

Министерство образования и науки Российской Федерации

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

**Приоритетный национальный проект «Образование»
Национальный исследовательский университет**

В. Н. ВОЛКОВА

СИСТЕМНЫЙ АНАЛИЗ ИНФОРМАЦИОННЫХ КОМПЛЕКСОВ

*Рекомендовано Учебно-методическим объединением
по университетскому политехническому образованию в качестве
учебного пособия для студентов высших учебных заведений,
обучающихся по направлению подготовки
магистров «Системный анализ и управление»*

Санкт-Петербург
Издательство Политехнического университета
2014

УДК 002.531.5.681.3(075.8)

ББК 32.97 я 73

В

Рецензенты

Доктор технических наук, профессор кафедры «Математика и моделирование»

Петербургского университета путей сообщения

В.В. Изранцев

Кандидат технических наук, доцент кафедры «Информационные системы в экономике и менеджменте» Инженерно-экономического института Санкт-Петербургского

государственного политехнического университета

С.В. Широкова

В.Н. Волкова. **Системный анализ информационных комплексов**: учеб. пособие / Волкова В. Н. – СПб.: Изд-во Политехн. ун-та, 2014. – 336 с..

Излагаются теоретические основы информатики и информационного поиска, лежащие в основе создания и функционирования информационных комплексов различного вида и назначения. Рассматриваются особенности разработки Автоматизированных информационных систем (АИС) как первой стадии Автоматизированных систем управления (АСУ) предприятиями и организациями, особенности документальных и документально-фактографических информационно-поисковых систем (ИПС) научно-технической информации (НТИ); структура и принципы функционирования автоматизированных систем нормативно-методического обеспечения управления (АСНМОУ) предприятиями. Дается представление о разработке методик организации проектирования и сравнительного анализа готовых программных продуктов при создании информационных систем (ИС) для конкретной организации.

Проводится анализ состояния теории информационных систем как научного направления и кратко характеризуются перспективные направления развития теории.

Учебное пособие предназначено для студентов вузов, обучающихся по направлению подготовки магистров «Системный анализ и управление». Оно может быть также использовано при обучении студентов направлений подготовки «Прикладная информатика», в системах повышения квалификации, в учреждениях дополнительного профессионального образования и пр.

Работа выполнена в рамках реализации программы развития национального исследовательского университета «Модернизация и развитие политехнического университета как университета нового типа, интегрирующего мультидисциплинарные научные исследования и надотраслевые технологии мирового уровня с целью повышения конкурентоспособности национальной экономики»

Печатается по решению редакционно-издательского совета Санкт-Петербургского государственного политехнического университета.

© В.Н. Волкова, 2013

© Санкт-Петербургский государственный политехнический университет, 2013

ISBN 978-5-7422-4306-9

ОГЛАВЛЕНИЕ

Сокращения, принятые в тексте	5
Введение	6
Глава 1. Понятие об информации и информационных ресурсах предприятий и организаций	13
1.1. Понятие об информации как ресурсе предприятия (организации).....	13
1.2. Информация: понятие и основные свойства	18
1.3. Подходы к измерению и оценке ценности информации	38
1.4. Понятие об информатике как основе теории информационных систем и комплексов	50
Глава 2. Понятие об информационных процессах	67
2.1. Информационные процессы и их виды	67
2.2. Модели информационных процессов, основанные на методе алгебраического агрегирования	74
2.3. Модели информационных процессов, базирующиеся на теории выдвигения и проверке статистических гипотез А. Вальда	76
2.4. Исследование информационных процессов на основе метода постепенной формализации модели.....	80
2.5. Информационные процессно-ориентированные технологии	95
2.6. Закономерности информетрии и их применение для исследования информационных потоков	101
Глава 3. Понятие об информационных системах и классификации ИС.	116
3.1. Из истории разработки информационных систем	116
3.2. Виды и классификации информационных систем	120
Глава 4. Теоретические основы построения и проектирования АИС и АСУ.....	130
4.1. Принципы построения АИС и АСУ	130
4.2. Принципы организации управления разработками АИС и АСУ	133
4.3. Виды и принципы разработки структуры функциональной части АСУ	138
4.4. Разработка структуры обеспечивающей части АИС ...	161
4.5. Разработка методики для проектирования АИС предприятия	167
4.6. Выбор для предприятия готовых программных продуктов	171

4.7. Интегрированные информационные системы управления (ИАСУ)	173
4.8. Общегосударственная автоматизированная система	174
Глава 5. Теоретические основы построения и проектирования документальных информационно-поисковых систем научно-технической информации	184
5.1. Виды научно-технической информации и структура Государственной системы НТИ	184
5.2. Информационные потребности и информационное обслуживание	193
5.3. Понятие об информационном поиске	199
5.4. Информационно-поисковая система. Структура ИПС	201
5.5. Информационно-поисковый язык	207
5.6. Система индексирования	228
5.7. Логика ИПС. Критерии смыслового соответствия	231
5.8. Критерии оценки ИПС	233
5.9. Разработка и отладка ИПС	243
Глава 6. Принципы построения и проектирования документально-фактографических информационно-поисковых систем	246
6.1. Понятие о документально-фактографических ИПС	246
6.2. Автоматизированные системы нормативно-методического обеспечения управления предприятиями и организациями	248
Глава 7. Перспективные направления развития теории информационных систем	256
7.1. Концепция архитектуры предприятия и ее применение для развития теории информационных систем	256
7.2. Концепция ситуационных центров	277
7.3. Подходы к анализу и проектированию систем и их применение для развития теории проектирования информационных систем	288
7.4. Анализ определений системы и их применение для развития теории информационных систем	297
7.5. Концепция многоуровневой информационной системы	304
7.6. Проблемы и перспективы создания единой системы информации на основе объединения идей ГСНТИ и ОГАС	319
Заключение	326
Библиографический список	334

СОКРАЩЕНИЯ, ПРИНЯТЫЕ В ТЕКСТЕ

АДПАЦФ	–	автоматизированная диалоговая процедура анализа целей и функций
АДФИПС	–	автоматизированная документально-фактографическая информационно-поисковая система
АИС	–	автоматизированная информационная система
АСНТИ	–	автоматизированная система научно-технической информации
АСНМОУ	–	автоматизированная система нормативно-методического обеспечения управления
АСУ	–	автоматизированная система управления
АСУП	–	автоматизированная система управления предприятием
ГСНТИ	–	Государственная система научно-технической информации
ДИП	–	документальные информационные потоки
ДЦ	–	«дерево целей»
ДИПС	–	документальная информационно-поисковая система
ИАСУ	–	Интегрированная АСУ
ИДМ	–	имитационное динамическое моделирование
ИЛС	–	информационно-логическая система
ИПС	–	информационно-поисковая система
ИПЯ	–	информационно-поисковый язык
ИСС	–	информационно-семантическая система
ИУС	–	информационно-управляющая система
КИС	–	корпоративная информационная система
КСС	–	критерий смыслового соответствия
МАИС	–	методы активизации интуиции специалистов
МОСЭ	–	методы организации сложных экспертиз
МФПС	–	методы формализованного представления систем
НМД	–	нормативно-методические документы
НПД	–	нормативно-правовые документы
НТД	–	нормативно-технические документы
НТИ	–	научно-техническая информация
ОАСУ	–	отраслевая автоматизированная система управления
ОЧ	–	обеспечивающая часть
ОТП	–	организационно-технологические процедуры
СТК	–	сложный технический комплекс
СНМОУ	–	система нормативно-методического обеспечения управления
ТАСУ	–	территориальная автоматизированная система управления
ФЧ	–	функциональная часть
ЯАМ	–	язык автоматизации моделирования

ВВЕДЕНИЕ

Первыми информационными комплексами, которые стали разрабатывать в середине XX в. с помощью электронно-вычислительных машин (ЭВМ), были комплексы, разрабатываемые для космических исследований. Для их названия вводились разные термины. Так, для информационного комплекса радиоуправления ракетами был введен термин *системы обработки данных* – СОД.

Как информационные комплексы разрабатывались электро-механические аналоговые вычислительные машины (АВМ), с помощью которых был решен ряд задач динамики движения летательных аппаратов и управления морскими объектами, автоматизированной цифровой информационной машины, предназначенной для определения параметров траектории орбитальных объектов и передачи данных по каналам связи ¹.

Длительный период истории развития работ по созданию автоматизированных информационных комплексов связан с разработкой человеко-машинных систем разного рода, информационно-измерительных и информационно-управляющих систем (ИУС) – цифровых систем контроля или управления различными объектами, решающих разные задачи в реальном масштабе времени и имеющих определенную специфику (специфические требования по надёжности и безопасности функционирования; эксплуатационные и инструментальные особенности; непрерывный режим функционирования; возможность работы без оператора; разрешение нештатных ситуаций самой ВС). Наряду со специальной техникой для космических исследований к ИУС относят системы моделирования атмосферы, прогнозирования погоды и т. п.

¹ В частности, первая серия таких машин типа «Кварц» была изготовлена в 1957-1958 гг. на Ленинградском заводе им. М.И. Калинина. 15 мая 1958 г. машины «Кварц», обслуживаемые сотрудниками и студентами ЛПИ, впервые были использованы для определения параметров траектории третьего искусственного спутника Земли, а 12 апреля 1961 г. с их помощью обеспечивался полет в космос Ю.А. Гагарина.

В последующем термины СОД и ИУС были распространены и на объекты организационного управления предприятиями, территориями и т. д.

В частности, СОД использовался и применительно к системам обработки экономической информации. Но по мере увеличения памяти ЭВМ стали создаваться более объемные информационные массивы, введены понятия *база* и *банк данных* (БД), а в настоящее время – и *хранилище данных*, *хранилище информации*.

Начиная с 1960-х гг. в истории развития информационных систем нашей страны относительно независимо сформировались два направления: разработка *автоматизированных информационных систем* (АИС) как первой очереди *автоматизированных систем управления* (АСУ); разработка *автоматизированных систем научно-технической информации* – АСНТИ.

Первоначально разрабатывались АСУ технологическими процессами (АСУТП) и АСУ организационного управления предприятием (АСУП).

По мере расширения сфер автоматизации стали создаваться автоматизированные систем научных исследований (АСНИ); систем автоматизированного проектирования изделий (САПР) и автоматизированных систем технологической подготовки производства (АС ТПП); автоматизированных систем общезаводского управления, ориентированных на автоматизацию функций управления объединением, корпорацией (АСУО) и предприятием (АСУП); автоматизированных систем для комбинированного организационного и технологического управления (АСУ ОТ); автоматизированных систем управления гибкими производствами (АСУ ГПС), включая АСУ отдельными производствами, цехами, участками, гибкими автоматизированными линиями (ГАЛ), гибкими производственными модулями (ГПМ); автоматизированных систем контроля качества и испытания изделий (АСКИ).

Для объединения в единую систему разнообразных автоматизированных систем в 1980-1990-е гг. стали создаваться интег-

рированные автоматизированные системы управления (ИАСУ) разрабатывавшихся на предприятии.

По мере расширения сфер и средств автоматизации опыт показал, что в реальных условиях создать единую информационную систему (ИС) даже для крупных предприятий крайне сложно. ИС развивается постепенно, начиная с автоматизации наиболее значимых функций предприятий и организаций.

А современные предприятия и организации малого и среднего бизнеса, как правило, вообще не имеют средств для внедрения единой автоматизированной системы, поэтому существует практика выбора имеющихся на рынке программных продуктов, автоматизирующих соответствующие области управления (бухгалтерский учет, материально-техническое обеспечение и т. п.) с последующей интеграцией в единую систему.

Программные продукты, предназначенные для автоматизации различных сфер управления организацией, обеспечивающие автоматизацию достаточно полной совокупности функций, в настоящее время принято также называть информационными системами.

В их числе [11]: бухгалтерские информационные системы (БуИС), маркетинговые информационные системы (МИС), информационные системы мониторинга (ИСМ); системы сопровождения жизненного цикла изделий типа ИПИ-системы (ИПИ – информационная поддержка изделий) и системы управления проектными и инженерными данными предприятия – так называемые PDM-системы (Product Data Management Systems), включая САД-системы (Computer-Aided Design – компьютерная поддержка проектирования), САМ-системы (Computer-Aided Manufacturing – компьютерная поддержка изготовления), САЕ-системы (Computer-Aided Engineering – поддержка инженерных расчетов) и т. п.; системы управления взаимоотношениями с клиентами (CRM – Customer Relationships Management); SFA (Sales Force Automation) – автоматизации продвижения продаж; SCM (Supply Chain Management) – системы управления цепочкой по-

ставок (логистические информационные системы); ИС документооборота и делопроизводства (ИСДД).

Для комплексной автоматизации организационной деятельности фирм применяются ИС планирования потребности в материалах – MRP (Material Requirements Planning); ERP (Enterprise Resources Planning) системы, направленные на усовершенствование таких процессов, как планирование, изготовление, учет и контроль. Такие комплексные системы, создаваемые для корпораций, называют корпоративными. Наибольшее распространение из зарубежных КИС получили SAP R/3, Baan IV, Oracle Application, Ахарта 2.1; из российских – ИС: Предприятие, БОСС-Корпорация, Парус-Корпорация,

Для фирм различного профиля разрабатываются специализированные ИС: автоматизированные банковские информационные системы (АБИС), ИС фондового рынка; ИС в контроллинге, ИС бюджетирования; ИС в кадровом менеджменте; ИС в налоговых органах, в страховании и пенсионном обеспечении; ИС в туристическом бизнесе; ИС в управлении недвижимостью; ИС таможенной и внешнеэкономической деятельности и т. п.

С развитием систем телекоммуникаций стал активно развиваться информационный рынок, получивший название «пятый рынок». Развиваются информационные технологии (ИТ или IT) – способы и средства сбора, обработки, хранения и распределения информации. Для оперативного и интеллектуального анализа хранимых данных используются технология комплексного многомерного анализа данных OLAP (On-Line Analytical Proctssing) и технология интеллектуального анализа информации Data Mining.

На основе новых информационных технологий развиваются *процессо-ориентированные информационные системы*.

С учетом наличия на рынке готовых систем такого рода предприятия и организации, как правило, приобретают различные системы и средства автоматизации информационного обеспечения их

деятельности и формируемые наборы таких систем удобнее называть *информационными комплексами*.

В настоящее время существуют понятия: *информационно-производственный комплекс, информационно-технический комплекс, информационно-территориальный комплекс, информационно-библиотечный комплекс, сеть информационных комплексов, информационно-инфраструктурный комплекс и т.п.*

Каждый из этих видов информационных комплексов имеет определенную специфику. Для их проектирования и реализации разрабатываются соответствующие принципы, подходы, рекомендации, которые изучаются в различных дисциплинах с учетом соответствующих специальностей.

В то же время для любой технической и экономической специальности необходимы знания систем управления предприятиями и организациями, занимающимися созданием и реализацией разного рода продукции и услуг.

Поэтому пособие ориентировано в основном на разработку информационных комплексов для обеспечения управления предприятиями и организациями. С учетом различных средств автоматизации и терминов для их наименования специалистам различного профиля требуются знания в области информационных систем для их управления, умение разрабатывать или выбирать готовые программные продукты.

Подготовка специалистов по информационным системам разного рода ведется в вузах технического и экономического профиля уже более 40 лет. Существуют несколько Федеральных государственных образовательных стандартов высшего профессионального образования по направлениям и специальностям, связанным с разработкой и эксплуатацией информационных систем, определяющих требования к содержанию и уровню подготовки специалистов, бакалавров и магистров.

Имеются монографии и учебные пособия по различным направлениям информатики. В то же время нет единой трактовки понятий «информация», ее видов, способов измерения (мер ин-

формации), существуют различные подходы и теории проектирования информационных систем различного вида и назначения.

Ц е л ь данного учебного пособия – обобщить сведения об информационных комплексах для управления предприятиями и организациями; об основных особенностях и принципах их построения и проектирования.

Реализация этой цели требует анализа предшествующего опыта разработки информационных систем, основных понятий и теорий, применяемых при их разработке, и в частности, понятия информации, информационных процессов.

Несмотря на то, что пособие ориентировано в основном на разработку информационных комплексов для обеспечения управления предприятиями и организациям, есть и общие принципы, определяемые теорией систем, которые могут быть полезны специалистам, разрабатывающим и эксплуатирующим информационные комплексы различного вида и назначения.

Поэтому вначале излагаются некоторые сведения об информации, способах ее измерения, информационных ресурсах систем управления, информатике как научном направлении, ориентированном на исследование и разработку информационных комплексов разного вида и назначения (гл. 1); об информационных процессах и закономерностях информационных потоков (гл. 2), кратко рассматривается история разработки автоматизированных информационных систем различного вида и назначения, предлагается классификация ИС (гл. 3); рассматриваются особенности и теоретические основы разработки автоматизированных информационных систем (АИС) и автоматизированных систем управления (АСУ) предприятий (организаций), включая принципы формирования структур функциональной и обеспечивающей части, принципы выбора готовых программных продуктов для автоматизации деятельности предприятий и организаций (гл. 4); приводятся основные сведения о теории информационного поиска и информационно-поисковых систем (ИПС), рассматриваются принципы функционирования и структура докумен-

тальных информационно-поисковых систем (ДИПС), включая понятия об информационно-поисковом языке (ИПЯ), классификациях ИПС, системе индексирования (СИ), логике ИПС, критериях смыслового соответствия (КСС), критериях оценки ИПС, релевантности информационного поиска (гл. 5); дается представление о документально-фактографических информационно-поисковых системах (ДФИПС), о системе нормативно-методического обеспечения управления (СНМОУ) предприятия (организации) и ее автоматизированном варианте (АСНМОУ) как ИС документально-фактографического типа (гл. 6).

Для создания предпосылок формирования теории информационных комплексов рассматриваются концепции архитектуры информационной системы; ситуационных центров, концепции многоуровневых информационных систем (гл. 7), проводится анализ состояния теории информационных систем как основы разработки и реализации информационных комплексов.

Учебное пособие предназначено для студентов вузов, обучающихся по профилю «Системный анализ и управление» направления подготовки бакалавров и магистров «Системный анализ и управление». Оно может быть также использовано при обучении студентов направлений подготовки «Прикладная информатика» и другим направлениям, связанным с изучением и разработкой информационных систем, в том числе в системах повышения квалификации, в учреждениях дополнительного профессионального образования и пр.

Автор выражает благодарность рецензентам пособия доктору технических наук, профессору кафедры математики и моделирования Петербургского университета путей сообщения **Виталию Васильевичу Изранцеву** и кандидату технических наук, доценту кафедры Информационные системы в экономике и менеджменте Инженерно-экономического института СПбГПУ **Светлане Владимировне Широковой**.

Глава 1. ПОНЯТИЕ ОБ ИНФОРМАЦИИ И ИНФОРМАЦИОННЫХ РЕСУРСАХ ПРЕДПРИЯТИЙ И ОРГАНИЗАЦИЙ

1.2. Понятие об информации как ресурсе предприятия (организации)

В течение длительной истории развития цивилизации деятельность, выходящая за пределы материального производства и обслуживания, относилась к непроизводительным затратам. Экономический статус государства определялся имеющимися материальными и энергетическими ресурсами. Однако в настоящее время осознано, что основным ресурсом производственных предприятий и других организаций становится *информация*.

Переход к наукоемкому производству во второй половине XX в. сопровождался возникновением новых видов информационного обмена. Возросла зависимость экономики от источников, объемов и качества информации (научно-технической, экономической, политической и др.), от уровня развития средств доступа к этой информации, что привело к формированию на рубеже 80-х гг. XX в. принципиально новой экономической категории – *национальные информационные ресурсы*. Интенсивное развитие экономики становится невозможным без эффективного информационного сопровождения. Информация превратилась в стратегический ресурс, существенно возросла роль знаний.

В социально-экономических системах управления в зависимости от формы различают *документальную* и *фактографическую* информацию, в зависимости от приложений – *экономическую*, *социальную* информацию, классифицируемую по различным признакам.

Информационное пространство предприятия (организации) должно включать в себя весь спектр различных видов информации, отображающей состояние и функционирование конкретного предприятия или организации:

- сведения о состоянии производства (или обслуживания) и его обеспечении материальными, финансовыми и иными ресурса-

ми, представляемые в виде данных или фактов (*фактографическая информация*);

- научно-техническую информацию, представляемую в виде текстовых документов (*документальная информация*);
- нормативно-правовую информацию, также представляемую в виде текстов, в которых нужно искать элементы, при этом элементы – нормы, статьи, т. е. фрагменты текста;
- нормативно-техническую, нормативно-методическую, маркетинговую, мониторинговую информацию, являющуюся документально-фактографической, т. е. представляемой в форме текстов, из которых нужно извлекать конкретные данные о производственных нормативах, состоянии среды, сведения о наличии и ценах товаров на рынках и т. п.

Разными исследователями предлагались различные способы классификации информационного обеспечения.

Так, с точки зрения взаимодействия предприятия (организации) с окружающей средой всю информацию принято делить на *входящую* и *исходящую*. В зависимости от сроков хранения различают *постоянную*, *условно-постоянную* (иногда обновляемую) и *переменную* (регулярно изменяющуюся). Классифицируют информацию и *по уровням управления* (корпоративная, внутризаводская, цеховая, внутрицеховая), *по характеру деятельности* (конструкторско-технологическая, учетно-отчетная, плановая и т. п.). В автоматизированных системах информационное обеспечение делят на *машинное* (в памяти компьютера) и *внемашинное* (на бумажных носителях).

Различные классификации предлагались и использовались в системах управления, как правило, для информации, создаваемой и хранящейся в форме документов (приказов, планов, писем, справочно-табличных форм статистической отчетности и т. п.).

Однако по мере развития автоматизированных средств появилась возможность регистрации и хранения информации в виде отдельных фактов (характеристик предметов, событий, операций и т. п.), т. е. в виде массивов *фактографической* информации, в которых данные могут сортироваться по различным признакам и выводиться в различных формах, удобных для решения той или иной управленческой или проектной задачи. При создании банков данных или баз данных фактографической информации использовались более разнообразные классификации.

Например, в создаваемых для отображения производственного процесса

банках данных формировались БАНК, СИОД²: главный предметный массив (ГП), массив состава изделия (СИ), массив рабочих мест или рабочих центров (РМ, РЦ), массив пооперационных трудовых нормативов (ПТН), т. е. выделялись группы данных, характеризующих предмет, средства, условия труда при производстве конкретных изделий. При создании информационных массивов о кадрах (сотрудниках предприятия) данные классифицировали в соответствии с личным делом работника, выделяя социально-демографическую, производственную, общественную и другую информацию о личности.

Разделение на документальную и фактографическую информацию носит более принципиальный характер, когда речь идет о научно-технической информации (монографиях, статьях, отчетах, патентах, законодательных актах и т. п.). Такая информация формируется человеком всегда в виде текстов, т. е. в форме документальной информации, а тексты (даже относительно структурированные) имеют ряд принципиальных особенностей (синонимия, омонимия, парадоксы), которые затрудняют извлечение из них фактографической информации, необходимой для решения практических или управленческих задач.

При этом иногда усложняется собственно понятие факта, который должен представлять собой минимальную единицу информации, принципиально значимую для того или иного вида деятельности. В частности, применительно к нормативно-правовой информации фактом следует считать не отдельный термин, а юридически значимую норму, регламентирующую управленческую деятельность, и эта форма факта должна иметь определенную структуру. Иными словами, при фактографическом анализе текста речь идет о расчленении целостного объекта (текста) на минимально значимые части (факты), и здесь следует учитывать рекомендации о расчленении системы на элементы, т. е. применять системный анализ.

Для научно-технической информации также введены и используются различные классификации: по видам источников документальной информации (первичные и вторичные), видам

² *Келехсаев А. А.* Системы интеграции и обработки данных СИОД1, СИОД2 / А. А. Келехсаев, А. П. Беляев. – М.: 1977. – 208 с.

информационных изданий (общегосударственные, отраслевые, внутрифирменные).

В последнее время появился еще один важный вид информации – *программные продукты* (программное обеспечение, сертифицированное и предложенное к распространению, в том числе и к продаже). Это программное обеспечение для станков с ЧПУ, автоматических линий и т. п., компьютерные программы для планирования, учета и обслуживания производственных процессов, программная поддержка научных исследований, менеджмента и маркетинга, программные средства автоматизации поиска информации, анализа текстов и т. п.

Научно-техническая информация и программные продукты могут быть ориентированы не только на внутреннего потребителя, но и быть предметом экспорта.

Пример структуризации информации для производственной системы приведен на рис. 1.1, на котором информация структурирована по сферам деятельности предприятия: производственная информация, информация для научно-исследовательской деятельности, информация для организационного управления, научно-техническая информация.

Разработка структуры информационного обеспечения предприятий или организаций является важной и сложной задачей, от решения которой во многом зависит эффективность их деятельности. Эту задачу следует решать с учетом конкретных особенностей предприятия (организации).

Для более эффективного удовлетворения производственных потребностей сотрудников предприятия необходимы знания о сути и особенностях информации различных видов, об информационных процессах, информационных потоках и их закономерностях, целесообразно разрабатывать классификаторы не только по видам информации, но и по ее содержанию, т. е. по функциям производственного процесса, организационного управления и т. д.

Эти вопросы с учетом кратко излагаемой истории развития теорий проектирования фактографических и документальных информационных систем рассматриваются в последующих разделах.



Рис. 1.1. Структура видов информации для производственной системы

1.2. Информация: понятие и основные свойства

Строгого определения информации в настоящее время нет.

Примем вначале наиболее распространенное определение и рассмотрим особенности информации.

Информация – это сведения об окружающем мире (объекте, процессе, явлении, событии), которые являются объектом преобразования (включая хранение, передачу и т. д.) и используются для выработки поведения, для принятия решения, для управления или для обучения.

Характерные черты информации:

1) является основой развития научных исследований, которые, в свою очередь, обеспечивают развитие любой организации и общества в целом;

2) представляет собой наиболее важный ресурс современного производства: снижает потребность в земле, труде, капитале, уменьшает расход сырья и энергии;

3) вызывает к жизни новые производства;

4) является товаром, причем продавец информации ее не теряет после продажи;

5) придает дополнительную ценность другим ресурсам, в частности, трудовым: действительно, работник с высшим образованием ценится больше, чем со средним;

6) может накапливаться.

С информацией всегда связывают три понятия (их взаимосвязь показана на рис. 1.2):

- *источник информации* – тот элемент окружающего мира, сведения о котором являются объектом преобразования;

- *потребитель информации* – тот элемент окружающего мира, который использует информацию;

- *сигнал* – материальный носитель, который фиксирует информацию для переноса ее от источника к потребителю и *код*, в форме которого информация передается.

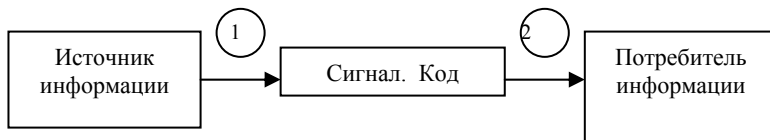


Рис. 1.2. Схема взаимосвязи основных понятий информации

Так, *источником* информации, которую получает читатель, является информатика как сфера человеческой деятельности; *потребителем* – читатель, а *сигналом* – бумага с текстом (в этом случае говорят, что информация имеет бумажный носитель).

Будучи прочитанной и освоенной потребителем информация приобретает еще один носитель – биологический, когда она «записывается» в память человека. Очевидно, что источник и потребитель в этом случае не меняются.

Исследованием феномена информации, ее видов и особенностей занимаются философы, специалисты по техническим наукам, экономисты, психологи.

Обобщая и преломляя результаты этих исследований применительно к проблеме использования информации как ресурса развития предприятий и организаций, выделяют ряд специфических особенностей информации, обуславливающие ее отличие от других видов ресурсов: *практически неубывающая потенциальная эффективность информации, тиражируемость и многократность использования, зависимость фактической реализуемости и эффективности от степени использования информации, абстрактный характер* и т. п.

Основные из них кратко охарактеризованы в табл. 1.1.

Таблица 1.1

Основные особенности информации

Особенности информации	Краткая характеристика
<i>Практически неубывающая потенциальная эффективность информации</i>	<p>Потенциальная эффективность информационных ресурсов не исчезает после не только однократного, но и многократного использования одной и той же информации (уменьшение эффективности может быть связано только с устареванием информации). В то время как основная характеристика материальных ресурсов (минеральных ресурсов, сырья и т. п.) – их потенциальная эффективность – уменьшается по мере их использования. В случае возобновляемых ресурсов их потенциальная эффективность может восстанавливаться, но для этого требуется определенный (часто значительный) период. Потенциальная эффективность искусственно создаваемых технических средств, в том числе оборудования, также имеет предел, определяемый их сроком службы, и исчезает после списания станка или другого технического средства. Особенность информационных ресурсов часто реализуется не сразу, а спустя многие годы (как бывало, например, с рядом открытий, изобретений или новых идей, не понятых поколениями при жизни их авторов).</p>
<i>Независимость информации от ее создателей</i>	<p>Свойство неубываемости информации является одним из условий потенциальной эффективности информации и приводит к относительной <i>независимости информации от ее создателей</i>.</p>
<i>Тиражируемость и многократность использования</i>	<p>Иногда говорят и о тиражируемости технических средств, например, станка. Но в этом случае речь идет о тиражируемости конструкции станка, идеи, т. е. информации, а на создание каждого экземпляра станка одного и того же типа снова нужно затратить труд, в то время как при тиражировании информации труд (умственный) на ее создание больше не тратится, а труд, затрачиваемый на размножение информации, незначителен и им, как правило, можно пренебречь при оценке эффективности использования информации.</p>

Особенности информации	Краткая характеристика
<i>Неаддитивность, некоммутативность и неассоциативность информации</i>	Отсутствие свойств <i>аддитивности, коммутативности</i> и <i>ассоциативности</i> у информации означает, что содержащаяся в каком-либо сообщении информация не есть просто арифметическая сумма составляющих это сообщение элементов информации, что эти элементы нельзя без искажения смысла располагать в сообщении в любой произвольной последовательности и группировать в произвольные сочетания. Эта особенность информации относится к научной и социальной информации.
<i>Кумулятивность информации</i>	Это свойство, характеризующее также в основном научную и социальную информацию, связано с одной из основных закономерностей развития науки и общества – ее преемственностью. Предшествующие достижения науки и культуры являются фундаментом для дальнейшего развития научной мысли и общества. Поэтому необходимы не только новые открытия, но и специальная работа по систематизации, оценке и обобщению информации. Со свойством кумулятивности тесно связано и явление концентрации информации во времени, т. е. переход знания о мире на все более высокие уровни абстракции.
Зависимость фактической <i>реализуемости</i> и <i>эффективности</i> информации от степени ее использования	В случае материальных ресурсов эффективность можно оценить коэффициентом использования материалов, сырья, электроэнергии и т. п., коэффициентом полезного действия оборудования и др. технических средств. Оценка реализуемости и эффективности информации – более сложная проблема, которая кратко будет рассмотрена ниже.
Необходимость наличия <i>источника, переносчика</i> и <i>приемника</i> (потребитель информации)	Сообщение становится информацией только в случае, когда есть <i>источник, переносчик</i> (в том числе передатчик, носитель) и <i>приемник</i> (потребитель), который должен хотеть воспринять информацию и быть способным ее понять и использовать. Эта особенность является определяющей при оценке как потенциальной, так и фактически реализуемой эффективности информационных ресурсов.

