в открытии ранее неизвестных свойств систем.

Чтобы познакомиться с основными из этих результатов, введем еще несколько понятий и параметров информационных процессов и систем.

ИЗБЫТОЧНОСТЬ

Одной из важнейших характеристик сигнала является содержащееся в нем количество информации. Однако по ряду причин количество информации, которое несет сигнал, обычно меньше, чем то, которое он мог бы нести по своей физической природе; информационная нагрузка на каждый элемент сигнала меньше той, которую элемент способен нести. Для описания этого свойства сигналов введено понятие избыточности и определена ее количественная мера.

Пусть сигнал длиной в n символов содержит количество информации I . Если это представление информации обладает избыточностью, то такое же количество информации I может быть представлено с помощью меньшего числа символов. Обозначим через n_0 наименьшее число символов, необходимое для представления I без потерь. На каждый символ в первом случае приходится $I_1 = I/n$ бит информации, во втором $I_{1\max} = I/n_0$ бит. Очевидно, $nI_1 = n_0I_{1\max} = I$. В качестве меры избыточности R принимается относительное удлинение сигнала, соответствующее данной избыточности:

$$R = \frac{n-n_0}{n} = 1 - \frac{n_0}{n} = 1 - \frac{I_1}{I_{1\max}}. \quad (1)$$

В дискретном случае имеются две причины избыточности: неравновероятность символов и наличие статистической связи между символами. В непрерывном случае – это неэкстремальность распределений (т.е. отклонение от распределений, обладающих максимальной энтропией), что в широком смысле сводится к отклонениям от экстремальности распределения первого порядка и от минимальности связи во времени (от равномерности спектра при его ограниченности).

Не следует думать, что избыточность – явление всегда отрицательное. При искажениях, выпадениях и вставках символов именно избыточность позволяет обнаружить и исправить ошибки.

СКОРОСТЬ ПЕРЕДАЧИ И ПРОПУСКНАЯ СПОСОБНОСТЬ

Следующим важнейшим понятием является **скорость передачи информации**. Так называется количество информации, передаваемое в единицу времени. Эта величина определяется по формуле

$$R = H(X) - H(X | Y), \quad (2)$$

где указанные энтропии исчисляются на единицу времени*. В дискретном случае единицей времени удобно считать время передачи одного символа, тогда в формуле (2) фигурируют априорная и апостериорная энтропии на один символ. Для непрерывных каналов единицей времени может служить либо обычная единица (например, секунда), либо интервал между отсчетами (см. § 5.5); тогда в формулу (2) входят соответствующие дифференциальные энтропии. Для более наглядного представления о величине R укажем, что темп обычной речи соответствует скорости порядка 20 бит/с, муравьи обмениваются информацией путем касания усиками со скоростью около 1/10 бит/с.

Скорость передачи информации по каналу связи зависит от многих факторов – от энергии сигнала, числа символов в алфавите избыточности, полосы частот, способа кодирования и декодирования. Если имеется возможность изменять некоторые из них, то, естественно, следует делать это так, чтобы максимально увеличить скорость. Оказывается, что обычно существует предел, выше которого увеличение скорости невозможно. Этот предел называется **пропускной способностью канала**:

$$C = \sup_{\{A\}} R_A, \quad (3)$$

где R_A – скорость передачи информации при условиях A , $\{A\}$ – множество вариантов условий, подлежащих перебору. Так как множество $\{A\}$ можно определить по-разному, то имеет смысл говорить о нескольких типах пропускных способностей. Наиболее важным является случай, когда мощность сигнала (объем алфавита) фиксирована, а варьировать можно только способ кодирования. Именно таким образом пропускную способность определил К.Шэннон [9]. С другой стороны, В.И.Сифоров показал, что целесообразно рассмотреть предел, к которому стремится шэнноновская пропускная способность C при стремлении мощности полезного сигнала к бесконечности. Оказалось, что все каналы связи разбиваются на два класса: каналы первого рода (терминология Сифорова), для которых указанный предел бесконечен, и каналы второго рода, имеющие конечную пропускную способность даже при бесконечной мощности передатчика. Этот предел называют *собственной пропускной способностью*. При других требованиях, предъявляемых к множеству $\{A\}$, мы придем к тем или иным *условным пропускным способностям*.

Для представления о порядках величин C приведем примеры. Прямыми измерениями установлено, что пропускные способности зрительного, слухового и тактильного каналов связи человека имеют порядок 50 бит/с (вопреки распространенному мнению о сильном отличии зрительного канала). Возможно, ограничивающим фактором являются не сами рецепторы, а нервные волокна, передающие возбуждения. Если включить в канал и “исполнительные” органы человека (например, предложить ему нажимать педаль или кнопку в темпе получения сигналов), то пропускная способность близка к 10 бит/с. Интересно отметить, что многие бытовые технические устройства слабо согласованы с органами чувств человека. Например, канал телевидения имеет пропускную способность в десятки миллионов бит/с.

КОДИРОВАНИЕ В ОТСУТСТВИЕ ШУМОВ

С помощью введенных понятий можно рассмотреть многие информационные процессы. Начнем с дискретных систем без шумов. Здесь главное внимание привлекает проблема эффективности: важно, чтобы данная информация заняла в запоминающем устройстве как можно меньше ячеек, при передаче желательно занимать канал связи на максимально короткий срок. В такой постановке задачи легко распознается проблема устранения всякой избыточности. Однако эта проблема не тривиальна.

Пусть алфавит системы состоит из m символов. Средствами этого алфавита требуется представить любой из M возможных сигналов $\{u_k\}$, $k = \overline{1, M}$, вероятности которых $\{p(u_k)\}$ за-

* Совпадение обозначений избыточности и скорости передачи информации установилось случайно: это первые буквы соответствующих английских терминов: redundancy – избыточность и rate – темп.

даны. Обычно $M > m$, поэтому каждый из сигналов, подлежащих передаче, невозможно обозначить только одним символом и приходится ставить им в соответствие некоторые последовательности символов; назовем их *кодовыми словами*. Так как возможна последовательная передача разных кодовых слов, то они не только должны различаться для разных u_k , но и не должны быть продолжением других, более коротких. Пусть сигналу u_k соответствует кодовое слово длиной l_k символов. При стационарной передаче сигналов с вероятностями $\{p(u_k)\}$ средняя длина кодового слова равна $L = \sum_k l_k p(u_k)$.

Возникает вопрос о том, как выбирать L и $\{l_k\}$. Он не имеет смысла, пока не задан критерий выбора этих величин. Определим этот критерий так: L и $\{l_k\}$ должны быть минимальными величинами, такими, при которых еще не происходит потери информации. Напомним, что в отсутствие шумов среднее количество информации на один элемент u_k ансамбля $\{u_k\}$ равно энтропии этого ансамбля, т.е. $H(U) = -\sum_{k=1}^M p(u_k) \log p(u_k)$, а индивидуальное количество информации в u_k есть $i(u_k) = -\log p(u_k)$. С другой стороны, на один символ придется максимальное количество i информации, если символы равновероятны и независимы; при этом $i = \log m$. Поскольку кодирование должно вестись без потерь информации, сразу получаем

$$\frac{H(U)}{\log m} \leq L, \quad (4)$$

$$\frac{i(u_k)}{\log m} \leq l_k \leq \frac{i(u_k)}{\log m} + 1. \quad (5)$$

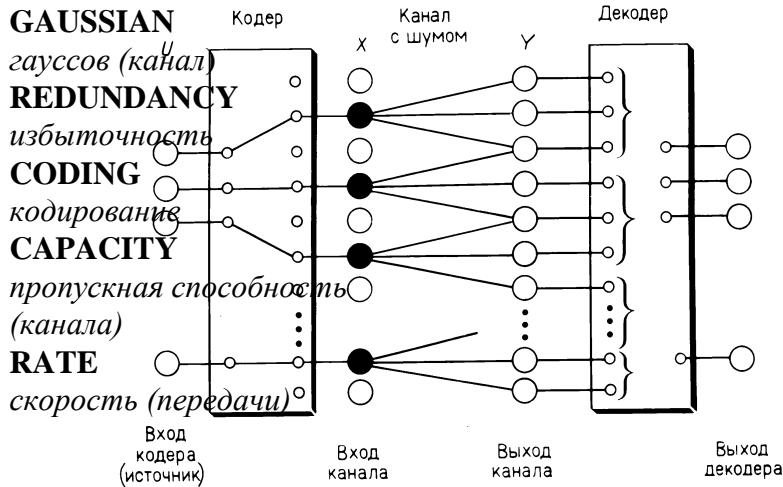
Так из общих соображений находим нижние границы для L и l_k . В теории информации доказываются теоремы об условиях достижимости этих границ. Мы не будем приводить эти теоремы; укажем лишь, что речь идет не только о принципиальной возможности: разработаны процедуры построения кодов, обеспечивающих безызбыточное кодирование (либо в случае невозможности этого – сколь угодно близкое к нему).

КОДИРОВАНИЕ ПРИ НАЛИЧИИ ШУМОВ

Наиболее интересные и важные результаты были получены при рассмотрении передачи информации по каналам связи с шумами. В этом случае безызбыточное кодирование приведет к безвозвратным потерям информации:искаженный символ нельзя ни обнаружить, ни исправить. Для борьбы с влиянием помех необходимо ввести избыточность в сигнал. Основываясь на интуитивных соображениях (например, на опыте многократного повторения), легко прийти к выводу, что при неограниченном повышении требований к малости вероятности ошибки избыточность и при любом способе кодирования должна неограниченно возрастать, а скорость передачи – стремиться к нулю. Здесь мы имеем яркий пример того, как сильно интуиция может привести к заблуждению. Шэннон показал, что *существуют такие способы введения избыточности, при которых обеспечиваются одновременно и сколь угодно малая вероятность ошибки, и конечная (отличная от нуля) скорость передачи информации, причем эта скорость может быть сколь угодно близкой к пропускной способности канала*. Это замечательное открытие и привлекло столь большое внимание к теории информации.

Воспроизведем упрощенное доказательство указанного утверждения. Рассмотрим схему передачи по каналу с шумом (рис. 5.5). Будем считать, что на вход кодера сигналы поступают закодированными безызбыточно. Кодер вносит в сигналы избыточность, увеличивая длительность кодовых слов. Число возможных последовательностей сразу резко увеличивается, но избыточность и состоит в том, что к отправке предназначаются не все из них, а лишь разрешенные (отмеченные на рис. 5.5 черными кружками). Согласно фундаментальному свойству энтропии (см. формулу (8) § 5.6), число всевозможных последовательностей* длины n равно $2^{nH(X)}$, а число разрешенных* к отправке равно $2^n < 2^{nH(X)}$ (считаем, что энтропия исчисляется в битах); H – энтропия на символ во множестве разрешенных к отправке последовательностей (“энтропия источника”, или “скорость создания информации”), $H(X)$ –

* В высоковероятной группе реализаций.



Для информационных процессов в отсутствие шумов Схема передачи информации по каналу с шумами

блемой является наиболее эффективное (безыбыточное) представление информации. Свойства таких оптимальных кодов легко определяются из эвристических соображений (хотя могут быть доказаны и строго), что иллюстрирует полезность введенных понятий (см. (5) и (6)).

Пожалуй, самым важ-

энтропия на символ во множестве всевозможных последовательностей. В результате воздействия шумов какие-то из символов отправленной последовательности подменяются другими и на приемный конец поступает другая, отличная от отправленной, последовательность. Поскольку $p(x / y)$ считается известным, каждой принятой последовательности соответствует $2^{nH(X/Y)}$ возможно отправленных*. Декодирование (т.е. принятие решения о том, какая последовательность была отправлена) можно выразить как разбиение всего множества Y принимаемых последовательностей на $2^n H$ подмножеств, сопоставляемых с разрешенными к отправке: если, например, принят сигнал i -й группы, то считается, что был послан i -й разрешенный сигнал, который тут же выдается в "чистом" виде получателю.

Итак, в построенной модели проблема выбора кода (способа передачи) состоит в размещении разрешенных последовательностей среди множества всевозможных на передающем конце и в разбиении этого множества из $2^{nH(X)}$ последовательностей на $2^n H$ подмножеств на приемном конце. Идея Шэннона состоит не в том, чтобы указать некоторый регулярный способ кодирования со сколь угодно малой вероятностью ошибки, а в том, чтобы показать, что такой код вообще существует. Рассмотрим класс всевозможных кодов, которые получаются, если разрешенные последовательности размещать среди всевозможных случайным образом, а в качестве декодирующего подмножества брать $2^{nH(X/Y)}$ последовательностей высоковероятностной группы, соответствующей принятому сигналу.

Вероятность ошибки при этом равна вероятности того, что среди $2^{nH(X/Y)}$ последовательностей окажется более одной разрешенной. Так как код получается в результате случайного (равновероятностного) выбора 2^{nH} последовательностей среди $2^{nH(X)}$, то вероятность того, что данная последовательность окажется разрешенной, равна $2^{nH}/2^{nH(X)} = 2^{n(H-H(X))}$. Средняя вероятность того, что в декодирующем подмножестве из $2^{nH(X/Y)}$ последовательностей имеется только одна разрешенная, выражается соотношением

$$P = (1 - 2^{-n(H-H(X))})^{2^{nH(X/Y)}} ; \quad (6)$$

это и есть вероятность безошибочного приема. Поскольку $H < C = H(X) - H(X/Y)$, имеем $H - H(X) = -H(X/Y) - \eta$, где $\eta > 0$. Отсюда (пренебрегая единицей по сравнению с $2^{nH(X/Y)}$) находим

$$P = (1 - 2^{-n[H(Y|I+\eta)]})^{2^{nH(Y|I)}}. \quad (7)$$

Логарифмируя и применяя правило Лопиталя, легко показать, что $\lim_{n \rightarrow \infty} P = 1$, т.е. что при кодировании достаточно длинными блоками средняя вероятность ошибки может быть сделана сколь угодно малой. Доказательство завершается утверждением, что существуют коды с вероятностями ошибок меньше средней.

ПРОПУСКНАЯ СПОСОБНОСТЬ ГАУССОВА КАНАЛА СВЯЗИ

Перейдем теперь к знакомству с основными результатами для систем с непрерывными сигналами. Наиболее известным выводом теории является формула для пропускной способности гауссова канала связи, которую мы сейчас получим.

Гауссовым каналом называется канал связи, для которого выполняются следующие условия:

- 1⁰) сигналы и шумы в нем непрерывны;
- 2⁰) канал занимает ограниченную полосу частот шириной F ;
- 3⁰) шум $n(t)$ в канале распределен нормально (“гауссов шум”);
- 4⁰) спектр мощности шума равномерен в полосе частот канала и равен N единиц мощности на единицу полосы частот;
- 5⁰) средняя мощность полезного сигнала $x(t)$ фиксирована и равна P_0 ;
- 6⁰) сигнал и шум статистически независимы;
- 7⁰) принимаемый сигнал $y(t)$ есть сумма полезного сигнала и шума: $y(t) = x(t) + n(t)$ (“шум аддитивен”).

Эти предположения позволяют вычислить пропускную способность гауссова канала. Во-первых, ограниченность полосы частот позволяет применить теорему отсчетов (см. § 5.5) и вести рассуждения для каждого отсчета в отдельности. Далее, аддитивность шума и его независимость от X позволяют представить количество информации в Y об X в виде

$$I(Y, X) = h(Y|X) - h(Y|N) = h(Y|N) = h(Y) - h(N), \quad (8)$$

где $h(N)$ – дифференциальная энтропия шума. Следовательно, пропускная способность такова:

$$C = \max_{\{p(x)\}} h(Y) - h(N).$$

Согласно условиям 3⁰ и 4⁰, имеем

$$h(Y) = F \log \frac{1}{2\pi e N} \int_{-\infty}^{\infty} |Y(f)|^2 df. \quad (9)$$

В силу условий 4⁰ – 7⁰ мощность принимаемого сигнала есть

$$P_0 = P_0 + NF. \quad (10)$$

Максимум $h(Y)$ при условии (10) достигается в случае нормального распределения (см. упражнение 2б) к § 5.6), т.е.

$$\max_{\{p(x)\}} h(Y) = F \log \frac{1}{2\pi e P_0 + NF} \int_{-\infty}^{\infty} |Y(f)|^2 df. \quad (11)$$

Так как шум имеет равномерный спектр (см. условие 4⁰) и спектр смеси $y(t)$ также равномерен (вследствие независимости отсчетов), то и полезный сигнал должен иметь равномерный спектр. Вводя спектральную плотность $P = P_0/F$ и вычитая равенство (9) из (11), получаем известную формулу Шэннона – Таллера

$$C = F \log \left(1 + \frac{P}{N} \right). \quad (12)$$

Таким образом, мы не только определили пропускную способность, но и заодно показали, что она реализуется, если полезный сигнал закодировать так, чтобы его спектр был равномерным в представленной полосе частот, а распределение мгновенных значений – нормальным.

Дальнейшие обобщения связаны с рассмотрением “небелых” сигналов и шумов (предполагается зависимость от частоты их спектральных мощностей $P(f)$ и $N(f)$), а также с допущением случайности величины P (например, в случае замираний радиосигналов). Решения возникающих при этом задач имеются в литературе по теории информации.

Подведем итог

Теоремы Шеннона о кодировании в каналах без шумов и при наличии шумов, приведенные в данном параграфе, относятся к числу основных результатов теории информации. Эти теоремы сильно расширили понимание природы информационных процессов, происходящих во всех системах.

Summary

Shannon's theorems of coding in noiseless and in noisy channels, given in this section, are among the main findings of information theory. These theorems have greatly broadened our understanding of the nature of information processes in various systems.

ЗАКЛЮЧЕНИЕ

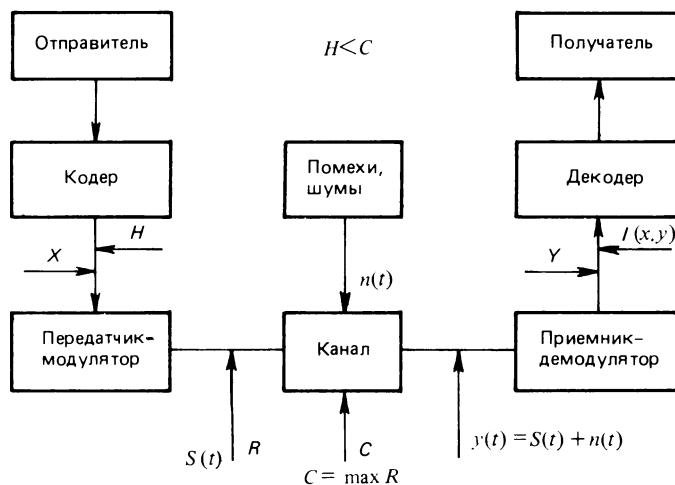
Отметим, что наиболее важные результаты теории информации – теоремы о кодировании – являются теоремами о существовании и носят асимптотический характер, т.е. не являются конструктивными. Однако уже само знание потенциальных возможностей имеет огромное значение: сравнение характеристик реальных систем с теоретическими пределами позволяет судить о достигнутом уровне и о целесообразности дальнейших затрат на его повышение. Прикладные же вопросы рассматриваются в специальном разделе теории информации – теории кодирования, которая изучает способы построения конкретных кодов и их свойства, в частности точные или граничные зависимости вероятностей ошибок от параметров кода. Для напоминания об основных понятиях теории информации приведем вариант блок-схемы системы передачи информации (рис. 5.6). На этом рисунке приведено большинство введенных в данной главе понятий и обозначений.

Значение теории информации выходит далеко за рамки теории связи, так как именно ее появление привело к широкому обсуждению новых понятий, к более глубокому пониманию открытых ранее закономерностей природы (например, второго закона термодинамики) и в конечном счете к тому, что понятие информации вошло теперь в число философских категорий, расширив и углубив тем самым наше видение и понимание мира.

Вместе с тем некоторые специалисты различных отраслей науки некритически распространяли методы и конкретные результаты теории информации на любые информационные процессы в реальном мире. Уже сам Шэннон заметил эту тенденцию и настоятельно предостерегал от того, чтобы ей безоглядно следовать [8]. Дело в том, что шэнноновское количество информации является характеристикой, лишь с одной стороны описывающей информационные отношения. Именно эта сторона – соответствие состояний – играет главную роль в технических устройствах. Однако в этом соответствии не проявляют себя (и, следовательно, могут не учитываться) такие свойства информации, как смысл, доброкачественность (степень истинности), ценность, полезность, временное старение, причинно-следственная направленность и т.д. – свойства, весьма существенные для информационных отношений с участием живых организмов, людей, коллективов.

Исследование проблем использования информации ведется в различных направлениях, достигнуты и успехи (семиотика, теория полезности, теория решений и выбора и т.д.). С некоторыми из них мы познакомимся в других разделах книги.

Необходимость расширения принципов и методов исследования информационных процессов вытекает не только из внутренней логики развития науки, стремящейся ко все большим обобщениям, снятию или ослаблению ограничений, предположений и т.п. К развитию новых подходов толкает сильное “внешнее” давление общественной практики: масштабы и значение информационных потоков в современном обществе так резко возросли в последние годы, что даже возникло образное выражение “информационный взрыв” [3]. Появилась новая отрасль – “индустрия обработки данных”, затраты на которую в промышленно развитых странах превосходят затраты на энергетику. Этой общественной потребности отвечает возникновение но-



5.6

вой отрасли науки – информатики.

Для системного анализа теория информации имеет двоякое значение. Во-первых, ее конкретные методы и результаты позволяют провести ряд количественных исследований информационных потоков в изучаемой или проектируемой системе (если в этом возникнет необходимость). Однако более важным является эвристическое значение основных понятий теории информации – неопределенности, энтропии, количества информации, избыточности, пропускной способности и пр. Их использование столь же важно для понимания системных процессов, как и использование понятий, связанных с временными, энергетическими процессами. Системный анализ неизбежно выходит на исследование ресурсов (наличных и необходимых), которые потребуются для решения анализируемой проблемы. Информационные ресурсы играют далеко не последнюю роль наряду с остальными ресурсами – материальными, энергетическими, временными, кадровыми.

ЛИТЕРАТУРА

1. *Бриллюэн Л.* Наука и теория информации. – М.: Физматгиз, 1960.
2. *Бриллюэн Л.* Научная неопределенность и информация. – М.: Сов. радио, 1970.
3. *Ефимов А.Н.* Информационный взрыв: проблемы реальные и мнимые. – М.: Наука, 1985.
4. *Майер А.Г., Леонович Е.А.* Об одном неравенстве, связанном с интегралом Фурье / ДАН СССР, 1934. Т. IV. № 7. С. 353 – 360.
5. *Тарасенко Ф.П.* Введение в курс теории информации. – Томск: ТГУ, 1963.
6. *Финк Л.М.* Сигналы. Помехи. Ошибки... – М.: Радио и связь, 1984.
7. *Френкс Л.* Теория сигналов. – М.: Сов. радио, 1974.
8. *Шэннон К.* Бандвагон. – В сб.: Работы по теории информации и кибернетике. – М.: ИЛ, 1963.
9. *Шэннон К., Уивер В.* Математическая теория связи. – В сб.: Работы по теории информации и кибернетике. – М.: ИЛ, 1963.

УПРАЖНЕНИЯ

§ 5.1

- Обсудите роль информации в парадоксе “демона Максвелла” [1, гл. 13].
- Некоторые философы считают, что информация в мозгу человека настолько качественно отличается от процессов в остальной природе, что их нельзя даже ставить в один ряд с процессами мышления. Обсудите аргументы за и против такого мнения. Обсудите истоки антропоцентризма.

§ 5.2

- При каких условиях можно отнести либо к статическим, либо к динамическим дымовые сигналы, запах, голограммическое изображение? Придумайте еще примеры трудно классифицируемых сигналов.

§ 5.3

- Обсудите подробнее сходство и различия между случайным процессом как моделью сигналов и реальными сигналами, приведите примеры.

§ 5.4

- Если вам интересно, попробуйте сами получить формулу для плотности распределения фазы смеси сигнала к шуму, вычислив интеграл (21).

§ 5.5

- Вычислите спектр прямоугольного импульса и убедитесь в его неограниченности по полосе частот.
- Проделайте подробно все выкладки, приводящие к ряду Котельникова (12).

§ 5.6

- Воспроизведите доказательства свойств $1^0 - 8^0$ энтропии, следуя указаниям, данным в конце описания каждого свойства.
- Задача нахождения функции $p(x)$, реализующей максимум функционала $\Phi_0[p(x)]$ при условии постоянства функционалов $\Phi_i[p(x)]$, $i=1, k$, называется изопериметрической задачей вариационного исчисления и сводится к решению уравнения Эйлера

$$\frac{\delta \Phi_0}{\delta p} + \sum_i \lambda_i \frac{\delta \Phi_i}{\delta p} = 0.$$

С помощью этого метода найдите распределения, обладающие максимальной дифференциальной энтропией при заданных условиях:

a) $\Phi_0 = - \int_a^b p(x) \ln p(x) dx,$

$$\Phi_1 = \int_a^b p(x) dx = 1.$$

Ответ: равномерное распределение, $p(x) = 1/(b-a)$.

b) $\Phi_0 = - \int_{-\infty}^{\infty} p(x) \ln p(x) dx,$

$$\Phi_1 = \int_{-\infty}^{\infty} p(x) dx = 1,$$

$$\Phi_2 = \int_{-\infty}^{\infty} x^2 p(x) dx = a,$$

$$\Phi_3 = \int_{-\infty}^{\infty} (x-a)^2 p(x) dx = \sigma^2.$$

Ответ: нормальное распределение, $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right)$.

v) $\Phi_0 = - \int_0^{\infty} p(x) \ln p(x) dx,$

$$\Phi_1 = \int_0^{\infty} p(x) dx = 1,$$

$$\Phi_2 = \int_0^{\infty} x p(x) dx = a.$$

Ответ: экспоненциальное распределение, $p(x) = \frac{1}{a} \exp\left(-\frac{x}{a}\right)$.

Вычислите дифференциальные энтропии экстремальных распределений, полученных в предыдущем упражнении.

§ 5.7

- Докажите справедливость свойств $1^0 - 5^0$ количества информации.
- Используя формулу (15), определите оптимальный алгоритм обработки выборки наблюдений x_1, \dots, x_N для проверки гипотезы H_0 : $p|_{H_0}$ $\text{d}H_0 = \frac{1}{\sqrt{2\pi\sigma^2}} \times \exp\left\{-\frac{x^2}{2\sigma^2}\right\}$ против альтернативы H_1 :

$$p|_{H_1} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\}$$

Ответ: $\Delta i = \frac{1}{N} \sum_{i=1}^N x_i$.

§ 5.8

- Пропускная способность человеческого зрения – около 50 бит/с, а пропускная способность телевизионного канала – около 50 млн. бит/с. Обсудите этот факт.

- Докажите, что предел вероятности, определенной формулой (7), равен 1 в соответствии с указанием в тексте.

**Вопросы
для самопроверки**

1. Каково обязательное условие того, чтобы один объект содержал информацию о другом объекте?
2. Может ли информация не иметь материального носителя?
3. Почему заданная функция времени не может быть адекватной моделью сигнала?
4. Какое главное свойство сигнала отображается математической моделью случайного процесса?
5. Какие расхождения между реальным сигналом и математической моделью случайного процесса вы можете назвать?
6. Какой смысл вы видите в дискретном представлении непрерывных сигналов?
7. Каковы различия в свойствах энтропии дискретных случайных объектов и дифференциальной энтропии и чем объясняются эти различия?
8. Не кажется ли вам удивительным, что доля реализаций высоковероятной группы неграниченно убывает с ростом длины реализации и в то же время именно эта группа определяет свойства случайного процесса в целом?
9. Почему энтропию и количество информации можно измерять в одинаковых единицах?
10. При каких условиях избыточность вредна и при каких полезна?
11. Что такое пропускная способность канала связи?
12. В чем вам видится ограниченность теории информации при описании реальных информационных процессов?

QUALITATIVE
качественный
QUANTITATIVE
количественный
OBSERVATION
наблюдение

Отношение между экспериментом и теоретической моделью двойное. С одной стороны, эксперимент позволяет проверить и при необходимости уточнить модель, т.е. эксперимент является источником информации для моделирования. С другой стороны, модель диктует, какой именно эксперимент следует проводить, т.е. модель является источником информации для организации эксперимента.

Современное понимание измерений стало шире классического, предусматривающего лишь количественные и однозначные измерения:
1) измерения могут но-

вопроса и перебора всех следствий, вытекающих из лежащей в основе его теоретической схемы, всех ответов, которые могла бы дать природа на выбранном экспериментатором теоретическом языке (...). Каков бы ни был ответ природы – “да” или “нет”, – он будет выражен на том же теоретическом языке, на котором был задан вопрос. Однако язык этот не остается неизменным, он претерпевает сложный процесс исторического развития, учитывающий прошлые ответы природы и отношения с другими теоретическими языками (...). Все это приводит к сложной взаимосвязи между (...) экспериментальным методом ведения диалога с природой (...) и культурной сетью, к которой, иногда неосознанно, принадлежит ученый (...). Сколько бы отрывочно ни говорила природа в отведенных ей экспериментом рамках, высказавшись однажды, она не берет своих слов назад: природа никогда не лжет” [5].

Д.С. Котари: “Простая истинна состоит в том, что ни измерение, ни эксперимент, ни наблюдение невозможны без соответствующей теоретической схемы” [5, с. 364].

Общая мысль этих высказываний ясна. Мы вернулись (правда, уже с другой стороны) к проблеме соотношения реальности и созданных нами ее моделей. Отличие от сказанного по этому поводу в § 2.6 состоит в том, что не только опыт является критерием истинности мо-

РОЛЬ ИЗМЕРЕНИЙ В СОЗДАНИИ МОДЕЛЕЙ СИСТЕМ

§ 6.1. ЭКСПЕРИМЕНТ И МОДЕЛЬ

В изначальном смысле отношение между экспериментом и моделью такое же, как и между курицей и яйцом: они находятся в одном цикле, и нельзя определить, что было “в самом начале”. Эксперимент с некоторым объектом проводится, чтобы уточнить модель этого объекта, поэтому постановка эксперимента определяется имеющейся до опыта моделью. Это полностью относится и к экспериментальному исследованию систем.

КЛАССИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ ОБ ЭКСПЕРИМЕНТЕ

Природу эксперимента хорошо понимали и понимают выдающиеся естествоиспытатели древности и современности. Приведем некоторые высказывания крупных ученых по этому поводу. Леонардо да Винчи: “...мне кажется, что пусты и полны заблуждений те науки, которые не порождены опытом, отцом всякой достоверности, и не завершаются в наглядном опыте (...). Опыт никогда не ошибается, ошибаются ваши суждения, ожидая от него такого действия, которое не является следствием ваших экспериментов (...). Мудрость есть дочь опыта” [4].

А. Розенблют и Н. Винер: “Любой эксперимент – всегда некий вопрос. Если вопрос неточен, получить точный ответ на него трудно. Глупые ответы, т.е. противоречивые, расходящиеся друг с другом или не относящиеся к делу иррелевантные результаты экспериментов, обычно указывают на то, что сам вопрос был поставлен глупо” [8].

И. Пригожин и И. Стенгерс [15]: “Природа, как на судебном заседании, подвергается с помощью экспериментирования перекрестному допросу именем априорных принципов. Ответы природы записываются с величайшей точностью, но их правильность оценивается в терминах той самой идеализации, которой физик руководствуется при постановке эксперимента (...). Экспериментальный метод есть искусство постановки интересного

ГЛАВА ШЕСТАЯ

дели, но и сама постановка эксперимента диктуется моделью, так как вытекает из необходимости ее проверки или уточнения.

Рассмотрим теперь возможности опытов (т.е. практического взаимодействия) с системами. Начнем обсуждение с модели “черного ящика” (см. § 3.3), т.е. с информации о входах и выходах системы. Выбор именно этих входов и выходов и есть построение модели, которая и будет определять организацию опыта. Если мы только регистрируем события на выбранных входах и выходах, то опыт называется *пассивным экспериментом* (или *наблюдением*). Если же мы не только созерцаем (и фиксируем) происходящее на входах и выходах, но и воздействуем на некоторые из них (одни намеренно поддерживая неизменными, другие изменения нужным образом), то опыт называется *активным* (или *управляемым*) экспериментом.

Результаты опыта регистрируются, фиксируются с помощью *измерений*, т.е. изображения результатов опыта в виде символов, номеров или чисел. Способы осуществления такого отображения будут рассмотрены в § 6.2. Важно, что современное понимание измерения существенно шире только количественного измерения. Не так уж давно Гейзенберг настаивал на идее, согласно которой не нужно говорить о том, что все равно нельзя измерить. Точку зрения физиков на эту идею поясняет Р. Фейнман:

“Дело в том, что об этом толкуют многие, по-настоящему не понимая смысла этого утверждения. Его можно интерпретировать следующим образом: ваши теоретические построения или открытия должны быть такими, чтобы выводы из них можно было сравнивать с результатами эксперимента, т.е. чтобы из них не получалось, что “один тук равняется трем нукам”, причем никто не знает, что такое эти самые тук и нук. Ясно, что так дело не пойдет. Но если теоретические результаты можно сравнить с экспериментом, то это все, что нам и требовалось. Это вовсе не значит, что ваши туки и нуки не могут появляться в первоначальной гипотезе. Вы можете впихнуть в вашу гипотезу сколько угодно хлама при условии, что ее следствия можно будет сравнить с результатами эксперимента. А это не всем до конца понятно” [10, с. 180, 181].

То, что Фейнман имел в виду именно количественное сравнение, видно из приводимого им примера:

“(…) Если ваша догадка сформулирована плохо или достаточно неопределенно и если метод, которым вы пользуетесь для оценки последствий, достаточно расплывчат (...), то ваша теория всем хороша – ведь ее нельзя опровергнуть. Кроме того, если ваш метод расчетов последствий достаточно нечеток, при некоторой ловкости всегда можно сделать так, чтобы результаты экспериментов были похожи на предполагаемые последствия. Возможно, вы знаете об этом по своему опыту в других областях. Некто ненавидит свою мать. Причина, конечно, в том, что она не заботилась о нем и не любила его достаточно, когда он был маленьким. Но если вы начнете раскапывать прошлое, то окажется, что на самом деле мать его очень любила и все у них было хорошо. Ну, тогда, ясно, она его слишком баловала! Как видите, расплывчатая теория позволяет получать любой результат. Поправить ее можно было бы следующим образом. Если бы вы смогли в точности и заранее определить, сколько любви недостаточно, а сколько чересчур много, то мы могли бы построить совершенно законную теорию, пригодную для экспериментальной проверки. Но стоит об этом заикнуться, как вам скажут: “Такие точные определения невозможны, когда речь идет о психологии”. Но раз так, то нельзя утверждать, что вы что-нибудь знаете” [10, с. 174].

Эти довольно пространные выдержки приведены не только как пример живого рассказа о непростых вещах, мастером которого является Фейнман, но и как иллюстрация того, насколько развилось понимание измерений за двадцать с лишним лет (книга [10] вышла в США в 1965 г.) Оставив незыблемым принцип проверки адекватности модели на опыте, современный подход позволил расширить понятие измерений по крайней мере в четырех отношениях.

СОВРЕМЕННОЕ ПОНЯТИЕ ЭКСПЕРИМЕНТА

1. Стало ясно, что есть наблюдаемые явления, в принципе не допускающие числовой меры (например, “количество материнской любви”), но которые можно фиксировать в “слабых”, “качественных” шкалах и эти результаты учитывать в моделях, получая качественные, но вполне научные выводы (см. § 6.2).

2. Расплывчатость некоторых наблюдений также признана их неотъемлемым природным свойством, которому придана строгая математическая форма, и разработан формальный аппарат “работы” с такими наблюдениями (см. § 6.3).

3. Хотя по-прежнему считается, что чем точнее измерения, тем лучше, теперь осознано, что погрешности измерений являются не только чем-то побочным, чуждым для измерений (сторонние помехи, результат небрежности или ошибок оператора и т.п.), но и неотъемлемым, естественным и неизбежным свойством самого процесса измерения (“шумы квантования”, соотношения неопределенности, собственные шумы аппаратуры). Проверяемые на практике модели должны быть не только гипотезами об исследуемом объекте, но и гипотезами об ошибках измерения.

4. Широкое распространение получили статистические измерения, т.е. оценивание функционалов распределений вероятностей по реализации случайного процесса; этот класс измерений важен потому, что большинство временных зависимостей входов и выходов носит сигнальный характер (см. гл. 5). Для таких измерений требуется специфическая методика и техника (см. § 6.4).

В данной главе будет затронута еще одна важная тема. Хотя для проведения эксперимента необходима модель объекта, с которой мы экспериментируем, а для уточнения модели объекта необходим эксперимент, здесь нет порочного круга: после завершения очередного цикла следующий начинается с новой, измененной модели. Ситуация напоминает не вращающееся колесо, а катящийся снежный ком: с каждым оборотом он становится больше, весомее. Вопрос о том, как именно происходит переход от модели “черного ящика” к модели “белого ящика” при использовании результатов измерений, мы обсудим в § 6.5.

Подведем итог

Summary

В современное понятие измерений включаются: пассивные наблюдения и активные эксперименты; количественные и качественные данные; точные, расплывчатые и зашумленные результаты опыта.

The current understanding of measurement includes the following: both passive observations and active experiments; either quantitative or qualitative data; exact and fuzzy and noisy experimental results.

§ 6.2. ИЗМЕРИТЕЛЬНЫЕ ШКАЛЫ

Измерение – это алгоритмическая операция, которая данному наблюдаемому состоянию объекта, процесса, явления ставит в соответствие определенное обозначение: число, номер или символ. Такое соответствие обеспечивает то, что результаты измерений содержат информацию о наблюдавшемся объекте, количество же информации зависит от степени полноты этого соответствия и разнообразия вариантов (см. § 5.7). Нужная нам информация получается из результатов измерения с помощью их преобразований, или, как еще говорят, с помощью обработки экспериментальных данных.

Совершенно ясно, что чем теснее соответствие между состояниями и их обозначениями, тем больше информации можно извлечь в результате обработки данных. Менее очевидно, что степень этого соответствия зависит не только от организации измерений (т.е. от экспериментатора), но и от природы исследуемого явления, и что сама степень соответствия в свою очередь определяет допустимые (и недопустимые) способы обработки данных.

В данном параграфе мы будем рассматривать только такие объекты, про любые два состояния которых можно сказать, различимы они или нет, и только такие алгоритмы измере-

ния, которые различным состояниям ставят в соответствие разные обозначения, а неразличимым состояниям – одинаковые обозначения*. Это означает, что как состояния объекта, так и их обозначения удовлетворяют следующим **аксиомам эквивалентности**:

- 1⁰. $A = A$ (рефлексивность).
- 2⁰. Если $A = B$, то $B = A$ (симметричность).
- 3⁰. Если $A = B$ и $B = C$, то $A = C$ (транзитивность).

Здесь символ $=$ обозначает отношение эквивалентности; в том случае, когда A и B – числа, он означает их равенство.

ШКАЛЫ НАИМЕНОВАНИЙ

Предположим, что число различимых состояний (математический термин – число классов эквивалентности) конечно. Каждому классу эквивалентности поставим в соответствие обозначение, отличное от обозначений других классов. Теперь измерение будет состоять в том, чтобы, проведя эксперимент над объектом, определить принадлежность результата к тому или иному классу эквивалентности и записать это с помощью символа, обозначающего данный класс. Такое измерение называется измерением *в шкале наименований* (иногда эту шкалу называют

также *номинальной* или *классификационной*); указанное множество символов и образует шкалу.

Особенности шкалы наименований рассмотрим на примерах. Естественнее всего использовать шкалу наименований в тех случаях, когда классифицируются дискретные по своей природе явления (например, различные объекты). Для обозначения классов могут быть использованы как слова естественного языка (например, географические названия, собственные имена людей и т.д.), произвольные символы (гербы и флаги государств, эмблемы родов войск, всевозможные значки и т.д.), номера (регистрационные номера автомобилей, официальных документов, номера на майках спортсменов), так и их различные комбинации (например, почтовые адреса, экслибрисы личных библиотек, печати и пр.) Все эти обозначения эквивалентны простой нумерации (в некоторых странах человек при рождении получает номер, под которым он фигурирует в государственных информационных системах всю жизнь), но на практике часто предпочитают другие обозначения (вообразите, что вместо имен и фамилий ваших друзей и знакомых вы должны будете использовать номера!).

Поскольку присваиваемое классу объектов обозначение в принципе произвольно (хотя после присвоения и однозначно), эту свободу в выборе можно использовать для удобства. Так, при большом и/или нефиксированном числе классов их конкретизация упрощается и облегчается, если обозначения вводятся иерархически. Примером могут служить почтовые адреса: страна – территориальная административная единица (республика, штат, кантон, графство, область) – населенный пункт – улица – дом – квартира – адресат. Другой пример – автомобильные номера: в их символике есть обозначение как территории, так и принадлежности машины (государственная или личная).

Необходимость классификации возникает и в тех случаях, когда классифицируемые состояния образуют непрерывное множество. Задача сводится к предыдущей, если все множество разбить на конечное число подмножеств, искусственно образуя тем самым классы эквивалентности. Теперь принадлежность состояния к какому-либо классу снова можно регистрировать в шкале наименований. Однако условность введенных классов (не их шкальных обозначений, а самих классов) рано или поздно проявится на практике. Например, возникают трудности точного перевода с одного языка на другой при описании цветовых оттенков: в английском языке голубой, лазоревый и синий цвета не различаются; не исключено, что англичане иначе видят мир (например, в одном английском толковом словаре слово “синий” объясняется как “цвет чистого неба, древесного дыма, снятого молока, свинца”, а в другом –

* В дальнейшем мы убедимся, что существуют не только такие типы измерений.

как “цвет неба или моря, а также вещей намного бледнее или темнее, как дым, удаленные холмы, лунный свет, синяк”).

Аналогичная ситуация имеет место в профессиональных языках. Вспомним примеры с наименованиями коров у африканского племени масаев, различных состояний снега у эвенков (см. § 2.3).

Названия болезней также образуют шкалу наименований. Психиатр, ставя больному диагноз “шизофрения”, “паранойя”, “маниакальная депрессия” или “психоневроз”, использует номинальную шкалу; и все же иногда врачи не зря вспоминают, что “нужно лечить больного, а не болезнь”: название болезни лишь обозначает класс, внутри которого на самом деле имеются различия, так как эквивалентность внутри класса носит условный характер.

Перейдем теперь к вопросу о допустимых операциях над данными, выраженными в номинальной шкале. Подчеркнем еще раз, что обозначения классов – это только символы, даже если для этого использованы номера. Номера лишь внешне выглядят как числа, но на самом деле числами не являются. Если у одного спортсмена на спине номер 4, а другого – 8, то никаких других выводов, кроме того, что это разные участники соревнований, делать нельзя: так, нельзя сказать, что второй “в два раза лучше” или что у одного из них форма новее. С номерами нельзя обращаться как с числами, за исключением определения их равенства или неравенства: только эти отношения определены между элементами номинальной шкалы (см. приведенные выше аксиомы 1⁰ – 3⁰).

Поэтому при обработке экспериментальных данных, зафиксированных в номинальной шкале, непосредственно с самими данными можно выполнять только операцию проверки их совпадения или несовпадения. Изобразим эту операцию с помощью символа Кронекера: $\delta_{ij} = \{1: x_i = x_j; 0: x_i \neq x_j\}$, где x_i и x_j – записи разных измерений.

С результатами этой операции можно выполнять более сложные преобразования: считать количества совпадений (например, число наблюдений k -го класса равно $n_k = \sum_{j=1}^n \delta_{kj}$, n – общее число наблюдений), вычислять относительные частоты классов (например, частота k -го класса есть $p_k = n_k/n$), сравнивать эти частоты между собой (находя, например, моду – номер наиболее часто встречающегося класса $k_{\max} = \arg \max_k p_k$), выполнять различные статистические процедуры, строго следя, однако, чтобы в этих процедурах с исходными данными не выполнялось ничего, кроме операции проверки их на совпадение (например, можно использовать χ^2 -тест, другие тесты на относительных частотах, коэффициент согласия и т.д.).

В тех случаях, когда наблюдаемый (измеряемый) признак состояния имеет природу, не только позволяющую отождествить состояния с одним из классов эквивалентности, но и дающую возможность в каком-то отношении сравнивать разные классы, для измерений можно выбрать более сильную шкалу, чем номинальная. Если же не воспользоваться этим, то мы откажемся от части полезной информации. Однако усиление измерительной шкалы зависит от того, какие именно отношения между классами существуют в действительности. Это и явилось причиной появления измерительных шкал разной силы.

ПОРЯДКОВЫЕ ШКАЛЫ

Следующей по силе за номинальной шкалой является *порядковая шкала* (используется также название *ранговая шкала*). Этот класс шкал появляется, если кроме аксиом тождества 1⁰ – 3⁰ классы удовлетворяют следующим **аксиомам упорядоченности**:

4⁰. Если $A \neq B$, то либо $A > B$, либо $B > A$.

5⁰. Если $A > B$ и $B > C$, то $A > C$.

Обозначив такие классы символами и установив между этими символами те же отношения порядка, мы получим *шкалу совершенного порядка*. Примерами применения такой шкалы являются нумерация очередности, воинские звания, призовые места в конкурсе.

Иногда оказывается, что не каждую пару классов можно упорядочить по предпочтению: некоторые пары считаются равными. В таком случае аксиомы 4⁰ и 5⁰ видоизменяются.

4'. Либо $A \leq B$, либо $A \geq B$.

5'. Если $A \geq B$ и $B \geq C$, то $A \geq C$.

Шкала, соответствующая аксиомам 4' и 5', называется *шкалой квазипорядка*. Примером шкалы квазипорядка служит упорядочение по степени родства с конкретным лицом (мать = отец > сын = дочь, дядя = тетя < брат = сестра и т.п.).

Иная ситуация возникает, когда имеются пары классов, не сравнимые между собой, т.е. ни $A \leq B$, ни $B \leq A$ (это отличается от условия квазипорядка, когда одновременно $A \geq B$ и

RANK, NUMBER,	$B \geq A$, т.е. $A = B$) . В таком случае говорят о <i>шкале частичного порядка</i> . Шкалы частичного порядка часто возникают в социологических исследованиях субъективных предпочтений. Например, при изучении покупательского спроса субъект часто не в состоянии оценить, какой именно из двух разнородных товаров ему больше нравится (например, клетчатые носки или фруктовые консервы, велосипед или магнитофон и т.д.); затрудняется человек и упорядочить по предпочтению любимые занятия (чтение литературы, плавание, вкусная еда, слушание музыки...).
POINT, SCORE	
балл	
MEASUREMENT	
измерение	
INTERVAL	
интервал	
RATIO	
отношение (чисел)	
NOMINAL	
SCALE	
шкала наименований	

Измерение – операция, ставящая наблюдаемому явлению в соответствие один из элементов подходящей измерительной шкалы.
Измерительная шкала может

Как видим, порядковые шкалы могут быть различными. В зависимости от того, каким аксиомам упорядоченности отвечают рассматриваемые объекты, мы должны пользоваться либо шкалой совершенного, либо шкалой частичного порядка. Однако разнообразие порядковых шкал этим не исчерпывается. Иногда число градаций в шкале задается заранее, и эксперимент лишь определяет, к какому из упорядоченных классов относится наблюдаемый объект (например, оценка на экзамене, сила землетрясения, воинское звание и т.п.) . В других случаях эталонные классы отсутствуют, а упорядочение проводится непосредственным попарным сравнением самих рассматриваемых объектов (например, выстраивание солдат в шеренгу по росту, определение мест в результате спортивных соревнований, музыкальных конкурсов и т. д.).

Очень важно обратить внимание на то, что отношение порядка ничего не говорит о “дистанциях” между сравниваемыми классами или объектами. Это придает порядковым шкалам характерную особенность: наблюдения, зафиксированные в таких шкалах, не являются числами. Даже если экспериментальные данные представлены цифрами (как школьные баллы, номера мест, занятых в соревновании, и т.п.), эти данные нельзя рассматривать

как числа. Над ними нельзя производить арифметические операции и вообще любые действия, результат которых изменится при преобразованиях шкалы, не нарушающих порядка. Например, нельзя вычислять среднее арифметическое порядковых измерений ($x_1, x_2, \dots, x_i, \dots, x_n$), т.е. величину $\frac{1}{n} \sum_i^n x_i$, так как переход к монотонно преобразованной шкале $x' = f(x)$ (сохраняющей данную упорядоченность) после усреднения даст величину $\frac{1}{n} \sum_{i=1}^n x'_i \neq \frac{1}{n} \sum_{i=1}^n x_i$. Между тем не все это знают; ярким примером такого широко распространенного заблуждения являлось

использование в школах и вузах в недавнем прошлом (а кое-где продолжающееся и сейчас) “средних баллов”. Правда, сразу было замечено, что средний балл, худо-бедно “работавший” в руках одного учителя, в рамках одного класса переставал играть роль объективного показателя при сравнении выпускников разных школ. Во всяком случае непродуманно введенный ранее учет средних школьных баллов при проведении конкурса для поступления в высшие учебные заведения был недавно отменен.

Допустимые операции над порядковыми наблюдениями вытекают из отношений, определяющих эти шкалы, т.е. из отношений эквивалентности и предпочтения. Допустимые операции представляют собой только операции проверки выполнимости этих отношений. Операция проверки принадлежности наблюдения к заданному классу (или неразличимости двух наблюдений) была уже введена выше, при рассмотрении номинальной шкалы, как символ Кронекера δ_{ij} , где один индекс – номер наблюдения, а другой – номер класса или другого наблюдения (в зависимости от типа порядковой шкалы). Операция проверки отношения предпочтения тоже может быть формализована. Введем индикаторную функцию C_{ij} предпочтения для упорядоченной* пары индексов (i, j) , а именно: $C_{ij} = \{ 1, \text{ если объект с индексом } i \text{ предпочтительнее объекта с индексом } j \text{ или эквивалентен ему; } 0, \text{ если верно обратное предпочтение } \}$. В результате по значению бинарной функции C_{ij} мы можем однозначно судить о порядке предъявленных объектов. Как и ранее, в зависимости от типа шкалы, один объект – данное наблюдение, а другой – либо некоторый класс, либо другое наблюдение.

Итак, непосредственно над порядковыми данными можно производить только операции по определению величин δ_{ij} и C_{ij} . Результаты этих операций являются двоичными числами; над ними уже можно производить арифметические и логические операции. Например, если i и j – номера наблюдений в совокупности данных $(x_1, \dots, x_i, \dots, x_n)$, то мы можем установить номер i -го наблюдения в упорядоченном ряду: $R_i = \sum_{j=1}^n C_{ij}$. Этот номер называется *рангом i -го объекта*; отсюда про-

исходит специальное название для данного типа порядковых шкал – *ранговые*. Если имеет место квазипорядок, то часть наблюдений может совпадать (в статистике такая группа наблюдений называется *связкой*), и все члены связки получают одинаковый (старший для них) ранг. Когда это неудобно, прибегают либо к присвоению ранга, среднего для данной связки (*мидранга*), либо присваивают ранги от младшего до старшего случайным образом.

С числами δ_{ij} и C_{ij} можно выполнять и другие необходимые операции. Кроме нахождения частот и мод (как и для номинальной шкалы), появляется возможность определить выборочную медиану (т.е. наблюдение с рангом R_i , ближайшим к числу $n/2$); можно разбить всю выборку на части в любой пропорции, находя выборочные квантили любого уровня p , $0 < p < 1$ (т.е. наблюдения с рангом R_i , ближайшим к величине pr); можно определить коэффициенты ранговой корреляции между двумя сериями порядковых наблюдений (r_s Спирмэна, ρ Кендальла); строить другие статистические процедуры.

Подчеркнем еще раз, что даже в тех случаях, когда состояния, которые допускают только порядковые сравнения, в эксперименте измеряются через величины, связанные с ними косвенно, но фиксируемые в числовых шкалах, эти измерения все равно остаются измерениями в порядковой шкале. Пфанцагль [6] приводит наглядные примеры, иллюстрирующие сказанное.

Первый пример взят из медицины. Известно, что за показатель интенсивности патологического процесса принимается скорость выпадения осадка при добавлении в пробирку с кровью цитрата натрия; скорость осаждения измеряется в миллиметрах в единицу времени.

* Напомним, что упорядоченность пары объектов означает, что их нельзя менять местами; точнее, что пара с обратным порядком – это другая пара.

Идея такого измерения основана на том, что увеличение интенсивности патологического процесса приводит к повышению содержания глобулина, что увеличивает скорость выпадения осадка. Поэтому высота слоя осадка за данный интервал времени монотонно связана с интенсивностью исследуемого патологического процесса. Функциональный вид этой связи неизвестен, для разных лиц различен и нелинеен: изменение количества цитрата натрия или времени осаждения приводят к непропорциональным изменениям высоты осадка. Теперь предположим, что для одного больного лекарство *A* привело к уменьшению осадка с 75 до 60 мм, а для другого лекарство *B* – с 65 до 55 мм. Отсюда нельзя заключать, что лекарство *A* эффективнее, так как оно привело к уменьшению осадка на 15 мм, а лекарство *B* – только на 10!

В качестве второго примера рассматривается испытание умственных способностей, при котором измеряется время, затрачиваемое испытуемым на решение тестовой задачи. В таких экспериментах время хотя и измеряется в числовой шкале, но как мера интеллекта принадлежит порядковой шкале.

Выше мы не без умысла к названию порядковой шкалы присоединяли слова “в строгом смысле”. Суть состоит в том, что порядковые в строгом смысле шкалы определяются только для заданного набора сравниваемых объектов, у этих шкал нет общепринятого, а тем более абсолютного стандарта. Поэтому при определенных условиях правомерно выражение “первый в мире, второй в Европе” – просто чемпион мира занял второе место на всеевропейских соревнованиях.

МОДИФИЦИРОВАННЫЕ ПОРЯДКОВЫЕ ШКАЛЫ

По-видимому, опыт работы с сильными числовыми шкалами и желание уменьшить относительность порядковых шкал, придать им хотя бы внешнюю независимость от измеряемых величин побуждают исследователей к различным модификациям, придающим порядковым шкалам некоторое (чаще всего кажущееся) усиление. Другая важная причина попыток усиления шкалы состоит в том, что многие измеряемые в порядковых (принципиально дискретных) шкалах величины имеют действительный или мыслимый непрерывный характер: сила ветра или землетрясения, твердость вещества, глубина и прочность знаний, овладение навыками и т.п. Сама возможность введения между любыми двумя шкальными значениями третьего способствует тому, чтобы попытаться усилить шкалу.

Все это вместе взятое привело к появлению и использованию на практике ряда порядковых шкал, но не в таком “строгом смысле”, как те, о которых мы говорили выше. При этом иногда с полученными данными начинают обращаться как с числами, даже если произведенная модификация не выводит шкалу из класса порядковых. Это сопряжено с ошибками и неправильными решениями. Рассмотрим некоторые из известных модификаций.

Шкала твердости по Моосу. Из двух минералов тверже тот, который оставляет на другом царапины или вмятины при достаточно сильном соприкосновении. Отношение “*A* тверже *B*” – типичное отношение порядка. В 1811 г. немецкий минералог Ф. Моос предложил установить стандартную шкалу твердости, постулируя только десять ее градаций. За эталоны приняты следующие минералы с возрастающей твердостью: 1 – тальк, 2 – гипс, 3 – кальций, 4 – флюорит, 5 – апатит, 6 – ортоклаз, 7 – кварц, 8 – топаз, 9 – корунд, 10 – алмаз. Шкала Мооса устанавливает искусственно квазипорядок, так как промежуточных единиц градаций твердости эта шкала не имеет. Градации твердости все равно не носят числового характера: нельзя говорить ни что алмаз в два раза тверже апатита, ни что разница в твердостях флюорита и гипса такая же, как у корунда и кварца; измерения твердости методом царапания не дают оснований для оправдания таких утверждений.

Шкала силы ветра по Бофорту. В 1806 г. английский гидрограф и картограф адмирал Ф. Бофорт предложил балльную шкалу силы ветра, определяя ее по характеру волнения моря: 0 – штиль (безветрие), 4 – умеренный ветер, 6 – сильный ветер, 10 – штурм (буря), 12 – ураган. Кроме штиля, градации силы ветра имеют условный, качественный характер.

Шкала магнитуд землетрясений по Рихтеру. В 1935 г. американский сейсмолог Ч. Рихтер предложил 12-балльную шкалу для оценки энергии сейсмических волн в зависимости от последствий прохождения их по данной территории. Затем он развел метод оценки силы землетрясения в эпицентре по его магнитуде на поверхности земли и глубине очага.

Балльные шкалы оценки знаний учащихся. Слушая ответы учащихся или сравнивая их письменные работы, опытный преподаватель может обнаружить разницу между ними и установить, чьи ответы лучше; это типичное отношение порядка. Методом сравнения можно определить, кто в классе лучше других знает данный предмет; сложнее, но иногда возможно (это зависит от состава класса) определить лучшего ученика в классе. Сравнение старшеклассника с младшеклассником по степени овладения знаниями проблематично.

Потребность общества в официальном определении степени квалифицированности проходящих обучение независимо от того, где, когда и как они получают образование, способствовала введению общепринятых шкал для оценивания знаний учащихся в виде баллов (такие шкалы введены повсеместно). Все испытывают, в том числе и на собственном опыте, неточность, приблизительность этой шкалы. Одна из попыток “улучшить” шкалу баллов состоит в увеличении числа градаций. В наших школах принята 5-балльная, в вузах – 2-балльная (для зачетов) и 4-балльная (для экзаменов) системы оценок, в некоторых европейских странах – 10-балльная, а в англоязычных странах – 100-балльная система. Это не спасет положения, и преподаватели неофициально (“для себя”) вводят дополнительные градации – присоединяют к баллам плюсы, минусы, точки. Примечательно, что и при 100-балльной шкале некоторые преподаватели используют дробные баллы. Все это происходит потому, что не существует ни абсолютного стандарта, единого для всех людей, ни даже условного общедоступного стандарта, наподобие эталонов твердости или высоты волн, и знания могут оцениваться только в порядковой шкале. Тем не менее мало кто (не только учащиеся, но и преподаватели) понимает, что балльная шкала принадлежит к классу порядковых. Дело доходит до того, что даже в официальных вопросах, влияющих на судьбы людей, учитывают среднеарифметический балл – величину, не имеющую смысла в порядковой шкале! Некоторый оттенок объективности и количественности балльной шкале пытаются придать директивным определением того, каким требованиям должен удовлетворять учащийся, чтобы иметь право на тот или иной балл, т.е. ввести независимые стандарты. Однако преподаватели неизбежно по-разному понимают и выполняют инструкции, и оценки все равно получаются относительными: известно, что уровень знаний отличников разных школ или вузов заметно различается. Именно поэтому в ответственных случаях устраивают не конкурсы документов об успеваемости, а конкурсы самих претендентов, т.е. возвращаются к порядковому измерению, непосредственному сравнению обладателей знаний.

ШКАЛЫ ИНТЕРВАЛОВ

Если упорядочивание объектов можно выполнить настолько точно, что известны расстояния между любыми двумя из них, то измерение окажется заметно сильнее, чем в шкале порядка. Естественно выражать все расстояния в единицах, хотя и произвольных, но одинаковых по всей длине шкалы. Это означает, что объективно равные интервалы измеряются одинаковыми по длине отрезками шкалы, где бы они на ней ни располагались. Следствием такой равномерности шкал этого класса является независимость отношения двух интервалов от того, в какой из шкал эти интервалы измерены (т.е. какова единица длины интервала и какое значение принято за начало отсчета).

Сказанное можно выразить вполне формализованно. Пусть M – множество совершенно упорядоченных элементов, для каждой пары c, d которых задано вещественное число $?(c, d)$, удовлетворяющее следующим условиям:

- (1) если $c < d$, то $?(c, d) > 0$;
- (2) если $c \in M$ и r – вещественное число, то найдутся такие $d, e \in M$, что $?(c, d) = r$, $?(c, e) = -r$;
- (3) для любых $(c, d, e) \in M$ верно равенство $?(c, d) + ?(d, e) = ?(c, e)$.

Множество M с таким бинарным отношением назовем *интервальной шкалой*. В шкале интервалов можно ввести систему координат. Выберем для этого любую пару точек (“репер”)

ORDER, ORDINAL,	$c, d \in M$; точка c играет роль начала координат, а
ORDER-TYPE	интервал (c, d) – роль единичного интервала. Каждой
порядковый	точке $e \in M$ поставим в соответствие координату $x_e = ?(c, e)/?(c, d)$. Тогда точка c будет иметь координату 0, а точка d – координату 1.
TIE	
связка (наблюде-	
ний)	
WEAK, FEEBLE	
слабый	
PARTIAL	
частичный	

Если ввести в M другую систему координат, построенную на репере $c_1 n d_1$, то координаты x_e и x_{e_1} точки e в этих двух системах координат будут связаны линейным соотношением $x_e = x_{e_1} = ax_{e_1} + b$. Действительно,

$$x_e = \frac{\rho_{e,g}}{\rho_{d,g}} f, \quad x_{e_1} = \frac{\rho_{e,g}}{\rho_{d_1,g}} c_1$$

Следовательно,

$$\begin{aligned} x_e &= \frac{\rho_{e,g} \rho_{g,b} g}{\rho_{d,g} \rho_{d,g}} = \frac{\rho_{e,d_1} \rho_{d_1,c_1} \rho_{c_1,g} \rho_{g,e}}{\rho_{d,d_1} \rho_{d_1,d_1} g} = \frac{\rho_{e,d_1} g}{\rho_{d,d_1} g} + \\ &+ \frac{\rho_{e,c_1} \rho_{c_1,g}}{\rho_{d,d_1} g} x_{e_1} + \frac{\rho_{e,c_1} \rho_{c_1,b} b}{\rho_{d,d_1} g} = ax_{e_1} + b, \end{aligned}$$

где a и b – очевидные обозначения. Несмотря на то, что координата x_e и разности $x_e - x_f$ меняются при смене репера, для любых $e, f, g, h \in M$ отношение интервалов $\frac{x_e - x_f}{x_g - x_h}$ не зависит от выбора репера.

Итак, интервальные шкалы могут иметь произвольные начала отсчета и единицы измерения, что можно выразить словами: “шкала интервалов единственна с точностью до линейных преобразований”.

Примерами величин, которые по физической природе либо не имеют абсолютного нуля, либо допускают свободу выбора в установлении начала отсчета и поэтому измеряются в интервальных шкалах, являются температура, время, высота местности.

Начало летоисчисления у христиан установлено от рождения Христова, а у мусульман – на 622 г. позднее – от переезда Мухаммеда в Медину; единицы летоисчисления привязаны к относительным перемещениям Солнца и Луны, но в астрономии существует целых шесть различных определений года. Высоту принято отсчитывать от уровня моря, но это привело к тому, что большая часть территории Голландии имеет... отрицательную высоту, так как расположена ниже уровня моря.

Несмотря на произвольность начала отсчета, в обыденной жизни координаты интервальной шкалы иногда абсолютизируются (вспомните, как много эмоций и реальных событий мы связываем с Новым годом, началом нового века, и т.д.).

Данные, зафиксированные в протоколе эксперимента принадлежат определенной измерительной

412

Название “шкала интервалов” подчеркивает, что *в этой шкале только интервалы имеют смысл настоящих чисел* и только над интервалами следует выполнять арифметические операции: если произвести арифметические операции над самими отсчетами по шкале, забыв об их относительности, то имеется риск получить бессмысленные результаты. Например, если сказать, что температура воды увеличилась в два раза при ее нагреве от 9 до 18° по шкале Цельсия, то для тех, кто привык пользоваться шкалой Фаренгейта, это будет звучать весьма странно*, так как в этой шкале температура воды в том же опыте изменится от 48,2 до 64,4.

Подобно тому как определение значения символа Кронекера является единственной допустимой операцией над наблюдениями в номинальной шкале, а вычисление ранга наблюдения – в порядковой шкале, в интервальной шкале единственной новой допустимой операцией над наблюдениями является определение интервала между ними. Над интервалами же можно выполнять любые арифметические операции, а вместе с ними – использовать подходящие способы статистической и иной обработки данных. Например, центральные моменты (в том числе дисперсия) имеют объективный физический смысл, а начальные моменты (в том числе среднее значение) являются относительными наряду с началом отсчета. Поэтому понятие относительной погрешности (коэффициента вариации, т.е. отношения стандартного отклонения к математическому ожиданию) не имеет смысла для интервальной шкалы. Это не означает, что вообще нельзя суммировать показания в шкале интервалов, например вычислять выборочное среднее $\bar{x} = \frac{1}{n} \sum x_i$. Однако с такой величиной нужно обращаться так же, как и с другими исходными наблюдениями, – она остается интервальной величиной и приобретает числовой смысл только в процессе определения интервалов. Поэтому выборочная дисперсия имеет объективный смысл, хотя и определяется через \bar{x} по формуле $S^2 = \overline{\overline{b}} - \bar{x} \overline{c}$; дело в том, что $X - \bar{x}$ является интервалом.

ШКАЛЫ ОТНОШЕНИЙ

Пусть наблюдаемые величины удовлетворяют не только аксиомам 4⁰ и 5⁰, но и **аксиомам аддитивности**:

6⁰. Если $A = P$ и $B > 0$, то $A + B > P$.

7⁰. $A + B = B + A$.

8⁰. Если $A = P$ и $B = Q$, то $A + B = P + Q$.

9⁰. $(A + B) + C = A + (B + C)$.