Особенности информации	Краткая характеристика
<i>Материя и информация</i> – парные философские категории	<p>Появление новой информации всегда сопутствует появлению новых форм существования материальных объектов и процессов, независимо от того, осознают ли этот факт их создатели.</p> <p>Поэтому вклад в формирование информационных ресурсов вносят практически все профессиональные группы работников предприятия: рабочие, которые создают новые образцы изделий и принимают участие в совершенствовании технологии (рационализаторская и изобретательская деятельность), инженерно-технические работники, которые проектируют изделия, системы, новые технологические процессы и т. п., ученые и научные работники, которые изучают явления и процессы естественного и искусственного (созданного руками человека) мира, делают открытия, разрабатывают фундаментальные основы будущих новых изделий, систем, технологий, а также руководители предприятий (организаций) и управленческие работники, которые занимаются непосредственно регистрацией, хранением и обработкой производственной и управленческой информации разного рода.</p>
<i>Наличие ценности</i>	<p>Ценность или полезность информации влияет на поведение получателя информации, на принятие им управленческих решений. Информация имеет тем большую ценность, чем в большей мере она <i>влияет на достижение целей</i>, стоящих перед получателем информации.</p>

Имеется и ряд других особенностей информации, специфичных для научно-технической информации.

Научно-техническая информация (НТИ) или в широком смысле научная информация – это «получаемая в процессе познания логическая информация, которая адекватно отображает явления и законы природы, общества и мышления и используется в общественно-исторической практике»³.

³ *Словарь* современного русского литературного языка. – М. – Л., Изд-во АН СССР, 1956. – Т. 5. – С. 418.

Это определение отражает четыре наиболее важных признака, необходимых для раскрытия понятия научная информация:

1. Научной может быть лишь информация, полученная в процессе познания объективных закономерностей природы, общества и мышления. При этом основу процесса познания составляет практика, производственная деятельность людей. К практике относятся материальное производство, научный эксперимент, сельскохозяйственная деятельность, деятельность по преобразованию природы и т. п.

2. Не всякая информация, полученная в процессе познания, есть научная. Чувственное познание дает представление лишь об отдельных, внешних сторонах вещей. Для того чтобы стать руководством к практической деятельности, необходимо знание сущности вещи, законов природы и общественной жизни. А для этого необходимо абстрактно-логическое мышление, облеченное в языковую форму. Именно этим *научная информация* отличается от *сведений* или *данных*, получаемых в процессе чувственного (эмпирического) познания. Данные и сведения – это «сырье» для создания научной информации.

3. Необходимо адекватное отображение явлений и законов природы, общества и мышления. При этом определение адекватности новых гипотез или теорий является сложной проблемой, по-разному решаемой на различных этапах развития общества с учетом конкретно-исторических условий.

4. Научная информация непременно должна использоваться в общественно-исторической практике. Это обеспечивает проверку истинности логической информации.

А. Д. Урсул уточняет приведенное определение, дополняя его возможностью *использования информации для предвидения и преобразования действительности*⁴ и отмечает необходимость *фиксации получаемой в процессе познания информации в системе точных понятий* (суждений, умозаключений, теорий, гипотез).

В то же время в [9, 10] научную информацию отличают от научных знаний. В *знаниях* информация представлена уже в наиболее обобщенном и систематизированном виде и выражается в *системе понятий, суждений, умозаключений, в теориях*. Таким образом, научные знания – не вся научная информация, а лишь ее определенная часть. Предлагается различать ряд уровней, на которых совершаются акты мышления: сбор научных фактов,

⁴ *Урсул А. Д.* Информация / А. Д. Урсул. – М.: Наука, 1971. – 295 с.; *Урсул А. Д.* Отражение и информация / А. Д. Урсул. – М.: Мысль, 1973. – 231 с.

систематизация и обобщение, формулирование и проверка гипотез, концепций, законов, построение теорий.

Первоначально различали *научную* и *научно-техническую* информацию, относя последний термин к области «наука и техника» [9, 10]. Однако в последующем эти термины используются как синонимы, поскольку информация, полученная и используемая в технике (как и в любой другой области человеческой деятельности) обязательно относится к сфере науки, получена в результате научных исследований. Поэтому в настоящее время эти термины используются как синонимы, и даже в большей мере принят термин «научно-техническая информация».

Основные специфические особенности научно-технической информации [9, 10, 11] приведены в табл. 1.2.

Таблица 1.2

Специфические особенности научно-технической информации

Особенности НТИ	Краткая характеристика
1. <i>Неотрывность научной информации от физического носителя</i>	Несмотря на то, что научная информация по своей природе идеальна (нематериальна), она не может существовать вне той или иной материальной оболочки (носителя). Эта особенность отражает более общую: материя и информация – парные категории. Информация не существует без отражения материи, как и без других атрибутов материи – движения, пространства, времени и т. д.; она является содержанием отражения.
2. <i>Неаддитивность, некоммуникативность и неассоциативность</i>	Отсутствие свойств <i>аддитивности, коммутативности и ассоциативности</i> у научной информации, на что впервые обратил внимание В. И. Сифоров ⁵ означает, что содержащаяся в каком-либо сообщении информация не есть просто арифметическая сумма составляющих это сообщение элементов информации (например, слов), что эти элементы нельзя без искажения смысла располагать в сообщении в любой произвольной последовательности и группировать в произвольные сочетания.

⁵ **Сифоров В. И.** Научная информация и повышение эффективности научных исследований / В. И. Сифоров // Управление, планирование и организация научных и технических исследований: кн. – М.: 1970. – Т. 3. – С. 291–292.

Особенности НТИ	Краткая характеристика
3. <i>Наличие ценности субъективно</i>	<p><i>Ценность</i> научной информации <i>субъективна</i>. То, что очень важно для одного субъекта, может быть совсем не важно для другого и наоборот. Производство информации всегда должно быть <i>адресным</i>, иначе оно бессмысленно.</p>
4. <i>Общественная природа</i>	<p>Источником научно-технической информации является познавательная деятельность людей. Явления и законы природы познаются совокупностью индивидуумов в их взаимодействии, обществом. Общество обладает памятью, более емкой и разносторонней, чем память любого индивидуума.</p> <p>Совокупность научных знаний о явлениях и законах мира и деятельности людей, накопленная обществом, называют общечеловеческим <i>тезаурусом</i>. В этом смысле совокупность понятий и закономерностей функционирования предприятия иногда называют <i>тезаурусом предприятия</i> [Шемакин Ю. И. <i>Тезаурус в автоматизированных системах управления и обработка информации</i> / Ю. И. Шемакин. – М., 1974. – 172 с.].</p>
5. <i>Семантический характер и языковая природа</i>	<p>Термин «семантический» характеризует научную информацию в плане содержания, а не формы (план выражения). Научная информация является смысловой, семантической, т. е. имеет понятийный характер. Именно понятия составляют смысл слов, в них обобщаются существенные признаки явлений и процессов.</p> <p>При этом понятия формируются и существуют в языковой оболочке. Под языком понимается знаковая система любой физической природы, служащая средством человеческого общения, мышления и выражения. Язык может быть естественным и искусственным, т. е. созданным для узких потребностей (математическая символика, язык структур, информационно-поисковый язык и т. п.). Язык – явление социальное. Термин «языковый» (от греч. <i>λογος</i>, <i>logos</i> – слово, понятие, смысл, разум). Поэтому научную информацию можно считать логической. В то же время семантика НТИ значительно более изменчива, чем форма ее выражения – язык.</p>

Особенности НТИ	Краткая характеристика
6. <i>Независимость от языка и носителя информации</i>	Инвариантность научной информации относительно языка, на котором она выражается, и относительно ее физических носителей характерна для научной информации, но не является абсолютно универсальным свойством. Например, эстетическая информации существенно зависит от языка, и даже от физического носителя. Аналогично конструкторская информации не может быть адекватно отображена в вербальной форме.
7. <i>Дискретность</i>	<p>Дискретность является диалектически противоречивым свойством. Семантическая научная информации: в плане языкового выражения представляется единицами языка (словами, предложениями, абзацами, т. е. дискретными носителями); а в плане содержания – понятиями, описанием фактов, законов, гипотез, концепций, теорий, которые можно считать превращением дискретной информации в непрерывную, целостную.</p> <p>В то же время «научные произведения» имеют логическую завершенность, и отчуждение этих произведений от их творцов происходит не непрерывно, а дискретно, в виде относительно законченных творческих актов, получающих материальное воплощение, и их можно считать «квантом» научной информации.</p> <p>В случае устных научных произведений (лекции, доклады, выступления) содержательная дискретность проявляется в виде смысловых фрагментов, а материальная дискретность – в виде риторических периодов, характеризующихся паузами.</p> <p>Свойством дискретности обладают не все виды научной информации. Например, данные, получаемые с помощью некоторых измерительных приборов, могут быть и непрерывными (аналоговые устройства).</p>
8. <i>Независимость от создателей</i>	Это свойство характерно в большей мере для научной социальной информации, и особенно для информации, используемой при организационном управлении. В этой сфере передача опыта часто осуществляется без фиксации авторов нововведений.

Особенности НТИ	Краткая характеристика
	<p>Для регистрации авторства в области научно-технической информации существуют центры регистрации и выдачи патентов и авторских свидетельств, в которых проводится работа по установлению патентной чистоты на основе сопоставления предлагаемых изобретений и открытий с предшествующими; патентные библиотеки; центры регистрации научных отчетов, диссертаций и др. неопубликованных материалов, специальные издания, представляющие обобщенные сведения о новых научно-технических материалах и их авторах.</p> <p>В нашей стране для обеспечения работы по поиску авторов научно-технической информации в отраслях науки и техники существуют специальные издания – Бюллетень изобретений и открытий, сигнальные и обзорные информации, экспрес-информации, библиографические указатели, реферативные журналы и т. д.</p> <p>По мере развития науки, увеличения числа открытий и изобретений возникла проблема авторства и в сфере научных исследований. Все больше наука становится не индивидуальной, а общечеловеческой. Возрастает число крупномасштабных коллективных научных работ. Усложняется проблема выявления и учета цитируемости опубликованных материалов, особенно статей и тезисов докладов. В XIX в. этой работой активно занимался Всесоюзный институт научно-технической информации (ВИНИТИ).</p> <p>Однако эта работа может проводиться в основном за счет государственного финансирования, и в современных условиях рынка и сети Интернет становится все более затруднительной. Существует и проблема охраны авторских прав (особенно в сети Интернет).</p> <p>Это свойство связано с одной из основных закономерностей развития науки и общества – ее преемственностью и интернациональным характером.</p>

Особенности НТИ	Краткая характеристика
	<p>Предшествующие достижения науки и культуры являются фундаментом для дальнейшего развития научной мысли и общества.</p>
<p>9. <i>Кумулятивность</i></p>	<p>Необходимы не только новые открытия, но и специальная работа по систематизации, оценке и обобщению информации. Со временем все второстепенное, частное отсеивается, а существенное, основное получает простое и краткое выражение. Со свойством кумулятивности тесно связано и явление <i>концентрации</i> информации во времени, переход знания на более высокие уровни абстрагирования.</p>
<p>10. <i>Концентрация и рассеяние (интеграция и дифференциация) информации</i></p>	<p>Явление <i>концентрации</i> информации во времени – это переход знания о мире на все более высокие уровни абстракции: А. Научные факты. ⇒ В. Научные гипотезы, концепции. ⇒ С. Совокупность гипотез, концепций теорий, законов, образующие основы определенной науки, парадигмы, отрасли знаний. ⇒ D. Научная информация мировоззренческого характера.</p> <p>Условно явление концентрации представлено на рис. Т2.2.1.</p> <div data-bbox="453 973 940 1228" style="text-align: center;"> </div> <p>Рис. Т1.2.1</p> <p>Это явление тесно связано со свойством кумулятивности.</p>

Особенности НТИ	Краткая характеристика
	<p>Свойство научной информации <i>рассеиваться</i> по многим научным произведениям непосредственно связано с рассмотренными выше свойствами дискретности, кумулятивности, независимости от создателей и свойством старения. Оно проявляется в том, что одни и те же содержательные единицы научной информации (факты, понятия, гипотезы, теории и т. п.) одних авторов могут по-разному и в разных контекстах использоваться в произведениях других авторов, могут менять языковое выражение.</p> <p>Это неизбежно, способствует развитию научной мысли. Заимствованная информация, располагаясь в новых произведениях в соответствии с замыслом их авторов, формирует новый целостный смысл, что и обеспечивает развитие науки.</p> <p>Явления <i>концентрации</i> и <i>рассеяния</i> представляют собой проявление явлений <i>интеграции</i> и <i>дифференциации</i>, которые являются закономерностями развития науки. Изучения этих явлений – важнейшая проблема информатики. В результате исследования этой проблемы были открыты законы Г. Ципфа, Б. Мандельброта, С. Брэдфорда, Б. Викери, закономерность концентрации – рассеяния В. И. Горьковой (см. гл. 7).</p>
11. <i>Старение</i>	<p>Старение научно-технической информации не связано непосредственно только со временем ее возникновения. В строгом понимании устаревают только те факты, гипотезы теории и т. п., которые с появлением новой научной информации оказываются неверными, перестают адекватно отображать явления и закономерности материального мира, общества и мышления, т. е. перестают быть научной информацией. Примером устаревания информации служит замена геоцентрической концепции Аристотеля–Птолемея на гелиоцентрическую концепцию Коперника–Бруно.</p> <p>Наибольшему старению подвержена информация класса В, в наименьшей – классов А и С.</p>

Особенности НТИ	Краткая характеристика
	<p>К проблеме старения информационных изданий нужно относиться аккуратно. Критерием здесь не может служить год издания, поскольку первоисточник может содержать более точные и глубокие сведения по сравнению с его пересказом в других изданиях.</p> <p>В то же время старение информационных изданий может быть связано не с устареванием фактов или теорий, а с более сжатым их изложением, с изменением формализованного языка представления этих теорий (что происходило, например, с алгеброй, статистикой и др. науками), т. е. со свойством кумулятивности научной информации. Старение может быть связано с изменением господствующей концепции, парадигмы на том или ином этапе развития научного знания.</p>
<p>12. <i>Зависимость получаемой информации от априорной информации</i></p>	<p><i>Потребление информации</i> требует определенных <i>навыков и усилий</i>: чем сложнее информация, тем большим интеллектом должен обладать ее потребитель. Усвоение многих видов информации требует специальных знаний, расширения <i>тезауруса</i> потребителя информации. Проводятся исследования зависимости получаемой информации от априорной, т. е. от тезауруса потребителя. Например, исследования Ю. А. Шрейдера (см. раздел 2.2).</p>

Особенности научно-технической информации необходимо учитывать при разработке информационно-поисковых систем, при комплектовании и актуализации информационных фондов.

Виды и формы научно-технической информации будут рассмотрены при характеристике систем научно-технической информации (гл. 4).

Рассмотренные особенности научно-технической информации позволяют охарактеризовать ее соотношение с другими

видами экономической и социальной информации. А это, в свою очередь, помогает решать вопросы сбора, хранения, предоставления, защиты информации.

В настоящее время достаточно широко распространено мнение экономистов о том, что научно-техническая информация является не только важным ресурсом экономической системы, но и товаром. В то же время есть точки зрения, в соответствии с которыми научно-техническая информация не является товаром в классическом смысле, а если и имеет ценность на рынке, то как товар особого свойства.

Действительно, научно-техническую информацию можно отнести к категории товара на основе следующих свойств: 1) НТИ есть продукт труда; 2) НТИ продается и покупается, т. е. имеет цену на рынке. При этом товар есть продукт труда, который удовлетворяет какую-либо потребность людей и производится не для собственного потребления, а для продажи. И этим свойствам тоже в принципе может удовлетворять научная информация.

Правда, по Марксу «Для того, чтобы стать товаром, продукт должен быть передан в руки того, кому он служит в качестве потребительной стоимости посредством обмена».⁶ В то же время, хотя изначально научно-техническая информация создается не для продажи, она может использоваться посредниками (реализация книг фирмами-посредниками, продажа патентов и т. п.).

Стремлению рассматривать научно-техническую информацию как товар способствуют такие ее свойства, как практически неубывающая потенциальная эффективность информации, тиражируемость и многократность использования, зависимость фактической реализуемости и эффективности от степени использования информации, наличие ценности.

Если научно-техническую информацию признать товаром, то следует изменить отношение к ее распространению, к ее представлению в сети Интернет.

В то же время научно-техническая информация – особый товар, и нужно, видимо, по-разному относиться к хранению и организации доступа к различным ее видам, учитывать, что НТИ является основой развития общества в целом, основой развития науки и образования.

⁶ Маркс К., Энгельс Ф. Соч. – Т. 23. – С. 49.

По мере исследования свойств информации разного вида стали осознавать, что существуют различные виды информации.

Так, **Е. К. Войшвилло**⁷ предложил различать информацию *восприятия* (знаки) и *информацию-значение*, смысловое содержание информации (означаемое знака).

В дальнейшем с точки зрения пользователя информацию стали рассматривать в нескольких аспектах:

синтактический (или материально-энергетический, знаковый, технологический) – с точки зрения техники передачи информации;

семантический – с точки зрения смыслового содержания и правильности толкования;

прагматический – для достижения целей потребителя.

В последующем стали рассматривать эти аспекты и как уровни существования информации и вводить для разных уровней разные оценки (подробнее см. в следующем разделе).

Развивалось понятие *информации* и как философской категории. В частности, **А. Д. Урсул**⁸ стал рассматривать информацию как форму существования материи, подобно массе и энергии, а в последующем как категорию, парную материи, как отражение материи.

В *теории информационного поля А. А. Денисова*⁹ информация рассматривается как *структура материи*, как категория, парная по отношению к материи. В этой теории вводятся понятия *чувственной* и *логической* (семантической и прагматической) информации, которые для удобства практических приложений на-

⁷ **Войшвилло Е. К.** Понятие / Е. К. Войшвилло. – М.: МГУ, 1967. – 286 с.; **Войшвилло Е. К.** Символическая логика: Классическая и релевантная: Философско-методологические аспекты: учеб. пособие / Е. К. Войшвилло. – М.: Высшая школа, 1989. – 149 с.

⁸ **Урсул А. Д.** Информация / А. Д. Урсул. – М.: Наука, 1971. – 295 с.; **Урсул А. Д.** Отражение и информация / А. Д. Урсул. – М.: Мысль, 1973. – 231 с.

⁹ **Денисов А. А.** Теоретические основы кибернетики: Информационное поле / А. А. Денисов. – Л.: ЛПИ, 1975. – 40 с.

званы *информацией восприятия* и *информационным потенциалом* ([6] и разделы, подготовленные **А. А. Денисовым** в [1, 13, 16].

Долгое время в предыдущих теориях эти информации рассматривались как отдельно существующие, или как уровни информации, и поэтому единое определение понятия *информации* являлось предметом дискуссий.

А. А. Денисов показал, что *информацией* для потребителя является пересечение информации восприятия или *чувственной информации* и ее потенциала или *логической информации*, в результате чего формируется единое понятие – *информационный смысл, информационная сложность*.

При этом он определяет информацию следующим образом:

«... информация – это понятие, не поддающееся анализу средствами формальной логики и требующее применения к нему диалектической логики, которая обеспечивает возможность анализа не только абсолютно, но и относительно истинных высказываний» [1, с. 239].

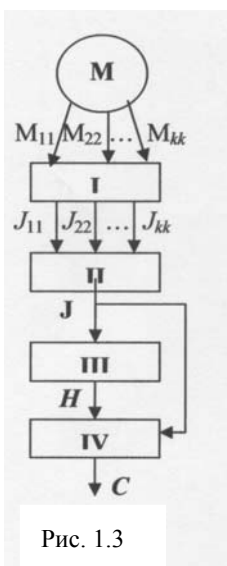


Рис. 1.3

В отличие от этой точки зрения **А. А. Денисов** в теории информационного поля и ее дискретном варианте, известном как информационный подход к анализу систем [1, 6, 13], рассматривает информацию как категорию, парную по отношению к материи, и вводит понятия *чувственной J* и *логической H* информации, которые для удобства практических приложений названы *информацией восприятия* и *информационным потенциалом*.

Согласно формуле познания «От живого созерцания к абстрактному мышлению, и от него – к практике»¹⁰ можно выделить три этапа отражения действительности: два пассивных – *чувственное* и *логическое* отражение, и один активный – этап *прагматическо-*

¹⁰ **В.И. Ленин**. Философские тетради / Полн. собр. соч. – Т. 29. – С. 153.

го отражения (осмысления), приведенные на рис. 1.3.

На рис. 1.3 обозначено: I – блок измерений (ощущений); II – коррелятор; III – осмысление значимости информации H (семантики); IV – мультипликатор ($C = J \times H$).

Соответственно продуктами этих этапов являются *чувственная, логическая и прагматическая информация, и информационная сложность*.

Этап живого созерцания (чувственного отражения) начинается с выделения объекта нашего интереса M – *восприятия информации*

I – блок измерений (ощущений);

Опираясь на введенные И. Кантом понятия «вещь в себе» и «вещь для нас» А.А. Денисов вводит соотношение В первом варианте представления АА были понятнее:

$$J_n = R_k M, \quad (1.1)$$

где M – измеряемое материальное свойство (масса, цвет, заряд и т. п.), создающее J_c («информация в себе»);

J_n – *чувственная информация (информация для нас) или информация восприятия*, которую в дальнейшем для краткости будем использовать без индекса;

R_k – *относительная информационная проницаемость среды*.

С учетом того, что восприниматься может не только материя, но и информация, соотношение (1.1), можно записать так:

$$J_n = R_k J_c = R_k M, \quad (1,1, a)$$

Соотношение (1.1) реализует *первый* из постулат теории информационного поля – *об адекватности отражения материи, закон чувственно-го отражения*, в соответствии с которым информация есть функция материи, которая, по меньшей мере, для ограниченных приращений носит характер пропорциональной зависимости

Если объект сложный – то он отражается вектором информации. И тогда – нужно как-то синтезировать компоненты вектора. Просто сумма – не годится даже по Аристотелю – «Сумма свойств частей не есть свойство целого», т.е. имеет место эмерджентность.

Поэтому А.А. обращается к теории поля – теорема Гаусса:

В формализованном представлении поле можно отобразить следующим образом. Если говорить об отражении материального объекта или поля некой произвольной замкнутой вокруг него поверхностью, то полная информация составит из суммы потоков информации, проходящих на единицу dS площади этой поверхности, т. е. из $\mathbf{O} = dJ/dS$. В таком случае должна иметь место теорема Гаусса, являющаяся математическим выражением философского положения о познаваемости мира:

$$M = \oint_S \mathbf{O} dS \quad \text{или} \quad J_c = \oint_S \mathbf{O} dS, \quad (1.2)$$

где \mathbf{O} – вектор интенсивности потока существования (отражения); интеграл берется по замкнутой поверхности S , охватывающей изучаемое явление или объект.

Соотношение (1.2) означает, что всякая информация в себе создает поле существования, суммарный поток которого адекватен этой информации, т. е. материи, служащей источником поля. Это соотношение подобно теореме Гаусса. Из теоремы Гаусса в форме (1.2) следует, что источник поля информации J принципиально полностью идентифицируем по реакции тех или иных пробных материальных объектов на изучаемое им поле существования без непосредственного контакта с самим источником.

С учетом (1.1) теорему Гаусса можно представить в форме:

$$J_u = \oint_S R_k \mathbf{O} dS = \oint_S \mathbf{O}_u dS, \quad (1.3)$$

где $\mathbf{O}_u = R_k \mathbf{O}$ – вектор интенсивности отражения.

Посредством измерительных приборов в блоке I идет процесс расчленения M на элементарные свойства M_{kk} и их измерение с получением чувственной информации J_{kk} .

Далее идет процесс логического синтеза вектора восприятия J в подходящей к случаю системе координат путем формирования взаимных и системных информаций о материальных свойствах объекта в блоке II.

Системно-структурные представления J об объекте переносятся на сам объект M . Этим и завершается этап чувственного восприятия, когда реальный материальный объект воспроизводится в нашей психике как идеальный объект J (восприятие).

Обратим внимание, что по описанию восприятие и представление ничем не отличаются друг от друга, но по содержанию им можно приписать разли-

чие, состоящее в том, что восприятие подразумевает сиюминутное отражение одновременно с ним существующего объекта, а представление может воспроизводиться по памяти, даже в отсутствие объекта, или даже отражать фантастические образы, которые, однако, строятся по общим с восприятием правилам.

Судя по рис. 1.3, чувственное отражение требует наличия двух образований: измерительного блока I (блока органов чувств) и коррелятора II (синтезатора), хотя и одновременных, но разрозненных ощущений. Первый, по всей видимости, присущ всей природе как живой, так и неживой. Второй же вероятно свойствен только всей живой природе, способной воспринимать как целое те или иные объекты.

Логическая информация (сущность) H в отличие от *J*, всегда относящейся к конкретным объектам, или свойствам, характеризует целый класс однородных в определенном отношении объектов или свойств, являясь семантическим синтезом законов логики, правил функционирования системы и ее элементов, образующих функционал ее существования.

Под логикой материального объекта в данном случае имеется в виду объективная логика природы, логика причинно-следственных связей источника и приемника информации.

Тогда закон логического отражения, олицетворяющий адекватность отражения в отсутствие априорного знания, можно записать следующим образом

$$\mathbf{E} = \mathbf{OR}(\mathbf{O}), \quad (1.4)$$

или, в линейном приближении $\mathbf{E} = R\mathbf{O}, \quad (1.5)$

$$R = R_k R_0, \quad (1.6)$$

$$\mathbf{E} = \mathbf{J}/M, \quad (1.7)$$

где \mathbf{E} – вектор интенсивности логики (напряженности поля логики); \mathbf{J} – вектор логики; R_0 – безразмерная константа, характеризующая логическую реакцию (поведение) отражающего объекта на поток \mathbf{O} чувственной информации об отражаемом объекте.

В общем случае R_0 зависит от \mathbf{O} , но всегда

$$\mathbf{E} = -\text{grad } H; \quad H = \int_r \mathbf{E} dr, \quad (1.8)$$

где H – потенциал поля (сущность воспринимаемой информации);

r – расстояние от объекта до изучаемой точки пространства в сферических координатах.

Закон логического отражения – это *вторая аксиома* излагаемой теории универсального моделирования (*отражения*).

Из определения вектора интенсивности логики (1.4) следует, что хотя материальные объекты различной природы в принципе получают одинаковый поток информации об отражаемом материальном свойстве, но их реакция на этот поток различна в зависимости от величины R_o , характеризующей природу соответствующего объекта. При прочих равных условиях различные объекты по-разному реагируют на один и тот же поток отражения.

Любое распределение информации на фоне наложенных на нее логических связей должно обладать определенным содержанием. При анализе тех или иных ситуаций мы нередко говорим о том, что они имеют больший или меньший *смысл* с точки зрения определенных целей. Тем самым мы признаем измеримость содержания, смысла ситуации, хотя и не имели до сих пор способа для соответствующих измерений. Концепция информационного поля позволяет найти количественную оценку *содержания, смысла* на основе прослеживания путей реализации логических связей. При этом «содержание» выступает как «смысл» взаимодействия неживых объектов в соответствии с «целями» законов природы.

Информационная сложность или содержание (смысл) C определяется пересечением (логическим произведением, а в частных случаях – декартовым произведением)

$$C = J \cap H \text{ или } C = J \times H. \quad (1.9)$$

В зависимости от того, применительно к характеристике всей системы или ее элементов используется C , можно говорить о *системной* C_s , *собственной* C_o и *взаимной* C_v сложности; при этом

$$C_s = C_o + C_v. \quad (1.10)$$

Для конструктивного использования понятий чувственная и логическая информация вводятся соответствующие детерминированные и статистические меры.

Измерение информационной проницаемости R_k и плотности информации \mathbf{O} не всегда возможно¹¹. Поэтому в теории **А. А. Денисова** предлагаются рассматриваемые в следующем разделе *детерминированные* и *вероятностные* меры, более удобные для практических приложений.

1.3. Подходы к измерению и оценке ценности информации

В различных науках об информации предпринимались попытки ее измерения.

Первое научное осмысление понятия *количество информации* началось в теории связи. Были введены меры информации.

В 1924 г. **Г. Найквист** показал, что скорость передачи сигналов W по каналу связана с числом n различных кодовых символов зависимостью

$$W = k \log n, \quad (1.11)$$

где k – константа, определяемая скоростью передачи последовательных символов. Он первым предложил логарифмическую меру информации.

В 1928 г. **Р. Хартли**¹² определил информацию сообщения как

$$H = m \log n, \quad (1.12)$$

где m – количество символов в сообщении, а n – количество доступных для использования различных символов.

С точки зрения современной теории информации эта мера применима только тогда, когда символы выбираются независимо один от другого, и выбор любого из символов равновероятен.

¹¹ Найдены способы измерения информации в физических полях, существуют попытки измерения биоинформации с помощью приборов на жидких кристаллах. Однако когда речь идет о социально-экономической, научно-технической информации, то проблема измерения чувственной информации усложняется.

¹² **Хартли Р.** Передача информации / Р. Хартли // Теория информации и ее приложения. – М.: 1959. – С. 5–35.

Наиболее полное развитие теории информации применительно к проблемам ее передачи было осуществлено в сороковые годы **К. Шенноном**¹³.

Шеннон связал понятие информации с понятием *энтропии* и предложил следующую энтропийную меру:

$$H = -\sum_{i=1}^n p_i \log p_i, \quad (1.13)$$

где p_i – вероятность того, будет или нет выбран i -й символ из полного набора в n символов, которые вырабатывает источник сообщений; величина H (энтропия) измеряется в битах (от англ. *binary digit* – двоичная система).

Работы **К. Шеннона** и его последователей нашли широкое применение на практике – при оптимизации каналов связи, в системах обработки данных, при разработке электронно-вычислительной техники и т. п.

Важные аспекты теории информации применительно к теории связи были решены **А. Н. Колмогоровым**, **Н. Винером**.

В частности, они независимо решили задачу *фильтрации* (выделения полезного сигнала из комбинации «сигнал–шум» при заданных статистических характеристиках сигнала) и *предсказания* значения полезного сигнала. Эта задача находит широкое использование при проектировании непрерывных динамических систем, в системах управления технологическими процессами.

С развитием кибернетики и теории систем понятие информации развивалось.

Для оценки возможностей различных технических средств сбора, передачи и обработки информации были предложены: теоретико-игровой подход к оценке информации **Р. Л. Стратоновича**¹⁴ алгоритмический подход **М. М. Бонгарда**¹⁵.

¹³ **Шеннон К. Э.** Работы по теории информации и кибернетике: сб. статей / К. Э. Шеннон. Пер. с англ. с предисловием А. Н. Колмогорова / Под ред Р. Л. Добрушина и О. Б. Лупанова. – М.: Иностр. лит., 1963. – 829 с.

¹⁴ **Стратонович Р. Л.** О ценности информации / Р. Л. Стратонович // Изв. АН СССР: Техническая кибернетика. – 1965. – № 5. – С. 25–38.

¹⁵ **Бонгард М. М.** О понятии «полезность информации» / М. М. Бонгард // Проблемы кибернетики: сб. – М.: Физматгиз, 1963. – Вып. 9. – С. 3–23.

В дальнейшем информацию стали связывать не только с *источником* и *приемником* сигналов, но и с ее *значением* для потребителя.

Вначале имелись попытки применения рассмотренных выше мер и для оценки ценности социальной информации. Однако при этом возникают трудности доказательства адекватности таких мер. Предпринимались также попытки создать *семантическую* теорию информации. Из них наиболее известной является концепция *Р. Карнапа* и *И. Бар-Хиллела*¹⁶, базирующаяся на понятии *логической вероятности* как степени подтверждения гипотезы. В соответствии с этой концепцией наивысшей ценностью обладают гипотезы, подтвержденные достоверным знанием, экспериментом. В этом случае логическая информация приравнивается единице, а семантическая – нулю. С уменьшением степени подтверждения гипотезы количество содержащейся в ней семантической информации увеличивается. Следует отметить, что при этом гипотезы формулируются на специальном языке, предложенном в рассматриваемой теории.

Развивая концепцию Карнапа–Бар-Хиллела, *Л. Бриллюэн*¹⁷ предложил разновидность статистической меры для измерения семантической информации, основанную на измерении и уменьшении неопределенности (энтропии). Опираясь на второе начало термодинамики (принцип Карно), *Л. Бриллюэн* вводит для измерения информации следующую меру снижения энтропии:

$$I_1 = K \ln(P_0/P_1) = K \ln P_0 - R \ln P_1, \quad (1.14)$$

где P_0 – число равновероятных исходов, возникновение которых *априори* считается равновероятным; при этом нет исходной информации о задаче, т. е. априорная информация $I_0 = 0$; P_1 – число равновероятных исходов при наличии информации об измерениях

¹⁶ *Bar-Hillel Y.* Semantic Information / Y. Bar-Hillel, R. Carnap // British Journal of the Philosophy of Science, 1953. – V. 4. – № 14.

¹⁷ *Бриллюэн Л.* Наука и теория информация / Л. Бриллюэн. – М.: Физматлит, 1960. – 392 с.

в аналогичных, родственных ситуациях, и тогда информация, снижающая неопределенность, равна I_1 .

Бриллюэн по аналогии с термодинамикой использовал натуральный логарифм, а для того, чтобы выбрать систему единиц ввел коэффициент K , который для приведения к принятым к тому времени битам равен

$$K = 1/\ln 2.$$

А. А. Харкевич¹⁸ связал ценность информации с целью деятельности, предложив рассматривать энтропийную меру Шеннона как меру вероятности непопадания в цель, т. е. как меру целесообразности.

Ю. А. Шрейдер предложил построить теорию семантической информации на основе концепции разнообразия, а не концепции снятия неопределенности, и в частности, на основе учета такого свойства информации как *зависимость получаемой информации от априорной информации*¹⁹. Опираясь на идею **Н. Винера** о том, что для понимания и использования информации получатель должен обладать определенным запасом знаний, т. е. в терминах математической лингвистики и теории языков – *тезаурусом Θ* , **Ю. А. Шрейдер** определяет количество семантической информации, содержащейся в тексте T , как изменение тезауруса $\Theta I(T, \Theta)$.

Такой подход существенно отличается от концепции выбора при статистическом подходе, при котором предполагается, что получаемая информация тем больше, чем меньше априорных сведений содержится в приемнике информации. Напротив, по Шрейдеру, чем более сложную структуру имеет тезаурус, тем больше существует возможностей для изменений под воздействием одного и того же сообщения. Это хорошо согласуется с законом необходимого разнообразия **У. Р. Эшби**, согласно которому управляющая (осмысливающая, понимающая, принимающая решения) система должна обладать большим необхо-

¹⁸ **Харкевич А. А.** О ценности информации / А. А. Харкевич // Проблемы кибернетики: сб. – Вып 4. – М.: Физматгиз, 1960. – С. 33–41.

¹⁹ **Шрейдер Ю. А.** О семантических аспектах теории информации / Ю. А. Шрейдер // Информация и кибернетика: сб. – М.: Изд-во «Сов. Радио», 1967. – С. 22–28.

димым разнообразием (сложностью) по сравнению с разнообразием поступающей в нее информации от управляемой (понимаемой) системы.

Предлагаемый **Ю. А. Шрейдером** подход согласуется также с наблюдениями за процессами обмена информацией: по мере роста наших знаний об изучаемом объекте растет и количество извлекаемой информации об этом объекте. Динамику насыщения индивидуального тезауруса Ю. А. Шрейдер иллюстрирует некоторой условной зависимостью (рис. 1.4), характер которой зависит от конкретного потребителя информации или накапливающей информацию информационной системы.

Эта кривая характеризует одну из особенностей информации – *зависимость получаемой информации от априорной информации.*

Для количественной оценки ценности семантической информации предлагались различные меры.

Так, в [9] предлагается мера для общечеловеческого тезауруса, под которым понимается совокупность понятий, вырабатываемых человечеством в процессе научного познания природы, общества и мышления, связанных друг с другом.

Процесс научного познания представляется как уточнение старых и формирование новых научных понятий, в выявлении и корректировке связей между понятиями, т. е. тезаурус трактуется как определенная структура, которая изменяется под воздействием новой информации в процессе познания и предлагается гипотетический подход, в соответствии с которым информация не одинаково зависит от сложности структуры.

Если под воздействием сообщения внутренняя структура общечеловеческого тезауруса Θ становится менее сложной, чем до этого сообщения, то в таком сообщении содержится больше информации, чем в сообщении, вызывающем усложнение структуры Θ . Этот вывод согласуется с опытом: известно, что крупные научные открытия в целом упрощают структуру зна-

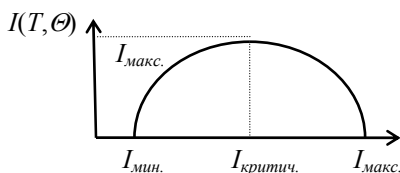


Рис. 1.4. Зависимость информации от априорной

ния – благодаря введению новых, более общих понятий, что и является следствием и проявлением свойства *кумулятивности* информации.

В [9] оговаривается гипотетичность предлагаемого подхода к оценке ценности информации в зависимости от структуры представления общечеловеческого тезауруса, и высказывается идея о необходимости поиска иных мер оценки ценности информации.

Для оценки удовлетворения информационных потребностей в теории научно-технической информации введены меры *релевантности* и *пертинентности* (см. гл. 5).

Затем, когда стали осознавать, что существуют различные виды информации, **Е. К. Войшвилло**²⁰ предложил различать информацию *восприятия*, измеряемую в знаках, и *информационно-значение*, смысловое содержание для потребителя информации – означаемое знака, для оценки которой вначале не было мер.

В дальнейшем, как было отмечено выше, информацию стали рассматривать в нескольких аспектах с точки зрения потребителя информации, пользователя: *прагматический* – для достижения целей пользователя; *семантический* – с точки зрения смыслового содержания и правильности толкования; *синтактический* (знаковый, технологический) – с точки зрения техники передачи информации. При этом считали, что либо наиболее общим является понятие *прагматической информации*, а *семантический* и *синтактический* аспекты информации имеют подчиненное значение, либо эти аспекты можно считать *уровнями* существования информации, и тогда – вводить для разных уровней разные оценки (иногда с достаточно экзотическими названиями).

Так, в системе мер оценки экономической информации, изложенной в работах **Ю. И. Черняка** предлагается различать несколько аспектов (уровней) представления и измерения экономической информации (с точки зрения ее *полезности* для решения задачи, *смысла* или *семантики* текста, *синтактики* знакового отображения, *морфологии* образования знаков – слов и словосочетаний, и передачи их по каналам связи) и для каждого

²⁰ **Войшвилло Е. К.** Понятие / Е. К. Войшвилло. – М.: МГУ, 1967. – 286 с.; **Войшвилло Е. К.** Символическая логика: Классическая и релевантная: Философско-методологические аспекты: учеб. пособие / Е. К. Войшвилло. – М.: Высшая школа, 1989. – 149 с.

уровня вводятся свои меры информации в терминах, принимаемых на этом уровне решений (прагма, сем, знак или символ и т.п.)²¹.

Развивалось понятие *информации* и как философской категории. В частности, **А. Д. Урсул**²² стал рассматривать информацию как форму существования материи, подобно массе и энергии, и тогда возникает потребность во введении соответствующих количественных мер.

В отличие от этой точки зрения **А. А. Денисов** в теории информационного поля и ее дискретном варианте, известном как информационный подход к анализу систем [1, 6, 13], рассматривает, как было сказано в предыдущем разделе, информацию как категорию, парную по отношению к материи, и вводит понятия *чувственной J* и *логической H* информации, которые для удобства практических приложений названы *информацией восприятия* и *информационным потенциалом*.

При этом **А. А. Денисов** показал, что *J* и *H* – не отдельно существующие виды информации (как рассматривалось в предыдущих теориях), а дополняющие друг друга понятия: *информацией* для потребителя является пересечение информации восприятия или *чувственной информации* и ее потенциала или *логической информации*, в результате чего формируется единое понятие – *информационный смысл, информационная сложность*, которая в частных случаях представляет собой декартово произведение) *J* и *H* (1.9).

Для конструктивного использования понятий чувственная и логическая информация **А. А. Денисов** ввел соответствующие детерминированные и статистические меры (табл. 1.3).

²¹ **Черняк Ю. И.** Информация и управление / Ю. И. Черняк. – М.: Наука, 1974. – 184 с.; **Черняк Ю. И.** Система мер информации как инструмент оптимизации языка управления / Ю. И. Черняк, А. И. Чухнов, Ю. Ю. Финкельштейн // Языки экономического управления и проектирования систем. – М.: Наука, 1973. – С. 24–27.

²² **Урсул А. Д.** Информация / А. Д. Урсул. – М.: Наука, 1971. – 295 с.; **Урсул А. Д.** Отражение и информация / А. Д. Урсул. – М.: Мысль, 1973. – 231 с.

Таблица 1.3

Меры информации А.А. Денисова

<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="border: 1px solid black; border-radius: 50%; padding: 10px; text-align: center;"> Информация восприятия (чувственная информация) J (элементная база) </div> <div style="border: 1px solid black; border-radius: 50%; padding: 10px; text-align: center;"> C </div> <div style="border: 1px solid black; border-radius: 50%; padding: 10px; text-align: center;"> Логическая информация H (информационный потенциал) (сущность) </div> </div>			
Детерминированный способ измерения	$J_i = A_i / \Delta A_i,$ где A_i – значение измеряемой величины; ΔA_i – «квант», с точностью до которого ЛПП интересуется воспринимаемая информация (единица измерения, разрешающая способность прибора)	$H = \gamma \sqrt[\gamma]{\frac{1}{n} \sum_{i=1}^n J_i^\gamma},$ где J_i – результаты измерения A_i ; n – объем понятия об охватываемых измерениями объектах; γ – параметр усреднения	При $\gamma = 1$ получим среднее арифметическое $H = \sum_{i=1}^n \frac{1}{n} J_i.$ При $\gamma = 0$ получим среднее геометрическое $H = \sqrt[n]{\prod_{i=1}^n J_i}.$ При $\gamma = -1$ – среднее гармоническое $H = n / \sum_{i=1}^n \frac{1}{J_i}$
Вероятностный способ измерения	$J_i = -\log_2 p_i,$ где p_i – вероятность события. В случае использования информации для достижения цели p_i называют вероятностью недостижения цели или степенью нецелесоответствия	$H = \int f(J_i) dJ_i \Rightarrow$ $H = \sum_{i=1}^n q_i J_i = -\sum_{i=1}^n q_i \log p_i,$ где q_i – вероятность использования элемента информации. При $q_i = p_i$ $H = -\sum_{i=1}^n p_i \log p_i.$ При равновероятном выборе элемента $p_i = 1/n$ и $H = -\sum_{i=1}^n \frac{1}{n} \log \frac{1}{n} = \log n$ Для прагматической информации $H_y = -\sum_{i=1}^n q_i \log(1 - p_i'),$ где p_i' – вероятность достижения цели, степень целесообразности	

Дискретное представление теории помогло ввести детерминированные меры.

Чувственная информация J вводится как мера отраженной в нашем сознании объективной реальности, элементной базы системы в форме

$$J = A/\Delta A, \quad (1.15)$$

где A – общее количество каких-либо знаков, воспринимаемых измерительными приборами или нашими органами чувств, ΔA – «квант», с точностью до которого нас интересует воспринимаемая информация, или разрешающая способность прибора.

Действительно, во-первых, J существенно зависит от разрешающей способности выбранного измерительного прибора и во-вторых, определяется значимостью для нас измеряемой величины, т. е. целью измерений.

Так, наличие или отсутствие личного автомобиля у конкретного гражданина фиксируется с точностью до $\Delta A = 1$ автомобиль, поскольку это обстоятельство для автолюбителя является весьма существенным. Напротив, сводные данные по производству автомобилей в стране приводятся с точностью до $\Delta A = 10\,000$ автомобилей, поскольку меньшее число автомобилей для страны в целом несущественно.

Аналогично длину отрезка можно измерять с точностью до 1 м, до 1 дм, до 1 см, и соответственно получать разные значения J .

Таким образом, в прагматическом аспекте информация всегда несет в себе весьма значительный элемент субъективности и различна для разных людей при одном и том же A .

Логическая информация (сущность) H в отличие от J , всегда относящейся к конкретным объектам, или свойствам, характеризует целый класс однородных в определенном отношении объектов или свойств, являясь семантическим синтезом законов логики, правил функционирования системы и ее элементов, образующих функционал ее существования.

Согласно основному закону классической логики *Аристотеля* *собственная сущность (суть) системы обратна объему понятия n о ней*, т. е.

$$H = J/n. \quad (1.16)$$

Объем понятия зависит от аспекта рассмотрения системы (элемента) и обычно предполагает родовую их принадлежность.

Например, объемом понятия «производственное предприятие» является общее количество вообще всех производственных предприятий в городе (области, стране), а объем понятия «это производственное предприятие» равен единице.

Если система характеризуется множеством своих состояний, подобно рабочей неделе, которая состоит из понедельника, вторника, среды и т. д., то это множество и составляет объем понятия «неделя», который равен $n = 7$, а сущность трех дней недели $H = 3/7$.

Способ опосредования (усреднения) J может быть и иным, для чего вводится параметр γ , который может выбирать постановщик задачи. Тогда

$$H = \gamma \sqrt{\frac{1}{n} \sum_{i=1}^n J_i \gamma}, \quad (1.17)$$

где J_i – результаты измерения A_i согласно (1.15); n – объем понятия, т. е. число, охватываемых понятием объектов; γ – параметр логики усреднения, при различных значениях которого получаются различные выражения для определения H , приведенные в табл. 1 (в таблице знак \prod – знак произведения).

Обратим внимание на тот факт, что рассматриваемый закон в основе своей непрерывен (в силу обратной пропорциональности n) и его действие должно распространяться на непрерывную (многозначную) логику.

Конструируя сущность понятия H , *А.А. Денисов* вначале пользовался аналогиями теории цепей, а затем сделал доклад на семинаре *В.С. Тюхтина* в Институте философии АН СССР, и Виктор Степанович увидел в модели *А.А. Денисова* возможность вывода его символических представлений из последовательности требований, предъявляемых законами диалектики применительно к суждению.

В результате получилась модель, отображающая процесс отражения, восприятия и осмысления информации с учетом вре-

мени реакции и ригидности (сопротивляемости новой информации) воспринимающего субъекта ²³.

Логическую информацию H можно определить не только через параметры синтезирующей ее системы (человека, автоматизированной информационной системы).

Если учесть, что, как было отмечено ранее, H характеризует не единичный объект, а класс однородных в определенном смысле объектов или свойств, то H можно определить через плотность вероятности $f(J_i)$ того, что J имеет значение J_i

$$H = \int f(J_i) dJ_i \quad (1.18)$$

В частном случае вместо плотности вероятности можно охарактеризовать класс однородных объектов просто вероятностью q_i и представить J_i в логарифмической форме; тогда получим

$$H = - \sum_{i=1}^n q_i \log p_i \cdot \quad (1.18, a)$$

Значения q_i и p_i могут быть не равны, но возможны ситуации, когда $q_i = p_i$, что имеет место в формуле **К. Шеннона**

$$H = - \sum_{i=1}^n p_i \log p_i \quad (1.18, б)$$

Прагматическая (целевая) информация H_u описывается моделью, аналогичной (12,б), только для практических приложений удобнее заменить вероятность недостижения цели p_i на сопряженную $(1 - p_i')$

$$H_u = - \sum_{i=1}^n q_i \log(1 - p_i') \quad (1.18, в)$$

²³ **Денисов А.А.** Теория больших систем управления: Учеб. пособие для студентов вузов / А.А. Денисов, Д.Н. Колесников. – Л.: Энергоиздат, 1982. – 288 с.; **Денисов А.А.** Современные проблемы системного анализа: Информационные основы / А.А. Денисов. – СПб.: Изд-во Политехн. ун-та, 2005. – 296 с. И др.

где p_i' – вероятность достижения цели; q_i – вероятность того, что оцениваемая компонента будет использована для достижения цели.

При вероятностном способе измерения

$$J = -\log_2 p \quad (1.19)$$

Таким образом, из вышесказанного следует, что J и H могут измеряться различными способами – детерминированно и с помощью вероятностных характеристик.

Поскольку в некоторых приложениях могут быть использованы одновременно обе формы представления информационных характеристик – и детерминированная, и вероятностная, а также – переход от одной формы к другой, то удобно пользоваться сопоставительной табл. 1.3, в которой приведены основные способы измерения J и H .

В концепции *А. А. Денисова* учитывается, что далеко не всегда может быть строго определено вероятностное пространство. В этом случае можно использовать понятие размытой вероятности в смысле Заде. Поэтому предложена информационная мера, учитывающая степень влияния информации как ресурса на реализацию целей (1.18, *а*).

Эта мера позволяет учесть не только вероятность p_i , влияющую на достижение (или недостижение) цели, но и вероятность q_i того, что данная информация будет использована лицом, принимающим решение.

Мера *К. Шеннона* получается из соотношения (2.5) при $q_i = p_i$ как частный случай.

Существуют и другие меры количества информации, учитывающие процесс ее передачи и изменение вероятности появления сообщения от априорного значения $p(x)$ на входе канала передачи информации к апостериорному значению $p(x/y)$ на выходе канала, связанные с искажением информации в канале, т. е. основанные на условной вероятности [11, с. 16–21].

1.4. Понятие об информатике как основе теории информационных систем и комплексов

Из истории развития термина «информатика»²⁴

Термин «**информатика**» использовался первоначально в зарубежных научных публикациях (нем. *Informatik*, фр. *Informatique*, англ. *informatics*) для названия научно-практического направления, занимающегося автоматизированной обработкой данных.

Этот термин впервые был введен в 1957 г. в Германии **К. Штейнбухом** (K. Steinbuch)²⁵. Затем – в 1962 г. во Франции **Ф. Дрейфузом** (F. Dreyfus)²⁶, который образовал его как слияние французских слов *information* и *automatique*. В том же 1962 г. этот термин использовал и в США **У. Ф. Бауэр**²⁷ для названия фирмы, занимающейся автоматизированной переработкой информации.

В отечественной научной литературе термин «информатика» был первоначально использован в 1963 г. профессором Московского энергетического института **Ф. Е. Темниковым**²⁸, который

²⁴ При подготовке данного раздела использовались следующие работы: **Колин К. К.** Теоретические проблемы информатики. Т. 1. Актуальные философские проблемы информатики / под общ. ред. К. И. Курбакова. – М.: КОС•ИНФ, 2009. – 222 с.; **Черный Ю. Ю.** Полисемия в науке: когда она вредна? (на примере информатики) / Ю. Ю. Черный // Открытое образование = Open education. – М., 2010. – № 6. – С. 97-107 [электронный ресурс]. – Режим доступа: http://www.e-joe.ru/i-joe/i-joe_01/files/chorniy.pdf.

²⁵ **Steinbuch K.** Informatik: Automatische Informationsverarbeitung / K. Steinbuch // SEG-Nachrichten (Technische Mitteilungen der Standard Elektrik Gruppe). – Berlin, 1957. Nr. 4. S. 171.

²⁶ **Dreyfus Ph.** L'informatique / Ph.Dreyfus // Gestion, 1962. Vol. 5. June. P. 240–241.

²⁷ Informatics and (et) Informatique. IEEE Annals of the History of Computing, 18(2). Available at: <http://www.softwarehistory.org/history/Bauer1.html>.

²⁸ **Темников Ф. Е.** Информатика / Ф. Е. Темников // Известия ВУЗов. Электромеханика, 1963. – № 11. – С. 1277.

определил информатику как науку об информации вообще, состоящую из 3-х основных частей:

- 1) теория информационных элементов;
- 2) теория информационных процессов;
- 3) теория информационных систем.

В журнале «Известия вузов: Электромеханика», № 11, 1963 г.] Федор Евгеньевич опубликовал всего одну страничку, начав ее фразой. *«Давно ощущается потребность в интегральной научной дисциплине, связывающей воедино многочисленные вопросы сбора, передачи, обращения, переработки и использования информации»* и предложил программу создания такой дисциплины, *«... могущей послужить важным теоретическим стержнем автоматике, телемеханики, измерительной и вычислительной техники, связи и радиолокации, бионики и кибернетики»*. Проект программы Ф.Е. Темников представил в виде таблицы со столбцами *«Теория информационных элементов»*, *«Теория информационных процессов»* и *«Теория информационных систем»*.

В таком значении термин использовался в начальный период создания автоматизированных информационных систем в технической литературе. Однако в последующем это определение долгое время оставалось лишь историческим фактом, и не было оценено должным образом. Возможно, потому что было опубликовано только в специальном журнале «Известия вузов: Электромеханика».

В 1966 г. независимо термин «информатика» был введен вместо термина «теория научной информации» директором Всесоюзного института научной и технической информации (ВИНИТИ) АН СССР профессором **А. И. Михайловым** и научными сотрудниками ВИНТИ **А. И. Черным** и **Р. С. Гиляревским**. Информатикой была названа *«... научная дисциплина, изучающая структуру (не конкретное содержание) и свойства научной информации, а также закономерности научно-информационной деятельности, ее теорию, историю, методику*

и организацию»²⁹. Появлению этого термина предшествовала рекомендация директора Института проблем передачи информации (ИППИ) АН СССР чл.-корреспондента **А. А. Харкевича**, содержащаяся в его отзыве 11 октября 1962 на подготовленную коллективом сотрудников ВИНТИ Проблемную записку «Научная информация»³⁰ (Вопросы советской науки)».

В таком значении в советской научно-технической литературе термин «информатика» стал широко известен. Благодаря работам **А. И. Михайлова**, **А. И. Черного** и **Р. С. Гиляревского** он использовался на протяжении достаточно длительного времени в сфере бурно развивающегося в тот период направления «Научно-техническая информация» [14, 15, 24 и др.].

В последующем термин «информатика» стал употребляться в нашей стране в более узком смысле применительно к техническим и программным средствам хранения и обработки данных на электронно-вычислительных машинах. Новое значение соответствовало немецкому и французскому, а в США и Великобритании применялся термин *computer science* – *компьютерная наука*.

В 1978 г. Международным конгрессом в Японии было принято следующее определение: *«Понятие информатики охватывает области, связанные с разработкой, созданием, использованием и материально-техническим обслуживанием системы обработки информации, включая машины, оборудование, математическое обеспечение, организационные аспекты, а так-*

²⁹ **Михайлов А. И.** Информатика – новое название теории научной информации / А. И. Михайлов, А. И. Черный, Р. С. Гиляревский // Научно-техническая информация. – М., 1966. – № 12. – С. 35–39.

³⁰ **Черный А. И.** Всероссийский институт научной и технической информации: 50 лет служения науке / А. И. Черный. – М.: ВИНТИ, 2005. – С. 190.

же комплекс промышленного, коммерческого, административного, социального и политического воздействия»³¹.

Академик **А. А. Дородницын**³² определил информатику как науку о преобразовании информации, которая базируется на вычислительной технике.

В таком понимании информатика включает дисциплины, относящиеся к обработке информации в вычислительных машинах и вычислительных сетях: как абстрактные, вроде анализа алгоритмов, так и довольно конкретные, например, разработка языков программирования.

Широкому распространению термина в таком значении способствовал академик **А. П. Еришов**, основавший школьную информатику, обеспечившую широкое распространение компьютерной грамотности в школе³³.

Термин оказался удобным как краткое название курса о применении ЭВМ для обработки данных в учебном процессе школы, поскольку школьникам сложно объяснить понятия информационных процессов, информационных систем, поиск, хранение и обработку текстовой информации.

Однако такое сужение смысла термина «информатика» нежелательно в учебном процессе вуза, поскольку уже достаточно давно осознано, что информация – важнейший ресурс социально-экономических организаций, обеспечивающий их развитие, и нужно уметь оценивать содержание, смысл информации как интеллектуального ресурса.

Это осознавалось постепенно.

³¹ **Беликов Е. П.** Об организации в Академии наук СССР работ по информатике, вычислительной технике и автоматизации / Е. П. Беликов // Вестник АН СССР. – 1983. – № 6.

³² **Дородницын А. А.** Информатика: предмет и задачи / А. А. Дородницын // Кибернетика. Становление информатики: сб. – М.: Наука, 1986.

³³ **Деятели** компьютерики и информатики: биографический справочник / Т. М. Аскеров, Р. С. Гиляревский, К. И. Купбаков. – М., 2006ю - С. 11.

В 1985 г. академик **А. А. Самарский** обращает внимание на новую научную методологию, возникшую благодаря информатике: «Она основана на развитии в широком применении методов математического моделирования и вычислительного эксперимента и служит ближайшим стратегическим резервом ускорения научно-технического прогресса. Сущность математического моделирования и его главное преимущество состоит в замене исходного объекта соответствующей математической моделью и в дальнейшем ее изучении (экспериментирование с нею) на ЭВМ с помощью вычислительно-логических алгоритмов. Математическое моделирование представляет собой естественное развитие и обобщение методов научного исследования, соединенных с современной информационной технологией. Цикл вычислительного эксперимента объект – модель – алгоритм – программа – ЭВМ – управление объектом отражает основные этапы процесса познания в нынешнем компьютерном воплощении. Здесь органично соединяются сильные стороны теоретических методов и натурального эксперимента. Работа с моделью, а не с объектом, оборачивается оперативным получением подробной и наглядной информации, вскрывающей его внутренние связи, качественные характеристики и количественные параметры. Многократно уменьшаются материальные и трудовые затраты, присущие традиционным экспериментальным подходам, дающим, как правило, лишь крупицы нужной информации. Вычислительный эксперимент не подвластен каким-либо ограничениям – математическая модель может быть безопасно испытана в любых мыслимых и немыслимых условиях»³⁴.