Это существенное усиление шкалы: измерения в такой шкале являются “полноправными” числами, с ними можно выполнять любые арифметические действия, так как вычитание, умножение и деление – лишь частные случаи сложения. Введенная таким образом шкала называется *шкалой отношений*. Этот класс шкал обладает следующей особенностью: отношение двух наблюдаемых значений измеряемой величины не зависит от того, в какой из таких шкал произведены измерения: $x_1/x_2 = y_1/y_2$. Этому требованию удовлетворяет соотношение вида $y = ax$ ($a \neq 0$). Таким образом, величины, измеряемые в шкале отношений, имеют естественный, абсолютный нуль, хотя остается свобода в выборе единиц. В самом деле, при наличии абсолютного нуля свобода в выборе начала отсчета исчезает, и в формуле связи между разными системами координат, выведенной для интервальных шкал, второй член равен нулю (так как $?(\mathbf{c}, \mathbf{c}_1) = ?(\mathbf{c}, \mathbf{c}) = 0$), откуда и следует $y = ax$.

Примерами величин, природа которых соответствует шкале отношений, являются длина, вес, электрическое сопротивление, деньги.

ШКАЛЫ РАЗНОСТЕЙ

К числу шкал, единственных с точностью до линейных преобразований, относятся шкала интервалов ($y = ax + b$, $a > 0$ и b произвольны) и шкала отношений ($y = ax$, $a > 0$ – преобразование растяжения). Рассмотрим особенности шкал, инвариантных к сдвигу: $y = x + b$.

* Связь между шкалами Фаренгейта и Цельсия выражается формулой $F = -\frac{9}{5}C + 32$.

Повторно применяя сдвиг к $y(z = y + b = x + 2b)$, затем к z и т.д., обнаруживаем, что в такой шкале значение не изменяется при любом числе сдвигов: $y = x + nb$, $n = 0, 1, 2, \dots$. Постоянная b является характерным параметром шкалы и называется ее *периодом*. Полученную шкалу будем называть *шкалой разностей* (иногда ее также называют *циклической* или *периодической*).

В таких шкалах измеряется направление из одной точки (шкала компаса, роза ветров и т.д.), время суток (циферблат часов), фаза колебаний (в градусах или радианах).

Циклические шкалы являются частным случаем интервальных шкал. Однако соглашение о хотя и произвольном, но едином для нас начале отсчета шкалы позволяет использовать показания в этой шкале как числа, применять к ним арифметические действия (до тех пор, пока кто-нибудь не забудет об условности нуля, например при переходе на летнее время или обратно, пересечении линии смены дат и т.д.).

АБСОЛЮТНАЯ ШКАЛА

Рассмотрим такую шкалу, которая имеет и абсолютный нуль, и абсолютную единицу. Эта шкала не единственна с точностью до какого-либо преобразования, а просто единственна, уникальна. Именно такими качествами обладает числовая ось, которую естественно назвать *абсолютной шкалой*. Важной особенностью абсолютной шкалы по сравнению со всеми остальными является отвлеченность (безразмерность) и абсолютность ее единицы. Указанная особенность позволяет производить над показаниями абсолютной шкалы такие операции, которые недопустимы для показаний других шкал, – употреблять эти показания в качестве показателя степени и аргумента логарифма. Числовая ось используется как измерительная шкала в явной форме при счете предметов, а как вспомогательное средство присутствует во всех остальных шкалах. Внутренние свойства числовой оси, при всей кажущейся ее простоте, оказываются чрезвычайно разнообразными, и теория чисел до сих пор не исчерпала их до конца. А некоторые безразмерные числовые отношения, обнаруживаемые в природе, вызывают восхищение и изумление (явления резонанса; гармонические отношения размеров, звуков; законы теории подобия и размерности; квантование энергии элементарных частиц и т.п.).

СОГЛАСОВАНИЕ ШКАЛЫ С ПРИРОДОЙ НАБЛЮДЕНИЙ

В табл. 6.1 приведены основные сведения о всех рассмотренных в данном параграфе измерительных шкалах. Можно сказать, что чем сильнее шкала, в которой производятся измерения, тем больше сведений об изучаемом объекте, явлении, процессе дают измерения. Поэтому так естественно стремление каждого исследователя провести измерения в возможно более сильной шкале. Однако важно иметь в виду, что выбор шкалы измерения должен ориентироваться на объективные отношения, которым подчинена наблюдаемая величина, и лучше всего производить измерения в той шкале, которая максимально согласована с этими отношениями. Можно измерять и в шкале, более слабой, чем согласованная (это приведет к потере части полезной информации), но применять более сильную шкалу опасно: полученные данные на самом деле не будут иметь той силы, на которую ориентируется их обработка.

Аналогичная ситуация имеет место и после того, как проведены измерения. У исследователя могут быть причины, побуждающие его преобразовать протокол наблюдений, переведя их из одной шкалы в другую. Если при этом данные переводятся в более слабую шкалу, то обычно исследователь отдает себе отчет в том, что в результате происходит некоторое ухудшение качества выводов. Иногда же исследователи усиливают шкалы; типичный случай – “оцифровка” качественных шкал: классам в номинальной или порядковой шкале присваиваются номера, с которыми дальше “работают” как с числами. Если в этой обработке не выходят за пределы допустимых преобразований, то “оцифровка” – это просто перекодировка в более удобную (например, для ЭВМ) форму. Однако применение других операций сопряже-

но с заблуждениями и ошибками, так как свойства, навязываемые подобным образом, на самом деле не имеют места.

О ДРУГИХ ШКАЛАХ

Обширный опыт наблюдений в разнообразных областях науки и практики нередко приводил к целесообразности использования шкал, отличающихся от рассмотренных выше. Обсудим наиболее важные и интересные из них.

Т а б л и ц а 6.1. Итоговая таблица измерительных шкал

Назва- ние школь- ных отношений	Определяю- щие отношения	Эквивалент- ное преоб- разование школ	Допустимые операции над данными (пер- вичная обработка)	Вторичная обработка
Номи- наль- ная	Эквивалент- ность	Перестанов- ки наимено- ваний	Вычисление символа Кро- некера $?_{ij}$	Вычисление относитель- ных частот и операций над ними
Поряд- ковая	Эквивалент- ность; пред- почтение	Не изменя- ющее по- рядка (мо- нотонное)	Вычисление $?_{ij}$ и рангов R_i	Вычисление относитель- ных частот и выбороч- ных кванти- лей, опера- ции над ни- ми
Интер- вальная	Эквивалент- ность предпочте- ние; сохранение отношения интервалов	Линейное преобразо- вание $y = ax + b$, $a > 0$, $b \in R$	Вычисление $?_{ij}$, рангов R_i и ин- тервалов (раз- ностей между наблюдениями)	Арифмети- ческие дей- ствия над интервала- ми
Цикли- ческая	Эквивалент- ность; пред- почтение; сохранение отношения интервалов; периодич- ность	Сдвиг $y = x + nb$, $b = \text{const}$, $n = 0, 1, 2, \dots$	То же, что и для интервал- ной шкалы	То же, что и для интер- вальной шкалы
Отно- шений	Эквивалент- ность; пред- почтение; сохранение отношения интервалов; сохранение отношения двух значе- ний	Растяжение $y = ax$, $a > 0$	Все арифмети- ческие opera- ции	Любая под- ходящая обработка

Абсо- лютная ноль; пред- почтение; сохранение отношения интервалов; сохранение отношения двух значе- ний; абсолю- тная и безразмерная единица; абсолю- тный нуль	Эквивалент- ность; пред- почтение; сохранение отношения отношения двух значе- ний; абсолю- тная и безразмерная единица; абсолю- тный нуль	Шкала уникальна	Все арифмети- ческие опера- ции; использо- вание в каче- стве показателя степени, осно- вания и аргу- мента лога- рифма	Любая не- обходимая обработка
--	--	--------------------	---	-------------------------------------

Очень распространены измерения непрерывных величин, возможные значения которых образуют континуум. По ряду причин результат наблюдения такой величины всегда фиксируется с “округлением”, с конечной точностью, т.е. так, как будто наблюдаемая величина дискретна. Иногда эта точность связана лишь с выбором числа разрядов в записи наблюдения, и ее можно увеличить, просто наращивая число значащих цифр (что часто делается в компьютерных расчетах). Однако в научных и технических измерениях эта точность ограничивается не тем, на сколько еще частей можно разделить каждое деление шкалы, а классом точности самого прибора.

В связи с этим следует различать шкалы, в которых измеряются величины, дискретные по своей природе (например, измерение энергии с точностью до квантов в физике элементарных частиц, подсчет поголовья скота, количества деревьев на данной площади и т.п.), и шкалы, в которых измеряются с конечной точностью непрерывные величины (масса, длина, напряжение, время и т. д.). Первые будем называть *дискретными* шкалами, вторые – *дискретизованными*. Обращение с данными в дискретизованной шкале имеет свои особенности. Важно, что в таких измерениях нельзя фиксировать дробные части деления шкалы, даже если стрелка прибора остановилась между метками: класс точности прибора не гарантирует различия внутри интервала точности. Казалось бы, последующая статистическая обработка совокупности неточных наблюдений может дать повышение точности за счет усреднения. Однако это верно не во всех случаях: все зависит от того, какой параметр распределения данных мы оцениваем, и не изменяет ли погрешность этот параметр. Другими словами, вопрос о влиянии дискредитации на конечный результат не является тривиальным.

Известны случаи, когда “ухудшение” измерения в дискретизованной шкале на самом деле улучшает результат. Например, при определении координат звезд при проектировании изображения участка неба на светочувствительную матрицу, рекомендуется расфокусировать изображение так, чтобы световое пятно от звезды охватывало несколько пикселов (ячеек) матрицы. Это увеличивает точность определения положения звезды. Другой интересный эффект дискретизации заключается в том, что при искажении измеряемого параметра распределения наблюдений существует некоторое число усредняемых наблюдений, превышение которого может лишь ухудшить точность конечного результата.

Еще один практически важный класс шкал – *нелинейные*. Так называются шкалы, в которых интервалы не удовлетворяют условиям аддитивности. Иначе говоря, “цена” единичного деления такой шкалы зависит от того, в какой части шкалы находится это деление. Примерами могут служить квадратичная, логарифмическая, экспоненциальная шкалы, “вероятностная бумага”, многие номограммы. Причины введения нелинейных шкал могут быть как объективными (например, нелинейность измерительного прибора, большой динамический диапазон измеряемой величины), так и субъективными (удобство и наглядность представления данных, желание подчеркнуть некоторые детали полученной зависимости и т.д.). Обработка данных, зафиксированных в таких шкалах, требует учета конкретного характера нелинейности. Следует иметь в виду и то, что некоторые из нелинейных преобразований могут ненамеренно изменить силу шкалы. Например, в акустике и радиотехнике часто отношение мощностей сигналов представляется в децибелах: $N [\text{дБ}] = 10 \lg(P_2/P_1)$. Мощности P_1 и P_2 измеряются в шкале отношений, следовательно, все операции, необходимые для получения количества децибел, допустимы. Однако величина N принадлежит шкале интервалов, что должно учитываться при дальнейшем оперировании с нею (например, нельзя говорить, что мощность данного сигнала равна такому-то количеству децибел и не указать, в сравнении с чем).

Обратим внимание на еще одну особенность типовых шкал. Бросается в глаза резкое, принципиальное отличие между “слабыми” качественными шкалами – номинальной и порядковой – и “сильными” количественными шкалами – интервалов, разностей, отношений, абсолютной. Кажущееся очевидным различие между “качеством” и “количеством”, несводи-

мость одного к другому, применительно к измерительным шкалам вдруг теряет очевидность. Во всяком случае возникает интригующий вопрос: является ли переход от качественных шкал к количественным принципиально скачкообразным, или существует возможность путем пошагового усиления порядковой шкалы “плавно” дойти до интервальной? Возможность постепенного усиления открывается в связи с введением, наряду с упорядочением альтернатив, понятия *силы предпочтения*. Речь идет о возможности сравнения “расстояний” между упорядоченными альтернативами. Ясно, что как только эти расстояния начнут изменяться в числовых шкалах, так и шкала в целом станет числовой (напомним, что формальное определение интервальной шкалы вводилось именно через определение интервала). Отказ от сравнения сил предпочтения оставляет нас в рамках порядковой шкалы. Существует и промежуточный вариант: сравнивать различия между альтернативами в порядковой шкале. Это – явное усиление шкалы, не переводящее ее, однако, в разряд количественной. Такую шкалу называют *шкалой гиперупорядочения*. Дальнейшее усиление состоит в том, чтобы упорядочивать силы предпочтения сил предпочтения и т.д. Сходится ли такая последовательность порядковых шкал хотя бы в пределе к числовой шкале, и при каких требованиях к нечисловому характеру сил предпочтения – остается пока математически открытым вопросом.

Как пример нетривиального усиления порядковой шкалы при числовом характере сил предпочтения, изложим идею шкалы Черчмена и Акоффа на частном примере.

Пусть имеется четыре предмета. Сначала опрашиваемый упорядочивает их в порядке предпочтения: $A \geq B \geq C \geq D$. Затем его просят поставить в соответствие (приписать) предметам любые числа между нулем и единицей, выразив грубо “силу” предпочтения. Пусть результат таков:

$$\begin{array}{cccc} A & B & C & D \\ \hline 1,00 & 0,85 & 0,75 & 0,20 \end{array} .$$

Целью является уточнение с помощью дальнейших вопросов действительной силы предпочтений опрашиваемого. Например, что он предпочитает, A или B , C и D вместе взятые. Результат необходимо как-то отразить в весовых коэффициентах. Делается предположение, что весовой коэффициент совокупности альтернатив равен сумме их весовых коэффициентов. Если, например, $A > (B \cap C \cap D)$, приписываются новые коэффициенты:

$$\begin{array}{cccc} A & B & C & D \\ \hline 1,00 & 0,65 & 0,20 & 0,10 \end{array} .$$

Далее спрашивают, как упорядочиваются B и $(C \cap D)$. Если, по мнению опрашиваемого, $C \cap D > B$, то уменьшают вес B так, чтобы он был меньше суммы весов C и D :

$$\begin{array}{cccc} A & B & C & D \\ \hline 1,00 & 0,25 & 0,20 & 0,10 \end{array} .$$

Другие начальные веса при указанных вопросах и ответах могут остаться неизменными, если они сразу отвечали указанным требованиям. Например:

$$\begin{array}{cccc} A & B & C & D \\ \hline 1,00 & 0,33 & 0,33 & 0,33 \end{array} \text{ и } \begin{array}{cccc} A & B & C & D \\ \hline 1,00 & 0,04 & 0,03 & 0,03 \end{array} .$$

и

Чтобы сократить перебор комбинаций при уточнении шкалы, авторы метода предлагают наиболее предпочтительной альтернативе приписывать единичный вес, а остальные группировать по три и действовать по указанной методике. Если и при этом количество перебираемых комбинаций окажется большим (что неизбежно при большом числе упорядочиваемых объектов), то можно прибегнуть к неполному перебору, применив случайный механизм выбора троек и установив критерий прекращения пересчета весов.

Основным предметом критики порядковой шкалы Черчмена и Акоффа является тот факт, что предположение об аддитивности весов предпочтения в психологии нередко не выполняется: скажем, опрашиваемый может оценивать смесь меда с дегтем иначе, чем суммой весов

меда и дегтя в отдельности; то же может относиться и к оценке хлеба с маслом и хлеба и масла в отдельности.

Интересно отметить, что несмотря на задание сил предпочтения в шкале отношений (так как к ним применяются операции сложения и деления), результирующая шкала определяется только приблизительно, остается оценочной. Это прежде всего связано с тем, что опрашиваемый не в состоянии точно описать силы своих предпочтений, и задаваемые им числа являются лишь интуитивными приближениями.

Подведем итог

Summary

Основные результаты данного параграфа сведены в табл. 6.1, которая отражает главные особенности каждой измерительной шкалы.

The gist of this section is in Table 6.1, which summarized all the important traits of measuring scales of all possible types.

§ 6.3. РАСПЛЫВЧАТОЕ ОПИСАНИЕ СИТУАЦИЙ

Обратим внимание на то, что все измерительные шкалы, рассмотренные в § 6.2, имеют одно общее свойство: они основаны на справедливости отношения эквивалентности (см. табл. 6.1). Это отношение имеет силу как отдельно на множестве состояний наблюдаемого объекта и множестве наблюдений, зафиксированных в любой из шкал (два состояния или два измерения либо тождественны, либо различны), так и на их совокупности (состояния и соответствующие им измерения находятся во взаимно однозначном соответствии). Использование рассогласованной (т.е. более слабой, чем можно) шкалы приводит к образованию на множестве состояний новых классов эквивалентности, внутри которых состояния неразличимы в данной шкале (хотя их и можно различить в более сильной шкале). Однако и в этом случае отношение эквивалентности соблюдается.

ПОНЯТИЕ РАСПЛЫВЧАТОСТИ

В действительности встречаются (и гораздо чаще, чем кажется) случаи, когда тождество или различие двух состояний и/или наблюдений нельзя утверждать с полной уверенностью. Наиболее явно это видно на примере шкал, в которых классы обозначаются конструкциями естественного языка. “В комнату вошел высокий молодой человек” – класс, к которому принадлежит человек, назван (т.е. измерение состоялось), но какого он роста и сколько ему лет? “В руках он держал довольно тяжелый сверток” – какого веса была его ноша? Если разобраться, то почти каждое наше слово обозначает некоторое не вполне определенное множество. (“Почти” – какой процент? “Наше” – чье именно? “Некоторое” – какое же? “Не вполне” – насколько? “Определенное” – кем и как? и т.д.) Это свойство естественного языка, природное и неотъемлемое, безусловно, полезное (иначе бы оно не закрепилось в процессе развития языка), но приводящее к затруднениям, когда сопровождающая его неопределенность мешает. Древние логики дискутировали вопрос о том, сколько песчинок должно быть собрано вместе, чтобы получилась куча песка; сегодня мы просто говорим, что слово “куча” – это лишь метка нечетко определенного множества. Спор о том, сколько песчинок в “куче”, эквивалентен спору о том, в каком возрасте человек становится “старым” или сколько волосинок должно у него выпасть, чтобы он был “лысым”.

Эта неопределенность смысла языковых конструкций является одной из основных трудностей автоматизации анализа и синтеза речи, автоматического (и не только автоматического) перевода с одного языка на другой. Например, одному английскому предложению, состоящему из пяти слов, можно дать пять разных (!) смысловых интерпретаций [7]:

TIME FLIES LIKE AN ARROW
ВРЕМЯ ЛЕТИТ СТРЕЛОЙ
ВРЕМЯ ЛЕТИТ В НАПРАВЛЕНИИ СТРЕЛЫ

МУХАМ ВРЕМЕНИ НРАВИТСЯ СТРЕЛА*
ИЗМЕРЯЙ СКОРОСТЬ МУХ ТАК ЖЕ, КАК СКОРОСТЬ СТРЕЛЫ**
ИЗМЕРЯЙ СКОРОСТЬ МУХ, ПОХОЖИХ НА СТРЕЛУ

Неизвестно, действительный ли это факт или научно-фольклорная история, основанная на потенциальной возможности, но в литературе по автоматизации перевода приводится рассказ о кольцевой работе программ, переводящих с одного языка на другой: фраза “плоть слаба, а дух силен” после нескольких переводов превратилась в “мясо тухлое, но водка крепкая”.

Все сказанное выше мотивирует введение понятия **лингвистической переменной как переменной, значение которой расплывчано по своей природе, как метки размытого, расплывчатого множества***. Хотя теория размытых множеств, построенная Л. Задэ, прекрасно иллюстрируется языковыми примерами и имеет интересные приложения в области искусственного интеллекта, размытость оказывается свойством не только естественного языка. Например, в математике с успехом применяются понятия “значительно больше” (символ \gg) и “приблизительно равно” (символ \approx или $\&$), являющиеся типично расплывчатыми.

ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ РАСПЛЫВЧАТЫХ МНОЖЕСТВ

Изложим основные понятия теории расплывчатых множеств [1]. Расплывчатое множество A состоит из неопределенного числа элементов x : признаки, по которым элементы включаются в расплывчатое множество, не позволяют однозначно отделить все элементы, входящие в него, от элементов, ему не принадлежащих; по крайней мере некоторые элементы можно считать как относящимися к множеству, так и не входящими в него.

Важным является понятие функции принадлежности $\gamma_A(x)$. Считается, что для каждого элемента x можно задать число $\gamma_A(x)$, $0 \leq \gamma_A(x) \leq 1$, выражающее степень принадлежности этого элемента к расплывчатому множеству A . Если $\gamma_A(x) = 0$, то элемент x определенно не принадлежит множеству A , если $\gamma_A(x) = 1$ – определенно входит в него. Величина $\gamma_A(x)$, рассматриваемая как функция аргумента x , и называется функцией принадлежности. Если $\gamma_A(x)$ принимает значения только либо 0, либо 1, то множество A является нерасплывчатым (например, множеству A чисел, не превосходящих 5, соответствует функция $\gamma_A(x) = \{1: x \leq 5; 0: x > 5\}$). Характерным признаком расплывчатости множества является наличие хотя бы одного элемента с функцией принадлежности, отличной от 0 и 1 (например, множество R^+ положительных чисел становится размытым, если положить $\gamma_{R^+}(0) = 1/2$, так как есть основания считать нуль “отчасти положительным, а в чем-то отрицательным” числом).

Итак, *расплывчатое множество A в X* определяется как совокупность упорядоченных пар вида

$$A = \{x, \gamma_A(x)\}, x \in X.$$

Пустое расплывчатое множество \emptyset определяется как такое, для которого $\mu_\emptyset(x) := 0$.

Иногда удобно использовать понятие *носителя* $S(A)$ расплывчатого множества A , который определяется как такое множество, для которого

$$[x \in S(A) \subseteq X] \Leftrightarrow [\gamma_A(x) > 0].$$

Расплывчатое множество A называется *номинальным* тогда и только тогда, когда $\sup_x \gamma_A(x) = 1$, в противном случае – *субнормальным*. Непустое субнормальное множество можно нормализовать, разделив $\gamma_A(x)$ на $\sup_x \gamma_A(x)$. В связи с возможностью субнормальности следует дополнить определение нерасплывчатого множества следующим образом, когда $\gamma_A(x) = \text{const} < 1$ для всех $x \in S(A)$.

Равенство двух расплывчатых множеств A и B определяется условием

* Слово FLIES имеет смысл не только глагола “летит”, но и существительного “мухи”.

** Слово TIME может означать “измерять время и/или величины, связанные со временем”.

* Английский термин fuzzy sets, предложенный Л. Задэ и прижившийся в научной литературе, на русский язык переводится как “размытые”, “нечеткие”, “расплывчатые” и т.д. множества. Термин пока не стандартизирован.

$$(A = B) \Leftrightarrow (?_A(x) = ?_B(x)) \forall x \in X.$$

Включение расплывчатого множества A в множество B определяется следующим образом:
 $(A \subseteq B) \Leftrightarrow [?_A(x) \leq ?_B(x)] \forall x \in X.$

Например, множество очень больших чисел является подмножеством больших чисел.

Расплывчатое множество A' называется *дополнением* к расплывчатому множеству A тогда и только тогда, когда $?_{A'}(x) = 1 - ?_A(x)$. Например, множества “высокие люди” и “невысокие люди” могут быть как дополнительными друг к другу, если их функции принадлежности в сумме тождественно равны единице, так и не являться дополнительными при другом зада-

SUPPLEMENT,	
COMPLEMENT	
<i>дополнение</i>	
LABEL	
<i>метка</i>	
SET	
<i>множество</i>	
MEMBERSHIP	
<i>принадлежность</i>	
<i>(множеству)</i>	
FUZZINESS	
<i>расплывчатость</i>	

Расплывчатость – это такое свойство явлений, при котором не выполняется отношение эквивалентности: явление одновременно может принадлежать данному классу и не принадлежать ему.
Неопределенность

соответствующие задачи (некоторые из таких задач будут рассмотрены в гл. 7). Цель данного параграфа – дать представление о том, как можно построить математическую модель наблюдений, не удовлетворяющих аксиомам тождества. Иными словами, каждая измерительная шкала может быть “размыта”. Для размытия шкал наименований и порядка достаточно тех понятий, которые приведены выше; количественные шкалы требуют некоторых дополнительных определений.

Самым “узким” местом теории (и практики) размытых множеств является задание функций принадлежности. Существует несколько подходов к определению функции $?_A(x)$:

нии этих функций.

Пересечение размытых множеств A и B определяется соотношением

$$A \cap B \Leftrightarrow ?_{A \cap B}(x) = \min[?_A(x), ?_B(x)], x \in X.$$

Объединением размытых множеств A и B называется расплывчатое множество $A \cup B$, удовлетворяющее условию

$$A \cup B \Leftrightarrow ?_{A \cup B}(x) = \max[?_A(x), ?_B(x)], x \in X.$$

В некоторых приложениях удобно определить такие составные множества, которые соответствуют конкретным арифметическим операциям над функциями принадлежности составляющих множеств.

Так, *алгебраическое произведение* расплывчатых множеств A и B обозначается через AB и определяется равенством

$$?_{AB}(x) = ?_A(x) \cdot ?_B(x), x \in X;$$

алгебраическая сумма $A \oplus B$ соответствует равенству

$$?_{A \oplus B}(x) = ?_A(x) + ?_B(x) - ??_A(x) \cdot ?_B(x), x \in X.$$

Говорят, что имеет место *расплывчатое отношение* R между элементами x и y множеств X и Y , если множество пар (x, y) , удовлетворяющих этому отношению xRy , образует расплывчатое множество в $X \times Y$, т.е. можно задать $?_R(x, y)$ – функцию принадлежности (x, y) к R .

Например, пусть отношение R есть $x \gg y$:

$$?_R(x, y) = \{ 0: x \leq y; [1 + (x - y)^{-2}]^{-1}: x > y \}.$$

Пусть C – расплывчатое множество в пространстве $X \times Y$ с функцией принадлежности $?_C(x, y)$. Множество C называется *разложимым* по X и Y в том и только в том случае, если C допускает представление $C = A \cap B$, или, что то же самое,

$$?_C(x, y) = \min[?_A(x), ?_B(y)].$$

Мы привели основные (не все) понятия, с помощью которых строится теория размытых множеств и решают-

- 1) *эвристический подход*, когда субъект сам определяет, как он понимает степень принадлежности (например, числа n к множеству “несколько”); функции, задаваемые разными людьми для одного множества, могут различаться, что отражает разницу в понимании расплывчатого термина;
- 2) *статистически подход*, при котором $\hat{\gamma}_A(x)$ определяется усреднением функций, задаваемых разными экспертами;
- 3) *частичное задание* $\hat{\gamma}_A(x)$ *поясняющими примерами* (например, для нескольких значений x) и последующее доопределение всей функции подходящим методом;
- 4) *интервальное определение* типа задания пессимистической и оптимистической границ для функции $\hat{\gamma}_A(x)$;
- 5) *кратная расплывчатость*, т.е. задание $\hat{\gamma}_A(x)$ как размытого множества с помощью функции принадлежности второго порядка $\hat{\gamma}_A^2(\hat{\gamma}_A(x))$.

Подведем итог

Расплывчатость является специфическим видом неопределенности. Ее главная особенность состоит в том, что в результате наблюдения конкретизируется лишь сам наблюдаемый объект, а неопределенность его принадлежности к известной заранее, сохраняется. Это описывается с помощью функции принадлежности. Другие особенности расплывчатых ситуаций моделируются аксиомами теории расплывчатых множеств.

Summary

Fuzziness is a specific type of uncertainty. Its main feature is that observation concretely defines merely the object itself, but uncertainty of its membership in a fuzzy set remains the same as before. This is represented by introducing the concept of the membership function. The other characteristics of fuzzy situations are modeled by the axioms of fuzzy set theory.

§ 6.4. ВЕРОЯТНОСТНОЕ ОПИСАНИЕ СИТУАЦИЙ.

СТАТИСТИЧЕСКИЕ ИЗМЕРЕНИЯ

Говоря о наблюдениях над изучаемым объектом и о фиксации результатов этих наблюдений (измерений), а именно это является основной темой данной главы, еще раз напомним, что сама необходимость обращения к эксперименту вытекает из того, что нужно устранить некоторую неопределенность, свойственную нашим знаниям об объекте до проведения этого эксперимента. В некоторых случаях эксперимент устраняет неопределенность полностью (как при бросании монеты или контролльном замере уровня масла в двигателе); в других случаях неопределенность лишь уменьшается до некоторого предела, относительного (т.е. в принципе преодолимого) или абсолютного (неуменьшаемого). Ясно, что и организация эксперимента, и обработка экспериментальных данных, определяющие степень уменьшения неопределенности, должны исходить из природы, существа, причины неопределенности.

ПОНЯТИЕ СЛУЧАЙНОЙ НЕОПРЕДЕЛЕННОСТИ

Оказывается, что неопределенность бывает разного происхождения. Один из ее видов – неизвестность – рассматривается теорией познания и философией; такого типа неопределенность характеризует ситуацию, когда мы задаемся вопросом “есть ли жизнь на Марсе?” (посадка советской автоматической станции на эту планету уменьшила неопределенность, но не сняла ее совсем) или “существуют ли внеземные цивилизации?” (поиск возможных искусственных радиосигналов, пока, к сожалению, безуспешен). Другой вид неопределенности – расплывчатость – был обсужден в предыдущем параграфе; для нее характерно, что экспери-

DATA
данные
EXPERIMENT
DESIGN
планирование
эксперимента
ASSUMPTION,
SUPPOSITION,
PREMISE
предположение
DISTRIBUTION
распределение
STATISTICAL
статистический

Самая полная
информация, ко-
торой можно рас-
полагать о слу-
чайном объекте,
содержится в рас-
пределении веро-
ятностей по воз-
можным состоя-

О. Курно писал, что “случайность вмешивается во все, что творится на свете”, что “мир управляет случай, или, говоря точнее, случай имеет свою часть, и притом весьма значительную, в управлении вселенной”.

3. Промежуточная позиция признает как существование вполне детерминированных явлений, так и в принципе случайных, описываемых статистическими закономерностями (большая частота рождения мальчиков по сравнению с девочками; законы Менделея; статистические законы физики, химии, термодинамики; законы квантовой механики и т.д.). Таким образом, случайность признается объективным свойством лишь некоторых явлений.

4. В последние годы представители школы И. Пригожина разрабатывают подход, согласно которому случайные и детерминированные периоды сменяют друг друга в истории любой системы. Детерминированные процессы постепенно сменяются процессами, все более удаленными от равновесия, все более хаотическими, пока в период сильной неравновесности случайность не становится определяющей причиной того, в какое из возможных новых равновесных состояний придет система.

мент в принципе не снимает ее полностью. Третий вид неопределенности – случайность – мы кратко рассмотрим сейчас; при этом будем исходить из того, что читателю знакомы элементы теории вероятностей.

Говоря о случайных явлениях, прежде всего обращают внимание на их непредсказуемость, противопоставляют случайность детерминированности, хаотичность – упорядоченности. Имеющее определенный смысл, такое противопоставление является односторонним, так как оставляет в тени тот факт, что *под случайностью понимается вид неопределенности, подчиняющийся строгой закономерности, которая выражается распределением вероятностей*. Зная распределение (например, плотность $p(x)$) вероятностей, можно ответить на любой вопрос о случайной величине: в каком интервале находятся ее возможные значения (определим носитель распределения X – множество элементов x , для которых $p(x) > 0$); около какого значения рассеиваются ее реализующие значения (найдем параметр положения распределения, например среднее, моду или медиану); насколько сильно разбросаны эти значения (найдем масштабный параметр – дисперсию или стандартное отклонение, средний модуль разности, энтропию); какова связь между разными реализациями (вычислим заданную меру зависимости) и т.д.

О ПРИРОДЕ СЛУЧАЙНОСТИ

Различные определения вероятности рассматриваются в соответствующих курсах; мы же остановимся на природе случайности. Существует несколько точек зрения на этот счет, причем каждая из них имеет достаточные основания.

1. Согласно первой точке зрения, случайным нам представляется нечто такое, в чем мы пока не уловили закономерности. По мере познания явления в последнем остается все меньше и меньше случайного. Ярким выразителем такой позиции был Лаплас, считавший, что случайность не присуща самим объектам, а связана только с незнанием, в принципе устранимым.

2. Противоположная точка зрения состоит в том, что случайность является объективным свойством всех явлений, а детерминированность – лишь предельный случай случайности. Более ста лет назад

СТАТИСТИЧЕСКИЕ ИЗМЕРЕНИЯ

При всем отличии этих точек зрения они не столько несовместимы, как это представляется на первый взгляд. Рассмотрим наглядную в этом отношении простейшую задачу обнаружения постоянного сигнала в нормальном шуме. Сигнал может принимать одно из двух (известных заранее) значений, S или 0 , с вероятностями P и Q соответственно (это и есть случайность незнания). В любом из этих случаев мы можем наблюдать только аддитивную смесь сигнала с гауссовым шумом, т.е. иметь выборку x_1, \dots, x_N либо из распределения $N_x(0, ?^2)$, либо из распределения $N_x(S, ?^2)$. Здесь шум представляет собой объективную и неустранимую случайность, подчиненную закономерности нормального распределения. Зная априори величины $P, Q, S, 0$, функцию $N_x(a, ?^2)$ при любых a и $?^2$, а также используя наблюдения x_1, \dots, x_N , мы можем уменьшить неопределенность того, какое же из возможных значений, S или 0 , имеет место, т.е. уменьшить случайность незнания. Однако объективная случайность шума не позволяет сделать это безошибочно; даже при оптимальных методах обработки измерений x_1, \dots, x_N вероятности ошибок отличны от нуля, хотя при неограниченном увеличении N они стремятся к нулю.

Как видим в практических задачах объективные и субъективные случайности неразделимо переплетены. Такое слияние может быть еще более тесным: например, в непараметрической статистике распределения, характеризующие объективную случайность, считаются лишь существующими, но функционально неизвестными, т.е. субъективное незнание распространяется и на описание объективной случайности.

Итак, как и любые эксперименты, измерения случайных величин и процессов выполняются для уточнения их моделей, снятия или уменьшения неопределенности незнания. Обычно достаточно знать не все распределение, а лишь какой-то из его параметров, и тогда задача сводится к оценке этого параметра по наблюдаемой выборке. Хотя это уже “вторичная” обработка данных, измерение выборочных значений и вычисление оценки в совокупности можно трактовать как “измерение параметра”. То же относится и к определению по выборке более сложных характеристик – самих распределений, регрессий, корреляций, спектров и т.д. Такое совместное рассмотрение непосредственных измерений и их обработки оказывается полезным еще и потому, что можно проводить общую оптимизацию этого процесса, и она далеко не всегда совпадает с оптимизацией компонент в отдельности.

Все эти соображения и дают основания ввести понятие *статистических измерений*, рассматривать эту проблематику как самостоятельный раздел метрологии со своей теорией и измерительной техникой [2; 12].

В заключение подчеркнем еще раз, что статистический, вероятностный подход относится к неопределенности, описываемой распределениями вероятностей. На то, что методы статистики надо применять осторожно, что многие экспериментальные ситуации могут быть хотя и хаотическими, но не иметь вероятностного характера, обращали внимание многие исследователи. В учебной и популярной литературе этот момент настойчиво и очень эмоционально подчеркивает В.Н. Тутубалин [9].

Еще один важный момент состоит в том, чтобы по возможности ослабить или хотя бы учесть влияние измерений на наблюдаемый объект. Особенно это существенно при социальных исследованиях, наблюдениях за людьми: сам факт осознания, что они являются объектом внимания, заметно меняет их поведение. Воздействие измерительного устройства на измеряемый объект должно также учитываться при физических и химических экспериментах.

Подведем итог

Случайная неопределенность характеризуется предположением о том, что распределение вероятностей существует, хотя и неизвестно. После наблюдений над случайной величиной требуется снять неопределенность ее распределения (или его заданной характеристики). Чем больше произведено наблюдений, тем больше имеется возможностей снять неопределенность. Как именно это сделать и от чего зависит оставшаяся неопределенность – на эти вопросы отвечает математическая статистика (см. § 7.7).

Summary

Under stochastic uncertainty we are able to suppose that there exists a probability distribution, even if it is unknown. Having observed a random variable we must remove the uncertainty of the distribution (or of its characteristic). The larger the size of a sample, the more possibilities there are for doing this. How to do this, and what the remaining uncertainty depends on – these questions are answered by mathematical statistics (sf. § 7.7).

§ 6.5. РЕГИСТРАЦИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ И ЕЕ СВЯЗЬ С ПОСЛЕДУЮЩЕЙ ИХ ОБРАБОТКОЙ

Результаты любого эксперимента фиксируют в той или иной форме, а затем используют для той цели, ради которой и проводился эксперимент. Иногда эти операции практически совмещены во времени, например при автоматическом управлении производственным процессом, при автоматизации экспериментов в реальном масштабе времени и т.п. В некоторых же видах человеческой практики (научные исследования; системный анализ; контрольная, ревизионная, следственная и другие виды административной деятельности; учебные эксперименты и пр.) обработка экспериментальных данных является отдельным, самостоятельным этапом, промежуточным между этапами получения информации (измерения) и ее использования (принятия решений и их выполнения). В таких случаях исходной информацией для обработки являются *протоколы наблюдений* (называемые также *матрицами данных, экспериментальными таблицами*).

Характер самих протоколов наблюдений и методы их обработки зависят от того, какова модель, для уточнения которой ставится эксперимент: фактически обработка данных – это просто преобразование информации к виду, удобному для использования, перевод ответов природы с языка измерений на язык уточняемой модели. Наши знания могут быть как первоначальными, грубыми, так и далеко продвинутыми, хорошо структурированными, хотя и требующими уточнения. Соответствующие два типа моделей разные авторы называют по-разному, в зависимости от того, какой их аспект они хотят подчеркнуть: дескриптивные и конструктивные, качественные и количественные, декларативные и процедуральные, *классификационные и числовые*. Мы в данном параграфе будем пользоваться последними терминами; отметим, что в них отражено и то различие, что классификационные модели описывают множество различных объектов, а числовые – один объект (или множество схожих объектов). Отметим также, что в классификационной модели могут участвовать количественные переменные, и это не меняет ее качественного характера (например, диагноз больному ставится с учетом количественных анализов); аналогично, в числовых моделях часть переменных может измеряться в слабых шкалах. Рассмотрим кратко особенности экспериментальных данных и их обработки для обоих типов моделей.

OBSERVATION	
<i>наблюдение</i>	
DIMENSIONALITY	
<i>размерность</i>	
DISTRIBUTION	
<i>распределение</i>	
ORDERING	
<i>упорядочивание</i>	

Всякая статистическая задача состоит в том, чтобы, несмотря на неустранимую

рядка между $x_{10}, x_{20}, \dots, x_{N_0}$ (или некоторой их частью) по определенному критерию предпочтения.

Уменьшение размерности модели. Классификационные модели как первоначальные, “сырые”, учитывают множество предположений, которые еще надо проверять. Так, сам список признаков X формируется эвристически, часто “с запасом”, и оказывается довольно длинным, а главное, избыточным, содержащим “дублирующие” и “шумящие” признаки. Поэтому одна из важных задач совершенствования классификационных моделей состоит в уменьшении размерности модели с помощью отбора наиболее информативных признаков, “склеивания” нескольких признаков в один и т.п. Как следует из практики, информативные признаки могут оказаться различными для разных классов [3].

ЧИСЛОВЫЕ МОДЕЛИ

Числовые модели отличаются от классификационных тем, что: 1) целевые признаки x_0 измеряются в числовых шкалах; 2) числа x_0 представляют собой функционалы или функции признаковых переменных (которые не обязательно все являются числовыми); 3) в них гораздо чаще учитываются связи переменных во времени (в классификационных задачах время иногда даже называют “забытой” переменной). В связи с этим и протоколы наблюдений мо-

КЛАССИФИКАЦИОННЫЕ МОДЕЛИ

Классификационные модели являются основополагающими, первичными, исходными формами знания. Узнавание окружающих предметов – типичный пример классификационных процессов в мыслительной деятельности человека (и животных). И в науке познание начинается с соотнесения изучаемого объекта с другими, выявления сходства и различия между ними. Поэтому протокол наблюдений на классификационном уровне эксперимента содержит результаты измерения ряда признаков X для подмножества A объектов, выбранных из множества Γ : каждый объект $a_i \in A \subseteq \Gamma$ обладает значениями признаков $x_i = (x_{i0}, x_{i1}, \dots, x_{in}) \in \{X_0, X_1, \dots, X_n\} = X$, $i = \overline{1, N}$, n – число признаков, N – число объектов в A . Признак характеризует конкретное свойство объекта, поэтому иногда такой протокол называют таблицей “объект – свойство”.

Как уже отмечалось, способ обработки протокола зависит от цели обработки. Часто оказывается, что задача может быть сформулирована как определение по наблюдавшимся значениям признаков $x = (x_1, \dots, x_n)$ значений ненаблюданного (“целевого”) признака x_0 . Как правило, целевыми признаками являются те параметры модели, которые требуется уточнить по экспериментальным данным.

Рассмотрим различные типы задач для классификационных моделей.

Кластеризация (поиск “естественной” группировки объектов). Не заданы ни границы классов в пространстве признаков, ни число классов. Требуется их определить исходя из “близости”, “похожести” или “различия” описаний объектов $x_i = (x_{i1}, \dots, x_{in})$. Компоненты вектора X_0 – признаки кластера, значения которых подлежат определению.

Классификация (распознавание образов). Число классов задано. Если также заданы границы между классами, то имеем априорную классификацию; если границы требуется найти, оценить по классифицированным примерам, то задача называется *распознаванием образов по обучающей выборке*. Целевой признак X_0 имеет значения вnominalной шкале (имена классов).

Упорядочивание объектов. Требуется установить отношения порядка между $x_{10}, x_{20}, \dots, x_{N_0}$ (или некоторой их частью) по определенному критерию предпочтения.

гут не обязательно относиться к множеству объектов; модель можно уточнять и по экспериментам с одним объектом в разные моменты времени.

Отметим, что числовые модели могут задавать связь между переменными как в параметризованной форме (т.е. в виде функции с конечным числом параметров), так и в непараметризованной форме (в виде функционала). Например, зависимость между входом $x = \{x_i\}$ и выходом y некоторой системы может задаваться в виде параметризованной линейной регрессии $y = ?a_i x_i + + E$ либо в непараметризованной форме как функционал линии регрессии $y(x) = \int y \cdot p(y/x) dy$, где $p(y/x)$ – неизвестная плотность условного распределения вероятностей.

Приведем типичные задачи для числовых моделей.

Косвенные измерения (оценка параметра). Требуется определить значение x_0 по заданному множеству $\{x_{ij}\}$. В отличие от классификации x_0 измеряется не в номинальной, а в числовой шкале. Если $\{x_{ij}\}$ определены до некоторого момента t_0 , а x_0 требуется оценить для $t > t_0$, то задача называется *прогнозированием*. (Прогнозирование имеет смысл и в задаче классификации; например, ранняя диагностика заболевания.)

Поиск экстремума (планирование эксперимента). Считается, что имеется возможность пошагового изменения величин $\{x_{ij}(t_k)\}$, $t_k = t_0 + k?t$, $k = 0, 1, 2, \dots$. Требуется изменять их так, чтобы в конце концов получить экстремальное значение целевого признака x_0 .

ОСОБЕННОСТИ ПРОТОКОЛОВ НАБЛЮДЕНИЙ

Отметим встречающиеся на практике особенности реальных протоколов наблюдений, которые следует учитывать при их обработке.

Большая размерность. Во многих исследованиях число объектов N и число признаков n велики, так что произведение $n \times N$ достигает нескольких десятичных порядков. Учет времени приводит к еще большему увеличению размерности блока данных. В настоящее время применение ЭВМ существенно расширяет количественные возможности обработки данных, но “проклятие размерности” остается в силе и для ЭВМ.

Разнотипность данных. Разные признаки могут измеряться в различных шкалах. Многие алгоритмы построены для обработки однотипных переменных, что часто вызывает необходимость приводить разнотипные данные к одной шкале. Ясно, что более правильной стратегией поведения является разработка алгоритмов, специально построенных так, чтобы имела возможность обрабатывать разнотипные данные, не внося в протокол никаких изменений, не связанных с экспериментом.

Пропущенные значения. Незаполненная ячейка таблицы данных – не такой уж редкий случай, особенно если эксперимент производится не в лабораторных, а в естественных условиях. Исключить из таблицы строку и столбец, на пересечении которых находится пустая ячейка, – выход далеко не всегда приемлемый. Можно, используя избыточность таблицы, некоторым образом “восстановить” пропущенные значения, а затем обрабатывать таблицу так, будто их и не было. Однако критерий “восстановления” и цель обработки должны быть согласованы, поэтому не может быть универсального способа “восстановления” пропусков. Хотя этот путь в ряде случаев вполне допустим, перспективным представляется конструирование алгоритмов обработки, позволяющих использовать таблицы с пробелами без их предварительного заполнения.

Зашумленность. Довольно часто измерение, занесенное в протокол, на самом деле отличается от измеряемого значения на некоторую случайную величину. Статистические свойства этой добавочной помехи могут не зависеть от измеряемой величины, и тогда мы говорим об аддитивном шуме. В противном случае имеет место неаддитивная или зависимая помеха. Все эти варианты должны по-разному учитываться при обработке.

Искажения, отклонения от предположений. Приступая к обработке протокола наблюдений, мы всегда исходим из определенных предположений о природе величин, занесенных в протокол. Любой способ обработки дает результаты ожидаемого качества только в том случае, если данные отвечают определенным предположениям. Далеко не всегда в ходе обра-

ботки данных обращают внимание на то, действительно ли данные отвечают предположениям, заложенным в алгоритм обработки.

Например, данные могут выглядеть как неразмытые, но быть на самом деле расплывчатыми (см. § 6.3). Цифры в действительности могут быть символами, а мы можем считать, что они числа. Числовые шкалы предполагают одинаковость единиц измерения вдоль всей шкалы (см. § 6.2), а измерительный прибор может обладать нелинейной характеристикой, и если это не отражено в протоколе, то мы будем обрабатывать искаженные данные. Измеряемая величина может быть непрерывной, но в протоколе она неизбежно приводится с округлением, и это также является искажением.

Чтобы повысить качество выводов, получаемых при обработке данных, мы должны обеспечить соответствие свойств данных и требований к ним алгоритмов либо максимально обезопасить себя от возможного несоответствия: контролировать условия эксперимента; вносить допустимые поправки в протоколы (например, производить перерасчет нелинейности); наконец, разрабатывать алгоритмы, либо содержащие возможный минимум предположений (например, процедуры непараметрической статистики), либо специально разработанные с расчетом на возможные отклонения (как в робастной статистике).

Способы обработки экспериментальных данных с целью извлечения из них полезной информации и отсеивания ненужной, мешающей, мы рассмотрим в следующей главе.

Подведем итог

Summary

Данный параграф посвящен тому, как обрабатывать результаты измерений. Рассмотрены два аспекта этой проблемы: связь способа обработки с целию измерений (т.е. в конечном счете с проверяемой моделью) и его связь с условиями измерений, приводящими к тому, что реальные протоколы наблюдений обычно далеки от желаемого идеала.

The processing of experimental data is the subject of this section. There are many problems connected with this question. In this section we have singled out only two: the relationship between data processing and the purpose of the experiment; and data processing's connection with experimental conditions that result in various errors in real-life data.

ЗАКЛЮЧЕНИЕ

Системный анализ часто приводит к необходимости экспериментального исследования изучаемой системы. Такие эксперименты должны проводиться обязательно с участием прямых специалистов в данной конкретной области. Однако системный аналитик, оказавшийся в каком-то случае не прямым специалистом, не может ограничиваться лишь ролью заказчика на выполнение эксперимента. Очень многое зависит от постановки цели опытов, но часто не менее важно правильно извлечь информацию из результатов опыта. Поэтому предметом специального внимания должны стать такие вопросы, как выбор шкал максимально допустимой силы; употребление методов обработки, содержащих только допустимые преобразования исходных данных; учет реальных особенностей протоколов наблюдений в алгоритмах их обработки.

ЛИТЕРАТУРА

1. *Беллман Р., Задэ Л.* Принятие решений в расплывчатых условиях. – В. кн.: Вопросы анализа и процедуры принятия решений. – М.: Мир, 1976,
2. *Бендарт Дж., Пирсол А.* Измерение и анализ случайных процессов. – М.: Мир, 1974.
3. *Лбов Г.С.* Методы обработки разнотипных экспериментальных данных –Новосибирск: Наука, 1981.
4. *Леонардо да Винчи.* Избр. естественно-научные произв. – М.: АН СССР, 1965.
5. *Пригожин И., Стенгерс И.* Порядок из хаоса. – М.: Прогресс, 1986.
6. *Пфанцагль И.* Теория измерений. – М.: Мир, 1976.
7. *Рафаэл Б.* Думающий компьютер. – М.: Мир, 1979.
8. *Розенблют А., Винер Н.* Роль моделей в науке. – В кн.: Неуймин Я.Г. Модели в науке и технике. – Л.: Наука, 1984.
9. *Тутубалин В.Н.* Теория вероятностей. Краткий курс и научно-методические замечания. – М.: МГУ, 1972.
10. *Фейнман Р.* Характер физических законов. – М.: Мир, 1968.
11. *Черчмен У., Акофф Р., Арноф Л.* Введение в исследование операций. – М.: Наука, 1968.
12. *Gray R.M., Davisson L.D.* Random Processes; A Mathematical Approach for Engineers. –New Jersey: Prentice – Hall Inc., 1986.

УПРАЖНЕНИЯ

§ 6.1

- Темы для обсуждения: соотношение априорных знаний (моделей) и практических действий в постановке и проведении: а) активного эксперимента; б) пассивного наблюдения.

§ 6.2

- Приведите примеры наблюдений в каждой из измерительных шкал.
- Тема для обсуждения: что происходит при рассогласовании между природой наблюдаемого явления и силой измерительной шкалы? как обеспечить их согласование?
- Тема для обсуждения: когда недопустимые преобразования результатов наблюдений безвредны?

§ 6.3

- Постройте графики функций принадлежности чисел $x \in \mathbf{R}^+$ к расплывчатым множествам “мало”, “много”, “насколько”. Сравните ваши графики с графиками соседа и обсудите природу различий.

• Докажите, что $A \cap A' \neq \emptyset$, если A – расплывчатое множество, и $A \cap A' = \emptyset$, если A – нерасплывчатое множество.

• Покажите, что $A \cup B = (A' \cap B')'$.

• Покажите, что $A \oplus B = (A' \cdot B')'$.

§ 6.4

• Тема для обсуждения: причины, приведшие к разным определениям вероятности. Что общего и что различно для этих определений?

§ 6.5

• Обсудите, какие из особенностей реальных протоколов наблюдений, перечисленные в тексте, в действительности встречались в вашей практике измерений.

**Вопросы
для самопроверки**

1. Почему верными оказываются оба противоположных утверждения: “опыт определяет модель” и “модель определяет опыт”?
2. Что такое измерение?
3. Почему над наблюдениями в некоторой шкале можно производить не любые, а только допустимые операции?
4. Каковы возможные последствия “усиления” и “ослабления” наблюдений, т.е. пересчета протокола наблюдений в шкалу, отличающуюся от той, в которой производилось измерение?
5. Чем отличается расплывчатая неопределенность от вероятностной?
6. Какими способами можно задать функцию принадлежности?
7. Как описывается вероятностная неопределенность?
8. Каковы основные отклонения свойств реальных протоколов наблюдений от желаемых?

§ 7.1. МНОГООБРАЗИЕ ЗАДАЧ ВЫБОРА

Главная цель курса системного анализа – раскрытие системности любой целенаправленной деятельности. Для этого необходимо построить систему моделей, с помощью которых можно обобщать, передавать и совершенствовать опыт такой деятельности. В предыдущих главах мы уже выделили некоторые из операций, входящие во всякую целенаправленную деятельность: моделирование (см. гл. 2), перенос информации во времени и пространстве (см. гл. 5), получение новой информации (см. гл. 6). Предстоит построить еще достаточно полный список действий, из которых складывается всякая успешная деятельность, и только после этого можно будет приступить к обсуждению ее структуры и принципов организации. В данной главе рассмотрим еще одну операцию, обязательно входящую в целенаправленные процессы, – выбор.

ВЫБОР КАК РЕАЛИЗАЦИЯ ЦЕЛИ

Выбор является действием, придающим всей деятельности целенаправленность. Именно выбор реализует подчиненность всей деятельности определенной цели или совокупности целей. Рано или поздно наступает момент, когда дальнейшие действия могут быть различными, приводящими к разным результатам, а реализовать можно только одно действие, причем вернуться к ситуации, имевшей место в этот момент, уже (как правило) нельзя.

Способность сделать правильный выбор в таких условиях – очень ценное качество, которое присуще людям в разной степени. Великие полководцы, выдающиеся политики, гениальные инженеры и ученые, талантливые администраторы отличались и отличаются от своих коллег или конкурентов прежде всего умением принимать лучшие решения, делать лучший выбор.

Естественно стремление понять, что такое “хороший выбор”, выработать рекомендации, как приблизиться к наилучшему решению, а если возможно, то и предложить алгоритм получения такого решения. Работа многих исследователей в этом направлении выявила характерную ситуацию, типичную для моделирования (в данном случае – моделирования процессов принятия решений): полная формализация нахождения наилучшего решения возможна, но лишь для хорошо изученных (хорошо структурированных) задач; для решения слабо структурированных задач полностью формальных алгоритмов не существует (если не считать тривиального и далеко не всегда приемлемого алгоритма перебора, т.е. метода проб и ошибок), но опытные и способные специалисты часто делают выбор, оказывающийся хорошим. Современная тенденция практики выбора в естественных ситуациях состоит в сочетании способности человека решать неформализованные задачи с возможностями формальных методов и компьютерного моделирования (например, диалоговые системы поддержки решений, экспертные системы, информационно-поисковые системы, системы управления базами данных, автоматизированные системы управления и т.п.).

Задачи выбора чрезвычайно многообразны, различны и методы их решения. Прежде всего введем понятия, общие для всех задач выбора.

Будем представлять *принятие решения как действие над множеством альтернатив, в результате которого получается подмножество выбранных альтернатив* (обычно это одна альтернатива, что не обязательно, а иногда и невозможно). Сужение множества альтернатив возможно, если имеется способ сравнения альтернатив между собой и определения наиболее предпочтительных. Каждый такой способ будем называть **критерием предпочтения**. Обратим внимание на то, что при таком описании выбора считают сами собой разумеющимися, уже пройденными, два чрезвычайно важных этапа: 1) порождение множества альтернатив, на котором предстоит осуществлять выбор; 2) определение целей, ради достижения которых производится выбор. В практике системного анализа реализация этих этапов связана с определенными трудностями, для преодоления которых необходимы свои приемы и методы. В

гл. 9 мы вернемся к этим действиям, отнеся их к числу этапов системного анализа, а пока будем считать, что исходное множество альтернатив, из которых требуется выбрать наиболее предпочтительные, уже задано и преследуемые нами цели определены настолько детально, что уже имеются критерии оценки и сравнения любых альтернатив.

МНОЖЕСТВЕННОСТЬ ЗАДАЧ ВЫБОРА

Даже в такой упрощенной постановке проблема выбора не тривиальна и допускает существенно различающиеся математические постановки задач. Дело в том, что каждая компонента ситуации выбора может реализоваться в качественно различных вариантах. Отметим основные из этих вариантов:

множество альтернатив может быть конечным, счетным или континуальным;

оценка альтернатив может осуществляться по одному или по нескольким критериям, которые в свою очередь могут иметь как количественный, так и качественный характер;

режим выбора может быть однократным (разовым) или повторяющимся, допускающим обучение на опыте;

последствия выбора могут быть точно известны (*выбор в условиях определенности*), иметь вероятностный характер, когда известны вероятности возможных исходов после сделанного выбора (*выбор в условиях риска*), или иметь неоднозначный исход, не допускающий введения вероятностей (*выбор в условиях неопределенности*);

ответственность за выбор может быть односторонней (в частном случае индивидуальной) или многосторонней. Соответственно различают индивидуальный и групповой выбор;

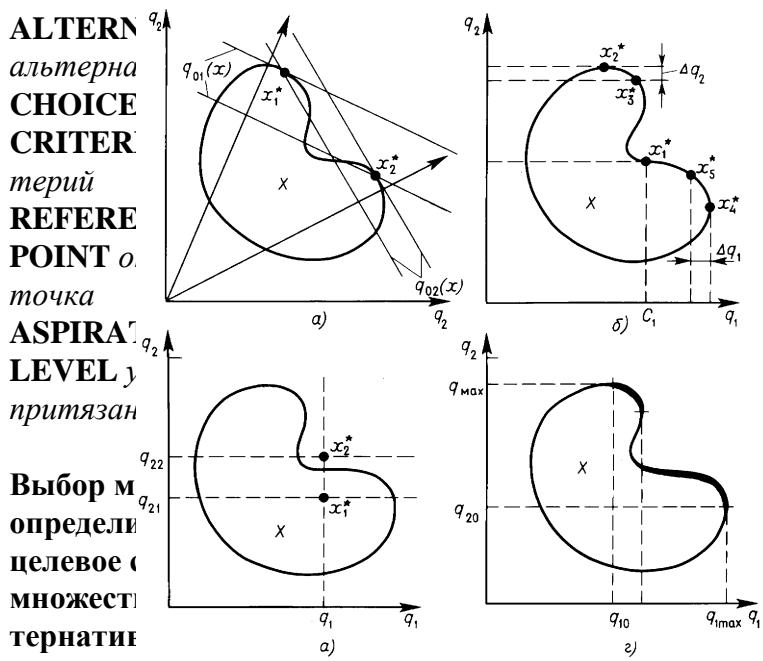
степень согласованности целей при многостороннем выборе может варьироваться от полного совпадения интересов сторон (*кооперативный выбор*) до их противоположности (*выбор в конфликтной ситуации*). Возможны также промежуточные случаи, например *компромиссный выбор*, *коалиционный выбор*, *выбор в условиях нарастающего конфликта* и т.д.

Различные сочетания перечисленных вариантов и приводят к многообразным задачам выбора, которые изучены не в одинаковой степени. В данной главе дадим краткий обзор состояния теории выбора в настоящее время, а также рассмотрим некоторые подходы к решению слабо формализованных задач выбора. При этом главное внимание будем уделять постановке задач и важным результатам и лишь упоминать, какие именно теории дают методы решения (иначе бы резко возрос объем книги, а для ряда специальностей имелось бы место дублирование материала других дисциплин – теории оптимизации, исследования операций, вариационного исчисления, математического программирования, теории игр, математической статистики и т.д.)

§ 7.2. КРИТЕРИАЛЬНЫЙ ЯЗЫК ОПИСАНИЯ ВЫБОРА

На примере описания выбора видно, как об одном и том же явлении можно говорить на языках различной общности. К настоящему моменту сложилось три основных языка описания выбора. Самым простым, наиболее развитым (и, быть может, поэтому чаще употребляемым в приложениях) является **критериальный язык**. Это название связано с основным предположением, состоящим в том, что каждую отдельно взятую альтернативу можно оценить конкретным числом (значением критерия), и сравнение альтернатив сводится к сравнению соответствующих им чисел.

Пусть x – некоторая альтернатива из множества X . Считается, что для всех $x \in X$ может быть задана функция $q(x)$, которая называется *критерием* (*критерием качества, целевой функцией, функцией предпочтения, функцией полезности* и т. д.) и обладает тем свойством, что если альтернатива x_1 предпочтительнее альтернативы x_2 (будем обозначать это $x_1 > x_2$), то $q(x_1) > q(x_2)$ и обратно.



вой критерий –
ального языка оценивания многокритериальных задач: а) санция выбора максимизация по одному “суперкритерию”, ляется предположение о возможностях оценить

каждую отдельно взятую альтернативу определенным числом. При этом выбор сводится к отысканию альтернативы с наибольшим значением критериальной функции

Иллюстрация методов

решения многокритериальных задач: а) санция выбора максимизация по одному “суперкритерию”, ляется предположение о возможностях оценить

ВЫБОР КАК МАКСИМИЗАЦИЯ КРИТЕРИЯ

Если теперь сделать еще одно важное предположение, что выбор любой альтернативы приводит к однозначно известным последствиям (т.е. считать, что выбор осуществляется в условиях определенности) и заданный критерий $q(x)$ численно выражает оценку этих последствий, то наилучшей альтернативой x^* является, естественно, та, которая обладает наибольшим значением критерия:

$$x^* = \arg \max_{x \in X} q(x) \quad (1)$$

Задача отыскания x^* , простая по постановке, часто оказывается сложной для решения, поскольку метод ее решения (да и сама возможность решения) определяется как характером множества X (размерностью вектора x и типом множества X – является ли оно конечным, счетным или континуальным), так и характером критерия (является ли $q(x)$ функцией или функционалом и какой или каким именно).

Однако сложность отыскания наилучшей альтернативы существенно возрастает, так как на практике оценивание любого варианта единственным числом обычно оказывается неприемлемым упрощением (см. § 3.3). Более полное рассмотрение альтернатив приводит к необходимости оценивать их не по одному, а по нескольким критериям, качественно различающимся между собой. Например, при выборе конструкции самолета проектировщикам следует учитывать множество критерии: технических (высотность, скорость, маневренность, грузоподъемность, длительность полета и т.д.), технологических (связанных с будущим процессом серийного изготовления самолетов), экономических (определенющих затраты на производство, эксплуатацию и обслуживание машин, их конкурентоспособность), социальных (в частности, уровень шума, загрязнение атмосферы), эргономических (условия работы экипажа, уровень комфорта для пассажиров) и пр. Даже в обыденной жизни при выборе мы почти никогда не используем единственный критерий: вспомните хотя бы затруднения при выборе подарка ко дню рождения или при выборе места для стоянки в турпоходе.

Итак, пусть для оценивания альтернатив используется несколько критериев $q_i(x)$, $i = 1, \dots, p$. Теоретически можно представить себе случай, когда во множестве X окажется одна альтернатива, обладающая наибольшими значениями всех p критериев; она и является наилучшей. Однако на практике такие случаи почти не встречаются, и возникает вопрос, как же тогда осуществлять выбор (так, например, на рис. 7.1 множеству X соответствуют внутренние точки фигуры на плоскости значений двух критериев q_1 и q_2 ; оба критерия желательно максимизировать).

СВЕДЕНИЕ МНОГОКРИТЕРИАЛЬНОЙ ЗАДАЧИ К ОДНОКРИТЕРИАЛЬНОЙ

Рассмотрим наиболее употребительные способы решения многокритериальных задач. Первый способ состоит в том, чтобы *многокритериальную задачу свести к однокритериальной*. Это означает введение **суперкритерия**, т.е. скалярной функции векторного аргумента: $q_0(x) = q_0(q_1(x), q_2(x), \dots, q_p(x))$. (2)

Суперкритерий позволяет упорядочить альтернативы по величине q_0 , выделив тем самым наилучшую (в смысле этого критерия). Вид функции q_0 определяется тем, как мы представляем себе вклад каждого критерия в суперкритерий; обычно используют аддитивные или мультипликативные функции:

$$q_0 = \sum_{i=1}^p \frac{a_i q_i}{s_i}; \quad (3) \quad 1 - q_0 = \prod_{i=1}^p \mathbf{E}^{-\frac{b_i q_i}{s_i}}. \quad (4)$$

Коэффициенты s_i обеспечивают, во-первых, безразмерность числа q_i/s_i (частные критерии могут иметь разную размерность, и тогда некоторые арифметические операции над ними, например сложение, не имеют смысла) и, во-вторых, в необходимых случаях (как в формуле (4)) выполнение условия $\sum_i q_i/s_i \leq 1$. Коэффициенты a_i и b_i отражают относительный вклад частных критериев в суперкритерий.

Итак, при данном способе задача сводится к максимизации суперкритерия:

$$x^* = \arg \max_{x \in X} q_0(q_1(x), \dots, q_p(x)). \quad (5)$$

Очевидные достоинства объединения нескольких критериев в один суперкритерий сопровождаются рядом трудностей и недостатков, которые необходимо учитывать при использовании этого метода. Оставив в стороне трудности построения самой функции и вычислительные трудности ее максимизации, обратим внимание на следующий очень важный момент. Упорядочение точек в многомерном пространстве в принципе не может быть однозначным и полностью определяется видом упорядочивающей функции. Суперкритерий играет роль этой упорядочивающей функции, и его даже “небольшое” изменение может привести к тому, что оптимальная в новом смысле альтернатива окажется очень сильно отличающейся от старой. На рис. 7.1, *a* видно, как изменяется выбор наилучшей альтернативы при простой смене коэффициентов в линейной упорядочивающей функции (3), что отражается в изменении наклона соответствующей прямой: $q_{01}(x_1^*) > q_{01}(x_2^*)$, но $q_{02}(x_1^*) < q_{02}(x_2^*)$. Заметим, что линейные комбинации частных критериев придают упорядочению следующий смысл: “чем дальше от нуля в заданном направлении, тем лучше”. На рис. 7.1, *a* направления, соответствующие суперкритериям q_{01} и q_{02} , изображены стрелками. Идея такого упорядочивания в многомерном пространстве заложена в некоторых балльных системах оценки вариантов [34]. Другой вариант поиска альтернативы, самой удаленной от нуля в заданном направлении, дает максимизация минимального критерия [23]:

$$x^* = \arg \max_{x \in X} \min_i \frac{|a_i q_i|}{s_i} \quad (6)$$

что означает поиск вокруг направления $\sum_i a_i q_i/s_i = \text{const}$ методом “подтягивания самого отстающего”.