Академик **Н. Н. Моисеев** считает, что «информатика – это некая синтетическая дисциплина, которая включает в себя и разработку новой технологии научных исследований и проектирования, основанные на использовании электронной вычисли-

³⁴ Самарский А. А. Проблема использования вычислительной техники и развитие информатики / А. А. Самарский // Вестник АН СССР, 1985. – № 3.

тельной техники, и несколько крупных научных дисциплин, связанных с проблемой общения с машиной, и, наконец, с созданием машины»³⁵.

В ряде определений информатики на первое место ставили собственно *информацию*.

Директор Института проблем передачи информации АН СССР **В. И. Сифоров** считал, что: «*В основу определения должны быть положены действия над информацией. Информатика развивается под действием потребности общества и согласно внутренней логике развития. В основе этого развития лежат закономерности процессов в ЭВМ, закономерности развития ЭВМ. Информатика имеет дело не с конкретными формами материи, а с категориями: информация, модель и т. п. Информатика – комплексная дисциплина – это наука (фундаментальные исследования) и отрасль производства (опытно-конструкторские работы и совершенствование технологий), а кроме того, и инфраструктурная область (эксплуатация информационных систем)*»³⁶.

Заместитель директора ВИНТИ в 1970-е гг. **Ю.И. Шемакин** в книге «Введение в информатику»³⁷ отмечает, что «*основной задачей информатики является изучение закономерностей, в соответствии с которыми происходят создание, преобразование, хранение, передача и использование информации всех видов, в том числе с применением современных технических средств*».

³⁵ **Моисеев Н. Н.** Информатика: новые пути познания законов природы и общества / Н. Н. Моисеев // Вестник АН СССР 1985. – № 5.

³⁶ **Сифоров В. И.** Информатика и ее взаимодействие с философией и другими науками / В. И. Сифоров // Философская наука, 1984. – № 2.

³⁷ **Шемакин Ю. И.** Введение в информатику / Ю. И. Шемакин. – М.: Финансы и статистика, 1985.

Академик **Б. Н. Наумов** в предисловии к сборнику ³⁸ подчеркивает, что информатика – это *«естественная наука, изучающая общие свойства информации, процессы, методы и средства ее автоматизированной обработки»*. При этом под обработкой информации понимаются процессы ее восприятия, хранения, преобразования, перемещения и вывода (ввода) с применением средств вычислительной техники».

В документах ЮНЕСКО 1986–1988 гг. термину «информатика» дается широкое толкование ³⁹. Утверждается, что этот термин охватывает собственно информацию, ее сбор, анализ и обработку, а также соответствующие аппаратные средства, включая микропроцессоры или другие электронные системы. Информатика рассматривается как крупное научное направление, заслуживающее активного развития в интересах всего человечества, которое способно (при соответствующем освоении ее методов и средств) помочь человеку полнее использовать информационные ресурсы в интересах научно-технического прогресса и социального развития.

В 1988г. академик **А. П. Ершов** в Математическом энциклопедическом словаре ⁴⁰ дал следующее определение информатики как науки, отрасли промышленности и разновидности человеческой деятельности: *«Информатика – 1) находящаяся в становлении наука, изучающая законы и методы накопления, передачи и обработки информации с помощью ЭВМ; 2) родовое понятие, охватывающее все виды человеческой деятельности, связанные с применением ЭВМ»*.

³⁸ *Информатика* и компьютерная грамотность. ИПИ АН СССР / Отв. ред. академик Б. Н. Наумов. – М.: Наука, 1988.

³⁹ **Свириденко С. С.** Современные информационные технологии / С.С. Свириденко. – М.: Радио и связь, 1989.

⁴⁰ **Математический** энциклопедический словарь / Гл. ред. Ю. В. Прохоров. – М.: Сов. энциклопедия, 1988.

В последнее время некоторыми авторами, в связи с определением современных задач информатики, особый упор делается на обработку знаний.

Роль теории искусственного интеллекта в информатике была обоснована академиком **Г. С. Поспеловым**⁴¹.

В. Д. Ильин предложил дать следующее определение информатики: «*Предметом информатики как науки будем считать процесс создания, накопления и применения знаний*»⁴².

К. К. Колин считаете: «*Информатика является общенаучной дисциплиной, которая изучает свойства, закономерности, процессы, методы и средства формирования, хранения и распространения знаний в природе и обществе*»⁴³, и в 2000 г. издал учебное пособие⁴⁴, в котором развивает представление об информатике как междисциплинарной науке о закономерностях и формах движения информации в природе и обществе.

В некоторых работах⁴⁵ на основе сопоставления различных определений информатики предлагается рассматривать ее как науку о формализованном общении.

Во многих учебниках и учебных пособиях информатика трактуется как наука о законах и методах получения и измерения, накопления и хранения, переработки и передачи информации с применением математических и технических средств. И в качестве краткого определения: используется исходное опреде-

⁴¹ **Поспелов Г. С.** Искусственный интеллект – новая информационная технология / Г. С. Поспелов. – М.: Наука, 1988.

⁴² **Ильин В. Д.** Система порождения программ / В. Д. Ильин. – М.: Наука, 1989.

⁴³ **Колин К. К.** О структуре научных исследований по комплексной программе «Информатика» / К. К. Колин // Науч. труды «Социальная информатика»: сб. – М.: КВШ при ЦК ВЛКСМ, 1990.

⁴⁴ **Колин К. К.** Фундаментальные основы информатики: Социальная информатика / К. К. Колин. – М.: Академический проект; Екатеринбург: Деловая книга, 2000. – 350 с.

⁴⁵ **Белошапка В.О.** О языках моделей и информатике / В.О. Белошапка // Информатика и образование. – 1987. – № 6.

ление **Ф. Дрейфуса** – это *информация* плюс *автоматика*, что сужает представление о современной информатике.

В силу ряда объективных причин в учебном процессе вузов практически не используется понятие информатики в смысле **А. И. Михайлова – А. И. Черного – Р. С. Гиляревского**, предложенного в 1960-е гг. для названия науки о научно-технической информации.

Очевидно, что важная функция информатики состоит в разработке методов и средств преобразования информации с использованием компьютера, а также в применении их при организации технологического процесса преобразования информации. В то же время, выполняя эту функцию, информатика как прикладная наука должна решать более широкий спектр задач исследовать информационные процессы в технических и социально-экономических системах; разрабатывать и/или адаптировать технические средства и создавать новые технологии для преобразования информации на основе результатов, полученных в ходе исследования информационных процессов; решать научные и инженерные проблемы создания, внедрения и обеспечения эффективного использования компьютерной техники и технологии во всех сферах человеческой деятельности.

При этом информатика исследует разнородные аспекты:

- *технические*, связанные с изучением методов и средств надежного сбора, хранения, передачи, обработки и выдачи информации;
- *синтаксические*, связанные с решением задач по формализации и автоматизации некоторых видов научно-информационной деятельности, в частности индексирование, автоматическое реферирование, машинный перевод;
- *семантические*, определяющие способы описания смысла информации, изучающие языки ее описания;
- *прагматические*, т.е. методы анализа и преобразования информации для решения конкретных прикладных задач.

На основе обобщения различных точек зрения в современном представлении можно дать следующее обобщающее определение: *информатика – научная дисциплина, изучающая структуру и общие свойства информации, закономерности процессов обмена информацией от непосредственного устного и письменного общения специалистов до формальных процессов обмена посредством различных носителей информации. Важной сферой информатики является научно-информационная деятельность по сбору, переработке, хранению, поиску и распространению научно-технической информации.*

Структуризация информатики

Идеи современной информатики возникали в недрах математики и кибернетики, электроники и системотехники, логики и лингвистики.

Информатика обслуживает науку, технику, производство и другие виды человеческой деятельности на основе информационных технологий. Расширение сферы научных методов и практических приложений информатики привело к необходимости ее структуризации. По поводу структуры информатики существуют различные точки зрения.

В 1991 г. научные сотрудники Института проблем информатики АН СССР **И. А. Мизин** и др. рассматривали информатику, развивая приведенное выше определение **В. И. Сифорова**, как комплексную дисциплину⁴⁶:

во-первых, это *естественная наука* (фундаментальные и прикладные исследования);

во-вторых, *отрасль промышленности* (опытно-конструкторские работы и производство);

⁴⁶ **Развитие** определений «информатик» и «информационные технологии»: Препринт Института проблем информатики АН СССР // И.А. Мизин, И. Н. Сеницын, Б. Г. Доступов, В. Н. Захаров, А. Н. Красавин; под ред. чл.-корреспондента АН СССР И.А. Мизина. – М.: ИПИ АН СССР, 1991.

в-третьих, *инфраструктурная область* (профессиональная деятельность и эксплуатация систем информатизации).

Как *естественная наука* информатика изучает общие свойства информации (данных и знаний), методы и системы для ее создания, накопления, обработки, хранения, передачи и распределения с помощью средств вычислительной техники и связи. Как *фундаментальная наука* информатика связана с философией – через учение об информации и теорию познания; с математикой – через теорию математического моделирования; математическую логику и теорию алгоритмов; с лингвистикой – через учение о знаковых системах (семиотику) и формальных языках. Она также тесно связана с теорией информации и управления.

Как *отрасль промышленности* информатика занимается проектированием, изготовлением, сбытом и развитием систем информатизации и их компонентов.

Как *инфраструктурная область* информатика занимается сервисом и эксплуатацией систем информатизации, обучением и др.

В ряде работ принята следующая структуризация: *теоретическая* информатика, *техническая* информатика; *прикладная* информатика.

Подобным образом, но в несколько другой терминологии структурируется информатика в учебнике **Н. В. Макаровой** и **В. Б. Волкова**⁴⁷ (рис. 1.5), в котором с информатикой связывают одно из следующих понятий: это либо *отрасль производства*, либо *фундаментальная наука*, либо *прикладная дисциплина*. И эта точка зрения широко представлена на различных сайтах в сети Интернет.

При этом обращается внимание на то, что все три направления информатики опираются на средства для преобразования информации и развивают их.

⁴⁷ **Макарова Н. В.** Информатика: учебник для вузов / Н. В. Макарова, В. Б. Волков. – СПб.: Питер, 2011. – 576 с.

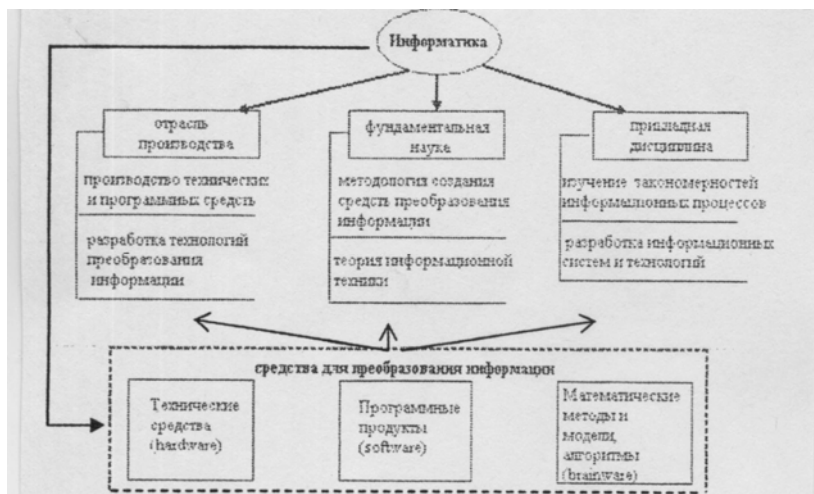


Рис 1.5. Структура информатики

Информатика как *фундаментальная наука* занимается разработкой методов, моделей и алгоритмов, а также связанных с ними математических теорий. Ее прерогативой является исследование процессов преобразования информации и на основе этих исследований разработка соответствующих теорий, моделей, методов и алгоритмов, которые затем применяются на практике.

Практическое использование результатов исследований информатики как фундаментальной науки воплощается в информатике – *отрасли производства*. Широко известны западные фирмы по производству программных продуктов, такие как Microsoft, Lotus, Borland, и технических средств – IBM, Apple, Intel, Hewlett Packard и другие. Помимо производства самих технических и программных средств разрабатываются также и технологии преобразования информации.

Как *прикладная дисциплина* информатика изучает закономерности протекания информационных процессов в конкретных

областях и методологию разработки конкретных информационных систем и технологий.

Информатика – наука об информационных элементах, информационных процессах и информационных системах

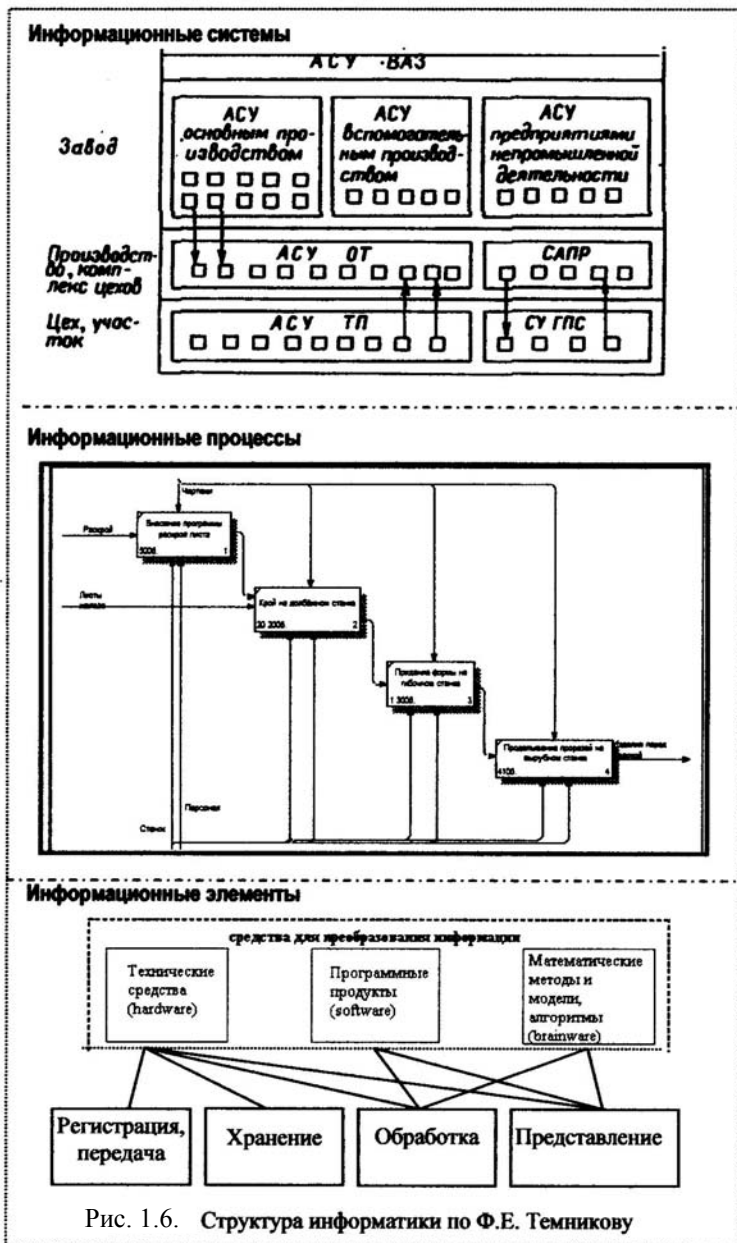
На основе первоначального определения информатики, предложенного в 1963 г. **Ф. Е. Темниковым**, в качестве способа систематизации разрозненно возникающих сведений об информации и информатике можно принять выделение 3-х основных частей: (рис. 1.6):

- 1) *информационные элементы*;
- 2) *информационные процессы*;
- 3) *информационные системы*.

С учетом осознания широкого спектра проблем, включаемых в настоящее время в направление, обобщаемое термином «информатика», ряд ученых ⁴⁸ считает, что именно определение **Ф. Е. Темникова** следовало бы возродить и принять в качестве обобщающего определения науки об информации как науки об *информационных элементах, информационных процессах и информационных системах*. Такое определение особенно важно для подготовки бакалавров и магистров по направлениям «Системный анализ и управление», которое ориентировано на необходимость анализа смысла, содержания информации, разработки и выбора соответствующих методов и средств исследования информационных процессов сбора, хранения, обработки и представления информации применительно к конкретным областям и ситуациям принятия решений.

Интерпретация этих частей применительно к настоящему периоду развития наук об информации приведена в табл. 1.4.

⁴⁸ Например, **Юсупов Р. М.** Информатика в системе научного знания XX и XXI веков / Р. М. Юсупов, Б. В. Соколов // Кибернетика и информатика: сб. науч. трудов: – СПб.: Изд-во Политехн. ун-та, 2006. – С. 6–21.



Информационные элементы	Информационные процессы	Информационные системы
<p><i>Информационный элемент</i> – предел членения системы с точки зрения аспекта её рассмотрения, цели исследования или создания, решаемой задачи.</p> <p><i>Информационный элемент</i> – это отражение существующих элементарных объектов, учитываемых при решении прикладных задач</p> <p><i>Информация и средства её обработки.</i></p> <p><i>Данные.</i></p> <p><i>Буквы, цифры, слова.</i></p> <p><i>Сведения о деталях, блоках и других компонентах изделий.</i></p> <p><i>Показатели, характеризующие состояние объекта, процесса.</i></p> <p><i>Информационный массив, база данных (БД), хранилище данных.</i></p> <p><i>Документальная информация, т.е. тексты.</i></p> <p><i>Нормы, статьи в законах и др. нормативных документах, т.е. фрагменты текста.</i></p> <p><i>Данные о производственных нормативах, состоянии среды, сведения о наличии и ценах товаров.</i></p> <p><i>Технологии, технические средства.</i></p> <p><i>Программные средства.</i></p>	<p><i>Информационные процессы</i> – это последовательности операций во времени при проектировании, производстве изделий, в экономике – бизнес-процессы, в управлении – организационно-технологические процедуры (ОТП)</p> <p><i>подготовки и реализации управленческих решений по организации производства, обслуживания, торговли</i></p> <p><i>Информационные процессы</i></p> <p><i>Процессы передачи</i></p> <p><i>Компьютерная технология: ввод, хранение, обработка, поиск, представление..</i></p> <p><i>Функционально-ориентированные и объектно-ориентированные CASE- и RAD- технологии.</i></p> <p><i>IDEF-технологии – компьютерная реализация методологии SADT (Structured Analysis and Design – структурный анализ и проектирование).</i></p> <p><i>Процессы возникновения и распространения научно-технической информации – информационные потоки.</i></p>	<p><i>Информационные системы</i> – информационные комплексы, включающие в себя в качестве основы технические объекты, обслуживаемые человеком (человеко-машинные системы), социальные и социально-экономические комплексы.</p> <p><i>Автоматизированные информационные системы (АИС) как первая очереди автоматизированных систем управления – АСУ.</i></p> <p><i>Автоматизированные системы научно-технической информации – АС НТИ</i></p> <p><i>Государственной системы научно-технической информации (ГСНТИ).</i></p> <p><i>ИС библиотечной сети РФ.</i></p> <p><i>Информационно-поисковые системы.</i></p> <p><i>Системы нормативно-методического обеспечения управления (СНМОУ) и их автоматизированный вариант – АСНМОУ.</i></p> <p><i>Справочно-правовые системы (СПС).</i></p> <p><i>ИС специального назначения (ИС государственной статистики; ИС природных ресурсов, явлений и процессов; экономические, финансовые внешне-экономической деятельности и др.).</i></p>

Такое определение особенно важно для подготовки бакалавров и магистров по направлениям «Системный анализ и управление», которое ориентировано на необходимость анализа смысла, содержания информации, разработки и выбора соответствующих методов и средств исследования информационных процессов сбора, хранения, обработки и представления информации применительно к конкретным областям и ситуациям принятия решений.

Виды и классификации информационных систем, определения информационных систем разного вида, принципы их разработки рассматриваются в главах 3–6.

Вопросы для самопроверки

1. Понятие об информации.
2. Основные особенности информации.
3. Материя и информация.
4. Неаддитивность, некоммутативность и неассоциативность информации.
5. Кумулятивность информации.
6. Условия неубывающей потенциальной эффективности информации
7. Зависимость фактической реализуемости и эффективности информации от степени ее использования.
8. Зависимость ценности информации от степени влияния на достижение целей получателя.
9. Специфические особенности научно-технической информации.
10. Меры информации.
11. Мера Р. Хартли.
12. Мера информации как снижения энтропии Л. Бриллюэна.
13. Теория насыщения тезауруса Ю. А. Шрейдера.
14. Энтропийная мера К. Шеннона.
15. Информационная мера А. А. Харкевича.
16. Прагматический, семантический и синтаксический аспекты информации.
17. Система мер оценки экономической информации Ю. И. Черняка.

18. Понятие об информации восприятия (знаки) и информации-значении, смысловом содержании для потребителя информации (означаемое знака) Е. К. Войшвилло.
19. Понятие о чувственной (информации восприятия), логической информации (содержания, потенциала) и смысле информации в теории А. А. Денисова.
20. Детерминированные и вероятностные меры А. А. Денисова.
21. Понятие об информации как ресурсе развития предприятия (организации).
22. Информационные ресурсы предприятия (организации).
23. Краткая история развития термина «информатика».
24. Понятие об информатике как науке об информационных элементах, информационных процессах и информационных системах.

Практикум:

Студент выполняет часть курсового проекта (курсовой работы): разрабатывает структуру информационного обеспечения конкретного предприятия или организации (по выбору студента).

Глава 2. ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ И МЕТОДЫ ИХ МОДЕЛИРОВАНИЯ

Проведенный в гл. 1 анализ понятия информации как отражения объективной реальности позволяет сделать вывод о том, что информационные процессы должны сопровождать все процессы, ориентированные на преобразования материальных и иных компонентов объективной реальности. Поэтому вопросы исследования и моделирования информационных процессов вынесены в самостоятельную главу.

2.1. Информационные процессы и их виды

Любые процессы – это последовательности операций во времени при проектировании, производстве изделий, обслуживании, управлении. Информационные процессы связаны с регистрацией, хранением, передачей, обработкой, представлением, поиском, использованием информации,

В технических дисциплинах (технической информатике) изучаются процессы передачи, хранения, преобразования информации с использованием технических средств, компьютерных технологий, В социально-экономической сфере – информационные процессы сопровождают все этапы: от возникновения бизнес-идеи, разработки вариантов ее реализации и выбора приемлемого, до реализации замысла, оценивания эффективности, анализа результатов реализации бизнес-идеи. При разработке любого проекта информационные процессы сопровождают все этапы процесса проектирования: от возникновения идеи, определения конфигурации проекта – до его реализации и анализа результатов использования.

При исследовании процессов в информационных системах вначале учитывали, что информация может переноситься в пространстве и во времени, для чего необходимы не только средства передачи, но средства хранения информации. Затем было

осознано, что для принятия решений в системах управления необходимо укрупнять, обобщать, преобразовывать первичную информацию с применением операций сортировки, упорядочения, математических вычислений, поиска.

Наиболее полное исследование функций информационной системы предложил в 1970 г. **Ф. Е. Темников**, изобразив их в форме окружностей и назвав «контурами» *связи, хранения, расчета, рассудка* и *политики*, которые должны формироваться в любой сложной системе (рис. 2.1):

C – связь (communication), регистрация, передача информации, перемещение ее в пространстве

G ; M – память (memory), хранение информации, перенос ее во времени t ;

K – расчет (от «калькулятор», «компьютер») обработка, получение новой информации, R – рассудок (reason), разум; P – «политика».

R – рассудок (reason), разум; P – «политика».

Этот набор функций – утверждал

Ф. Е. Темников – отличительная особенность любой сложной живой системы, необходимая и достаточная для ее реализации.

Этот набор функций – утверждал

Ф. Е. Темников – отличительная особенность любой сложной живой системы, необходимая и достаточная для ее реализации.

Этот набор функций – утверждал

Ф. Е. Темников – отличительная особенность любой сложной живой системы, необходимая и достаточная для ее реализации.

Этот набор функций – утверждал

Ф. Е. Темников – отличительная особенность любой сложной живой системы, необходимая и достаточная для ее реализации.

Этот набор функций – утверждал

Ф. Е. Темников – отличительная особенность любой сложной живой системы, необходимая и достаточная для ее реализации.

Этот набор функций – утверждал

Ф. Е. Темников – отличительная особенность любой сложной живой системы, необходимая и достаточная для ее реализации.

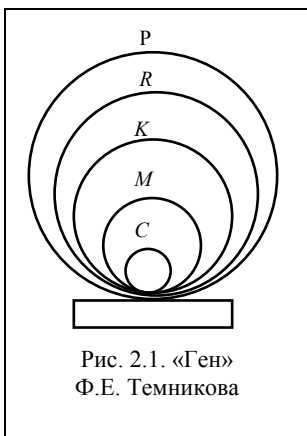


Рис. 2.1. «Ген»
Ф.Е. Темникова

Для создания простой *информационной* системы достаточно первых трех контуров, функции которых приведены на рис. 2.2.

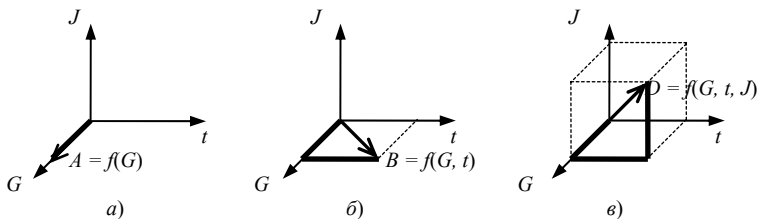


Рис. 2.2. Функции C , M , K

Для относительно простой технической системы, в которой происходит движение информации в какой-либо форме, достаточно только переноса информации в пространстве (рис. 2.2,а), т. е. функции C , после применения которой система попадает в точку $A = f(t)$ (например, системы передачи информации, телефонной связи и т. п.).

Более сложными являются технические системы с *памятью*, в которых, наряду с передачей информации предусматриваются блоки ее *задержки* во времени, хранения, т. е. выполняются функции C и M . В результате (рис. 2.2, б) система попадает в точку $B = f(G, t)$.

Для того, чтобы систему можно было назвать *информационной*, в ней должна осуществляться еще и обработка информации, т. е. выполняться три функции – C , M и K (рис. 2.2, в), после применения которых система попадает в точку $D = f(G, t, J)$.

Для реализации функции C и M разрабатывают соответствующие средства, для исследования процессов сбора и передачи C информации разрабатывают и применяют различные модели, для реализации функции K разрабатывают модели, алгоритмы и программы, ориентированные на решение конкретных задач.

В живых системах недостаточно способности к обработке информации типа расчетов, появляется еще и *рассудок*, *разум*. R . А в более сложно организованных социальных системах – еще и функция *политики* P , для реализации которой нужна соответствующая информация. Моделирование функций R и P является предметом исследования специальных методов и моделей, базирующихся на идеях искусственного интеллекта, извлечения знаний и др. научных направлений, ориентированных на формализацию исследования семантической информации.

Функции, приведенные на рис. 2.2. могут выполняться в различной последовательности:

Примеры приведены на рис. 2.3 и 2.4..

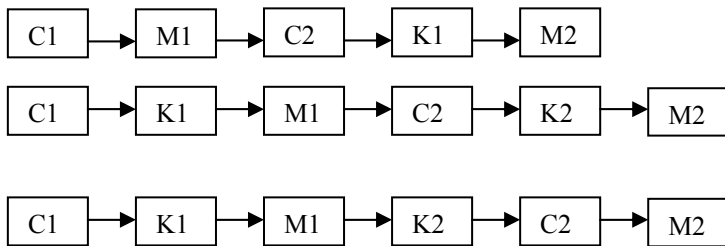


Рис. 2.3 Пример процессов

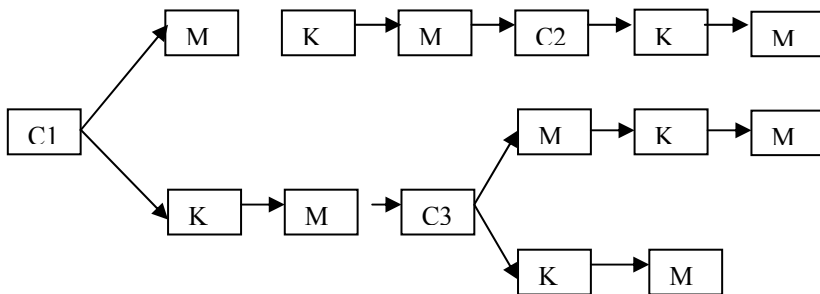


Рис. 2.4. Пример процессов

Последовательности функций определяются конкретными задачами и условиями, влияющими на их решение, а названия функций в структурных схемах, отображающих эти последовательности, формулируются на естественном языке с учетом конкретных условий моделирования информационных процессов. При определении этих последовательностей используется понятие «жизненный цикл», исследуемое в теории систем.

Для исследования рассмотренных функций разрабатывались модели, использующие различные методы, кратко рассмотренные в данной главе.

В начальный период развития теории информации разрабатывались и исследовались модели передачи данных. Для моделирования процессов передачи информации применялись методы алгебраического агрегирования (раздел 2.2) и теории выдвижения и проверки статистических гипотез *А. Вальда* (см. раздел 2.3).

Методы алгебраического агрегирования как один из методов, реализующих подход теории систем, основанный на анализе «пространства состояний» первоначально пытались применить и при разработке информационных систем. Эти методы позволяли формировать модели, применяющиеся при проектировании систем из укрупненных компонентов, что частично решало проблему «перечисления» системы (подробнее см. в гл. 5). Однако попытки реализации такого подхода на практике все же показали, что более предпочтительным подходом при проектировании АИС и АСУ является подход, основанный на формировании целей и функций системы управления, который в первых работах называли дезагрегированием (противопоставляя агрегированию), а в последующем системно-целевым (гл. 3), проблемно-ориентированным ⁴⁹.

Параллельно предпринимались попытки применения для исследования информационных процессов сетевых технологий. Первыми стали использоваться технология временных диаграмм, чаще называемых графиками Ганта, являющаяся одной из старейших технологий планирования и реализации работ во времени, которая была предложена в начале XX в. независимо друг от друга *К Адамецким* (K. Adamiecki) в Польше и *Г. Л. Гантом* (G.L. Gant) в США.

⁴⁹ *Никаноров С. П.*. Опыт прикладного применения системного анализа / С. П. Никаноров. – М.: Концепт, 2006. – 348 с.

В последующем ленточные диаграммы были разделены на дискретные операции, и возник ряд сетевых технологий, отличающихся различными формами представления графов (графы работ, графа событий, смешанные), и способами обработки графов, предложенные разными авторами. Наибольшее распространение в первые годы проектирования АСУ имела технология PERT (*Program Evaluation and Review Technique* – Методика оценки и контроля программ.)⁵⁰, теория *сетевого планирования и управления* (СПУ), а позднее и ряд методов, в том числе *статистического сетевого моделирования* с использованием вероятностных оценок графов. В настоящее время существуют модифицированная технология PERT-COST, технология стохастической сети GERT и ее модификация GERTS, технология CPM-COST и др.⁵¹

Когда в теории организационных структур был предложен подход, названный функционально-технологическим, его стали применять и для исследования информационных процессов при проектировании автоматизированных информационных систем предприятий (организаций). Один из способов реализации этого подхода, позволяющий формировать варианты прохождения информации при ее сборе и первичной обработке, был предложен и реализован в Проблемной лаборатории АСУ ВШ при кафедре Системотехники Московского энергетического института⁵².

⁵⁰ *Миллер Р. В.* ПЕРТ – система управления / Р.В. Миллер. Пер. с англ. Под ред. И. М. Сыроежина. – М.: Экономика, 1955. – 204 с.

⁵¹ *Троцкий М.* Управление проектами / М. Троцкий, Б. Гручо, К. Огонек. Перевод с польского. – М.: Финансы и статистика, 2006. – 302 с.

⁵² *Волкова В. Н.* Вопросы автоматизации проектирования АИС вуза / В. Н. Волкова // Опыт разработки и внедрения элементов автоматизированной информационной системы управления ВУЗом: сб. – М., 1971. – С. 39 – 42.

Этот подход основан на постепенной формализации модели информационных процессов⁵³. Принципы, положенные в его основу и пример реализации рассматривается в разделе 2.4.

Однако реализации функционально-технологического подхода была трудоемкой и сдерживалась ограниченными возможностями вычислительной техники того периода.

В 1990-е гг. для представления и исследования процессов при проектировании информационных систем и разработана методология SADT и семейство IDEF-технологий, что будет кратко рассмотрено в разделе 2.5.

В последнее время начинает осознаться, что исследование процессов возникновения и распространения информации является весьма значимым не только для научно-технической информации, но и для анализа мониторинговой информации, формирования ядра бизнеса. Этим исследованиям начинают уделять все больше внимания, в том числе в материалах, представленных в сети Интернет.

В теории научно-технической информации о таких процессах принято говорить как об *информационных потоках* и исследовать закономерности информетрии – законы *Дж. Ципфа, С. Брэдфорда, Б. Викери*, закономерность концентрации – рассеяния *В. И. Горьковой* [6], рассматриваемые в разделе 4.6.

⁵³ *Волкова В. Н.* Некоторые вопросы автоматизации проектирования АИС: автореф. дис. на соискание ученой степени канд. техн. наук по специальности 05.13.01 – Техническая кибернетика и теория информации / В. Н. Волкова. – М., 1973. – 24 с.; *Волкова В. Н.* К методике проектирования автоматизированных информационных систем / В. Н. Волкова // Автоматизация управления и вычислительная техника: кн. – Вып. 11. – М.: Машиностроение, 1975. – С. 189–300.

2.2. Модели информационных процессов, основанные на методе алгебраического агрегирования

При исследовании информационных процессов с информацией всегда связывают три понятия (их взаимосвязь показана на рис. 1.2): *источник информации* – элемент окружающего мира, сведения о котором являются объектом исследования; *потребитель информации* – элемент окружающего мира, который использует информацию; *сигнал* – материальный носитель, который фиксирует информацию для переноса ее от источника к потребителю и *код*, в форме которого информация передается.

При моделировании процессов передачи информации следует иметь в виду рассмотренные в гл. 1 свойства информации, и в частности, учитывать, что она переносится в пространстве с помощью носителей различной физической природы (звуковые, световые, радио волны, магнитные диски, ленты, бумажные носители т. п.) и передается в виде отдельных сообщений, которые могут тиражироваться без существенных энергетических затрат. При этом применяют *кодирующее* и *декодирующее* устройства. В процессе передачи информация может теряться и искажаться: Помехи, или, как их называют специалисты, шумы, искажают информацию. Существует наука, разрабатывающая способы защиты информации – *криптология*.

Каналы передачи сообщений характеризуются *пропускной способностью* и *помехозащищенностью*. Каналы делятся на *симплексные*, с передачей информации только в одну сторону (телевидение), и *дуплексные*, по которым возможно передавать информацию в оба направления (телефон, телеграф и др.). По каналу может одновременно передаваться несколько сообщений. Каждое из этих сообщений выделяется (отделяется от других) с помощью специальных фильтров. Например, возможна фильтрация по частоте передаваемых сообщений, как это делается в радиоканалах. Пропускная способность канала определяется макси-

мальным количеством символов, передаваемых ему в отсутствие помех. Эта характеристика зависит от физических свойств канала. Для повышения помехозащищенности канала используются специальные методы передачи сообщений, уменьшающие влияние шумов.

В начальный период развития теории информации и исследования информационных процессов достаточно широкое применение получили модели, теоретической основой которых являлся один из методов математики – *метод алгебраического агрегирования*, являющийся методом построения моделей, одинаковых по форме, но допускающих различную степень детализации.

Построение агрегированных моделей на множествах элементов реализует подход, основанный на анализе «пространства состояний», рассмотренный в предыдущей главе 3. Термин «агрегирование» противопоставлялся термину «деагрегирование» подобно терминам «композиция» – «декомпозиция».

Целью агрегирования может быть уменьшение числа компонентов, вводимых в математическую модель, уменьшение размерностей, углубление анализа некоторых свойств системы (например, устойчивости, помехозащищенности и т.п.) и характера взаимодействия системы с внешней средой (см., напр., в [10]).

Наряду с алгебраическим агрегированием применялись модели, основанные на вероятностном агрегировании.

Примером построения агрегированных моделей являются модели, основанные на принципе формально заданного «черного ящика», под которым понимают любое преобразующее устройство, если его внутреннее строение и функционирование не учитывается, а исследуются только выходные реакции на входные воздействия.

Например, в моделях массового обслуживания конкретные физические процессы обслуживания не исследуются, а учитываются только временные затраты, связанные с организацией обслуживания.

Важным при переходе от моделей элементов к агрегированным моделям является изменение параметров, определяющих состояние системы, и число учитываемых состояний. При анализе компонент моделируемого процесса можно учитывать макропараметры. Например, типа спектральных состояний сигнала, характеристик элементов принципиальных радиоэлектронных схем, значения внешних воздействия. По мере объединения элементов осуществлялся переход от макропараметров к обобщенным характеристикам. Вместо детального анализа процессов функционирования переходили к учету продолжительности выполнения этапов преобразования, к оценке вероятности сохранения параметров в установленных пределах и т.д.

При агрегативном моделировании начальный этап исследования базировался на формализованном задании элементов системы, алгоритмов их функционирования, параметров состояний, входных и выходных сигналов. В последующем для отображения учитываемых элементов, параметров, алгоритмов на начальном этапе моделирования использовались теоретико-множественные представления, а по мере углубления анализа формализованным отображениям постепенно придавался более формальный физический и математический смысл.

В последующем для моделирования информационных процессов стали применяться идеи и средства математической логики, математической лингвистики, разрабатываться языки моделирования.

2.3. Модели информационных процессов, базирующиеся на теории выдвигания и проверке статистических гипотез А. Вальда

Для исследования процессов передачи информации разрабатываются и применяются модели, базирующиеся на *теории выдвигания и проверки статистических гипотез А. Вальда*.

Идея метода заключается в следующем.

Имеется два векторных пространства – пространство априорной информации и пространство апостериорной информации. Первое из них иногда называют просто пространством информации S_I ; второе – пространством восприятия S_B или пространством решения.

В случае детерминированного (достоверного) восприятия точке $x \in S_I$ соответствует только одна точка $x' \in S_B$.

В случае статистического (вероятностного) восприятия каждой точке $x \in S_I$ в пространстве S_B соответствует распределение (x/x') , которое называется решающей функцией.

Что собой представляет решающая функция, можно пояснить на примере двувальтернативного решения с событиями x_1 и x_2 и их априорными вероятностями p_1 и p_2 ,

События x_1 и x_2 в пространстве S_I могут представлять собой два состояния точки x : x_1 – отсутствие точки \bigcirc x_2 – наличие точки \otimes .

В пространстве S_B событиям x_1 и x_2 соответствуют события x_1' и x_2' и условные вероятности $p(x_i'/x_i)$.

Способы восприятия информации приведены на рис. 2.5.

Для x_1



Для x_2



Рис. 2.5

Условные вероятности $p(x_1'/x_1)$ и $p(x_2'/x_2)$ характеризуют случаи правильного восприятия:

$p(x_1'/x_1)$ – отсутствие точки x в пространстве восприятия S_B в случае отсутствия ее в пространстве информации S_I ;

$p(x_2' / x_2)$ – обнаружение точки x в пространстве S_B в случае наличия ее в пространстве S_I .

Условная вероятность $p(x_2' / x_1)$ характеризует восприятие события x_2 при условии, что в пространстве S_I произошло событие x_1 , т. е. обнаружение точки x в пространстве S_B в случае ее отсутствия в пространстве S_I . Этот случай называют «ложная тревога».

Условная вероятность $p(x_1' / x_2)$ характеризует отсутствие точки x в пространстве S_B в случае ее наличия в пространстве S_I . Этот случай называют «пропуск сигнала» (например, самолет появляется, но его не обнаружили).

Для оценки рассмотренных случаев вводится понятие «функция потерь W ».

В случае правильного восприятия, т. е.

при $p(x_1' / x_1)$ и $p(x_2' / x_2)$ $W(x_1' / x_1) = W(x_2' / x_2) = 0$.

В случае ложной тревоги или пропуска сигнала

$W(x_2' / x_1) = 1$; $W(x_1' / x_2) = 1$.

Пользуясь этими оценками потерь, можно ввести понятие *условного риска*:

Для случая x_1 : $\bigcirc \longrightarrow \bigcirc$, α – принятое решение

$$r(x_1, \Delta) = W(x_1' / x_1) p(x_1' / x_1) + W(x_2' / x_1) p(x_2' / x_1) =$$

$$W(x_2' / x_1) p(x_2' / x_2) = p(x_2' / x_2) = \alpha,$$

$$\text{так как } W(x_1' / x_1) = 0, W(x_2' / x_1) = 1,$$

где Δ – условный знак, характеризующий возможность отклонения от правильного решения, т. е. условный риск равен вероятности ложной тревоги.

Для случая x_2 : $\bullet \longrightarrow \bigcirc$ β

$$r(x_2, \Delta) = W(x_2' / x_2) p(x_2' / x_2) + W(x_1' / x_2) p(x_1' / x_2) =$$

$$p(x_1' / x_2) = \beta.$$

Общий риск (имеют место решения α и β):

$$R = M(r) = p_1 r(x_1, \Delta) + p_2 r(x_2, \Delta) = p_1 \alpha + p_2 \beta.$$

При решении некоторых задач удобнее оперировать не непосредственно с функционалами, а с их отношениями. В частности, с отношением правдоподобия, представляющем собой отношение функционалов правдоподобия условных вероятностей при наличии и при отсутствии сигнала.

$$\text{Для рассматриваемого случая } L(x) = \frac{p(x_2' / x_1)}{p(x_1' / x_2)} = \frac{\alpha}{\beta}.$$

Решение о наличии или отсутствии сигнала принимается путем проверки, превышает ли отношение правдоподобия некоторый порог L_0 , т. е. если $L(x) > L_0$, то принимается решение о наличии сигнала. В противном случае принимается решение об отсутствии сигнала.

Рассмотренный случай – это критерий Зигерта–Котельникова, называемый часто критерием идеального наблюдения (минимизируется общий риск).

В теории статистических решений используются также следующие критерии:

а) критерий Байеса – критерий минимального риска; ищется минимальный риск из нескольких максимальных общих рисков;

б) критерий минимакса; априорные вероятности неизвестны и минимизируется значение максимально возможного риска, т. е. ищется $\min R \max$;

в) критерий Неймана–Пирсона; минимизируется β при $\alpha < \varepsilon$;

г) критерий А. Вальда – последовательный анализ: минимизируется число испытаний n , достаточное для принятия определенного решения. В этом случае производится непрерывный анализ отношения правдоподобия и

сравнения его с двумя порогами, нижним $L_1 = \frac{\beta}{1-\alpha}$ и верхним $L_2 = \frac{1-\beta}{\alpha}$.

Если $L(x) < L_1$, то принимается решение об отсутствии сигнала; а если $L(x) > L_2$, то принимается решение о наличии сигнала; если же $L_1 < L(x) < L_2$, то имеющихся данных недостаточно для принятия решения, и испытания продолжают.

В рассматриваемом примере пространство априорной информации содержит 2 события, и решающая функция включает 4 апостериорных условных вероятностей с соответствующими им функциями потерь. В случае большего объема событий рас-

пределение $\Delta(x'/x)$ усложняется. При выборе решающей функции следует руководствоваться возможностью минимизации среднего риска.

Основы общей теории статистических решающих функций разработаны **А. Вальдом**⁵⁴. Им же предложен общий подход к постановке задачи принятия статистических решений с интерпретацией ее как антагонистической игры. Проверка гипотез, теория статистических оценок, вопросы планирования эксперимента могут быть рассмотрены в рамках этой теории как частные случаи.

Независимо от теории решающих функций в теории связи развивалась теория потенциальной помехоустойчивости, начало которой положено работами **В. А. Котельникова**⁵⁵.

Слившись в работах **Д. Миддлтона**, эти два направления образовали теоретико-техническую дисциплину – *теорию статистических решений*, которая занимается разработкой методов раскрытия неопределенности пространства апостериорной информации.

2.4. Исследование информационных процессов на основе метода постепенной формализации модели

Принципиальной особенностью модели постепенной формализации задачи является то, что она ориентирована на развитие представлений исследователя об объекте или процессе принятия решений, на постепенное «выращивание» решения задачи. Поэтому предусматривается не одноразовый выбор методов моделирования, а смена методов по мере развития у ЛПП представлений

⁵⁴ **Вальд А.** Статистические решающие функции / А. Вальд // Теория игр. Позиционные игры: сб. / Под ред. Н. Н. Воробьева и И. Н. Врублевской. – М.: Наука, 1967.

⁵⁵ **Котельников В.А.** Теория потенциальной помехоустойчивости // В. А. Котельников. – М.: Госэнергоиздат, 1956.

об объекте и проблемной ситуации в направлении все большей формализации модели принятия решения.

Для формирования и анализа модели постепенной формализации была разработана методика системного анализа, сочетающая методы из групп МАИС и МФПС, которые помогают создать язык моделирования, процедуру оценки вариантов решения и автоматизировать процесс «выращивания» решения.

Возможный вариант смены методов по мере развития модели проиллюстрируем на упрощенном примере моделирования процессов прохождения информации при выборе структуры обеспечивающей части автоматизированной информационной системы.

Основная идея постепенной формализации иллюстрируется рис. 2.6. При этом на рис. 2.6 показаны последовательные переходы от методов работы с ЛПР (из группы МАИС) к методам формализованного представления и обратно.

В методике системного анализа для решения данной задачи удобно предусмотреть два основных этапа:

1. Формирование модели, отображающей возможные варианты прохождения информации в АИС (этот этап иллюстрируется рис. 2.6 и 2.7).

2. Оценка модели и выбор наилучшего варианта пути прохождения информации (рис. 2.8).

Охарактеризуем кратко эти этапы.

Этап 1. Формирование модели, отображающей возможные варианты прохождения информации в АИС.

- 1.1. Отграничение системы от среды («перечисление» элементов системы).*

Подэтап может выполняться с применением метода «мозговой атаки», а в реальных условиях – методов типа комиссий, семинаров и других форм коллективного обсуждения, в результате которого определяется некоторый перечень элементов будущей системы. В состав таких комиссий должны входить и разработчики, и будущие пользователи АИС.

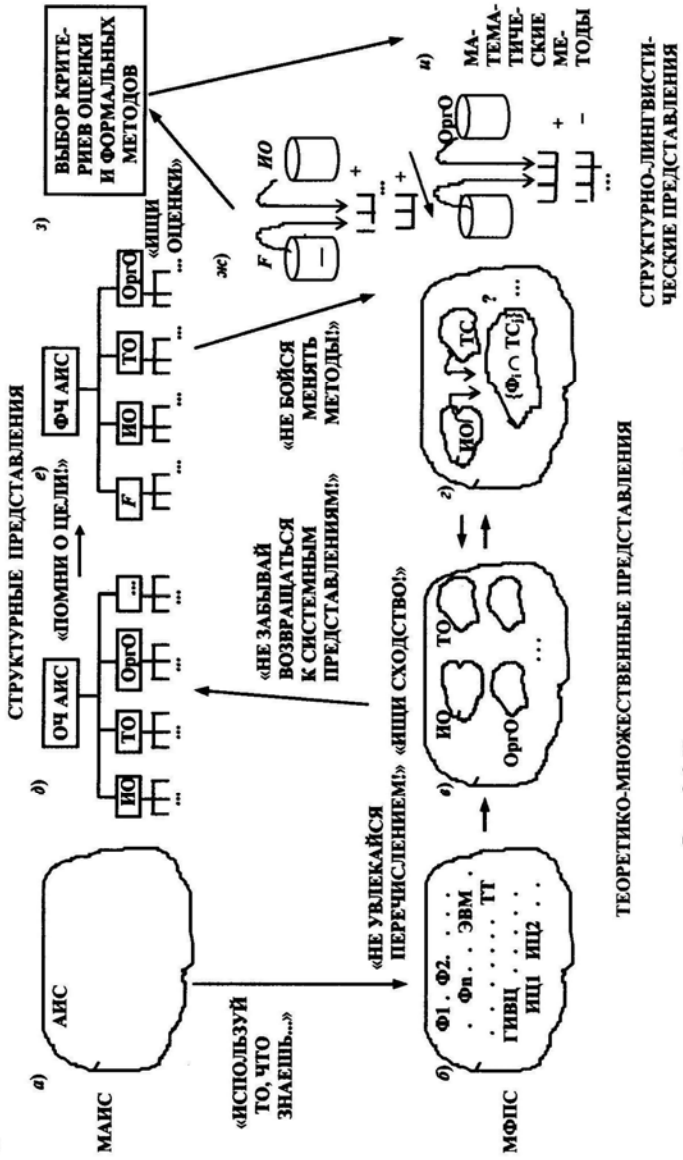


Рис. 2.6. Иллюстрация идеи постепенной формализации

Задачу «перечисления» можно представить на языке теоретико-множественных методов как переход от названия характеристического свойства, отраженного в названии формируемой системы и множества ее элементов, к перечислению элементов, которые отвечают этому свойству и могут быть включены в исходное множество.

На рис. 2.6, б и 2.7, а перечислено для примера небольшое число исходных элементов: *ГИВЦ, ИЦ1, ИЦ2, ... , А1, А2, ...* – пункты сбора и обработки информации; *Ф1, Ф2, ...* – формы сбора и представления информации (документы, массивы); *ЭВМ, ТТ* (телетайп), *Т* (телефон), ... и т. д. Понятно, что в реальных условиях конкретных видов подобных элементов существенно больше, и они будут названы более конкретно – не *ЭВМ*, а тип *ЭВМ*; аналогично – тип *ТТ*, регистраторов производства (*РП*), наименование или код документов и массивов и т. д.

1.2. Объединение элементов в группы.

Не следует слишком увлекаться подэтапом «перечисления» (хотя на практике это и имело место). Сложную реальную развивающуюся систему невозможно «перечислить» полностью. Следует, набрав некоторое множество элементов, попытаться объединить их в группы, найти меры сходства, «близости» и предложить способ их объединения.

Если в качестве метода формализованного отображения совокупности элементов выбраны теоретико-множественные представления, то этот подэтап можно трактовать как образование из элементов исходного множества некоторых подмножеств путем перехода от перечисления сходных по какому-то признаку элементов к названию характеристического свойства этого подмножества. В результате в приводимом примере могут быть образованы подмножества элементов по соответствующим видам обеспечения – *ИО, ТО, ОрГО* (рис. 2.6, в и 2.7, б).

Примечание. Можно было бы предусмотреть еще один подэтап разделения подмножеств на более мелкие.

1.3. Формирование из элементов подмножеств новых множеств, состоящих из «пар», «троек», «n-ок» элементов исходных подмножеств.

В рассматриваемом примере объединяя элементы подмножеств *ИО*, *ТО*, *ОргО* в «пары» и «тройки» можно получить, например: *Ф1_ЭВМ*, *Ф1_ТТ*, *Ф2_ЭВМ* и т. п.; *ЭВМ_ГИВЦ*, *ЭВМ_ИЦ1*, *ЭВМ_А1*, *ТТ_ГИВЦ*, *ТТ_ИВЦ1*, *ТТ_А1* и т. п.; *Ф1_ЭВМ_ГИВЦ*, *Ф1_ТТ_А1* и т. д.

Иногда в задачах моделирования на этом этапе можно получить новый результат, который подсказывает путь дальнейшего анализа.

Но в данном примере, как правило, интерпретация получаемых компонентов затруднена и ввести какое-либо формальное правило сравнения элементов новых множеств для принятия решения о выборе наилучших не удастся. В таких случаях, согласно рассматриваемому подходу, нужно обратиться к системно-структурным представлениям и попытаться поискать дальнейший путь развития модели.

Примечание. Отметим, что отражением данного подэтапа в истории развития проектных работ по созданию АСУ был период, когда формировались матрицы «документы – технические средства», «информационные службы – технические средства» и на основе экспертной оценки элементов этих матриц пытались принимать решения, связанные с формированием структуры ОЧ АИС.

1.4. Содержательный анализ полученных результатов и поиск новых путей развития модели.

Для проведения содержательного анализа следует вернуться к системным представлениям и использовать один из методов группы МАИС – структуризацию (в данном случае в форме иерархической структуры – рис. 2.6, д и 2.7, в).

Такое представление более удобно для руководителей работ по созданию АСУ, чем теоретико-множественные представления,

и помогает им вначале распределить работу между соответствующими специалистами по *ИО*, *ТО* и т. д., а затем найти дальнейший путь развития модели на основе содержательного анализа сути полученных «пар» и «троек» с точки зрения формулировки решаемой задачи.

В формулировке задачи, помимо упоминания видов обеспечения, говорится, что все эти компоненты нужно разместить так, чтобы обеспечить наиболее эффективную реализацию подсистем и задач, входящих в ФЧ АИС. А поскольку любая задача представляет собой последовательность действий (*функций*) по сбору, хранению и первичной обработке информации, то становится очевидной необходимость внесения в модель нового подмножества «Функции-операции (*ФО*)», добавление элементов которого к прежним «парам» и «тройкам» позволяет получить новое их осмысление (рис. 2.6, *е*).

Для того, чтобы глубже понять развитие модели постепенной формализации, проиллюстрируем изложенную идею на конкретном примере.

Предположим, что нужно принять решение о структуре ОЧ АИС отрасли (первой очереди ОАСУ), предприятия которой расположены в разных городах.

Предварительно рассматривались 2 основных варианта:

1) создание единого Главного информационно-вычислительного центра (*ГИВЦ*) отрасли и организации централизованного сбора от всех предприятий посредством установленных на них периферийных средств сбора информации (*A1, A2, ... , Ak*);

2) наряду с *ГИВЦ* и периферийными средствами сбора на предприятиях, создать региональные *ИВЦ* (обозначенные на рис. 2.7 *ИЦ1, ИЦ2, ... , ИЦn*), которые будут расположены в городах.

Необходимо выбрать вариант и определить вычислительную мощность *ГИВЦ* и региональных *ИВЦ* (в случае выбора второго варианта), типы *ЭВМ* для *ГИВЦ* и *ИВЦ*, типы периферийных средств регистрации информации, объемы информационных массивов в *ГИВЦ* и *ИВЦ*, формы документов *Ф1, Ф2, ... , Фm* сбора и передачи информации между пунктами, принятыми в соответствующем варианте. При этом следует иметь в виду, что в случае выбора первого варианта возникают проблемы диспет-

через прием-передачи информации от достаточно многочисленных пунктов первичного сбора информации на предприятиях.

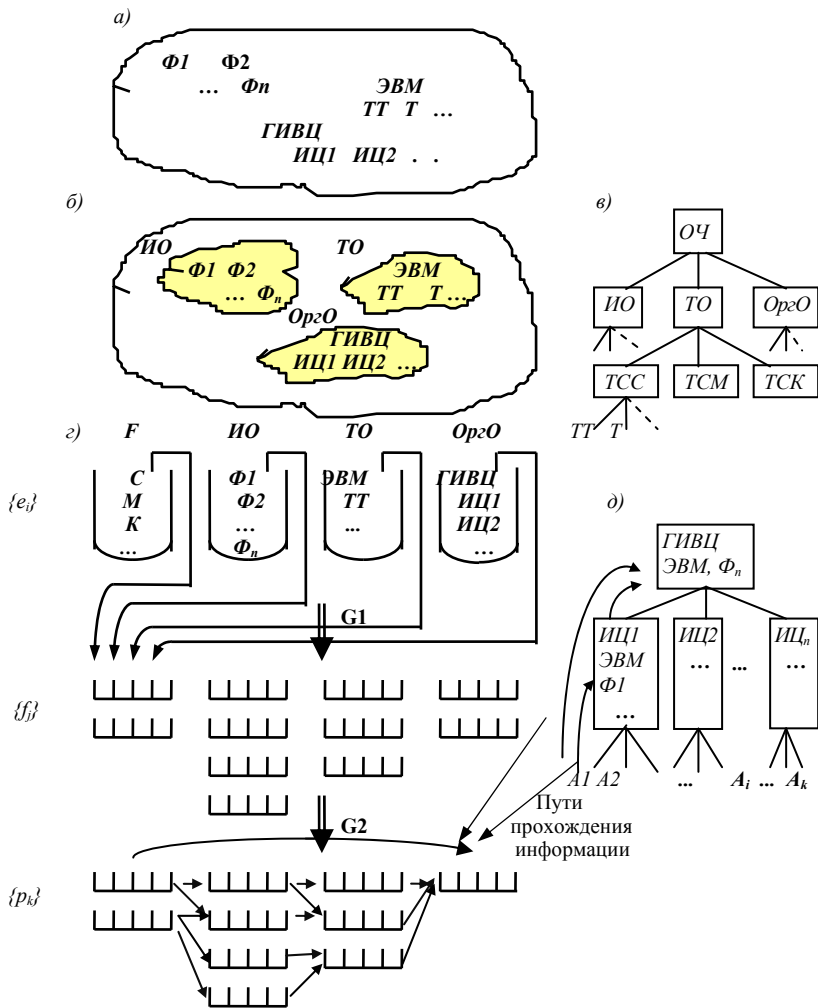


Рис. 2.7. Иллюстрация формирования графо-семиотической модели

Аналогично может быть поставлена задача для корпорации, фирмы которых расположены в разных городах (как, например, в Объединении

АвтоВАЗ), или для предприятия, крупные производства которого расположены в разных корпусах.

Для ответа на требуемые вопросы и выбора структуры ОЧ АИС необходимо исследовать информационные потоки.

Для простоты на рис. 2.6 и рис. 2.7 показаны только принципиально отличающиеся друг от друга функции – связи C , хранения M (от «*memory*» – «память») и обработки K (от «компьютер»). После их добавления получают комбинации, которые ЛПР могут не только сравнивать, но и оценивать.

Например, комбинации типа $C_ФI_TT$ и $C_ФI_T$ отличаются друг от друга скоростью передачи информации, которые в конкретных условиях можно измерить или вычислить.

Если бы задача не была сформулирована, то при содержательном осмыслении результатов предшествующих этапов нужно было бы уточнить формулировку задачи, что и помогло бы сделать дальнейший шаг в развитии модели.

Примечание. Важно отметить, что применение на практике матричных представлений, упомянутых в предыдущем пункте, обратило внимание разработчиков на тот факт, что анализировать названные матрицы практически невозможно, если не задумываться над тем, для решения каких задач предназначены технические средства и документы. Тогда стали формировать матрицы «задачи – методы», «задачи – средства», т. е. вводить фактически новое множество – «Задачи» или «Функции (F)».

1.5. Разработка языка моделирования.

После того, как найдено недостающее подмножество, в принципе можно было бы продолжить дальнейшее формирование модели, пользуясь теоретико-множественными представлениями. Однако, когда осознана необходимость формирования последовательностей функций-операций, конкретизированных путем дополнения их видами обеспечения – конкретизированных функций ($KФ$), то целесообразнее выбрать лингвистические представления, которые удобнее для разработки языка моделирования последовательностей $KФ$.

Тогда в терминах лингвистических представлений данный этап можно представить следующим образом:

- разработка тезауруса языка моделирования;
- разработка грамматики (или нескольких грамматик, что зависит от числа уровней модели и различии правил).

Структура тезауруса языка моделирования, приведенная на рис. 2.6, *ж* и 2.7, *з*, включает три уровня:

- уровень первичных терминов (или слов), которые представлены в виде списков, состоящих из элементов $\{e_i\}$ подмножеств ΦO , $И O$, $Т O$, $О р г O$;
- уровень фраз $\{f_j\}$, который в этом конкретном языке можно назвать уровнем $K\Phi$, так как абстрактные функции C , M , K , объединяясь с элементами подмножеств $И O$, $Т O$, $О р г O$, конкретизируются применительно к моделируемому процессу;
- уровень предложений $\{p_k\}$, отображающий варианты прохождения информации при решении той или иной задачи.