УСЛОВНАЯ МАКСИМИЗАЦИЯ

Недостатки свертывания нескольких критериев заставляют искать другие подходы к решению задач многокритериального выбора. Рассмотрим теперь второй способ решения таких задач. Он заключается в ином, нежели при свертывании, использовании того факта, что

частные критерии обычно неравнозначны между собой (одни из них более важны, чем другие). Наиболее явное выражение этой идеи состоит в выделении основного, главного критерия и рассмотрении остальных как дополнительных, сопутствующих. Такое различие критериев позволяет сформулировать задачу выбора как *задачу нахождения условного экстремума основного критерия*:

$$x^* = \arg \left\{ \max_{x \in X} q_1 \mathbf{b} \mathbf{g} \mathbf{b} \mathbf{g} | C_i, i = 2, 3, \dots, p \right\} \quad (7)$$

при условии, что дополнительные критерии остаются на заданных им уровнях. На рис. 7.1, б приведено решение задачи

$$x_1^* = \arg \left\{ \max_x q_2 \mathbf{b} \mathbf{g} \mathbf{b} \mathbf{g} | C_1 \right\}.$$

В некоторых задачах оказывается возможным или даже необходимым задавать ограничения на сопутствующие критерии не столь жестко, как в задаче (7). Например, если сопутствующий критерий характеризует стоимость затрат, то вместо фиксации затрат разумнее задавать их верхний уровень, т.е. формулировать задачу с ограничениями типа неравенств:

$$x^* = \arg \left\{ \max_{x \in X} q_1 \mathbf{b} \mathbf{g} \mathbf{b} \mathbf{g} | q_i \leq C_i, i = 2, \dots, p \right\}. \quad (8)$$

На рис. 7.1, б приведено решение задачи $x_2^* = \arg \left\{ \max_x q_2 | q_i \leq C_1 \right\}$. Отметим, что такое, казалось бы, незначительное изменение постановки задачи требует принципиально иных методов ее решения. Мы пока не будем касаться этой стороны вопроса и рассмотрим лишь различия в постановках задач выбора.

ВАРИАНТЫ ОПТИМИЗАЦИИ ПРИ РАЗНОВАЖНЫХ КРИТЕРИЯХ

Условная оптимизация, изложенная в предыдущем разделе, является не единственным возможным подходом к рассмотрению задач с разноважными критериями. Возможны и другие варианты, отличие между которыми проистекает из того, что степень разноважности критериев может быть слабо выраженной, а может быть и весьма сильной.

Встречаются случаи, когда пользователь готов на некоторое снижение величин более важных критериев, чтобы повысить величину менее важных. В таких ситуациях можно пользоваться *методом уступок*. Идею этого метода можно изложить следующим образом.

Пусть частные критерии могут быть пронумерованы в порядке убывания их важности. Возьмем первый из них и найдем наилучшую по этому критерию альтернативу (на рис. 7.1, б это x_2^* , если самым важным критерием является q_2 , и x_4^* , если им является q_1). Затем определим “уступку” $?q_i$, т.е. величину, на которую мы согласны уменьшить достигнутое значение самого важного критерия, чтобы в пределах этой уступки попытаться увеличить, насколько возможно, значение следующего по важности критерия (на рис. 7.1, б полученные таким образом альтернативы изображены точками x_3^* и x_5^*). Далее (если число критериев более двух) определяется уступка по только что максимизированному критерию и максимизируется следующий; процедура повторяется до тех пор, пока перечень критериев не закончится.

Как видим, в методе уступок предполагается, что разница в важности критериев не слишком велика; можно предположить, что величина уступок как-то связана с нашим ощущением этой разницы.

Противоположным крайним случаем является ситуация, в которой разница между упорядоченными критериями настолько велика, что следующий в этом ряду критерий рассматривается только в том случае, если сравниваемые альтернативы неразличимы по старшим критериям. Ни о каких уступках при этом не может быть и речи. В этой ситуации выбор довольно часто заканчивается на первом же шаге, а до последнего критерия дело обычно не доходит (точнее, он “изобретается” в том чрезвычайно редком экзотическом случае, когда принятые ранее критерии не выделили единственной альтернативы). Такой выбор получил название *лексикографического упорядочивания* альтернатив, поскольку этот метод используется

при упорядочении слов в различных словарях (предпочтительность определяется алфавитным рангом очередной буквы в данном слове).

ВЫБОР МЕЖДУ УПОРЯДОЧЕНИЯМИ

Для рассматриваемых методов многокритериальной оптимизации существенным является исходное упорядочение критериев. Иногда их порядок очевиден (“кошелек или жизнь?”) или общепризнан (как порядок букв в алфавите), но бывает, что этот вопрос не тривиален, а привлекаемые для его решения эксперты дают несовпадающие упорядочения критериев. Выход состоит в том, чтобы установить, какое из предложенных экспертами упорядочений является “средним”, “типичным” для данной группы. Это опять-таки можно делать по-разному. Среди специалистов пользуется признанием упорядочение, называемое *медианой Кемени*.

Обозначим через R_i упорядочение критериев, предложенное i -м экспертом. Введем некоторую меру расхождения между двумя (i -й и j -й) ранжировками: $d(R_i, R_j)$. Медианой Кемени R^* среди n предложенных упорядочений R_1, R_2, \dots, R_n называется то из них, которое отвечает условию

$$R^* = \arg \min \sum_{j=1}^n d(R_i, R_j),$$

т.е. то, сумма “расстояний” до которого от всех остальных минимальна. Ясно, что многое зависит от того, как определить расстояние d . Например, если $R_i = q_1^{(i)}, \dots, q_p^{(i)}$, то $d_p(R_i, R_j)$ можно определить как $d(h_i, R_j) = p - \sum_{k=1}^p \delta(q_k^{(i)}, q_k^{(j)})$, где $\delta(x^{(i)}, x^{(j)})$ – символ Кронекера.

Однако следует отметить, что с медианой Кемени связано несколько трудностей. Во-первых, оптимизационная задача по нахождению R^* решается методами дискретной оптимизации (динамического программирования, ветвей и границ, и др.), трудоемкость которых экспоненциально растет с увеличением размерности задачи. Во-вторых, иногда решение задачи не единственное, и в этом случае возникают трудности: в литературе приводится пример, когда в одной из оптимальных ранжировок конкретная альтернатива стоит на первом месте, а в другой – на последнем). Поэтому используют и другие способы упорядочения, наиболее известным из которых является *метод строчных сумм*. Пусть критерии сравниваются попарно: $a_{ij} = 1$, если k -й эксперт считает, что q_i важнее q_j ; 0, если наоборот; 1/2, если он считает их равноценными. Для каждого критерия вычисляют величины $S_i = \sum_{k=1}^m \sum_{j=1}^n a_{ij}^k$, $i = I, n$, и критерии упорядочиваются по этой “сумме очков” – совсем как по турнирной таблице в спорте. Однако и этот метод может давать сбои, как и всякое голосование (см. § 7.5).

ПОИСК АЛЬТЕРНАТИВЫ С ЗАДАННЫМИ СВОЙСТВАМИ

Третий способ многокритериального выбора относится к случаю, когда заранее могут быть указаны значения частных критериев (или их границы), и задача состоит в том, чтобы найти альтернативу, удовлетворяющую этим требованиям, либо, установив, что такая альтернатива во множестве X отсутствует, найти в X альтернативу, которая подходит к поставленным целям ближе всего. Характеристики решения такой задачи (сложность процесса вычислений, скорость сходимости, конечная точность и пр.) зависят от многих факторов. Снова оставив в стороне вычислительные и количественные аспекты (что является далеко не простой и в ряде случаев нерешенной задачей), обсудим некоторые принципиальные моменты данного подхода.

Удобным свойством является возможность задавать желательные значения \bar{q}_i критериев как точно, так и в виде верхних или нижних границ; назначаемые значения величин \bar{q}_i иногда называют *уровнями притязаний* [48], а точку их пересечения в p -мерном пространстве критериев – *целью* [23] или *опорной точкой* [48], *идеальной точкой* [22]. Поскольку уровни притязаний задаются без точного знания структуры множества X в пространстве частных

критериев, целевая точка может оказаться как внутри, так и вне X (достижимая или недостижимая цель; на рис. 7.1, в приведены оба варианта, соответственно x_1^* и x_2^*).

Теперь идея оптимизации состоит в том, чтобы, начав с любой альтернативы, приближаться к x^* по некоторой траектории в пространстве X . Это достигается введением числовой меры близости между очередной альтернативой x и целью x^* , т.е. между векторами $q(x) = (q_1(x), \dots, q_p(x))$ и $\bar{q} = (\bar{q}_1, \dots, \bar{q}_p)$. Можно по-разному количественно описать эту близость. Например [23], используют расстояния типа

$$d_k b_{\bar{q}} g \sum_{i=1}^{V_k} w_i |q_i - \bar{q}_i|^k \quad (9)$$

либо [48] расстояния типа

$$S b_{\bar{q}} g \min_i \alpha_i |q_i - \bar{q}_i| + \alpha_{p+1} \sum_{i=1}^p \alpha_i |q_i - \bar{q}_i| \quad (10)$$

где считается, что $q_i \geq \bar{q}_i$, α_i – коэффициенты, приводящие слагаемые к одинаковой размерности и одновременно учитывающие разноважность критериев, α_{p+1} выражает наше отношение к тому, что важнее – уменьшать близость к цели любого из частных критериев или суммарную близость всех критериев к целевым значениям. Если часть уровней притязания ограничивают критерии снизу ($q_i \geq \bar{q}_i$, $i = 1, \dots, p'$), часть ограничивают их сверху ($q_i \leq \bar{q}_i$, $i = p' + 1, \dots, p''$), а остальные задают их жестко ($q_i = \bar{q}_i$, $i = p'' + 1, \dots, p$), то функцию (10) модифицируют:

$$S b_{\bar{q}} g \min_i Z b_{\bar{q}_i} g \alpha_{p+1} \sum_{i=1}^p Z b_{\bar{q}_i} g \quad (11)$$

где

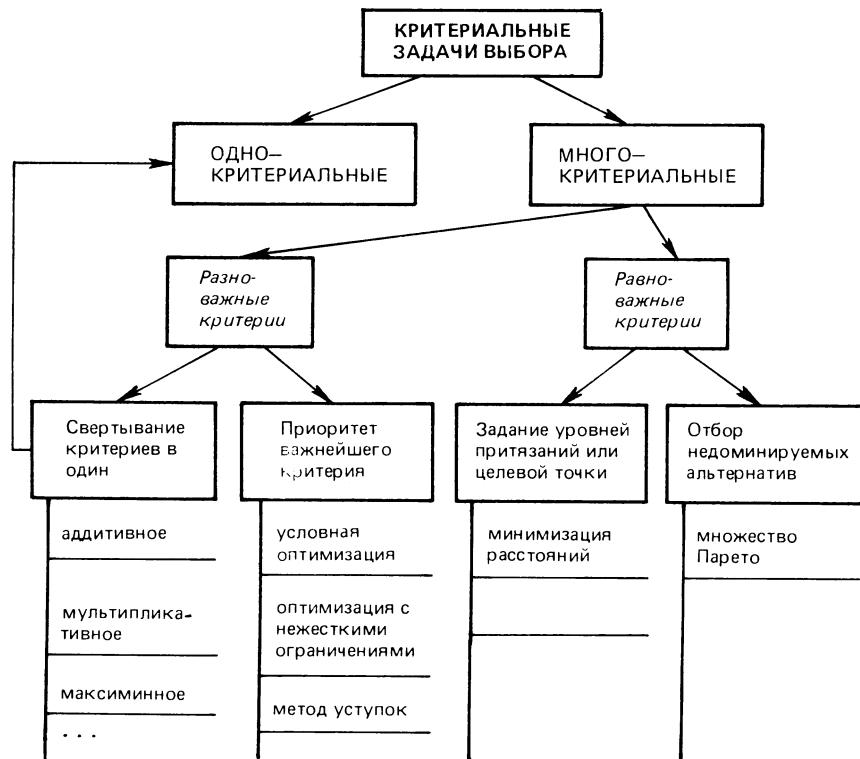
$$Z b_{\bar{q}_i} g = \begin{cases} \alpha_i |q_i - \bar{q}_i| & \text{при } 1 \leq i \leq p', \\ \alpha_i |q_i - \bar{q}_i| & \text{при } p'+1 \leq i \leq p'', \\ \alpha_i \min |q_i - \bar{q}_i| & \text{при } p''+1 \leq i \leq p. \end{cases}$$

Конечно, возможны и другие меры близости, но для функций (9) и (11) проведены подробные исследования их математических свойств, что важно для обеспечения сходимости процесса минимизации этих функций, в ходе которого обеспечивается приближение к x^* .

НАХОЖДЕНИЕ ПАРЕТОВСКОГО МНОЖЕСТВА

Четвертый полностью формализуемый способ многокритериального выбора состоит в отказе от выделения единственной “наилучшей” альтернативы и соглашении о том, что *предпочтение одной альтернативе перед другой можно отдавать только если первая по всем критериям лучше второй*. Если же предпочтение хотя бы по одному критерию расходится с предпочтением по другому, то такие альтернативы признаются несравнимыми. В результате попарного сравнения альтернатив все худшие по всем критериям альтернативы отбрасываются, а все оставшиеся несравнимые между собой (недоминируемые) принимаются. Если все максимально достижимые значения частных критериев не относятся к одной и той же альтернативе, то принятые альтернативы образуют **множество Парето** и выбор на этом заканчивается. На рис. 7.1, г жирной линией выделено множество Парето для рассматриваемого примера. При необходимости же выбора единственной альтернативы следует привлекать дополнительные соображения: вводить новые, добавочные критерии и ограничения, либо бросать жребий, либо прибегать к услугам экспертов.

Мы обсудили наиболее употребительные способы описания выбора в терминах критериального языка. Возможны и другие постановки задач на этом языке; наша цель состояла в том, чтобы дать лишь общее представление об их многообразии. Математические аспекты решения изложенных и других задач оптимизации рассматриваются в ряде монографий и учебников (см., например, [22]). Для обозримости и облегчения запоминания приведем схему совокупности изложенных способов (рис. 7.2)



7.2 Классификация задач выбора и способов их решения при их описании на критериальном языке

Подведем итог

Главный результат данного параграфа состоит в том, что для общей задачи многокритериальной оптимизации не существует единственного решения, а ее частные постановки, имеющие единственное решение, приводят к разным результатам. Поэтому лицо, принимающее решения на основе использования оптимационных методов, должно с наибольшим вниманием относиться прежде всего к постановке задачи, к тому, в какой степени именно такая постановка соответствует стоящей перед ним проблеме.

Summary

The gist of this section is that the multicriterial optimization problem have no single solution and that solutions of its particular formulations are generally different for different formulations. This is why a decision maker using the optimization technique must pay great attention to the problem formulation, being especially sensitive to the correspondence between this particular mathematical formulation and the nature of the real-life problem.

§ 7.3. ОПИСАНИЕ ВЫБОРА НА ЯЗЫКЕ БИНАРНЫХ ОТНОШЕНИЙ

Второй, более общий язык, на котором описывается выбор, – это **язык бинарных отношений**. Его большая, нежели у критериального языка, общность основана на учете того факта, что в реальности дать оценку отдельно взятой альтернативе часто затруднительно или невозможно; однако если рассматривать ее не в отдельности, а в паре с другой альтернативой, то находятся основания сказать, какая из них более предпочтительна.

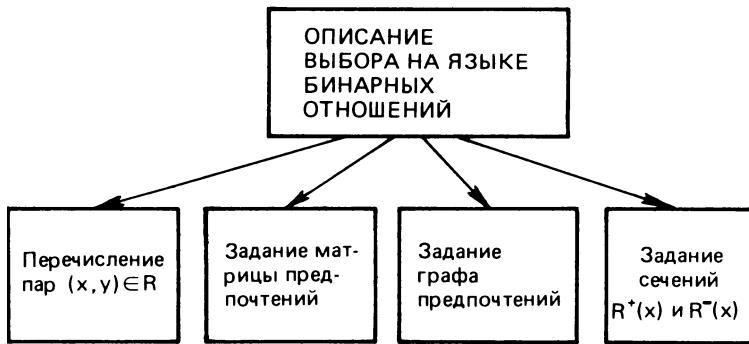
Таким образом, основные предположения этого языка сводятся к следующему:

- 1) отдельная альтернатива не оценивается, т.е. критериальная функция не вводится;
- 2) для каждой пары альтернатив (x, y) некоторым образом можно установить, что одна из них предпочтительнее другой либо они равноценны или несравнимы (чаще всего последние два понятия отождествляются);
- 3) отношение предпочтения внутри любой пары альтернатив не зависит от остальных альтернатив, предъявленных к выбору.

Математически бинарное отношение R на множестве X определяется как определенное подмножество упорядоченных пар (x, y) . Удобно использовать обозначение xRy , если x находится в отношении R с y , и $x\bar{R}y$ – в противном случае. Множество всех пар $\{(x, y), x, y \in X\}$ называется *полным* (“универсальным”) *бинарным отношением*. Поскольку в общем случае не все возможные пары (x, y) удовлетворяют условиям, накладываемым отношением R , бинарное отношение является некоторым подмножеством полного бинарного отношения, т.е. $R \subseteq X \times X$.

Задать отношение – это значит тем или иным способом указать все пары (x, y) , для которых выполнено отношение R .

СПОСОБЫ ЗАДАНИЯ БИНАРНЫХ ОТНОШЕНИЙ



7.3 ————— Способы описания выбора на языке бинарных отношений

Существует четыре разных способа задания отношений (рис. 7.3); преимущества каждого проявляются при разных характеристиках множества X .

Первый, очевидный, способ состоит в *непосредственном перечислении таких пар*. Ясно, что он приемлем лишь в случае конечного множества X .

Второй удобный способ задания отношения R на конечном множестве – *матричный*. Все элементы нумеруются, и матрица отношения R определяется своими элементами $a_{ij}(R) = \{1: x_i R x_j; 0: x_i \bar{R} x_j\}$ для всех i и j . Известным примером такого задания отношений являются турнирные таблицы (если ничьи обозначить нулями, как и проигрыш, то матрица изобразит отношение “ x_i – победитель x_j ”).

Третий способ – задание отношения *графом*. Вершинам графа $G(R)$ ставят в соответствие (пронумерованные) элементы множества X , и если $x_i R x_j$, то от вершины x_i проводят направленную дугу к вершине x_j ; если же $x_i \bar{R} x_j$, то дуга отсутствует.

Для определения отношений на бесконечных множествах используется четвертый способ – задание отношения R *сечениями*. Множество

$$R^+(x) = \{y \in X \mid (y, x) \in R\}$$

называется **верхним сечением** отношения R , а множество

$$R^-(x) = \{y \in X \mid (x, y) \in R\}$$

– **нижним сечением**. Иначе говоря, верхнее сечение – это множество всех $y \in X$, которые находятся в отношении $y R x$ с заданным элементом $x \in X$, а нижнее сечение – множество всех $y \in X$, с которыми заданный элемент x находится в отношении R . Отношение однозначно определяется одним из своих сечений.

Приведенные ниже примеры иллюстрируют все четыре способа представления конкретных отношений.

Пример 1. Полное бинарное отношение U :

1) в U входят все пары (x_i, x_j) , $x_s \in X$;

2) $a_{ij}(U) = 1$ для всех i и j ;

3) график $G(U)$ такой, что его дуги соединяют любую пару вершин (стрелки направлены в обе стороны, поскольку $x_i U x_j$ и $x_j U x_i$, а каждая вершина имеет петлю: $x_i U x_i$);

4) $R^+(x) = R^-(x) = X$ для любого $x \in X$.

Пример 2. Диагональное отношение E :

1) в E входят только пары с одинаковыми номерами: $x_i E x_j$ верно только при $i = j$;

2) $a_{ij}(E) = \{1: i = j; 0: i \neq j\}$;

3) график $G(E)$ такой, что каждая его вершина имеет петлю, а остальные дуги отсутствуют;

4) $R^+(x) = R^-(x) = x$ для любого $x \in X$.

ОТНОШЕНИЯ ЭКВИВАЛЕНТНОСТИ, ПОРЯДКА И ДОМИНИРОВАНИЯ

Для теории выбора особое значение среди всех бинарных отношений имеют отношения, которые соответствуют предпочтению одной альтернативы перед другой или случаю невоз-
442

можности отдать предпочтение одной из двух альтернатив. Эти отношения можно задать через строго определяемые отношения эквивалентности, порядка и доминирования.

Для их определения нам понадобятся некоторые свойства отношений вообще*. Бинарное отношение R на множестве X называется:

рефлексивным, если xRx для каждого $x \in X$;

антирефлексивным, если $x\bar{R}x \forall x \in X$ (т.е. R может выполняться только для несовпадающих элементов);

симметричным, если $xRy \Rightarrow yRx \forall x, y \in X$;

асимметричным, если $xRy \Rightarrow y\bar{R}x \forall x, y \in X$ (ясно, что асимметричное отношение R антирефлексивно);

антисимметричным, если для всех $x, y \in X$ ($xRy, yRx \Rightarrow x = y$);

транзитивным, если для всех $x, y, z \in X$ ($xRy, yRz \Rightarrow xRz$);

отрицательно транзитивным, если отношение \bar{R} транзитивно;

сильно транзитивным, если отношение R одновременно транзитивно и отрицательно транзитивно.

Теперь можно охарактеризовать отношения, используемые в теории выбора.

Отношение R на множестве X называется *отношением эквивалентности* (обозначение \sim), если оно рефлексивно, симметрично и транзитивно. Примеры отношений эквивалентности: “быть четным”, “иметь одинаковый остаток от деления на 3” – на множестве натуральных чисел; “быть одноклассниками” – на множестве учеников данной школы; “быть подобными” – на множестве многоугольников. Задание отношения эквивалентности равносильно разбиению множества X на непересекающиеся классы ($X = \bigcup_i X_i, X_i \cap X_j = \emptyset$ при $i \neq j$) эквивалентных элементов: $x \sim y$ тогда и только тогда, когда $x, y \in X_i$ (т.е. если x и y принадлежат одному классу эквивалентности)

Отношением нестрогого порядка (обозначение \leq) называется рефлексивное, антисимметричное и транзитивное отношение. **Отношением строгого порядка** (обозначение $<$) называется антирефлексивное, асимметричное и транзитивное отношение. Отношение нестрогого порядка можно рассматривать как объединение отношений $<$ и \sim .

Наконец, **отношением доминирования** называется отношение, обладающее антирефлексивностью и асимметричностью. Говорят, что “ x доминирует y ” (обозначается $x \gg y$), когда x в каком-то смысле превосходит y . (Очевидно, строгий порядок – частный случай доминирования, при котором имеет место еще транзитивность.)

BINARY RELATION
бинарное
отношение

PREFERENCE PATTERN граф
предпочтений
RELATION MATRIX матрица
отношения

ORDERED PAIRS
упорядоченные пары

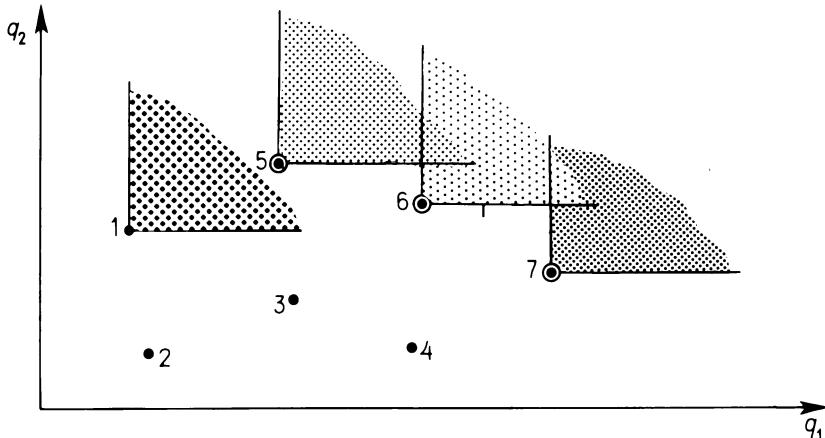
UTILITY FUNCTION
функция полезности

Обратите внимание на очень важ-

* Существует развитая теория отношений на языке теории множеств; см., например, [22; 38]. Мы не ставим задачу воспроизводить ее здесь, а лишь приведем те ее факты, которые позволяют изложить идею описания выбора на языке бинарных отношений.

Хотя при подробном рассмотрении выбора потребуются и другие факты теории отношений, введенные понятия позволяют составить представление о возможностях данного языка.

В случае конечных множеств X очень удобно находить наилучшие альтернативы с помощью графа предпочтений, стрелки которого направлены в сторону менее предпочтимой альтернативы (рис. 7.4). Выделив вершины графа, из которых стрелки только исходят (альтернативы 6 и 10 на рис. 7.4), мы находим недоминируемые, т.е. наилучшие, альтернативы. Можно показать, что если граф сильно транзитивен (т.е. транзитивен и по наличию, и по отсутствию стрелок) и антирефлексивен (отсутствуют петли), то описываемый выбор сводится к однокритериальному выбору. Другие типы графов описывают другие ситуации вы-



7.5 ————— Описание паретовского множества как множества таких альтернатив, для которых верхнее сечение $P^+(x)$ пусто.

бора.

Несмотря на то что язык бинарных отношений введен для описания более общих и сложных ситуаций, нежели те, которые описываются критериальным языком, в чисто познавательных целях поучительно проследить, как уже известная нам ситуация выглядит в новом представлении.

Например, многомерное критериальное пространство (с теми оговорками о соотношении размерностей критериев, которые были сделаны в предыдущем параграфе) может быть поставлено в соответствие евклидову пространству. Введение на этом пространстве бинарных отношений требует учета его свойств. В частности, начинают играть роль отношения и н - в а р и а н т н ы е (относительно переноса), для которых верхнее сечение в любой точке может быть получено параллельным переносом верхнего сечения в любой другой точке. Примером инвариантного отношения является *отношение Парето* P :

$$\exists_{x, y \in X} \forall_{y \in P} \Rightarrow \left\{ \exists_{j=1, m} | x_j \geq y_j \text{ и } \exists_{j=1, m} | x_{j_0} > y_{j_0} \right\}.$$

Верхнее сечение отношения P есть первый квадрант с началом в точке x ; теперь понятно, как находится паретовское множество альтернатив (на рис. 7.5 приведен случай конечного множества X ; сравните этот рисунок с рис. 7.1, 2): в паретовское множество включаются альтернативы, верхнее сечение которых пусто (на рис. 7.5 они отмечены кружками).

В общем же случае выделение наиболее предпочтительных альтернатив возможно с помощью понятия *оптимальности по отношению* R , позволяющего придавать разный смысл понятию “наилучший” (задавая разные отношения R). Элемент $x \in X$ называется *мажорантой по отношению R на X*, если для всех $y \in X$ выполнено условие $y \overline{R} x$. Множество $X_+(R)$ всех мажорант называется *множеством R-оптимальных элементов*.

ОБ ОЦИФРОВКЕ ПОРЯДКОВЫХ ШКАЛ

Важно обсудить ситуацию, возникшую при описании выбора на языке бинарных отношений в результате создания теории полезности [38]. П. Фишберн строго доказал теорему,

смысл которой довольно ясен: *если множество X конечно и между его элементами имеется отношение строгого порядка, то можно построить такую вещественную функцию $u(x)$ на X , для которой*

$$(x < y) \Rightarrow [u(x) < u(y)]$$

(в левой части $<$ означает отношение предпочтения, в правой – знак “меньше”).

Функция $u(x)$ называется *функцией полезности*. Ясно, что такая функция не единственна: произвольное монотонное преобразование сохраняет ее упорядочивающее свойство. Этот результат затем был обобщен на счетные и континуальные множества X , на нестрогий порядок и на многокритериальный случай (аддитивные функции полезности). Определение функции $u(x)$ позволяет перейти от языка бинарных отношений к критериальному языку, взяв $u(x)$ в качестве критериальной функции. Были развиты методы, позволяющие сузить класс функций полезности, например благодаря рассмотрению иерархических парных предпочтений (см. § 6.2), повышая тем самым “точность определения $u(x)$ ”.

Создается впечатление, что от качественных порядковых измерений можно перейти к количественным. На самом деле мы здесь вновь сталкиваемся с такой ситуацией, когда “оцифровка” порядковой шкалы не делает ее числовой шкалой. Для воспроизведения упорядочения фиксированного попарно упорядоченного множества X , конечно, можно воспользоваться числовой функцией $u(x)$; однако стоит дополнить X альтернативами, которые не рассматривались при первом упорядочении, как функцию $u(x)$ потребуется определять заново. Более того, если два разных эксперта дадут разные упорядочения множества X , то можно доопределить функции полезности для каждого из них, но сравнивать их численно иначе как в отношении порядка не имеет смысла (см. пример в § 6.2), хотя обе они определены на одном множестве.

В тех случаях, когда количественная величина по каким-то причинам измеряется в порядковой шкале, оцифровка порядковых данных могла бы иметь смысл. Однако во многих приложениях теории полезности мы имеем дело с измерениями, которые в принципе не могут выйти из разряда порядковых.

Подведем итог

Summary

В ряде практических случаев In many cases no criterion function exists because an evaluation критериальная функция не существует, т.е. оценку данной of a given alternative may be альтернативе можно дать только made only by comparing it with ко в результате ее сравнения с another alternative. Это a more general description of the потребовало более общего choice. The language of binary описания выбора. Первым таким обобщением и является язык бинарных отношений.

§ 7.4. ЯЗЫК ФУНКЦИЙ ВЫБОРА

Некоторые особенности выбора привели к построению третьего, еще более общего языка его описания. Во-первых, нередко приходится сталкиваться с ситуациями, когда предпочтение между двумя альтернативами зависит от остальных альтернатив. Например, предпочтение покупателя между чайником и кофеваркой может зависеть от наличия в продаже кофемолки. Во-вторых, возможны такие ситуации выбора, когда понятие предпочтения вообще лишено смысла. Например, по отношению к множеству альтернатив довольно обычными являются правила выбора “типичного”, выбора “среднего”, выбора “наиболее отличного, оригинального”, теряющие смысл в случае двух альтернатив.

ФУНКЦИИ ВЫБОРА КАК МАТЕМАТИЧЕСКИЙ ОБЪЕКТ

Язык функций выбора описывает выбор как операцию над произвольным множеством альтернатив X , которая ставит этому множеству в соответствие некоторое его подмножество

$C(X)$: $C(X) \subseteq X$. (Обозначение связано с первой буквой английского слова choice – “выбор”.) Функция выбора как отображение совокупности множеств в совокупность множеств (поскольку для выбора могут предлагаться любые подмножества $X_i \subseteq X$) без поэлементного отображения одного множества на другое и без отображения множеств на числовую ось является своеобразным и пока еще не полно изученным математическим объектом. Конечно, накладывая на функцию выбора определенные требования, мы можем на этом языке описывать и те варианты выбора, которые отражаются в предыдущих языках. Однако главное достоинство нового языка – возможность рассмотрения более сложных правил выбора. На такую возможность указывает хотя бы различие числа возможных функций выбора и числа возможных графов предпочтения на множестве n альтернатив. Число графов, отличающихся наличием или отсутствием хотя бы одной дуги, равно 2^{n^2} . Если для выбора предлагаются k из n альтернатив, то число функций выбора равно 2^k (каждая из альтернатив может либо входить в $C(X_k)$, либо нет). Так как число возможных вариантов предъявления альтернатив равно C_n^k , то общее число функций выбора равно $\prod_{k=1}^n 2^{C_n^k} = 2^{n^{2^{n-1}}}$. Как видим, разнообразие функций выбора намного превосходит разнообразие графов предпочтения. Кроме того, здесь сразу допускается отказ от выбора, т.е. пустой выбор $C(X_i) = \emptyset$ что также расширяет множество правил выбора.

ОГРАНИЧЕНИЯ НА ФУНКЦИИ ВЫБОРА

Различие между классами правил выбора можно выразить через различные ограничения, которым подчиняется тот или иной тип функции выбора. Отдельные ограничения и их комбинации дают уже известные нам правила выбора, другие определяют новые правила, которые предстоит изучить. (По разным причинам наиболее подробному рассмотрению подвергались именно те функции выбора, которые идентичны выбору, описываемому на предыдущих языках [22].) Приведем некоторые из таких ограничений:

Аксиома наследования (Н):

$$X' \subseteq X \Rightarrow C(X') \supseteq C(X) \cap X'.$$

Смысл этой аксиомы сводится к требованию, чтобы в выбор на подмножестве X' вошли все те альтернативы из X' , которые входили в выбор на X (возможно, еще и другие; рис. 7.6, а).

Аксиома согласия (С):

$$\bigcap_i C(X_i) \subseteq C(\bigcup_i X_i).$$

Это означает, что в выбор из объединения множеств обязательно должны входить альтернативы, общие для выборов из всех множеств (и, возможно, другие альтернативы; рис. 7.6, б).

Оказывается, совместное подчинение функции выбора аксиомам Н и С дает выбор, описываемый в языке бинарных отношений.

Аксиома отбрасывания (О):

$$C(X) \subseteq X' \subseteq X \Rightarrow C(X') = C(X).$$

Это означает, что если отбросить любую часть отвергнутых при выборе альтернатив, то выбор на оставшемся множестве не изменится (рис. 7.6, в); поэтому данную аксиому называют также *условием независимости от отвергнутых альтернатив*.

Совместное наложение на выбор аксиом Н, С и О приводит к случаю выбора паретовского множества.

Аксиома Плотта (КС):

$$C(X_1 \cup X_2) = C(C(X_1) \cup C(X_2)).$$

Это отражает требования, накладываемые при многоступенчатых выборах, когда считается, например, что определить чемпиона мира можно путем соревнований между чемпионами стран и результат окажется тем же, если соревноваться будут не только чемпионы (рис. 7.6, г). Поэтому эту аксиому называют еще *условием независимости от пути*. Функции выбора, удовлетворяющие ей, называются *квазисумматорными*.

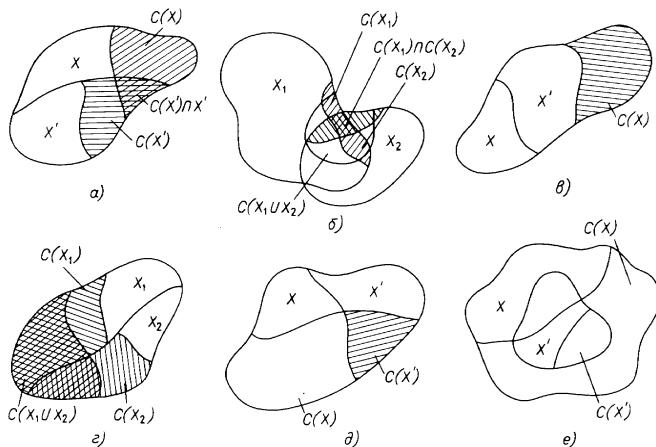
Можно показать, что требование КС эквивалентно совместному выполнению Н и О; следовательно, соединение требований КС и С тоже приводит к паретовскому выбору.

Аксиома предпочтения (П):

$$X' \subseteq X \Rightarrow C(X) \cap X' = C(X').$$

Она требует, чтобы при сужении множества альтернатив в выборе оставались только те альтернативы, которые входили в выбор ранее (рис. 7.6, д). Это столь жесткое ограничение, что оно эквивалентно скалярному критериальному выбору.

Ясно, что некоторые из введенных аксиом можно ослаблять или усиливать (например, П есть усиление Н). Аксиому Плотта можно усилить до **аксиомы сумматорности**: $C(X_1 \cup X_2)$



7.6 Иллюстрация различных аксиом, накладываемых на функции выбора

$= C(X_1) \cup C(X_2)$; можно накладывать новые, независимые требования (например, **аксиома мультипликаторности** $C(X_1 \cap X_2) = C(X_1) \cap C(X_2)$, **аксиома монотонности** $X_1 \subseteq X_2 \Rightarrow C(X_1) \subseteq C(X_2)$; рис. 7.6, е), получая при этом различные типы выбора. Наоборот, можно, изучив ограничения того или иного реального правила выбора, искать свойства класса функций выбора, удовлетворяющего этим ограничениям (попробуйте, например, установить, каким аксиомам отвечает и противоречит правило определения победителя шахматного турнира по числу набранных очков). Айзерман с сотрудниками начал изучение классов функций выбора [2]; эту работу пока нельзя считать законченной.

Подведем итог

Summary

Язык функций выбора является весьма общим и потенциально может описать любой выбор. Однако его теория находится в начальной стадии развития и пока еще занимается преимущественно описанием старых ситуаций в новых терминах

The language of choice functions is very general and can potentially describe any type of choice. However, its theory is only beginning to be developed and is still occupied with describing old situations in new terms.

§ 7.5. ГРУППОВОЙ ВЫБОР

В человеческом обществе единоличное принятие решений является не единственной формой выбора. “Ум – хорошо, а два – лучше”, гласит поговорка, имеющая в виду тот случай, когда оба ума с одинаковыми намерениями пытаются найти хороший выбор. Этот случай мы и рассмотрим в данном параграфе (выбор в условиях конфликта будет рассмотрен в дальнейшем).

ОПИСАНИЕ ГРУППОВОГО ВЫБОРА

Итак, пусть на множестве альтернатив X задано n в общем случае различных индивидуальных предпочтений (для определенности будем говорить о бинарных отношениях) R_1, R_2, \dots, R_n . Ставится задача о выработке некоторого нового отношения R , которое согласует индивидуальные выборы, выражает в каком-то смысле “общее мнение” и принимается за групповой выбор. Очевидно, что это отношение должно быть какой-то функцией индивидуальных выборов: $R = F(R_1, \dots, R_n)$. Различным принципам согласования будут отвечать разные функции F . В принципе, т.е. теоретически, функции F могут быть совершенно произвольными, учитывать не только индивидуальные выборы, но и другие факторы, в том числе

VOTING голосование

GROUP CHOICE групповой выбор

CONSENSUS консенсус

MAJORITY RULE правило большинства

CHOICE FUNCTION функция выбора

Язык функций выбора (“глобальных функций множеств”) описывает результат выбора как некоторое подмножество множества альтернатив. Такое соответствие двух множеств без их поэлементного соответствия является новым понятием, расширяющим смысл термина “функци-

и исход некоторых случайных событий (например, бросания жребия), и главный вопрос состоит в том, чтобы правильно отобразить в функции F особенности конкретного варианта реального группового выбора.

РАЗЛИЧНЫЕ ПРАВИЛА ГОЛОСОВАНИЯ

Один из наиболее распространенных принципов согласования – **правило большинства**: *принятой всеми считается альтернатива, получившая наибольшее число голосов*. Правило большинства привлекательно своей простотой и демократичностью, но имеет особенности, требующие осторожного обращения с ним. Прежде всего оно лишь обобщает индивидуальные предпочтения, и его результат не является критерием истины. Только дальнейшая практика показывает, правильным или ошибочным было решение, принятое большинством голосов; само голосование – лишь форма согласования дальнейших действий*. Во-вторых, даже в простейшем случае выбора одной из двух альтернатив легко представить себе ситуацию, когда правило большинства не срабатывает: разделение голосов поровну при четном числе голосующих. Это порождает варианты: “председатель имеет два голоса”, “большинство простое (51 %)”, “подавляющее большинство (около 3/4)”, “абсолютное большинство (близкое к 100 %)”, наконец, “принцип единогласия (консенсус, право вето)”.

Подчеркнем, что при любом из этих вариантов подразумевается отказ от принятия решения, если ни одна из альтернатив не получила необходимого процента голосов. Поскольку в реальной жизни отказ от дальнейших действий, следующих за решением, бывает недопустим, а переход к принятию за групповой выбор выбора отдельного лица (“диктатора”) – нежелательным, разрабатываются различные приемы, сокращающие число ситуаций, приводящих к отказу.

Например, если два эксперта дали противоположные предпочтения между двумя вариантами a и b , то можно сделать выбор, сравнивая “ силу предпочтения ” каждого эксперта. При возможности введения количественного критерия оценки это сводится к арифметической операции, но и при порядковом сравнении есть воз-

* Этот вопрос, однако, формально находится вне нашего рассмотрения: ведь пока мы обсуждаем выбор в условиях определенности, а здесь речь зашла о выборе определенной альтернативы с неопределенными последствиями. Такие задачи мы рассмотрим позже.

можность оценки “силы предпочтения”. В криминалистической практике в таких случаях экспертам предлагается в одном ряду с a и b упорядочить по предпочтению еще несколько альтернатив, скажем c , d и e . Пусть первый эксперт дал упорядочение (c, d, a, b, e) , а второй – (b, c, d, e, a) . Тогда можно сделать вывод, что степень предпочтения b по сравнению с a у второго эксперта больше, чем степень предпочтения a перед b у первого, и принять решение в пользу b (за этим приемом стоит ряд предположений – сравнимость интенсивностей предпочтений, одинаковая компетентность экспертов и т.д., требующих проверки в ответственных случаях).

Даже для консенсуса, требующего единогласия, разработаны приемы, облегчающие его достижение. Так, Р. Акофф отмечает:

“Консенсус часто трудно достичь, но редко невозможен. Я обнаружил, что в трудных случаях очень эффективна следующая процедура. Первое – максимально уточнить формулировки альтернатив, между которыми консенсус не допускает выбора. Второе – коллективно построить тест эффективности альтернатив и принять консенсусом решение, что данный тест справедлив и что все согласны следовать его результату. Третье – провести тест и использовать его результат. Я смог успешно применить эту процедуру даже в таком случае, когда законодатели одного государства не могли прийти к согласию, вводить или нет смертную казнь за убийство. В результате обсуждения члены законодательного органа пришли к согласию, что они все имеют одну цель – минимизировать число жертв убийств. Как только такое согласие было достигнуто, проблема была сведена к конкретному вопросу: уменьшает ли введение смертной казни число убийств? Все согласились, что необходимо провести исследование, отвечающее на этот вопрос. Такое исследование было проведено и его результаты использованы (оно показало, что число убийств в ряде государств до и после отмены или введения смертной казни заметно и значимо не изменилось)” [44].

Если же не удавалось достичь консенсуса не только по поводу самих альтернатив, но и относительно способа их проверки, то, по мнению Акоффа, следует найти консенсусное решение, что же делать дальше. Интересно его наблюдение, что в таких случаях обычно принималось решение поручить выбор одному из авторитетных и ответственных лиц (мы еще вернемся к этому моменту). Фактически это переход от демократического, но не давшего решения правила голосования к недемократическому, но приводящему к какому-то решению “диктаторскому” принципу. Такое экспериментальное наблюдение имеет и теоретическое объяснение, излагаемое ниже на качественном уровне.

ПАРАДОКСЫ ГОЛОСОВАНИЯ

Итак, следующая особенность правила голосования – это *возможность отказа от выбора из-за недостижения требуемого большинства*. Казалось бы, исключив такую возможность, можно обеспечить принятие решения в любых случаях. Например, пусть три эксперта большинством голосов решают вопрос, какая из двух альтернатив более предпочтительна. При такой постановке вопроса они действительно не могут не сделать выбор. Однако здесь мы приходим к еще одной особенности правила голосования – его нетранзитивности.

Пусть, например, каждая из трех группировок законодателей, образующих большинство лишь попарно, выдвинули свой вариант законопроекта: a , b и c . Или три парня заспорили, чья девушка лучше, и намереваются решить спор голосованием. Чтобы гарантировать большинство на каждом шаге процедуры, альтернативы предъявляются попарно. Каждая сторона руководствуется при этом своим набором предпочтений; пусть это соответственно последовательности $(a > b > c)$, $(b > c > a)$ и $(c > a > b)$.

После голосования по паре (a, b) в результате получаем два голоса против одного: $a > b$; по паре (b, c) имеем $b > c$; по паре (c, a) имеем $c > a$. Голосование большинством не привело к выяснению “общепризнанного” порядка альтернатив: $a > b > c > a$. В случае же применения процедуры, при которой после рассмотрения очередной пары отвергаемая альтернатива заменяется новой, окончательно принятое решение зависит от порядка предъявления альтер-

натив: при порядке (a, b, c) выбирается c ; при порядке (b, c, a) выбор остановится на a ; при порядке (a, c, b) – на b . Если таким образом принять законопроект, то чье мнение он будет выражать – большинства или организатора голосования? Очевидно, что такие решения не отвечают идеалу согласованного группового выбора.

Причина данного парадокса нетранзитивности группового выбора состоит, конечно, в цикличности совокупности исходных индивидуальных предпочтений. Однако это лишь частный пример более общего явления, получившего название **парадокса Эрроу (или теоремы о невозможности)**. Не вдаваясь в подробности этой теоремы и ее доказательства, изложим ее смысл.

Из всевозможных функций F индивидуальных выборов R_1, \dots, R_n выделим те, которые отвечают требованиям, выражющим наше понимание того, какой выбор можно считать *согласованным*. Кроме формальных требований

$1^0)$ “ $n \geq 2$ ”, “число альтернатив ≥ 3 ”, “ F определена для любых $\{R_i\}$ ”
естественно также потребовать, чтобы:

$2^0)$ если в результате группового выбора предпочтение было отдано альтернативе x , то это решение не должно меняться, если кто-нибудь из ранее отвергавших x изменил свое предпочтение в его пользу (*условие монотонности*);

$3^0)$ если изменения индивидуальных предпочтений не коснулись определенных альтернатив, то в новом групповом упорядочении порядок этих альтернатив не должен меняться (*условие независимости несвязанных альтернатив*);

$4^0)$ для любой пары альтернатив x и y существует такой набор индивидуальных предпочтений, для которого $F(R_1, \dots, R_n) = (x > y)$ (*условие суверенности*; без него возможно навязывание y независимо от порядков предпочтений индивидуумов);

$5^0)$ не должно быть такого индивидуума, для которого из его предпочтения $x > y$ (при любых x и y) вытекает, что $F(R_1, \dots, R_n) = (x > y)$ независимо от предпочтений других индивидуумов (*условие отсутствия диктаторства*).

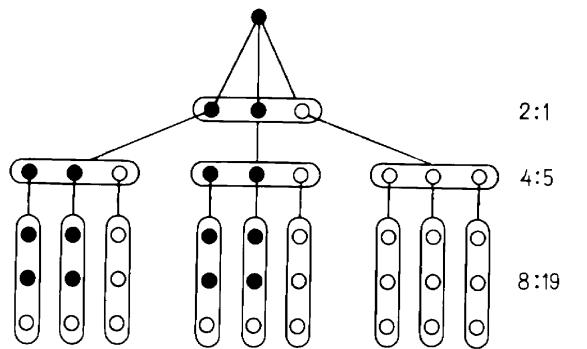
Парадокс Эрроу состоит в том, что первые четыре условия противоречат пятому; не существует правила F , удовлетворяющего всем пяти требованиям. Анализ причин такого обескураживающего следствия из столь невинных на вид предположений показывает [21; 24], что основную роль играет возможность циклических множеств ранжирований, что характерно для бинарных отношений, удовлетворяющих условию 3^0 .

Нетранзитивность мажоритарного отношения может проявляться и в других неожиданных формах. Рассмотрим такую задачу [24]. Пусть каждый из n субъектов имеет свою долю a_i общего ресурса $a = \sum_{i=1}^n a_i$. Вектор $\bar{a} = (a_1, \dots, a_n)$ назовем состоянием системы. Другое состояние $\bar{b} = (b_1, \dots, b_n)$ с точки зрения i -го субъекта хуже \bar{a} , если $a_i \geq b_i$. Будем теперь производить перераспределение ресурсов на основе очень сильного большинства: система перейдет из \bar{a} в \bar{b} , если состояние \bar{b} не хуже \bar{a} для всех, кроме одного (“тотально-мажоритарное правило”). Последовательность состояний $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_k$ будем называть тотально-мажоритарным путем из \bar{a}_1 в \bar{a}_k , если переходом в очередное состояние удовлетворены все участники, кроме, естественно, того, чей ресурс в данный момент перераспределяется. Пусть теперь заданы два произвольных состояния системы: \bar{a} и \bar{b} . При каких условиях существует тотально-мажоритарный путь из \bar{a} в \bar{b} ? Оказывается, что такой путь существует всегда. Снова имеем дело с парадоксом: возможны любые перераспределения, и все они выражают мнение “всего общества”, кроме одного субъекта (правда, эти “несогласные” на разных этапах различны).

Задачи группового выбора часто все же могут быть разрешены. Во-первых, в ряде случаев циклические ранжирования могут отсутствовать, либо они не охватывают “наиболее важные” альтернативы, либо принимаются меры по их обнаружению и устраниению. Во-вторых, во многих случаях “диктаторский” принцип согласования не является неприемлемым. Это

илюстрируется примером оптимизации по “главному” из нескольких критериев. В других случаях это единственно возможный принцип (например, единоначалие в армии). В-третьих, переход (когда это возможно) к использованию единой числовой, а не порядковых индивидуальных шкал предпочтений может вообще аннулировать проблему нетранзитивности. В-четвертых, в реальных ситуациях мажоритарные правила применяются в комбинации с другими правилами, так что, образовав, например, коалицию, группы субъектов могут блокировать действие голосования.

Здесь мы приходим к еще одной особенности голосования, которую следует иметь в виду на практике. Речь идет о вмешательстве коалиций в механизм голосования, фактически ме-



7.7 ————— Иллюстрация парадокса многоступенчатого голосования при наличии коалиции

няющем его характер. Например, при многоступенчатом голосовании по правилу большинства коалиция, находящаяся в меньшинстве, может добиться принятия своего решения. На рис. 7.7 изображено голосование по три большинством в 2/3 на каждой ступени. Видно, что уже на второй ступени меньшинство может навязывать свое мнение большинству. Если число ступеней не ограничивать, то теоретически побеждающее таким образом меньшинство может быть сколь угодно малым. То, что при многоступенчатом голосовании может побеждать кандидат, не набравший действительного большинства голосов, происходит и в действительности. Например, в 1876 г. президентом США был избран Р.Б. Хейес (185 голосов выборщиков), а не С. Дж. Тилден (184 голоса), хотя на долю последнего пришлось 51% голосов всех избирателей. Такие же ситуации имели место в президентских выборах 1874 и 1888 гг.

PAYOUT MATRIX матрица платежей
SADDLE POINT седловая точка
MIXED STRATEGY смешанная стратегия
GAMES THEORY теория игр
LOSS FUNCTION функция потерь

Если вплоть до момента осуществления выбора остается неизвестным точно, какой именно из возможных исходов реализуется после выбора, то что же выбирать?

Подведем итог

Общественная практика ставит проблему группового выбора, в котором могли бы принять участие все члены сообщества. Один из простых и популярных способов коллективного выбора – голосование. Это широко применяемая и во многих случаях успешная процедура наряду с очевидными достоинствами обладает рядом скрытых особенностей, которые могут ослабить и даже извратить демократический характер голосования. Знание таких “парадоксов” голосования необходимо не только специалистам по системам.

Summary

The ideals of democracy give rise to the problem (among others) of group choice, in which all members of society can take part. One simple and popular procedure of collective choice is voting. This widely used and often successful procedure has, besides its evident virtues, some latent features that could weaken and even pervert the democracy of voting. Knowledge of such “paradoxes” of voting is necessary, and not only for systems scientists.

§ 7.6. ВЫБОР В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

До сих пор мы обсуждали подходы к описанию и осуществлению выбора в таких условиях, когда последствия сделанного выбора были определены однозначно. Выбор одной из альтернатив $x \in X$ был связан с известным выбирающему однозначным следствием, и вся проблема выбора – это проблема сравнения разных следствий (или, что в данном случае то же самое, альтернатив).

ЗАДАНИЕ НЕОПРЕДЕЛЕННОСТИ С ПОМОЩЬЮ МАТРИЦЫ

В реальной практике нередко приходится иметь дело с более сложной ситуацией, когда выбор альтернативы неоднозначно определяет последствия сделанного выбора: имеется набор возможных исходов $y \in Y$, из которых один окажется совмещенным с выбранной альтернативой, но какой именно – в момент выбора неизвестно, а станет ясным позже, когда выбор уже сделан и изменить ничего нельзя. Хотя с каждой альтернативой x связано одно и то же множество исходов Y , для разных альтернатив одинаковые исходы имеют разное значение. В случае дискретного набора альтернатив и исходов такую ситуацию можно изобразить с помощью следующей матрицы:

$X \setminus Y$	y_1	y_2	...	y_j	...	y_m
x_1	q_{11}	q_{12}	...	q_{1j}	...	q_{1m}
?
x_i	q_{i1}	q_{i2}	...	q_{ij}	...	q_{im}
?
x_n	q_{n1}	q_{n2}	...	q_{nj}	...	q_{nm}

В этой матрице все возможные исходы образуют вектор $y = (y_1, \dots, y_m)$, числа q_{ij} выражают оценку ситуации, когда сделан выбор альтернативы x_i и реализовался исход y_j . В разных случаях числа q_{ij} могут иметь различный смысл: иногда это “выигрыши”, иногда “потери”, “платежи”; в литературе употребляются также и другие названия.

Если все строки $q_i = (q_{i1}, \dots, q_{im})$ при любых i одинаковы, то проблемы выбора между альтернативами нет. Если же строки матрицы различны, то возникает вопрос, какую альтернативу предпочесть, не зная заранее, какой из исходов реализуется.

Аналогично, в случае непрерывных множеств X и Y ситуация описывается с помощью задаваемой на этих множествах функции $q(x, y)$, $x \in X$, $y \in Y$ с соответствующей постановкой вопроса о выборе x .

Сказанного до сих пор недостаточно для формальной постановки задачи выбора. При различной конкретизации этой задачи она приобретает различный смысл и требует различных методов решения. Исторически сложилось так, что первыми были formalизованы искусственные, игровые задачи, что придало всей терминологии несколько легкомысленное звучание (взаимодействующие стороны называются “игроками”, выбираемые ими альтернативы – “ходами”, правила выбора – “стратегиями”, величины q_{ij} – “выигрышами”, а вся теория – “теорией игр”).

Один класс задач называется “играми против природы”. В таких задачах считается, что исходы y_1, \dots, y_m есть возможные “состояния природы”. Желательность каждой альтернативы x_i зависит от того, каково состояние природы, но узнать, каково оно, мы сможем лишь после того, как сделаем выбор.

В другом классе задач предполагается, что исходы Y – это множество альтернатив, на котором выбор осуществляет второй игрок. В отличие от бесстрастной Природы второй игрок преследует свои интересы, отличные от интересов первого игрока. При этом матрица $Q = \|q_{ij}\|$, характеризующая оценки ситуаций с точки зрения игрока, выбирающего x , уже недостаточна для описания всей игры. Необходимо задать вторую матрицу $U = \|u_{ij}\|$, описывающую игру с позиций второго игрока. Задание X, Y, Q и U называется *нормальной формой игры*. Расхождения между матрицами Q и U определяют степень антагонизма игроков. Если $a_{ij} + u_{ij} = \text{const}$ для всех i и j , то соперничество называется *строгим*. В случае $q_{ij} + u_{ij} = 0$ имеем *игру с нулевой суммой*. Можно представить себе игры, где выигрыши и проигрыши сторон не связаны линейно и это будет отражать усиление или ослабление конфронтации сторон. Можно также рассматривать изменение матриц платежей после очередного хода. Например, интерес исследователей привлекли игры с нарастающей конфликтностью; примером может служить игра “в цыпленка”*, но, конечно, у такой задачи есть и более серьезные приложения. Возможны и другие обобщения, например рассмотрение игр с участием большего числа участников, с образованием коалиций между ними и т.д.

Разнообразие задач выбора в условиях неопределенности существенно возрастает в связи с тем, что и сам характер неопределенности может быть различным. Например, в “игре против природы” можно считать, что состояние природы “совершенно неизвестно”, а можно ввести на множестве Y вероятностную меру, что даст основания для усиления различий между исходами; такие разные постановки задач дают, естественно, и различные их решения.

КРИТЕРИИ СРАВНИВАНИЯ АЛЬТЕРНАТИВ ПРИ НЕОПРЕДЕЛЕННОСТИ ИСХОДОВ

Вряд ли возможно (да и целесообразно) в кратком обзоре рассмотреть все важнейшие результаты теории игр (опубликовано много монографий; интересующимся можно рекомендовать начать изучение с книг [7; 21; 29]). Однако об основных идеях и подходах к решению задач теории игр желательно иметь представление всем, кому придется проводить исследования систем.

Центральным моментом является введение критерия для оценки выбранного варианта. В силу неопределенности исхода нужно дать оценку сразу целой строке платежной матрицы; имея такие оценки для всех строк и сравнивая их, мы и можем делать выбор.

Самым распространенным является критерий выбора “наименьшего из зол”, называемый **максиминным критерием**. В каждой из строк матрицы платежей находится наименьший выигрыш $\min_j q_{ij}$, который характеризует гарантированный выигрыш в самом худшем случае и считается оценкой альтернативы x_i . Теперь остается найти альтернативу x^* , обеспечиваю-

* Эта игра хулиганствующих водителей в США начинается с выхода автомобиля на разделяющую шоссе полосу. Желающий принять вызов встречный водитель также выходит на полосу. Проигравший (“цыпленок”) – тот, кто отворачивает с полосы на свою сторону первым.

шую наибольшее значение этой оценки: $x^* = \arg \max_i \min_j q_{ij}$. Эта альтернатива и называется оптимальной по максиминному критерию. Поскольку часто платежную матрицу определяют не через выигрыш, а через проигрыш, тот же принцип приводит к **минимаксному критерию**.

Минимаксный критерий является крайне осторожным, очень пессимистическим, поэтому были предложены другие критерии. Таков, например, **критерий минимаксного сожаления**, предложенный Сэвиджем. При этом по платежной матрице Q вычисляется “матрица сожалений” S , элементы которой определяются как $s_{ij} = q_{ij} - \min_i q_{ij}$, и минимаксный критерий применяется к матрице S : $x^* = \arg \min_i \max_j s_{ij}$.

Дальнейшее ослабление пессимистичности оценки альтернатив дает **критерий пессимизма – оптимизма (критерий Гурвица)**, который сводится к взвешенной комбинации наилучшего и наихудшего исходов. А именно: за оценку альтернативы x_i в критерии Гурвица принимается величина

$$g(x_i) = \alpha \min_j q_{ij} + (1 - \alpha) \max_j q_{ij}, \quad 0 \leq \alpha \leq 1.$$

Показатель α называется показателем пессимизма – оптимизма (при $\alpha=1$ имеем снова максиминный критерий); оптимальная альтернатива есть $x^* = \arg \max_i g(x_i)$.

ОБЩЕЕ ПРЕДСТАВЛЕНИЕ О ТЕОРИИ ИГР

Некоторые особенности игровых ситуаций хорошо видны на простейшем примере. Пусть имеется игра с континуальными множествами X и Y , строгим соперничеством сторон и нулевой суммой. Это делает достаточным рассмотрение лишь одной функции платежей $q(x, y)$, которую один игрок старается максимизировать по x , а другой – минимизировать по y .

В тех случаях, когда $\max_{x \in X} \min_{y \in Y} q(x, y) = \min_{y \in Y} \max_{x \in X} q(x, y)$, точка (x^*, y^*) , в которой достигается это равенство, одновременно удовлетворяет амбиций обоих игроков. Эта точка равновесия интересов сторон называется *седловой*. Отход от этой точки невыгоден обеим сторонам, так что ее выбор решает игру.

Однако существуют игры без седловой точки. В такой ситуации становится выгодным скрывать от противника свой выбор и даже свой способ выбора. Это достигается введением *смешанной* стратегии. В отличие от *чистой* стратегии, при которой альтернатива выбирается однозначно по детерминированному правилу, смешанная стратегия состоит в том, что задаются лишь вероятности выбора альтернатив, а сам выбор осуществляется случайным механизмом, подчиняющимся заданному распределению. В результате получаемый выигрыш становится случайной величиной и сравнение стратегий можно вести через средние значения выигрыша. Оказывается (теорема фон Неймана), что любые матричные игры со строгим соперничеством имеют решение в смешанных стратегиях. Кроме того, матричную игру можно свести к задаче линейного программирования, что дает не только практические методы численного решения игр, но и позволяет перенести ряд теоретических результатов из теории программирования в теорию игр.

Подведем итог

Summary

Неопределенность в момент выбора характеризуется распределением потерь и выигрышей по исходам, связанным с каждой альтернативой. Вводя подходящую числовую характеристику этого распределения, мы получаем возможность упорядочения (сравнения) альтернатив. Разнообразие задач теории игр связано с разными степенями конфликтности между сторонами, с другими особенностями конкретных задач.

The distribution of gains and losses among the outcomes for each alternative defines the choice under uncertainty. The introduction of a proper numerical characterization of the distribution enables us to order (i.e., to compare) the alternatives. The variety of game theory problems stems from differences in the numerical characterizations of loss distributions, in the various levels of conflict between the playing sides, and in other characteristics of particular problems.

§ 7.7. О ВЫБОРЕ В УСЛОВИЯХ СТАТИСТИЧЕСКОЙ НЕОПРЕДЕЛЕННОСТИ

Существует класс задач выбора, особенностью которых является наличие неопределенности даже после того, как проведена серия наблюдений, измерений. Дело в том, что данные, полученные в результате эксперимента, связаны с интересующим нас аспектом явления не непосредственно, не однозначно, а в совокупности с другими, неконтролируемыми факторами.

СТАТИСТИЧЕСКИЕ РЕШЕНИЯ КАК ВЫБОР

Пусть, например, требуется знать высокоточное значение веса некоторого предмета. Неоднократное его взвешивание на аналитических весах даст хотя и близкие, но разные значения, так как на показания весов оказывают влияние не только вес самого взвешиваемого предмета, но и трение, неидеальность геометрической формы опорной призмы, течение струй воздуха, тепловой режим и т.д. Аналогично, при радиолокационном зондировании Луны в результате каждого измерения получают отличающиеся цифры для расстояния до нашего спутника, чему имеется множество причин. В таких задачах возникает нетривиальная проблема выбора, какое именно из значений интересующей нас величины принять за истинное – с учетом имеющихся данных.

Аналогичная ситуация – выбор в условиях статистической неопределенности – имеет место не только при оценке некоторой величины, но и при классификации объектов (чем болен пациент, если его состояние характеризуется такими-то данными анализов? На какой глубине находится нефть или вода, если в результате геофизической разведки получена некоторая совокупность чисел?), а также при необходимости подобрать математическую модель явления (на какую кривую лучше всего ложатся эти экспериментальные данные?) и при необходимости обнаружить какую бы то ни было закономерность (влияет ли солнечная активность на здоровье людей? С какими наблюдаемыми характеристиками жизни человека связана возможность заболевания шизофренией? Есть ли закономерность в том, как в течение 30 лет поражала молния опоры данной линии электропередачи?).

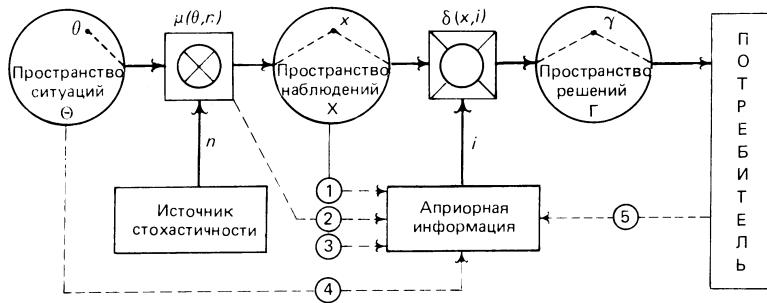
Во всех таких задачах есть общее – необходимость выбора на основании косвенных или прямых, но обязательно “зашумленных” данных. Основным, центральным, самым важным предположением для формализации решения таких задач является предположение о статистичности экспериментальных данных. Оно состоит в том, что связь между истинной, но неизвестной искомой альтернативой θ (будем обозначать этой буквой любую закономерность, отыскиваемую в протоколе наблюдений, считая, что она принадлежит множеству Θ возможных закономерностей, на котором и надо сделать выбор) и наблюдаемыми данными

x_1, x_2, \dots, x_N адекватно описывается распределением вероятностей (например, функцией распределения $F(x_1, \dots, x_N | \theta)$ или, если x_i – непрерывные величины, а функция F дифференцируема, – плотностью вероятностей $f(x_1, \dots, x_N | \theta)$). Другими словами, считается, что, во-первых, выборка наблюдений принадлежит статистическому ансамблю всевозможных выборок, на котором задано распределение вероятностей, и, во-вторых, это распределение различно для разных θ , что и обеспечивает наличие информации о θ в выборке x_1, \dots, x_N (см. гл. 5). Вопрос состоит в том, как извлечь эту информацию, т.е. как сделать выбор на множестве Θ или как принять статистическое решение.

Естественно, напрашивается идея – свести задачу к уже решенной ранее. Такую возможность предоставляет теория “игр против природы” (см. § 7.6): выбор $\hat{\theta}$ на Θ и действительное состояние θ природы можно в совокупности охарактеризовать функцией потерь $l(\hat{\theta}, \theta)$, которую и рассматривать как платежную функцию игры. Использование такого представления в самом деле позволяет перенести ряд результатов теории игр в теорию статистических решений [5]. Однако одни статистические задачи не принадлежат к числу решенных задач теории игр, для других задач существуют более прямые и короткие способы решения, третьи имеют настолько сильно выраженную специфику обработки экспериментальных данных, что игровая терминология лишь “затемняет” суть дела; наконец, теория синтеза и анализа процедур для решения статистических задач – математическая статистика – начала развиваться задолго до возникновения теории игр, достигла значительных результатов и продолжает успешно развиваться самостоятельно. По этим причинам теоретико-игровой подход к статистике, сохраняя свое методологическое, междисциплинарное значение, не оказал существенного влияния на прикладную статистику.