Грамматика языка включает правила двух видов:

- G1 – преобразования элементов $\{e_i\}$ первого уровня тезауруса в компоненты $\{f_j\}$ второго уровня, которые имеют характер правил типа «помещения рядом» (конкатенации, сцепления, R_I на рис. 2.7);

• G2 – преобразования компонентов $\{f_j\}$ в предложения $\{p_k\}$ (варианты прохождения информации) – правила типа «условного следования за» (R_{II} на рис. 2.85); правила этого вида исключают из рассмотрения недопустимые варианты следования информации: например, после функции $C1_Ф2_A1-ИЦ1_ТТ$ (передача документа $Ф2$ из $Ф1$ в $ИЦ1$ с помощью $ТТ$) не может следовать функция $M1_Ф2_ГИВЦ_МН$, так как в результате выполнения предшествующей функции документ $Ф2$ в $ГИВЦ$ не поступил (здесь $МН$ – машинный носитель).

Используя разработанный язык, процедуру формирования модели можно автоматизировать. При этом правила типа G1 и G2 относительно несложно реализуются с помощью языков логического программирования, и в частности языка Турбо-Пролог.

Э т а н 2. Оценка модели и выбор наилучшего варианта пути прохождения информации.

После формирования вариантов следования информации необходимо их оценить. Для этого могут быть приняты также разные варианты – от содержательной оценки путей сбора и первичной обработки информации (нижний уровень рис. 2.6, ж) до поиска алгоритмов последовательного преобразования оценок компонентов предшествующих уровней модели в оценки компонентов последующих уровней, что осуществляется путем анализа сформированной графо-семиотической модели. Варианты оценки модели иллюстрируются рис. 2.8.

2.1. Выбор способа оценки графо-семиотической модели.

В рассматриваемом примере можно проводить оценку тремя способами (рис. 2.8):

а) на уровне вариантов прохождения информации $\{p_k\}$, что иногда могут сделать компетентные специалисты путем коллективного обсуждения предложенных им вариантов (если число этих вариантов не очень велико – не более 7 ± 2);

б) на уровне $K\Phi \{f_{ij}\}$ с последующим преобразованием этих оценок $W \{f_{ij}\}$ в оценки вариантов $W' \{p_k\}$;

в) на уровне элементов $\{e_{ij}\}$ с последующим преобразованием оценок $W \{e_{ij}\}$ в оценки $W \{f_{ij}\}$, а их – в оценки $W' \{p_k\}$.

При втором способе можно выделить на модели «сферы компетентности» и поручить оценку $K\Phi$ по сферам соответствующим специалистам; оценки $K\Phi$ в большинстве случаев также получают экспертно, однако в некоторых случаях они могут быть измерены (например, время передачи информации с помощью TT и T в приведенных выше функциях $C_ФI_TT$ и $C_ФI_T$); этот способ подобен оценке сетевой модели, и при определении алгоритма преобразования оценок Φ_l можно пользоваться опытом сетевого моделирования (для большинства критериев оценки алгоритм преобразования – суммирование, а для критерия надежности передачи или хранения информации, оцениваемых с помощью вероятностей, алгоритм более сложный).

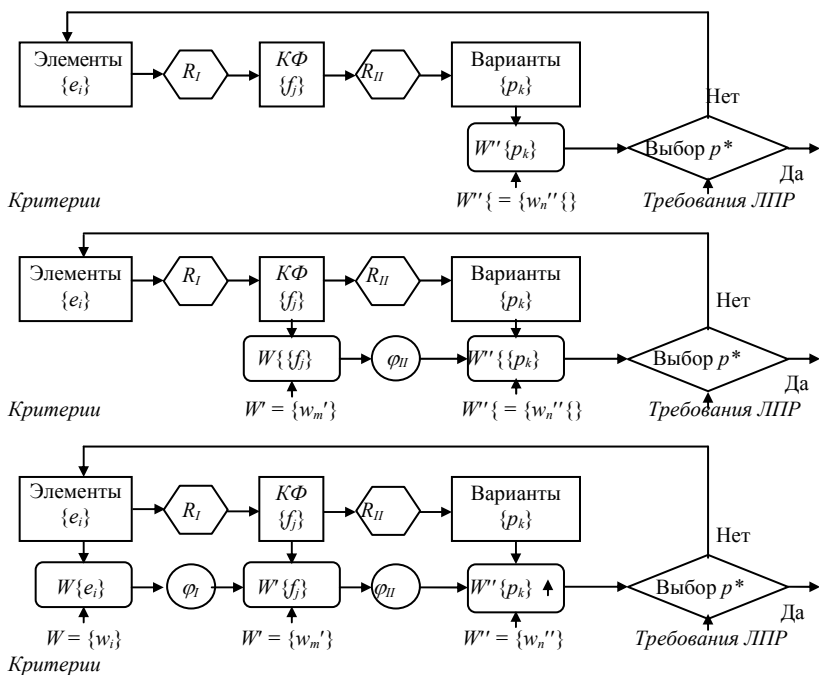


Рис. 2.8. Способы оценки графо-семиотической модели

При третьем способе алгоритмы преобразования φ могут быть найдены путем анализа различных $K\Phi$ с точки зрения влияния на их оценку по тому или иному критерию элементов соответствующего вида. Например, оценка $K\Phi$ передачи информации «С...» по критерию времени t может быть получена на основе выяснения, что в структуре $K\Phi$ «С...» влияет на оценку по t . Если используются технические средства связи, то, зная принципы передачи информации с их помощью, можно определить v_{TC} и зависимости $t = r_{\varphi}/v_{TC}$, где r_{φ} – объем передаваемой информации (например, измеряемых в числе знаков), т. е. оценка элементов, принадлежащих подмножеству $ИО$; v_{TC} – скорость передачи информации с помощью соответствующего технического средства, т. е. оценка элемента, принадлежащего подмножеству $ТО$. Таким образом, в данном примере на оценки $K\Phi$ «С...» влияют элементы подмножеств $ИО$ и $ТО$, и следует предусмотреть оценку этих элементов в исходных списках элементов. Аналогично можно определить, какие из эле-

ментов влияют на оценки КФ по стоимости, надежности, срокам внедрения и другим учитываемым критериям оценки.

С учетом сказанного следует сделать вывод о том, что выбор способа оценки модели зависит от результатов этапа 1, т. е. от вида графосемиотической модели, а алгоритмы преобразования оценок φ и φ_l определяются на основе анализа этой модели.

2.2. Выбор критериев оценки модели.

Выбор критериев оценки зависит от выбранного способа оценки модели.

Например, при первых двух способах оценки (на уровне $\{p_i\}$ и на уровне $\{f_{ij}\}$) могут быть приняты такие оценки, как оперативность (время), достоверность (вероятность сбоя при передаче информации, ошибок при ее обработке и т. п.), трудоемкость, затраты на внедрение, эксплуатационные расходы, сроки внедрения и т. д., а при оценке модели на уровне элементов $\{e_{ij}\}$ – оценки типа r_{φ} , v_{TC} и т. п., на основе которых могут быть вычислены оценки КФ, или оценки трудоемкости, скорости заполнения форм или ввода информации и т. п.

2.3. Оценки модели.

Способ оценки модели на уровне вариантов $\{p_k\}$ – экспертный; на уровне $\{f_{ij}\}$ для экспертного оценивания могут быть выделены сферы компетентности и привлечены соответствующие специалисты, знающие особенности конкретных технических средств и т. п.; и кроме того, наряду с экспертным оцениванием могут быть проведены эксперименты по той или иной КФ.

Оценки элементов $\{e_{ij}\}$, необходимые для вычисления оценок соответствующих КФ, могут быть в большинстве случаев получены из справочной литературы или измерены.

2.4. Выбор наилучшего варианта структуры ОЧ.

Основой структуры ОЧ являются выбранные варианты следования информации по задачам и подсистемам ФЧ.

Задачу определения наилучшего варианта можно поставить как многокритериальную. Но можно организовать процесс выбора вариантов путем постепенного ограничения области допустимых решений, вначале исключить все p_k , которые не удовлетворяют граничным значениям учитываемых критериев, затем

предложить рассмотреть оставшиеся варианты ЛПР, которые могут либо сразу выбрать из них наиболее предпочтительный, либо ввести весовые коэффициенты критериев, либо исследовать область допустимых решений по Парето.

Можно также добавить новые критерии качественного характера, не включенные в первоначально выбранный перечень критериев из-за невозможности их количественной оценки.

В заключение отметим, что после того, как для какого-то класса задач пройдены все этапы постепенной формализации задачи и найдены основы языка моделирования, можно применять не всю методику, а сразу начинать с подэтапа 1.5.

В то же время в случае, когда нужно поставить задачу для принципиально нового объекта или процесса, полезно при обосновании модели выполнять все подэтапы методики. Это может обосновать адекватность модели и принципы разработки языка автоматизации моделирования.

В рассматриваемом примере при разработке структуры ОЧ учитываются только задачи сбора, предварительной обработки информации и формирования первичных информационных массивов (что и включалось в основные функции АИС как первой очереди АСУ).

Если предполагается, что об АИС первоначально ничего неизвестно, кроме ее назначения, то модель формируется, как показано выше.

В качестве первого шага системного анализа предлагается принять «отграничение» системы от среды путем «перечисления» ее возможных элементов (рис. 2.7, б).

Затем (рис. 2.7, в и д) для анализа некоторого полученного множества выбраны теоретико-множественные представления, помогающие найти на сформированном пространстве состояний «меры близости» для объединения элементов в группы.

При этом вначале может быть использован эффект получения нового смысла у элементов, сформированных из «пар», «троек», «*n*-ок» элементов исходных подмножеств, на которые предварительно разделено общее множество элементов АИС).

Далее, когда возможности теоретико-множественных представлений в познании взаимодействия элементов в системе исчерпываются, следует вернуться к системно-структурным представлениям, с помощью которых активизируется использование интуиции и опыта ЛПР, перечень множеств анализируется и при необходимости дополняется (рис. 2.7, е) принципиально важными подмножествами для дальнейшего моделирования (в частности, в рассматриваемой задаче на этом этапе перечень исходных подмножеств *ИО*, *ТО*, *ОргО* дополнен подмножеством функций *F*).

Для дальнейшей реализации идеи комбинирования элементов в поисках вариантов решения задачи (в рассматриваемом примере – путей прохождения информации при ее сборе и первичной обработке) могут быть выбраны более удобные и подсказывающие правила формирования вариантов лингвистические представления, являющиеся основой разработки языка моделирования путей прохождения информации.

В рассматриваемом примере использовано сочетание *лингвистических*, *семиотических* и *графических* представлений и разработан язык *графо-семиотического моделирования* (который в первоначальных вариантах использования рассматриваемого подхода иногда носил и другие названия – *структурно-лингвистического*, *сигнаурного* (знакового) моделирования).

Таким образом, с помощью языка моделирования разрабатывается многоуровневая модель. В нашем примере – двухуровневая, если считать уровень исходных множеств нулевым (рис. 2.7, ж).

Осмысление этой модели (на уровне МАИС) приводит к преобразованию структуры: первоначально структура ОЧ формировалась как структура-состав, в которой были представлены виды обеспечения и их детализация (рис. 2.7, г и е), и на той основе осознана необходимость анализа структуры-функционирования, т. е. вариантов структуры информационных потоков (рис. 2.7, з).

Многоуровневую модель, отображающую взаимосвязи между компонентами и целями системы, в обобщенном виде можно представить в виде аналитических зависимостей (2.1, а) и (2.1, б).

Для варианта оценок, приведенного на рис. 2.8, а:

$$\begin{aligned}
 W^{n*}(p_{jn}) &= \text{opt } W^n(p_{jn}); \\
 & \quad p_{jn} \in P_n, \quad P_n \subset S_I \\
 W^n(p_{jn}) &= \varphi^n \{W^{n-1}(p_{j\ n-1})\}; \\
 & \quad p_{j\ n-1} \in P_{n-1}, \quad P_{n-1} \subset S_I \\
 & \dots\dots\dots \\
 W^k(p_{jk}) &= \varphi^k \{W^{k-1}(p_{j\ k-1})\}; \\
 & \quad p_{jk} \in P_k, \quad P_k \subset S_I, \quad p_{j\ k-1} \in P_{k-1}, \quad P_{k-1} \subset S_I \\
 & \dots\dots\dots \\
 W^1(p_{j1}) &= \varphi^1 \{W(e_i)\}; \\
 & \quad p_{j1} \in P_1, \quad P_1 \subset S_I, \quad e_i \in E, \quad E \subset S_I
 \end{aligned}
 \tag{2.1, а}$$

Для варианта, приведенного на рис. 2.8, б:

$$\begin{aligned}
 W^{n*}(p_{jn}) &= \text{opt } W^n(p_{jn}); \\
 & \quad p_{jn} \in P_n, \quad P_n \subset S_I \\
 P_{jn} &= \bigcup_{j\ n-1=1}^m p_{j\ n-1}; \\
 & \quad p_{j\ n-1} \in P_{n-1}, \quad P_{n-1} \subset S_I \\
 & \dots\dots\dots \\
 P_{jk} &= \bigcup_{j\ k-1=1}^I p_{j\ k-1}; \\
 & \quad p_{jk} \in P_k, \quad P_k \subset S_I, \quad p_{j\ k-1} \in P_{k-1}, \quad P_{k-1} \subset S_I \\
 & \dots\dots\dots \\
 P_{j1} &= \bigcup_{j=1}^r e_i; \\
 & \quad p_{j1} \in P_1, \quad P_1 \subset S_I, \quad e_i \in E, \quad E \subset S_I.
 \end{aligned}
 \tag{2.1, б}$$

Знаком \bigcup обозначено любое взаимодействие компонент «условное следование за», сложное взаимодействие или просто «помещение рядом»; $W^n(p_{jn})$ – функционал, связывающий критерии оценки выбираемого решения с компонентами p_{jn} , которые зависят от компонентов предыдущего уровня $p_{j\ n-1}$; в общем случае p_{jk} зависят от компонентов $p_{j\ k-1}$; $E, P_1, \dots, P_k, \dots, P_{n-1}, \dots, P_n$ – множества смысловыражающих элементов (тезаурус) задачи;

$W(e_i), W^1(p_{il}), W^k(p_{jk}), W^n(p_{jn})$ – критериальные отображения элементов (компонентов) структурных уровней тезауруса языка моделирования; $\varphi^1, \varphi^k, \varphi^n$ – алгоритмы преобразования критериальных отображений одного структурного уровня в другой.

В результате получается система алгоритмов, обеспечивающая возможность автоматизации, и благодаря этому – повторяемость процесса формирования и анализа модели при изменении наборов первичных элементов и их оценок. Эта система алгоритмов как бы обеспечивает взаимосвязь между компонентами и целями системы (или при моделировании потоков информации по отдельным задачам АИС – между компонентами и этой задачей), т. е. получается в результате формальная, аналитическая модель, только представленная не в виде привычных для такого рода моделей формул или уравнений, а в виде алгоритмов и их программной реализации в памяти ЭВМ.

Однако получить такой сложный алгоритм взаимодействия элементов в системе, позволяющий отобразить конкретную ситуацию и выбрать лучшее решение, практически невозможно без организации направленной постепенной формализации задачи.

2.5. Информационные процессо-ориентированные технологии

Процессный подход (который можно считать развитием *функционально-технологического подхода*) основан на структуризации во времени, на представлении процессов в форме графов.

При несомненной привлекательности применение функционально-технологического подхода долгое время было практически нереализуемым из-за большой трудоемкости, отсутствия правил и средств автоматизации формирования графов, отображающих процессы в системах.

В 1990-е годы была разработана методология SADT⁵⁶ (предложена *Дугласом Россом*⁵⁷), представляющая собой совокупность методов, правил и процедур, предназначенных для построения функциональной модели объекта какой-либо предметной области. На ее основе разработаны и стали широко применяться функционально-ориентированные и объектно-ориентированные CASE-⁵⁸ и RAD-⁵⁹ технологии. Компьютерная реализация методологии SADT получила название IDEF (Icam Definition). Основными структурными моделями являются модели процессов IDEF0 и IDEF3, модель данных IDEF1X. Созданы стандарты IDEF и DFD,⁶⁰ ориентированные на анализ процессов (в том числе бизнес-процессов). Для реализации моделей применяются автоматизированные средства – BPWin, ARIS, язык UML (Unified Modeling Language – Унифицированный язык моделирования)⁶¹.

Популярность CASE-методологии и RAD-технологий базируется на разработке принципов и автоматизации формирования процессов, на развитии методов их формирования на основе анализа «жизненного цикла» производства, обслуживания или других процессов, причинно-следственных связей и т.п., что и обеспечило развитие *процессного подхода*, преимущества которого заключаются в возможности учитывать особенности конкретного объекта и условий его деятельности.

⁵⁶ SADT – Structured Analysis and Design – структурный анализ и проектирование

⁵⁷ *Ross D.* Applications and extension of SADT // IEEE. Computer. – April, 1995.

⁵⁸ CASE – computer aided software engineering – компьютерное проектирование программных систем/

⁵⁹ RAD – rapid application development – быстрая разработка приложений/

⁶⁰ DFD – Data Flow Diagram – диаграммы потоков данных.

⁶¹ См., напр., *Фаулер М.* UML в кратком изложении: Применение стандартного языка объектного моделирования / Пер. с англ./ М. Фаулер, К. Скотт. – М.: Мир, 1996. – 112 с.

Впервые термин «реинжиниринг бизнес-процессов» (от англ. *business process reengineering*, BPR) был введен **М.Хаммером**⁶², который определяет этот вид деятельности как «*фундаментальное перепроектирование бизнес-процессов компаний для достижения коренных улучшений в основных актуальных показателях их деятельности: стоимость, качество, услуги и темпы*».

М.Хаммер рассматривает BPR как революцию в бизнесе, отход от базовых принципов построения предприятий, который превращает конструирование бизнеса в инженерную деятельность. BPR возник на стыке двух различных сфер деятельности управления (менеджмента) и информатизации. Именно поэтому реинжиниринг потребовал новых специфических средств представления и обработки проблемной информации, понятных как менеджерам, так и разработчикам информационных систем. Одной из основных особенностей BPR является ориентация реинжиниринга *не на функции, а на процессы*. Видимо поэтому возникло название «*процессный подход*».

Для отображения документооборота применяют различные средства, одним из которых являются диаграммы потоков данных (DFD), входящие в семейство средств компьютерного проектирования программных систем (CASE), предназначенные для профессионалов информационных систем.

Для реализации процесса моделирования БП применяют семейство методологий IDEF⁶³, являющееся государственным стандартом в США. С помощью методологии семейства IDEF можно эффективно отображать и анализировать модели деятельности широкого спектра сложных систем в различных разрезах.

⁶² **Хаммер М.** Реинжиниринг корпорации: Манифест революции в бизнесе. Пер. с англ. / М. Хаммер Дж. Чампи. – СПб.:Изд-во СПбГУ, 1999. – 287 с.

⁶³ IDEF – Integrated DEFinition.

В настоящий момент в семейство IDEF включают следующие стандарты:

IDEF0 – методология функционального моделирования. С помощью наглядного графического языка IDEF0 представляет изучаемую систему в виде набора взаимосвязанных функций («функциональных блоков»). Как правило, моделирование средствами IDEF0 является первым этапом изучения любой системы;

IDEF1 – методология моделирования информационных потоков внутри системы. Позволяет отображать и анализировать их структуру и взаимосвязь;

IDEF1X (IDEF1 Extended) – методология построения реляционных структур. IDEF1X относится к типу методологий «Сущность-взаимосвязь» (ER – Entity-Relationship) и, как правило, используется для моделирования реляционных баз данных, имеющих отношение к рассматриваемой системе;

IDEF2 – методология динамического моделирования развития систем. Из-за серьезных сложностей, связанных с анализом динамических систем, от этого стандарта сейчас практически отказались, и его развитие приостановилось на самом начальном этапе. Существующие алгоритмы и их компьютерные реализации позволяют превращать набор статических диаграмм IDEF0 в динамические модели, построенные на базе «раскрашенных сетей Петри» (CPN – Color Petri Nets);

IDEF3 – методология документирования процессов, происходящих в системе. С помощью IDEF3 описываются сценарий и последовательность операций для каждого процесса. IDEF3 напрямую связана с методологией IDEF0: каждая функция (функциональный блок) может быть представлена средствами IDEF3 в виде отдельного процесса;

IDEF4 – методология построения объектно-ориентированных систем. Средства IDEF4 позволяют наглядно отображать структуру объектов и принципы их взаимодействия,

позволяя анализировать и оптимизировать сложные объектно-ориентированные системы;

IDEF5 – методология онтологического исследования сложных систем. С помощью словаря терминов и правил позволяет описать онтологию системы. В итоге могут быть сформированы достоверные утверждения о состоянии системы в некоторый момент времени, на основе которых делаются выводы о дальнейшем развитии системы и производится её оптимизация.

Методологию IDEF0 можно считать конечным этапом развития известного графического языка описания функциональных систем SADT..

В нашей стране эту методологию развивают *Г. Н. Калянов*⁶⁴ и *С. В. Черемных*⁶⁵.

Особенностью рассматриваемого семейства методологий является, во-первых, возможность «задавать вопросы» в процессе моделирования, а, во-вторых, неразрывная связь графических средств (нотации), методологии и технологии. С этой точки зрения семейство IDEF является единственным средством, которое предоставляет не только возможности отображения бизнес-процессов, но и методологию взаимодействия «аналитик-специалист», и, кроме того, дает технологию создания проектов, охватывающую все стадии «жизненного цикла» – от первичного анализа до формы представления окончательного проекта.

Для различных классов систем используются свои методы разработки. Они определяются типом создаваемой системы, средствами реализации. Вероятно, самыми распространенными

⁶⁴ *Калянов Г. Н.* Моделирование, анализ, реорганизация и автоматизация бизнес-процессов / Г. Н. Калянов. – М.: Финансы и статистика, 2007. – 240 с.

⁶⁵ *Черемных С. В.* Структурный анализ систем: IDEF-технологии / С. В. Черемных, И. О. Семенов, В. С. Ручкин. – М.: Финансы и статистика, 2003. – 208 с. Черемных С.В. Моделирование и анализ систем IDEF-технологии: практикум. / С. В. Черемных, И. О. Семенов, В. С. Ручкин. – М.: Финансы и статистика, 2006. – 192 с.

по объемам разработок являются информационные системы бизнес-класса. Спецификация этих систем в большинстве случаев состоит из двух основных компонентов: функционального и информационного. По способу сочетания этих компонентов подходы к представлению информационных систем можно разбить на два основных типа – *структурный* и *объектно-ориентированный*.

Разумеется, объектно-ориентированные методы также являются структурными. Но исторически в программной инженерии этот термин закрепился за рядом дисциплин: структурное программирование, структурный дизайн, структурный анализ. В структурных технологиях функциональная и информационная модели строятся отдельно, чаще всего в виде диаграмм потоков данных и диаграмм «сущность-связь». Объектно-ориентированные технологии рассматривают информацию неотделимо от процедур ее обработки.

Объектно-ориентированные технологии доминируют в области создания операционных систем, средств разработки и исполнения приложений, программных систем реального времени. Структурные технологии – в области создания бизнес-систем, поскольку они максимально приспособлены для взаимодействия с заказчиками и пользователями, не являющимися специалистами в области информационных технологий.

Анализ опыта разработок информационных систем показал, что активное привлечение пользователей на этапах выявления требований и постановки задачи является крайне важным. При разработке систем бизнес-класса основные усилия затрачиваются именно на понимание и спецификацию требований пользователя.

В качестве инструментальных средств выполнения проектных работ используются пакеты программ, приведенные в табл. 2.1. С краткой характеристикой этих пакетов, приведенных в табл. 2.1, можно познакомиться, например, в [17].

Таблица 2.1

Инструментальные средства моделирования бизнес-процессов

Пакеты программ	Функциональные возможности
Design/IDFF компании Meta Software	Для функционального и информационного моделирования, анализа и проектирования бизнес-процессов.
EasyABC Plus компании ABC	Для функционально-стоимостного анализа бизнес-процессов.
ServiceModel компании PROMODEL	Для имитационного моделирования выполнения бизнес-процессов, анализа зависящих от времени характеристик распределения ресурсов.
Пакет S-Designor компании Powersof	Для создания концептуальных и физических моделей структур базы данных.
Пакет BPwin и Erwin компании Platinum	Для функционального моделирования и построения модели данных.
Семейство ARIS Toolset	Включает не только средства моделирования, но и мощный инструментальный анализ, а также модули-интерфейсы, обеспечивающие интеграцию с другими системами – ERwin, Designer-2000 и др.

2.6. Закономерности информетрии и их применение для исследования процессов организации и распространения информационных потоков

Понятие об информетрии

Термин *информетрия*⁶⁶ был введен в начале 80-е гг. XX в. по аналогии с наукометрией, библиометрией для краткой характеристики количественных методов исследования научно-технической информации. Наиболее полно этот термин был раскрыт **В.И. Горьковой** [4].

Закономерности информетрии определяют распределения информации в документальных информационных потоках

⁶⁶ **Bonitz M.** Scientometrie, Bibliometrie, Informetrie / M. Bonitz // Zbl. Bibliotheksw. – 1982. – Vol. 86. – № 1. – S. 19–24.

(ДИП). количественные и качественные параметры организации частотных словарей, использования слов в текстах документов. ДИП формируют официальные, периодические и продолжающиеся издания и др. опубликованные и неопубликованные документы *научно-технической информации*

Первые результаты исследований лингвистических закономерностей естественного языка были получены *Дж. Эсту (J. B. Estoup, 1916 г.)*, *А. Лоткой (A. J. Lotka, 1926 г.)*.

Характеристику качественных свойств частотных словарей определил в 1916 г. Дж. Эсту, который обнаружил, что частота использования слов в тексте обратно пропорциональна его номеру в частотном словаре.

Наиболее исследованы закономерности информетрии *Г. Ципфа, Б. Мальденброта, С. Брэдфорда, Б. Викери*. которые в связи с этим часто называют законами.

Закон Ципфа

George K. Zipf (Дж. Ципф), или в некоторых современных переводах – *Г. Зунф*) в начале 30-х гг. XX в. на основе статистических исследований получил следующую закономерность⁶⁷.

Пусть есть текст длиной N слов и словарь объемом m слов с указанием частоты появления слова в тексте. Слова в словаре расположены в порядке убывания их по частоте и пронумерованы от 1 до m . Ранг, равный 1 присваивается слову, частота появления которого наибольшая; ранг, равный m – наименее употребляемому слову. Тогда

$$p_{r_i} = \frac{f_{r_i}}{N},$$

⁶⁷ *Zipf G. K. Human behavior and principle of least effort / G. K. Zipf // Cambridge (Mass.): Addison – Wesley, 1949. – Vol. XI. – 573 p. Zipf G. K. Selected studies of the principle of relative frequency in language. / G. K. Zipf // Cambridge (Mass.): Harvard Univ. Press., 1932. – 51 p.*

где p_{r_i} – относительная частота появления слова в тексте;
 f_{r_i} – абсолютная частота появления слова r_i ранга в тексте
определенной длины;

N – число слов в тексте; r_i – ранг слова, где $1 \leq i \leq m$.

Если умножить вероятность или относительную частоту обнаружения слова в тексте на ранг r_i слова, то получим:

$$p_{r_i} r_i = f_{r_i} \frac{r_i}{N} = k,$$

где k – константа; $1 \leq r_i \leq m$.

Если преобразовать эту формулу, то получим $p_{r_i} = \frac{k}{r_i}$,

т. е. функцию типа $y = k/x$, график которой – равносторонняя гипербола.

Таким образом, на основе анализа полученных зависимостей Ципф предложил эмпирическую формулу, устанавливающую связь между частотой появления слов в тексте и его рангом в словаре:

$$p_{r_i} = k r_i^{-1},$$

где k – эмпирически определяемая константа, изменяющаяся для разных текстов.

При этом $1 \leq r_i \leq m$;

p_1 – частота наиболее употребимого слова;

p_m – частота наименее употребимого слова;

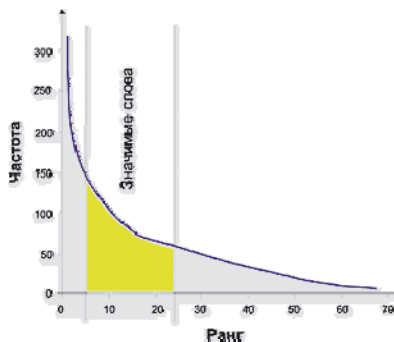


Рис. 2.9. Закон Ципфа

$p_{r_i} = \varphi(r_i)$ – «гиперболическая лестница», поскольку ранговое распределение имеет ступенчатый характер (ряд слов появляется с одинаковой частотой), но при аппроксимации можно считать распределение Ципфа гиперболой (рис. 2.9).

Значение константы k в разных языках различно, но внутри одной языковой группы остается неизменным, какой бы текст мы ни взяли. Так, имеются исследования, показывающие, что, например, для английских текстов константа Ципфа приблизительно равна 0,1; а для русского языка – примерно 0,06–0,7.

Поэтому Ципф приводил также запись этой закономерности в форме

$$p_{r_i} = kr_i^{-1} = 0,1r_i^{-1}, \quad 1 < r_i < m,$$

где $k = 0,1$ (для естественных языков).