ОБЩАЯ СХЕМА ПРИНЯТИЯ СТАТИСТИЧЕСКИХ РЕШЕНИЙ

Вернемся к рассмотрению черт, общих для задач выбора в условиях статистической неопределенности. Обстоятельства принятия статистических решений иллюстрирует схема, приведенная на рис. 7.8 [33]. На этой схеме точкой $\theta \in \Theta$ изображено то, что нам неизвестно, но необходимо определить; Θ – множество всех предполагаемых возможностей относительно θ . Точкой $x \in X$ изображена выборка (протокол наблюдений) $x = (x_1, \dots, x_N)$; X – множество всех возможных выборок. Тот факт, что на реализовавшееся значение выборки оказывает влияние не только искомая закономерность θ , но и совокупность случайных факторов, изображен на схеме как результат совместного отображения θ и некоторого случайного воздействия n в пространство X с помощью некоторого оператора ?: $x = ?(\theta, n)$. Зная x , мы должны сделать выбор относительно θ , принять решение, какую из множества альтернатив Θ мы примем за истинную. Чтобы не путать принимаемое решение и “истинное” состояние θ , обозначим пространство, на котором производится выбор, через Γ . Очевидно, что в Γ входят все элементы множества Θ , но могут войти и дополнительные решения (типа отказа от выбора, требования увеличить число наблюдений или провести рандомизацию и т.п.). Процедура выбора изображена как действие некоторого оператора ? над выборкой x : каждой выборке x этот оператор, называемый *решающей функцией*, ставит в соответствие решение ??= ?(x, i). Здесь аргумент i введен, во-первых, для того чтобы подчеркнуть, что одну и ту же выборку можно обрабатывать по-разному, получая решения различного качества, и, во-вторых, чтобы сделать акцент на том, что качество решения зависит не только от того, какой протокол обрабатывается, но и от того, какие априорные предположения вошли в структуру алгоритма.



7.8 ————— Общая схема принятия статистических решений

Итак, и проблема синтеза статистических процедур (построения решающих функций), и проблема анализа их качества (оценивания степени близости между $\hat{\theta}$ и θ) тесно связаны с ролью априорной информации.

Определим конкретнее, что именно в статистике понимается под априорной информацией. В нее включают любые сведения, имеющиеся до того, как мы приступили к синтезу новой процедуры $\hat{\theta}$, в том числе и любую информацию о природе наблюдений (но не саму выборку x , считающуюся информацией апостериорной). Конкретнее априорные сведения характеризуют: 1) пространство ситуаций Θ ; 2) природу случайных факторов n ; 3) оператор μ , определяющий характер взаимодействия Θ и n ; 4) пространство наблюдений X ; 5) требования потребителя к качеству решений (нумерация та же, что и на рис. 7.8).

ПОНЯТИЕ ОБ ОСНОВНЫХ НАПРАВЛЕНИЯХ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Априорная информация может быть более или менее полной и точной; в зависимости от этого по-разному ставятся и решаются статистические задачи выбора. Можно даже утверждать, что разным уровням априорной информации соответствуют различные специфические ветви математической статистики.

Самое полное описание случайного объекта состоит в задании распределения вероятностей на множестве возможных состояний этого объекта, поэтому наиболее подробное и полное задание априорной информации состоит в том, что считаются известными: распределение $P(\cdot)$, $\cdot \in \Theta$; условное распределение выборочных значений $F(x|\cdot)$, $x \in X$, $\cdot \in \Theta$; функция потерь $l(\cdot, \cdot)$, выражающая отношение потребителя решений к расхождению между $\hat{\theta}$, т.е. тем, что он должен использовать вместо истинного θ , и действительным состоянием θ . Такой уровень априорной информации соответствует байесову направлению статистики (Т. Байес – известный английский статистик). Среднее значение потерь R , связанное с конкретным алгоритмом $\hat{\theta}$ обработки наблюдений x , называемое *байесовым риском*, принимается за меру качества этого алгоритма. Оптимальная в этом смысле процедура $\hat{\theta}^*$ (также называемая байесовой) и считается наилучшим решением задачи:

$$\begin{aligned}\hat{\theta}^*(x) &= \arg \min_{\hat{\theta}(x)} R(\hat{\theta}(x)) = \arg \min_{\hat{\theta}(x)} M_{X\Theta} l(\hat{\theta}(x), \theta) = \\ &= \arg \min_{\hat{\theta}(x)} \int \int l(x, \theta) dF(x|\theta) dP(\theta).\end{aligned}$$

Наибольшее количество споров относительно байесовых задач вызывала необходимость задавать априорное распределение $P(\cdot)$. Постулат Лапласа – Байеса, предлагающий при неизвестности $P(\cdot)$ считать его равномерным в Θ , приводит к противоречиям в случае деформаций пространства Θ [39, § 17]. Не помогает и предположение о том, что неизвестное $P(\cdot)$ принадлежит некоторому классу распределений, с тем чтобы взять в этом классе “наихудшее” распределение и для него найти байесову процедуру. Такая минимаксная процедура гарантирует, что “хуже не будет”, если только $P(\cdot)$ действительно входит в заданный класс.

Спор между сторонниками байесова подхода и его противниками можно считать историческим недоразумением. В конце концов было признано, что могут существовать и другие уровни априорной информации, для которых требуется создание своих методов синтеза про-

цедур. Следующим уровнем стал отказ от необходимости знать $P(?)$; на этом уровне в синтезе алгоритмов участвует только информация о семействе функций $F(x|?)$. Оказалось, что если подставить в функцию плотности $f(x|?)$ выборочные значения x_1, \dots, x_N , и рассматривать ее зависимость от $?$, то такая зависимость $L(?) | x_1, \dots, x_N = f(x_1, \dots, x_N | ?)$ обладает замечательными свойствами, из-за которых ее и назвали *функцией правдоподобия*. Например, если $? -$ неизвестный числовой параметр распределения, то $\hat{\theta}_{\text{пп}} = \arg \max_{\theta} L(\theta | x_1, \dots, x_N)$ является очень хорошей оценкой рассматриваемого параметра (этот метод оценивания называется *методом максимального правдоподобия*). В том случае, когда по выборке x_1, \dots, x_N следует принять

BAYESIAN RISK	<i>байесов риск</i>
SAMPLE выборка	
NON-PARAMETRIC	
<i>непараметрический</i>	
DECISION решение	
LIKELIHOOD FUNCTION	<i>функция правдоподобия</i>

решение в пользу одной из конкурирующих гипотез H_0 и H_1 , т.е. решить – это выборка из распределения с плотностью $f(x|H_0)$ или $f(x|H_1)$, лучшей процедурой является вычисление отношения правдоподобия $f(x_1, \dots, x_N | H_1) / f(x_1, \dots, x_N | H_0)$ и выбор гипотезы H_1 , если это отношение превышает заданный порог, и гипотезы H_0 , если ниже его. В рамках этого уровня возможны и другие методы принятия статистических решений; обычно их использование вызвано соображениями простоты реализации, но по качеству получаемых с их помощью решений они не превосходят процедур, основанных на функции правдоподобия.

Априорное знание функции $F(x|?)$ не всегда достижимо, и возникает вопрос, как осуществить выбор, если $F(x|?)$ неизвестно. Если оставаться в рамках параметрических моделей, то можно снова рассмотреть класс распределений и воспользоваться минимаксной методикой. Подобные алгоритмы также используются в статистике и дают хорошие результаты в рамках принятых предположений. Как и в аналогичном расширении класса байесовых задач, о свойствах процедур минимаксного правдоподобия можно говорить, только если имеется уверенность, что реальные распределения действительно находятся в заданном классе.

Можно сразу и полностью отказаться от необходимости знать распределение $F(x)$, конечно, не от того, что выборка подчинена какому-то, пусть неизвестному, но существующему распределению: предположение о статистичности наблюдений остается в силе. Такая позиция соответствует новому уровню априорной информации и новому направлению в статистической теории и практике – *непараметрической статистике**.

Незнание функционального вида распределения не означает, что вообще ничего неизвестно о распределении

Так как вся информация о случайном объекте содержится в его распределении вероятностей, то любая статистическая задача, по существу, может быть сведена к выбору определенного распределения из некоторого множества распределений.

и о свойствах выборки. То, что все-таки мы знаем, и служит основой для построения непараметрических процедур, решающих задачу выбора в этих условиях. Информация, не делающая известной функцию распределения, но полезная для принятия решения, может быть весьма разнообразной. Например, мы можем знать, что распределение непрерывно или симметрично (в частности, при измерении физических полей); из природы явления можно узнать некоторые числовые характеристики – моменты, квантили (так, для сигналов в ряде

* Название связано с тем, что в этом разделе статистики рассматриваются классы распределений, несводимые к параметрическим семействам функций.

технических систем известна их мощность); обычно известно, чем отличаются альтернативы, из которых следует выбирать (скажем, если новый метод диагностики эффективен, то процент верных диагнозов должен существенно превышать процент случайных угадываний); иногда даже неизвестно, чем именно будут различаться альтернативы, просто известен факт, что они чем-нибудь отличаются (как в задачах проверки правильности теоретических моделей). Все подобные сведения могут быть использованы при синтезе решающей процедуры.

Имеется несколько подходов к синтезу непараметрических процедур. Кроме эвристического, изобретательского подхода (часто достаточно удачного, но всегда оставляющего место для сомнений) развиваются и теоретически обоснованные подходы. Например, теория инвариантности [11; 20] предлагает методы, использующие свойства симметрии некоторых задач. Большой универсальностью обладает метод интерпретации статистик как функционалов от оценок распределений. Этот метод основан на двух следующих идеях. Во-первых, всякая (или почти всякая) статистическая задача может быть сведена к выбору значения некоторого функционала $J(F)$ от неизвестного распределения $F(x)$ (главной трудностью этого метода является построение функционала, содержащего всю априорную информацию, имеющуюся в постановке задачи [12]. Например, в задачах сравнения двух распределений необходимо построить удачную меру их различия $J(F, G)$). Во-вторых, оценку интересующего нас функционала (т.е. выбор альтернативы) предполагается получать, подставляя в него вместо неизвестных распределений их непараметрические (т.е. не использующие знание их функционального вида) оценки этих распределений. Хотя при этом можно использовать разные оценки одного и того же распределения, что позволяет иметь несколько процедур решения одной задачи, сами процедуры часто возможно сравнивать теоретически и решать, в каких условиях одна лучше другой.

Переходя с одного уровня априорной информации на другой, отказываясь от учета недоступной информации, мы в результате получаем в общем случае решения все более “низкого” качества: байесовы процедуры лучше процедур правдоподобия, которые, в свою очередь, обычно лучше непараметрических. Однако, во-первых, это отношение нестрогого порядка (в некоторых задачах непараметрические процедуры не хуже байесовых), а во-вторых, и это главное, такое упорядочение верно только при верной априорной информации. В случае неверной априорной информации ситуация резко меняется: чем меньше априорной информации заложено в процедуре, тем слабее ухудшает решение ее ложность. Именно поэтому непараметрические процедуры часто оказываются предпочтительнее остальных.

Все же полный отказ от знания распределений выглядит как чересчур сильная реакция на слишком жесткие требования параметрической статистики. Существует множество практических ситуаций, в которых мы не знаем распределения точно, но знаем его приблизительно. В таких случаях применение непараметрических процедур равносильно отказу от этой хотя и неточной, но правильной и поэтому полезной информации. В последние годы развивается новая ветвь математической статистики, соответствующая такому уровню априорной информации, – *робастная статистика** [40]. Ее основная идея состоит в том, чтобы несколько поступиться оптимальностью на точно задаваемом опорном распределении, но зато обеспечить некоторый гарантированный уровень качества решения на всех других распределениях, лежащих в окрестности опорного решения. Математическая сторона этого вопроса связана с теорией непрерывности и дифференцируемости функционалов от распределений и требует соответствующих знаний, но в ряде случаев получаемые алгоритмы легко воспринимаются и на уровне здравого смысла. Например, для получения робастных оценок среднего рекомендуется вычислять среднее арифметическое, отбросив из выборки самое большое и самое малое наблюдения. Именно так и поступают при определении средней урожайности за несколько лет; аналогично начисляют баллы в фигурном катании. Конечно, это простейший случай, когда требуется обеспечить устойчивость к случайным и не относящимся к делу от-

* От англ. robust – загруженность, малочувствительность, устойчивость.

клонениям, таким, как субъективное оценивание или нетипичные погодные условия. Теория робастных процедур дает рекомендации и в более сложных условиях [40]. Еще один вариант, при котором распределения изначально считаются задаваемыми интервалами, а не точными значениями, только начал развиваться.

ПРАВИЛА “СТАТИСТИЧЕСКОЙ ТЕХНИКИ БЕЗОПАСНОСТИ”

В предыдущих параграфах установлено, что любой вариант задачи выбора сопряжен со сложностями, парадоксами, “ловушками” и опасностями, когда теоретические методы применяются на практике. Не является исключением и выбор в условиях статистической неопределенности. Неудачное или неправильное применение статистических методов к решению реальных проблем дало повод к появлению довольно зловещей шутки: “Есть три вида лжи – просто ложь, наглая ложь и статистика”.

Причины неправильного применения статистических методов немногочисленны, и их знание совершенно необходимо. Перечень условий, связанных с отрицательными последствиями использования статистических решений, можно рассматривать как своего рода “инструкцию по технике статистической безопасности”.

1. *Статистический вывод по своей природе случаен, он может иметь высокую надежность и точность, но почти никогда не может быть абсолютно достоверным.* В статистике этот факт и не скрывается, каждая процедура сопровождается характеристикой ее качества: оценка параметра имеет фиксированную точность (задаваемую, например, дисперсией); принятие одной из гипотез связано с вероятностями ошибок, сообщаемыми пользователю заранее. Когда необходимо ужесточить требования к качеству, это обычно можно сделать увеличивая объем выборки. Итак, статистический вывод может быть ошибочным, но мы можем варьировать характеристики этих ошибок (в пределах наличных ресурсов).

2. *Качество решения на выходе статистической процедуры зависит от того, что подается на ее вход.*

Известны случаи, когда проспавший лаборант утром сам сочинял “протокол”очных наблюдений и измерений; когда в таблицы и протоколы вносились “поправки”, угодные лицу, не заинтересованному в истине, и т.д. Ясно, что статистическая обработка таких “данных” выдаст некий результат, но стоит ли обвинять статистику в его качестве?

3. Следующий, более сложный случай – *добросовестное заблуждение относительно статистичности серии наблюдений, когда этого на самом деле нет.* Тутубалин [36] приводит много примеров, когда статистической обработке подвергались данные, вообще не имеющие статистической природы. Иногда этот факт трудно проверить, особенно при небольших объемах выборки. Этому моменту следует уделять специальное внимание при организации экспериментов, а при малейшем сомнении не полагаться на “паспортные” характеристики качества процедуры, относиться к результатам обработки как к ориентировочным данным, как к поводу для дальнейших исследований.

4. Утрата ожидаемого качества статистических решений может быть *следствием использования процедуры, не соответствующей действительному уровню априорной информации* (например, применение дисперсионного анализа к негауссовым данным). Если “действительный” уровень априорной информации не очень ясен, то полезно обработать данные несколькими способами. Расхождение выводов должно стать сигналом к поиску причин расхождения. Следует осторегаться принимать решение “голосованием процедур”, так как опасность нетранзитивности (см. § 7.5) – свойство голосования вообще, а не только голосования людей.

Нарушение априорных предположений может происходить довольно неожиданным образом. Поучителен пример из реальной жизни: непараметрическая процедура классификации была “обучена” эффективно распознавать нефтяные слои от водяных по геофизическим данным одного месторождения. Внедренная на другом, соседнем месторождении нефти программа стала давать слишком большие вероятности ошибок распознавания: непараметрич-

ность процедуры не означает, что на выборках из другого распределения она сохранит те же свойства.

5. Причиной необоснованных претензий к статистике может служить *неверная содержательная интерпретация правильного статистического вывода*. Например, одно из английских статистических исследований прошлого века установило, что здоровье людей, носящих котелки, значительно лучше здоровья мужчин в кепках; в другом случае оказалось, что главным фактором, влияющим на различие в урожайности клевера в соседних деревнях, было... число старых дев. Таким (и подобным) правильным выводам можно придать совершенно разный смысл. (Кстати, оказалось, что английские старые девы держат по несколько кошек, а мыши любят разорять гнезда шмелей – основных опылителей клевера.) Действительные причины обнаруженной статистической связи могут оставаться неясными. Например, новосибирские статистики установили очень сильную статистическую связь между урожайностью зерна в данном году и яйценоскостью кур. Интерпретация зависимостей лежит вне статистики, и за неправильную интерпретацию нельзя осуждать саму статистику.

Несмотря на все предосторожности, как и в любой практической деятельности, в статистической практике возможны нарушения правил безопасности и неизбежны связанные с этим потери. Видимо, именно поэтому юристы одной крупной фирмы, поставляющей очень развитую программную систему для статистического анализа на ЭВМ, включили в договор о поставке этой системы пункт о том, что “фирма не несет ответственности за возможный ущерб от использования данной системы”.

Подведем итог

Неопределенность в статистических задачах имеет “двуэтажную” природу. Наблюдаемые данные подчинены конкретному вероятностному распределению, и связанная с этим распределением неопределенность образует “первый этаж”. Имеется и другая неопределенность относительно того, какое же именно распределение из некоторого множества порождало экспериментальные данные. Этую “вторую” неопределенность и требуется снять, осуществив выбор на данном множестве альтернативных распределений.

Алгоритм такого выбора самого распределения (или значения некоторого его признака) называется статистической процедурой. Использование статистических выводов со-

пряжено с определенными сложностями, “ловушками”, опасностями, риском; требуется знание и соблюдение правил “статистической безопасности”.

§ 7.8. ВЫБОР ПРИ РАСПЛЫВЧАТОЙ НЕОПРЕДЕЛЕННОСТИ

Любая задача выбора является задачей целевого сужения множества альтернатив. Как описание альтернатив (перечень их признаков, параметров и т.п.), так и описание правил их сравнения (критериев, отношений) даются в терминах той или иной измерительной шкалы (см. § 6.2). Известно, что любая измерительная шкала допускает размытие (см. § 6.3). Точнее говоря, в жизни мы часто сталкиваемся с ситуациями, описать которые можно лишь в размытых шкалах. Это, разумеется, относится и к ситуациям, приводящим к выбору. В результате мы приходим к задачам выбора в условиях расплывчатой неопределенности. Каждой из задач, рассмотренных в предыдущих параграфах, можно поставить в соответствие несколько расплывчатых задач, поскольку размытыми могут оказаться все или только некоторые компоненты задачи. До настоящего времени рассмотрено лишь незначительное число таких задач, однако ведется работа в этом направлении.

МНОГОКРИТЕРИАЛЬНЫЙ ВЫБОР В РАСПЛЫВЧАТОЙ СИТУАЦИИ

Уже в первой работе по принятию решений в расплывчатой ситуации Беллман и Задэ [4] выдвинули идею, состоящую в том, чтобы и цели, и ограничения представлять как размытые множества на множестве альтернатив (в случае одной цели и одного ограничения это соответствует заданию множеств $G = \{x, ?_G(x)\}$ и $C = \{x, ?_C(x)\}$). Следующий важный шаг состоял в определении размытого решения D как пересечения размытой цели G и размытого ограничения C , т.е. (см. § 6.3)

Summary

Uncertainty in statistical problems occurs on two “levels”. Observed data are generated from a definite probability distribution, and uncertainty of the latter forms the first “level”. There is also another form of uncertainty – that of which particular distribution from a certain set of distributions is the “parent” of the sample. Ultimately this “second” uncertainty is removed by making a choice from the given set of alternative distributions. An algorithm of this choice of a distribution (or of its characteristic value) is called a statistical procedure. The practical use of statistical procedures is accompanied by difficulties, pitfalls, dangers, and risks; knowledge of and compliance with certain “rules of statistical safety” are also needed.

CONSTRAINT	$\exists_D(x) = \min [\exists_G(x), \exists_C(x)].$
<i>ограничение</i>	Обобщение на случай большего числа условий очевидно. Если из размытого множества D требуется выделить какую-то одну альтернативу, то можно поступать по-разному (вплоть до рандомизации выбора), но возможный вариант состоит в максимизации $\exists_D(x)$:
OPTIMAL опти- мальный	$x^* = \arg \max_{x \in X} \mu_D(x).$
FUZZY расплы- ватый	(2)
DISTANCE рас- стояние	
COMPARISON сравне- ние	

Правила выбора в расплывчатой ситуации, естественно, являются различными в зависимости от того, что именно расплывчато в этой ситуации.
Задача выбора решается просто и изящно, если кри- включение $E(p_2) \subseteq E(p_1) \subseteq \text{PM}$, где PM – паретовское множество (см. § 7.2).

Функции принадлежности вообще находить непросто (см. § 6.3), а при использовании изложенного подхода, кроме того, требуется, чтобы они еще имели и смысл критериальных функций в задаче выбора. Это может оказаться и неудобным, и бессмысленным. С.А. Орловский [27] предложил не изменять содержательного смысла критериев качества альтернатив и не отождествлять критериальные функции с принадлежностями, а отразить в модели расплывчатость шкал, в которых эти критерии фиксируются (если такая расплывчатость имеет место). Предполагается, что критериальные функции $q_i(x)$ относятся к параметрическим семействам, т.е. $q_i(x) = J_i(x, \bar{q})$, и считается, что расплывчатость критериальных функций сводится к расплывчатости в описании параметров \bar{q} : $Q = \{ \bar{q}, \exists_Q(\bar{q}) \}$. Теперь для каждой альтернативы x значение критерия $J_i(x, \bar{q})$ принадлежит размытому множеству, функция $\exists_i(J_i(x))$ принадлежности к которому зависит от x и от конкретного вида функций $J_i(x, \bar{q})$ и $\exists_Q(\bar{q})$. Носитель этого множества может быть как ограничен сверху величиной $J_i^0(x)$, так и не ограничен (чем больше J_i , тем лучше; ограниченность J_i сверху несущественна). Если естественного ограничения снизу нет, то его можно ввести искусственно, задав некоторый уровень $\underline{\exists}$ ($0 < \underline{\exists} < 1$) для функции принадлежности и взяв в качестве $J_i^0(x)$ наименьший корень уравнения $\exists_i(J_i(x)) = \underline{\exists}$. В результате величины $J_i^0(x), \dots, J_m^0(x)$ можно рассматри-

$$\exists_D(x) = \min [\exists_G(x), \exists_C(x)]. \quad (1)$$

Обобщение на случай большего числа условий очевидно. Если из размытого множества D требуется выделить какую-то одну альтернативу, то можно поступать по-разному (вплоть до рандомизации выбора), но возможный вариант состоит в максимизации $\exists_D(x)$:

$$x^* = \arg \max_{x \in X} \mu_D(x).$$
 (2)

При таком изложении задачи выбора напрашивается идея о том, чтобы вообще функцию принадлежности i -му условию интерпретировать как i -й критерий качества и вернуться к многокритериальным задачам. Тогда соотношение (1) оказывается формой суперкритерия (1) § 7.2, которая далеко не единственна.

Интересны исследования в этом направлении, сделанные Эстером [45]. Он рассмотрел суперкритерий вида

$$Z_p(x) = \left\{ \sum_{i=1}^m g_i[\mu_i(x)] \right\}^{1/p}, \quad (3)$$

где $0 \leq g_i \leq 1$, $\sum_{i=1}^m g_i = 1$; m – число размытых условий; $\exists_i(x)$ – функция принадлежности i -му условию; p – параметр суперкритерия. Представление (3) интересно не только наличием свойств, облегчающих математическое рассмотрение задачи (например, монотонность и непрерывность по всем компонентам), но и тем, что оно охватывает широкий класс частных суперкритериев. Так, при $p \rightarrow -\infty$ получается оператор нахождения минимального элемента из заданной совокупности (т.е. снова приходим к формуле (1)), при $p = 0$ – оператор умножения, при $p = 1$ – оператор сложения, при $p \rightarrow +\infty$ – оператор нахождения максимального элемента. Итак, задача нахождения наилучшей альтернативы x^* сводится к максимизации $Z_p(x)$:

$$x_g^* = \arg \max_{x \in X} Z_p(x).$$
 (4)

Очевидно, что при этом решение зависит от конкретного набора коэффициентов $g = \{g_i\}$. Обозначим через $E(p)$ множество $\{x_g^*\}$, соответствующее разным g при фиксированном p . Эстер обнаружил [46] интересные свойства множеств $E(p)$: для всех $-\infty < p_1 \leq p_2 < \infty$ справедливо

вать как новые (и уже неразмытые!*) критериальные функции, и мы возвращаемся к стандартной многокритериальной задаче, которую можно решать любым из стандартных методов (см. § 7.2). Орловский, в частности, предлагает находить паретовское множество альтернатив.

Заканчивая обзор расплывчатых вариантов критериальных задач выбора, рассмотрим еще задачи, связанные с использованием расстояний между точками в пространстве альтернатив. При расплывчатом описании альтернатив предлагается “расстояние” определять через модули разностей функций принадлежности, например

$$d(x_i, x_j) = \left(\frac{1}{m} \sum_{r=1}^m |\mu_r(x_i) - \mu_r(x_j)|^p \right)^{\frac{1}{p}}, \quad (5)$$

где $\mu_r(x)$ – функция принадлежности по r -му признаку к интересующему нас множеству. Такие расстояния используются в задачах классификации [17].

НЕКРИТЕРИАЛЬНЫЕ ЗАДАЧИ РАСПЛЫВЧАТОГО ВЫБОРА

Некоторые успехи имеются и в рассмотрении расплывчатых вариантов выбора, описываемого на языке бинарных отношений (см. § 7.3). Во-первых, сделано расплывчатое обобщение отношения предпочтения [49]. *Размытое отношение R слабого порядка* определяется как удовлетворяющее размытым условиям связности и транзитивности:

$$x_i \neq x_j \Rightarrow \mu_R(x_i, x_j) > 0 \text{ или } ?_R(x_j, x_i) > 0 \quad (6)$$

(связность);

$$\mu_R(x_i, x_k) = \max_{x_j} \{\min[\mu_R(x_i, x_j), \mu_R(x_j, x_k)]\} \quad (7)$$

(транзитивность).

Если условие (6) заменить условием

$$\mu_R(x_i, x_j) > 0 \Rightarrow ?_R(x_j, x_i) = 0 \quad (8)$$

(асимметрия), то такое упорядочение называется *сильным*.

Во-вторых, Л. Задэ показал [30], что любое расплывчатое отношение R допускает разложение по $?$ в виде объединения неразмытых множеств $R_?$ с функциями принадлежности

$$\mu_{R_\alpha}(x_i, x_j) = \begin{cases} \alpha & \text{при } (x_i, x_j) \in R_\alpha, \\ 0 & \text{при } (x_i, x_j) \notin R_\alpha, \end{cases} \quad (9)$$

где $0 < ? < 1$ (что можно представить как расслоение объема под $?_R$ на горизонтальные пластины уровней $?$).

Это, например, позволяет перейти от расплывчатого описания коллективного упорядочения альтернатив к нерасплывчатому множеству альтернатив, отобранных “со степенью согласия на уровне $?$ ” [49].

О других классах задач выбора кратко можно сказать следующее.

Расплывчатой версии языка глобальных функций множеств (описанного в § 7.4) пока не создано.

Начато исследование различий и аналогий между статистической и размытой неопределенностями. Некоторые особенности, возникающие при одновременном наличии обеих неопределенностей, рассмотрены, например, в [26]. Не углубляясь в детали (так как для этого понадобилось бы использовать достаточно сложные построения, связанные с понятием случайных множеств), отметим, что в целом идеи теории расплывчатых множеств привлекают все больший интерес, поскольку в этой модели удалось отразить многие особенности языковых моделей и действий человека на их основе.

Подведем итог

Summary

Ряд ситуаций выбора характеризуется расплывчатой неопределенностью. Рассмотрено несколько различных вариантов ситуаций, описанных как неопределенность. Several variants of such situations are described as fuzziness. Several variants of such situations are considered, with

* При отсутствии размытости $J_i^0(x)$ совпадает с $J_i(x, q)$.

тов таких ситуаций; функции different meanings of the fellowship function (it can either be разный смысл (либо сами слу- identified with the criterion function, or it can describe the fuzziness of a certain parameter of the параметра критериальной criterion function). Naturally, this функции). Естественно, это leads to different algorithms of приводит к разным алгоритмам choice.
выбора.

§ 7.9. ДОСТОИНСТВА И НЕДОСТАТКИ ИДЕИ ОПТИМАЛЬНОСТИ

Во всех рассмотренных в предыдущих параграфах вариантах задачи выбора проблема состояла в том, чтобы в исходном множестве найти *наилучшие в заданных условиях, т.е. оптимальные, альтернативы*. Здесь важное значение имеет каждое слово. Говоря “наилучшие”, мы предполагаем, что нам известен критерий, способ сравнения вариантов и нахождения самого лучшего из них. Однако этого мало: важно учесть условия, ограничения, так как их изменение может привести к тому, что при том же критерии наилучшим окажется другой вариант.

ДОСТОИНСТВА ОПТИМИЗАЦИОННОГО ПОДХОДА

Идея оптимальности является центральной идеей кибернетики. Понятие оптимальности получило строгое и точное представление в математических теориях, прочно вошло в практику проектирования и эксплуатации технических систем, сыграло важную роль в формировании современных системных представлений, широко используется в административной и даже общественной практике, стало понятием, известным почти каждому человеку. И это не удивительно: стремление к повышению эффективности труда, творчества, любой целенаправленной деятельности, это естественное стремление человека как бы нашло свое выражение, ясную и понятную форму в идее оптимальности. Различие между строго научным и “общепринятым”, житейским пониманием оптимальности совсем невелико. Правда, встречающиеся выражения типа “наиболее оптимальный” или “добьемся максимального эффекта при минимуме затрат” математически некорректны, но лица, использующие эти выражения, на самом деле просто нестрого и неудачно выражают правильную мысль: как только дело касается конкретной оптимизации, они быстро и легко исправляют формулировки.

Нхождение оптимальных вариантов особенно важно для оценки состояния современной техники и определения перспектив ее дальнейшего развития. Знание параметров оптимальной альтернативы позволяет составить представление о принципиально непревосходимых пределах возможности техники; сравнение с этими параметрами часто помогает решить вопрос о целесообразности дальнейших усилий по улучшению того или иного показателя качества изделия. Часто оказывается, что в результате оптимизации уже имеющейся аппаратуры значение критерия качества можно повысить всего на несколько процентов; это означает, что и без оптимизации достигнутое значение уже находится в окрестности оптимума. Однако нередко оптимизация вскрывает значительные резервы улучшения: если изменение критерия качества на несколько процентов на практике трудно уловить и реализовать, то разрыв в десятки процентов уже обещает существенное улучшение, отвечающее затрачиваемым усилиям. Иногда же этот разрыв настолько велик, что возникает вопрос о том, нет ли принципиально новых путей развития данной отрасли техники. Например, пропускная способность человеческого глаза составляет несколько десятков битов в секунду, а пропускная способность существующих телевизионных каналов – несколько миллионов битов в секунду. Такая пропускная способность телеканалов обусловлена необходимостью передавать каждую деталь в каждом кадре заново, хотя подавляющее большинство деталей не изменяется на протяжении многих кадров. Но как сократить столь высокую избыточность телевизионного сигнала, пока не придумано.

ОГРАНИЧЕННОСТЬ ОПТИМИЗАЦИОННОГО ПОДХОДА

При всей очевидной полезности идеи оптимизации практика требует необходимости осторожного обращения с ней. Для такого заключения имеются достаточно веские основания.

1. Оптимальное решение часто оказывается очень “хрупким”: *незначительные на первый взгляд изменения в условиях задачи могут привести к выбору существенно отличающихся альтернатив*. В связи с этим в теории оптимизации в последнее время уделяется все большее внимание таким модификациям понятия оптимальности, которые придают решениям определенную устойчивость [25; 40].

2. Оптимизация всегда опирается на предположение, что участвующие в задаче критерии достаточно хорошо отображают поставленную цель. Даже если это и так, то обычно рассматриваемая система является частью некоторой большей системы, и тогда локальная оптимизация совсем не обязательно приведет к тому же результату, который потребуется от подсистемы при оптимизации системы в целом. *Это приводит к необходимости увязывать критерии подсистем с критериями системы, часто делая ненужной локальную оптимизацию*.

3. Максимизация критерия оптимальности часто отождествляется с целью, а на самом деле это разные вещи. Фактически критерий и цель относятся друг к другу как модель и оригинал, со всеми вытекающими отсюда особенностями (см. гл. 2). Многие цели трудно или даже невозможно количественно описать. Однако при необходимости это более или менее удачно делается. Конечно, количественный критерий является лишь суррогатом цели. Например, уровень работы пожарной команды можно оценивать по скорости ее прибытия в случае вызова. Согласно стандартам ЮНЕСКО, принято оценивать уровень медицинского обслуживания населения по показателям детской смертности. Качество работы ряда предприятий бытового обслуживания оценивается по количеству жалоб от населения. Не только объем, но и уровень работы научного сотрудника оценивается по количеству опубликованных им статей. В более сложных системах оценок о качестве научных публикаций судят по их цитируемости другими авторами. Подобные примеры можно продолжить. Общее для них сводится к тому, что *критерии характеризуют цель лишь косвенно, иногда лучше, иногда хуже, но всегда приближенно*.

4. Как мы уже отмечали, в понятии оптимальности кроме критериев не менее важную роль играют ограничения. Даже небольшие их изменения существенно сказываются на решении. Еще более разительный эффект можно получить, снимая одни ограничения и добавляя другие. В этом моменте содержится серьезная “ловушка” при оптимизационном подходе к сложным системам, на что обратил внимание Н. Винер в первых же публикациях по кибернетике [9; 10]. Опасность состоит в том, что, *не задав всех необходимых ограничений, мы можем одновременно с максимизацией основного критерия получить непредвиденные и нежелательные сопутствующие эффекты*. Для иллюстрации этого Н. Винер любил приводить английскую сказку об обезьяньей лапке. Обладатель этого талисмана мог с его помощью выполнить любое желание (согласитесь, что таким свойством может обладать лишь чрезвычайно сложная система, хотя управление ею внешне весьма просто: стоит только выразить желание). Когда он однажды получил таким образом большую сумму денег, то оказалось, что за это необходимо было пожертвовать жизнью любимого сына... Мысль Н. Винера о том, что по отношению к сложным системам мы принципиально не в состоянии заранее определить все условия и ограничения, гарантирующие отсутствие нежелательных последствий оптимизации, позволила ему сделать мрачное предположение о катастрофических последствиях кибернетизации общества.

PITFALL ловушка

POLL опрос

DESIGN проектирование

CONDITION условие

EFFICIENCY эффективность

Идею оптимальности, чрезвычайно плодотворную для систем, поддающихся адекватной математической формализации, нельзя перенести на сложные системы.

Конечно, математические модели, которые удается иногда предложить для таких систем, можно оптимизировать; однако всегда следует

оптимизации.

Итак, с позиций системного анализа отношение к оптимизации можно сформулировать следующим образом: это мощное средство повышения эффективности, но использовать его следует все более осторожно по мере возрастания сложности проблемы.

Многие задачи проектирования технических систем могут быть достаточно хорошо formalизованы, сведены к математическим моделям, позволяющим ставить и решать оптимационные задачи. Однако даже после успешного преодоления сложностей formalизации системотехнических проблем остаются две “ловушки”: неустойчивость оптимальных решений, т.е. их сильная чувствительность к изменениям условий, и неоднозначность постановки многокритериальных задач. Возможны следующие меры предосторожности: исследование чувствительности, пользование устойчивыми (робастифицированными) процедурами; выделение паретовских множеств альтернатив с последующим их сужением за счет дополнительной информации либо рассмотрение оптимальных альтернатив по нескольким различным сверткам критериев и т.п.

При переходе от чисто технических или технологических проблем к проблемам, включающим организационные и социальные вопросы, ситуация существенно усложняется. Именно на разработку методов рассмотрения таких проблем направлены сейчас основные усилия в системном анализе. Мы будем подробнее говорить об этом в последующих главах, а сейчас затронем только проблему оптимизации. Сложные системы потому и сложны, что не поддаются полной formalизации. Следовательно, оптимационные задачи, которые удается поставить при исследовании сложных систем, неизбежно имеют частичный, подчиненный характер, если описывают хорошо структурированные подсистемы, либо являются заведомо приближенными, если относятся к системе в целом. Поэтому оптимизация в таких исследованиях не конечная цель, а средство, промежуточный, а иногда и начальный этап работы. Заостряя этот момент, некоторые авторы пишут: “вопрос не в том, могут ли люди максимизировать, а в том, нужно ли это делать” [48]. Меры предосторожности в подобных ситуациях состоят в рассмотрении результатов оптимизации как предварительных, вспомогательных данных для дальнейшего анализа, для изменения самих задач

Подведем итог

Оптимизационные проблемы являются строго формальными математическими задачами. Практическое значение решений таких задач прямо зависит от того, насколько хороша исходная математическая модель. Высокая практическая оптимизация в технических системах не должна порождать иллюзии, что тот же эффект даст оптимизация сложных систем; в сложных системах математическое моделирование является затруднительным, приблизительным, неточным (см. § 4.4). Чем сложнее система, тем осторожнее и скептическое следует относиться к ее оптимизации.

Optimization problems are strongly formal mathematical problems. The practical significance of these problems directly depends on how good the mathematical model is. Efficiency of optimization of technical systems should not create the illusion that it is equally efficient in more complex systems; for complex systems modeling is generally difficult, and mathematical modeling is even more difficult, approximate, and imprecise (cf. section 4.4). The more complex a system is, the more precautions must be taken and the more skeptical one should be with its optimization.

Summary

§ 7.10. ЭКСПЕРТНЫЕ МЕТОДЫ ВЫБОРА

При исследовании сложных систем возникают проблемы, выходящие за пределы формальных математических постановок задач. В таком случае прибегают к услугам экспертов, т.е. лиц, чьи суждения и интуиция могут уменьшить сложность проблемы. В последующих главах мы подробнее рассмотрим работу экспертов – системных аналитиков; сейчас обсудим вопросы привлечения экспертов к решению конкретной и частной задачи системного анализа – задачи выбора. Правда, в этой частной задаче имеются и некоторые общие черты экспертных методов (например, подходы к оценке компетентности экспертов, к интерпретации даваемых ими результатов и пр.).

Основная идея экспертных методов состоит в том, чтобы использовать интеллект людей, их способность искать и находить решение слабо формализованных задач. Однако особенность интеллектуальной деятельности людей состоит в том, что она во многом зависит от внешних и внутренних условий. Поэтому в методиках организации экспертных оценок специальное внимание уделяется созданию благоприятных условий и нейтрализации факторов, неблагоприятно влияющих на работу экспертов.

ФАКТОРЫ, ВЛИЯЮЩИЕ НА РАБОТУ ЭКСПЕРТА

Очень важную роль играют факторы психологического характера. Прежде всего эксперты должны быть освобождены от ответственности за использование результата экспертизы. Дело не только в том, что лицо, принимающее решения, не хочет или не должно возлагать ответственность на других, но и в том, что сама ответственность накладывает психологические ограничения на характер выбора, а этого на стадии оценки альтернатив желательно избежать. Однако это еще не все. Приходится также учитывать, что оценка, даваемая экспертом, может зависеть от межличностных отношений с другими экспертами и иногда даже от того, известна ли его оценка другим лицам. На ход экспертизы могут повлиять и личная заинтересованность эксперта, т.е. его необъективность, и другие личностные качества (самолюбивость, конформизм, сила характера и пр.). С другой стороны, обычно сложность проблемы выходит за рамки возможностей одного человека. Было бы неразумно не использовать тот очевидный факт, что коллективная деятельность открывает дополнительные возможности для взаимного стимулирования экспертов.

Поскольку взаимодействие между экспертами может как стимулировать, так и подавлять их деятельность, в разных случаях используют методики экспертиз, имеющие различные степень и характер взаимного влияния экспертов друг на друга: анонимные и открытые опросы и анкетирование, совещания (комитеты, коллегии, комиссии, штабы), дискуссии (консилиумы, суды, ученые советы), деловые игры, мозговой штурм и т.д. Рассмотрим наиболее типичные процедуры экспертного выбора.

МЕТОДЫ ОБРАБОТКИ МНЕНИЙ ЭКСПЕРТОВ

Простейший вариант состоит в следующем. Если эксперты предлагают различающиеся упорядочения альтернатив, то возникает вопрос о том, как использовать мнения всех экспертов для окончательного упорядочения. Это далеко не тривиальная задача. Фактически мы возвращаемся к проблеме коллективного выбора со всеми его особенностями, в том числе – с возможностями парадоксов (см. § 7.5) [24; 31].

Предположим, например, что эксперты оценивают альтернативы в числовых шкалах (см. § 6.2). Пусть $q_j(x_i)$ – оценка i -й альтернативы j -м экспертом ($i = \overline{1, m}$, $j = \overline{1, n}$). Оценки $q_1(x_i)$, ..., $q_n(x_i)$ можно рассматривать как “измерения” искомой “истинной характеристики” $q(x_i)$, считая отклонения $q_j(x_i) - q(x_i)$ случайными величинами. В качестве приближения можно использовать некоторую статистику $\hat{q}(x_i) = \hat{q}(q_1(x_i), \dots, q_n(x_i))$; обычно это выборочное среднее

$$\hat{q}(x_i) = \frac{1}{n} \sum_{j=1}^n q_j(x_i),$$

но можно использовать и другие статистики (см. § 7.7).

Сложнее обстоит дело, когда альтернативы нельзя оценить сразу одним числом и экспертам предлагается дать оценки отдельно по каждому показателю. Например, оценка качества промышленного изделия складывается из оценок признаков социальных (уровень потребности), функциональных (степень соответствия назначения), экономических, эстетических, эргономических и др. В этом случае имеем набор чисел $q_{jk}(x_i)$, где k – номер признака. Кроме этих чисел экспертов просят оценить степень важности γ_{jk} каждого показателя (если это не выполнено другим способом). Тогда

$$\hat{q}(x_i) = \frac{1}{n} \sum_j \sum_k \lambda_{jk} q_{jk}(x_i).$$

Следующее уточнение вводят в случае неоднородности группы экспертов. Естественно придать различные (а не одинаковые, равные $1/n$) веса мнениям экспертов, имеющих разную квалификацию. Определение коэффициента α_j компетентности j -го эксперта можно поручить самим экспертам. Пусть каждый из них (l -й) оценивает компетентность других числами $0 \leq \gamma_{lj} \leq 1$ (при этом и свою – числом γ_{ll}). Усреднение* дает $\alpha_j = \sum_l (\gamma_{lj} / \sum_s \gamma_{ls})$. В результате получают итоговую оценку

$$\hat{q}(x_i) = \sum_j \sum_k \alpha_j \lambda_{jk} q_{jk}(x_i).$$

В тех случаях, когда эксперты лишь упорядочивают альтернативы, т.е. используют только порядковую шкалу, возможность арифметических операций отпадает (см. § 6.2). Тогда переходят к обработке либо относительных частот предпочтений данной альтернативы, либо рангов; иногда используют “медианную” альтернативу или расстояния между ранжированием [24].

Более глубокие сведения о сравниваемых альтернативах можно получить, изучая “тонкую структуру” суждений экспертов. Дело в том, что оценки экспертов “разбросаны” в пространстве критериев или упорядочений неравномерно, и чаще всего можно выделить несколько кластеров, т.е. кучно расположенных оценок, можно обнаружить и “выбросы”, т.е. оценки, далеко отстоящие от кластеров. Значительный интерес представляет выяснение общего меж-

* При этом, очевидно, предполагается, что компетентность в оценке компетентности у экспертов одинаковая. По-видимому, здесь разумнее использовать простейший вариант с отбрасыванием наивысшей и наименее высокой оценок.

ду близкими оценками и различий между отдаленными оценками. Это выполняют уже в ходе уточняющих опросов, в других терминах. Например, при сравнении качества промышленных изделий [24] выяснилось, что кластеры связаны с такими отсутствовавшими в исходной анкете признаками, как различная жесткость правил приемки для различных альтернатив, технологические различия между моделями изделий и т.д. Возможен и более детальный анализ расхождений экспертов во мнениях [29; 37].

МЕТОД “ДЕЛФИ”

Следующим шагом в развитии экспертных оценок является **метод “Делфи”** [43]. Основная идея этого метода состоит в том, что *критика благотворно влияет на эксперта, если она психологически не связана с персональной конфронтацией*. Поэтому если проводить оценку альтернатив в несколько туров, сообщая после каждого его полные итоги и сохраняя анонимность участников, то эксперты склонны не только критиковать, но и прислушиваться к критике, относящейся к ним лично. Устранение психологических трудностей, связанных с персональной критикой, придает самой критике большую деловитость, объективность, она легче воспринимается. Все это приводит к тому, что обычно оказывается достаточно следующих четырех этапов:

- 1) раздача анкет, сбор оценок, их обобщенное представление с указанием разброса мнений;
- 2) сообщение итогов и запрос объяснений причин индивидуального отклонения от средней или медианной оценки первой итерации;
- 3) сообщение всех объяснений и запрос контрагументов на них;
- 4) сообщение возражений и запрос новых оценок альтернатив, если эксперт пожелает их изменить; нахождение окончательного итога.

Вся работа проводится под руководством отдельной управляющей группы, в которую входят системный аналитик и лицо, принимающее решения; анонимность экспертов сохраняется до конца работы (а по желанию экспертов – и после ее окончания). Методика “Делфи” показала на практике хорошую эффективность.

Остальные экспертные процедуры, упомянутые в начале параграфа, предназначены не столько для выбора, сколько для генерирования альтернатив. Мы вернемся к ним в гл. 9.

Подведем итог

Всякий выбор основан на оценивании или на сравнении альтернатив. В ряде случаев не удается сделать такую оценку “объективно”, т.е. произвести нужные измерения с помощью измерительных приборов. Тогда прибегают к “субъективным измерениям”, т.е. к экспертным оценкам. Такой способ оценивания альтернатив имеет важные особенности не только в самом процессе получения оценок, но и в методах их обработки. Эти особенности необходимо учитывать в ходе проведения системного анализа с привлечением экспертов.

Summary

Any decision is based on comparison of alternatives. In some cases it is impossible to compare them “objectively”, i.e., to take “objectively”, i.e., to take “objective measurements with the use of measuring devices. In such cases one usually turns to “subjective measurements”, i.e., to expert opinions. This type of estimate has certain important features, both in the process of obtaining the estimates and in processing them as well. These features should be taken into account in the course of systems analysis that involves expert opinions.

§ 7.11. ЧЕЛОВЕКО-МАШИННЫЕ СИСТЕМЫ И ВЫБОР

Основной причиной возникновения системного анализа является необходимость разрешения сложных проблем, управления сложными системами. Многие существенные особенности преодоления сложности можно проследить и на примере конкретного этапа, представляющего собой хотя и важную, но лишь составную часть управления, – этапа выбора (принятия решения).

Как бы ни понималась сложность (см. § 4.4), простота понимается одинаково: простым является случай, когда посторонняя помощь не требуется. В сложных случаях, особенно если принимающий решение сталкивается со сложностью в отягчающих условиях дефицита време-

SELECTION отбор

REPEATED, ITERATED повторный

DECISION – FEEDBACK SYSTEM система, адаптивная по решениям

DECISION – SUPPORT SYSTEM система поддержки решений

ELITE GROUP элитная группа

Создание человеко-машинных систем является основным направлением кибернетизации общества. Типы таких систем различаются “пропорцией участия” в них чело-

ми или других экстремальных обстоятельств, ему требуется квалифицированная помощь в оценке возможных альтернатив.

Помощь экспертов неоценима; каждый военачальник имеет штаб, ректор вуза или директор НИИ – ученый совет, министр – коллегию; в отдельных случаях образуют разовую группу экспертов для рассмотрения конкретной ситуации (см. § 7.10).

Однако существуют естественные пределы человеческих способностей при восприятии и обработке информации. Работу экспертов лимитируют не только межличностные отношения (см. § 7.10), но и внутренние психологические и физиологические причины. Оказывается, человек одновременно может оперировать лишь с небольшим числом операндов (понятий, идей, моделей, альтернатив и т.д.) – психологи, говоря о пределе возможностей, иногда называют это законом “семь плюс – минус два”. Кроме того, столкнувшись, например, с многокритериальной задачей, эксперт часто проявляет непостоянство, неуверенность, нелогичность, стремление к резкому упрощению задачи. Наконец, в ряде случаев играет роль и низкое быстродействие нервной и мышечной системы человека.

Во всех этих отношениях возможности ЭВМ превосходят способности человека, и возникает простая, но очень плодотворная идея создания единой системы, которая объединила бы достоинства человека и машины и компенсировала их недостатки. Так, Н. Винер [9] отмечал:

“Вычислительная машина очень хороша при быстрой работе, проводимой однозначным образом над полностью представленными данными. Вычислительная машина не может сравниться с человеческим существом при обработке еще не выкристаллизовавшихся данных. Если

назвать это интуицией, то я не сказал бы, что интуиция недоступна вычислительной машине, но у нее она меньше, а экономически невыгодно заставлять машину делать то, что человек делает намного лучше”.

Вряд ли возможно, да и не стоит создавать одну универсальную систему на все случаи жизни. На практике идут по пути создания человеко-машинных систем, называемых *проблемно-ориентированными*. Даже в сравнительно конкретной сфере принятия решений

наблюдается разветвление типов систем по типам задач выбора. К настоящему моменту существует несколько самостоятельных направлений этого развития.

ПАКЕТЫ ПРИКЛАДНЫХ ПРОГРАММ ДЛЯ ВЫБОРА

К первому относятся программы и пакеты программ для решения конкретных хорошо определенных задач выбора. Примером может служить математическое обеспечение ЭВМ для статистической обработки данных (т.е. выбора в условиях стохастической неопределенности; см. § 7.7). В дополнение к стандартному математическому обеспечению ЭВМ в ряде коллективов созданы пакеты прикладных программ статистики (например, в Москве [1], Ташкенте [18], Вильнюсе [30], Тарту [35], Киеве [8], Ленинграде [3], Новосибирске [19], Харькове [32] и др.). Некоторые пакеты прикладных программ основаны на идеях, позволяющих расширить классические постановки статистических задач [19; 32]; в этом отношении, например, интересны методы “машинного” обнаружения закономерностей, развивающие новосибирской школой [16]. К этому же направлению относятся системы программного обеспечения оптимизационных задач; современные базы данных и пр.

БАЗЫ ЗНАНИЙ, ЭКСПЕРТНЫЕ СИСТЕМЫ

Второе направление – создание баз знаний и экспертных систем. В настоящее время это, пожалуй, главный путь движения к “искусственному интеллекту”. Приведем высказывания наиболее активных сторонников такого мнения:

“А есть ли еще что-нибудь в искусственном интеллекте, кроме экспертных систем?” Действительно, экспертная система определяется как “воплощение в ЭВМ компоненты опыта эксперта, основанной на знании в такой форме, что машина может дать интеллектуальный совет или принять интеллектуальное решение относительно выполняемой функции. Желательно дополнительное свойство (которое многие считают главным) – способность системы по требованию объяснять ход своих рассуждений понятным для пользователя образом” [42].

Такие поистине интеллектуальные свойства экспертных систем реализуются благодаря двум их особенностям:

1) наличию полученных от человека (эксперта) знаний в определенной предметной области в форме набора фактов (предметное знание) и эвристических приемов (эмпирических правил), вводимых в машинную базу данных и базу знаний. Например, наиболее употребительный формат правила в базе знаний – это формат “ЕСЛИ <условие>, ТО <действие>“. При этом <действие> может быть и действием по изменению содержимого базы данных;

2) в отличие от программных систем прошлых поколений машина оперирует не только с “данными”, но и с понятиями, выраженными в терминах естественного языка, а также со знаниями о классах объектов, обозначенных этими терминами, и отношениями между ними. Это достигается созданием специальных программ; в последние годы разработан специальный язык Пролог, операндами которого служат как элементы данных, так и правила-операторы.

Экспертные системы имеют широкие перспективы: известны их многочисленные практические реализации в разнообразных предметных областях. Некоторые важные принципы организации экспертных систем, учитывающие расплывчатость терминов естественного языка, были заложены Д.А. Поспеловым еще в системах ситуационного управления [28]; примеры решения ряда практических задач, в которых используются эти принципы, см., например, в [41].

Если первое направление ориентировано на полную автоматизацию хорошо формализованных задач, а второе – на создание систем, накапливающих опыт экспертов и, по существу, впоследствии заменяющих самих экспертов, то в третьем современном направлении развития человеко-машинных систем выбора делается основной акцент на участие самого лица, принимающего решения, в попытках формализовать задачу выбора, в самостоятельном сравнении и оценивании с помощью ЭВМ различных альтернатив разными способами.

СИСТЕМЫ ПОДДЕРЖКИ РЕШЕНИЙ

Это третье направление представлено системами “интерактивной оценки решений” [6] и особенно “системами поддержки решений” (DSS – Decision Support Systems). Разработка систем поддержки решений ведется, в частности, в рамках интернационального проекта, осуществляемого учеными ряда стран под эгидой Международного Института Прикладного Системного Анализа в Лаксенбурге (Австрия) [48]. Идеология этого подхода хорошо выражена Хатри [47]. Рассматривая случай, когда системный аналитик помогает лицу, принимающему решение, сформировать однокритериальную задачу оценки альтернатив (см. § 7.2), он пишет:

“Без сомнения, работа принимающих решения была бы легче, если бы можно было использовать пусть приблизительную, но единственную меру эффективности. Однако я настаиваю на том, что такие процедуры ставят аналитика в положение, когда он неоднократно делает суждения о назначении коэффициентов, которые по существу, по праву должны назначаться в процессе политического вынесения решений, а не самим аналитиком. Эти назначения коэффициентов закопаны в процедурах, используемых аналитиком, и редко известны принимающему решения или понятны ему. Такие трюки в перспективе ведут к дискредитации системного анализа и к заметному уклонению от того, что является его главным предназначением: представлять принимающим решения альтернативные пути достижения целей, оценивать и показывать все основные взаимосвязи между затратами и эффективностями для этих альтернатив”.

Системы поддержки решений ориентированы не на автоматизацию функций лица, принимающего решения, а на предоставление ему помощи в поиске хорошего решения. Поэтому в таких системах особое внимание уделяется диалогу и его “дружественности” лицу, принимающему решения. Конечно, в математическое и программное обеспечение систем поддержки решений входят и формализованные процедуры, которые лицо, принимающее решения, может использовать в любой нужной ему степени.

Подведем итог

Выбор в реальных ситуациях требует выполнения ряда операций, одни из которых более эффективно выполняет человек, а другие – машина. Эффективное объединение их достоинств воплощается в создании человеко-машинных систем. Поэтому наряду с созданием чисто машинных программ и пакетов для полностью автоматического решения задач в последние годы развиваются экспертные системы и системы поддержки решений.

§ 7.12. ВЫБОР И ОТБОР

До настоящего параграфа речь шла о процедурах одноразового выбора. Даже тогда, когда приходилось проводить выбор в несколько этапов (многокритериальная оптимизация, кол-

Summary

Decision making in real-life situations demands the execution of certain operations; some of these are fulfilled more effectively by man, others by machine. It is a self-evident idea to combine the advantages of each into a man-machine system. The results are the creation of specialized program systems for the fully automated solution of problems and, more recently, the development of such man-machine systems as expert systems and decision support systems.

лективный выбор, метод “Делфи” и т.д.), это были этапы промежуточные, подготовительные перед последним, окончательным выбором.

ПОВТОРНЫЙ ВЫБОР

Однако возможны ситуации, в которых выбор повторяется многократно, причем каждый последующий выбор происходит в условиях, отличающихся от тех, в которых происходил предыдущий. Это придает динамику самому процессу выбора и его последствиям. Конкретный характер происходящих при этом изменений зависит от многих факторов: самой природы множества альтернатив, степени влияния предыдущего выбора на последующий, от того, насколько и как именно учитываются произошедшие изменения на очередном шаге выбора, и т.д. При этом возможные постановки задач весьма разнообразны, но очень немногие из этих задач на сегодняшний день рассмотрены.

Наиболее подробно изучены процессы принятия статистических решений с адаптацией, т.е. с обратной связью по решениям, иначе говоря, принятие решения на очередном шаге с учетом решений, принятых на предыдущих шагах. Примером могут служить радиолокационные станции, постоянно ведущие обзор заданной зоны, накапливающие информацию о помеховой обстановке в зоне обзора и использующие эту информацию при обработке принятых сигналов для обнаружения целей. Главный результат состоит в том, что такая адаптация может улучшить качество решений.

Другой пример процессов многократного выбора дает естественный отбор. Свообразие таких процессов изучается теорией эволюции, математической биологией.

Для нас основной интерес представляют процессы сознательного выбора, поэтому, обращаясь к многократному выбору, мы приходим, в частности, к задачам целенаправленного многократного выбора, т.е. искусственного отбора, селекции. Как показали исследования А.Н. Ефимова и В.М. Кутеева, тенденции, возникающие в ходе селекции, сильно зависят от конкретных способов формирования и пополнения отборных (“элитных”) групп. Даже простейшие модели селекции обнаруживают интересные эффекты в эволюции элитных групп [14; 15]. Эти эффекты следует иметь в виду при комплектовании любых групп элементов, в чем-то лучших, чем остальные: в промышленности – при изготовлении высокосортной продукции; в сельском хозяйстве – при выводе высокопроизводительных пород животных и сортов растений; в управлеченческой деятельности – при комплектовании групп исполнителей особо ответственных дел и т.д.

ОСНОВНЫЕ ИДЕИ ТЕОРИИ ЭЛИТНЫХ ГРУПП

Рассмотрим модель, предложенную А.Н. Ефимовым*. Предположим, что имеется некоторая совокупность элементов. Пусть интересующее нас свойство элемента выражается некоторой критериальной величиной x ; для определенности будем считать, что чем больше значение x , тем лучше, и что $0 \leq x \leq 1$. В исходной совокупности присутствуют элементы с любыми значениями величины x , и задача отбора возникает, если для достижения некоторой цели потребуется, чтобы показатель качества был не ниже некоторой заданной величины $a < 1$. Предположим, что из исходной совокупности с помощью определенного эталона (носителя величины a) отбирается заданное количество n элементов. Для общности можно предположить, что процедура отбора изредка дает сбои, так что в элитную группу с небольшой вероятностью попадают и “сорные” элементы, для которых $x \leq a$. С некоторым количеством “сорных” элементов или без них, но элитная группа сформирована и может приступить к выполнению стоящей перед ней задачи. Если элементы для отбора выбираются случайно, $F(x)$ – функция распределения качества x в исходной группе, $f(x)$ – соответствующая ей плотность, то распределение качества x в сформированной элитной группе характеризуется плотностью

$$f_*(x) = \left\{ \begin{array}{ll} \frac{\beta}{F(a)} f(x): & x < a; \\ \frac{1-\beta}{1-F(a)} f(x): & x \geq a \end{array} \right. . \quad (1)$$

* Авторы благодарят проф. А.Н. Ефимова за материалы, предоставленные им для данного параграфа.

Очевидно, что среднее качество \bar{x}_e элитной группы зависит от величин \bar{x} и $F(a)$. Так как обычно \bar{x} достаточно мало, а $F(a)$ достаточно велико, т.е. $\bar{x} < F(a)$, то $\bar{x}_e > \bar{x}$. При $\bar{x} = F(a)$ среднее качество “элиты” не отличается от среднего качества всей совокупности, а при $\bar{x} > F(a)$ становится хуже него (эти случаи не представляют практического интереса).

Если в силу любых причин (старения, разрушения, изъятия, отчисления, смерти и т.д.) какие-то элементы выбывают из элитной группы, а ее численность требуется сохранить, то возникает задача повторного выбора элементов из оставшейся основной совокупности в состав элитной группы. Ответ на вопрос, как будут изменяться свойства элитной группы при многократном повторении этого действия, зависит от ряда обстоятельств и может быть различным. Главные факторы, очевидно, таковы:

доля элитной группы в совокупности элементов, хотя и отвечающих требованиям элитности, но оставшихся вне элиты (если в первичную элитную группу вошли все элементы, для которых $x > a$, то ее придется пополнять за счет элементов, для которых $x \leq a$);

характер изменения качества x каждого элемента со временем не только в элитной группе, но и в остальной совокупности (оно может как оставаться постоянным, так и изменяться);

правило отсева из элитной группы (происходит ли это случайно, без учета величины x , либо выбывают лучшие или худшие элементы);

правило включения новых элементов в элитную группу (в соответствии с прежним эталоном a либо с измененным эталоном, либо при невозможности дальнейшего использования эталона после первого отбора);

временные отношения между моментами очередных пополнений элитной группы (эти отношения становятся важными при изменении качества x у элементов со временем).

Различные сочетания этих условий приводят к возникновению большого количества задач, приводящих к разным типам эволюции качества элитной группы. Рассмотрим некоторые из них. Во всех случаях будем считать, что выделение элитной группы практически не оказывается на свойствах оставшейся совокупности (математически это соответствует, например, конечной численности элитной группы и несчетности исходной совокупности, что позволяет считать распределение $F(x)$ неизменным при выделении элитных элементов).

ПРОЦЕДУРА “ПРЕТЕНДЕНТ – РЕКОМЕНДАТЕЛЬ”

В качестве первой задачи рассмотрим **правило “претендент – рекомендатель”**. Это правило состоит в том, что при наличии вакансии в элите взятый наугад из общей совокупности элемент (“претендент”) сравнивается с наугад взятым из элиты элементом (“рекомендателем”); если значение x у претендента не меньше, чем у рекомендателя, то претендент становится членом элиты; если меньше, то образуется новая пара “претендент – рекомендатель”. В этом случае направление изменения качества элитной группы определяется тем, какие элементы (худшие или лучшие) дольше существуют в группе. Если дольше “живут” худшие (как шарик в подшипнике, имеющий наименьший диаметр), то элитная группа неминуемо деградирует, поскольку худшие чаще выступают как рекомендатели. Наоборот, при увеличении времени $T(x)$ жизни элемента с ростом величины x худшие элементы выбывают в первую очередь, в рекомендатели чаще попадают элементы повышенного качества, что приводит к непрерывному росту среднего качества элитной группы. При некоторых ограничениях на временные интервалы между заменами существуют невырожденные предельные распределения качества x в элитной группе.

ПРОЦЕДУРЫ “ПРОПОЛКА” И “СНЯТИЕ УРОЖАЯ”

Рассмотрим теперь вторую задачу – **правило “прополка”**. Оно состоит в удалении из элитной группы m наихудших элементов и замене их взятыми наугад m элементами из основной группы. При этом в элита могут попасть как элементы лучшие, чем удаленные при прополке, так и худшие. Однако на следующем шаге прополки снова удаляются m худших элементов элитной группы, так что при $m < n$ худшие в элите не задерживаются, а лучшие из новых остаются. В результате элитная группа прогрессирует, ее распределение $F_k^3(x|m, n)$ с

ростом числа прополок k сходится к некоторому предельному распределению $F_\infty^3(x|m, n)$, зависящему от величины m , $0 < m < n$. Интересно, что наилучшая стратегия прополки, дающая наивысшее предельное среднее качество элитной группы, состоит в удалении из элитной группы только одного наихудшего элемента на каждом шаге.

Процедуру, обратную “прополке”, когда из группы удаляются t наилучших элементов, А.Н. Ефимов называет **правилом “снятия урожая”**. При этом эффекты противоположны тем, которые мы наблюдали при прополке: наибольший вред группе приносит удаление на каждом шаге только одного наилучшего элемента.

Процедура “претендент – рекомендатель” реализует в некотором смысле одинаковое обращение с исходной совокупностью и с элитной группой. Процедуры “прополка” и “снятие урожая” предполагают основную активность внутри элитной группы.

ПРОЦЕДУРА “ДЕЛЕГИРОВАНИЕ”

Рассмотрим теперь возможность внешней активности в формировании элиты. Вариант такой процедуры, называемой **правилом “делегирования”**, состоит в следующем:

- 1) из исходной совокупности случайным образом выбирают N элементов – делегирующую выборку;
- 2) делегирующую выборку упорядочивают по величине x ;
- 3) элемент с наибольшим рангом зачисляют в формируемую элитную группу.

На этапе формирования элитной группы процедуру повторяют n раз – столько, сколько вакансий должно быть заполнено. С помощью теории порядковых статистик удается связать величины a , ? и N :

$$N \geq \frac{\ln \beta}{\ln F(a)}, \quad (2)$$

т.е. чем выше требования к качеству элитной группы, тем больше должен быть объем делегирующей выборки. Принципиальное отличие “делегирования” от процедуры “претендент – рекомендатель” состоит в отказе от сравнения делегатов с членами элиты. Далее, оказывается, что “делегирование” не только позволяет сформировать элиту любого нужного качества, но и предохранить элитную группу от деградации даже при большей жизнеспособности плохих элементов.

Результаты решения приведенных задач имеют некоторые аналогии в общественных процессах. Однако рассмотренные модели слишком просты, чтобы можно было говорить о количественной теории социальных систем. А.Н. Ефимов пишет [13]: “Во-первых, такую теорию еще неплохо было бы создать объединенными усилиями социологов, экономистов, юристов и математиков. Пока что есть лишь несколько простейших моделей. Их анализ показывает, что построение элитных групп, не подверженных деградации, в принципе возможно. Это – во-вторых, и это вселяет надежды”.