Опираясь на экспериментальные данные, собранные в результате статистического исследования многих текстов на различных языках, Ципф обнаружил также, что распределение слов естественного языка подчиняется единому простому закону, который он назвал «принципом наименьшего усилия»: выражая мысли с помощью языка, мы подвергаемся действию двух противоположных сил – силе унификации и силе диверсификации, проявляющихся, с одной стороны, в необходимости быть понятыми, а с другой – в желании выразить мысль покороче.

Ципф установил, что частота и количество слов, входящих в текст с этой частотой, связаны между собой. Если построить зависимость количества слов в данной частоте от частоты вхождения слова, то получится кривая аналогичная рис. 2.9, которая будет сохранять свои параметры для всех без исключения созданных человеком текстов с некоторыми отклонениями для разных естественных языков (рис. 2.10).

Эту закономерность считают иногда *вторым законом Ципфа*.

Исследования показали, что наиболее значимые слова лежат в средней части гиперболы (см. рис. 2.9). Слова, которые попадают слишком часто, в основном оказываются предлогами, местоимениями, в английском языке – артиклями и т. п. Редко встречающиеся слова тоже в большинстве случаев не имеют решающего смыслового значения

От того, как будет выставлен диапазон значимых слов, зависят свойства информационно-поисковой системы.

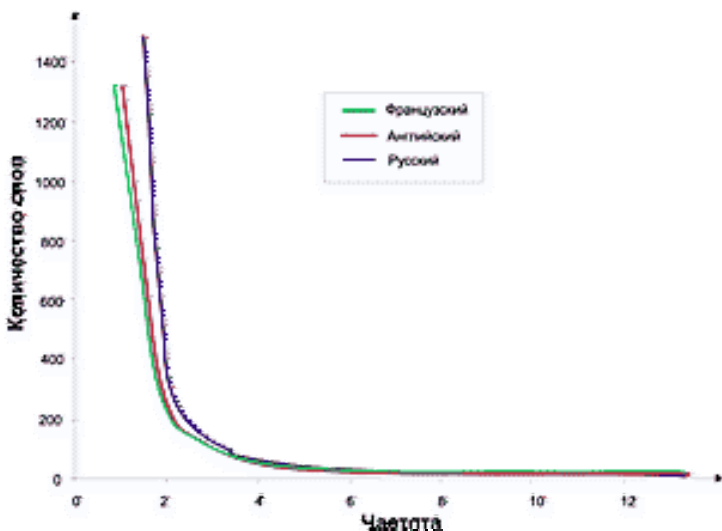


Рис. 2.10. Второй закон Ципфа

Если поставить широко – нужные термины потонут в море вспомогательных слов; если установить узкий диапазон – можно потерять смысловые термины. В каждой поисковой системе эта проблема решается по-своему, с учетом общего объема текста, специальных словарей и т. п.

Таким образом, закономерности Ципфа отражают некоторое общее свойство, присущее разным языкам. Это свойство заключается в том, что в каждом тексте на любом естественном языке имеется некоторое количество наиболее употребимых слов. Причем число этих слов значительно меньше общего числа слов, используемых в тексте.

Законы Ципфа универсальны. В принципе, они применимы не только к текстам.

В аналогичную форму выливается, например, зависимость количества городов от числа проживающих в них жителей. Характеристики популярности узлов в сети Интернет тоже отвечают законам Ципфа.

Закономерность Ципфа проявляется и при исследовании документальных информационных потоков (ДИП). В этом случае закон Ципфа представляют через абсолютную частоту появления слов:

$$f_i = Cr_i^{-1},$$

где f_i – абсолютная частота появления слова в текстах документального потока; r_i – ранг слова в ранговом распределении; C – частота появления слова 1-го ранга, которую для данного ДИП можно считать эмпирической константой.

Закон Мандельброта

Benoit Mandelbrot (Б. Мандельброт) в 1954 г.⁶⁸ предложил теоретическое обоснование эмпирически открытого закона Ципфа. Представляя слова в виде последовательности букв, разделенных пустыми промежутками, и присваивая знакам определенную «стоимость» (затраты усилий, времени), Мандельброт показывает, что словам можно приписать априорные вероятности, так чтобы их общая «стоимость» в среднем была минимальной и при этом количество информации оставалось инвариантным. На основе этих представлений математическим путем Мандельброт показал, что результирующее соотношение между частотой слова и его рангом соответствует эмпирическому закону Ципфа с небольшой поправкой:

$$P_{r_i} = kr_i^{-\gamma},$$

где P_{r_i} – относительная частота появления слова в тексте; r_i – ранг слова; k – эмпирическая постоянная; γ – величина, близкая к единице, но изменяющаяся в зависимости от свойств текста.

⁶⁸ **Мандельброт Б.** Теория информации и психологическая теория частот слов / Б. Мандельброт // Математические методы в социальных науках. – М.: Прогресс, 1973. – С. 326–337.

Коэффициент γ характеризует определенные свойства языка – степень его формализованности, при этом с уменьшением γ степень формализованности языка уменьшается.

Закон Брэдфорда

Этот закон открыт английским химиком и библиографом **С. Брэдфордом** в 1948 г.⁶⁹ на основе обнаружения общих принципов распределения публикаций по изданиям в разных областях (на примере изданий по геофизике и химии).

Основной смысл закона состоит в следующем: если научные журналы расположить в порядке убывания числа статей по конкретной проблеме, то журналы можно разбить на три зоны таким образом, чтобы количество статей в каждой зоне по заданной теме было одинаковым.

При этом в *первую зону*, названной Брэдфордом зоной *ядра*, входят профильные журналы, непосредственно посвященные рассматриваемой тематике. Количество журналов в зоне ядра невелико. *Вторую зону* образуют журналы, частично посвященные заданной области, причем их число существенно возрастает по сравнению с числом журналов в ядре. *Третья зона* – самая большая по количеству изданий – объединяет журналы, тематика которых весьма далека от рассматриваемой предметной области.

При равном числе публикаций в каждой зоне число источников (в данном примере – наименований журналов) резко возрастает при переходе от одной зоны к другой. Брэдфорд установил, что число наименований журналов в третьей зоне примерно во столько раз больше, чем во второй зоне, во сколько раз число наименований журналов во второй зоне больше, чем в ядре:

⁶⁹ **Bradford S. C.** Documentation / S. C. Bradford. – London: Grosby Lockwo, 1948. – 156 p.; **Brookes B. C.** Theory of Bradford law / B. C. Brookes // J. Doc. – 1977. – Vol. 33. – № 3. – P. 180–209.

$$\frac{P_3}{P_2} \approx \frac{P_2}{P_1} \approx a \quad \text{и} \quad \frac{P_3}{P_4} \approx a^2, \quad \text{т. е.} \quad P_3 : P_2 : P_1 \approx 1 : a : a^2 : a^3,$$

где $P_1 : P_2 : P_3$ – число наименований журналов в 1-й, 2-й и 3-й зонах соответственно.

Закон Викери

Б. Викери⁷⁰ уточнил модель **С. Брэдфода**. Он выяснил, что журналы, проранжированные в порядке уменьшения в них статей по конкретному вопросу, можно разбить не на три зоны, а на любое число зон. Основной смысл закономерности Викери (рис. 2.11): если периодические издания расположить в порядке уменьшения в них количества статей по конкретному запросу, то в полученном списке можно выделить ряд зон, каждая из которых содержит одинаковое количество статей.

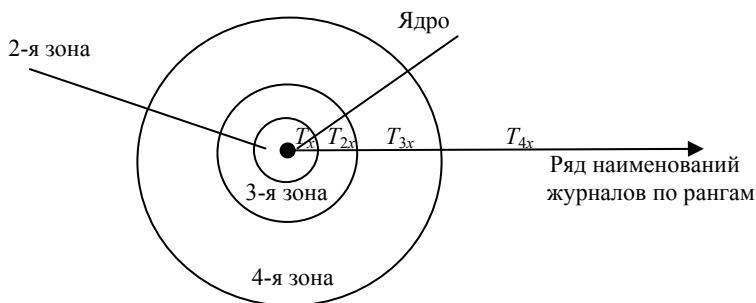


Рис. 2.11. Закон Викери

При этом число журналов в первой зоне и нарастающее их число в последующих зонах соотносятся следующим образом:

$$T_x : T_{2x} : T_{3x} : T_{4x} : \dots = 1 : a : a^2 : a^3 : a^4 : \dots$$

⁷⁰ **Vickery B. C.** Bradford's law of scattering / B. C. Vickery // J. Doc. – 1948. – Vol. 4. P. 198 – 203.

где x – количество статей в каждой зоне; T_x – количество журналов, содержащих x статей; T_{2x} , T_3 , T_{4x} , ... – количество журналов, содержащих $2x$, $3x$, $4x$ и т.д. статей соответственно.

Часто этот закон называют законом Брэдфорда в толковании Викери.

Закономерность концентрации–рассеяния В. И. Горьковой

В работах **В. И. Горьковой** [4] было показано, что количественные закономерности строения ДИП могут быть выражены как посредством количественных параметров ранговых распределений (закономерность Ципфа), так и посредством параметров упорядоченных потоков элементов ДИП (закономерность Брэдфорда–Викери), т.е. что эти закономерности отображают одно феноменологическое явление.

Смысловый анализ эмпирических данных строения подсистем ДИП, проведенный в работах **В. И. Горьковой** и ее учеников ⁷¹, позволил зафиксировать наличие специфических

⁷¹ **Горькова В. И.** Информетрия: Количественные методы в научнотехнической информации / В. И. Горькова // Итоги науки и техники. – Сер. Информатика. – Т. 10. – М.: ВИНТИ, 1988. – 328 с.; **Горькова В. И.** Анализ документальных информационных потоков и изучение запросов потребителей информации: Лекции / В. И. Горькова, Т. И. Гусева. – М.: ИПКИР, 1974. – 60 с.; **Горькова В. И.** Совершенствование системы информационного обеспечения на основе статистического анализа информационных потребностей специалистов / В. И. Горькова, Б. В. Петренко. – Минск: БелНИИНТИ, 1973. – 543 с.; **Горькова В. И.** Ранговое распределение на множествах научнотехнической информации / В. И. Горькова // НТИ. – 1969. – Сер. 2 – № 7. – С. 5–11.; **Горькова В. И.** Статистические оценки статистических совокупностей документальных информационных потоков / В.И. Горькова // НТИ. – 1971. – Сер. 2. – № 12. – С 14–20.; **Горькова В. И.** Закономерности распределения публикаций в периодических и продолжающихся изданиях по электротехнике и энергетике / В. И. Горькова, С. П. Меллион // НТИ. – 1968. – Сер. 2. – № 11 – С. 3–7.; **Горькова В. И.** Частотное распределение множества ключевых слов / В. И. Горькова, К. И. Нумычева // НТИ. – 1970. – Сер. 2. – № 6. – С. 3–8.; **Горькова В. И.** Системное исследование документального информационного потока / В. И. Горькова // Системные исследования: Ежегодник, 1979. – М.: Наука, 1980. – С. 240–266.

функциональных свойств различных частотных зон ранговых распределений и упорядоченных потоков. Научные журналы, наиболее продуктивные по числу научных статей для соответствующего тематического раздела, сосредоточены в зоне рангового распределения, где частоты появления наименований элементов наибольшие, т. е. в ядре рангового распределения.

Структурное *подмножества ядра* рангового распределения состоит из *профильных научных журналов*, которые составляют ~10 % от всех наименований журналов, содержащих публикации по данному тематическому разделу; при этом в научных журналах, входящих в ядро рангового распределения, содержится 50–60 % всех публикаций по данной тематике.

Тематическое содержание профильных журналов ядра рангового распределения определяет тематические признаки данного упорядоченного потока.

Для ядерного подмножества терминов рангового распределения характерно вхождение в него заглавных понятий, тематическое содержание которых определяет классификатор данного тематического раздела.

В зону ядра рангового распределения авторов публикаций входят имена исследователей, которым принадлежат основополагающие положения тематического раздела ⁷².

На основе исследования этих специфических особенностей ядра **В. И. Горьковой** был сделан вывод о том, что явление образования ядра рангового распределения системы ДИП можно трактовать как свойство системы ДИП концентрировать в ядре рангового распределения логическую информацию, определяющую основные понятия предметной области, объектов и методов исследования соответствующей отрасли науки (техники), научной дисциплины.

Явление образования ядра Горькова рассматривает как организацию высокого уровня, когда между элементами подмножества ядра существуют взаимосвязи, играющие более важную

⁷² **Горькова В. И.** Частотное распределение множества ключевых слов / В. И. Горькова, К. И. Нумычева // НТИ. – 1970. – Сер. 2. – № 6. – С. 3–8.

роль, чем связи между элементами других подмножеств. В качестве одного из важных параметров взаимосвязанности для формирования ядра **В. И. Горьковой** предложены меры «совстречаемости» понятий.

Элементы ядра выступают в роли «организаторов» подсистемы ДИП в конкретной отрасли науки или научной дисциплине, т.е. выполняют функции организующих отношений системы ДИП. Наименования элементов зоны ядра ДИП определяют условия включения элементов всех других частотных зон ранговых распределений в данную упорядоченную совокупность.

Одновременно с явлением образования ядра рангового распределения элементов ДИП по тематическому разделу наблюдается рассеяние элементов.

Так, если ядро содержит ~10 % от всех наименований элементов ДИП, что составляет 50–60 % упорядоченного потока подсистем ДИП, то в зоне рассеяния при наименовании элементов ~90 % сосредоточено всего 40–50 % упорядоченного потока документов.

Зона рассеяния рангового распределения определяет структуру связи подсистем ДИП различных отраслей науки (техники) или научных дисциплин, взаимно обогащающих друг друга семантической информацией.

Рассмотренные функциональные свойства двух зон рангового распределения и упорядоченного потока огрубленно трактуют функциональные свойства ранговых распределения. В зоне рассеяния можно выделить третье структурное подмножество элементов с малой частотой появления, обладающее специфическими свойствами. В дальнейшем, видимо, возможна и более детальная структуризации зоны рассеяния.

Рассмотренное феноменологическое явление концентрации и рассеяния элементов ДИП было названо *закономерностью концентрации–рассеяния*.

Действие закономерности концентрации–рассеяния основано на том, что упорядоченное множество элементов ДИП имеет два структурно-функциональных свойства.

Во-первых, концентрировать подмножества ядерных элементов, функция которых состоит в том, чтобы:

а) сосредотачивать логическую информацию, определяющую основные предметы, объекты и методы исследований в отрасли науки;

б) выполнять роль «организующих отношений».

Во-вторых, рассеивать подмножество элементов, функция которых состоит в том, чтобы:

а) определять структурные связи подсистем ДИП различных отраслей науки и научных дисциплин, обогащающих друг друга семантической информацией;

б) обеспечивать определенную целостность всей подсистемы ДИП.

Закономерность концентрации–рассеяния позволяет использовать параметры аппроксимирующих функций, описывающих ранговые распределения и упорядоченные потоки в качестве параметров, определяющих функциональные свойства упорядоченных совокупностей элементов ДИП.

Следует еще раз обратить внимание на тот факт, что численные меры упорядоченности строения ДИП, определяемые статистическими методами только тогда могут использоваться в практике научно-информационной деятельности, когда они сочетаются с методами исследования функциональных свойств упорядоченных совокупностей ДИП. С примерами статистических оценок параметров закономерностей ДИП можно познакомиться в работах В. И. Горьковой ⁷³.

Перспективы развития информетрии

На основе идей законов Ципфа–Мандельброта и Брэдфорда–Викери, закономерности концентрации–рассеяния, сформулирован-

⁷³ *Горькова В. И.* Информетрия: Количественные методы в научно-технической информации / В. И. Горькова // Итоги науки и техники. – Сер. Информатика. – Т. 10. – М.: ВИНТИ, 1988. – 328 с.; *Горькова В. И.* Анализ документальных информационных потоков и изучение запросов потребителей информации: Лекции / В. И. Горькова, Т. И. Гусева. – М.: ИПКИР, 1974. – 60 с.

ной **В. И. Горьковой**, развиваются методики автоматизации индексирования и анализа текстов, введения весовых коэффициентов терминов (см. ссылки в <http://www.medialingvo.ru>).

Вводятся меры веса ключевых слов.

Так, в работах **Спарка Джонса** экспериментально показано, что если N число документов и n число документов, в которых встречается данный индексный термин (ключевое слово), то вычисление его веса по формуле:

$$W = \log \frac{N}{n} + 1$$

приводит к более эффективным результатам поиска, чем без использования оценки значимости индексного термина, т.е. определенное значение имеет не только частота применения слова в конкретном документе, но и число документов, в которых это слово встречается.

Вводятся логарифмические меры.

Например, чтобы избавиться от лишних слов и в тоже время поднять рейтинг значимых слов, вводят инверсную частоту термина

$$i_x = \log (N / n_i),$$

где N – количество документов в базе данных; n_i – количество документов с термином i .

А затем каждому термину присваивают весовой коэффициент, отражающий его значимость в форме:

$$j = j_x / i_x,$$

где j – вес термина i в документе; j_x – частота термина i в документе; i_x – инверсная частота термина.

В новом смысле используется термин «ядро».

В 1995 г. на симпозиуме в Дублине предложена интересная и полезная для совершенствования информационного поиска идея «Дублинского ядра» (Dublin Core)⁷⁴, основанная на формировании метаданных, зафиксированных в спецификации определенного стандарта, и представлении k -го документа множеством пар $D_k = \{N_{ik}, V_{ik}\}$, где N_{ik} – имя i -го элемента метаданных Дублинского ядра в описании содержания k -го документа; V_{ik} – значение этого элемента метаданных. Аналогично описывается запрос.

⁷⁴ **Когаловский М. Р.** Перспективные технологии информационных систем / М. Р. Когаловский. – М.: ДМК Пресс; М.: Компания АйТи, 2003. – 288 с.

Перспективным представляется использование для формирования «Дублинского ядра» закономерности концентрации–рассеяния.

Возрастает интерес и к способам оценки текстов. Например, к работам *Г. Луна*⁷⁵, в которых предложения текста оцениваются в соответствии с параметром:

$$V_{np} = \frac{N_{zc}^2}{N_c},$$

где V_{np} – значимость предложения; N_{zc} – число значимых слов в предложении; N_c – полное число слов в предложении.

Используя этот критерий, из любого документа можно отобрать некоторое число предложений. Понятно, что они не будут составлять членораздельного текста. Нужно учитывать также, что значимые слова должны браться из тематического тезауруса или отбираться экспертом. По той причине методика может лишь помочь человеку, а не заменить его (во всяком случае, на современном этапе развития вычислительной техники).

Закономерности организации ДИП, введения количественных мер терминов, предложений и др. компонентов текста полезно использовать на всех этапах создания информационно-поисковых систем: при комплектовании информационных фондов, создании информационно-поисковых языков и логико-семантического аппарата ИПС, при организации справочно-информационного обслуживания в библиотеках и отделах научно-технической информации, при создании и совершенствовании классификационных систем, выявлении тенденций роста и старения ДИП, при аналитико-синтетической обработке текстовой информации.

В последнее время на основе идеи закономерности концентрации–рассеяния разрабатываются методы выявления информационного ядра предметной области при построении информационной системы для реорганизации бизнес-процессов, при создании виртуальных предприятий.

⁷⁵ *Luhn H. P.* Automatic creation of literature abstracts BM / H. P. Luhn // Journal of Research and Development, 2, 1958, p. 159–165.

Вопросы для самоконтроля

1. Виды информационных процессов.
2. Система функций Ф.Е. Темникова.
3. Понятие о методах алгебраического агрегирования и их применении при моделировании процессов передачи информации.
4. Понятие о теории выдвижения и проверки статистических гипотез А. Вальда и ее применении при моделировании процессов передачи информации.
5. Принципы и критерии принятия решения при моделировании информационных процессов, базирующиеся на теории выдвижения и проверки статистических гипотез А. Вальда;
6. Основные информационные процессно-ориентированные технологиями и направления их применения.
7. Методика формирования и исследования информационных процессов на основе применения метода постепенной формализации модели;
8. Методика исследования информационных процессов на основе метода постепенной формализации модели.
9. Понятие об информетрии.
10. Закономерности информетрии и возможности их применения при анализе информационных потоков;
11. Законы Г. Ципфа.
12. Какое уточнение в закон Г. Ципфа внес Б. Мандельброт?
13. Закон С. Брэдфорда.
14. Закон Б. Викери.
15. Закономерность концентрации – рассеяния В.И. Горьковой.

Глава 3. ПОНЯТИЕ ОБ ИНФОРМАЦИОННЫХ СИСТЕМАХ И КЛАССИФИКАЦИИ ИС

3.1. Из истории разработки информационных систем

В середине XX в., когда появилась возможность обработки информации с помощью электронно-вычислительных машин (ЭВМ), был распространен термин *системы обработки данных* – СОД (термин был введен в системах радиоуправления ракетами). По мере увеличения памяти ЭВМ основное внимание стали уделять проблемам создания *баз данных* (БД). В последующем появились термины *автоматизированные информационные системы* (АИС), *информационно-поисковые системы* (ИПС), *информационные системы поддержки принятия решений* (СППР).

Начиная с 1960-х гг. в истории развития информационных систем нашей страны относительно независимо сформировались два направления:

- разработка *автоматизированных информационных систем* как первой очереди *автоматизированных систем управления* (АСУ);
- разработка *автоматизированных систем научно-технической информации* – АСНТИ.

Работы по их созданию начались практически одновременно.

Первое направление – разработка АИС и АСУ – было инициировано научно-техническим прогрессом и возникшими в связи с этим проблемами организационного управления.

Зарубежная практика шла по пути разработки отдельных программных процедур для бухгалтерии, учета материальных ценностей и т. п., и основные работы проводились в направлении исследования и совершенствования возможностей вычислительной техники, разработки средств, обеспечивающих наиболее рациональную организацию информационных массивов, удобный для

пользователя интерфейс, наращивание памяти компьютера и повышение скорости обработки информации.

В нашей стране задача обеспечения информацией управленческих работников с самого начала поставлена системно. Была разработана классификация АСУ, ориентированная на разные уровни управления предприятиями, включая их структурные подразделения, отраслями и государством в целом. Разрабатываемые отраслевые АСУ (ОАСУ), региональные или территориальные АСУ (РАСУ, ТАСУ) и АСУ предприятий (АСУП) объединялись в Общегосударственную автоматизированную систему (ОГАС) для сбора, хранения и поиска информации.

Для управления разработками автоматизированных систем были подготовлены и изданы соответствующие руководящие методические материалы ⁷⁶, в которых отражались теория и основная терминология, обязательная для использования при представлении отчетных материалов по разработке и внедрению АИС и АСУ.

Следует отметить, что АИС создавались как *фактографические информационные системы* с представлением информации пользователям в виде регламентированных форм, в которых фактографическая информация была сгруппирована в соответствии с алгоритмами решаемых на ее основе прикладных задач управления.

Второе направление – разработка АСНТИ – инициировалось научной общественностью как объективная необходимость возрастания роли информации во всех сферах деятельности, и прежде всего, в науке и технике.

Зарубежные работы в этой области первоначально проводились в форме создания разрозненных локальных центров сбора и анализа

⁷⁶ **Общепромышленные** руководящие методические материалы по созданию автоматизированных систем управления предприятиями и производственными объединениями (АСУП). – М.: Статистика, 1977. – 284 с.; **Общепромышленные** руководящие методические материалы по созданию многоуровневых интегрированных АСУ производственным объединением (предприятием). – М.: ГКНТ, 1986. – 22 с.; и др.

информации, но в последующем стала наблюдаться тенденция к их объединению, по крайней мере, в рамках одной отрасли.

В нашей стране задача обеспечения научно-исследовательской и управленческой деятельности научно-технической информацией сразу, так же, как и при создании АСУ, была поставлена системно. Были подготовлены и утверждены Техническое задание ⁷⁷ и Единый порядок ⁷⁸ разработки общегосударственной *Автоматизированной системы научно-технической информации* – АС-НТИ, в дальнейшем – Государственной автоматизированной системы научно-технической информации (ГАСНТИ) в стране [9, 19]), в структуре которой предусмотрены общегосударственные, отраслевые и региональные органы НТИ, отделы или бюро научно-технической информации (ОНТИ, БТИ) на предприятиях, в научно-исследовательских институтах и в других организациях.

Для исследования проблем создания систем НТИ, семантического анализа текста, реферирования и перевода с одного языка на другой развивалась теория научно-технической информации, базирующаяся на применении математической лингвистики и семиотики. Был введен ряд понятий – *информационно-поисковая система* (ИПС) и ее структура, *информационно-поисковый язык* (ИПЯ), *дескриптор*, *тезаурус*, *грамматика*, *семантика*, понятия *релевантности* и *пертинентности* информационного поиска как характеристики ИПС, на основе которых их можно сравнивать и выбирать, и другие понятия, которые играют важную роль в развитии теории и практики информационного обеспечения предприятий и организаций.

Благодаря возросшим возможностям вычислительной техники, и особенно в результате появления персональных ЭВМ, рассмотренные направления сближаются. По мере роста объемов и усложнения структуры информационных массивов АИС возникает необходимость в заимствовании понятий теории информационно-

⁷⁷ *Техническое задание на совершенствование Единой системы научно-технической информации в стране.* – М.: ГКНТ СМ СССР, 1972. – 34 с.

⁷⁸ *Единый порядок и технические условия разработки и внедрения автоматизированных подсистем обработки, поиска, хранения, выдачи и передачи информации.* – М.: ГКНТ при СМ СССР, 1972. – 35 с.

го поиска, разработанной ранее для систем НТИ. В свою очередь, в решении проблемы разработки документальных АС НТИ полезны понятия *функциональной части (ФЧ)* и *обеспечивающей части (ОЧ)*, введенные в нормативных документах по разработке АИС и АСУ. Кроме того, для управления современными предприятиями и организациями необходим спектр информации разного вида и назначения: от *фактографической* до *документальной*.

При введении принципов самокупаемости и самофинансирования в период косыгинских реформ возросла роль еще одного важного вида информации – *нормативно-методической*, регламентирующей деятельность предприятий, для чего стали разрабатывать системы нормативно-методического обеспечения управления (СНМОУ) и их автоматизированный вариант – АСНМОУ. АСНМОУ занимают промежуточное место между фактографическими и документальными информационными системами, поскольку законодательные нормативно-методические документы (законы, постановления, уставы, стандарты, инструкции и т. п.) – тексты, но, как правило, достаточно хорошо структурированные, что облегчает извлечение из них фактографической информации – правовых норм, статей, функций управления и т. п. Поэтому СНМОУ разрабатываются как *документально-фактографические* информационные системы.

Работы по созданию централизованных АСУ и АСНТИ были приостановлены в связи с политическими и экономическими преобразованиями 1990-1991 гг. Вплоть до 1999 г., на предприятиях и в организациях в основном продолжали развиваться существовавшие локальные информационные системы и базы данных, приобретались и адаптировались зарубежные программные продукты, создавались отдельные страницы документальной и фактографической информации в сети Интернет.

С 2000-х гг. происходит постепенная адаптация к новым экономическим условиям, и в частности, информационные ресурсы и некоторые принципы организации ГСНТИ сохранились в системе,

которую в настоящее время условно называют Российской ГСНТИ (РГСНТИ). Структура этой системы приведена в разделе 5.1.

С возрастанием роли нормативно-правовой информации при внедрении рыночных принципов управления экономикой создаются автоматизированные системы *нормативно-правовой документации* (АС НПД) разного рода и назначения (системы «Консультант +», «Кодекс», «Гарант» и др.), которые чаще называют справочно-правовыми системами (СПС).

На новой технической базе с использованием современных технологий развиваются информационные системы (ИС) государственного управления на федеральном и муниципальном уровнях, в частности, ИС государственной статистики; ИС природных ресурсов; государственные системы экономической, финансовой и научно-технической информации, внешнеэкономической деятельности; ИС библиотечной сети РФ и другие ИС специального назначения (см. примеры таких систем в [11]).

3.2. Виды и классификации информационных систем

В настоящее время не существует единой классификации информационных систем. Это связано с тем, что различные направления (СОД, АИС и АСУ, АСНТИ) долгое время развивались относительно независимо, поэтому и классификации в каждом из направлений предлагались также независимые.

Обобщенная классификация, базирующаяся на основных видах информации – *документальной* и *фактографической*, приведена на рис. 3.1.

При классификации АСУ и АИС применяют различные признаки классификации: степень автоматизации (форма участия человека в системе при выполнении функций управления); назначение системы (тип автоматизируемого процесса); степень использования технических средств (ТС) человеком для принятия управленческих решений; структурированность проблем, решаемых системой; степень централизации обработки информации; уровень управления.