Подведем итог

Итеративный вариант выбора – There are different options for селекция, отбор – имеет не- iterative choice or selection. сколько постановок. Интересно, Interestingly, even in the case of что даже в том случае, когда special selection intended to отбор направлен на получение create a group of high quality групп повышенного качества (an “elite group”), the actual (“элитных”), действительные evolution of group quality can изменения группы могут про- tend either toward improvement исходить как в сторону улуч- or toward deterioration of the шения ее среднего качества, так group. Even the simplest models и в сторону ухудшения. Уже reveal the causes for this phe- простейшие модели указывают nomenon: changes in the quality на причины этого: изменение of each specimen with time, качества каждого элемента со временем, правила выбытия из sion from the elite group, кон- элитной группы и вхождения в tamination of the elite group, нее, засорение элитной группы and so on. These characteristics и пр. Эти особенности следует must be taken into account in учитывать и в реальных ситуа- real-life selection as well. циях отбора.

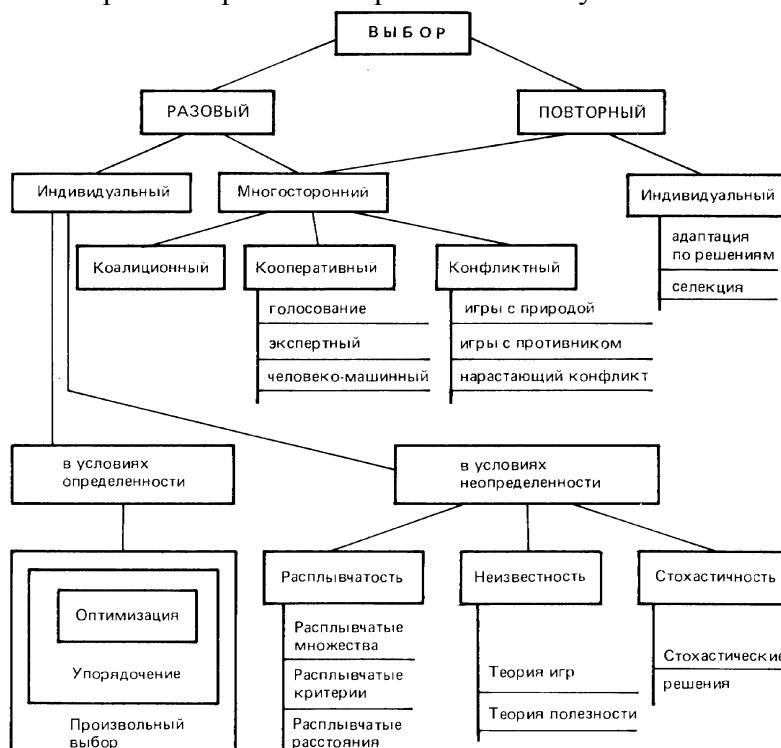
Summary

ЗАКЛЮЧЕНИЕ

На рис. 7.9 приведена схема, на которой изображены все упомянутые в данной главе задачи выбора, указаны их общность и различия в иерархической классификации. Кроме наглядности взаимосвязи этих задач схема выявляет тот факт, что большое количество задач осталось вообще нерассмотренным, поскольку иерархия далеко не полная. Причины того, что те или иные задачи не вошли в схему, могут быть различными.

ЛИТЕРАТУРА

1. Айвазян С.А., Енюков И.С. О содержании и структуре пакета программ по прикладному статистическому анализу. – В сб.: Материалы Всесоюзн. школы “Программно-алгоритмическое обеспечение прикладного многомерного статистич. анализа”. – Ереван: ВЦ Госплана Арм. ССР, 1979. С. 50 – 78.
2. Айзерман М.А., Завалишин Н.В., Пятницкий Е.С. Глобальные функции множеств в теории выбора альтернатив // Автоматика и телемеханика. 1977. № 3. С. 111 – 125; № 5. С. 103 – 113.
3. Александров В.В., Горский Н.Д., Поляков А.О. Пакет прикладных программ АЛПОГОР. – Л.: Физ.-техн. ин-т, 1978.
4. Беллман Р., Задэ Л. Принятие решений в расплывчатых условиях. – В кн.: Вопросы анализа



7.9 ————— Взаимосвязь задач выбора, рассмотренных в гл. 7

и процедуры принятия решений. – М.: Мир. 1976. С. 172 – 215.

5. Блэкuell Д., Гиришик М.А. Теория игр и статистических решений. – М.: ИЛ, 1958.
6. Борисов А.Н., Левченков А.С. Методы интерактивной оценки решений. – Рига: Зинатне, 1982.
7. Ватель И.А., Ерецко Ф.И. Математика конфликта и сотрудничества. – М.: Знание, 1973.
8. Ващенко Н.Д., Гладун В.П., Стогний А.А. Применение системы АНАЛИЗАТОР в научно-исследовательских работах // УС и М. 1978. № 3. С. 104 – 107.
9. Винер Н. Кибернетика. – М.: Сов. радио, 1968.
10. Винер Н. Кибернетика и общество. – М.: ИЛ, 1958.

11. Гаек Я., Шидак З. Теория ранговых критериев. – М.: Наука, 1971.
12. Дмитриев Ю.Г., Устинов Ю.К. Статистическое оценивание распределений вероятностей с учетом дополнительной информации. – Томск: ТГУ, 1988.
13. Ефимов А.Н. Элитные группы, их возникновение и эволюция // Знание – сила. 1988. № 1. С. 56 – 64.
14. Ефимов А.Н., Кутеев В.М. Исследование и моделирование некоторых свойств элитных групп // Изв. АН СССР. Техническая кибернетика. 1980. № 3. С. 177 – 185.
15. Ефимов А.Н., Кутеев В.М. Ранговые процедуры управления эволюцией элитных групп // Изв. АН СССР. Техническая кибернетика. 1980. № 6. С. 3 – 12.
16. Загоруйко Н.Г., Елкина В.Н., Лбов Г.С. Алгоритмы обнаружения эмпирических закономерностей. – Новосибирск: Наука, 1985.
17. Каипов В.Х., Селюгин А.А., Дубровский С.А. О решении некоторых задач управления методами прикладной теории расплывчатых множеств. Podstawy Sterowania, T. 15 (1985). Z. 1 – 2, Р. 19 – 40.
18. Ким А.Н., Бузурханов В.Б., Камилов М.М. Программно-распознающий комплекс ПРАСК. – “Алгоритмы и программы”. Ташкент, 1975. Вып. 17. С. 3 – 12.
19. Лбов Г.С. Методы обработки разнотипных экспериментальных данных. – Новосибирск: Наука, 1981.
20. Леман Э. Проверка статистических гипотез. – М.: Наука, 1964.
21. Льюс Р.Д., Райфа Х. Игры и решения. Введение и критический обзор. – М.: ИЛ., 1961.
22. Макаров И.М. и др. Теория выбора и принятия решений. – М.: Наука 1987.
23. Машунин Ю.К. Методы и модели векторной оптимизации. – М.: Наука, 1986.
24. Миркин Б.Г. Проблема группового выбора. – М.: Наука, 1974.
25. Молодцов Д.А. Устойчивость принципов оптимизации. – М.: Наука, 1987.
26. Орлов А.И. Задачи оптимизации и нечеткие переменные. – М.: Знание, 1980.
27. Орловский С.А. Проблемы принятия решений при расплывчатой информации. – М.: Наука, 1981.
28. Поспелов Д.А. Логико-лингвистические модели в системах управления. – М.: Энергоиздат, 1981.
29. Райфа Г. Анализ решений. Введение в проблемы выбора в условиях неопределенности. – М.: Наука, 1977.
30. Раудис Ш., Пикялис В., Юшкевичус К. Система оперативной разработки распознающих алгоритмов (СОПРА) // Статистические проблемы управления, 1977. Вып. 27. С. 3 – 27.
31. Растрогин Л.А. Современные принципы управления сложными объектами. – М.: Сов. радио, 1980.
32. Сироджса И.Б. и др. Пакет прикладных программ классификационной обработки данных (ППП КОД-2). – В кн.: Матем. методы анализа динамических систем. Харьков, 1983. Вып. 7. С. 127 – 134.
33. Таракенко Ф.П. Непараметрическая статистика. – Томск: ТГУ, 1976.
34. Таракенко Ф.П. О принципиальных трудностях балльной оценки научной деятельности // Вестник АН СССР. 1976. № 6. С. 69 – 75.
35. Тийт Э.А., Тоодинг Л.М. Опыт анализа научных исследований при помощи пакетов программ в ТГУ. – В сб.: Материалы Всесоюзн. школы “Программно-алгоритмическое обеспечение прикладного многомерного статистич. анализа”. – Ереван: ВЦ Госплана Арм. ССР, 1979. С. 120.
36. Тутубалин В.Н. Теория вероятностей. Краткий курс и научно-методические замечания. – М.: МГУ, 1972.
37. Тюрин Ю.Н. Статистические методы анализа экспертных оценок. – М.: Наука, 1977.
38. Фишберн П. Теория полезности для принятия решений. – М.: Наука, 1978.
39. Худсон Д. Статистика для физиков. – М.: Мир, 1967.

40. Хьюбер П. Робастность в статистике. – М.: Мир, 1984.
41. Шапиро Д.И. Принятие решений в системах организационного управления: использование расплывчатых категорий. – М.: Энергоатомиздат, 1983.
42. Элти Дж., Кумбс М. Экспертные системы: концепции и примеры. – М.: Финансы и статистика, 1987.
43. Янч Э. Прогнозирование научно-технического процесса. – М.: Прогресс, 1970.
44. Ackoff R.L. A Theory of Practice in the Social Systems Sciences. Paper to an International Roundtable, IIASA, Laxenburg, Austria, 6 -8 Nov., 1986.
45. Ester J. Concept of efficiency and fuzzy aggregation rules. In: Large-Scale Modelling and Interactive Decision Analysis. (Eds.: Fandel G., Grauer M., Kurzhanski A., Wierzbicki A.), Berlin, Springer, 1986.
46. Ester J., Troeltzsch F. On generalized notions of efficiency in MCDM. // Systems Anal. Model. Simul., 1986, 3, Heft 2.
47. Hatry H.P. Measuring the Effectiveness of Nondefence Public Programs. Operations Research, 1970, 18(5), 774.
48. Lewandowski A., Werzbicki A. Theory, Software and Testing Examples in Decision Support Systems. Working paper WP-88-071, International Institute for Applied System Analysis, Laxenburg, Austria, 1988.
49. Seo F., Sakawa M. Fuzzy Multivariate Utility Analysis for Collective Choise. IEEE Trans., 1985, SMC, vol. 15, N 1, 45 – 53.
50. Zadeh L.A. Similarity relations and fuzzy orderings.//Inform. Sci., 1971. Vol. 3, P. 177 – 200.

УПРАЖНЕНИЯ

7.1

- Составьте полный граф дерева классификации задач выбора; отметьте на нем задачи, которые уже были рассмотрены; попробуйте объяснить, почему не решались неотмеченные задачи.

§ 7.2

- Рассмотрите какую-нибудь многокритериальную задачу и установите, какой из методов ее решения лучше отвечает поставленной вами цели.

§ 7.3

- Покажите, что если график предпочтения сильно транзитивен и антирефлексивен, то выбор сводится к однокритериальной задаче.

§ 7.4

- Каким ограничениям должна удовлетворять функция выбора, чтобы она могла описывать какой-нибудь определенный вами тип выбора?

§ 7.5

- Найдите в литературе доказательство теоремы о невозможности и проследите в нем, где именно существенно используется то, что предпочтения даются в порядковой шкале. Возможна ли тупиковая ситуация при выражении индивидуальных предпочтений в числовой шкале?

§ 7.6

- Придумайте правдоподобную жизненную ситуацию, которая укладывается в рамки теоретико-игровой модели. Проанализируйте, какие упрощения и допущения пришлось при этом сделать.

§ 7.7

- Обсудите подробно, когда реальную неопределенность можно считать вероятностной и что можно сделать, чтобы проверить статистичность реальных данных.

§ 7.8

- Попробуйте сформулировать (и по возможности решить) различные варианты многокритериальных расплывчатых задач.
§ 7.9
- Тема для обсуждения: роль и место оптимизационных задач в системном анализе.
§ 7.10
- Сравните разные формы экспертных процедур по степени их пригодности в различных условиях.
- Попробуйте провести вашей учебной группой оценку альтернатив методом “Делфи”.
§ 7.11
- Обсудите различие между базой знаний и системой поддержки решений.
§ 7.12
- Попробуйте придумать свои, отличные от описанных процедуры формирования и пополнения элитных групп. Какова при этом, по вашему представлению, динамика элитной группы?

Вопросы для самопроверки

1. Что значит “сделать выбор”?
2. В чем главные отличия в описании выбора на трех языках: критериальном, бинарных отношений, функций выбора?
3. Почему разные постановки задачи многокритериального выбора приводят в общем случае к различным решениям?
4. Как определяется оптимальность по отношению R ?
5. В чем заключается парадокс Эрроу?
6. При каких условиях меньшинство может навязать свою волю, несмотря на принятие решений большинством голосов?
7. Что позволяет выбор в условиях неопределенности исхода рассматривать как игру?
8. На каком множестве осуществляется выбор в случае статистической неопределенности?
9. Каковы основные правила статистической “техники безопасности”?
10. Как решается задача выбора при расплывчатой неопределенности, если критериальные функции отождествляются с функциями принадлежности?
11. Какие причины сужают возможности оптимизации в решении реальных проблем?
12. Какие факторы влияют на работу экспертов?
13. Какими достоинствами обладают человеко-машические способы выбора?
14. Почему элитная группа может деградировать?

AGGREGATION
агрегирование
DECOMPOSITION
декомпози-
ция
COMPLETENESS
полнота
PROBLEM-
CONTAINING
SYSTEM проблемо-
содержащая
система
PROBLEM-
SOLVING SYS-
TEM проблемораз-
решающая систе-
ма

ДЕКОМПОЗИЦИЯ И АГРЕГИРОВАНИЕ КАК ПРОЦЕДУРЫ СИСТЕМНОГО АНАЛИЗА

ГЛАВА ВОСЬМАЯ

§ 8.1. АНАЛИЗ И СИНТЕЗ В СИСТЕМНЫХ ИССЛЕДОВАНИЯХ

В § 1.3 мы уже рассматривали вопрос о соотношении анализа и синтеза в человеческом познании. Их единство позволяет понять окружающий мир. Это относится ко всем отраслям знаний и, в частности, к получившей название системного анализа. В данной главе будут рассмотрены технические аспекты аналитического и синтетического методов исследования систем, т.е. будет акцентировано внимание на том, как выполняются операции разделения целого на части и объединения частей в целое и почему они выполняются именно так. Иными словами, мы обсудим, в какой степени анализ и синтез на сегодняшний день могут быть алгоритмизированы.

Аналитический метод, изначально, органически присущий человеческому мышлению, в явной форме был осознан, выделен и сформулирован как самостоятельный технический прием познания в XVII в. представителями рационализма. Так, Р. Декарт писал:

“Расчлените каждую изучаемую вами задачу на столько частей (...), сколько потребуется, чтобы их было легко решить” [3].

Успех и значение аналитического метода состоит не только и не столько в том, что сложное целое расчленяется на все менее сложные (и в конечном счете простые) части, а в том, что, будучи соединены надлежащим образом, эти части снова образуют единое целое. Этот момент агрегирования частей в целое является конечным этапом анализа, поскольку *лишь только после этого мы можем объяснить целое через его части – в виде структуры целого* (см. § 3.5).

Аналитический метод имеет колоссальное значение в науке и на практике. Разложение функций в ряды, дифференциальное и интегральное исчисление, разбиение неоднородных областей на однородные с последующим “сшиванием” решений – в математике, анализаторы спектров, всевозможные фильтры, исследования атомов и элементарных частиц – в физике; анатомия и нозология – в медицине; значительная часть схемотехники, конвейерная технология производства – все атрибуты эффективности анализа. Успехи аналитического метода привели к тому, что термины “анализ” и “научное исследование” стали восприниматься как синонимы. Важнейшей формой познания стала считаться причинно-следственная закономерность, причина является необходимым и достаточным условием осуществления процесса.

СОЧЕТАНИЕ АНАЛИЗА И СИНТЕЗА В СИСТЕМНОМ ИССЛЕДОВАНИИ

Многие философы и естествоиспытатели обращали внимание на то, что роль синтеза не сводится только к “сборке деталей”, полученных при анализе. Среди специалистов по системному анализу особенно настойчиво выделяет эту мысль Р. Акофф. Он подчеркивает значение целостности системы; эта целостность нарушается при анализе, при расчленении системы утрачиваются не только существенные свойства самой системы (“разобранный автомобиль не поедет, расчлененный организм не способен жить”), но исчезают и существенные свойства ее частей, оказавшихся отделенными от нее (“оторванный руль не рулит, отделенный глаз не видит”). Поэтому, отмечает Р. Акофф, результатом анализа является лишь вскрытие структуры, знание о том, как система работает (“ноу – хай”), но не понимание того, почему и зачем она это делает.

“Синтетическое мышление требует объяснить поведение системы. Оно существенно отличается от анализа. На первом шаге анализа вещь, подлежащая объяснению, разделяется на части; в синтетическом мышлении она должна рассматриваться как часть большего целого. На втором шаге анализа объясняются содержимые части; в синтетическом мышлении объясняется содержащее нашу вещь целое. На последнем шаге анализа знание о частях агрегируется в знание о целом; в синтетическом мышлении понимание содержащего целого дезагрегируется для объяснения частей. Это достигается путем вскрытия их ролей или функций в целом. Синтетическое мышление открывает не структуру, а функцию; оно открывает, почему система работает так, а не то, как она делает это” [24].

Таким образом, не только аналитический метод невозможен без синтеза (на этом этапе части агрегируются в структуру), но и синтетический метод невозможен без анализа (необходима дезагрегация целого для объяснения функций частей). Анализ и синтез дополняют, но не заменяют друг друга. Системное мышление совмещает оба указанных метода.

Такое положение приводит к некоторой “терминологической напряженности”, имеющейся в преподавании системных знаний и в общении между специалистами. Акофф предлагает преодолеть противоречивость в названиях “аналитического метода”, содержащего синтетическую стадию установления структуры, и “синтетического метода”, включающего анализ функций частей, употребляя обобщающие термины “редукционизм” для первого и “экспансионизм” для второго. Существуют и более радикальные терминологические предложения: “Сам термин “системный анализ” несостоителен, поскольку слово “анализ” противоречит понятию целостности, содержащемуся в термине “система”. Настало время заменить устаревший термин, пустив в оборот новый, непротиворечивый по существу. Возможно, подойдет название “исследование систем”, как не содержащее противоречия” [25].

Так как в настоящее время все еще преобладает аналитический (редукционистский) подход в исследованиях, то имеет смысл привести дополнительные аргументы, привлекающие внимание к синтетическим (экспансионистским) методам.

ОСОБЕННОСТИ СИНТЕТИЧЕСКИХ МЕТОДОВ

Во-первых, аналитический метод приводит к достижению наивысших результатов, если целое удается разделить на независимые друг от друга части, поскольку в этом случае их отдельное рассмотрение позволяет составить правильное представление об их вкладе в общий эффект (как в случае функциональных ортогональных рядов, интегрального исчисления, мозаики, накопления денег и пр.). Однако случаи, когда система является “суммой” своих частей, не правило, а редчайшее исключение. Правилом же является то, что вклад данной части в общесистемный эффект зависит от вкладов других частей. Поэтому, например, если заставить каждую часть функционировать наилучшим образом, то в целом эффект не будет наивысшим (см. § 7.9). Можно сказать [24], что, отобрав лучшие в мире карбюратор, двигатель, фары, колеса и т.д., мы не только не получим самого лучшего автомобиля, но вообще не сможем собрать машину, так как детали машин разных марок не подойдут друг к другу. Итак, при анализе “неаддитивных” систем следует делать акцент на рассмотрение не отдельных частей, а их взаимодействия. Это существенно более трудная задача. Примером является управление “неаддитивной” системой, которое окажется более эффективным, если управлять не действиями ее частей отдельно, а взаимодействиями между ними.

Во-вторых, идеалом, конечной целью аналитического метода является установление причинно-следственных отношений между рассматриваемыми явлениями. Нечто считается公认的, полностью понятым лишь в том случае, если известна его причина (совокупность условий, необходимых и достаточных для реализации следствия). Однако это далеко не всегда достижимо. Даже в тех случаях, когда имеет место причинно-следственное описание (т.е. когда условия, входящие в причину, действительно перечислимы), все остальное должно быть исключено. Для причинно-следственного отношения не существует понятия окружающей среды, так как для следствия ничего, кроме причины, не требуется. Примером служит

закон свободного падения тел, справедливый, если отсутствуют все другие силы, кроме силы тяготения. Однако когда мы имеем дело со сложными системами, исключить “ненужные”, “неинтересные” взаимодействия бывает невозможно не только практически, но и абстрактно (при необходимости сохранить адекватность модели; см. § 2.6). Имеется два способа описать такую ситуацию. Один состоит в отображении “беспринной” компоненты поведения системы либо “объективной случайностью”, либо “субъективной неопределенностью” (происходящей от незнания), либо их сочетанием (см. § 6.4 и 7.7). Другой вытекает из синтетического, экспансионистского метода и состоит в признании того, что отношение “причина – следствие” является не единственным возможным и приемлемым описанием (объяснением) взаимодействия. Более адекватной моделью взаимодействия оказывается отношение “продуцент – продукт”, характеризуемое тем, что продуцент является необходимым, но не достаточным условием для осуществления продукта. Например, желудь является для дуба продуцентом, а не причиной, поскольку кроме желудя для произрастания дуба необходимы почва, влага, воздух, свет, тепло, сила тяготения и т.д. Таким образом, *для получения продукта необходимы и другие условия, которые и образуют окружающую среду*. Причинное, свободное от среды объяснение является предельным случаем продуцентного, идеалом, к которому можно приближаться, но достичь которого можно не всегда и не всегда необходимо.

Как бы то ни было, и при аналитическом, и при синтетическом подходе наступает момент, когда необходимо разложить целое на части либо объединить части в целое. Будем называть эти операции соответственно **декомпозицией и агрегированием**. Далее рассмотрим технические аспекты выполнения этих операций.

Подведем итог

Summary

Анализ и синтез являются не-
элементарными действиями, nonelementary themselves, and
которые содержат более про- both include the more simple
стые операции декомпозиции и operations of decomposition and
агрегирования. Эти операции, aggregation. The latter operations
в свою очередь, можно алго- can in their turn be algorithmized
ритмизировать, что и сделано в and it is done in this chapter.
данной главе.

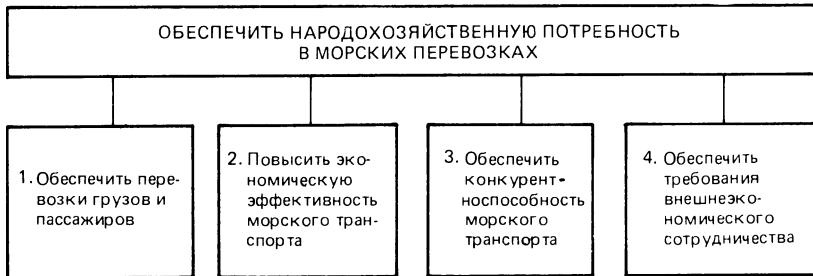
§ 8.2. МОДЕЛИ СИСТЕМ КАК ОСНОВАНИЯ ДЕКОМПОЗИЦИИ

Основной операцией анализа является разделение целого на части. Задача распадается на подзадачи, система – на подсистемы, цели – на подцели и т.д. При необходимости этот процесс повторяется, что приводит к иерархическим древовидным структурам. Обычно (если задача не носит чисто учебного характера) объект анализа сложен, слабо структурирован, плохо формализован, поэтому операцию декомпозиции выполняет эксперт. Если поручить анализ одного и того же объекта разным экспертам, то полученные древовидные списки будут различаться. Качество построенных экспертами деревьев зависит как от их компетентности в данной области знаний, так и от применяемой методики декомпозиции.

Обычно эксперт легко разделяет целое на части, но испытывает затруднения, если требуется доказательство полноты и безызбыточности предлагаемого набора частей. Стремясь перейти от чисто эвристического, интуитивного подхода к более осознанному, алгоритмическому выполнению декомпозиции, мы должны объяснить, почему эксперт разделяет целое именно так, а не иначе, и именно на данное, а не на большее или меньшее число частей. Объяснение состоит в том, что *основанием всякой декомпозиции является модель рассматриваемой системы* [13].

СОДЕРЖАТЕЛЬНАЯ МОДЕЛЬ КАК ОСНОВАНИЕ ДЕКОМПОЗИЦИИ

Остановимся на этом важном соображении подробнее. Операция декомпозиции представляется теперь как сопоставление объекта анализа с некоторой моделью, как выделение в



8.1 ————— Первый уровень дерева целей из примера 1

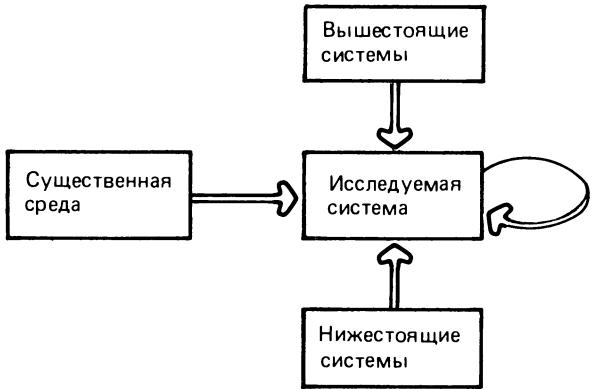
нем того, что соответствует элементам взятой модели. Поэтому на вопрос, сколько частей должно получиться в результате декомпозиции, можно дать следующий ответ: столько, сколько элементов содержит модель, взятая в качестве основания. Вопрос о полноте декомпозиции – это вопрос завершенности модели.

Пример 1 [20]. В начале 70-х годов проводились работы по системному анализу целей развития морского флота. Первый уровень дерева целей выглядел в виде схемы, изображенной на рис. 8.1.

Декомпозиция проведена по модели входов организационной системы (рис. 8.2), которая включает входы: от “нижестоящих” систем (здесь клиентуры – подцель 1); от “вышестоящих” систем (здесь народного хозяйства в целом – подцель 2); от “существенной среды” (в данном случае от флотов капиталистических государств – подцель 3 и социалистических государств – подцель 4). Очевидно, что такая декомпозиция неполна, поскольку отсутствует подцель, связанная с собственными интересами морского флота. Это, по-видимому, не столько ошибка экспертов-аналитиков (в [20, с. 120] абстрактно упоминается дерево целей социально-экономического развития), сколько результат преобладавшего тогда “остаточно-го” подхода к проблемам быта, присущего тогдашнему стилю руководства. Через 15 лет пришлось говорить о том, что неучет компоненты соцкультбыта создал серьезные проблемы в работе флота в целом [9].

Итак, объект декомпозиции должен сопоставляться с каждым элементом модели-основания. Однако и сама модель-основание может с разной степенью детализации отображать исследуемый объект. Например, в системном анализе часто приходится использовать модель типа “жизненный цикл”, позволяющую декомпозировать анализируемый период времени на последовательные этапы от его возникновения до окончания. С помощью такой декомпозиции шахматную партию можно разбить на дебют, миттельшпиль и эндшпиль; в жизни человека принято различать молодость, зрелость и старость, но можно выделять и более мелкие этапы, например детство, отрочество и юность. Такое же разнообразие может иметь место и при декомпозиции жизненного цикла любой проблемы. Разбиение на этапы дает представление о последовательности действий, начиная с обнаружения проблемы и кончая ее ликвидацией (иногда такую последовательность рассматривают как “алгоритм системного анализа”; насколько это справедливо, мы обсудим в следующей главе).

Пример 2. В табл. 8.1 приведены декомпозиции жизненного цикла проблем, разработанные крупными специалистами по системному анализу – С.Л. Оптнером [10], С. Янгом [22], Н.П. Федоренко [19], С.П. Никаноровым [10, вступительная статья], Ю.И. Черняком [20]. В качестве упражнения рекомендуем обсудить различия приводимых декомпозиций. При этом следует иметь в виду, что каждый из авторов впоследствии проводит дальнейшее разбиение каждого этапа (т.е. выполняет декомпозицию следующих уровней). Приводимые декомпозиции являются наглядным примером различий, возникающих на эвристических этапах системного анализа.



8.2 ————— Схема входов организационной системы

Установив, что декомпозиция осуществляется с помощью некоторой модели, сквозь которую мы как бы рассматриваем расчленяемое целое, далее следует ответить на естественно возникающие вопросы: 1) модели какой системы следует брать в качестве оснований декомпозиций? 2) какие именно модели надо брать?

Выше уже упоминалось, что основанием декомпозиции служит модель “рассматриваемой системы”, но какую именно систему следует под этим понимать? Всякий анализ проводится для чего-то, и именно эта цель анализа и определяет, какую систему следует рассматривать. Система, с которой связан объект анализа, и система, по моделям которой проводится декомпозиция, не обязательно совпадают, и хотя они имеют определенное отношение друг к другу, это отношение может быть любым: одна из них может быть подсистемой или надсистемой для другой, они могут быть и разными, но как-то связанными системами.

Таблица 8.1. Этапы ликвидации проблем

По С.Л. Оптнеру П

По С. Янгу

Іо. Н.П. Фе-
доренко

По С.П. Никанорову

По. Ю.И. Черняку

1.Идентификация симптомов	1.Определение цели	1.Формулирование проблем	1.Обнаружение проблемы	1.Анализ проблемы
2.Определение актуальности проблемы	2.Выявление организаций	2.Определение целей	2.Оценка актуальности проблемы	2.Определение системы
3.Определение целей	3.Сбор информации о проблемах	3.Сбор информации о проблемах	3.Анализ ограниченной проблемы	3.Анализ структуры системы
4.Определение структуры системы и ее дефектов	3.Диагноз	4.Разработка максимального количества решений	4.Определение критериев	4.Формулирование общей цели и критерия
5.Определение возможностей	4.Поиск альтернатив	5.Оценка и выбор альтернатив	5.Анализ существующей системы	5.Декомпозиция цели, выявление потребности в ресурсах,
6.Нахождение альтернатив	5.Отбор альтернатив	6.Согласование альтернатив	6.Поиск возможностей (альтернатив).	6.Выявление ресурсов, композиция целей
7.Оценка альтернатив	6.Построение модели в виде уравнений, программ или сценария	7.Выбор альтернатив	7.Выбор альтернатив	7.Прогноз и анализ будущих условий
8.Выработка решения	7.Утверждение решения	8.Подготовка к вводу в действие	8.Обеспечение признания	8.Оценка целей и средств
9.Признание решения	8.Приятие решения	9.Оценка затрат	9.Принятие решения (принятие формальной ответственности)	9.Отбор вариантов
10.Запуск процесса решения	9.Испытание чувствительности решения	10.Реализация решения	10.Реализация решения	10.Диагноз существующей системы
11.Управление процессом реализации решения	9.Управление при менением решения	11.Определение результатов решения	11.Определение результатов решения	11.Построение комплексной программы развития
12.Оценка реализации и ее последствий	10.Проверка эффективности	12.Проектирование организации для достижения целей	12.Проектирование организации для достижения целей	

Например, анализируя цель “выяснить этиологию и патогенез ишемической болезни сердца”, в качестве исследуемой системы можно взять сердечно-сосудистую систему, а можно выбрать конкретный кардиологический институт. В первом случае декомпозиция будет порождать перечень подчиненных подцелей научного, во втором – организационного характера.

Отметим также, что иногда в качестве оснований декомпозиции полезно не только перебирать разные модели целевой системы, но и брать сначала модели надсистемы, затем самой системы и, наконец, подсистем. Например, при системном анализе функций Минвуза РСФСР декомпозиция глобальной цели высшего образования страны сначала проводилась по моделям вузовской системы в целом, а в конце – по моделям функционирования министерского аппарата. Можно также рассматривать и такую процедуру анализа, когда перед каждым очередным актом декомпозиции заново ставится вопрос не только о том, по какой модели проводить декомпозицию, но и о том, не следует ли взять модель иной системы, нежели ранее.

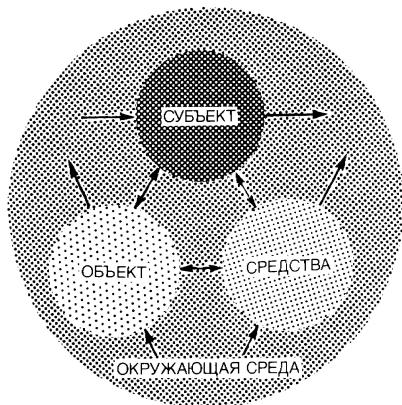
Однако чаще всего в практике системного анализа в качестве глобального объекта декомпозиции берется нечто, относящееся к проблемосодержащей системе и к исследуемой проблеме, а в качестве оснований декомпозиции берутся модели проблеморазрешающей системы.

Перейдем теперь к рассмотрению вопроса о том, какие модели брать за основания декомпозиции. Прежде всего напомним, что при всем практических неизбежном многообразии моделей формальных типов моделей немного: это модели “черного ящика”, состава, структуры, конструкции (структурной схемы) – каждая в статическом или динамическом варианте (см. гл. 3). Это позволяет организовать нужный перебор типов моделей, полный или сокращенный, в зависимости от необходимости.

СВЯЗЬ МЕЖДУ ФОРМАЛЬНОЙ И СОДЕРЖАТЕЛЬНОЙ МОДЕЛЯМИ

Однако основанием для декомпозиции может служить только конкретная, содержательная модель рассматриваемой системы. Выбор формальной модели лишь подсказывает, какого типа должна быть модель-основание; формальную модель следует наполнить содержанием, чтобы она стала основанием для декомпозиции. Это позволяет несколько прояснить вопрос о полноте анализа, который всегда возникает в явной или неявной форме.

Полнота декомпозиции обеспечивается полнотой модели-основания, а это означает, что прежде всего следует позаботиться о полноте формальной модели. Благодаря формальности, абстрактности такой модели часто удается добиться ее абсолютной полноты.



8.3 ————— Общая схема деятельности

Пример 3. Схема входов организационной системы на рис. 8.2 является полной: к ней нечего добавить (перечислено все, что воздействует на систему), а изъятие любого элемента лишит ее полноты.

Пример 4. Формальный перечень типов ресурсов состоит из энергии, материи, времени, информации (для социальных систем добавляются кадры и финансы). При анализе ресурсного обеспечения любой конкретной системы этот перечень не дает пропустить что-то важное.

Пример 5. К числу полных формальных моделей относится марксова схема любой деятельности человека, которая в “Капитале” применялась для анализа процесса труда (рис. 8.3). В схеме выделены: субъект деятельности; объект, на который направлена деятельность; средства, используемые в процессе деятельности; окружающая среда; все возможные связи между ними.

Пример 6. Если в качестве модели жизненного цикла принять формулировку “все имеет начало, середину и конец”, то такая модель является формально полной. Конечно, эта модель настолько обща, что оказывается мало полезной во многих конкретных случаях. Так, при рассмотрении жизненного цикла проблем (см. табл. 8.1) приходится использовать более детальные модели.

Итак, полнота формальной модели должна быть предметом особого внимания. Поэтому одна из важных задач информационного обеспечения системного анализа и состоит в накоплении наборов полных формальных моделей (в искусственном интеллекте такие модели носят название *фреймов*).

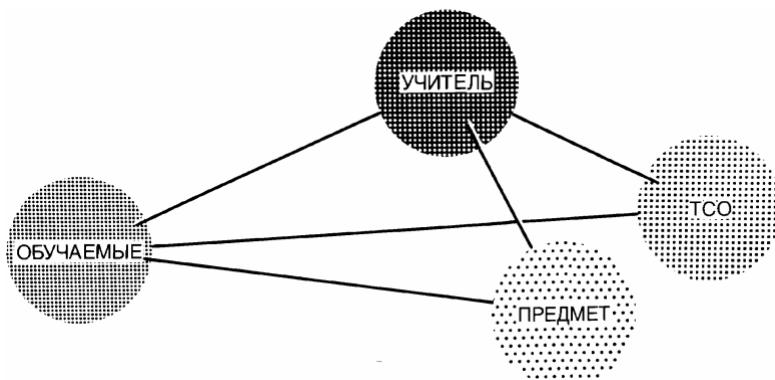
ПРОБЛЕМА ПОЛНОТЫ МОДЕЛЕЙ

Полнота формальной модели является необходимым, но не достаточным условием для полноты декомпозиции. В конечном счете все зависит от полноты содержательной модели, которая строится “по образу” формальной модели, но не тождественна ей. Фрейм лишь привлекает внимание эксперта к необходимости рассмотреть, что именно в реальной системе соответствует каждому из составляющих фрейм элементов, а также решить, какие из этих элементов должны быть включены в содержательную модель. Это очень ответственный момент (ведь то, что не попадет в модель-основание, не появится в дальнейшем анализе) и

очень трудный (заранее не всегда “очевидно”, что данная компонента должна войти в основание).

Для иллюстрации вернемся к примеру 1. Фреймовая модель входов оргсистемы (см. рис. 8.2) рекомендует, в частности, определить конкретно, что именно понимается под “существенной средой”, т.е. взаимодействие с какими реальными системами не своего ведомства должно войти в основание. Судя по результату анализа, его авторы учитывали только взаимодействие морского флота с флотами других государств. Для каких-то целей этого достаточно, но ясно также, что в других случаях может потребоваться учет взаимодействий с суходутным транспортом (железнодорожным и автомобильным в отдельности), речным и воздушным флотами. Если возникнет вопрос о ресурсах, то потребуется учет связей с ведомствами, производящими топливо и энергию, продукты питания, всевозможную технику, услуги и т.д. Таким образом, вопрос достаточной степени детализации содержательных моделей в отличие от фреймовых всегда остается открытым. Чтобы сохранить полноту и возможность расширения содержательной модели, можно рекомендовать осуществлять логическое замыкание перечня ее элементов компонентой “все остальное”. Эта компонента, как правило, окажется “молчашей”, поскольку к ней отнесено все, что кажется несущественным, но ее присутствие будет постоянно напоминать эксперту, что, возможно, он не учел что-то важное.

С проблемой степени детализации модели-основания связан и вопрос удобства – понятие трудно формализуемое, но вполне ощущимое. Поясним это на конкретном примере. Для целей анализа проблем преподавания необходима модель педагогического процесса. В качестве фрейма для нее можно взять модель деятельности вообще (рис. 8.3), придав соответствующую интерпретацию входящим в нее элементам. С помощью такой модели педагогического процесса удалось упорядочить и сопоставить ряд современных направлений в методической работе высшей школы [2]. Однако для анализа организационных аспектов учебного процесса в вузе более удобной оказалась модель, в которой из фреймового элемента “средства” выделены в отдельно учитываемый элемент не только информационное средство “изучаемый предмет” но и “технические средства обучения” (рис. 8.4). Например, можно усмотреть связь “граней пирамиды” на рис. 8.4 с такими организационными аспектами учебного процесса, как аудиторные практические занятия, самостоятельная работа студентов, методическая работа преподавателей, лекционное преподавание. Это лишний раз иллюстрирует целевую предназначность моделей (см. § 2.2): изменение цели моделирования требует изменения модели.



8.4 ————— Схема компонент учебного процесса

Подведем итог

Основанием для декомпозиции является содержательная модель проблеморазрешающей системы. Это означает, что в разделяемом целом мы должны найти часть, соответствующую каждому из элементов модели-основания. Ориентиром для построения содержательной модели (т.е. основания декомпозиции) служат формальные модели известных типов (см. гл. 3). Предметом особого внимания является полнота модели, и существует несколько приемов, помогающих добиться полноты в формальных моделях и повысить полноту содержательных моделей.

§ 8.3. АЛГОРИТИЗАЦИЯ ПРОЦЕССА ДЕКОМПОЗИЦИИ

В предыдущем параграфе мы рассмотрели некоторые аспекты того, каким образом эксперт осуществляет единичный акт разложения целого на части. Теперь можно дать дальнейшие рекомендации по осуществлению всего многоступенчатого процесса декомпозиции, от начальной декомпозиции первого уровня до последнего, завершающего данный этап анализа уровня [13].

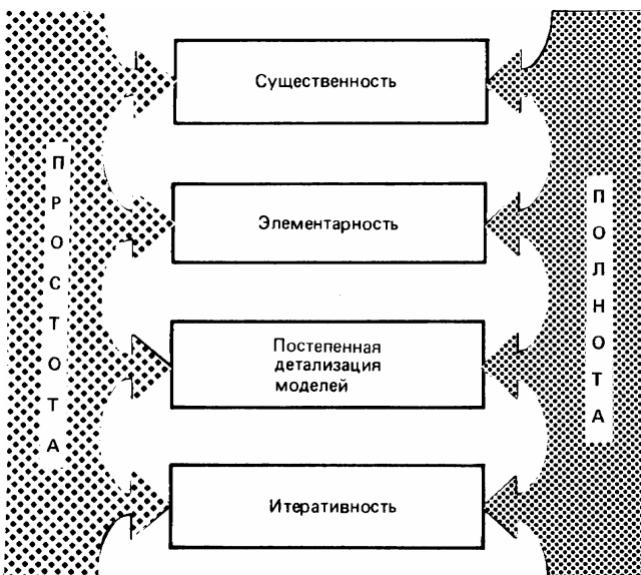
Summary

The basis for decomposing a studied problem is the content model of the problem-solving system. This means that we must find in a decomposed whole the parts corresponding to each element of the basic model. The formal model serves as the pattern for construction of the content model, which is used as the basis for decomposition, and there are several heuristics that help to attain completeness in formal models and to make content models more complete.

КОМПРОМИССЫ МЕЖДУ ПОЛНОТОЙ И ПРОСТОТОЙ

Начнем с обсуждения требований к древовидной структуре, которая получится как итог работы по всему алгоритму. С количественной стороны эти требования сводятся к двум противоречивым принципам: *полноты* (проблема должна быть рассмотрена максимально всесторонне и подробно) и *простоты* (все дерево должно быть максимально компактным – “вширь” и “вглубь”). Эти принципы относятся к количественным характеристикам (размерам) дерева. Компромиссы между ними вытекают из качественного требования – главной цели: свести сложный объект анализа к конечной совокупности простых подобъектов либо (если это не удается) выяснить конкретную причину неустранимой сложности (рис. 8.5).

Принцип простоты требует сокращать размеры дерева. Мы уже знаем, что размеры “вширь” определяются числом элементов модели, служащей основанием декомпозиции. Поэтому принцип простоты вынуждает брать как можно более компактные модели-основания. Наоборот, принцип полноты заставляет брать как можно более развитые, подробные модели. Компромисс достигается с помощью понятия *существенности*: в модель-основание включаются только компоненты, существенные по отношению к цели анализа (релевантные). Как видим, это понятие неформальное, поэтому решение вопроса о том, что же является в данной модели существенным, а что – нет, возлагается на эксперта. Чтобы облегчить работу эксперта, в алгоритме должны быть предусмотрены возможности внесения (в случае необходимости) поправок и дополнений в модель-основание. Одна из таких возможностей заключается в дополнении элементов, которые эксперт считал существенными, еще одним элементом “все остальное”; он может не использоваться экспертом для декомпозиции, но будет постоянно пробуждать у эксперта сомнение в полноте предложенной им модели. Другая возможность состоит в разукрупнении, разбиении отдельных элементов модели-основания в случае необходимости, которая может возникнуть на последующих стадиях анализа. Позже мы вернемся к этому моменту.



8.5 ————— Схема компромиссов между принципами простоты и полноты анализа

Перейдем теперь к вопросу о размерах дерева “вглубь”, т.е. о числе “этажей” дерева, числе уровней декомпозиции. Конечно, желательно, чтобы оно было небольшим (принцип простоты), но принцип полноты требует, чтобы в случае необходимости можно было продолжать декомпозицию как угодно долго до принятия решения о ее прекращении по данной ветви (разные ветви иногда могут иметь различную длину). Такое решение принимается в нескольких случаях. Первый, к которому мы обычно стремимся, наступает, когда композиция привела к получению результата (подцели, подфункций, подзадач и т.п.), не требующего дальнейшего разложения, т.е. результата простого, понятного, реализуемого, обеспеченнего, заведомо выполнимого; будем называть его *элементарным*. Для некоторых задач (например, математических, технических и т.п.) понятие элементарности может быть конкретизировано до формального признака, в других задачах анализа оно неизбежно остается неформальным и проверка фрагментов декомпозиции на элементарность поручается экспертом.

Неэлементарный фрагмент подлежит дальнейшей декомпозиции по другой (не использовавшейся ранее) модели-основанию. Очевидно, что эффективность работы эксперта, размеры получающегося дерева и в конечном счете качество анализа в определенной мере зависят от последовательности, в которой эксперт использует имеющиеся модели. Например, алгоритм декомпозиции, встроенный в компьютерную диалоговую систему, должен, ради удобства эксперта, допускать предъявление моделей в том порядке, который определит сам эксперт. Вместе с тем должен быть предусмотрен и режим совета эксперту, рекомендующий определенный порядок взятия оснований, упрощающий дело (например, тот, в котором излагаются модели систем в гл. 3).

Если эксперт перебрал все фреймы, но не достиг элементарности на какой-то ветви дерева, то прежде всего выдвигается предположение, что дальнейшая декомпозиция может все-

таки довести анализ до получения элементарных фрагментов, и следует дать эксперту возможность продолжить декомпозицию. Такая возможность состоит во введении новых элементов в модель-основание и продолжении декомпозиции по ним. Поскольку новые существенные элементы могут быть получены только расщеплением уже имеющихся, в алгоритме декомпозиции должна быть заложена возможность возврата к использованным ранее основаниям. При этом нет необходимости рассматривать заново все элементы модели, так как обрабатываемый фрагмент находится на ветви, соответствующей только одному элементу каждого основания. Тогда следует рассмотреть возможность расщепления именно этого элемента (например, при рассмотрении системы “вуз” вход “абитуриенты” можно разделить на абитуриентов со стажем и без него, выход “научная информация” – на выходы “монографии”, “статьи”, “отчеты по НИР”, “заявки на изобретения” и т.п.). На этой же стадии можно рекомендовать эксперту решить, не наступила ли пора выделить из “всего остального” и включить в число существенных еще один элемент. Пройдя таким образом всю предысторию неэлементарного фрагмента, мы получаем новые основания для его декомпозиции, а значит, и возможность продолжить анализ, надеясь достичь элементарности по всем ветвям.

Итак, указанная *итеративность алгоритма декомпозиции* придает ему вариабельность, возможность пользоваться моделями различной детальности на разных ветвях, углублять детализацию сколько угодно (если это потребуется).

ТИПЫ СЛОЖНОСТИ

Несмотря на возможности, предоставляемые сменой моделей и итеративностью, может наступить момент, когда эксперт признает, что его компетентности недостаточно для дальнейшего анализа данного фрагмента и что следует обратиться к эксперту другой квалификации (например, построение новых содергательных моделей требует знаний по иной специальности). По существу, сложность такого типа есть *сложность из-за неинформированности* (“невежества”), которую можно преодолеть с помощью информации, рассредоточенной по разным экспертам и источникам. Однако случай, когда декомпозиция заканчивается элементарными фрагментами на всех ветвях дерева, является простейшим. Не имеет значения, один или несколько экспертов довели анализ до конца, а важно, что это оказалось возможным, и, следовательно, первоначальная сложность была вызвана не столько недостатком информации, сколько большой размерностью проблемы. В действительно сложных случаях получение вполне завершенной декомпозиции должно не только радовать, но и настораживать: не связана ли реальная сложность с пропущенной ветвью дерева, соченной экспертами несущественной? Опасность неполноты анализа следует иметь в виду всегда (примеры последних

IGNORANCE *не-
знание, невежес-
тво*

**WASTE PROD-
UCTS** *отходы*

**CONTENTFUL
MODEL** *содер-
жательная модель*

**RELEVANT, ES-
SENTIAL** *суще-
ственный*

FRAME *формаль-
ная модель*

**Алгоритм деком-
позиции опреде-
ляет последова-
тельность дей-
ствий над объек-
том анализа. В
результате полу-
чается древовид-
ная иерархия все
более простых
составных частей
анализируемого
объекта, пока
анализ не будет**

лет – проблема поворота северных рек, проблемы Байкала и Ладожского озера и т.д.). Один из приемов (не дающий полной гарантии, но иногда полезный) – предлагать экспертам выявлять не только доводы в пользу рассматриваемого проекта, но и сопровождать его обязательным указанием возможных отрицательных последствий. В частности, в классификатор выходов (конечных продуктов) любой системы помимо полезных продуктов обязательно должны быть включены *отходы*.

Невозможность доведения декомпозиции до получения элементарного фрагмента, которая либо эвристически констатируется экспертом на ранних стадиях анализа, либо обнаруживается в виде “затягивания” анализа по данной ветви, является не отрицательным, а также положительным результатом. Хотя при этом сложность не ликвидируется полностью, но ее сфера сужается, обнаруживается и локализуется истинная причина этой сложности. Знание о том, что именно мы не знаем, быть может, не менее важно, чем само позитивное знание.

Правда, вокруг таких результатов часто возникает атмосфера неприятия. Даже физики, говоря “отрицательный результат – тоже результат”, чаще желают просто утешить коллегу-неудачника, а сам отрицательный результат стараются обходить стороной. Так было в начале века с “ультрафиолетовой катастрофой” до возникновения квантовой механики, похожая ситуация сейчас сложилась с объяснением природы шаровой молнии. Однако если в науке *сложность из-за непонимания* расценивается как временно неустранимое и терпимое явление, то в управлении (т.е. в деловых, административных, политических вопросах) она часто воспринимается как неприемлемый вариант, ведущий к недопустимой отсрочке решения. Не потому ли именно в управлении нередко прибе-

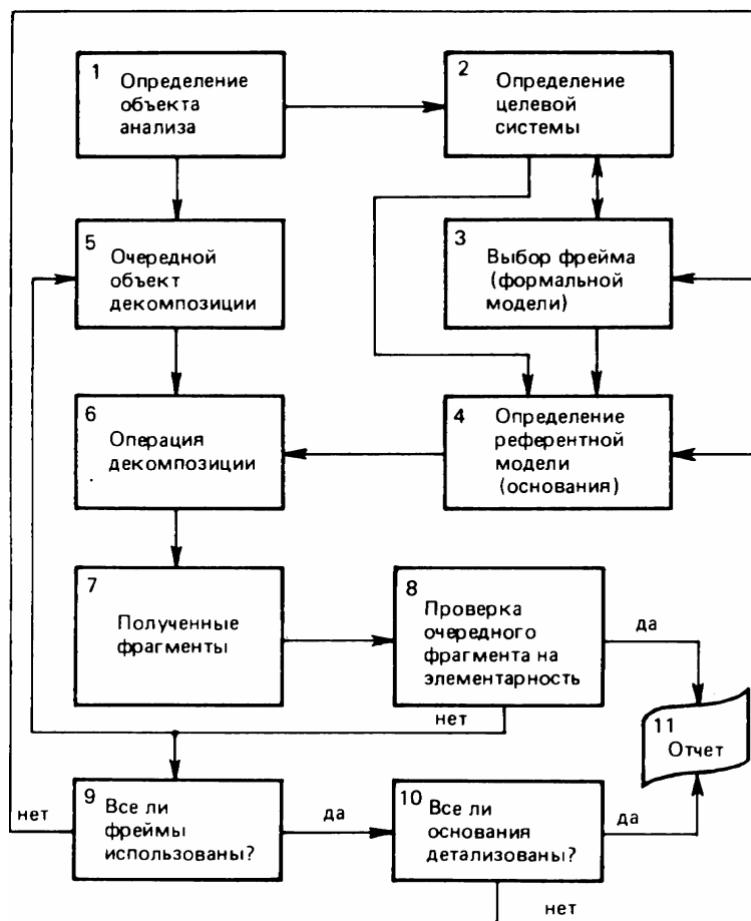
гают к интуитивным и волевым решениям? И не из-за отрицательного ли (в целом) опыта таких решений в последнее время наблюдается быстрое сближение образа мышления управленцев и ученых, повышение роли научных методов в управлении?

Итак, если рассматривать анализ как способ преодоления сложности, то полное сведение сложного к простому возможно лишь в случае сложности из-за неинформированности; в случае сложности из-за непонимания анализ не ликвидирует сложность, но локализует ее, позволяет определить, каких именно сведений нам не хватает. Поэтому (с некоторой натяжкой) можно сказать, что метод декомпозиции не дает новых знаний, а лишь “вытягивает” знания из экспертов, структурирует и организует их, обнажая возможную нехватку знаний в виде “дыр” в этой структуре. Дело в том, что в действительности не только обнаружение нехватки конкретных знаний все-таки является новым знанием (раньше нам было неизвестно, что именно мы не знали), но и по-иному скомбинированные фрагменты старых знаний также обладают новыми качествами.

АЛГОРИТМ ДЕКОМПОЗИЦИИ

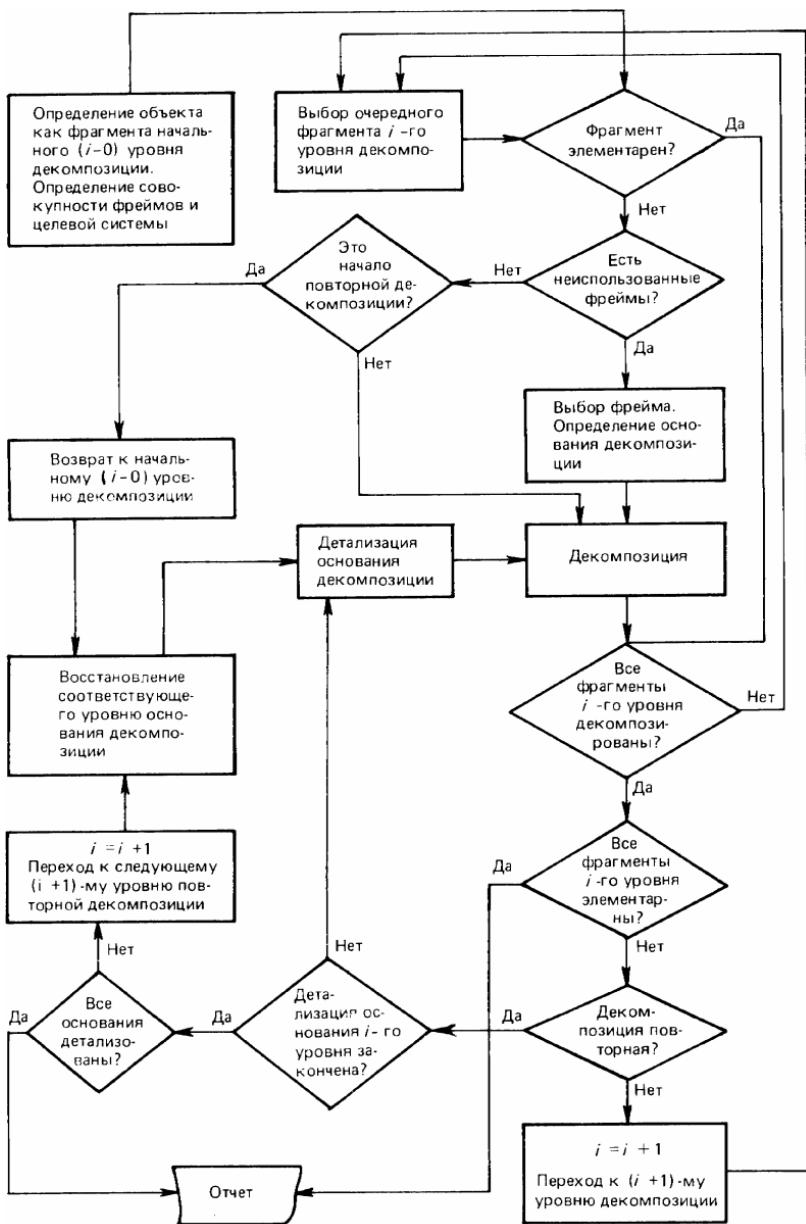
Сам алгоритм декомпозиции, описанный в данном параграфе, представлен в виде блок-схемы (рис. 8.6). К тому, что было уже сказано об изображаемых блоками операциях алгоритма, добавим следующее.

Блок 1. Объектом анализа может стать все, что угодно, – любое высказывание, раскрытие смысла которого требует его структурирования. На определение объекта анализа иногда затрачиваются весьма значительные усилия. Когда речь идет о действительно сложной проблеме, ее сложность проявляется и в том, что сразу трудно правильно сформулировать объект анализа. Даже в таком основательно регламентированном документами случае, как работа министерства, формулировка гло-



8.6 ————— Укрупненная блок-схема алгоритма декомпозиции

8.7 —————> Развёрнутая блок-схема алгоритма декомпозиции



бальной цели возглавляемой им отрасли требует неоднократного уточнения и согласования, прежде чем она станет объектом

анализа (более подробно о сложностях определения цели мы будем говорить в следующей главе). Это относится не только к формулировкам цели, но и к определению любого высказывания, подлежащего анализу. От правильности выбора объекта анализа зависит, действительно ли мы будем делать то, что нужно.

Блок 2. Этот блок определяет, зачем нужно то, что мы будем делать. В качестве целевой системы выступает система, в интересах которой осуществляется весь анализ. Снова подчеркнем, что более формального определения целевой системы дать нельзя, что многое зависит от конкретных условий. Например, опыт построения деревьев целей для Минморфлота [20] и Минвуза показал, что, хотя результаты анализа будут использовать министерство, целевой системой для верхних уровней дерева должна быть отрасль в целом и лишь на нижних уровнях потребовались модели самого министерства.

Блок 3. Этот блок содержит набор фреймовых моделей и рекомендуемые правила их перебора либо обращение к эксперту с просьбой самому определить очередной фрейм.

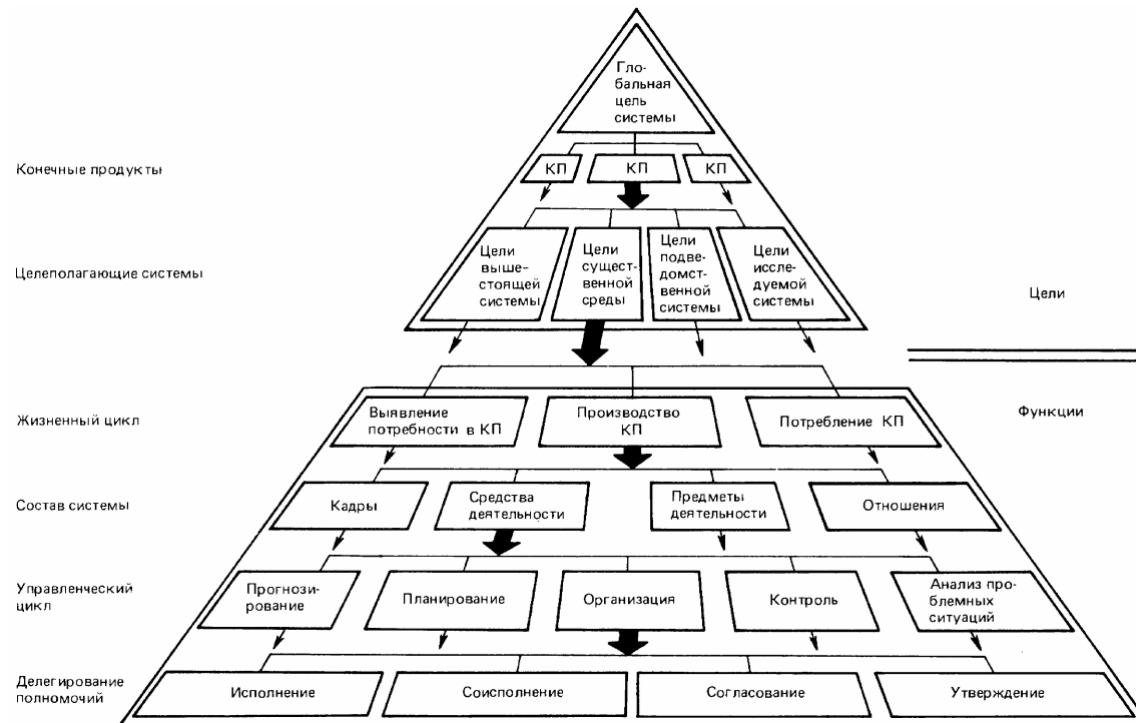
Блок 4. Содержательная модель, по которой будет произведена декомпозиция, строится экспертом на основании изучения целевой системы. Хорошим подспорьем ему могут служить различные классификаторы, построенные в различных областях знаний [12; 13], а также собранные в справочниках и специальных энциклопедиях.

Блоки 5 – 10 были достаточно пояснены ранее.

Блок 11. Окончательный результат анализа оформляется в виде дерева, конечными фрагментами ветвей которого являются либо элементарные фрагменты, либо фрагменты, признанные экспертом сложными, но не поддающимися дальнейшему разложению. Причины такой сложности могут состоять либо в ограниченности знаний данного эксперта или данной группы экспертов (сложность из-за неинформированности), либо в том, что нужные знания существуют, но еще не объединены в объясняющие модели (сложность из-за непонимания), либо в принципиальном отсутствии нужных знаний (сложность из-за незнания).

Блок-схема, изображенная на рис. 8.6, является, конечно, слишком укрупненной; она предназначена для разъяснения лишь основных идей алгоритма декомпозиции. Если потребуется большая конкретизация формальных операций в алгоритме, то можно обратиться к блок-схеме на рис. 8.7.

Для иллюстрации использования алгоритма декомпозиции приведем пример (рис. 8.8) формальной структуры дерева целей и функций



8.8 ————— Фрагмент (одна ветвь) дерева, получающегося при декомпозиции глобальной цели системы

организационной системы, которая была взята за основу декомпозиции глобальной цели Минвуза РСФСР (работа выполнена в Томском политехническом институте под руководством проф. В.З. Ямпольского). Полученное дерево дало возможность внести улучшающие корректиры в структуру министерства, провести аттестацию рабочих мест и составить полные служебные инструкции для них.

Подведем итог

Один из способов упрощения сложного – метод декомпозиции – состоит в разложении сложного целого на все более мелкие (и простые) части. Компромиссы между требованием не упустить важного (принцип полноты) и требованием не включать в модель лишнего (принцип простоты) достигаются с помощью понятий существенного (необходимого), элементарного (достаточного), а также постепенной нарастающей детализации базовых моделей и итеративности алгоритма декомпозиции (см. рис. 8.5).

One method of simplifying complexity – the decomposition method – consists of dividing a whole into lesser (and simpler) parts. Between the requirement not to miss any important aspect (the completeness principle) and the requirement not to include in the model anything irrelevant (the simplicity principle), some compromises are achieved through (1) the concept of relevant (necessary) and elementary fragments, (2) a stepwise detailing of basis models, and (3) making the decomposition algorithm iterative in nature (see Fig. 8.5).

§ 8.4. АГРЕГИРОВАНИЕ, ЭМЕРДЖЕНТНОСТЬ, ВНУТРЕННЯЯ ЦЕЛОСТНОСТЬ СИСТЕМ
Операцией, противоположной декомпозиции, является операция агрегирования, т.е. объединения нескольких элементов в единое целое. Необходимость агрегирования может вызываться различными целями и сопровождаться разными обстоятельствами, что приводит к различным (иногда принципиально различным) способам агрегирования (см. § 8.5). Однако у всех агрегатов (так мы будем называть результат агрегирования) есть одно общее свойство, получившее название *эмержентность*. Это свойство присуще всем системам, и ввиду его важности остановимся на нем подробнее.

ЭМЕРДЖЕНТНОСТЬ КАК ПРОЯВЛЕНИЕ ВНУТРЕННЕЙ ЦЕЛОСТНОСТИ СИСТЕМЫ

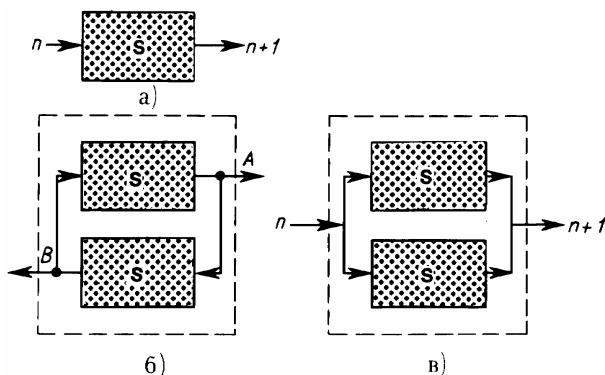
Будучи объединенными, взаимодействующие элементы образуют систему, которая обладает не только внешней целостностью, обособленностью от окружающей среды, но и внутренней целостностью, природным единством. Если внешняя целостность отображается моделью “черного ящика” (см. § 3.3), то внутренняя целостность связана со структурой системы (см. § 3.5). Наиболее яркое проявление внутренней целостности системы состоит в том, что свойства системы не являются только суммой свойств ее составных частей. Система есть нечто большее, система в целом обладает такими свойствами, которых нет ни у одной из ее частей, взятой в отдельности. Модель структуры подчеркивает главным образом связанность элементов, их взаимодействие. Мы же стремимся сейчас сделать акцент на том, что *при объединении частей в целое возникает нечто качественно новое, такое, чего не было и не могло быть без этого объединения*.

Красивый пример проявления этого свойства привел М. Арбид [1]. Пусть имеется некоторый цифровой автомат S , преобразующий любое целое число на его входе в число, на единицу большее входного (рис. 8.9, *a*). Если соединить два таких автомата последовательно в кольцо (рис. 8.9, *б*), то в полученной системе обнаружится новое свойство: она генерирует возрастающие последовательности на выходах A и B , причем одна из этих последовательностей состоит только из четных, другая – только из нечетных чисел.

Summary

ЭМЕРДЖЕНТНОСТЬ КАК РЕЗУЛЬТАТ АГРЕГИРОВАНИЯ

Такое “внезапное” появление новых качеств у систем и дало основание присвоить этому их свойству название эмерджентности. Английский термин emergence означает возникновение из ничего, внезапное появление, неожиданную случайность. В специальной литературе на русском языке не делалось попыток найти эквивалентный русский термин. Однако сам термин имеет обманчивый смысл. Какие бы удивительные свойства ни возникали при объединении элементов в систему, ничего мистического, взявшегося “ниоткуда”, здесь нет: новые свойства возникают благодаря конкретным связям между конкретными элементами. Другие связи дадут другие свойства, не обязательно столь же очевидные. Например, парал-



8.9 Иллюстрация внутренней целостности систем (свойства эмерджентности)

лельное соединение тех же автоматов (рис. 8.9, в) ничего не изменяет в арифметическом отношении, но увеличивает надежность вычислений, если на выход поступает сигнал только от исправного автомата.

Свойство эмерджентности признано и официально: при государственной экспертизе изобретений патентоспособным признается и новое, ранее не известное соединение хорошо известных элементов, если при этом возникают новые полезные свойства.

Возникновение качественно новых свойств при агрегировании элементов есть частное, но яркое проявление всеобщего закона диалектического материализма – закона перехода количества в качество. Чем больше отличаются свойства совокупности от суммы свойств элементов, тем выше организованность системы. Так, физик А. Эддингтон писал:

“Нередко думают, что, изучив один какой-то объект, знают уже все о двух точно таких же объектах, так как “два” – это “один и один”. При этом, однако, забывают, что необходимо исследовать еще и то, что скрывается за этим “и”. Изучением этого “и”, т.е. рассмотрением организации, занимается, можно сказать, вторичная физика” [23].

Кибернетик У. Эшби показал, что у системы *там больше возможностей в выборе поведения, чем сильнее степень согласованности поведения ее частей* [21, § 21.10]. Это утверждение легко доказывается на примере системы, состоящей из p частей, каждая из которых может находиться в любом из s состояний (p и s конечны).

Будем исходить из того, что система полностью согласована, если возможен неслучайный переход между любыми двумя ее состояниями. Считая, что каждая из p частей полностью согласована, получим, что число возможных переходов (т.е. число возможностей в выборе поведения) для каждой из частей равно s^s . Объединение p частей в одну систему приводит к тому, что число k возможных состояний становится равным s^p . Однако теперь возможность перейти от одного произвольного состояния к другому неслучайным образом зависит от того, насколько согласованы между собой части системы. Рассмотрим два крайних случая. При полном согласовании частей число возможных переходов равно k^k . Если же система состоит из p несвязанных частей, то число возможных переходов есть $(s^s)^p = (s^p)^s = k^s$. Так как $s < k$, то $k^s < k^k$, что и доказывает приведенное утверждение.

Итак, агрегирование частей в единое целое приводит к появлению новых качеств, не связанных с качествами частей в отдельности. Это свойство и является проявлением внутренней целостности систем, как еще говорят, *системообразующим фактором*. Новые качества систем определяются в очень сильной степени характером связей между частями и могут варьироваться в весьма широком диапазоне – от полного согласования до полной независимости частей.

Подведем итог

Summary

Главная особенность систем The main feature of systems is состоит в том, что свойства that the properties of the whole целого не сводятся к простой cannot be reduced to a simple совокупности свойств его ча- sum of the properties of its parts, стей, что система в целом об- that a system possesses some ладает принципиально новым fundamentally new quality. This качеством. Это новое качество new quality exists as long as the существует, пока существует whole exists and is therefore целое, являясь, таким образом, manifestation of the inner integri- проявлением внутренней це- ty of a system. The phenomenon лостности системы, или, как of the appearance of a new qual- еще говорят, ее “системообра- ity is called emergence. зующим фактором”. Само яв- ление возникновения нового качества называют эмерджент-ностью.

§ 8.5. ВИДЫ АГРЕГИРОВАНИЯ

Как и в случае декомпозиции (см. § 8.2), техника агрегирования основана на использовании определенных моделей исследуемой или проектируемой системы. Именно избранные нами модели жестко определяют, какие части должны войти в состав системы (модель состава; см. § 3.4) и как они должны быть связаны между собой (модель структуры, см. § 3.5). Разные условия и цели агрегирования приводят к необходимости использовать разные модели, что в свою очередь определяет как тип окончательного агрегата, так и технику его построения.

В самом общем виде *агрегирование можно определить как установление отношений на заданном множестве элементов*. Благодаря значительной свободе выбора в том, что именно рассматривается в качестве элемента, как образуется множество элементов и какие отношения устанавливаются (т.е. выявляются или навязываются) на этом множестве, получается весьма обширное количественно и разнообразное качественно множество задач агрегирования. Отметим здесь лишь основные агрегаты, типичные для системного анализа: конфигуратор, агрегаты-операторы и агрегаты-структуры.

КОНФИГУРАТОР

Всякое действительно сложное явление требует разностороннего, многопланового описания, рассмотрения с различных точек зрения. Только совместное (агрегированное) описание в терминах нескольких качественно различающихся языков позволяет охарактеризовать явление с достаточной полнотой. Например, автомобильная катастрофа должна рассматриваться не только как физическое явление, вызванное механическими причинами (техническим состоянием автомобиля и дорожного покрытия, силами инерции, трения, ударов и т.д.), но и как явление медицинского, социального, экономического, юридического характера. Даже движение планет имеет не только механические аспекты, но и социальные, – вспомним, какие потрясения вызвал переход от геоцентрического к гелиоцентрическому описанию этого движения. В реальной жизни не бывает проблем чисто физических, химических, экономических, общественных или даже системных – эти термины обозначают не саму проблему, а выбранную точку зрения на нее. По образному выражению писателя-фантаста П. Андерсона,

AGGREGATE аг-
регат
CONFIGURATOR
конфигуратор
SELF-
ORGANIZATION
самоорганизация
STRUCTURE
структура
EMERGENCE
эмерджентность

Эмерджентность
яв-ляется важ-
нейшей особенно-
стью систем и со-
стоит в том, что
свойства системы
не сводятся к со-
вокупности
свойств ее частей,
что целое облада-
ет качественно
новыми свой-
ствами. Свойство
эмерджентности

проблема, сколь бы сложной она ни была, станет еще сложнее, если на нее правильно посмотреть.

Эта многоплановость реальной жизни имеет важные последствия для системного анализа. С одной стороны, системный анализ имеет междисциплинарный характер. Системный аналитик готов вовлечь в исследование системы данные из любой отрасли знаний, привлечь эксперта любой специальности, если этого потребуют интересы дела; с другой стороны, перед ним встает неизбежный вопрос о допустимой минимизации описания явления. Однако если при декомпозиции этот вопрос решался компромиссно с помощью понятия существенности, что давало некоторую свободу выбора, сопровождающую риском недостаточной полноты или излишней подробности (см. § 8.2), то при агрегировании этот вопрос обостряется: риск неполноты становится почти недопустимым, поскольку при неполноте речь может идти вообще не о том, что мы имеем в виду; напротив, риск переопределения связан с большими излишними затратами.

Приведенные соображения приводят к понятию *агрегата*, состоящего из качественно различных языков описания системы и обладающего тем свойством, что число этих языков минимально, но необходимо для заданной цели. Следуя В.А. Лефевру [5], будем называть такой агрегат **конфигуратором**. Продемонстрируем смысл этого понятия на примерах.

Пример 1. Конфигуратором для задания любой точки n -мерного пространства является совокупность ее координат. Обратим внимание на эквивалентность разных систем координат (разных конфигураторов) и на предпочтительность ортогональных систем, дающих независимое описание на каждом “языке” конфигуратора.

Пример 2. Конфигуратором пространственной перспективы (т.е. обнаружения различий в направлениях и удаленностих источников сигналов от наблюдателя) является фиксация каждого сигнала в двух “разнесенных” точках: бинауральный эффект в акустике, бинокулярное зрение (стереоскопический эффект). Отметим приспособленность животных и человека к жизни на “плоской” поверхности Земли, что делает достаточным замеры в двух точках. Не исключено, что малые глазки стрекоз, расположенные на ее “лбу” в вершинах равностороннего треугольника, обеспечивают “наведение” стрекозы-хищницы на “цель”, движущуюся по трехмерной траектории. Летучим мышам, ориентирующимся с помощью ультразвуковой локации, это не потребовалось благодаря возможности поворота головы, чего лишена стрекоза.

Пример 3. Конфигуратором для описания поверхности любого трехмерного тела на “плоскостных” языках является совокупность трех ортогональных проекций, принятая в техническом черчении и начертательной геометрии. Обратим внимание на невозможность уменьшения числа проекций и на избыточность большего числа “точек зрения”.

Пример 4 [7]. В радиотехнике для одного и того же прибора используется конфигуратор: блок-схема, принципиальная схема, монтажная схема. *Блок-схема* может определяться теми технологическими единицами, которые выпускаются промышленностью, и тогда прибор членится на такие единицы. *Принципиальная схема* предполагает совершенно иное расчленение: она должна объяснить функционирование этого прибора. На ней выделены функциональные единицы, которые могут не иметь пространственно локализованных аналогов. Приборы могут иметь различные блок-схемы и одинаковые принципиальные схемы, и наоборот. Наконец, *монтажная схема* является результатом расчленения прибора в зависимости от геометрии объема, в пределах которого производится его монтаж.

Подчеркнем (в отличие от [7]), что главное в конфигураторе не то, что анализ объекта должен производиться на каждом языке конфигуратора отдельно (это разумеется само собой), а то, что синтез, проектирование, производство и эксплуатация прибора возможны только при наличии всех трех его описаний. Однако этот пример дает возможность подчеркнуть еще и зависимость конфигуратора от поставленных целей. Например, если кроме целей производства мы будем преследовать и цели сбыта, то в конфигуратор радиоаппаратуры придется включить и язык рекламы, позволяющий описать внешний вид и другие потребительские качества прибора.

Пример 5. Когда обсуждаются кандидатуры на руководящую должность, каждый претендент рассматривается с учетом его профессиональных, деловых, идеально-политических, моральных качеств и состояния здоровья. Попробуйте в порядке упражнения обсудить такую структуру характеристики человека как конфигуратор.

Пример 6. При описании процессов, происходящих в народнохозяйственных комплексах областного масштаба, было признано необходимым [11; 13] для характеристики любого выходного продукта производственной или обслуживающей сферы использовать три типа показателей: натуральные (экономико-технологические), денежные (финансово-экономические) и социально-ценностные (идеологические, политические, этические и эстетические). Деятельность завода и театра, совхоза и школы, любого предприятия и организации описывается на этих трех языках, образующих конфигуратор по отношению к целям автоматизированной системы управления хозяйством области.

Пример 7. Опыт проектирования организационных систем [12] показывает, что для синтеза оргсистемы конфигуратор состоит из описания распределения власти (структуры подчиненности), распределения ответственности (структур функционирования) и распределения информации (организации связи и памяти системы, накопления опыта, обучения, истории). Все три структуры не обязаны совпадать топологически, хотя связывают одни и те же части системы.

Пример 8. В § 4.3 и 4.4 приведены классификации систем по следующим основаниям: описание происхождения – описание переменных – описание оператора – описание управления – описание ресурсов управления. В порядке упражнения обсудите аргументы за и против того, что эта совокупность может служить конфигуратором.

Заметим, что конфигуратор является содергательной моделью высшего возможного уровня. Перечислив языки, на которых мы будем говорить о системе, мы тем самым определяем, синтезируем тип системы, фиксируем наше понимание природы системы. Как всякая модель, конфигуратор имеет целевой характер (см. § 2.2) и при смене цели может утратить свойства конфигуратора (см. пример 4). Как каждая модель, конфигуратор в простых случаях очевидно адекватен (см. примеры 1 и 2), адекватность других подтверждает практика (см. примеры 3 – 6), в полноте третьих мы лишь более или менее уверены и готовы пойти на их дополнение новыми языками.

АГРЕГАТЫ-ОПЕРАТОРЫ

Одна из наиболее частых ситуаций, требующих агрегирования, состоит в том, что совокупность данных, с которыми приходится иметь дело, слишком многочисленна, плохо обозрима, с этими данными трудно “работать”. Именно интересы работы с многочисленной совокупностью данных приводят к необходимости агрегирования. В данном случае на первый план выступает такая особенность агрегирования, как уменьшение размерности: агрегат объединяет части в нечто целое, единое, отдельное.

Простейший способ агрегирования состоит в установлении отношения эквивалентности между агрегируемыми элементами, т.е. образования классов. Это позволяет говорить не только о классе в целом, но и о каждом его элементе в отдельности.

Можно рассматривать различные задачи, связанные с классификацией и ее использованием. Приведем примеры таких задач: классификация как самая простая, самая первая, “атомарная” форма моделирования (это проблематика философии и психологии); элемент как

представитель класса (совсем не простой вопрос; так, в теории случайных процессов это приводит к проблеме эргодичности); искусственная классификация и природная кластеризация (т.е. образование классов “для удобства” и естественная общность определенных явлений, возможные соотношения между ними; например, созвездия – первое или второе или отчасти то и другое?); иерархическая и фасетная (сетевая) классификация и т.д. Классификация является очень важным и многофункциональным, многосторонним явлением в человеческой практике вообще и в системном анализе в частности. С практической точки зрения одной из важнейших является проблема определения, к какому классу относится данный конкретный элемент. Обсудим ее подробнее.

КЛАССИФИКАЦИЯ КАК АГРЕГИРОВАНИЕ

Если признак принадлежности к классу является непосредственно наблюдаемым, то кажется, что особых трудностей классификации нет. Однако и в этих случаях на практике возникает вопрос о надежности, правильности классификации. Например, разложить окрашенные куски картона по цветам – трудная задача даже для ученых-психологов: отнести ли оранжевый кусок к “красным” или “желтым”, если между ними нет других классов? По одежду военных можно определить не только, к какому государству они принадлежат, но и в каком роде войск служат, в каком чине состоят и т.д. Но эту четкость и однозначность враг может использовать, переодевая своих диверсантов в форму противника, и тогда распознавание “своих” от “чужих” осуществляется с помощью других признаков. Если же непосредственно наблюдаемый признак принадлежности к классу формулируется на естественном языке, то, как известно, некоторая неопределенность (расплывчатость) становится неизбежной (например, кого отнести в класс “высоких людей”? – см. § 6.3 и 7.8).

Сложности классификации резко возрастают, если признак классификации не наблюдается непосредственно, а сам является агрегатом косвенных признаков. Типичным примером является диагностика заболевания по результатам анамнеза: диагноз болезни (ее название есть имя класса) представляет собой агрегат большой совокупности ее симптомов и характеристик состояния организма. Если классификация имеет природный характер, то агрегирование косвенных признаков может рассматриваться как обнаружение закономерностей в таблицах экспериментальных данных [4; 6], т.е. как поиск устойчивых, достаточно часто повторяющихся в обучающей выборке “цеплений” признаков (см. § 6.4 и 7.7). При этом приходится перебирать все возможные комбинации признаков с целью проверки их повторяемости в обучающей выборке (например, в таблице, строки которой соответствуют данным предварительных опроса и анализов для каждого пациента). Вообще, метод перебора вариантов – самый очевидный, простой и надежный способ поиска решения. Несмотря на трудоемкость, его нередко с успехом применяют. Т. Эдисон утверждал, что перебор – его основной метод изобретательской деятельности (хотя, скорее всего, это была шутка). Метод Ф. Цвикки морфологического анализа систем [8], машинное обнаружение закономерностей [5; 6; 76], дискретные задачи оптимизации – уже не шуточные, а типичные примеры использования перебора. Однако уже при совсем небольшом количестве признаков полный перебор становится нереальным даже при использовании ЭВМ. Успех в значительной степени зависит от того, удастся ли найти метод сокращения перебора, приводящий к “хорошим” решениям, и разработке таких методов посвящено значительное количество исследований.

Все сказанное свидетельствует о том, что агрегирование в классы является эффективной, но далеко не тривиальной процедурой. Если представлять класс как результат действия агрегата-оператора, то такой оператор имеет вид “ЕСЛИ <условия на агрегируемые признаки>, ТО <имя класса>“. Как было отмечено, иногда класс непосредственно задается совокупностью признаков, а в ряде случаев, наоборот, требуется доопределить оператор, выявив экспериментально, при каких условиях на признаки объект будет принадлежать заданному классу.

ФУНКЦИЯ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ КАК АГРЕГАТ

Другой тип агрегата-оператора возникает, если агрегируемые признаки фиксируются в числовых шкалах. Тогда появляется возможность задать *отношение на множестве признаков в виде числовой функции многих переменных, которая и является агрегатом*.

Свобода выбора в задании функции, агрегирующей переменные, является кажущейся, если этой функции придается какой-то реальный смысл. В этом отношении характерен случай перехода от многокритериальной оптимизационной задачи к однокритериальной с помощью агрегирования нескольких критериев в один суперкритерий. Построение суперкритериальной функции, по существу, является построением модели системы. Не зная “истинной” упорядочивающей функции, мы можем аппроксимировать ее гиперплоскостью (т.е. линейной комбинацией частных критериев), но должны стремиться к тому, чтобы эта гиперплоскость была “достаточно близка” к неизвестной суперповерхности, чтобы сравниваемые альтернативы находились “вблизи” точки касания суперплоскости с суперповерхностью [17]. Если обеспечить это мы не в состоянии, то можно использовать кусочно-линейные и другие нелинейные аппроксимации, т.е. другие агрегаты критериев, либо вообще отказаться от их агрегирования в один критерий (см. § 7.2). Отметим, что паретовская оптимизация в каком-то смысле аналогична отказу от агрегата-оператора и возврату к агрегату-конфигуратору.

Интересно подчеркнуть, что в тех (к сожалению, редких) случаях, когда агрегат-оператор является вполне адекватной моделью системы, мы вообще лишаемся свободы выбора функции, агрегирующей набор переменных. Именно этот случай имеет место, когда закономерности природы отображаются безразмерными степенными одночленами физических размерных величин [15]. Такое, казалось бы, тривиальное требование, как сохранение отношения двух числовых значений составных физических величин (т.е. зависящих от нескольких других величин) при изменении единиц измерения исходных величин, приводит к нетривиальному выводу: *если удалось построить безразмерный степенной одночлен из размерных физических величин, образующих конфигуратор рассматриваемого явления, то выявлена физическая закономерность данного явления*. Например, из того, что $F^{-1} ma = c$, где c – безразмерная постоянная, F – сила, m – масса, a – ускорение, следует второй закон Ньютона. Конечно, метод размерности может привести к уже известным, а иногда тривиальным закономерностям, но это не является недостатком метода.

Другой редкий пример однозначности агрегата-функции дает широко используемый стоимостный анализ экономических систем. Если все участвующие факторы удается выразить в терминах денежных расходов и доходов, то агрегат оказывается их алгебраической суммой. Вопрос состоит лишь в том, в каких случаях можно использовать этот агрегат, не обращаясь к другим системам ценностей, а когда следует вернуться к конфигуратору, включающему политические, моральные, экологические, а не только финансовые критерии.

Добавим, что числовую функцию можно задавать не только на числовых аргументах, и это позволяет рассматривать еще один вид агрегата-функции.

СТАТИСТИКИ КАК АГРЕГАТЫ

Важный пример агрегирования данных дает статистический анализ. Среди различных агрегатов (называемых в этом случае *статистиками*, т.е. функциями выборочных значений, см. § 7.7) особое место занимают *достаточные статистики*, т.е. такие агрегаты, которые извлекают всю полезную информацию об интересующем нас параметре из совокупности наблюдений. Однако при агрегировании обычно потери информации неизбежны, и достаточные статистики являются в этом отношении исключением. В таких условиях становятся важными *оптимальные статистики*, т.е. позволяющие свести неизбежные в этих условиях потери к минимуму в некотором заданном смысле. Наглядный пример статистического агрегирования представляет собой факторный анализ, в котором несколько переменных сводятся в один фактор. Именно потому, что при рассмотрении реальных данных самым важным является построение модели-агрегата при отсутствии информации, необходимой для теоретического синтеза статистики, Тьюки [18] предложил назвать эту область “анализом данных”,

оставляя за математической статистикой задачи алгоритмического синтеза и теоретического анализа статистик.

Наконец, еще раз подчеркнем, что с созданием агрегата-оператора связан не только выигрыш, ради которого он и создается, но и риск попасть в “ловушки”. Отметим основные из них:

потеря полезной информации. Агрегирование является необратимым преобразованием (например, по сумме нельзя восстановить слагаемые), что в общем случае и приводит к потерям (см. § 5.7); достаточные статистики – лишь счастливое исключение (если сумма есть достаточная статистика, то информация об отдельных слагаемых и не нужна);

агрегирование представляет собой выбор определенной модели системы, причем с этим выбором связаны непростые проблемы адекватности (см. § 2.5 и 2.6);

некоторым агрегатам-операторам присуща внутренняя противоречивость, сопряженная с отрицательными (по отношению к целям агрегирования) последствиями. Наиболее ярким примером этого является теорема о невозможности (см. § 7.5), но не присуще ли это свойство (хотя и выраженное в разной степени) всем агрегатам?

АГРЕГАТЫ-СТРУКТУРЫ

Важной (а на этапе синтеза – важнейшей) формой агрегирования является образование структур. К тому, что о моделях структур уже было сказано в § 3.5 – 3.6, можно добавить следующее.

Как и любой вид агрегата, структура является моделью системы и, следовательно, определяется тройственной совокупностью: объекта, цели и средств (в том числе среды) моделирования (см. § 2.8). Это и объясняет многообразие типов структур (сети, матрицы, деревья и т.д.), возникающих при выявлении, описании структур (познавательные модели; см. § 2.2).

При синтезе мы создаем, определяем, навязываем структуру будущей, проектируемой системе (прагматические модели; см. § 2.2). Если это не абстрактная, а реальная система, то в ней вполне реально (т.е. независимо от нашего желания) возникнут, устаноются и начнут “работать” не только те связи, которые мы спроектировали, но и множество других, не предусмотренных нами, вытекающих из самой природы сведенных в одну систему элементов. Поэтому при проектировании системы важно задать ее структуры во всех существенных отношениях, так как в остальных отношениях структуры сложатся сами, стихийным образом (конечно, не совсем независимо от установленных и поддерживаемых проектных структур). Совокупность всех существенных отношений определяется конфигуратором системы, и отсюда вытекает, что *проект любой системы должен содержать разработку стольких структур, сколько языков включено в ее конфигуратор*. Например, проект организационной системы должен содержать структуры распределения власти, распределения ответственности и распределения информации (см. пример 7). Подчеркнем, что, хотя эти структуры могут весьма сильно отличаться топологически (например, структура подчиненности иерархична, а функционирование организовано по матричной структуре), они лишь с разных сторон описывают одну и ту же систему и, следовательно, не могут быть не связанными между собой.

В современных системных науках все возрастающее внимание уделяется одному из специфических видов структур – так называемым *семантическим сетям*. Начало их исследованию и применению положил в 60-х годах Д.А. Поспелов в развиваемом им и поныне ситуационном управлении [14]. В настоящее время исследование таких сетей с разных позиций ведется во многих научных коллективах, поскольку *логико-лингвистические модели* (иное название семантических сетей) оказались в центре всех событий, происходящих в искусственном интеллекте и его приложениях. Такое положение вызвано тем, что указанные модели отображают структуру человеческих знаний, выражаемых на естественном языке, причем это отображение может быть осуществлено средствами ЭВМ.

Хотя можно перечислить, казалось бы, все мыслимые структуры как частные случаи полного графа, некоторые явления природы наводят на мысль, что и в этом вопросе не следует спешить с окончательными выводами. Отдельные особенности живых организмов, экономи-

ческих и социальных систем заставляют предположить, что даже сложнейшие существующие модели структурной организации в чем-то слишком просты.

Очевидным примером нерешенной задачи организации системы является работа человеческого мозга, хотя точно известно, что он состоит из 10^{10} нейронов, каждый из которых имеет $10^2 - 10^3$ нервных окончаний и может находиться лишь в одном из двух состояний.

Другие интересные примеры, показывающие, что существуют реальные системы, структуру которых не охватывает современная теория управления, приводил в одном из своих выступлений акад. Дородницын.

Так, долгое время считалось что у тропических ящериц-гекконов способность бегать по стенам и потолку обеспечивается микроскопическими присосками на лапках. Когда же обнаружилось, что геккон не может бегать по полированному стеклу, выяснилось, что никаких присосок нет, а есть многие тысячи мельчайших волосков, которые со всех сторон “обжимают”, а в нужный момент “отпускают” малейшие шероховатости поверхности стены или потолка. В процессе погони за мухой каждому волоску-щупу нужно отдать верную команду в подходящий момент! Ясно, что централизованная система с этим не справится, но как тогда синхронизируется нецентрализованное управление?

Не менее поразителен и другой пример. Многие головоногие способны менять свою окраску, отождествляя ее с фоном. Так, если только что пойманного осьминога положить на газету, то на его теле выступят полоски, имитирующие строчки, а в крупных заголовках можно даже различить некоторые буквы. Оказалось, что окраска осьминога зависит от того, насколько растянуты микроскопические упругие шарики, распределенные по всему его телу. “Цвет” шарика определяется той длиной волны, до размеров которой он растянут. Но ведь каждый шарик должен “знать”, до какой длины ему растянуться, чтобы создать свой фрагмент в общей мозаике! Как осуществляется вся эта координация?

Может быть, существуют еще не известные нам принципы самоорганизации? Может быть, имеется качественная, а не количественная разница между объединениями большого числа составляющих с малым и большим числом связей для каждой из них? Подобные вопросы пока остаются без ответа.

Подведем итог

Summary

Существует много форм агрегирования, т.е. соединения частей в целое. Их общность состоит в том, что: агрегирование диктуется выбранной моделью описываемой системы; агрегирование есть установление отношений между агрегируемыми элементами. Наиболее важными являются следующие виды агрегаторов: а) конфигуратор (совокупность языков описания системы); б) оператор (конкретизация отношения, в частности классификация, упорядочение, числовые функции, поиск закономерностей и пр.); в) структура (описания связей на всех языках конфигуратора, в частности

The are many forms of aggregation, or combining of parts into a whole. The common features of all of them are (1) any aggregation stems from a given model of the studied system, and (2) any aggregation is the establishment of relationships between aggregated elements. Perhaps the most important types of aggregates are the following: (i) the configurator (i.e., the set of languages for the description of the system); (ii) the operator (i.e., a definite relation – classification, ordering, numerical function, established dependence, etc.); (iii) the structure (i.e., a description of the connections in all the languages of the configurator, the definition

задание семантических сетей). of the semantic network, for example). There are also other less-studied types of aggregations, such as self-organization.

Отметим также существование малоизученных способов агрегирования (например, самоорганизацию).

ЗАКЛЮЧЕНИЕ

В данной главе наиболее важными являются три момента. Первый – в алгоритмизации декомпозиции удалось продвинуться достаточно далеко (см. § 8.2 и 8.3). Важность этого связана с тем, что во многих практических случаях построение соответствующего дерева оказывается ключевым моментом во всем анализе. Например, деревья целей и деревья функций организационных систем используются [13]: для выяснения происхождения возникшей проблемы; для построения количественных моделей изучаемой системы; для решения таких задач управления, как распределение ресурсов и разработка стимулов; для последующего синтеза оргструктур и т.д.

В практике распространения системного подхода в отношении к декомпозиции существуют две крайности: эйфорическая и пессимистическая. Одни вполне удовлетворялись ясностью, которая открывалась при построении соответствующего дерева, и полагали, что весь анализ, по существу, и сводится к такому построению. Другие, особенно специалисты, работающие на руководящих хозяйственных должностях, относились к этому с априорной предубежденностью.

Это, в частности, отчетливо наблюдается в ходе преподавания системного подхода на курсах повышения квалификации руководящих работников. В своей работе руководитель, хочет он того или нет, осознанно или интуитивно, вынужден своевременно выявлять проблемы, формулировать цели, планировать способы их реализации, организовывать и контролировать исполнение планов, т.е. постоянно заниматься системной деятельностью. Достаточно успешно справляясь со своими обязанностями (иначе они и не были бы руководителями), они с явной долей скептицизма и иронии поначалу относятся к попыткам учить их тому, что они “и так знают”. Например, если показать результаты системной декомпозиции некоторой проблемы, знакомой опытному руководителю, то он, как правило, говорит: “Все это правильно, не спорю. Но я и без вашего анализа это знаю. И вообще – чем системный подход отличается от простого здравого смысла?”. Бесполезно говорить ему, что в этом вопросе проявляется недооценка различия между естественным, интуитивным уровнем системности любой деятельности и сознательной системностью более высокого уровня: такая критика, как и любой “переход на личности”, является неэффективным способом убеждения. Вместо этого был придуман такой прием. Группе руководителей производства, повышающих квалификацию, предлагалось, исходя из опыта и здравого смысла, структурировать, разбить некоторую задачу на необходимое и, по их мнению, достаточное число подзадач, решение которых должно обеспечить решение всей задачи. После этого та же задача структурировалась по системной методике (изложенной в § 8.3), и каждый раз оказывалось, что полностью эвристическая декомпозиция давала на 25 – 30 % меньше составляющих, чем формализованная декомпозиция. Специалисты убеждались (и это поражало их больше всего), что дело не в незнании ситуации (недостающие компоненты также оказывались им знакомы), а в правильной организации имеющихся знаний. После этого скептическое отношение рассеивалось. Вообще складывается впечатление, что люди, на деле сталкивающиеся с необходимостью решать сложные проблемы, острее других понимают и активнее осваивают методы системного анализа.

Сделаем еще одно замечание по поводу алгоритмизации декомпозиции. Эта алгоритмизация требуется в связи с необходимостью структурировать “крупные” сложные объекты, анализ которых приводит к деревьям, содержащим многие десятки, а иногда тысячи фрагментов. Если для построения “маленького” дерева и для работы с ним обычно достаточно здравого смысла, то “крупное” дерево требует особой дисциплины в обращении с ним; такую дисциплину и обеспечивает алгоритм.

Остается подчеркнуть, что системный анализ в современном понимании не сводится к декомпозиции проблемы; это важный, но промежуточный этап.

Второй существенный момент – введение в § 8.4 понятия внутренней целостности (эмержентности) системы. Рекомендуем обратить на него внимание, так как это понятие связано с самым важным, самым характерным свойством систем.

Наконец, третий момент – это понятие конфигуратора, т.е. совокупности качественно различающихся точек зрения на проблему, подлежащую разрешению. Здесь главным является предположение о том, что такая совокупность точек зрения, “языков” (в кавычках и без них) исчерпывающим образом позволяет описать систему и проблему. Естественно, речь идет не об абсолютной, безотносительной полноте, а о полноте, связанной с целью анализа. К такой полноте следует стремиться в максимально возможной степени, а для ее обеспечения вводится понятие “проблематика”, которое будет рассмотрено в следующей главе.

ЛИТЕРАТУРА

1. *Арбид М.* Метафорический мозг. – М.: Мир, 1976.
2. *Вазина К.Я., Тарасенко Ф.П.* Эффективный метод решения проблем // Вестник высшей школы. 1987. № 7. С. 7 – 13.
3. *Декарт Р.* Избр. произв.– М., 1950.
4. *Загоруйко Н.Г., Елкина В.Н., Лбов Г.С.* Алгоритмы обнаружения эмпирических закономерностей. – Новосибирск: Наука, 1985.
5. *Закревский А.Д.* Логика распознавания. – Минск: Наука и техника, 1988.
6. *Лбов Г.С.* Методы обработки разнотипных экспериментальных данных. – Новосибирск: Наука, 1981.
7. *Лефевр В.Д.* Конфликтующие структуры. – М.: Сов. радио, 1973.
8. *Одрин М.В., Карташов С.С.* Морфологический анализ систем. – Киев: Наукова думка, 1977.
9. Океан начинается с берега // Известия. 1987. № 51, 203.
10. *Оптнер С.Л.* Системный анализ для решения деловых и промышленных проблем. – М.: Сов. радио, 1969.
11. *Перегудов Ф.И.* и др. Основы системного подхода. – Томск: ТГУ, 1976.
12. *Перегудов Ф.И.* Основы системного проектирования АСУ организационными комплексами. – Томск: ТГУ, 1984.
13. *Перегудов Ф.И.* и др. Системное проектирование АСУ хозяйством области. - М.: Статистика, 1977.
14. *Поспелов Д.А.* Ситуационное управление. Теория и практика. – М.: Наука, 1986.
15. *Седов Л.К.* Теория подобия и размерности в механике. – М.: ГИТТЛ, 1954.
16. *Сироджса И.Б.* и др. Пакет прикладных программ классификационной обработки данных (ППП КОД-2). – В кн.: Матем. методы анализа динамических систем.– Харьков, 1983. Вып. 7. С. 127 – 134.
17. *Тарасенко Ф.П.* О принципиальных трудностях балльной оценки научной деятельности // Вестник АН СССР. 1976. С. 69 – 75.
18. *Тьюки Дж.* Анализ результатов наблюдений. – М.: Мир, 1981.
19. *Федоренко Н.П.* О методах социально-экономического прогнозирования. – В кн.: Методология прогнозирования экономического развития СССР. – М.: Экономика, 1971.
20. *Черняк Ю.И.* Системный анализ в управлении экономикой. – М.: Экономика, 1975.
21. *Эшиби У.Р.* Конструкция мозга. – М.: ИЛ, 1962.
22. *Янг С.* Системное управление организацией. – М.: Сов. радио, 1972.
23. *Eddington A.* The Nature of the Physical World. Univ. of Mich. Press, 1958.
24. *Gharajedaghi J., Ackoff R.L.* Toward Systemic Education of System Scientists. System Research, 1985. Vol. 2. N 1. 21 – 27.
25. Future Methodology Based on Past Assumptions? In: “Rethinking the Process of Operational Research and System Analysis”: Pergamon Press, 1984. P. 1 – 17.

УПРАЖНЕНИЯ

§ 8.1

- Обсудите роль операций декомпозиции и агрегирования в анализе и синтезе.

§ 8.2

- Обсудите различия в декомпозициях процесса системного анализа, приведенных в табл. 8.1.
- Приведите примеры, показывающие, что именно берется в качестве объекта анализа и как именно система порождает модели – основания декомпозиции.

§ 8.3

- Попробуйте декомпозировать любое высказывание (например, “сходить группой в лыжный поход”, “организовать дискотеку”, “провести дискуссию” и т.п.) сначала интуитивно, а затем с использованием алгоритма. Сравните результаты и объясните полученные различия.
- С помощью алгоритма декомпозиции выявите структуру темы вашей курсовой работы.
- Если вы хорошо овладели программированием, то вас, возможно, заинтересует разработка программы, обеспечивающей диалог с экспертом в процессе декомпозиции в соответствии с изложенным алгоритмом. Это могло бы стать темой научного студенческого кружка или дипломной работы. При выполнении этой работы вам будут полезны [13, § 2.4] и рис. 8.6.

§ 8.4

- Приведите примеры эмерджентности.
- Обсудите, как управляемость системы зависит от управляемости каждой из ее частей в отдельности и от структуры систем.

§ 8.5

- Обсудите на примере какой-нибудь конкретной ситуации из реальной жизни ее конфигуратор. Убедитесь, что для разных целей могут понадобиться различные конфигураторы, хотя реальная ситуация остается прежней.
- Рассмотрите несколько примеров агрегатов-операторов. Определите, к какому типу они относятся.
- Попробуйте построить классификацию агрегатов-структур.
- Как интерпретировать тот факт, что структура одной и той же реальной системы оказывается различной в разных ее моделях?

Вопросы для самопроверки

1. Каково главное отличие причинно-следственного описания связи между явлениями от ее описания как отношения “продуцент – продукт” ?
2. Что конкретно имеется в виду, когда говорится, что основанием декомпозиции является содержательная модель целевой системы?
3. Как используются понятия существенности и элементарности в процессе декомпозиции?
4. В чем состоит свойство систем, называемое эмерджентностью?
5. Какая совокупность языков описания называется конфигуратором?
6. Почему классификацию можно рассматривать как агрегирование?
7. Чем ограничивается свобода выбора при задании агрегатов в виде числовых функций нескольких переменных?
8. Какие аспекты системы подчеркиваются при рассмотрении ее структуры как агрегата?

О НЕФОРМАЛИЗУЕМЫХ ЭТАПАХ СИСТЕМНОГО АНАЛИЗА

ГЛАВА ДЕВЯ- ТАЯ

§ 9.1. ЧТО ТАКОЕ СИСТЕМНЫЙ АНАЛИЗ

Системный анализ возник в ответ на требования практики, поставившей нас перед необходимостью изучать и проектировать сложные системы, управлять ими в условиях неполноты информации, ограниченности ресурсов, дефицита времени. До настоящего времени продолжаются споры, можно ли системный анализ считать наукой, искусством или “технологическим ремеслом”. Особенно остро дискутируются приложения системного анализа к проблемам, связанным с “социотехническими”, “социальными” системами, т.е. системами, в которых решающую роль играют люди. При решении таких проблем существенными оказываются не только вопросы построения и использования моделей, не только эвристические поиски решения слабо структурированных не полностью формализуемых задач, но и чисто психологические аспекты человеческих взаимоотношений, что еще более “удаляет” системный анализ от “чистых наук” типа физики и математики.

РАЗНОРОДНЫЕ ЗНАНИЯ И СИСТЕМНЫЙ АНАЛИЗ

Споры о “степени научности” системного анализа вызваны рядом причин. Во-первых, довольно часто недооценивается работа, связанная с формулировкой задач. Многие полагают, что пока не построены формальные модели, “настоящая” работа еще и не начиналась, а выражение “хорошо поставить задачу – значит наполовину ее решить” расценивают как шутку. В системном анализе акцентируется внимание на трудностях формулирования задач, на способах преодоления этих трудностей.

Во-вторых, преодоление сложности, природа которой связана с неполной формализуемостью (см. § 4.4), требует систематического применения неформальных знаний и методов. Это образно описывает Митрофф [30]. Сознательно упрощая классификацию знаний, он разделяет их на два основных типа – формализованные (“академические”) и неформализованные (“житейские”), а также рассматривает всего два уровня развитости (“высокие” и “низкие”) для каждого из этих типов знаний. В результате получается следующая таблица, характеризующая соотношения между двумя типами знаний:

Житейские знания	Академические зна- ния	
	Высокие	Низкие
Высокие	I	II
Низкие	IV	III

Далее, обсуждая особенности работы системных аналитиков, Митрофф пишет: “Иначе говоря, мы не имеем ни глубоких “житейских” знаний изнутри об организации, которую мы изучаем, ни очень хороших формальных теорий, которые при их применении к организации объяснили бы что-нибудь, кроме очевидного. Далее, традиционно научные круги делают упор на ячейке IV. Они ценят формальные теории выше знаний практика, даже если и соглашаются когда-либо, что практики вообще способны обладать чем-то, что называется “знанием”. Предполагается, что практики, конечно, должны предпочитать ячейку II другим ячейкам.

Остается еще одна ячейка – I. Думаю, что именно тут должны быть сосредоточены наши усилия. Именно тут мы должны работать. Я не верю, что мы можем создать нечто близкое к прикладной социальной науке, если оно не будет основано на обоих типах знаний и на уважительном отношении к ним обоим” [30, с. 35 – 36].

Все сказанное подчеркивает, что системный анализ намеренно объединяет теорию и практику, здравый смысл и абстрактную формализацию.

СИСТЕМНЫЙ АНАЛИЗ КАК ПРИКЛАДНАЯ ДИАЛЕКТИКА

В-третьих, современный системный анализ уже не является эмпирическим собранием философских установок, полезных советов и рецептов, снабженных арсеналом вспомогательных математических и технических средств, с привлечением знаний из любых предметных наук, которые имеют отношение к рассматриваемой проблеме. Все это объединено в систему, организованную в соответствии с единой идеей. Такой идеей является диалектика. Эту мысль в разной форме выражали многие советские и зарубежные авторы.

Итак, можно дать следующее определение: *системный анализ есть прикладная диалектика*. Материалистическая диалектика является методом познания, обеспечивающим согласование системности знаний и системности мира на любом уровне абстракции. Системный анализ в его современном понимании реализует диалектический метод при рассмотрении прикладных задач. Сила и эффективность прикладной диалектики наиболее ярко проявляются при анализе действительно сложных задач, что и наблюдается в практике системного анализа.

Подведем итог

При исследовании систем приходится ставить и решать как хорошо формализованные в математических терминах задачи, так и “слабо структурированные” задачи, выражаемые на естественном языке и решаемые эвристическими средствами. Однако более важно то, что главное достижение системного анализа состоит в разработке методов перехода от неформальных задач к формальным, от моделей типа “черного ящика” к моделям типа “белого ящика”. Большая часть этих методов имеет неформализуемый (в математическом смысле) характер, но они достаточно конкретны и пригодны для практического использования и могут называться не только “искусством” или “ремеслом”, но и технологией. Рассмотрению таких этапов системного анализа и посвящена данная глава.

Summary

In the course of studying systems, we are called upon to formulate and solve both well-defined mathematical problems and ill-defined problems that are formulated in terms of natural language and are solved with the help of heuristic methods. However, the main achievements of systems analysis lie in its capacity for transforming poorly defined problems into well-defined ones, of converting black-box models into “transparent” ones (if this is possible). Most systems analysis methods are nonformal (in the mathematical sense of the word), but they are sufficiently concrete and applicable to be called not merely an art or a craft, but a technology. The methods of this level of systems analysis are considered in this chapter.

§ 9.2. ФОРМУЛИРОВАНИЕ ПРОБЛЕМЫ

Постановка формальной задачи, которую надо решать, для традиционных наук – начальный, отправной этап работы. В исследовании же или проектировании сложной системы – это промежуточный результат, которому предшествует длительная кропотливая и сложная работа по структурированию исходной проблемы.

Проиллюстрируем основные особенности такой работы, составляющей первые этапы системного анализа, на примере “социотехнических” систем (название взято в кавычки, поскольку системологическая терминология в русском языке еще не совсем устоялась). Особо-

ANALYST *анализ*
DIALECTICS *диалектика*
STAKEHOLDERS *заинтересованные стороны*
APPLIED *прикладной*
PROBLEM *проблема*

Системный анализ, рассматриваемый как методология, является прикладной диалектикой. Законы диалектики, определяющие признаки правильности познания и практики, в системном анализе доводятся до конкретных практических рекомен-

бенностью социотехнической системы является то, что люди в ней играют важную роль. Имеются и другие названия таких систем (не совсем синонимичные): “организационные” (что означает в основном состоящие из людей), “автоматизированные” (т.е. состоящие из людей и машин), “человеко-машинные” (обычно так называют систему, состоящую из одного человека и одной машины). Типичными примерами социотехнических систем служат организации типа городской медицинской службы, завода, системы транспорта или связи, экологические системы. Участие в них многих людей, интересы которых различны, делает анализ таких систем особенно сложным. Разумеется, системный анализ применим и к менее сложным системам; при этом многие этапы анализа выполняются проще, быстрее, а иногда и вообще могут быть опущены как уже выполненные ранее; кроме того, уменьшается число итераций, возвращений от последующих этапов к предыдущему, что типично для анализа сложных систем. Чем проще анализируемая система, тем ближе реализуемый алгоритм ее анализа к линейному; чем система сложнее, тем больше циклов реализуется при ее анализе, что, кстати, может служить самостоятельным признаком сложности.

ПРЕВРАЩЕНИЕ ПРОБЛЕМЫ В ПРОБЛЕМАТИКУ

Итак, первые шаги в системном анализе связаны с *формулированием проблемы*. Хотя необходимость системного анализа возникает тогда, когда проблема уже не только существует, но и требует решения, когда инициатор системного анализа (“заказчик”, “клиент”) уже сформулировал свою проблему, системный аналитик знает, что первоначальная формулировка – лишь очень приблизительный намек на то, какой именно должна быть действительная рабочая формулировка проблемы. Это относится не только к случаям, когда заказчик лишь обозначает сферу интересов (“Как улучшить работу медицинских учреждений?”) или “Как повысить активность и самостоятельность студентов?”), но и когда он достаточно конкретен (“Какой из предложенных проектов принять к исполнению?” или “Какой должна быть модель следующего поколения данного изделия?”) или даже “совсем точен” (“Где в районе разместить новую больницу?”, “Каковы оптимальные параметры такого-то изделия?”).

Имеется ряд причин считать любую исходную формулировку проблемы лишь “нулевым приближением”. Главная из них состоит в том, что *проблемосодержащая система* (так называют систему, в деятельности которой проявилась данная проблема как некоторое отрицательное, нежелательное явление) не является ни изолированной, ни монолитной: она связана с другими системами и входит как часть в некоторую надсистему; сама она в свою очередь состоит из частей, подсистем, в различной степени причастных к данной проблеме. Если это действительно реальная проблема и мы намерены хотя бы ослабить ее остроту, то необходимо учитывать, как это скажется на тех, кого неизбежно затронут планируемые изменения.

Таким образом, к любой реальной проблеме необходимо априори относиться не как к отдельно взятой, а как к “клубку” взаимосвязанных проблем*. Используя для обозначения этой совокупности термин **проблематика**, можно сказать, что *этап формулирования проблемы состоит в определении проблематики* (техника этой операции будет рассмотрена позже).

Другой важной причиной того, чтобы относиться к первоначальной формулировке проблемы не как к безусловной отправной точке дальнейшего анализа, а как к начальному объ-

* В английской системной терминологии этому понятию соответствует введенный Р. Акоффом термин mess, во французской – problematique. Мы будем употреблять в этом же смысле слово “проблематика”.

екту, который сам подлежит изучению и уточнению, является тот факт, что предлагаемая заказчиком формулировка является его моделью реальной проблемной ситуации. Отсюда следует, что необходимо учитывать не только точку зрения заказчика, позиции других заинтересованных сторон (что, как уже было показано, приводит к необходимости расширения проблемы до проблематики), но и то, что его модель, как и всякая модель (см. гл. 2), неизбежно имеет целевой характер, является приблизительной, упрощенной. Поэтому следует проверять предложенную формулировку на адекватность, что обычно приводит к развитию, дополнению, уточнению первоначального варианта описания проблемы. Развитие описания состоит также и в том, что первоначальная формулировка может быть изложена в терминах не всех языков, включенных в конфигуратор (см. § 8.5).

Итак, системное исследование всякой проблемы начинается с ее расширения до проблематики, т.е. нахождения системы проблем, существенно связанных с исследуемой, без учета которых она не может быть решена. Это расширение происходит как “вширь” благодаря выявлению связей проблемосодержащей системы с над- и подсистемами, так и “вглубь” в результате рассмотрения данной проблемы с точки зрения каждого из языков конфигуратора и (если необходимо) детализации исходной проблемы.

МЕТОДЫ ПОСТРОЕНИЯ ПРОБЛЕМАТИКИ

Приведем некоторые рекомендации, как осуществлять построение проблематики.

Очевидно, что для расширения проблемы потребуется содержательная модель над- и подсистем относительно проблемосодержащей системы. В случае организационной системы формальная схема, приводящая к научной содержательной модели, изображена на рис. 8.2. В англоязычной литературе по анализу социотехнических систем часто используется подобная модель, имеющая название *stakeholders*, что означает “перечень заинтересованных лиц”. В этот перечень рекомендуется включать:

- 1) клиента, т.е. того, кто ставит проблему, заказывает и оплачивает системный анализ;
- 2) лиц, принимающих решения, т.е. тех, от полномочий которых непосредственно зависит решение проблемы;
- 3) участников как активных, т.е. тех, чьи действия потребуются при решении проблемы, так и пассивных – тех, на ком скажутся (положительным или отрицательным образом) последствия решения проблемы;
- 4) самого системного аналитика и его сотрудников, главным образом для того, чтобы предусмотреть возможность минимизации его влияния на остальных заинтересованных лиц, – своеобразная “мера безопасности”, с целесообразностью которой мы неоднократно столкнемся в дальнейшем.

Слово “заинтересованный” следует понимать в широком смысле, поскольку в перечень необходимо включить и тех, кто на самом деле не заинтересован в решении проблемы и будет сопротивляться возможным переменам.

Каждая из “заинтересованных” сторон имеет свое видение проблемы, отношение к ней, ее существование или исчезновение приведет к появлению их собственных проблем. Формулирование проблематики и состоит в описании того, какие изменения и почему хочет внести каждое из заинтересованных лиц. Диалектический метод предписывает рассматривать проблему всесторонне, в том числе и во временном, историческом плане. По сути дела, проблематика – это ответ на вопрос: “Какие существующие обстоятельства и прошлый опыт заставляют именно этих заинтересованных лиц, именно в данной культурной среде, включающей именно эти ценности, воспринимать данное состояние дел как проблему?” Чтобы ответ на этот вопрос был полным, его следует дать на всех языках конфигуратора.

Понятно, что при рассмотрении проблемосодержащих систем другой природы (сугубо технических, чисто биологических, экономических и др.) содержательные модели над- и подсистем окажутся другими, однако методика определения проблематики может оставаться той же.

Какова бы ни была природа рассматриваемой системы, ее проблематика включает спектр проблем: от допускающих формализацию в виде постановки математических оптимизационных задач (хорошо структурированных, формализуемых, формальных; английские термины – hard problems, well-defined problems до проблем “рыхлых”, слабо структурированных, неформализуемых, выражаемых на естественном языке (английские термины – soft problems, ill-defined problems). Естественно, эти проблемы следует рассматривать по-разному, но в практике системного анализа наблюдается тенденция сводить все проблемы к одному типу. Та же практика показывает, что исследовать “рыхлую” проблему как “жесткую” оптимизационную гораздо опаснее, чем наоборот: если во втором случае мы лишь частично отказываемся от некоторой полезной информации, то в первом привносим ложную информацию, вводя себя и других в заблуждение. Различать “жесткие” и “рыхлые” проблемы в ходе анализа – одно из условий хорошего анализа (но не его гарантия!).

Строя проблематику, системный аналитик дает развернутую картину того, **кто** из заинтересованных лиц и в **чем** заинтересован, **ка-кие** изменения и **почему** они хотят внести. При этом собственная его позиция должна быть нейтральной, он должен остерегаться повлиять на мнения обследуемых лиц. Например, в ходе выяснения проблематики системному аналитику следует избегать реплик типа “я согласен (не согласен) с вами”, лучше сказать “я слушаю вас”.

Подведем итог

Самый ответственный этап системного анализа – формулирование проблемной ситуации. Этот этап лишь начинается с постановки проблемы заказчиком. Необходимо выявить всех, кого затронут возможные изменения, и сформулировать вытекающие проблемы на всех языках конфигуратора. Полученное множество проблем, названное проблематикой, и является исходным пунктом для системного анализа.

Summary

The most important stage in systems analysis is representation of the problem situation. The client's statement of the problem is only the beginning. It is then necessary to determine who else will be influenced by the planned changes, and to present the stakeholders' problems that the changes will cause, in all languages of the configurator. The resulting set of problems – called a problematique or mess – is the starting point for further systems analysis.

§ 9.3. ВЫЯВЛЕНИЕ ЦЕЛЕЙ

Как хорошо формализованные, так и слабо структурированные проблемы должны быть приведены к виду, когда они становятся задачами выбора подходящих средств для достижения заданных целей. Поэтому прежде всего необходимо определить цели. На данном этапе системного анализа определяется, что надо сделать для снятия проблемы (в отличие от последующих этапов, определяющих, как это сделать).

Главная трудность выявления цели связана с тем фактом, что цели являются как бы антиподом проблемы. Формулируя проблему, мы говорим в явной форме, что нам не нравится. Сделать это сравнительно просто, поскольку то, чего мы не хотим, существует. Говоря же о целях, мы пытаемся сформулировать, что же мы хотим. Мы как бы указываем направление, в котором следует “уходить” от существующей и не устраивающей нас ситуации. Трудность в том и состоит, что возможных направлений много, а выбрать нужно только одно, действительно правильное, а не кажущееся таким.

Рассмотрим основные трудности выявления целей и методы его выполнения.

ОПАСНОСТЬ ПОДМЕНЫ ЦЕЛЕЙ СРЕДСТВАМИ

Как ни странно, в практике системного анализа первоначально сформулированные цели по мере выполнения анализа часто изменяются или отменяются совсем. Это вызвано тем, что субъект, цели которого должны быть выявлены, обычно сам не может их четко осознать, даже если и дает им четкие формулировки. Действительные цели, как правило, шире, чем объявленные. Например, в одном исследовании проблемы “где лучше разместить новую больницу” выяснилось, что действительная цель – улучшение медицинского обслуживания населения, и среди предложенных альтернатив нашлись более эффективные способы использования ресурсов, нежели строительство новой больницы. Это служит также примером того, как средства могут приниматься за цели. Еще одним примером этого является рассмотрение предложения о слиянии в одной из областей мелких деревообрабатывающих предприятий в одно крупное объединение, с очевидными выгодами технологического и экономического характера. Местные власти, провозглашавшие заинтересованность в повышении эффективности производства, на деле “заблокировали” этот вариант. Оказалось, что при образовании мощного объединения деревообрабатывающие предприятия перейдут из системы местной промышленности в подчинение союзного министерства, что резко снизит как отчисления в местный бюджет, так и процент производимых материалов и изделий, распределяемых в области.

Итак, исследование целей заинтересованных в проблеме лиц должно предусматривать возможность их уточнения, расширения или даже замены. В этом и состоит одна из основных причин итеративности системного анализа.

ВЛИЯНИЕ ЦЕННОСТЕЙ НА ЦЕЛИ

На выбор даже конкретных, частных целей субъекта решающее влияние оказывает его общая идеология, система ценностей, которой он придерживается. Одно из направлений исследования в ходе выявления цели может (в необходимых случаях) состоять в изучении системы ценностей, прежде всего той, которой обладают лица, принимающие решения; однако во многих случаях может оказаться необходимым принять во внимание системы ценностей и остальных заинтересованных сторон.

В научной литературе и в публицистике нередко обращается внимание на определенную противоположность “технократического мышления” и “гуманистического мышления” как разных подходов к формированию целей. Главное их различие образно можно выразить так: “Человек – царь природы” и “Человек – часть природы”. Зарубежные [27] и советские [8] авторы делают более подробные их противопоставления:

Технократическая система ценностей	Гуманистическая система ценностей
Природа как источник неограниченных ресурсов	Природные ресурсы ограничены
Превосходство над природой	Гармония с природой
Природа враждебна или нейтральна	Природа дружественна
Управляемая окружающая среда	Окружающая среда в хрупком равновесии
Информационно-технологическое развитие общества	Социокультурное развитие
Рыночные отношения	Общественные интересы
Риск и выигрыш	Гарантии безопасности
Индивидуальное самообеспечение	Коллективистская организация
нисе	

Разумность средств
Информация, запоминание
Образование

Разумность целей
Знания, понимание
Культура

Приведенные перечни не претендуют на полноту, они лишь иллюстрируют разницу между двумя стилями мышления. Хотя такое сравнение проводится обычно, чтобы подчеркнуть (вполне справедливо) опасные для человека последствия исключительно технократического подхода к выбору целей, полный отказ от всех технократических ценностей был бы не оправдан. Например, научно-технический прогресс является не альтернативной целям соци-

HUMANISTIC
гуманистический
CONFUSION OF
GOALS смешение
целей
PERFECTION OF
MEANS совершен-
ствование средст
MEANS AND
ENDS средства и
цели
TECHNOCRATIC
технократический

ального развития, а средством его; образование можно рассматривать как “антипод” культуры только в отрыве от нее.

Вопрос о системе ценностей, по существу, предвосхищает выбор конфигуратора, т.е. языков, на которых будет проводиться системный анализ. Об этом говорил С. Бир:

“Значение кибернетики в настоящее время определяется не приписываемой ей способностью создать сверхмозг, а необходимостью в кратчайшее время расстаться с архаической системой взглядов на сущность управления вообще. (...) Существуют ли люди, которые готовы отрицать роль этики и эстетики только потому, что создана механическая черепаха? Как это ни абсурдно, действительно существуют. Пусть же тот, кто хоть в какой-то мере понимает истинное могущество кибернетики и ее возможности, не следует их примеру” [2, с. 271].

МНОЖЕСТВЕННОСТЬ ЦЕЛЕЙ

В данном параграфе мы рассматриваем вопросы выявления целей только самого верхнего уровня, поскольку цели нижних уровней могут быть получены из верхних алгоритмически, методом декомпозиции (см. § 8.3). Однако и на самом верхнем уровне, как правило, бывает несколько целей (даже если текстуально они объединены одной формулировкой – ”глобальная цель”) и важно не упустить какую-нибудь существенную из них. Один из приемов уже был использован в предыдущем параграфе: мы воспользовались моделью над- и подсистем относительно проблемосодержащей системы. Этот позволяет учесть цели всех заинтересованных сторон. Но по отношению к каждой из сторон также возникает проблема полноты задания ее целей. Вместе с неясностью первоначальной цели это приводит к необходимости каких-то рекомендаций относительно поиска существенных целей.

Полезными оказываются следующие способы: включать в рассмотрение цели, противоположные заявленным (как в приведенном выше

примере – ”не строить больницу”), и цели “двойственные” (например, “минимизировать страдания” не то же самое, что “максимизировать удовольствия”); выявлять не только “желаемые”, но и “нежелательные” по последствиям цели (чтобы как можно раньше предвидеть возникновение новых проблем – типа загрязнения окружающей среды); допускать вообще всякие цели (их критика будет позже); единственным ограничением может служить то, что цели должны излагаться в номинальной шкале, т.е. быть названиями. Употребление более сильных шкал (см. § 6.2) – признак целей более низкого уровня, по существу, переход от целей к критериям, а это следующий этап анализа.

ОПАСНОСТЬ СМЕШЕНИЯ ЦЕЛЕЙ

Различие между целями не всегда очевидно, и существует опасность ошибочно принять одни за другие.

Например, П. Риветт [31] подчеркивает, что такая ситуация возникает обычно, когда специалисты-профессионалы, участвующие в решении проблем, навязывают свое видение мира и тем самым подменяют главные цели своими. “Операция прошла успешно, но пациент умер”, – это не дурная шутка, а действительно встречающееся среди хирургов высказывание. Риветт приводит также пример внешне очень элегантного здания, построенного в городке университета графства Сассекс. Несмотря на то, что это здание в 1965 г. получило золотую медаль Королевского Общества Архитекторов, его внутренняя планировка непригодна ни для учебных, ни для административных целей. Имеется большое количество отмеченных конкурсными призами рекламных плакатов, которые не оказали никакого влияния на сбыт рекламируемой продукции. Обследование Национальной Службы Здоровья в Англии показало, что менее одного процента времени подготовки врачей этой службы посвящено профилактической медицине, хотя организация создавалась именно с этой целью.

Смешение целей может и не быть столь явным. А. Вольштеттер иллюстрирует это на примере работы дорожной полиции:

“Отцы города хотели бы уменьшить число нарушений правил дорожного движения. Они хотели бы также наказывать штрафом или тюрьмой как можно больше нарушителей. Существует два хорошо известных альтернативных способа достижения этих целей – полицейская засада или открытое патрулирование. Первый способ увеличивает вероятность поимки нарушителя. Второй – отбивает охоту к нарушениям. Если наша цель – максимизировать число наказанных нарушителей или увеличить городской бюджет за счет штрафов, то, по-видимому, для этого лучше подойдет засада, как ни неприятно применять обманную тактику. Если же, с другой стороны, наша цель – уменьшение количества дорожных происшествий или, скажем, числа попыток нарушить правила (даже если сами такие попытки станут более успешными, поскольку нарушитель будет действовать, точно зная, присутствует ли полицейский), то вполне может оказаться, что более частое и открытое присутствие полицейских, способных немедленно покарать нарушителя, побудит водителей к осторожности и тем самым будет достигать цели наилучшим образом” [9, с. 184].

Примеры смешения целей можно продолжать. Именно это стало причиной экологических проблем Байкала, Кара-Богаз-гола и Онежского озера, проекта работ по переброске вод северных рек на юг и т.д.

Чтобы подчеркнуть важность проблемы смешения целей, П. Риветт ссылается на авторитетное мнение:

“На вопрос о том, что, по его мнению, станет главной проблемой в конце XX века, Эйнштейн ответил: совершенство средств и смешение целей. Очень похоже, что так оно и вышло” [31, с. 8].

ИЗМЕНЕНИЕ ЦЕЛЕЙ СО ВРЕМЕНЕМ

Диалектический принцип историзма относится и к целям. О возможном (и обычно происходящем) изменении целей в ходе анализа уже говорилось. Однако необходимо иметь в виду, что изменение целей во времени может происходить не только по форме, в силу все лучшего понимания действительных целей, но и по содержанию – вследствие изменения

объективных условий и/или субъективных установок, влияющих на выбор целей. Сроки стартования целей различны и зависят от многих причин. Цели более высоких уровней долговечнее. В социальных системах цели высших уровней часто формулируются как интересы будущих поколений, сроки целей нижних уровней связаны с настоящими действиями и с действиями в ближайшем будущем. Динамичность целей также должна учитываться в системном анализе.

Подведем итог

После определения проблемы After problem definition, clarification of objectives is the second следующим по важности этапом анализа становится выявление целей. Установить правильную цель важнее, чем найти наилучшую альтернативу. Не самая лучшая альтернativa всегда ведет к цели, пусть не оптимальным способом. Выбор же неправильной цели приводит не только к решению самой проблемы, сколько к появлению новых проблем.

Summary

To find the true purpose is more important than to find the best alternative. A less than optimal alternative can nevertheless satisfy the purpose, if not in the best way. The wrong choice of purpose leads not to the solution of the problem but to the appearance of new problems.

§ 9.4. ФОРМИРОВАНИЕ КРИТЕРИЕВ

Слово “критерий” будем употреблять не только в смысле “критериальная функция”, а в более широком смысле – как любой способ сравнения альтернатив. Это означает, что критериям качества альтернативы может служить любой ее признак, значение которого можно зафиксировать в порядковой или более сильной шкале (см. § 6.2). После того как такая характеристика найдена (“критерий сформирован”), появляется возможность ставить задачи выбора и оптимизации. Сложности и “ловушки”, связанные с этим выбором, были подробно рассмотрены в гл. 7. Здесь же мы обсудим, что происходит в ходе самого процесса формирования критерииев, еще до постановки задачи выбора.

КРИТЕРИИ КАК МОДЕЛЬ ЦЕЛЕЙ

Содержание процесса перехода от целей к критериям и многие особенности этого перехода становятся ясными, если рассматривать критерии как количественные модели качественных целей. В самом деле, сформированные критерии в дальнейшем в некотором смысле заменяют цели. От критерииев требуется как можно большее сходство с целями, чтобы оптимизация по критериям соответствовала максимальному приближению к цели. С другой стороны, критерии не могут полностью совпадать с целями уже хотя бы потому, что они фиксируются в различных шкалах: цели – в номинальных, критерии – в более сильных, допускающих упорядочение. Критерий – это подобие цели, ее аппроксимация, модель. Конкретнее, критерий является отображением ценностей (воплощенных в целях) на параметры альтернатив (допускающие упорядочение). Определение значения критерия для данной альтернативы является, по существу, косвенным измерением степени ее пригодности как средства достижения цели.

Начнем рассмотрение с того, что именно отображают критерии, т.е. с системы ценностей, выражаемых целями. Существуют не только альтернативные системы ценностей (как технократическая и гуманистическая; см. § 9.3), но и совместимые: философские, психологические, познавательные, моральные, экономические, политические, этические, эстетические, которых в каждой конкретной ситуации человек придерживается одновременно, в полном или любом частичном выборе, в виде конфигуратора (см. § 8.5). Эти системы ценностей несводимы друг к другу, несравнимы, неупорядочены, поэтому уже отсюда следует многокритериальность реальных задач, связанных со сложными системами.

Практическое построение критериев привело к необходимости обобщения такого опыта, т.е. к созданию “теории ценностей”, главной задачей которой была бы алгоритмизация перехода от ценности к критерию. Трудно предположить, что можно построить “общую теорию ценностей”, но частные такие теории существуют и развиваются. Наиболее известны теория экономической ценности (стоимости) [23, гл. 9], теория полезности (частично освещенная в гл. 6 и 7 данной книги) или психологическая теория ценности [23, гл. 10], казуистическая теория ценности (оценивающая ценности по прошлым прецедентам [23, гл. 12]). Однако даже в рамках частных теорий имеются значительные трудности, неясности и сложности. В целом в настоящее время картина такова, что построение критериев является более искусством, чем наукой. Правда, выражения подобного рода как бы подразумевают, будто в науке не может или даже не должно быть искусства. На деле же неформализуемые, творческие, эвристические этапы часто являются самыми важными в процессе научного исследования.

ПРИЧИНЫ МНОГОКРИТЕРИАЛЬНОСТИ РЕАЛЬНЫХ ЗАДАЧ

Многоокритериальность реальных задач связана не только с множественностью целей, но и с тем, что одну цель редко удается выразить одним критерием, хотя к этому обычно стремятся. Конечно, возможны случаи, когда единственный критерий отвечает требованиям практики. Например, по стандартам ЮНЕСКО уровень медицинского обслуживания оценивается по статистике детской смертности. Другой любопытный пример такого рода привел Дж. Моррисей:

“Самая трудная проблема заключается в том, чтобы преодолеть эмоциональное сопротивление части управляющих намерению “количественно оценить качество”. Они обычно заявляют: “Нет точных и надежных методов для количественной оценки творческого труда”. Это возражение справедливо, однако позвольте сослаться на личный пример.

В 1963 г. я работал преподавателем на курсах по управлению производством фирмы “Норт Америкен”. Тогда показателем для оценки эффективности учебного курса служило число студенто-часов, приходящихся на каждого преподавателя по данной дисциплине. Считая себя хорошим специалистом, я стремился повышать качество учебной работы, а не увеличивать ее объем, и вместе с другими работниками учебного отделения протестовал против метода оценки нашего труда по объему. “Вы хотите от нас получить мнение о количестве или о качестве учебной работы? Мы не можем одновременно отвечать установленным требованиям к качеству занятий и давать большие цифры!” Но когда наши жалобы и протесты стихли и нам пришлось примириться с заданным нормативом, произошло нечто интересное. Мы обнаружили, что существует четкая зависимость между числом студенто-часов, прочитанных по данному курсу, и качеством преподавания. Оказалось, что у преподавателей, которыми были довольны и слушатели, и их начальники, всегда было большее число слушателей, посетивших их занятия. Более того, требования о выполнении нормы по студенто-часам повышали чувство ответственности преподавателя за увеличение числа желающих заниматься у него и за сокращение числа пропусков занятий. Мы поняли, что преподаватель действительно неудовлетворительно выполняет порученную ему работу, если не привлекает нужного числа слушателей из контингента, на который рассчитан его курс. Мы поняли также, что самая лучшая учебная программа бесполезна, если нет желающих ею заниматься или если ею занимаются не те, на кого она рассчитана.

(...) Такая система, конечно, не безупречна. Но до сих пор я еще не встречал системы управления, которую управляющий не мог бы “разгромить”, если бы захотел. Однако, если вас чрезмерно заботит то, что предлагаемую вами систему жестоко раскритикуют, это значит, что вы еще не готовы к внедрению системы целевого управления” [10, с. 59 – 60].

Все же случаи, когда единственный критерий удачно отображает цель, скорее исключение. Это и понятно: критерий лишь приближенно (как и всякая модель) отображает цель, и адекватность одного критерия может оказаться недостаточной. Например, критерий быстроты прибытия пожарных не адекватен цели борьбы с пожарами: он не связан с уменьшением числа возгораний. Объем расходов на одного ученика не оценивает качества обучения в

COMPROMISE
компромисс
MULTICRITERIAL
многокритериальный
CONSTRAINTS,
CONDITIONS
ограничения
CORPORATION
оргсистема
VALUES *ценности*

Критерии являются количественными моделями качественных целей. Следовательно, расхождения критериев

моделям (входов оргсистемы, см. § 8.2; перечня заинтересованных лиц, см. § 9.2, и т.п.) добавим еще одну важную модель, которая широко используется в системном анализе (а не только для данной цели).

Это модель проблемной ситуации как совокупности следующих трех взаимодействующих систем:

проблемосодержащей системы, т.е. системы, в которой данная ситуация воспринимается как проблема (мы о ней уже говорили в § 9.2);

проблеморазрешающей системы, т.е. системы, которая может так повлиять на ход событий, чтобы проблема исчезла или ослабела;

окружающей среды, в которой существуют и с которой взаимодействуют обе системы.

Эта модель фактически уже была использована в алгоритме декомпозиции (см. гл. 8). Там “объект анализа” и выражал интересы проблемосодержащей системы, а в качестве “целевой” выступала именно проблеморазрешающая система. Было отмечено, что иногда (но на самом деле редко) обе эти системы могут оказаться совмещеными. Окружающая среда отображалась лишь в виде входов и выходов целевой системы.

Применительно к формированию критериев данная модель позволяет не только повысить полноту набора критериев, но и структурировать их совокупность, ввести такие различия

школе; число студентов на одного преподавателя не однозначно связано с качеством подготовки специалистов в вузе и т.д.

Решение может состоять не только в поиске более адекватного критерия (возможно, он и не существует), но и в использовании нескольких критериев, описывающих одну цель по-разному и дополняющих друг друга. Э. Квейд приводит интересный опыт формирования критериев для достаточно ясной цели: улучшить уборку мусора в большом городе. В результате анализа были отвергнуты как неадекватные следующие, на первый взгляд, подходящие критерии: расходы по уборке мусора в расчете на одну квартиру, число тонн убираемого мусора в расчете на один рабочий человеко-час, общий вес вывозимого мусора – эти критерии ничего не говорят о качестве работы. Более удачными были признаны такие критерии, как процент жилых кварталов без заболеваний, снижение числа пожаров из-за возгорания мусора, уменьшение числа укусов людей крысами, количество обоснованных жалоб жителей на скопление мусора. Впрочем, очевидно, что и эти критерии отражают только отдельные стороны качества уборки мусора в городе.

КРИТЕРИИ И ОГРАНИЧЕНИЯ

Подчеркнув, что многокритериальность является способом повышения адекватности описания цели, обратим теперь внимание на то, что *дело не только и не столько в количестве критериев, сколько в том, чтобы они достаточно полно “покрывали” цель*. Это означает, что критерии должны описывать по возможности все важные аспекты цели, но при этом желательно минимизировать число необходимых критериев.

Последнее требование удовлетворяется, если критерии являются независимыми, не связанными друг с другом (например, желательно не использовать в различных составных критериях одинаковые измеряемые величины или величины, выводимые друг из друга, и т.п.).

Что же касается выполнения требования полноты, всесторонности, то у нас есть уже испытанный способ: использовать достаточно полные модели рассматриваемой ситуации. Какие именно из этих моделей или какие их комбинации окажутся полезными, должен определить системный аналитик, в этом и состоит его работа. К уже рассмотренным ранее

моделям (входов оргсистемы, см. § 8.2; перечня заинтересованных лиц, см. § 9.2, и т.п.) добавим еще одну важную модель, которая широко используется в системном анализе (а не только для данной цели).

Это модель проблемной ситуации как совокупности следующих трех взаимодействующих систем:

проблемосодержащей системы, т.е. системы, в которой данная ситуация воспринимается как проблема (мы о ней уже говорили в § 9.2);

проблеморазрешающей системы, т.е. системы, которая может так повлиять на ход событий, чтобы проблема исчезла или ослабела;

окружающей среды, в которой существуют и с которой взаимодействуют обе системы.

Эта модель фактически уже была использована в алгоритме декомпозиции (см. гл. 8). Там “объект анализа” и выражал интересы проблемосодержащей системы, а в качестве “целевой” выступала именно проблеморазрешающая система. Было отмечено, что иногда (но на самом деле редко) обе эти системы могут оказаться совмещеными. Окружающая среда отображалась лишь в виде входов и выходов целевой системы.

Применительно к формированию критериев данная модель позволяет не только повысить полноту набора критериев, но и структурировать их совокупность, ввести такие различия

между ними, которые впоследствии облегчат постановку оптимизационных задач. Суть дела (упрощенно) заключается в том, что характер целей трех данных участников проблемной ситуации различен. Для проблемосодержащей системы это цели достижения (главное – разрешить проблему); цели проблеморазрешающей системы связаны с рациональным расходом ресурсов на разрешение проблемы (главное – экономно разрешить ее), а вмешательство среды в этот процесс носит, как правило, характер пассивный, но обязательный (главное – не предпринимать ничего, что противоречило бы законам природы). Так возникают критерии эффективности, подлежащие оптимизации (нередко противоречивые), а также критерии сохранения, ограничения, требующие соблюдения их постоянства. При постановке оптимизационных задач все эти критерии связываются воедино (формальная техника этого была рассмотрена в § 6.2). Содержательная сторона этого процесса состоит в том, что происходит согласование ресурсов проблеморазрешающей системы с потребностями проблемосодержащей системы при ограничениях, накладываемых средой. При этом может возникнуть необходимость в итерациях, чего не следует избегать в исследованиях сложных систем.

Между целевыми критериями и ограничениями имеется сходство и различие. На последующих этапах, например при решении некоторых типов оптимизационных задач, они могут выступать равноправно (как в уравнениях Эйлера или в терминах теории расплывчатых множеств). Однако на этапах формирования критериев и генерирования альтернатив (последний будет рассмотрен в следующем параграфе) они различаются существенно. Целевой критерий как бы открывает возможности для выдвижения все новых и новых альтернатив в поисках лучшей из них, а ограничение заведомо уменьшает их число, запрещая некоторые из альтернатив. Одними целевыми критериями можно жертвовать ради других, а ограничение исключить нельзя, оно должно жестко соблюдаться. В этом смысле ограничения упрощают, а не усложняют работу системного аналитика.

В практике системного анализа встречаются случаи, когда наложенные ограничения столь сильны, что делают нереальным достижение цели. Тогда системный аналитик должен ставить перед лицом, принимающим решение, вопрос о том, нельзя ли данные ограничения ослабить или снять совсем. Например, слишком жесткие требования на вероятность ложной тревоги приводят к неприемлемо длительным периодам накопления сигнала в радиолокаторе. Как выяснилось, столь жесткое требование вытекало лишь из нежелания “слишком часто” беспокоить зря вышестоящее начальство.

Итак, при формировании критериев ищется компромисс между полнотой (точностью) описания целей и количеством критериев. Кроме общих рекомендаций типа изложенных выше можно использовать и практический опыт исследования конкретных систем. Так, Э. Квейд [9] перечисляет критерии, наиболее часто встречающиеся в анализе сложных технических систем: финансовые (прибыль, стоимость и пр.), объемные (измеряющие количество продукта), технические качества (эффективность функционирования, надежность и т.д.), живучесть (совместимость с уже существующими системами, приспособляемость или гибкость, стойкость против морального старения, безопасность) и ряд других. Такие эмпирические перечни, безусловно, полезны, но их следует рассматривать лишь как основу дальнейшего поиска.

Не следует абсолютизировать критическое отношение к тому, что расхождение между критериями и целями неизбежно. Древняя поговорка гласит: “Можно много пройти в башмаках, которые немного жмут”.

Подведем итог

Summary

Считая критерии оценки альтернатив количественными моделями качественных целей, мы можем повысить полноту и адекватность созданного набора критериев для

Considering the criteria for evaluating alternatives as quantitative models of qualitative purposes enables us to improve the completeness and adequacy of the criteria set for a given problem.

конкретной задачи.

§ 9.5. ГЕНЕРИРОВАНИЕ АЛЬТЕРНАТИВ

Теория выбора (изложенная в гл. 7) исходит из того, что задано множество альтернатив, т.е. считается, что уже имеется то, из чего выбирать, и вопрос состоит в том, как выбирать. Это наглядный пример сугубо формальной постановки задачи: все главные, принципиальные трудности считаются уже преодоленными, и речь идет, можно сказать, о “технических” трудностях. А ведь именно формирование множества альтернатив и является наиболее трудным, наиболее творческим этапом системного анализа. Так, А. Холл отмечал:

“Стадия поисков идей, несомненно, представляет собой кульминационную точку процесса решения задачи, ведь без идей нечего анализировать и выбирать”. [23, 120].

Более того, все направлено на поиск самой лучшей альтернативы в заданном множестве альтернатив, и если в это множество мы по каким-то причинам не включили действительно наилучшую, то никакие методы выбора ее не “вычислят”.

Генерирование альтернатив, т.е. идей о возможных способах достижения цели, является настоящим творческим процессом. Поэтому все рекомендации относительно того, как самому генерировать новые альтернативы или как создать условия для того, чтобы другие участники анализа лучше генерировали эти идеи, – результат коллективного опыта теоретиков и методистов по решению творческих задач. Этот опыт изложен в ряде книг (например, [1; 2; 4; 5; 7; 11; 14; 15; 19; 20]), которые содержат много интересной и полезной информации.

Один из способов структурирования любой неформальной деятельности состоит в выявлении и формировании присущих ей эвристик, т.е. эмпирических правил, полезность которых обоснована лишь тем, что они во многих (хотя и не во всех) случаях приводят к успеху. Применительно к процессу генерирования альтернатив в литературе по системному анализу и методам творчества наиболее часто описываются следующие эвристики.

СПОСОБЫ УВЕЛИЧЕНИЯ ЧИСЛА АЛЬТЕРНАТИВ

Важно сознательно сгенерировать как можно большее число альтернатив. Для этого используют различные способы: а) поиск альтернатив в патентной и журнальной литературе; б) привлечение нескольких квалифицированных экспертов, имеющих разнообразную подготовку и опыт; в) увеличение числа альтернатив за счет их комбинирования, образования промежуточных вариантов между предложенными ранее (т.е. не “либо-либо”, а кроме того, еще и “от одной и от другой альтернативы”); г) модификация имеющейся альтернативы, т.е. формирование альтернатив, лишь частично отличающихся от известной; д) включение альтернатив, противоположных предложенным, в том числе и “нулевой” альтернативы (“не делать ничего”, т.е. рассмотреть последствия развития событий без нашего вмешательства); е) интервьюирование заинтересованных лиц и более широкие анкетные опросы; ж) включение в рассмотрение даже тех альтернатив, которые на первый взгляд кажутся глупыми или надуманными; з) генерирование альтернатив, рассчитанных на различные интервалы времени (долгосрочные, краткосрочные, экстренные), и т.д.

СОЗДАНИЕ БЛАГОПРИЯТНЫХ УСЛОВИЙ

При организации работы на этапе генерирования альтернатив следует помнить о существовании факторов, как тормозящих творческую работу, так и способствующих ей. Выделяют [23, § 16.6] внутренние (психологические) и внешние факторы.

К внутренним факторам относят: а) последствия неправильного восприятия действительности (крайние проявления: либо мы воспринимаем то, чего нет, либо не воспринимаем того, что есть); б) интеллектуальные преграды (инерционность мышления, довлеющие стереотипы, подсознательные самоограничения, связанные с убеждениями, лояльностью, и пр.); в) эмоциональные преграды; излишнее увлечение критикой других или, наоборот, боязнь критики со стороны других; опасение отрицательной реакции со стороны заказчика или начальства на предложенные альтернативы, субъективное отношение к “любимым” типам альтернатив (например, некоторые ярые сторонники теории массового обслуживания любые задачи стараются свести к задачам очередности) и т.д.

Большое влияние на творческие процессы оказывают и внешние факторы. Замечено, что даже физические (погодные и климатические) условия сказываются на продуктивности творческого труда. Так, некоторые исследователи считают, что существует связь между творческой деятельностью целых народов и географическими условиями их жизни. На индивидуальной деятельности также сказываются физические условия. Рассказывают, что Нильс Бор отпускал студентов с занятий, если было так жарко, что воск в пробирке расплывался; что Тимофеев-Ресовский однажды в жаркий день провел заседание международного симпозиума прямо в пруду, и это было, по воспоминаниям участников, самое плодотворное заседание. Известно отрицательное влияние посторонних шумов, различных неудобств и т.п.

Однако наиболее сильное влияние на индивидуальное творчество оказывают общественные условия, общий культурный фон, идеальная атмосфера. Одобрение определенной социальной группы – один из самых сильных стимулов для творчества человека.

СПОСОБЫ СОКРАЩЕНИЯ ЧИСЛА АЛЬТЕРНАТИВ

Если специально стремиться к тому, чтобы на начальной стадии было получено как можно больше альтернатив, то для некоторых проблем их количество может достичь многих десятков. Очевидно, что подробное изучение каждой из них приведет к неприемлемым затратам времени и средств. *В таких случаях рекомендуется провести “грубое отсеивание”, не сравнивая альтернативы количественно, а лишь проверяя их на при-сутствие некоторых качеств, желательных для любой приемлемой альтернативы.* К признакам “хороших” альтернатив относят устойчивость при изменении некоторых внешних условий, надежность, многоцелевую пригодность, адаптивность, другие признаки “практичности”. В отсеве могут помочь также обнаружение отрицательных побочных эффектов, недостижение контрольных уровней по некоторым важным показателям (например, слишком высокая стоимость) и пр. Предварительный отсев не рекомендуется проводить слишком жестко; для детального анализа необходимы хотя бы несколько альтернатив.

Рассмотрим теперь некоторые хорошо зарекомендовавшие себя на практике организационные формы генерирования альтернатив.

МОЗГОВОЙ ШТУРМ

Метод мозгового штурма специально разработан для получения максимального количества предложений. Его эффективность поразительна:

“Шесть человек за полчаса могут выдвинуть 150 идей. Бригада проектировщиков, работающая обычными методами, никогда не пришла бы к мысли о том, что рассматриваемая ею проблема имеет такое разнообразие аспектов” [7, с. 244].

Техника мозгового штурма такова. Собирается группа лиц, отобранных для генерации альтернатив; главный принцип отбора – разнообразие профессий, квалификации, опыта (такой принцип поможет расширить фонд априорной информации, которой располагает группа). Сообщается, что приветствуются любые идеи, возникшие как индивидуально, так и по ассоциации при выслушивании предложений других участников, в том числе и лишь частично улучшающие чужие идеи (каждую идею рекомендуется записать на отдельной карточке). *Категорически запрещается любая критика* – это важнейшее условие мозгового штурма: сама возможность критики тормозит воображение. Каждый по очереди зачитывает свою идею, остальные слушают и записывают на карточки новые мысли, возникшие под влиянием услышанного. Затем все карточки собираются, сортируются и анализируются, обычно другой группой экспертов.

“Общий выход такой группы, где идея одного может навести другого на что-то еще, часто оказывается больше, чем общее число идей, выдвинутых тем же количеством людей, но работающих в одиночку” [23, с. 418].

Число альтернатив можно впоследствии значительно увеличить, комбинируя сгенерированные идеи. Среди полученных в результате мозгового штурма идей может оказаться много глупых и неосуществимых, но

“глупые идеи легко исключаются последующей критикой, ибо компетентную критику легче получить, чем компетентное творчество” [23, с. 418].

Известно много замечательных примеров успешного мозгового штурма. Приведем лишь один из них, иллюстрирующий пользу запрещения критики.

Во время войны проблема противодействия минам и торпедам противника на море была подвергнута мозговому штурму. Одна из идей заключалась в следующем: “Пусть, как только будет обнаружена мина или торпеда, вся команда станет к борту и дует на нее!”. Эта несерьезная, на первый взгляд, идея не была отвергнута, и при дальнейшем анализе рациональное зерно, содержащееся в ней, было трансформировано в предложение с помощью мощных

BRAINSTORM-	насосов создавать потоки воды, отталкивающие опасный
ING мозговой	объект.
штурм	
MORPHOLOGI-	СИНЕКТИКА
CAL морфологиче-	Синектика предназначена для генерирования альтернатив путем ассоциативного мышления, поиска аналогий поставленной задаче. В противоположность мозговому
ский	штурму здесь целью является не количество альтернатив, а генерирование небольшого числа альтернатив (даже единственной альтернативы), разрешающих данную проблему. Эффективность синектики была продемонстрирована при решении конкретных технических проблем [7, § 10.2] типа “найти простой принцип устройства приводов с постоянной угловой скоростью”, “спроектировать усовершенствованный нож для открывания консервных банок”, “изобрести более прочную крышу”, “разработать герметичную застежку для костюма космонавта”; известен случай синектического решения более общей проблемы экономического плана: “разработать новый вид продукции с годовым потенциалом продаж 300 млн. долларов”. Имеются попытки применения синектики в решении социальных проблем типа “как распределить государственные средства в области градостроительства”. Суть синектики можно кратко изложить следующим образом. Формируется группа из 5 – 7 человек, отобранных по признакам гибкости мышления, практического опыта (предпочтение отдается людям, менявшим профессии и специальности), психологической совместимости, общительности, подвижности (последнее, как станет ясно из изложенного ниже, очень важно). Выработав определенные навыки совместной работы, группа ведет систематическое направленное обсуждение любых аналогий с подлежащей решению проблемой, спонтанно возникающих в ходе бесед. Перебираются не только три уже известных (см. § 2.3) вида подобия (прямое, косвенное и условное),
COMBINATOR-	
ICS перебор	
SYNECTICS си-	
нектика	
EURISTICS эври-	
стика	

Генерирование альтернатив является творческим процессом, на результативность которого влияют как внешние условия, так и в особенностях внутренние, психолого-психические условия. Разрабо-

но и чисто фантастические аналогии (последнее означает попытки представить себе вещи такими, какими они не являются на самом деле, но какими хотелось бы их видеть; например, можно вообразить, что дорога существует только непосредственно под колесами автомашин).

Особое значение синектика придает аналогиям, порождаемым двигательными ощущениями. Это вызвано тем, что наши природные двигательные рефлексы сами по себе высокоорганизованы и их осмысление может подсказать хорошую системную идею. Предлагается, например, вообразить свое тело на месте совершенствующего механизма, “почувствовать себя

им”, либо поставить себя на место фантастического организма, выполняющего функцию проектируемой системы, и т.п. Раскрепощенность воображения, интенсивный творческий труд создают атмосферу душевного подъема, характерную для синектики. Отмечаются и психологические затруднения, возникающие у новичков в случае применения этого метода: появление угрызений совести (“получаем деньги за приятное времяпрепровождение”); за- знайство после удачного решения первой проблемы; истощение нервной системы в результате столь интенсивной работы. Успеху работы синектических групп способствует соблюдение определенных правил, в частности: 1) запрещено обсуждать достоинства и недостатки членов группы; 2) каждый имеет право прекратить работу без каких-либо объяснений при малейших признаках утомления; 3) роль ведущего периодически переходит к другим членам группы и т.д.

В США создана специальная фирма “Синектика инкорпорейтед”, занимающаяся консультированием и обучением в области синектики. Подчеркнем, что в отличие от мозгового штурма при использовании синектики требуется специальная и длительная подготовка: “В течение года 5 или 6 человек должны затратить 1/4 своего рабочего времени на обучение. Группа обученных синекторов, работающих полный рабочий день, способна в течение года найти приемлемые решения примерно четырех небольших и двух крупных проблем” [7, с. 250].

РАЗРАБОТКА СЦЕНАРИЕВ

В некоторых проблемах (особенно в социотехнических) искомое решение должно определить реальное будущее течение событий. В таких случаях альтернативами являются различные (воображаемые, но правдоподобные) последовательности действий и вытекающих из них событий, которые могут произойти в будущем с исследуемой системой. Эти последовательности имеют общее начало (настоящее состояние), но затем возможные состояния различаются все сильнее, что и приводит к проблеме выбора. Такие гипотетические альтернативные описания того, что может произойти в будущем, называют сценариями, а рассматриваемый метод – разработкой сценариев. Сценарии-альтернативы представляют ценность для лиц, принимающих решения, только тогда, когда они не просто плод фантазии, а логически обоснованные модели будущего, которые после принятия решения можно рассматривать как прогноз, как приемлемый рассказ о том, “что случится, если...”. Создание сценариев относится к типичным неформализуемым процедурам, представляет собой творческую, научную работу. Тем не менее и в этом деле накоплен определенный опыт, имеются свои эвристики. Например, рекомендуется разрабатывать “верхний” и “нижний” сценарии – как бы крайние случаи, между которыми может находиться возможное будущее. Такой прием позволяет отчасти компенсировать или явно выразить неопределенности, связанные с предсказанием будущего. Иногда полезно включать в сценарий воображаемый активно противодействующий элемент, моделируя тем самым “наихудший случай”. Кроме того, рекомендуется не разрабатывать детально (как ненадежные и непрактичные) сценарии, слишком “чувствительные” к небольшим отклонениям на ранних стадиях. Важными этапами создания сценариев являются: составление перечня факторов, влияющих на ход событий, со специальным выделением лиц, которые контролируют эти факторы прямо или косвенно; выделение аспектов борьбы с такими факторами как некомпетентность, халатность и недисциплинированность, бюрократизм и волокита; учет наличных ресурсов и т.д.

МОРФОЛОГИЧЕСКИЙ АНАЛИЗ

Морфологический анализ – простой и эффективный способ генерирования альтернатив – предложен Ф. Цвикки [13]. Он состоит в выделении всех независимых переменных проектируемой системы, перечислении возможных значений этих переменных и генерировании альтернатив перебором всех возможных сочетаний этих значений.

Проиллюстрируем суть морфологического анализа на упрощенном примере разработки системы телевизионной связи [23]:

Независимая переменная	Значения переменной
1. Цвет изображения	1. Черно-белое 2. Одноцветное (например, все оттенки красно-белого) 3. Двухцветное 4. Трехцветное ? ? 8. Семицветное
2. Размерности изображения	1. Плоское изображение 2. Объемное изображение
3. Градация яркости	1. Непрерывные 2. Дискретные (оцифрованные)
4. Звуковое сопровождение	1. Без звука 2. Монауральный звук 3. Стереофонический звук
5. Передача запахов	1. Без передачи запахов 2. С сопровождением запахов
6. Обратная связь	1. Без обратной связи 2. С обратной связью

Эта таблица порождает $8 \times 2 \times 2 \times 3 \times 2 \times 2 = 384$ различные возможные системы. Современному телевизионному вещанию соответствует только одна альтернатива: 1.4 – 2.1 – 3.1 – 4.2 – 5.1 – 6.1. Имеется повод для размышления, почему остальные альтернативы до сих пор не привлекли внимания инженеров.

Отметим также, что количество вариантов можно увеличить, вводя новые независимые переменные (в рассмотренном примере ввести размеры изображения с градацией от обычных на сегодня до размеров во всю стену, ввести дополнительные каналы передачи информации, например кожно-электрический или тактильный; перейти от одноэкранной системы к многоэкранной и т.д.). Одна из основных проблем морфологического анализа при увеличении числа переменных – это проблема сокращения перебора. Она решается наложением различных ограничений, которые позволяют отбросить варианты, не подлежащие рассмотрению.

ДЕЛОВЫЕ ИГРЫ

Деловыми играми называется имитационное моделирование реальных ситуаций, в процессе которого участники игры ведут себя так, будто они в реальности выполняют порученную им роль, причем сама реальность заменяется некоторой моделью. Примерами являются штабные игры и маневры военных, работа на тренажерах различных операторов технических систем (летчиков, диспетчеров электростанций и т.д.), административные игры и т.п. Несмотря на то что чаще всего деловые игры используются для обучения, их можно использовать и для экспериментального генерирования альтернатив, особенно в слабо формализованных ситуациях. Важную роль в деловых играх кроме участников играют контрольно-арбитражные группы, управляющие моделью, регистрирующие ход игры и обобщающие ее результаты.

Мы рассмотрели далеко не все методы генерирования альтернатив. Так, например, кроме перечисленных методов можно успешно использовать модифицированный метод “Делфи” [23]; возможны и различные сочетания других методов. Важным моментом является также итеративность: на любой последующей стадии системного анализа должна иметься возможность порождения новой альтернативы и включения ее в анализ. Отметим, что при рассмотр-

рении слабо структурированных (“рыхлых”) проблем часто находят первую же подходящую альтернативу и производят ее пошаговое улучшение, совершенствование.

Подведем итог

Составление перечня возможных вари-антов решения про-blem (генерирование альтерна-tив) является самым на-пряженным творческим этапом систем-ного анализа. Для об-легчения и ускоре-ния этой работы разработано много спо-собов, владение которыми необхо-димо системному ана-литику.

Summary

The generation of alternatives (i.e., of feasible options for solving the problem) is perhaps the most creative stage of systems analysis. There are many ways of facilitating and stimulating this process that must be mastered by the analyst.

§ 9.6. АЛГОРИТМЫ ПРОВЕДЕНИЯ СИСТЕМНОГО АНАЛИЗА

Мы уже неоднократно отмечали, что совершенствование любой деятельности состоит в ее алгоритмизации, т.е. в совершенствовании технологии. Было бы по меньшей мере нелогично и странно не подходить с такой же меркой и к самому системному анализу: естественно поставить вопрос о том, насколько формализован сам системный анализ.

ТРУДНОСТИ АЛГОРИТМИЗАЦИИ СИСТЕМНОГО АНАЛИЗА

Если понимать формализацию узко (в идеале как математическую постановку задачи и вполне однозначную программу ее решения), то системный анализ в принципе не может быть полностью формализован, поскольку в нем большую и очень важную роль играют эта-пы, на которых системный аналитик и привлекаемые им эксперты должны выполнить твор-ческую работу. Такое положение не является особенностью только системного анализа, оно характерно для научной работы вообще. По этому поводу И. Пригожин и И. Стенгерс пишут: “Для нас экспериментальный метод является поистине искусством, т.е. мы считаем, что в основе его лежат особые навыки и умения, а не общие правила. Будучи искусством, экспери-ментальный метод никогда не гарантирует успех, всегда оставаясь на милости тривиальности или неверного суждения. Ни один методологический принцип не может исключить, напри-мер, риска зайти в тупик в ходе научного исследования. Экспериментальный метод есть ис-кусство постановки интересного вопроса и перебора всех его следствий, вытекающих из ле-жащей в его основе теоретической схемы, всех ответов, которые могла бы дать природа на выбранном экспериментатором теоретическом языке. Из конкретной сложности и многооб-разия явлений природы необходимо выбрать одно-единственное явление, в котором с наибольшей вероятностью ясно и однозначно должны быть воплощены следствия из рас-сматриваемой теории. Это явление затем надлежит абстрагировать от окружающей среды и “инсценировать” для того, чтобы теорию можно было подвергнуть воспроизведимой про-верке, результаты и методы которой допускали бы передачу любому заинтересованному ли-цу. Хотя такого рода экспериментальная процедура с самого начала вызывала (и продолжает вызывать) серьезные нарекания, (...) она пережила все модификации теоретического содер-жания научных описаний и в конечном счете определила новый метод исследования, вве-денный современной наукой. Экспериментальная процедура может становиться и орудием чисто теоретического анализа. Эта ее разновидность известна под названием мысленного эксперимента” [21, с. 86 – 87].

Из приведенной цитаты наглядно видно, что современный системный анализ – это просто одна из современных прикладных наук; и главное его отличие от других наук состоит в от-

существии (точнее, в расширении) субстратной специфики: системный анализ применим к системам любой природы.

Тот факт, что в арсенале системного анализа имеются как формальные (в том числе и математические) процедуры, так и операции, которые выполняются людьми неформально, и даже то, что в некоторых случаях анализ вообще не использует формализованные процедуры, тем не менее не означает, что нельзя говорить об алгоритмах системного анализа. Наоборот, требование системности (в первую очередь, целенаправленной структурированности) применимо прежде всего к самому системному анализу.

Неоднократно имели место попытки создать достаточно общий, универсальный алгоритм системного анализа (пять вариантов такого алгоритма приведены в табл. 8.1, еще несколько содержатся в [6; 12; 16; 17; 18; 22]). Бросаются в глаза общность в целом и различия в деталях приводимых алгоритмов. Было бы необоснованным утверждать, что какой-то из них “более правильный” или “более универсальный”, чем другой, что реализация одного из них – это системный анализ, а реализация другого – нет. Положение становится ясным, если вспомнить, что алгоритм является прагматической (нормативной) моделью деятельности. Выбрав конкретную модель, мы должны следовать предписаниям именно данного алгоритма, но это не означает, что нельзя было воспользоваться другой моделью. Из целевой предназначенности моделей вытекает, что для какого-то случая конкретный алгоритм предпочтительнее другого; однако отсюда не следует, что не может существовать еще одна, лучшая реализация алгоритма или что в каком-то другом случае порядок предпочтения алгоритмов не окажется противоположным. Впрочем, и полной независимости алгоритмов также нет: одни из них могут являться расширением других, частично совпадать и т.д.

Очень удобной и наглядной (и, кроме того, достаточно глубокой) аналогией служит программирование решения задачи на ЭВМ. Программист имеет в распоряжении все операторы некоторого языка и должен составить из них программу решения задачи. Разные программисты составят различные программы; одна из них будет написана изящнее; в какой-то из них использованы новинки; какая-то окажется “старомодной”. Однако все они решают одну и ту же задачу. Подобно этому, системный аналитик может в разной (хотя и не полностью произвольной) последовательности использовать различные операции исследования систем или спланировать свои действия заранее (для хорошо структурированных, например технических, задач), а может выбирать очередную операцию в зависимости от исхода предыдущей, либо использовать готовый алгоритм или “готовые” подпрограммы анализа.

КОМПОНЕНТЫ СИСТЕМНЫХ ИССЛЕДОВАНИЙ

Перечислим основные средства исследования систем (этапы системного анализа), т.е. блоки, из которых может состоять процедура анализа конкретной системы (в скобках указаны те места данной книги, где данная операция описана подробно):

- определение конфигуратора (§ 8.5 и 9.4);
- определение проблемы и проблематики (§ 9.2);
- выявление целей (§ 9.3);
- формирование критериев (§ 9.4);
- генерирование альтернатив (§ 9.5);
- построение и использование моделей (гл. 2 и 3, § 8.2);
- оптимизация (для простых систем; § 7.2 и 7.9);
- выбор (гл. 7);
- декомпозиция (§ 8.3);
- агрегирование (§ 8.5);
- исследование информационных потоков (гл. 5);
- исследование ресурсных возможностей (§ 4.4 и 9.4);
- наблюдения и эксперименты над исследуемой системой (гл. 6);
- реализация, внедрение результатов анализа (§ 9.7).

Приведенный перечень является укрупненным, может иметь другую последовательность операций, каждую указанную операцию можно разделить на более мелкие операции. Это и позволяет составлять алгоритмы системного анализа, имеющие различную степень подробности.

Итак, исследование каждой системы проводится с использованием любых необходимых методов и операций системного анализа (как формальных, в том числе с применением математических методов и ЭВМ, так и эвристических), а их конкретная последовательность определяется ведущим исследование системным аналитиком и во многом носит индивидуальный, приспособленный к данному случаю характер. Поэтому в системном анализе переплетаются особенности, присущие как науке, так и искусству. Д.К. Джонс, говоря о проектировании систем (частном случае системного анализа), выразил эту мысль так:

“Скажем сразу, что проектирование не следует путать ни с искусством, ни с естественными науками, ни с математикой. Это сложный вид деятельности, в котором успех зависит от правильного сочетания всех этих трех средств познания; очень мала вероятность добиться успеха путем отождествления проектирования с одним из них” [7, с. 49].

“При решении любой задачи проектирования необходимо определенное сочетание логики и интуиции. Пути такого сочетания интуитивного с рациональным не установлены; пожалуй, их и невозможно установить в общем виде, в отрыве от конкретной задачи и конкретного человека, так как они зависят от того, какое количество объективной информации имеется в распоряжении проектировщика, а также от его квалификации и опыта” [7, с. 90].

Подведем итог

Алгоритмы проведения системного анализа могут быть различными. В зависимости от степени сложности анализируемой проблемы употребляются “линейные” алгоритмы (в простейших случаях), алгоритмы с циклами (чем сложнее система, тем больше циклов и тем больше итераций осуществляется в каждом цикле), сложные “последовательные”, т.е. конструируемые в ходе исследования, алгоритмы (в том числе содержащие циклы, случайный поиск, адаптацию, самоорганизацию и т.д.).

Summary

The algorithms of systems analysis are varied. Depending on the complexity of the studied problem, they vary from linear algorithms (in the simplest cases), through algorithms involving cycles (and the more complex system, the more cycles, cycles within cycles, and repetition of cycles), and up to sequential algorithms designed in the course of study (they may contain such elements, random search, adaptation, self-organization, and so on).

§ 9.7. ПРЕТВОРЕНИЕ В ЖИЗНЬ РЕЗУЛЬТАТОВ СИСТЕМНЫХ ИССЛЕДОВАНИЙ

В отличие от теории систем системный анализ является прикладной наукой, его конечная цель – изменение существующей ситуации в соответствии с поставленными целями. Поэтому окончательное суждение о правильности и полезности системного анализа или его неправильности можно сделать на основании результатов его практического применения. Это же зависит не только от того, насколько совершенны и теоретически обоснованы методы, использованные в ходе анализа, но и от того, насколько правильно реализованы полученные рекомендации. В практике системного анализа имеется много случаев, когда хорошие и правильные (как впоследствии показывала практика) результаты системных исследований были отвергнуты лицами, принимающими решения; известны случаи, когда внедренные в практи-

ку рекомендации использовались лишь до тех пор, пока их автор сам участвовал в их реализации.

ВНЕДРЕНИЕ РЕЗУЛЬТАТОВ СИСТЕМНОГО АНАЛИЗА В ПРАКТИКУ

В связи с этим за последние годы в системном анализе усилилось внимание к проблемам внедрения его результатов в практику. Был выделен самостоятельный раздел системного анализа, получивший, по предложению Р. Акоффа [24; 25], название “теории практики”; его задачей является исследование условий эффективности претворения в жизнь результатов системного анализа. При этом подчеркивается, что никакая даже самая лучшая “теория практики” не приведет к успеху, если сама теория, результаты которой подлежат внедрению, недостаточно хороша. И все же, как образно выразился Акофф [24], теория практики и теория систем подобны противоположным полам, необходимым для производства потомства.

Следует подчеркнуть, что практика системных исследований и практика внедрения их результатов, хотя и имеют много общего, существенно различаются для систем разных типов. В этом отношении весьма продуктивной является классификация систем, введенная П. Чеклэндом [26]. Он делит системы на три типа. К первому относятся естественные системы, в которых связи образованы и действуют “природным образом” (физические, химические, биологические, экологические системы и т.п.). Ко второму типу принадлежат реальные искусственные системы, в которых связи образованы в результате человеческой деятельности, хотя эти связи имеют естественную природу (машины, приборы, другие технические системы и т.д.). К третьему типу систем относятся социотехнические системы (см. § 9.2). Самым важным является то, что в результате участия людей наиболее существенные связи в таких системах принадлежат не природе, а культуре (они условны, имеют знаковый характер), что смысл любой ситуации определяется отношением к ней субъекта. Хотя эти связи существуют лишь в мозгу человека (прежде всего в виде ценностей), его реальное, целенаправленное поведение делает их как бы реально существующими:

“Большинство реальных проблемных ситуаций относится именно к этому типу, как в масштабах мелких (например, как следует относиться к стареющим родителям?), так и в крупных (например, следует ли прекратить ядерное устрашение?)” [26, с. 58].

Системный анализ применяется для исследования систем всех трех типов. При этом содержательное, научное обоснование в первом случае дают естественные науки, во втором – естественные и технические, в третьем – связанные с данной проблематикой естественные, технические и гуманитарные науки, совокупность которых определяется природой исследуемой системы. Очевидно, что доля “рыхлых”, слабо структурированных проблем наиболее велика в исследованиях систем третьего типа. Ясно также, что наиболее сложна и практика их внедрения.

Свою “теорию практики” применительно к этому случаю Р. Акофф излагает [24], обобщая свой многолетний опыт системного анализа более чем для четырехсот частных и государственных организаций в различных странах.

НЕОБХОДИМОСТЬ МЕТОДОЛОГИИ ВНЕДРЕНИЯ

Сначала уточняются понятия. Социотехнические системы Акофф называет просто “социальными”, определяя их как системы, в которых люди играют наиболее существенную роль (например, это не водопровод, а водоканалтрест; не телефон, а управление телефонных сетей). Под “практикой” понимается работа на клиента, обладающего властью, достаточной для изменения системы теми способами, которые будут определены исследованием. Подчеркивается (на начальном этапе и неоднократно далее), что в этой работе обязательно должны непосредственно участвовать все заинтересованные стороны* или их представители. Поскольку предметом “теории практики” является эффективность практики, важно следующее определение: под “эффективной практикой” понимается улучшение работы организации клиента с точки зрения хотя бы одной из заинтересованных сторон и отсутствие ухудше-

* Напомним, что “заинтересованные стороны” (см. § 9.2) – это те, кто отвечает за решение проблемы, и те, кого эта проблема непосредственно касается.

IMPLEMENTATION внедрение (результатов)
CLIENT заказчик

DECISION-MAKER лицо, принимающее решение

ILL-DEFINED слабоструктурированный

WELL-DEFINED хорошо формализованный

С практической точки зрения системный анализ есть методология и практика улучшающего вмешательства в проблемные ситуации. Поскольку такое вмешательство затрагивает

ния этой работы с точки зрения всех остальных. (Следует обратить внимание на гуманистическую исходную позицию автора. Впоследствии именно она приведет Акоффа к проблеме этики системного анализа.)

Затем Акофф уточняет смысл термина “практикующий системный аналитик”:

“Не тот, кто, столкнувшись с проблематикой организации, собирает симптомы, ставит диагноз и выписывает рецепт, подобно врачу, имеющему дело с пациентом. Я имею в виду тех, кто, оказывая поддержку и помочь, делает других способнымиправляться со своими проблемами более успешно, чем они это могут сделать без его помощи. Таким образом, системный аналитик более похож на учителя, чем на доктора. Учителя знают, что они не могут научиться вместо своих учеников; ученики должны научиться сами. Но учителя могут помочь ученикам научиться большему и быстрее, чем это они могут без учительской помощи”.

Однако для Акоффа недостаточно аналогии с учителем. По его мнению, системная практика прежде всего является познавательным процессом не только для клиентов, но и для самих аналитиков:

“Если они ничему не учатся в ходе работы с другими, то они не практикуют, а консультируют, делясь тем, что они знали заранее. Скажем иначе: целью практики является поднять уровень развития системы клиента, всех ее заинтересованных сторон и самих аналитиков, вовлеченных в работу”.

Понятие развития играет в “теории практики” Акоффа важную роль, поэтому остановимся на нем подробнее. *Развитие определяется как увеличение желания и способности организации удовлетворять свои собственные и чужие нужды и оправданные желания.* Желания называются “оправданными”, если их удовлетворение ради одних не скажется отрицательно на развитии других. Нужды – это то, что необходимо для выживания. Возможны различные комбинации, например можно не хотеть нужного, можно желать ненужного.

РОСТ И РАЗВИТИЕ

В истории существования любой системы можно усмотреть периоды спада и подъема, стабильности и неустойчивости. Процессы, происходящие в системе, можно классифицировать по разным основаниям. Одна из полезных классификаций различает следующие типы динамики систем: деградация – спад – функционирование – рост – развитие.

Важно различать *развитие и рост*. Это далеко не одно и то же, и даже не обязательно одно связано с другим:

“Груда мусора может расти без развития. Человек развивается еще долго после того, как прекращается его рост. Это, конечно, очевидно. Менее очевидно, что многие проблемы, связанные с развитием, опираются на предположение, будто для развития экономический рост необходим, если не достаточен, и будто пределы роста ограничивают развитие”.

Развитие связано не столько с наличными ресурсами, сколько с умением их использовать. Оно больше зависит от информационных, чем от материальных, ресурсов. Недостаток ресурсов может ограничивать рост, но не развитие. По мере развития система все менее зависит от наличных ресурсов и все более способна добывать или производить недостающие ресурсы. Социальные системы могут чрезмерно вырасти, но предела их развитию не существует. Таким образом, Акофф связывает развитие с системностью, организованностью.

Развитие связано с обучением, а обучаться за других невозможно, поэтому невозможно и развивать другого человека или социосистему извне. Единственный способ развития –

саморазвитие. Тем не менее можно поощрить и поддержать развитие других, но это может быть сделано только при их участии. Поэтому Акофф считает абсолютно необходимым участие заинтересованных сторон в системной практике. Если некоторые из заинтересованных сторон представляют собой малодоступные или слишком многочисленные группы, то необходимо найти “хороших” представителей этих групп, что иногда не просто, но необходимо сделать. Лишь в крайнем случае аналитик должен брать на себя представительство отдельных заинтересованных сторон, и то лишь после специального изучения их интересов.

УСЛОВИЕ ДОБРОВОЛЬНОСТИ УЧАСТИЯ В АНАЛИЗЕ

Однако самой важной чертой участия заинтересованных сторон в системной практике Акофф считает его *добровольность* (если участие не является добровольным, то оно не может быть эффективным). По наблюдениям Акоффа, заинтересованные лица наиболее склонны участвовать в попытках эффективного разрешения проблем, когда они уверены, что выполнены следующие три условия:

- 1) их участие действительно повлияет на полученные результаты;
- 2) участие должно возбуждать интерес;
- 3) результаты действительно могут быть внедрены.

Первое условие наиболее полно реализуется, если каждая из сторон чувствует себя равноправной в принятии решений. Например, если решения принимаются большинством голосов и таким большинством обладает одна из заинтересованных сторон, то остальные стороны вряд ли добровольно будут участвовать в работе. Выход часто можно найти в консенсусе, т.е. принципе единогласного принятия решений (достичь которого часто трудно, но редко невозможно; см. § 7.5).

Второе условие добровольности участия (сделать его интересным) можно обеспечить несколькими способами. По опыту Акоффа эффективным приемом является предложение участникам анализа начать с “идеального перепроектирования” рассматриваемой системы, т.е. проектирования такой системы, которой, будь на то их воля, они немедленно заменили бы существующую. На такой “идеальный” проект накладывается не много ограничений; он должен в принципе быть реализуем (т.е. не чисто “научно-фантастическим”); удовлетворять основным внешним ограничениям (правовым, социальным, экономическим и т.п.), действующим в реальных условиях; предусматривать обучаемость и адаптивность системы при изменениях внешних условий. Составление такого проекта обычно вызывает оживленный интерес, поскольку содержит элементы творческой игры и забавы, но по мере его выполнения участникам становится все более ясно, что, во-первых, с наиболее важными проблемами и возможностями, стоящими перед организацией, они зачастую осознанно сталкиваются впервые, и, во-вторых, несмотря на не очень реалистичные предположения, спроектированная ими система обычно может быть реально аппроксимирована довольно близко.

РОЛЬ ОТНОШЕНИЙ МЕЖДУ УЧАСТНИКАМИ АНАЛИЗА

Третье условие (правдоподобность внедрения результатов) выполняется лишь тогда, когда лица, принимающие решения, проявляют готовность участвовать в системном исследовании и во внедрении его результатов. В отечественной литературе это требование известно как “принцип первого лица”. Акофф излагает его довольно категорично:

“Лично я не стану участвовать в проекте, в котором не хотят принять участия ответственные лица, и я не стану тратить на него больше времени, чем они. Проблемы их, а не мои; и если они считают, что эти проблемы не стоят их времени, почему я должен считать иначе?” [24, с. 11].

Конечно, дело здесь не в личных амбициях; как показывает опыт, вероятность внедрения результатов сильно зависит от персонального участия ответственных лиц в их получении. Кроме того, претворение ее результатов в жизнь зависит от многих обстоятельств, иногда неожиданных. Например, Акофф отмечает, что внедрение более вероятно, если за разработку рекомендаций было уплачено (“похоже, люди не очень ценят то, что достается бесплатно”); если окружение лица, принимающего решение (т.е. и его руководители, и подчиненные),

проявляет интерес к внедрению (этого можно добиться вовлечением их в “идеальное проектирование”). Но особо важное значение Р. Акофф придает отношениям взаимного доверия между ответственным лицом и системным аналитиком: он даже считает доверие абсолютно необходимым, если предстоит внедрение результатов. Руководитель нуждается в уверенности, что его собственные интересы не будут нарушены в процессе внедрения, и дружеские отношения с аналитиком дают такие гарантии. По наблюдениям Р. Акоффа, установление таких отношений облегчается при следующих условиях:

- 1) обе стороны могут отказаться от продолжения работы в любой момент и по любой причине; ни одна из них не обязана продолжать работу в случае неудовлетворенности ее ходом;
- 2) системный аналитик уделяет достаточное количество времени обучению персонала организации, с тем чтобы впоследствии работу организация могла выполнять и без него;
- 3) системный аналитик не стремится присвоить себе заслуги в получении положительных результатов, наоборот, он всячески подчеркивает заслуги других участников (это, кстати, быстрее приводит к росту его авторитета);
- 4) системный аналитик предъявляет профессиональные требования не только к качеству своей работы, но и к условиям, создаваемым ему для работы (в частности, если последние не отвечают профессиональным стандартам, например в отношении полного доступа к нужным лицам и информации, работа должна прекращаться);
- 5) системный аналитик должен открыто и искренне проявлять уважение к интеллекту ответственного лица, принимающего решения (разумеется, при этом исключается подхалимаж).

ПРОБЛЕМЫ И СПОСОБЫ ИХ РЕШЕНИЯ

Говоря о внедрении результатов системного анализа, важно подчеркнуть, что в реальной жизни очень редко, да и то лишь по отношению к простым системам, бывает так, что сначала проводят исследование, а затем его результаты внедряют в практику. При системном исследовании социальных и социотехнических систем они изменяются с течением времени как сами по себе, так и под влиянием самого исследования. Так, Риветт пишет:

“Бытовало мнение, что существуют четко определенные проблемы с ясно определенным набором целей и ограничений и что можно, так сказать, остановить мир, сойти с него, чтобы построить модель, а затем войти в него опять. Так было, и в некоторых случаях остается так, и именно в этих случаях формальное математическое моделирование является мощным и достаточным средством. Но существуют и другие случаи, в которых жизнь течет не останавливаясь и исследователь оказывается не на берегу, наблюдая, как река жизни течет мимо, а в хрупкой лодке, несущейся по быстрине” [31].

В процессе анализа изменяются состояние проблемы, цели, число и персональный состав участников, отношения между заинтересованными сторонами, а реализация принятых решений влияет на все факторы функционирования системы. Происходит фактическое слияние этапов исследования и внедрения, что придает системному анализу специфический характер: проблемы должны не “решаться”, а как бы “растворяться”, “исчезать” в ходе активного исследования.

Поясним эту мысль. Существует по крайней мере четыре способа обращения с любой проблемой реальной жизни. Первый (Акофф называет его *absolution*) – не решать проблему, надеясь, что она исчезнет сама собой. Второй (*resolution*) – сделать что-нибудь, частично решающее проблему, смягчающее ее до приемлемого состояния. Третий (*solution*) – решить проблему наилучшим в данных условиях (оптимальным) образом. Наконец, четвертый (*dissolution*) – ликвидировать, растворить проблему, изменив условия, произведя в системе и/или ее окружении такие изменения, чтобы не только исчезла сама проблема, но и будущие проблемы система могла бы преодолевать самостоятельно. Этот способ и реализуется в наиболее развитых формах системного анализа.

Пример [25]. В автобусной компании крупного города возникла проблема: после введения надбавок за качество работы водители начали конфликтовать с кондукторами. Дело в том, что качество работы водителей оценивалось по точности соблюдения графика движения, а кондукторов – по тому, насколько они успевают обслужить пассажиров. В часы пик кондукторы задерживали сигнал отправления (им надо было проверить у выходящих не только наличие билета, но и правильность оплаты, зависящей от расстояния, а входящим – продать билеты); это отрицательно сказывалось на надбавке водителей.

Сначала руководство компании игнорировало возникающую проблему (первый подход), но она не исчезла: враждебность стала нарастать, в конфликт были вовлечены профсоюзы. Тогда руководство попыталось вернуться к старой системе оплаты (второй подход); однако и водители, и кондукторы запротестовали, так как это уменьшило бы их зарплату. Затем руководство предложило делить причитающиеся надбавки поровну между водителями и кондукторами (третий подход); те отвергли предложение, поскольку они уже не желали сотрудничать друг с другом.

Проблема была “растворена” консультантом – системным аналитиком, обнаружившим, что в часы пик число автобусов на линии превышает число остановок. На эти часы кондукторов стали снимать с автобусов и закреплять по остановкам. Они продавали билеты до прихода автобуса, успевали проверять билеты у выходящих и стали вовремя давать сигнал отправления. По окончании часов пик кондукторы возвращались в автобусы, а лишние автобусы снимались с линий. К тому же компании потребовалось меньшее количество кондукторов.

РОЛЬ ЭТИКИ В СИСТЕМНОМ АНАЛИЗЕ

В заключение затронем еще один раздел “теории практики”, посвященный вопросам этики, т.е. морали и нравственности. Рассмотрение этих проблем неизбежно: не бывает этически нейтральных системных исследований. Дело не просто в том, что, как только заходит речь о субъективных целях и средствах их достижения, мы автоматически оказываемся причастными к проблемам этики (например, технократические и гуманистические парадигмы; см. § 9.3). В “теории практики” основной акцент делается на этике самих системных исследований.

С одной стороны, системные исследования имеют много общих черт с “обычными” научными исследованиями, в том числе и много общих этических норм: научная добросовестность, честность, объективность, преданность истине, требовательность к собственной компетентности, нормы общения с коллегами по профессии. С другой стороны, в системном анализе кроме истин фактических (“научных”) необходимо учитывать человеческие ценности, психологические моменты отношений между людьми, плохо изученные и далекие от формализации интересы конкретных людей. Это заметно усиливает значение этических аспектов в поведении системного аналитика.

Например, одна из опасностей (“ловушек”) в системном анализе состоит в навязывании системным аналитиком своего мнения лицу, принимающему решения. Этика поведения системного аналитика состоит [28] в том, чтобы не быть “серым кардиналом”, т.е.:

не скрывать альтернатив, которые почему-либо не нравятся ему самому; доводить и такие альтернативы до сведения лица, принимающего решения;

явно сообщать о предположениях, лежащих в основе полученных заключений;

обращать внимание лица, принимающего решения, на устойчивость или чувствительность альтернатив к изменениям условий.

В чем-то системный аналитик должен идти на компромиссы. Например, только для того, чтобы завоевать доверие заказчика, ему следует включать в модели детали, которые тот считает существенными, хотя сам аналитик придерживается противоположного мнения. Однако далеко не всегда компромиссы столь безболезненны и столь позитивны. Например, требуется определенная смелость, чтобы предложить ответственному лицу альтернативу, кото-

рая заведомо встретит его негативную реакцию, хотя анализ показывает ее высокую эффективность.

Системный аналитик оказывается перед этическим выбором, когда его принципы противоречат принципам заказчика. Дрор [28] формулирует несколько этических правил для системного аналитика в такой ситуации:

не работать на клиента, не дающего доступа к информации и не позволяющего публиковать результаты;

отказываться выполнять анализ только для обоснования уже принятого решения;

не работать на клиента, чьи цели и ценности противоречат гуманистическим целям и ценностям, а также собственным убеждениям аналитика.

Категоричность этих правил в процессе реализаций наталкивается на так называемые “сложности жизни”. Этика вообще дело не принудительное, а добровольное. Например, известный кибернетик С. Бир выполнял системные исследования проблем Чили по заказу правительства Альянде, но отказался работать по приглашению Пиночета, хотя ему после этого приходилось принимать меры личной безопасности. Однако сложности могут быть заложены даже не в персональных проблемах.

Исследования многих систем, обобщенные Черченом, показали, что реальные системы на деле служат не тем целям, ради которых они были созданы, а целям людей, работающих в данной системе. В качестве примеров приводятся отрасли промышленности, в работе которых интересы производителей стоят выше интересов потребителей; система торговли, в которой интересы продавцов выше интересов покупателей; можно привести аргументы за то, что в больницах интересы врачей часто выше интересов больных, а в вузах интересы преподавателей выше интересов студентов. Быть может, “эгоизм систем” столь же естествен и необходим, как и эгоизм личности? И вопрос состоит в том, как вести себя с системой, “эгоизм” которой выходит за “разумные рамки”.

Безоговорочный отказ работать на “неэтичную” систему не всегда оправдан. Бывают случаи, когда имеется надежда, что такая работа может изменить этику системы к лучшему.

Подведем итог

Системный анализ применяется при исследовании любых систем – природных, технических, социальных. Социальные системы наиболее трудны для исследования по многим причинам, одна из которых состоит в необходимости привлечения к анализу многих людей, даже тех, кто этого не желает. Включение в системный анализ этапа практического улучшения проблемной ситуации существенно увеличивает возникающие трудности. В последние годы большое внимание уделяется накоплению и обобщению опыта их преодоления. Один из вариантов представления этого опыта – “теория практики”, наибольший вклад в которую сделан

Summary

Systems analysis is applicable to the study of systems of all origins: natural, technical, and social. Social systems are the most difficult to study for many reasons, one of which is the necessity of involving many people, sometimes even against their will, in the analysis. The inclusion of a practical improvement of the problem situation into the analysis essentially increases the difficulties. In recent years great attention has been paid to the accumulation and generalization of the praxis of overcoming these difficulties. Ackoff's “theory of practice” is an example of such a generalization.

Р. Акоффом.

§ 9.8. О СПЕЦИФИКЕ СОЦИАЛЬНЫХ СИСТЕМ

Физика настолько успешно исследует и описывает явления неживой природы, что ее методы, модели и теории стали своеобразным эталоном того, как следует представлять научные знания. Однако попытки построить научные модели биологических, экономических и социальных систем по образу и подобию физических моделей показали, что очень немногие явления в этих системах поддаются адекватному математическому описанию.

Существуют разные причины того, что биологические и социальные явления принципиально качественно отличаются от явлений неживой природы.

НЕСВОДИМОСТЬ СОЦИАЛЬНЫХ ЗАКОНОВ К БИОЛОГИЧЕСКИМ И ФИЗИЧЕСКИМ

Прежде всего известна общесистемная закономерность, справедливая для систем любой природы: свойства системы не сводятся к свойствам ее частей и не выводятся из них. Система есть нечто большее, нежели простая совокупность составляющих ее частей, она обладает качественно новыми свойствами; в то же время у самих частей в системе проявляются такие качества, которые остаются скрытыми, не существующими вне системы. Это свойство систем в системологии называется *эмерджентностью*. Применительно к обсуждаемому вопросу свойство эмерджентности означает, что хотя каждая материальная часть любой сколь угодно сложной системы является физическим объектом и подчиняется всем законам физики, закономерности биологических и социальных систем не могут быть сведены к физическим закономерностям или выведены из них. Однако следует заметить, что данное соображение не является конструктивным: имея своеобразный нормативно-запретительный характер, оно не предлагает конкретных путей нахождения нового, более адекватного описания сложных систем.

Такие пути состоят в предметном, непосредственном изучении систем нового типа, в обнаружении природных отличий этих систем от чисто физических. По отношению к социальным системам таких отличий обнаружено достаточно много для того, чтобы осознать причины их качественно большей сложности.

В поведении системы определяющая роль принадлежит типу связей между элементами: совокупность этих связей образует структуру системы, а сама природа этих связей задает тип поведения системы и пределы ее возможностей. В физических системах, как природных, так и искусственных, связи существуют в виде естественных, задаваемых природой взаимодействий единообразно для всех физических систем: два электрона взаимодействуют одинаково в Африке и в Сибири, сейчас и всегда. В социальных системах все обстоит иначе. Механические взаимодействия не отменяются (и не могут быть отменены) и иногда играют важную роль в отношениях людей (например, в драке, спорте, совместной физической работе и других коллективных усилиях), но определяющую роль в обществе играют не они. На первый план выступают связи, имеющие конвенциональный, культуральный, условный, информационный характер. Эти связи имеют “нематериальную”, абстрактную природу; они существуют не в “первой” природе, а в сознании, в уме человека и в его культуре. Из того, что эти связи являются порождением воображения, не следует, что они не сказываются на физическом мире. Люди подчиняют свое поведение этим связям, и они становятся как бы реально существующими, и слова “как бы” можно вообще опустить: поведение человека в реальном мире оставляет вполне материальные следы. Заметим, что ту роль, которую играют объективные причины и причинно-следственные связи в физических системах, в социальных системах в значительной мере играют субъективные цели людей и их деятельность, направленная на осуществление своих целей.

Именно абстрактные информационные связи являются определяющими в социальных системах, задают их структуру. Ограничения, накладываемые этими связями на поведение элементов, имеют совершенно иную природу, чем подчиненность физическим законам. Вместо изучения физического мира, в который “погружены” все физические системы, мы должны

изучать культуру, в которой существует и которой обладает данная социальная система. При этом оказывается, что между двумя социальными объектами различия гораздо сильнее, чем между двумя физическими объектами: физическая природа едина, а культура у каждой социальной системы своя, да к тому же она хотя и консервативна, но и в какой-то степени изменчива.

СУЩЕСТВУЮТ ЛИ ИСТОРИЧЕСКИЕ ЗАКОНОМЕРНОСТИ?

Некоторые исследователи оценивают эту ситуацию весьма пессимистически, вообще отрицают существование таких утверждений, которые имели бы для истории ту же силу, которую имеют законы физики для неживой природы. Как выразил эту позицию Поппер, это означает, что мы должны отвергнуть возможность теоретической истории, так сказать, исторической науки об обществе, которая могла бы соответствовать теоретической физике. Не может быть научной теории исторического развития, которая служила бы основой для предсказания истории.

Принципиальные отличия социальных систем от технических проявляются также при попытках прогнозировать их поведение.

На один аспект сложности прогнозирования общественных явлений указывают противники “теоретической истории”. Вкратце их аргументация такова: события в социальных системах происходят под сильным влиянием роста человеческих знаний; будущее развитие знаний в принципе непредсказуемо, так как нельзя знать то, что еще не известно; поэтому невозможно предвидеть будущее социальных систем. Обсуждение этих доводов относится к компетенции историков; мы же обратим внимание на существование краткосрочных общественных явлений и проблем, не требующих прогнозирования слишком отдаленного будущего и допускающих использование любых современных научных достижений. Тем не менее остается фактом, что общественные науки заметно отстают от естественных наук по своей, так сказать, практической эффективности. Прикладной системный анализ способствует сокращению этого разрыва, в явной форме учитывая различия между системами разной природы.

Еще одна причина принципиального отличия в поведении социальных систем по сравнению с физическими состоит в следующем. В механической системе знание законов взаимодействия между ее частями и начальных условий позволяет рассчитать состояние системы в любой момент времени как в прошлом, так и в будущем. Система с участием человека (например, автомобиль) также подчиняется всем законам механики, но это не поможет предсказать все ее поведение, поскольку решение о том, тормозить или газовать, и в какую сторону поворачивать, принимает человек. Поведение социальной системы существенно определяется тем, какой выбор сделает каждый из людей, входящих в нее. Даже если точно зафиксировать начальные условия, и даже если лицо, принимающее решение, вполне определилось относительно своих будущих действий, у других людей, входящих в систему, всегда остается определенная свобода действовать по-разному в разные моменты времени (часто это даже не осознается), и в результате события перестают быть детерминированными.

Естествоиспытатели, обществоведы, философы и теологи давно размышляют и спорят о том, действительно ли живые существа имеют “свободу воли”, или это только им кажется, а на самом деле ход событий так или иначе (Бог или объективные законы) предопределен. Не вступая в этот непростой спор, отметим, что для понимания и верного описания того, что именно происходит в социальных системах, необходимо как можно глубже проникнуть в механизм принятия решений (выбора). Оказывается, что среди множества действий, которые образуют любую деятельность, акт выбора является единственной операцией, в которой проявляется, “овеществляется” цель того, кто осуществляет эту деятельность. Поэтому для прогноза социальных процессов очень важно знать действительные цели тех, кто принимает решения.

Выявление целей затруднено: часто люди затрудняются сформулировать свои действительные цели или заблуждаются относительно них, иногда некоторые цели сознательно

скрываются, не говоря уж о том, что цели могут изменяться со временем. Для объективизации выявления целей имеется две возможности. Первая – анализ актов выбора в прошлом: они содержат информацию об истинных целях принимающего решения. Вторая – выявление системы ценностей лица, принимающего решения: ведь цели субъекта вытекают из его ценностей.

Из таких и связанных с ними соображений постепенно сформировалась специфическая методология обращения (как изучения, так и внесения изменений) с социальными системами. Эта методология входит как часть в прикладной системный анализ, который предназначен для работы со сложными системами любой природы, и социосистемы образуют лишь один, хотя, возможно, наиболее трудный для работы, класс сложных систем.

“МЯГКАЯ” МЕТОДОЛОГИЯ В СИСТЕМНОМ АНАЛИЗЕ

Дадим краткий обзор современного состояния методологии, разрабатываемой в прикладном системном анализе специально для рассмотрения проблем, возникающих в социальных системах.

Исходным пунктом является признание того факта, что в отличие от технических систем, которые “есть только то, что они есть, и не могут быть ничем иным”, социальная система есть суперпозиция индивидуальных ее видений каждым участником социума. Вельтанштадт (видение мира) каждого индивида придает тому, что он видит вокруг себя, свой индивидуальный смысл, в соответствии с которым индивид и строит свое поведение. В результате, по выражению П. Чеклэнда, социальная система не задана априори, существует не в готовом, “данном виде”; это скорее процесс, в ходе которого непрерывно меняющееся сообщество постоянно воссоздается его членами.

Такой взгляд на социосистемы диктует остальные особенности методологии, начиная с самой постановки проблемы. В отличие от технических систем, проблемы которых – это то, что непосредственно подлежит решению, даже явно сформулированная проблема в социосистеме – это то, что еще подлежит изучению, предварительной структуризации. Иначе говоря, возможна “ошибка третьего рода” – можно начать решать “не ту” проблему.

Слабая структурированность социопроблем диктует резкое повышение требовательности к всесторонности, полноте, цельности и гармоничности системного исследования. Это находит выражение в специальных приемах.

Во-первых, конкретизируется, как следует выявлять “заинтересованные стороны”, т.е. всех участников проблемной ситуации. Обязательно включение в этот список как индивидов и организаций, страдающих в данной проблемной ситуации, так и, наоборот, получающих блага от нее; далее следует учсть всех, кто представляет собой активные силы в случае внесения изменений в существующую ситуацию; и, конечно, отдельно и специально нужно рассматривать организации и лиц, обладающих наибольшей властью среди сторон, вовлеченных в проблемную ситуацию (например, владельцев участвующих систем). Даже при таких, достаточно конкретных, рекомендациях эта работа не является простой. Из опыта системного анализа широкое признание получили рекомендации включать в число заинтересованных сторон также тех, кто по любым причинам не имеет возможности выразить свою причастность к проблемной ситуации, но окажется непосредственно задетым результатами ее изменения. К числу “нашумевших” примеров этого относятся: учет “интересов” маленькой рыбаки, обитающей в реке, на которой планировалось строительство плотины; включение в число участников проблемной ситуации прошлых и будущих поколений человечества. Рассмотрение интересов будущих поколений на таких реальных примерах, как проблема “что делать с радиоактивными отходами?” и т. п., привело У. Черчмена к идею о возможности интерпретации общественной морали как формы соблюдения интересов будущих поколений. В статье “Философия сложности” (в кн. “Future Research”, H. Linstone, W. Simmonds (Eds.), Addison – Wesley, London, 1977, pp. 82 – 90) он пишет:

“Что все “холистические” исследователи делают действительно полезного сегодня в нашем мире, так это то, что они заставляют нас уделять внимание будущим поколениям, делая это

нашей моральной обязанностью. И это очень серьезно. Когда-то в молодости меня мучил вопрос “Что такое мораль?” Я пришел к выводу, что морально то, что будущие поколения попросили бы нас делать, если бы они могли к нам обратиться. Я думаю, что голос моральной критики сегодня и является голосом будущих поколений, поскольку многое из того, что мы замышляем сегодня, будет иметь последствия для будущих поколений. Возьмем производство ядерной энергии и его отходы. Сейчас мы не знаем, что делать с этими отходами. Захоронение их в шахтах уже не представляется хорошим решением. Некоторые считают, что их надо вывезти в Антарктику и закопать в снег. Есть и такое предложение: раз уж мы не знаем, что делать с отходами, почему бы не забросить их на орбиту и предоставить какому-нибудь сообразительному инженеру будущего позаботиться о них. Пусть они ломают голову, что делать с нашими отходами! Это является вполне реальным использованием космоса. Я считаю это явно аморальным. Это аморально потому, что будущие поколения скорее всего попросили бы нас не делать этого, будь у них возможность говорить с нами. Может быть нам и придется сделать это, но по крайней мере мы должны признать, что при этом мы поступим аморально”.

Заканчивая обсуждение этапа выявления участников проблемной ситуации, подчеркнем, что сложность его выполнения связана с необходимостью глубокого изучения данного социума, так как многие из существенных связей в нем далеко не очевидны, а иногда и тщательно скрываются.

Второй прием, способствующий полноте восприятия и описания социосистемы, состоит в том, чтобы строить и использовать не какую-то одну, а одновременно несколько ее моделей, различающихся качественно, и даже противоречивых. Это оказывается эффективным и для понимания природы некоторых чисто физических объектов (вспомним корпускулярную и волновую теории света; квантовую механику), а при изучении социообъектов это является просто обязательным.

Множественность совместно рассматриваемых моделей “встроена” как в методику внешнего, так и внутреннего описания проблемной ситуации на всех уровнях рассмотрения. На самом верхнем уровне, как бы определяя, с каких точек зрения мы будем смотреть на все в ходе системного исследования, мы должны выбрать конфигуратор, т.е. набор качественно различных языков описания. Фактически выбор языка для включения его в конфигуратор определяет характер моделей, которые будут строиться средствами этого языка. Сейчас вряд ли возможно дать более четкие рекомендации по выбору конфигуратора, кроме того, что набор языков должен быть “необходимым и достаточным” для целей анализа.

Множественность моделей рекомендуется и на следующих уровнях описания социального явления. В качестве наглядного примера можно привести рассмотрение Дж. Эллисоном Кубинского ракетного кризиса. В своем исследовании он строил модели трех типов:

“квази-технологическая” модель (нация рассматривается как единое целое, выдвигаются альтернативные решения, целевые критерии для их оценки, поиск наилучшей альтернативы и т. д.);

модель “организационного процесса” (рассматриваются вовлеченные в ситуацию организации, учитываются их обычные процедуры действия, их цели и т. д.);

модель “бюрократической политики” (взгляд на проблему с точки зрения личностей руководителей, каждого со своими интересами, мнениями, стилем поведения в конфликтах и переговорах).

Как видим, в данном случае были выбраны модели, описывающие ситуацию на разных уровнях: государство в целом – отдельные государственные структуры – конкретные политические лидеры.

Подобная множественность моделей носит “сквозной” характер, присутствует на всех этапах и уровнях системного анализа. Возникает вопрос: можно ли предложить полный список типов моделей, придерживаясь которого в каждом исследовании, было бы реально расчитывать на требуемую полноту?

П. Чекленд предлагает в качестве такого набора моделей список, мнемоническая аббревиатура которого в английском оригинале выглядит как CATWOE. Буквам С, А и О соответствуют упомянутым выше обязательным членам списка заинтересованных сторон: страдающим и выигрывающим от проблемной ситуации (Customers), активным силам в функционировании системы и в устраниении проблемы (Actors), владельцам проблемосодержащей и проблеморазрешающей систем (Owners). В описании должна быть дана также модель окружающей среды (Environment) – прежде всего в виде ограничений, накладываемых на систему извне (правовых, экономических и пр.) Далее предлагается давать достаточно полное описание процессов преобразования (Transformation) входов системы в ее выходы, перечисляемые конкретно. Наконец, требуется дать ту картину мира (Weltanschauung), которая придает смысл функционированию исследуемой системы; это, в частности, может потребовать изучения систем ценностей определяющих лиц в системе. Например, важно определить, чьим интересам более всего предано лицо, принимающее решение: человечества в целом, собственного народа, членов своей организации, своего ближайшего окружения, своей семьи или только своим собственным? Ведь от этого будет кардинальным образом зависеть характер принимаемых им решений.

Как видим, CATWOE представляет собой структурную схему исследуемой системы. Правда, для реализации этого потенциала требуется соблюдать дополнительные требования. К их числу относится требование следования законам диалектики. Например, если в рассматриваемой модели отсутствуют противоречивые элементы, то модель может оказаться неполной, поскольку в реальном (моделируемом) объекте такие элементы обязательно присутствуют (это положение является одним из законов диалектики). Более того, в полном соответствии с диалектикой будут одновременно использованы противоречащие друг другу модели одной и той же системы!

Особенностью всей методологии системного анализа является преобладание синтетического подхода над аналитическим. Эта специфика существенно усиливается в той части методологии, которая направлена на работу с социальными системами. Здесь приходится быть особенно внимательным относительно воспитанного в нас, как бы само собой разумеющегося, предположения, что отдельная часть системы остается тем же, чем она была и в составе системы. Дисциплинарная структура наших знаний образует как бы “интеллектуальные полочки”, по которым следует раскладывать получаемые новые знания. В противоположность этому системное мышление имеет холистический характер, система постоянно рассматривается как нечто единое, целое. Бесперспективно решать любую проблему в отрыве от остальных: например, политика в любой области – будь то энергетика, связь, транспорт, оборона, сельское хозяйство – в реальности оказывается в сильной зависимости от политики в области образования. Невнимание к проблемам образования может свести на нет любые усилия по решению проблем в других областях.

СОГЛАСИЕ ПРИ РАЗНОГЛАСИЯХ

Один из источников значительных трудностей в работе с социальными системами – наличие в реальных системах различных, часто противоположных, несовместимых взглядов на мир, что и является причиной всевозможных конфликтов. Проявления этой трудности весьма разнообразны; рассмотрим некоторые из них.

Возьмем извечную проблему – как обеспечить целостность социосистемы при наличии разногласий у ее участников. Исторический опыт учит, что применение силы, принесение интересов меньшинства в жертву интересам большинства приводят к тому, что жертвой может стать каждый. Это относится и к демократическим формам выработки коллективных решений (см. § 7.5). Выясняется, что суть демократии не в принятии решений большинством голосов, а в защите интересов меньшинства, более того – в соблюдении интересов каждой отдельной личности. Интересы личности становятся столь же главенствующими, основными, сколь и сохранение социосистемы в целом. И дело не только в том, что поступиться одним – значит поступиться всеми. Не менее важно то, что без осознания своего внутреннего “я”

человек не способен осознать и общественные проблемы. Мысли эти не новы. Еще 200 лет назад Кант писал: “Две вещи наполняют мое сердце вечным благоговением: звездное небо над головой и моральный закон внутри человека”.

Тем не менее в конфликтных ситуациях, каковыми нередко оказываются проблемные ситуации, вопрос о совместном существовании противоположных взглядов переходит в практическую плоскость: какие же действия по изменению ситуации допустимы, а какие – нет? Высокоморальное требование Акоффа – считать улучшающими только такие изменения, которые улучшают ситуацию с точки зрения хотя бы одного ее участника и не ухудшают с точки зрения всех остальных – казалось бы, должно часто приводить к невозможности внесения каких-либо изменений вообще. Сам Акофф предлагает прибегать к консенсусу, достичь которого, по его словам, “часто трудно, но редко невозможно”. Известны случаи системного анализа, успешно решавшего проблему через консенсус. Если же консенсус невозможен или достигается на условиях, не позволяющих снять проблему, то можно вносить допустимые изменения, не решающие, а отодвигающие проблему и в то же время создающие новую ситуацию. При этом достигается какой-то выигрыш, что-то улучшается, ощущается движение вперед, накапливаются опыт и знания. Проблему, которую нельзя решить, мы вынуждены считать не “проблемой”, а “условием”, в рамках которого приходится действовать до тех пор, пока мы не окажемся в состоянии изменить его.

Таким образом, в отличие от технических проблем социальные проблемы, как правило, не “решаются окончательно”, а работа над ними является процессом непрерывного изучения и самообучения, который организуется с помощью системных понятий и методик. Системная методология не является сводом жестких инструкций; хотя в ней есть определенные правила, она оставляет место для проявления личного стиля и творчества.

УЧИТЬСЯ ВОЗДЕЙСТВОВАТЬ БУДУЩЕМУ

Одна из “ловушек” на пути такого постепенного решения проблем, сдвига их в будущее – явление, получившее название “недооценка будущего” (discount of futures). Оно состоит в том, что временная перспектива подобна пространственной: чем дальше от нас событие, тем меньший вес мы ему склонны придавать. Мы сейчас говорим о том, что запасы нефти иссякнут через 50 – 100 лет. Представьте, насколько изменился бы тон этих разговоров, если бы речь шла о 10 – 15 месяцах! На проблемы будущих поколений мы глядим словно через уменьшительное стекло.

Здесь дело не только в проявлении эгоизма (хотя и это имеет место). Существуют и объективные причины этого. Решение одних проблем порождает массу других, в результате чего в модели будущего приходится рассматривать интенсивно ветвящиеся деревья вариантов. (Схематический пример для иллюстрации. Решение проблемы безработицы потребует целой системы изменений. Одна из мер – снижение пенсионного возраста. Кроме позитивного эффекта, эта мера породит и ряд проблем: нагрузку на экономику; в совокупности с улучшением здравоохранения и удлинением средней продолжительности жизни эта нагрузка значительно возрастет; психологический дискомфорт, ощущение своей “ненужности” у пенсионеров и т.д.) Экспоненциальный рост числа возможных будущих проблем приводит к невозможности учета их всех, и мы просто вынуждены сокращать это число, отсекая какие-то варианты, упрощая задачу. Вместе с тем мы вносим в свои прогнозы и выводы погрешности и ошибки, последствия которых неизвестны.

Так как сокращение перебора является неизбежным, вынужденным шагом, то остается принять меры к снижению потерь от выполнения этого шага. Это значит – сделать сокращение перебора целенаправленным, осмысленным, по возможности менее субъективным и эгоистичным. Практика системного анализа дает по крайней мере два способа для этого. Первый состоит в рассмотрении (насколько удастся) утрачиваемых возможностей, связанных с отсекаемой ветвью, и их сопоставление с остающимися возможностями – так сказать, вариант “нечисловой оптимизации”. Второй способ – усиление учета интересов будущих поколений. При этом Черчен, например, рекомендует приписывать нарастающие веса по мере от-

даления поколений. “Мои дети в моей жизни значат намного больше меня самого, а дети их детей еще более важны, и т.д. Ценность каждого следующего поколения постоянно нарастает, и учет этого позволит как-то скомпенсировать эффект недооценки будущего”.

НЕОЖИДАННОСТЬ КАК СЛЕДСТВИЕ СЛОЖНОСТИ

Однако, как бы ни отрицали это одни и как бы ни старались уменьшить это другие, субъективность нельзя полностью исключить из любого научного исследования, а тем более из исследования социальной системы. Многие системщики отмечают, что “правильность” взятой проблемы и найденного решения зависит от того, кто формирует проблему, от времени, от организационной ситуации, от других факторов окружающей среды.

В частности, здесь уместно еще раз вернуться к вопросу о том, что, например, хирургическое вмешательство или наказание виновных является жизненными примерами явного отхода от определения “улучшающего вмешательства” по Акоффу.

Завершая обсуждение особенностей применения системного анализа к социальным системам, отметим еще раз, что это класс систем особой сложности: в каждой такой системе всегда есть нечто неизвестное, нечто не отраженное в сегодняшних ее моделях, нечто неожиданное. Как пример открытия подобных неожиданностей, можно привести интересное исследование проблемы голода на Земле, проведенное в 80-х годах в ИИАСА (“Hunger amidst abundance” ПАСА, 1987).

Один из этапов этого системного исследования состоял в имитационном моделировании мировой экономики по сценарию “Манна небесная”. Этот сценарий основан на том, что на Земле недоедает около 490 млн. людей, живущих в основном в слаборазвитых странах. Им нехватает 50 млн. тонн зерна в год, т.е. около 3 % мировых поставок зерна (первая удивительная цифра). Было сделано гипотетическое предположение, что некая фантастическая сторона (например, инопланетяне) из гуманистических побуждений решила, начиная с 1982 г., продавать на рынке за любую цену 50 млн. тонн зерна в год. Далее на модели мирового рынка прослеживалась динамика цен, спроса, предложения и производства зерна. Первоначальный спад цен (и производства) с минимумом в 1986 – 87 гг. сменяется их ростом, и после некоторых колебаний устанавливается к 2000 году практически на прежнем уровне. Неожиданным оказалось то, что рыночные отношения “спружинили” прибавку хлеба на рынке, частично снизив производство зерна и распределив остальное на кормовые нужды и в запасные фонды. Лишь незначительная часть маны небесной дошла до голодающих (их число снизилось всего на 3 – 6 %!). Вместо этого рынок изменил структуру производства и экспорта зерна, адаптировав ее к новым ценам.

Таким образом, обнаруживается, что проблема голода – это не просто проблема увеличения производства или улучшения хранения продуктов питания, а прежде всего проблема распределения, т.е. проблема политическая.

Подведем итог

Системный анализ применим для рассмотрения любых проблем, но применение его методологии должно согласовываться с природой конкретной системы. Наибольшую сложность для исследования представляют социальные системы, специфика которых вытекает из превалирования в них культуральных, информационных отношений, из превалирования субъективных при-

Summary

Systems analysis is feasible in treatment of any problem; however, its methodology, when applied, demands matching it to a nature of a system considered. Social systems are most complicated for the analysis, due to prevalence of cultural, information relations in them, to pre-dominance of subjective reasons (goals, duties, traditions) over objective causes. This is why so many special methods, designed

чин (целей, долга, традиций) and adaptable for the case, were над объективными. Поэтому elaborated in systems analysis. системный анализ социальных проблем включает много при-емов, созданных специально для этого случая.

ЗАКЛЮЧЕНИЕ

Если попытаться охарактеризовать современный системный анализ еще раз, очень укрупненно и несколько в ином ракурсе, то можно сказать, что он включает такие виды деятельности, как:

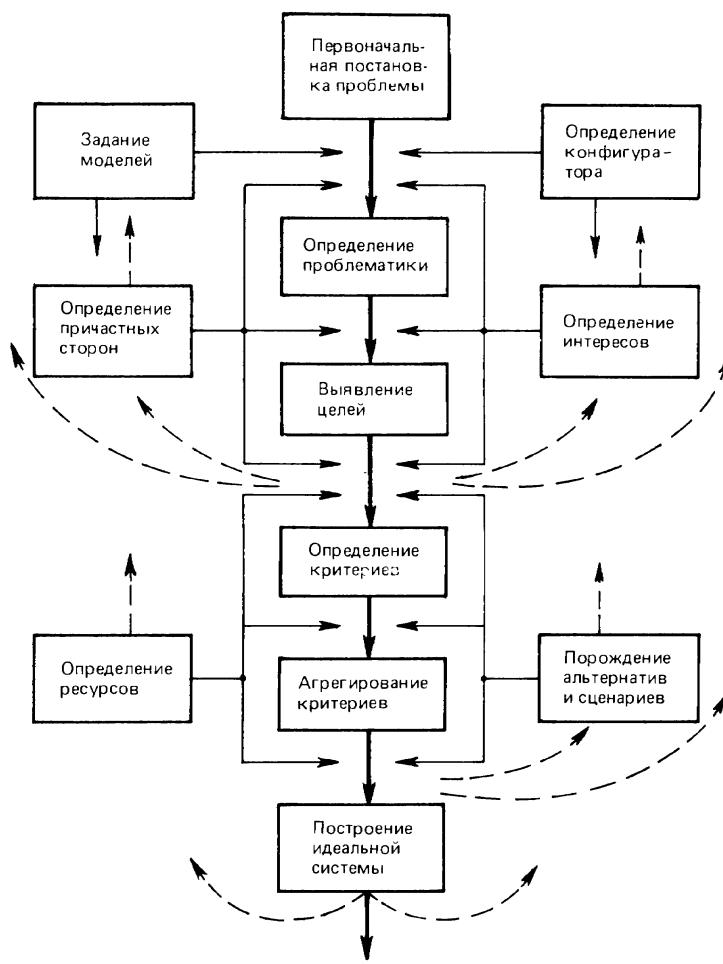
научное исследование (теоретическое и экспериментальное) вопросов, связанных с проблемой;

проектирование новых систем и изменений в существующих системах;

внедрение в практику результатов, полученных в ходе анализа.

Уже сам этот перечень, очевидно, лишает смысла спор о том, чего в системном исследовании больше – теории или практики, науки или искусства, творчества или ремесла, эвристики или алгоритмичности, философии или математики – это все в нем присутствует. Конечно, в конкретном исследовании соотношения между этими компонентами могут быть самыми различными. Системный аналитик готов привлечь к решению проблемы любые необходимые для этого знания и методы – даже те, которыми он сам лично не владеет; в этом случае он не исполнитель, а организатор исследования, носитель цели и методологии всего исследования. Жизнь разнообразна, и предлагаемые для исследования проблемы не всегда требуют использования всего арсенала системного анализа. Из трех типов систем (технических, природных и социотехнических) наибольшую трудность для анализа представляют последние из-за резкого преобладания в них субъективного над объективным, эвристического над формальным, знаковых отношений над физическими взаимодействиями. Однако еще более важным отличием социосистем является особое значение временного фактора: эти системы меняются в ходе исследования как сами по себе, так и под влиянием самого анализа. Только диалектический подход, лежащий в основе системного анализа, помогает создать динамическую модель текущих событий и с ее помощью спланировать и организовать действия всех участников анализа. Не следует обвинять в ненаучности сам системный анализ, если в жизни встречаются (а для социосистем нередко!) случаи, когда уже сбор и обработка информации вполне удовлетворяют заказчика или когда грубые, но быстрые исследования его устроят больше, чем глубокие, подробные и длительные.

Еще раз остановимся на проблеме алгоритмизации системного анализа. Любой процесс исследования, проектирования и целевого воздействия алгоритмичен: алгоритм является планом этого процесса. Составление



9.1—————Опорная схема постановки задач прикладного системного исследования реальной проблемы

такого плана – прерогатива системного аналитика. Для каждой проблемы может потребоваться особый, специально для нее приспособленный алгоритм анализа. Возвращаясь к аналогии с программированием для ЭВМ, можно сказать, что, подобно тому как программа составляется из операторов языка применительно к решаемой задаче, операции системного анализа реализуются в последовательности, удобной для аналитика применительно к данной конкретной ситуации. Чем выше квалификация аналитика, тем более разнообразны проводимые им исследования.

Несмотря на отсутствие универсального алгоритма, студенты и начинающие практику аналитики нуждаются в более определенных рекомендациях, в типовых схемах, которые в дальнейшем можно варьировать. К рекомендациям, подобным тем, которые были приведены в табл. 8.1, добавим блок-схему неформальных этапов первой стадии системного анализа (рис. 9.1). На схеме сплошными жирными стрелками изображена опорная последовательность действий, пунктирные стрелки обозначают возможность возврата к уже пройденным действиям, если это окажется необходимым. Нижний выход ведет к формализуемым этапам анализа.

Подводя окончательный итог, еще раз попытаемся дать определение системного анализа в его современном понимании. Нельзя сказать, что прикладной системный анализ в сегодняшнем состоянии вполне отвечает этому определению. Скорее, данное определение отражает направление развития прикладного системного анализа, которое осознанно осуществляется в последние годы.

Итак: *с практической стороны системный анализ есть теория и практика улучшающего вмешательства в проблемные ситуации; с методологической стороны системный анализ есть прикладная диалектика.*

ЛИТЕРАТУРА

1. Адамар Ж. Исследование психологии процесса изобретения в области математики. – М.: Сов. радио, 1970.
2. Альтшулер Г.С. Найти идею. – Новосибирск: Наука, 1986.
3. Бир С. Кибернетика и управление производством. – М.: Наука, 1965.
4. Бриллюэн Л. Наука и теория информации. – М.: Физматиз. 1960.
5. Буш Г.Я. Основы эвристики для изобретателей. – Рига: Зинатне, 1977. Ч. 1,2.
6. Гуд Г.Х., Макол Р.Э. Системотехника. Введение в проектирование больших систем. – М.: Сов. радио, 1962.
7. Джонс Дж. К. Методы проектирования. – М.: Мир, 1986.
8. Зинченко В.П. Человеческий интеллект и технократическое мышление // Коммунист. 1988. № 3. С. 96 – 104.
9. Квейд Э. Анализ сложных систем. – М.: Сов. радио, 1969.
10. Моррисей Дж. Целевое управление организаций. – М.: Сов. радио, 1979.
11. Мюллер И. Эвристические методы в инженерных разработках. – М.: Радио и связь, 1984.
12. Неуймин Я.Г. Модели в науке и технике. – Л.: Наука, 1984.
13. Одрин М.В., Карташов С.С. Морфологический анализ систем. – Киев: Наукова думка, 1977.
14. Орлов Н.А. Методологические основы конструирования. – Русе (НРБ): ВТУ, 1986.
15. Тельц Д., Эндрюс Ф. Ученые в организациях. Об оптимальных условиях для исследований и разработок. – М.: Прогресс, 1973.
16. Перегудов Ф.И. и др. Основы системного подхода. – Томск: ТГУ, 1976.
17. Перегудов Ф.И. Основы системного проектирования АСУ организационными комплексами. – Томск: ТГУ, 1984.
18. Перегудов Ф.И. и др. Системное проектирование АСУ хозяйством области. – М.: Статистика, 1977.
19. Пойа Д. Математика и правдоподобные рассуждения. – М.: ИЛ, 1957.

20. Половинкин А.И. Основы инженерного творчества. – М.: Машиностроение, 1988.
21. Пригожин И., Стенгерс И. Порядок из хаоса – М.: Прогресс. 1986.
22. Саами Т. Математические методы исследования операций. – М.: МО, 1963.
23. Холл А. Опыт методологии для системотехники. – М.: Сов. радио, 1975.
24. Ackoff R.L. A Theory of Practice in the Social Systems Sciences. Paper to an International Roundtable IIASA.– Laxenburg, Austria, 6 – 8, Nov., 1986.
25. Ackoff R.L. The mismatch between educational systems and requirements for successful management. // Wharton Alumni Magazine.– Spring, 1986. P. 10 – 12.
26. Checkland P. Rethinking a System Approach. In: Tomlison R., Kiss I. (Eds.) “Rethinking the Process of Operation Research and System Analysis”. – Pergamon Press, 1984. P. 43 – 66.
27. Cotgrove S. Catastrophe or Cornucopia: the Environment, Politics and the Future.– Wiley, Chichester, 1982.
28. Dror Y. Design for Policy Sciences.– N. – Y.: American Elsevier, 1971.
29. Gharajedaghi J, Ackoff R.L. Toward Systemic Education of System Scientists// System Research. 1985. Vol. 2. N 1. P. 21 – 27.
30. Mitroff I.I. Why our pictures of the world do not work anymore. In: Lawler et al. (Eds). “Doing Research that is Useful for Theory and Practice”.– San Francisco: Tossey – Boss, 1985.
31. Rivett P. Perfection of Means. Confusion of Goals. Paper to IIASA Roundtable, 1986.

УПРАЖНЕНИЯ

§ 9.1

- Тема для обсуждения: соотношение в системном анализе науки, искусства и ремесла; теории и практики; строгих рассуждений, эвристики и эксперимента.

§ 9.2

- Возьмите самую простую известную вам проблему и попробуйте построить проблематику в соответствии с данными рекомендациями.

§ 9.3

- Обсудите трудности выявления цели на конкретных примерах из вашей практики.
- Тема для обсуждения: рассмотрение допустимости эвтаназии (умерщвления из милосердия) с позиций различных систем ценностей.

§ 9.4

- Тема для обсуждения: критерии как модели цели.

§ 9.5

- Заготовьте несколько экземпляров кроссворда и раздайте их нескольким одиночкам и группам по 2 – 7 человек. Назначьте группу наблюдателей, установите лимит времени и соберите кроссворды по истечении этого периода. Из полученных данных извлеките информацию о факторах, влияющих на групповое творчество (размер группы; время, отпущенное на выполнение задачи; роль индивидов в группе; появление лидеров; взаимосвязи в группе и т.д.).

§ 9.6

- Попытайтесь составить свой вариант алгоритма системного анализа для какой-нибудь обозримой, очень несложной проблемы.

§ 9.7

- Обсудите пример с автобусной компанией. Видите ли вы другие варианты решения проблемы? Сравните их с приведенными.

- Тема для дискуссии: этика проводящего системный анализ.

Вопросы для самопроверки

1. Каковы основания для того, чтобы назвать системный анализ прикладной диалектикой?
2. Почему при исследовании реальной проблемы неизбежны неформализованные этапы?
3. Почему любую проблему не следует рассматривать изолированно, вне связи с другими проблемами и явлениями?
4. В чем различие между “рыхлой” и “жесткой” проблемами?
5. Каковы основные трудности выявления целей? Почему после каждого последующего этапа системного анализа следует возвращаться к проверке, уточнению и пересмотру целей?
6. Каково соотношение целей и критериев для оценки альтернатив?
7. В чем состоят главные особенности метода мозгового штурма?
8. Какова основная идея синектики?
9. Какое описание системы необходимо для морфологического анализа ее альтернатив?
10. Чем отличаются развитие и рост социосистем?
11. Почему исследовательский и внедренческий этапы системного анализа не могут быть разделены?
12. Каково значение этики в системном анализе?

Краткий словарь специальных терминов

В конце определений приводятся номера страниц книги, где встречаются данные термины. Поэтому “Словарь” может использоваться и как предметный указатель. Кроме того, даются английские эквиваленты терминов.

АВТОМАТ (*automatic machine*, от греч. “самодействующий”) – устройство, самостоятельно выполняющее некий процесс по заложенной в него программе. Программа может фиксироваться либо непосредственно в устройстве автомата, либо на вводимом в автомат носителе 11, 13, 92, 299

АВТОМАТИЗАЦИЯ (*automation*) – внедрение автоматов в практическую деятельность (например, автоматизация управления, нефтедобычи, медицинской диагностики, погрузочных работ и т.п.) 10–12, 192, 198, 266

АГРЕГАТ (*aggregate*, от лат. “присоединять”) – любая выделенная совокупность, от неструктурированной (множество, конгломерат) до высокоорганизованной системы 298, 301–303, 305–309

АГРЕГИРОВАНИЕ (*aggregation*) – 1) операция образования агрегата; 2) преобразование многомерной модели в модель меньшей размерности 277, 278, 280, 298–302, 305–309, 311, 344

АДЕКАВАТНОСТЬ (*adequacy*) – свойство модели; модель считается адекватной, если с ее помощью успешно достигается поставленная цель Адекватность не обязательно совпадает с истинностью: только в познавательных моделях эти понятия сливаются 19, 50, 53, 54, 143, 171, 242, 258, 280, 304, 307, 308, 321, 330, 331, 354, 355

АЛГОРИТМ (*algorithm*, от имени узбекского математика IX в. Аль–Хорезми) – 1) полное описание последовательности действий, выполнение которых в конце последовательности приводит к достижению цели; 2) конечный текст, записанный на алгоритмическом языке. Первоначально сугубо математическое понятие алгоритма в настоящее время расширено: допускается включение в алгоритм и указаний на неформализуемые действия, лишь бы они правильно понимались и выполнялись людьми (например, алгоритм изобретения). Алгоритм, воспринимаемый и исполняемый автоматом, называется программой 8–13, 244–246, 291–294, 342–345, 366–368

АЛГОРИТМИЗАЦИЯ (*algorithmization*) – 1) составление алгоритма для проектируемого процесса (например, алгоритмизация решения задачи); 2) выявление алгоритма, формализация существующего процесса (обычно в целях его изучения, совершенствования или автоматизации) 9, 57–59, 289, 342, 366

АЛЬТЕРНАТИВА (*alternative*) – вариант, одна из двух или более возможностей; то, что можно иметь, использовать и т.д. вместо чего–то еще. На множестве альтернатив осуществляется выбор 210–219, 223–235, 237–243, 326–329, 333–342

АНАЛИЗ (*analysis*, от греч. “расчленение”) – 1) мысленное или реальное разделение целого на части (например, химический анализ вещества, декомпозиция глобальной цели и т.д.); 2) до недавнего времени – синоним научного исследования вообще (“подвергнуть анализу” означало “изучить”); 3) метод познания, основанный на 1). Познание не сводится к анализу; только в сочетании, переплетении, единстве с синтезом становится возможным познание реальности 15–17, 242–244, 277–279, 281–287, 289–293, 317–320, 322–325

ВХОД (СИСТЕМЫ) (*input*) – 1) связь системы с окружающей средой, направленная от среды в систему, т.е. выражающая воздействия из среды на систему; 2) то, что преобразуется системой в выход 70, 73–78, 89–91, 286, 287, 360

ВЫХОД (СИСТЕМЫ) (*output*) – 1) связь системы с окружающей средой, выражающая воздействие системы на среду и направленная от системы к среде; 2) продукт системы; то, во что преобразуются входы; может иметь как реальный характер (например, материальная продукция), так и абстрактный (например, удовлетворение потребности) 70, 73–78, 89–91, 291, 368

ВЫБОР (*choice*) – 1) операция, входящая во всякую целенаправленную деятельность и состоящая в целевом сужении множества альтернатив (обычно, если позволяют условия,— до одной альтернативы); 2) принятие решения 206, 209–226, 228–255, 259–261, 263–269, 324–326

ГИПОТЕЗА (*hypothesis*) – 1) предположение; утверждение, требующее доказательства или проверки; 2) форма развития науки 56, 246

ГОЛОСОВАНИЕ (*vote*) – способ выражения коллективного мнения, вынесения коллективного решения, осуществления коллективного выбора с помощью одного из мажоритарных правил (простого или абсолютного большинства, консенсуса и т.п.) 218, 231–234, 236, 250

ГРАНИЦА СИСТЕМЫ (*boundary of a system*) – 1) поверхность в пространстве описания ситуации, разделяющая саму систему и окружающую ее среду, – в данной (целевой) модели системы, т.е. понятие относительное; 2) пределы, до которых распространяется и в которых исполняется управляющая информация системы 73, 80, 201, 217, 218

ГРАФ (*graph*, от греч. “записывать”) – 1) символическая диаграмма, состоящая из множества вершин и ребер (дуг), соединяющих некоторые из них (или все); 2) графическая модель структуры 85–87, 89, 92, 222–225, 228, 310

ДЕКОМПОЗИЦИЯ (*decomposition*) – 1) операция разделения целого на части с сохранением признака подчиненности, принадлежности; 2) повторное или многократное такое разделение, в результате чего получаются древовидные иерархические структуры 280–285, 289–298

ЗНАК (*sign*) – 1) сигнал, имеющий конкретное (обычно одиночное, элементарное) значение, воспринимаемое человеком (например, дорожный знак, знаки отличия, математические знаки, условные жесты и т.д.); 2) реальная модель абстрактного понятия 46, 50, 346, 366

ИЕРАРХИЯ (*hierarchy*) – 1) принцип организации, состоящий в том, что целое рассматривается как состоящее из частей, каждая из которых сама является целым, состоящим из своих частей, и т. д.; 2) многоуровневая древовидная структура с отношениями подчиненности сверху вниз. В реальных системах встречаются различные отступления от идеальной иерархической структуры: связь только с одним элементом нижнего уровня (“синекура”); связь более чем с одним элементом высшего уровня (“двойное подчинение”); связь с высшим уровнем помимо непосредственного верхнего (“дислокация”); более одного элемента на самом верху (“незавершенность”); связь сверху вниз с элементами разных уровней (“неоднородность”); связь между элементами одного уровня (“зависимость”); связь с окружающей средой помимо верхнего уровня (“нарушение субординации”), и их комбинации 36, 37, 42, 62, 79

ИЗБЫТОЧНОСТЬ (*redundancy*) – свойство сигналов и систем, обеспечивающее их устойчивость против разрушительного воздействия помех, шумов, отказов элементов, непредвиденных обстоятельств и т.п. Это свойство состоит во включении в структуру системы или сигнала большего числа элементов, чем это минимально необходимо при отсутствии помех. Если критерий эффективности связан с минимизацией числа элементов, то при отсутствии помех избыточность является излишней (отсюда – ее название), а при наличии помех – полезной, но возникает задача ее минимизации 155–160, 164, 202, 203, 256, 303

ИЗМЕРЕНИЕ (*measurement*) – действие по сопоставлению определенного состояния наблюдаемого явления или объекта с выбранной для регистрации этого состояния шкалой; результатом измерения является символ (принадлежащий выбранной шкале), обозначающий наблюдавшееся состояние (см. Шкалы измерительные) 123, 170–173, 175–177, 198–201

ИНГЕРЕНТНОСТЬ (*inherence*) – 1) согласованность модели с окружающей ее культурной средой; принадлежность модели этой среде; 2) условие, необходимое для проявления, реализации модельных свойств модели 47, 48, 50, 60, 62, 64, 69

ИНТЕЛЛЕКТ (*intellect*, от лат. “разум”, “рассудок”) – 1) интеллект *естественный*; внутренне – способность к абстракции; внешне – способность ориентироваться в незнакомых условиях и находить решение слабо формализованных задач; 2) интеллект *искусственный* – техническая имитация определенных возможностей естественного интеллекта (например, узнавания, образования понятий, принятия решений, синтеза речевых и эстетических сигналов и т.п.) 10, 12, 13

ИНФОРМАТИКА (*informatics, computer science*) – 1) наука о научной и технической информации и ее циркуляции в обществе; 2) в последние годы в России информатикой называют научное направление, акцентирующее внимание на использовании ЭВМ в самых разнообразных областях человеческой деятельности 14, 29, 46, 164

ИНФОРМАЦИЯ (*information*) – 1) в обыденной речи – любые сведения, известия, сообщения, новости и т.п.; 2) в научно-технических приложениях – то, что несет на себе сигнал; 3) как философская категория – всеобщее свойство материи, являющееся аспектом свойства отражения, допускающим количественное описание 123–175

КИБЕРНЕТИКА (*cybernetics*, от греч. “управлять”) – в широком смысле – наука об управлении в системах произвольной природы; наиболее полными и общими определениями кибернетики в современном понимании признаются определения А.И. Берга и А.Н. Колмогорова 12–15, 23, 26–30, 124

КЛАССИФИКАЦИЯ (*classification*) – 1) операция отнесения заданного объекта к одному из классов, внутри которых объекты считаются неразличимыми; результат этой операции; 2) простейший вид моделирования; в частности, самый слабый вид измерения 104–116, 200–203, 304–306, 311, 317, 346, 348

КОД (*code*) – совокупность условий и правил образования сигнала, использование которых на передающем и на приемном концах позволяет передавать и получать информацию с помощью сигнала 45, 47, 115, 125, 126, 156–160, 163, 186

КОЛИЧЕСТВО ИНФОРМАЦИИ (*quantity of information*) – числовая мера информации, содержащейся в одном случайном объекте о другом случайном объекте. Определяется как некий функционал от соответствующих распределений вероятностей (например, по Шэннону, по Фишеру, по Кульбаку – Лейблеру и др.), либо как объем вычислений, необходимых для алгоритмического определения состояния объекта (по Колмогорову) 66, 116, 149, 151–157, 161, 164

КОНФИГУРАТОР (*configurator* от англ. “формирователь”) – набор различных языков описания изучаемой системы, достаточный для проведения системного анализа данной проблемы. Определяется природой проблемосодержащей и проблеморазрешающей систем с целью анализа 301–304, 307–309, 311, 313, 321–323, 326, 329, 344, 360

КРИТЕРИЙ (*criterion*) – 1) средство для вынесения суждения; стандарт для сравнения; правило для оценки; 2) мера степени близости к цели; в этом смысле – модель цели 210–221, 225, 227, 230, 232, 239, 240, 251–258, 263, 327–334

МОДЕЛЬ (*model*, от лат. “образец”) – отображение: целевое; абстрактное или реальное, статическое или динамическое; ингерентное; конечное, упрощенное, приближенное; имеющее наряду с безусловно истинным условно-истинное, предположительно-истинное и ложное содержание; реализующееся и развивающееся в процессе его практического использования 28, 35–98, 280–292

- МОДЕЛЬ АБСТРАКТНАЯ (*abstract model*) – идеальная конструкция; модель, построенная средствами мышления, сознания (в частности – языковая модель) 36, 41, 42, 44–48, 50, 51, 56, 58, 64, 70, 71, 82, 103, 280, 285, 302, 309
- МОДЕЛЬ ДИНАМИЧЕСКАЯ (*dynamic model*) – модель, отображающая процессы, происходящие в системе со временем; в частности, модели функционирования и развития 40, 41, 43, 64, 87–90, 92–95, 366
- МОДЕЛЬ ЗНАКОВАЯ (*signary model*) – реальная модель, имеющая абстрактное содержание; модель, условно подобная оригиналу и предназначенная для непосредственного использования человеком 45, 46, 50
- МОДЕЛЬ ИНГЕРЕНТНАЯ (*inherent model*) – модель, согласованная с окружающей культурной средой, входящая в нее не как чуждый ей элемент, а как ее естественная часть 62, 64, 69
- МОДЕЛЬ КЛАССИФИКАЦИОННАЯ (*classificatory model*) – простейший вид модели, в которой фиксируются только отношения тождественности или различия 201, 202
- МОДЕЛЬ МАТЕМАТИЧЕСКАЯ (*mathematical model*) – абстрактная или знаковая модель, построенная средствами математики (например, в виде системы уравнений, графа, логической формулы и т.п.) 36, 43, 64, 88, 90–93, 127, 128, 130, 195, 242, 258
- МОДЕЛЬ МОДЕЛЕЙ (*model of models*) – иерархия моделей; многоуровневая абстракция; число уровней в иерархии моделей моделей предположительно связывается с развитостью интеллекта 51, 62
- МОДЕЛЬ ПОЗНАВАТЕЛЬНАЯ (*cognitive model*) – форма организации и представления знаний; средство соединения новых знаний с имеющимися 39, 40, 56, 309
- МОДЕЛЬ ПРАГМАТИЧЕСКАЯ (*pragmatic model*) – средство управления, организации практических действий; образец, эталон правильных действий (например, алгоритм) или их результата (например, модель цели) 39, 40, 56, 309, 343
- МОДЕЛЬ РЕАЛЬНАЯ (ВЕЩЕСТВЕННАЯ, ФИЗИЧЕСКАЯ, ПРЕДМЕТНАЯ) (*substantial model*) – модель,строенная из реальных объектов; подобие реальной модели и оригинала может быть прямым, косвенным и условным 36, 41, 43–46, 50, 58, 64
- МОДЕЛЬ СОСТАВА СИСТЕМЫ (*partition model*) – модель, описывающая, из каких подсистем и элементов состоит система 78–81, 89, 95
- МОДЕЛЬ СТАТИЧЕСКАЯ (*static model*) – модель, в которой отсутствует временной параметр 40, 64, 88, 89, 92, 94, 285
- МОДЕЛЬ СТРУКТУРЫ СИСТЕМЫ (*structural model*) – модель, описывающая все отношения (связи) между элементами модели состава системы 80–82, 94, 299, 301
- МОДЕЛЬ ФУНКЦИОНАЛЬНАЯ (*model of functioning*) – модель, описывающая процессы, которые характеризуют систему как часть более общей, охватывающей ее системы, т.е. связаны с назначением данной системы 41
- МОДЕЛЬ “ЧЕРНОГО ЯЩИКА” (*black-box-model*) – модель, описывающая только входы и выходы системы, но не внутреннее устройство системы. Например, математическая модель “черного ящика” – это просто совокупность множеств X и Y (X соответствует входам, Y – выходам); если оператор F , связывающий их ($Y = F(X)$), и предполагается существующим, то он считается неизвестным 70, 72–74, 76–78, 80
- МОДЕЛЬ ЯЗЫКОВАЯ (*conversational, or linguistic model*) – любая конструкция на естественном языке, рассматриваемая как описание чего-либо (например, определение как модель определяемого; имя, название как обозначение называемого и т.д.) 35, 41, 42, 44, 62, 69, 255

МОЗГОВОЙ ШТУРМ (*brain-storming*) – метод, предназначенный для неформального коллективного генерирования возможно большего числа альтернатив; основные идеи этого метода: а) полное запрещение критики на стадии генерирования; б) поощрение и провоцирование ассоциативного мышления на всех стадиях; в) на стадии оценки цель состоит не в отбрасывании “плохой” альтернативы, а в поиске рационального зерна в ней 60, 260, 336–339

МОРФОЛОГИЧЕСКИЙ АНАЛИЗ (*morphological analysis*) – формальный метод генерирования альтернатив с помощью перечисления всех возможных сочетаний значений заданных параметров альтернативы 306, 338, 340, 341

НАДСИСТЕМА (*wider system, containing system*) – система, содержащая в своем составе данную систему 283, 285, 321–323, 326

НЕОПРЕДЕЛЕННОСТЬ (*uncertainty*) – неоднозначность любого происхождения в описании системы 196–199, 237–280

НЕОПРЕДЕЛЕННОСТЬ РАСПЛЫВЧАТАЯ (*fuzziness*) – неопределенность, связанная с нарушением аксиом тождественности – неоднозначностью классификации. Описывается с помощью функции принадлежности; характерна для языковых моделей, но возможна в любых шкалах 42, 57, 170, 191, 195, 197, 204, 251–255, 265, 306

НЕОПРЕДЕЛЕННОСТЬ СТОХАСТИЧЕСКАЯ (*randomness*) – неопределенность, описываемая распределением вероятностей на множестве возможных состояний рассматриваемого объекта; случайность 57, 86, 264

НЕОПРЕДЕЛЕННОСТЬ ЧАСТОТНО–ВРЕМЕННАЯ (*time-and-frequency uncertainty*) – свойство функций времени, состоящее в невозможности неограниченного уменьшения произведения их длительности на ширину их спектра; существует лишь некий минимальный предел этого произведения, которого можно достичь выбором специальной формы сигнала 137, 139

НЕТРАНЗИТИВНОСТЬ ГОЛОСОВАНИЯ (*vote nontransitivity*) – одно из свойств коллективного выбора, состоящее в существовании набора альтернатив, на котором выбор единственной альтернативы методом голосования сделан быть не может; в частности, результат выбора может зависеть от последовательности предъявления альтернатив 234–236, 250

ОКРУЖАЮЩАЯ СРЕДА (*environment*) – то, что находится вне границ системы и взаимодействует с нею. Структурированность окружающей среды может выражаться с различной степенью подробности; как минимум – в виде входов и выходов системы 70–76, 327, 332

ОПТИМАЛЬНЫЙ (*optimal*) – наилучший в заданных условиях. Качество оценивается с помощью критерия оптимальности, а условия задаются в виде ограничений на дополнительные критерии. Оптимизация – центральная идея кибернетики 215, 255–258

ПОДСИСТЕМА (*sub-system, contained system*) – система, содержащаяся целиком в данной системе. Различают подсистемы разных уровней (подподсистемы, или подсистемы такого-то уровня) 48, 62, 79, 80, 89, 94, 104, 257, 259, 281, 283, 285, 321–323, 326

ПРИНЯТИЕ РЕШЕНИЯ (*decision making*) – целевой выбор на множестве альтернатив. Методы принятия решений разнообразны в зависимости от типа неопределенности и других условий выбора 159, 200, 209, 210, 231–233, 251, 263, 264, 267, 284, 290, 339, 349, 357, 362

ПРОБЛЕМА (*problem*, от греч. “задача”) – 1) проблема развития – неудовлетворительное состояние системы, изменение которого к лучшему является непростым делом; 2) проблема функционирования – удовлетворительное состояние системы, сохранение которого требует постоянных и непростых усилий (например, проблема выживания) 10, 12, 55–58, 70, 72, 312, 318–326, 350–352, 358–368

ПРОБЛЕМАТИКА (*mess*) – сплетение, клубок проблем, которые неразрывно связаны с проблемой, подлежащей разрешению. Необходимость рассмотрения проблематики вместо отдельной проблемы вытекает из того, что проблемосодержащая система сама состоит из подсистем и входит в надсистему, а устранение поставленной проблемы требует учета последствий для всех них 320–323

ПРОБЛЕМНАЯ СИТУАЦИЯ (*problem situation*) – такая ситуация, когда неудовлетворительность существующего положения осознана, но неясно, что следует сделать для его изменения 70, 321, 331, 332, 346, 348, 358–360, 362, 368

ПРОПУСКНАЯ СПОСОБНОСТЬ КАНАЛА СВЯЗИ (*channel capacity*) – максимальная скорость передачи информации по каналу, при которой еще возможна передача без потери информации, т.е. при сколь угодно малой вероятности ошибок 56, 155–158, 160–162, 164, 256

РАЗМЫТОЕ (РАСПЛЫВЧАТОЕ) МНОЖЕСТВО (*fuzzy set*) – множество, содержащее хотя бы один такой элемент, о котором нельзя однозначно сказать, принадлежит ли он или нет этому множеству (математическая модель расплывчатой неопределенности). Степень уверенности выражается функцией принадлежности, принимающей значения из интервала [0, 1] 193, 195, 251, 255, 333

РАНГ (*rank*) – 1) номер некоторого объекта в упорядоченном по некоторому признаку ряду объектов; 2) элемент порядковой (ранговой) шкалы 53, 86, 176, 178, 179, 184, 187, 217, 235, 261, 271, 326

РЕГУЛИРОВАНИЕ (*regulation, adjustment*) – способ управления с обратной связью, основанный на обнаружении ухода объекта с программной траектории и выработке регулирующего воздействия для возвращения объекта на эту траекторию 11, 111, 124

РЕСУРСЫ (*resources*) – 1) наличные или требуемые средства для реализации алгоритма достижения цели; сюда входят не только ресурсы, необходимые для функционирования управляемой системы, но и ресурсы, необходимые для выработки управления, т.е. затрачиваемые на актуализацию моделей, используемых в управляющей системе; 2) в философском смысле исчерпывающей квалификацией ресурсов является их деление на материальные, энергетические и информационные; однако в зависимости от целей возможны более подробные классификации. Например, материальные ресурсы иногда дробят на вещество и время (скажем, объем памяти ЭВМ и машинное время); в оргсистемах часто выделяют кадровые и финансовые ресурсы и т.д.

112–115, 165, 287, 317, 324, 344, 348

СВОЙСТВО (*property*) – 1) качество, постоянно присущее объекту; 2) абстракция отношения данного объекта с другими, “свернутое отношение”, модель отношения 83, 84, 90, 92, 93, 100, 101, 114, 116, 118, 123–128

СЕМАНТИКА (*semantics*, от греч. “обозначение”) – раздел семиотики, изучающий отношения между знаками и тем, что они обозначают 46

СЕМИОТИКА (*semiotics*, от греч. “знак”) – наука, исследующая знаки и знаковые системы 46, 164

СИГНАЛ (*signal*, от лат. “знак”) – материальный носитель информации 7, 45, 47, 75, 81, 124–143, 149, 152, 153, 155–162, 189, 198, 246, 250

СИНЭКТИКА (*synectics*) – метод генерирования альтернатив, основанный на догадках по ассоциации, возникающих в группе экспертов, специально подготовленных для поиска аналогий, в особенности аналогий двигательным ощущениям 60, 337–339

СИНТЕЗ (*synthesis*, от греч. “соединение”) – 1) мысленное или реальное соединение частей в единое целое; 2) метод познания, основанный на 1). Познание является единением, сочетанием анализа и синтеза 15–17, 21, 22, 28, 192, 243–245, 247, 277–279, 303, 304, 308, 309, 312

СИСТЕМА (*system*) – средство достижения цели; основные особенности систем: целостность, относительная обособленность от окружающей среды, наличие связей со средой, наличие частей и связей между ними (структурированность), подчиненность всей организации системы некоторой цели

СИСТЕМА БОЛЬШАЯ (*large-scale system*) – система, для актуализации модели которой в целях управления недостает материальных ресурсов (времени, емкости памяти, других материальных средств моделирования) 112, 114–117, 119

СИСТЕМА ЕСТЕСТВЕННАЯ (*natural system*) – система (т.е. многокомпонентный объект, обладающий всеми признаками системы), возникшая в природе в результате естественных процессов 18, 99, 101, 102, 104, 105, 118, 123, 346

СИСТЕМА ИСКУССТВЕННАЯ (*artificial, man-made system*) – система, созданная человеком как средство достижения поставленной цели 99, 101, 102, 105, 118, 125, 346

СИСТЕМА ПРОБЛЕМОРАЗРЕШАЮЩАЯ (*problem-solving system*) – система, обладающая возможностями (ресурсами, компетенцией и пр.), необходимыми для ликвидации проблемы; обязательный участник системного анализа; проблеморазрешающая и проблемосодержащая системы могут не совпадать 278, 285, 288, 332, 333, 360

СИСТЕМА ПРОБЛЕМОСДЕРЖАЩАЯ (*problem-containing system*) – система, в которой возникла проблема, подлежащая решению; обычно эта система является инициатором и заказчиком проведения системного анализа 278, 285, 320–323, 326, 332, 333, 360

СИСТЕМА СЛОЖНАЯ (*complex system*) – система, модель которой, используемая для управления системой, неадекватна заданной цели 26, 28, 102, 112, 114–117, 257–259, 317, 319, 320, 333, 345, 355, 358

СИСТЕМА СОЦИОТЕХНИЧЕСКАЯ (*socio-technical system*) – система, в составе которой имеются люди и коллективы, интересы которых существенно связаны с функционированием системы 317, 319, 322, 339, 346, 347, 351, 366

СИСТЕМА ЦЕННОСТЕЙ (*values system*) – идеологическая основа для постановки целей социотехнических систем; объект системного анализа на этапе выявления действительных целей лиц, причастных к решаемой проблеме 308, 322, 325, 326, 329, 357, 361

СИСТЕМНОСТЬ (*systematicity*) – 1) обладание всеми признаками системы; 2) всеобщее свойство материи, форма ее существования, а следовательно, неотъемлемое свойство человеческой практики, включая мышление 7–12, 14–19, 21–28, 35, 38, 64, 103, 118, 209, 312, 318, 343

СИСТЕМНЫЙ АНАЛИЗ (*systems analysis*) – 1) с практической стороны системный анализ есть система методов исследования или проектирования сложных систем, поиска, планирования и реализации изменений, предназначенных для ликвидации проблем; 2) с методологической стороны системный анализ является прикладной диалектикой, так как реализует идеи материалистической диалектики применительно к конкретным практическим задачам, особенность которых состоит в необходимости выяснения причин их сложности и устранения этих причин; 3) с методической стороны системный анализ отличается междисциплинарным и наддисциплинарным характером и вовлечением в работу как неформальных, эвристических, экспертных методов, так и эмпирических, экспериментальных методов, а также при возможности и необходимости – строгих формальных математических методов 277–279, 317–320, 322–326, 328, 331, 341–348, 351–354

СИСТЕМНЫЙ ПОДХОД (*systems approach*) – в настоящее время рассматривается либо как одна из ранних форм системного анализа, либо как начальная фаза современного системного анализа, этап первоначального, качественного анализа проблемы и постановки задач 7, 69, 118, 312

СЛОЖНОСТЬ (*complexity*) – свойство некоторого явления (объекта, процесса, системы), выражающееся в неожиданности, непредсказуемости, необъяснимости, случайности, “антиинтуитивности” его поведения 11, 24, 70, 114–117, 212, 218, 259, 260, 263, 292, 293, 296, 359

СТРУКТУРА (*structure*, от лат. “строение”) – совокупность связей между частями системы 80–82, 84, 86–90

СТРУКТУРНАЯ СХЕМА СИСТЕМЫ (*structural scheme*) – конструкция системы; объединение моделей “черного ящика”, состава и структуры 84, 85, 87, 88, 92, 94, 285, 361

СЦЕНАРИЙ (*scenario*) – воображаемая, но правдоподобная последовательность действий и вытекающих из них событий, которые могут произойти в будущем с исследуемой системой; модель будущего после принятия решения, представленная до его принятия 284, 339, 340, 364

УПРАВЛЕНИЕ (*control*) – целевое воздействие на системы (т.е. воздействие, приближающее к целевому состоянию) 109–114

ФУНКЦИЯ ВЫБОРА (*choice function*) – наиболее общая математическая модель выбора; отображение совокупности множеств в совокупность их подмножеств без поэлементного отображения одного множества на другое и без отображения множеств на числовую ось 228–232

ФУНКЦИЯ КРИТЕРИАЛЬНАЯ (*criterion function*) – функция от обозначения альтернатив, значения которой упорядочены в том же порядке, что и предпочтения альтернатив 212, 221, 227, 252, 253, 307, 328

ФУНКЦИЯ ПОТЕРЬ (*loss function*) – функция, выражающая потери, которые вынужден нести пользователь статистического решения, отличающегося от истинного суждения 238, 243, 245

ФУНКЦИЯ ПРИНАДЛЕЖНОСТИ (*membership function*) – функция, характеризующая расплывчатое множество и принимающая для каждой альтернативы значение из интервала [0, 1], выражающее степень принадлежности данного элемента этому расплывчатому множеству 57, 193–196, 252–254

ФУНКЦИЯ РЕШАЮЩАЯ (*decision function*) – функция, отображающая каждую выборку в пространство статистических решений 244

ЦЕЛЬ (*goal, end, purpose*) – 1) образ желаемого будущего (субъективная цель); 2) будущее реальное состояние (объективная цель) 8, 23, 37, 38, 41, 70, 71, 100–103, 209, 218, 220, 280, 285, 303, 306, 313, 357

ШКАЛЫ ИЗМЕРИТЕЛЬНЫЕ (*measurement scales*) – множество обозначений, используемых для регистрации состояния наблюдаемого объекта; в зависимости от введенных отношений на этом множестве, шкалы различаются по их силе; сила измерительной шкалы должна согласовываться с природой наблюдаемого явления 172, 176, 182, 185–187, 189, 191, 195, 251

ЭКСПЕРТНЫЕ МЕТОДЫ (*expert methods*) – методы системного анализа, в которых для выполнения тех или иных неформализуемых операций используются знания, опыт, интуиция, изобретательность, интеллект экспертов, специалистов в нужной области 259

ЭМЕРДЖЕНТНОСТЬ (*emergence*, от англ. “внезапное появление”) – 1) особенность систем, состоящая в том, что свойства системы не сводятся к совокупности свойств частей, из которых она состоит, и не выводятся из них; 2) внутренняя целостность систем 20, 298–302, 313

ЭНТРОПИЯ (*entropy*, от греч. “превращение”) – мера неопределенности случайного объекта 143–146, 151–153, 156–159, 161, 164, 197

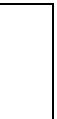
C O N T E N T S

Foreword **3**

Introduction **5**

ORIGIN AND DEVELOPMENT OF SYSTEMS

CONCEPTS 7

 § 1.1. Preliminary remarks **7**

§ 1.2. Role of systems ideas in practice **8**

§ 1.3. An intrinsic systematicity of cognition processes

14

§ 1.4. Systematicity as general property of matter **17**

§ 1.5. A brief outline of history of systems concepts **21**

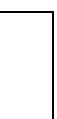
Conclusion **32**

Bibliography **33**

Exercises **33**

Test questions **34**

MODELS AND MODELING 35

 § 2.1. A broad interpretation of the concept of model **35**

§ 2.2. Modeling as a nonseparable stage of any
purposeful activity **37**

§ 2.3. Modes of model realization **41**

§ 2.4. Conditions for realization of model properties **47**

§ 2.5. Correspondence between a model and reality:
distinctions **49**

§ 2.6. Correspondence between a model and reality:
similarities **54**

§ 2.7. On evolution of models **57**

Conclusion **62**

Bibliography **65**

Exercises **65**

Test questions **68**

О ГЛАВЛЕНИЕ

Предисловие **3**

Введение **5**

ВОЗНИКНОВЕНИЕ И РАЗВИТИЕ СИСТЕМНЫХ ПРЕДСТАВЛЕНИЙ 7

-  § 1.1. Предварительные замечания **7**
§ 1.2. Роль системных представлений в практической деятельности **8**
§ 1.3. Внутренняя системность познавательных процессов **14**
§ 1.4. Системность как всеобщее свойство материи **17**
§ 1.5. Краткий очерк истории развития системных представлений **21**
Заключение **32**
Литература **33**
Упражнения **33**
Вопросы для самопроверки **34**
-

МОДЕЛИ И МОДЕЛИРОВАНИЕ 35

-  § 2.1. Широкое толкование понятия модели **35**
§ 2.2. Моделирование – неотъемлемый этап всякой целенаправленной деятельности **37**
§ 2.3. Способы воплощения моделей **41**
§ 2.4. Условия реализации свойств моделей **47**
§ 2.5. Соответствие между моделью и действительностью: различия **49**
§ 2.6. Соответствие между моделью и действительностью: сходство **54**
§ 2.7. О динамике моделей **57**
Заключение **62**
Литература **64**
Упражнения **65**
Вопросы для самопроверки **68**
-

SYSTEMS. MODELS OF SYSTEMS 69

- § 3.1. Multiplicity of models of systems **69**
 - § 3.2. The first definition of a system **69**
 - § 3.3. A “black box” model **72**
 - § 3.4. A partition model **78**
 - § 3.5. A structure model **81**
 - § 3.6. The second definition of a system. Structural scheme of the system **84**
 - § 3.7. Dynamical models of systems **87**
 - Conclusion **94**
 - Bibliography **96**
 - Exercises **97**
 - Test questions **98**
-

ARTIFICIAL AND NATURAL SYSTEMS 99

- § 4.1. The man – made systems and the natural objects **99**
 - § 4.2. Artificial and natural systems **101**
 - § 4.3. On classifications of systems **104**
 - § 4.4. Large and complex systems **112**
 - Conclusion **118**
 - Bibliography **120**
 - Exercises **120**
 - Test questions **122**
-

INFORMATION APPROACH TO STUDYING SYSTEMS 123

- § 5.1. Information as a property of matter **123**
- § 5.2. Signals in systems **124**
- § 5.3. Random process as a mathematical model of signals **127**
- § 5.4. Mathematical models of realizations of random processes **130**
- § 5.5. Some properties of continuous signals **137**
- § 5.6. Entropy **143**
- § 5.7. Quantity of information **149**
- § 5.8. Main results of information theory **154**
- Conclusion **163**
- Bibliography **165**
- Exercises **165**
- Test questions **168**

ROLE OF MEASUREMENTS IN DESIGN OF MODELS OF SYSTEMS 169

- § 6.1. Experiment and model **169**
§ 6.2. Measuring scales **172**
§ 6.3. Fuzzy description of situations **191**
§ 6.4. Probabilistic description of situations. Statistical measurements **196**
§ 6.5. Experimental data registering and its connection with succeeding data processing **200**
Conclusion **206**
Bibliography **206**
Exercises **207**
Test questions **208**
-

СИСТЕМЫ. МОДЕЛИ СИСТЕМ 69

- § 3.1. Множественность моделей систем **69**
§ 3.2. Первое определение системы **69**
§ 3.3. Модель “черного ящика” **72**
§ 3.4. Модель состава системы **78**
§ 3.5. Модель структуры системы **81**
§ 3.6. Второе определение системы. Структурная схема системы **84**
§ 3.7. Динамические модели систем **87**
Заключение **94**
Литература **96**
Упражнения **97**
Вопросы для самопроверки **98**
-

ИСКУССТВЕННЫЕ И ЕСТЕСТВЕННЫЕ СИСТЕМЫ 99

- § 4.1. Искусственные системы и естественные объекты **99**
§ 4.2. Обобщение понятия системы. Искусственные и естественные системы **101**
§ 4.3. Различные классификации систем **104**
§ 4.4. О больших и сложных системах **112**
Заключение **118**
Литература **120**
Упражнения **120**
Вопросы для самопроверки **122**
-

ИНФОРМАЦИОННЫЕ АСПЕКТЫ ИЗУЧЕНИЯ СИСТЕМ 123

- § 5.1. Информация как свойство материи **123**
§ 5.2. Сигналы в системах **124**
§ 5.3. Случайный процесс – математическая модель
сигналов **127**
§ 5.4. Математические модели реализаций
случайных процессов **130**
§ 5.5. О некоторых свойствах непрерывных сигналов
137
§ 5.6. Энтропия **143**
§ 5.7. Количество информации **149**
§ 5.8. Об основных результатах теории информации
154
Заключение **163**
Литература **165**
Упражнения **165**
Вопросы для самопроверки **168**
-

РОЛЬ ИЗМЕРЕНИЙ В СОЗДАНИИ МОДЕЛЕЙ СИСТЕМ 169

- § 6.1. Эксперимент и модель **169**
§ 6.2. Измерительные шкалы **172**
§ 6.3. Расплывчатое описание ситуаций **191**
§ 6.4. Вероятностное описание ситуаций.
 Статистические измерения **196**
§ 6.5. Регистрация экспериментальных данных и ее
связь с последующей их обработкой **200**
Заключение **206**
Литература **206**
Упражнения **207**
Вопросы для самопроверки **208**
-

CHOICE (DECISION MAKING) 209

- § 7.1. Multiplicity of choice problem statements **209**
§ 7.2. The criterial language for description of choice
211
§ 7.3. Description of choice in the binary relations
language **221**
§ 7.4. Language of choice functions **228**
§ 7.5. Group choice **231**
§ 7.6. Choice under uncertainty **237**
§ 7.7. Choice under statistical uncertainty **241**
§ 7.8. Choice under fuzzy uncertainty **251**
§ 7.9. Merits and demerits of the optimality idea **255**
§ 7.10. Expert methods of choice **259**
§ 7.11. Man – machine systems and choice **263**
§ 7.12. Choice and selection **267**
Conclusion **272**
Bibliography **273**
Exercises **275**
Test questions **276**
-

DECOMPOSITION AND AGGREGATION AS PROCEDURES OF SYSTEMS ANALYSIS 277

- § 8.1. Analysis and synthesis in systems studies **277**
§ 8.2. Models of systems as a decomposition basis **281**
§ 8.3. Algorithmization of the decomposition process
289
§ 8.4. Aggregation, emergence, intrinsic wholeness of
systems **298**
§ 8.5. Types of aggregates **301**
Conclusion **312**
Bibliography **314**
Exercises **315**
Test questions **316**
-

ON NON-FORMALIZABLE STAGES OF SYSTEMS ANALYSIS 317

- § 9.1. What is a systems analysis? **317**
§ 9.2. Defining a problem **319**
§ 9.3. Discovering of purposes **324**
§ 9.4. Formulation of criteria **328**
§ 9.5. Generating alternatives **334**
§ 9.6. On algorithms of systems analysis **342**
§ 9.7. Implementation of the systems analysis results
345
§ 9.8. On peculiarities of social systems **354**
Conclusion **366**
Bibliography **368**
Exercises **369**
Test questions **371**
-

Glossary **373**

ВЫБОР (ПРИНЯТИЕ РЕШЕНИЙ) 209

- § 7.1. Многообразие задач выбора **209**
§ 7.2. Критериальный язык описания выбора **211**
§ 7.3. Описание выбора на языке бинарных
отношений **221**
§ 7.4. Язык функций выбора **228**
§ 7.5. Групповой выбор **231**
§ 7.6. Выбор в условиях неопределенности **237**
§ 7.7. О выборе в условиях статистической
неопределенности **241**
§ 7.8. Выбор при расплывчатой неопределенности
251
§ 7.9. Достоинства и недостатки идеи
оптимальности **255**
§ 7.10. Экспертные методы выбора **259**
§ 7.11. Человеко–машинные системы и выбор **263**
§ 7.12. Выбор и отбор **267**
 Заключение **272**
 Литература **273**
 Упражнения **275**
 Вопросы для самопроверки **276**
-

ДЕКОМПОЗИЦИЯ И АГРЕГИРОВАНИЕ КАК ПРОЦЕДУРЫ СИСТЕМНОГО АНАЛИЗА 277

- § 8.1. Анализ и синтез в системных исследованиях
277
§ 8.2. Модели систем как основания декомпозиции
281
§ 8.3. Алгоритмизация процесса декомпозиции **289**
§ 8.4. Агрегирование, эмерджентность, внутренняя
целостность систем **298**
§ 8.5. Виды агрегирования **301**
 Заключение **312**
 Литература **314**
 Упражнения **315**
 Вопросы для самопроверки **316**
-

О НЕФОРМАЛИЗУЕМЫХ ЭТАПАХ СИСТЕМНОГО АНАЛИЗА 317

- § 9.1. Что такое системный анализ **317**
 - § 9.2. Формулирование проблемы **319**
 - § 9.3. Выявление целей **324**
 - § 9.4. Формирование критериев **328**
 - § 9.5. Генерирование альтернатив **334**
 - § 9.6. Алгоритмы проведения системного анализа
342
 - § 9.7. Претворение в жизнь результатов системных
исследований **345**
 - § 9.8. О специфике социальных систем **354**
 - Заключение **366**
 - Литература **368**
 - Упражнения **369**
 - Вопросы для самопроверки **371**
-

Краткий словарь специальных терминов **373**