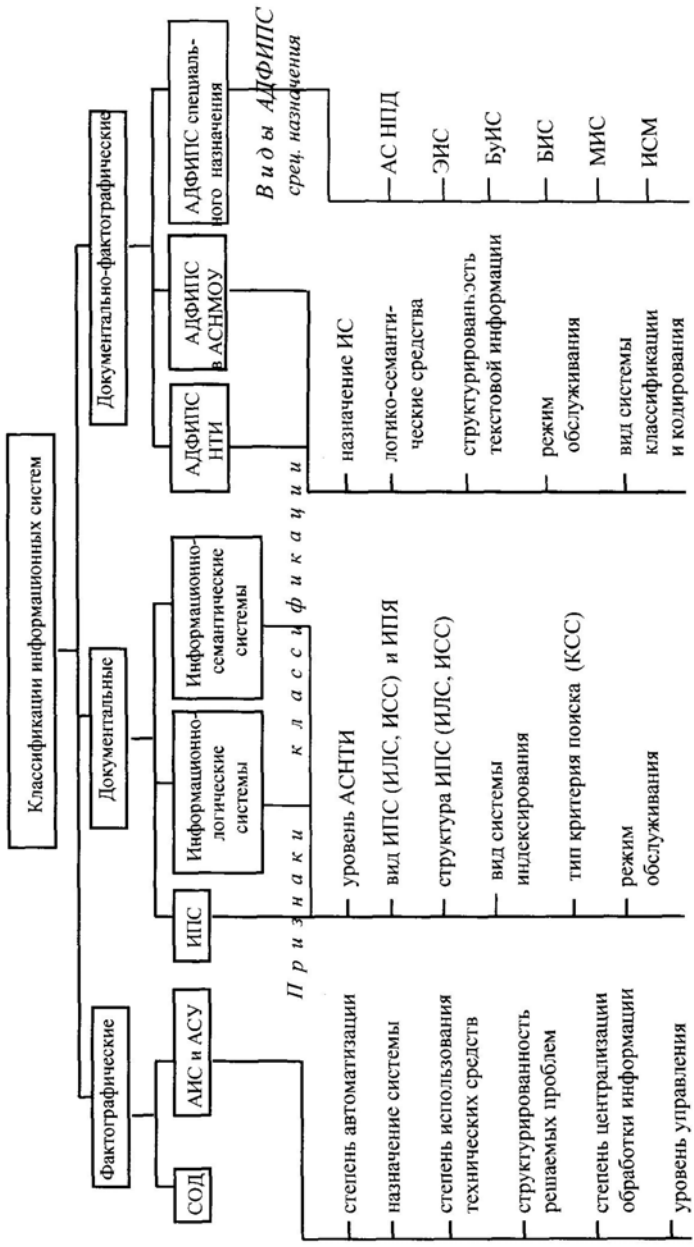


Рис. 3.1. Классификация информационных систем

По степени автоматизации АСУ и АИС делят на *автоматизированные* и *автоматические*.

В *автоматизированных* системах управления ТС используются для сбора, передачи, обработки информации и выдачи управленческих решений. При этом окончательное решение принадлежит человеку. ТС выполняют весь комплекс заранее предписанных действий. Управляющая информация непосредственно передается исполнительным механизмам, регулируя их работу без участия человека. За человеком остается функция контроля исправности ТС.

По назначению системы (характеру использования) различают такие типы АСУ и АИС:

Автоматизированные системы управления технологическими процессами (АСУТП), *системы организационного или административного управления* (АСОУ), *автоматизированные системы научных исследований* (АСНИ), *системы автоматизированного проектирования* (САПР).

По степени использования ТС человеком для принятия управленческих решений АСУ и АИС делят на:

информационные, в которых обеспечивается сбор и частичная систематизация первичной информации, и

управляющие, которые кроме этого обеспечивают выработку некоторых управленческих решений, передающихся непосредственно работникам или исполнительным механизмам по заранее заданным программам.

Информационные системы в свою очередь делят на *информационно-справочные*, которые выполняют задачу обеспечения руководства необходимыми справочными данными по запросам, и *информационно-советующие*, в которых кроме сбора, передачи и обработки информации подготавливаются рекомендации, используемые при принятии решений.

Управляющие системы делят на *информационно-управляющие* (например, система управления проектами), *управляющие системы с запрограммированными командами*, в которых решаются задачи регулирования (например, АСУТП), *самонастраивающиеся* и *самообучающиеся* системы, функционирование которых меняется в зависимости от воздействия внешней среды.

При классификации ИС можно учитывать *структурированность проблем*.

Задачи, решаемые в системе, можно считать *структурированными*, если их решения носят повторяющийся характер (учет материалов, расчет заработной платы и др.). Часть АИС предназначена для решения таких задач. *Плохо структурированные задачи* – это такие, характер решения

которых представляется не полностью определенным. Информационные системы, используемые для их решения, делят на системы, ориентированные на обработку данных и формирование специальных управленческих отчетов, и системы, в которых сведения из отчетов используются для принятия решения на основе предлагаемых альтернатив.

По степени централизации обработки информации выделяют системы, имеющие несколько уровней обработки информации (характерны для крупных объектов), системы с централизованной обработкой информации (характерны для средних объектов), системы коллективного пользования (характерны для малых объектов).

По уровню управления различают системы, относящиеся к: низшему уровню управления (АСУП – для уровня предприятий и организаций, АСОУ, АСУТП и т. д.);

среднему уровню управления (ОАСУ – отраслевые АСУ, РА-СУ – республиканские и региональные АСУ или ТАСУ – АСУ территориальных органов и др.);

высшему уровню управления (ОГАС – общегосударственная автоматизированная система).

Возможны и более детальные классификации [7, 11 и др.]: по уровню охвата пользователей (индивидуальные, корпоративные, региональные и т. п.); по назначению (управленческие, офисные, научно-исследовательские, редакторские); по характеру функционирования (ИС в режиме реального времени, ИС стратегического планирования); по уровню организации массивов данных и др.

Для систем **научно-технической информации** также разрабатывали различные классификации.

По уровням АСНТИ (*общегосударственные, отраслевые, региональные, ОНТИ и БТИ предприятий и организаций*).

По признаку *виды документальных ИС* информационно-поисковые системы прошли следующие стадии развития: собственно *информационно-поисковые системы (ИПС), информационно-логические системы (ИЛС), информационно-семантические системы (ИСС)*.

По режимам *информационного обслуживания*: по стандартным запросам (СЗ), системы избирательного распространения информации (ИРИ), дифференцированного обслуживания

руководителей (ДОР), ретроспективного поиска (РП) по произвольным запросам.

По видам ИПЯ различают ИПЯ, и соответственно ИПС, *без грамматики и с грамматикой*. ИПЯ классифицируют и более глубоко – по *парадигматическим отношениям, лексике языка и синтагматическим отношениям*.

Поскольку лексика и синтагматические отношения характеризуют текст, описанный на конкретном ИПЯ, а парадигматические отношения представляют собой внетекстовые смысловые отношения между лексическими единицами ИПЯ, которые устанавливаются на основании потребностей информационного поиска, то и ИПС классифицируют в зависимости от развитости парадигматических отношений ИПЯ.

По видам *критериев поиска* (критериев смыслового ответа); по тематическому профилю комплектования; формам носителей информации; уровням интеграции лексики и другим, специфическим для ИПС НТИ признакам классификации.

По видам *структур ИПС* различают:

- *ИПС иерархической структуры*, в которых все лексические единицы ИПЯ связаны сильными парадигматическими отношениями (подчинения и соподчинения) и образуют в совокупности иерархическую классификацию.

Иерархические классификации имеют вид древовидного графа или дерева понятий. На практике их представляют в табличной форме записи.

- *ИПС фасетной структуры*, в которых лексические единицы ИПЯ предварительно группируются в фасеты, а иерархические отношения устанавливаются внутри фасетов.

Фасеты, следующие друг за другом в определенной последовательности, образуют фасетную классификацию. Преимущество фасетной структуры по сравнению с иерархической заключается в *многоаспектности*, так как количество фасетов и субфасетов в принципе не ограничивается. Важной особенностью многоаспектной классификации является также то, что последовательность признаков, и соответственно фасетов, может быть произвольной.

- *ИПС неиерархической структуры*, в которых лексические единицы ИПЯ упорядочивают по внешним признакам, например, в алфавитном порядке.

Если иерархические и фасетные классификации строят в основном по заранее заданной схеме, то неиерархические языки получают эмпирически в процессе индексирования вводимых в ИПС документов.

Следует иметь в виду, что признаки, по которым выше были классифицированы системы управления, являются независимыми, вследствие чего каждая система, отнесенная к какому-то классу по одному признаку, может соответствовать нескольким классам по остальным признакам.

Предлагались и другие классификации.

Например, для того чтобы охарактеризовать систему НТИ более полно, разрабатывались многоаспектные классификации. Наиболее развитой из таких классификаций является *фасетная* классификация **А. В. Соколова**⁷⁹, в которой признаки классификации определяются *семантическими средствами* ИПС, т.е. по видам информационно-поисковых языков, методам (правилам, алгоритмам) индексирования, методам (правилам, алгоритмам) поиска по запросу.

В последующем представления о видах автоматизированных систем развивались. По мере развития вычислительных средств и технологий сети Интернет, методов и средств информационного поиска возникают новые термины, более точно характеризующие специфику информационных систем специального назначения. В развитие информационно-советующих и информационно-управляющих систем возникли *системы поддержки принятия решений, экспертные системы*, помогающие в принятии решений по проектированию и управлению.

Параллельно с *информационно-логическими* и *информационно-семантическими системами*, усиливающими интеллектуальные возможности информационно-поисковых систем научно-технической информации, развивается направление искусственного интеллекта, базирующееся на кибернетических исследованиях процессов в живых организмах, в том числе функционирования мозга, и моделировании этих процессов средствами

⁷⁹ **Соколов А. В.** Информационно-поисковые системы: учеб. пособие / А. В. Соколов. – М.: Радио и связь, 1981. – 161 с.

вычислительной техники, и на этой основе создаются *интеллектуальные информационные системы*.

На новой технической базе с использованием современных технологий развиваются информационные системы государственного управления на федеральном и муниципальном уровнях, системы экономической, финансовой и научно-технической информации, внешнеэкономической деятельности; и другие ИС специального назначения.

В их числе [11]: экономические информационные системы (ЭИС): *бухгалтерские информационные системы* (БуИС), *автоматизированные банковские информационные системы* (АБИС), *ИС фондового рынка*; *ИС в контроллинге*, *ИС бюджетирования*; *маркетинговые информационные системы* (МИС), *информационные системы мониторинга* (ИСМ); *ИС в кадровом менеджменте*; *ИС в налоговых органах, в страховании и пенсионном обеспечении*; *ИС в туристическом бизнесе*; *ИС в управлении недвижимостью*; *ИС таможенной и внешнеэкономической деятельности* и т. п.; *системы сопровождения жизненного цикла изделий* типа ИПИ-системы (ИПИ – информационная поддержка изделий) и системы управления проектными и инженерными данными предприятия – так называемые PDM-системы (Product Data Management Systems), включая CAD-системы (Computer-Aided Design – компьютерная поддержка проектирования), САМ-системы (Computer-Aided Manufacturing – компьютерная поддержка изготовления), САЕ-системы (Computer-Aided Engineering – поддержка инженерных расчетов) и т.п.; *системы управления взаимоотношениями с клиентами* (CRM – Customer Relationships Management); *ИС документооборота и делопроизводства* (ИСДД).

С развитием систем телекоммуникаций стал активно развиваться информационный рынок, получивший название «пятый рынок». Развиваются информационные технологии (ИТ или IT) – способы и средства сбора, обработки, хранения и распределения информации.

Для оперативного и интеллектуального анализа хранимых данных используются технология комплексного многомерного анализа данных OLAP (On-Line Analytical Proctssing) и технология интеллектуального анализа информации Data Mining.

На основе новых информационных технологий развиваются *процессо-ориентированные корпоративные информационные системы* – системы, обеспечивающие реализацию процессного

подхода к управлению предприятием. Менеджмент бизнес-процессов возник в рамках концепции CPI (Continuous Process Improvement) и его японского аналога TQM (Total Quality Management), согласно которой предполагается управление бизнес-процессом, как единым целым. Этот подход положен в основу стандартов серии ISO 9000, которые разработаны Международной организацией по стандартизации (International Standard Organization – ISO) и определяют базовый набор принципов менеджмента качества.

Для оптимизации и автоматизации внутренней деятельности фирмы применяются ИС планирования потребности в материалах – MRP (Material Requirements Planning); ERP (Enterprise Resources Planning) системы, направленные на усовершенствование таких процессов, как планирование, изготовление, учет и контроль.

В основе ERP-систем лежит принцип создания единого хранилища (репозитория) данных, содержащего всю корпоративную информацию, включая финансовую информацию, производственные данные, данные по персоналу и т.д. ERP-система – это набор интегрированных приложений, позволяющих создать интегрированную информационную среду (ИИС) для автоматизации планирования, учета, контроля и анализа всех основных бизнес-операций предприятия. В соответствии с современными требованиями APICS, ERP-система должна помимо ядра, реализующего стандарт MRPII, включать ряд модулей: управления логистическими цепочками SCM; усовершенствованного планирования и составления производственных графиков APS (Advanced Planning and Scheduling); CRM; электронной коммерции EC (Electronic Commerce); управления данными об изделии PDM; надстройку Business Intelligence на основе технологий OLAP и др.

В конце 1990-х гг. был разработан также стандарт CSRP (Customer Synchronized Resource Planning), который охватывает взаимодействие предприятия с клиентами: оформление заказа, техническое задание, проектирование и изготовление требуемого товара, поддержку клиентов и др. Если стандарты MRP/MRPII/ERP ориентированы только на внутреннюю организацию предприятия, то в стандарт CSRP включен полный жизненный цикл изделия: от его проектирования с учетом требований заказчика до гарантийного и сервисного обслуживания после продажи.

Специализированные информационные системы, примеры которых приведены на рис. 3.1, отнесены в этой классификации

к документально-фактографическим при использовании этого термина в широком смысле, поскольку даже те из этих ИС, которые начинали разрабатываться как фактографические, в последующем, как правило, дополняются нормативно-правовой и другой документальной информацией.

В то же время по мере увеличения числа специализированных ИС, по-видимому, целесообразно выделить их в особый класс специализированных ИС, классифицируемых в соответствии с эволюцией стандартов управления предприятиями, и разработать соответствующую классификацию.

Для целей анализа и создания ИС в сфере экономики удобна используемая на Западе классификация ИС, отличающаяся степенью сложности и участия лиц, принимающих решения⁸⁰:

- *Системы электронной обработки данных (СЭОД)* или просто *системы обработки данных (СОД)*.

Работают в автоматическом режиме с минимальным участием человека; предназначены для решения хорошо структурированных задач, по которым имеются входные данные, известны алгоритмы решения задач.

- *Информационные системы управления (ИСУ)* –

- Применяются при худшей структурированности задач; имеют возможность манипулирования данными за счет появления в их составе СУБД; осуществляют поиск и обработку входной информации; выходную информацию представляют в виде специальных управленческих отчетов; все решения принимает человек; используется на уровне стратегического планирования, управленческого и оперативного контроля.

- *Системы поддержки принятия решений (СППР)* .

Предназначены для решения в режиме диалога плохо структурированных задач, для которых характерна неполнота входных данных, частичная ясность целей и ограничений; обеспечивают значительное участие в работе системы человека, который может вмешиваться в ход решения, модифицировать входные данные, процедуры обработки, цели и ограничения задачи, выбирать стратегии оценки вариантов решений; включает помимо СУБД СППР базу моделей и систему управления этой базой, а также систему управления диалогом; используется на уровне стратегического планирования, оперативного и управленческого контроля.

⁸⁰ *Матвеев Л. А.* Системы поддержки принятия решений: учеб. пособие / Л. А. Матвеев. – СПб.: СПбГУиЭФ, 1993. – 47 с.

- *Экспертные системы (ЭС).*

Основываются на моделировании процесса принятия решения человеком-экспертом с использованием компьютера и разработок в области искусственного интеллекта, на использовании не только данных и информации, но и знаний; но не включают в себя математических моделей, улучшающих принимаемое человеком решение.

- *Гибридные экспертные системы (ГЭС).*

Являются гибридом ЭС и СППР; обеспечивает доступ человека к решению задачи на любой стадии; окончательное решение принимает человек; используются на уровне стратегического планирования и управленческого контроля.

- *Информационные системы мониторинга (ИСМ).*

Предназначены для целей контроля за деятельностью фирмы, обеспечивая высшие звенья управления важной укрупненной информацией; не предназначены для помощи в принятии решений, но полезны для выявления оперативных проблем, а также при анализе разного рода управленческих ситуаций за счет обеспечения текущей и ретроспективной информации.

Вопросы построения и разработки каждого из этих видов ИС могут стать предметом самостоятельного издания. Поэтому в соответствии с поставленной выше целью выявлены и рассмотрены некоторые общие принципы построения и функционирования основных из выделенных классов ИС – фактографических ИС на примере АИС (гл. 4) документальных ИС на примере АСНТИ (гл.5), документально-фактографических ИПС на примере АДФИПС для автоматизированных систем нормативно-методического обеспечения управления (гл. 6).

Вопросы для самоконтроля

1. История развития работ по автоматизации в СССР.
2. Понятие об информационных системах.
3. Определения информационной системы.
4. Направления разработки автоматизированных информационных систем: АИС и АСУ как фактографических ИС; системы научно-технической информации как документальных ИПС.
5. Классификации информационных систем.

Глава 4. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПОСТРОЕНИЯ И ПРОЕКТИРОВАНИЯ АИС И АСУ

4.1. Принципы управления разработками АИС и АСУ

Особенности построения и организации функционирования АСУ

Системы такой сложности, как АСУ, обладают рядом специфических особенностей, присущих открытым системам с активными элементами, к которым, в частности, относятся неоднозначность использования понятий «цели – средства», «система – подсистема»; трудность прогнозирования (а иногда и принципиальная непредсказуемость) поведения системы при внесении в нее изменений. Для обеспечения адаптивности системы, способности ее к самоорганизации необходимо предусмотреть соответствующие средства, обеспечивающие целеобразование, способность вырабатывать варианты поведения, а при необходимости изменять структуру системы управления и АСУ.

Эти особенности были осознаны с самого начала разработки АСУ и обусловили необходимость привлечения для их объяснения и обеспечения системных представлений, закономерностей функционирования и развития сложных систем (см. [1, 13, 16]).

Понимая неизбежность и необходимость проявления названных особенностей и обуславливающих их закономерностей, которые действуют в системе независимо от того, учитывают их или нет, и затрудняют управление разработками АИС (как первой очереди АСУ), некоторые специалисты уже на ранних стадиях их создания предлагали создавать системы проектирования и развития АСУ, разрабатывать единые принципы проектирования и терминологию. Были разработаны типовые положения, типовые структуры, порядок разработки и другие методические материалы, объединенные затем в единый документ – *Общепромышленные руководящие методические материалы (ОПММ)*.

С самого начала в конце 1960-х гг. при проектировании АСУ любого уровня (общегосударственного, регионального, предприятия и организаций, их подразделений и т. п.) возникает необходимость разделить сложную систему на части, чтобы распределить работы между исполнителями, занимающимися разработкой и организацией функционирования АСУ.

Столкнувшись со сложностью выбора первоочередных объектов и направлений автоматизации, распределения ресурсов на разработку их программного обеспечения, в 1970-е гг. для управления разработками АСУ и АИС стали создавать системы проектирования и развития АСУ, разрабатывать единые принципы проектирования и терминологию.

В понимании принципов построения и организации функционирования АСУ большую роль сыграло выделение *функциональной и обеспечивающей* частей (ФЧ и ОЧ), что не утрачивает актуальности и в настоящее время при разработке любых автоматизированных информационных и управляющих систем.

Понятие о структуре функциональной и обеспечивающей части

Эти понятия были введены по аналогии с понятиями «цель» и «средства», выделяемыми при постановке любой задачи принятия решения.

Структура ФЧ определяется на основе анализа целей и функций системы управления, для обеспечения деятельности которой создается АСУ, т. е. ФЧ определяет *цели* и основные *функции* АСУ и представляет собой совокупность *подсистем* и *задач*, выбранных для автоматизации.

Структура ОЧ включает различные виды обеспечения (организационное, информационное, техническое, программное, лингвистическое, эргономическое и т. п.), необходимые для реализации подсистем и задач ФЧ АСУ, т. е. ОЧ объединяет *средства* для достижения целей, реализации задач АСУ.

При этом под структурой ОЧ понимается не просто совокупность средств информационного, технического, алгоритмического, программного и других

видов обеспечения, а организационная форма взаимодействия всех видов обеспечения, необходимых для реализации подсистем и задач (входящих в структуру ФЧ) на всех уровнях иерархии структуры ОЧ.

Необходимость введения новых терминов – ФЧ и ОЧ – вместо привычных понятий «цели» и «средства» объясняется особенностями и закономерностями сложных многоуровневых систем.

В таких системах каждый уровень иерархической структуры ведет себя как «двуликий Янус» – как *средство* по отношению к вышестоящему уровню и как *цель* по отношению к нижележащему, т. е. имеет место *закономерность коммуникативности* (см., например, [1, 13, 16]). Соответственно составляющие любого промежуточного уровня в структуре АСУ можно рассматривать как подсистемы по отношению к вышестоящему уровню, к системе в целом, их объединяющей, и в то же время, взятые сами по себе, они могут рассматриваться как системы. Поэтому иногда при разработке АСУ возникали противоречивые ситуации; отдельные подсистемы «Управление кадрами», «Управление качеством» и т. п. на определенной стадии развития АСУ начинали называть «АСУ-Кадры», «АСУ-Качество», т. е. считали их как бы самостоятельными системами, в то время, как по отношению к общей системе управления предприятием они продолжали оставаться подсистемами.

Более того, на каждом уровне иерархической структуры ФЧ в силу *закономерности иерархической упорядоченности* всегда следует определять свои цели, функции, задачи, средства. Иными словами, понятия «цель» и «средства» в иерархической структуре всегда используются неоднозначно, и это должны осознать разработчики АСУ, не тратя времени на бессмысленные дискуссии по поводу терминов.

Следует также оговорить, что введенные термины ФЧ и ОЧ нельзя однозначно отождествлять с понятиями «цели» и «средства». Они являются более сложными понятиями.

Действительно, иногда говорят об *обеспечивающих подсистемах*, имея в виду определенным образом организованные совокупности информационных, программных, технических средств, используемых для автоматизации не укрупненной функции (для которой используется термин «функциональная подсистема»), а какой-либо вспомогательной функции нижележащих уровней или определенной их совокупности. Например, такого рода обеспечивающими подсистемами были банки данных опреде-

ленного назначения, типа ИНЭС ⁸¹, СИОД ⁸², БАНК и т. п., используемые как средства для хранения и предоставления лицам, принимающим решения, информации о состоянии производства. В качестве такого рода подсистем в настоящее время можно рассматривать программные продукты для автоматизации отдельных видов деятельности (бухгалтерского учета, материально-технического снабжения и т. п.).

С другой стороны, когда сдают в эксплуатацию *функциональную подсистему*, то имеют в виду не только совокупность задач и функций, включаемых в нее, но и технические средства, алгоритмы, программы, инструкции по их использованию, т. е. и совокупность средств реализации этой подсистемы.

Таким образом, термины ФЧ и ОЧ являются обобщающими условными понятиями, которые помогают охарактеризовать автоматизированную систему как целое, выделив в ней работы, в большей мере связанные с анализом и описанием целей системы (формирование структуры ФЧ АСУ), и работы, связанные с определением общей структуры средств реализации целей (разработка структуры ОЧ АСУ), что помогает организовать управление разработками АСУ и АИС.

4.2. Принципы организации управления разработками АСУ

Как было показано выше, в понимании принципов построения и организации функционирования АСУ большую роль играет выделение *функциональной* и *обеспечивающей* частей (охарактеризованных в предыдущем разделе), что не утрачивает актуальности и в настоящее время при разработке любых автоматизированных информационных и управляющих систем.

Соответственно при управлении разработками автоматизированных систем выделяют две основные проблемы:

⁸¹ *Авен О. И.* Что же такое АСУ / О. И. Авен. – М.: Наука, 1981. – 175 с.

⁸² *Келехсаев А. А.* Системы интеграции и обработки данных СИОД1, СИОД2 / А. А. Келехсаев, А. П. Беляев. – М.: 1977. – 208 с.

1. Формирование структуры ФЧ АСУ и выбор на ее основе первоочередных задач автоматизации.

2. Формирование структуры ОЧ АСУ как организационной формы взаимодействия всех видов обеспечения (информационного, технического, организационного и т. д.), необходимых для реализации подсистем и задач, входящих в структуру ФЧ АСУ.

Решение этих двух основных проблем взаимосвязано. С одной стороны, структура ОЧ определяется структурой ФЧ. В то же время выбор структуры ФЧ во многом зависит от имеющихся технических, программных и иных средств, т. е. определяется потенциальными возможностями ОЧ.

Основные проблемы управления разработками АСУ, в свою очередь, делят на задачи, которые часто можно решать параллельно.

Например, первую проблему делили следующим образом:

1.1. Прогнозирование структуры ФЧ АСУ. Разработка прогнозного варианта структуры ФЧ АСУ (на 15–20 лет) и основных направлений развития АСУ (на 5–10 лет).

1.2. Разработка структуры ФЧ АСУ последующей очереди (на 3–5 лет). Эту задачу называли также «Выбор первоочередных подсистем (комплексов задач) автоматизации для последующей очереди АСУ».

1.3. Выбор первоочередных (наиболее значимых) задач в подсистемах АСУ и последовательности их проектирования и внедрения.

1.4. Проектирование подсистем АСУ.

Вторую проблему можно представить следующими подпроблемами:

2.1. Выбор (обоснование) структуры ОЧ АСУ.

2.2. Проектирование отдельных видов обеспечения.

Кроме того, в действующей АСУ важно создавать *организационную структуру АСУ*, которая определяется составом и взаимосвязью отдельных структурных подразделений в условиях эксплуатации АСУ. Поэтому к названным основным проблемам в последующем была добавлена проблема управления разработками АСУ.

Систему управления разработками АСУ можно считать 3-й проблемой и разделить ее на следующие задачи:

3.1. Разработка структуры организационного обеспечения управлением разработками АСУ (на первых этапах эта задача сводится к определению структуры подразделения, разрабатывающего АСУ, а по мере развития – к определению взаимоотношений между подразделениями-разработчиками и подразделениями, использующими результаты разработки в практической деятельности, а также подразделениями, подготавливающими и контролирующими вводимую информацию). Иногда эту составляющую выделяют в самостоятельную проблему.

3.2. Создание информационной системы для обеспечения проектирования подсистем и задач АСУ.

Структура проблем разработки АСУ приведена на рис. 4.1.

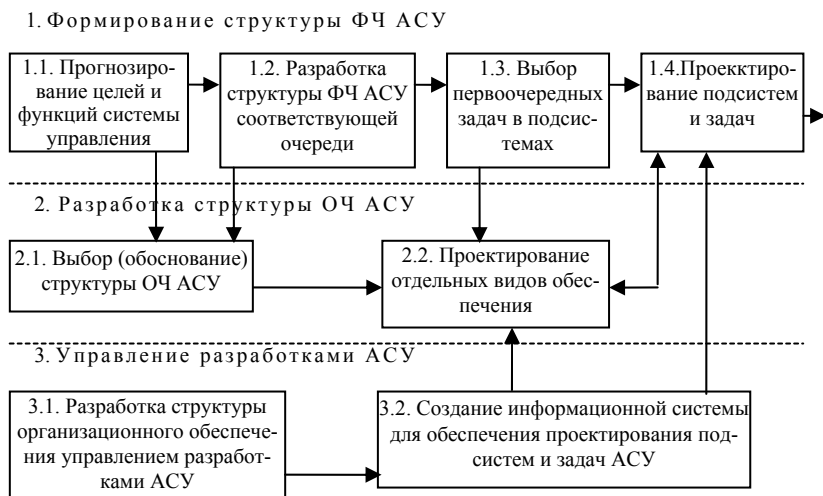


Рис. 4.1. Этапы разработки АСУ

Для управления разработками автоматизированных систем были подготовлены соответствующие руководящие методические материалы, в которых АСУ трактовалась как развивающаяся система и вводилось понятие *очереди*. АСУ первой очереди разрабатывалась как *автоматизированная информационная система – АИС*. В руководящих материалах оговаривался также порядок разработки соответствующей очереди АСУ (АИС) и ввода ее в эксплуатацию (рис. 4.2).

При реализации приведенной на рис. 4.1 структуры ориентировались на методику, приведенную на рис. 4.2. Получаемая структура ФЧ АСУ и выбор состава ОЧ проводились на основе обследования существующей системы управления предприятием и технико-экономического обоснования. (ТЭО) и являлись основой разработки технического задания (ТЗ). После согласования и утверждения ТЭО и ТЗ выполнялся технический проект (ТП) и рабочий проект (РП) и ввод разработанных подсистем и задач эксплуатацию.

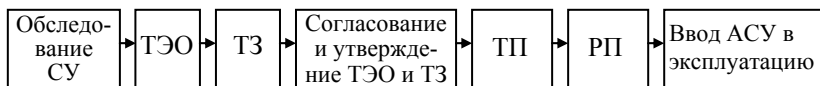


Рис. 4.2. Порядок разработки АСУП

Практика показала, что разработку АСУ невозможно реализовать на основе одного проекта. Было введено понятие очереди АСУ и в процессе реализации предшествующей очереди начиналась разработка структуры ФЧ и ТЗ следующей очереди (рис. 4.3), т.е. применялась одна из важных закономерностей теории систем – *закономерность историчности* (см., например, [1, 13, 16]).

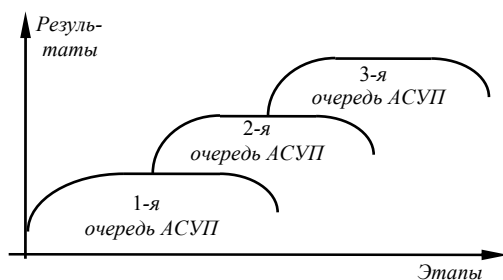


Рис. 4.3. Проявление закономерности историчности при разработке АСУ

При разработке каждой последующей очереди АСУ последовательность этапов, приведенная на рис. 4.2, повторялась. Более того, практика показала, что в период одной очереди при разработке подсистем этапы, приведенные на рис. 4.2, также повторялись применительно к подсистеме.

Были разработаны типовые положения, типовые структуры, порядок разработки и другие методические материалы, объединенные затем в единый документ – Общеотраслевые руководящие методические материалы (ОРММ), регламентирующий ход разработки АСУ и ОАСУ.

Четкий порядок, установленный ОРММ, ускорил распространение опыта организации работ по проектированию АСУ, облегчил учет и сравнительный анализ хода работ по созданию АСУ на предприятиях и в отраслях. Однако одновременно ограничил развитие АСУ ряда предприятий и НПО, что неизбежно в силу *закона «необходимого разнообразия»* У. Р. Эшби (см. в [1, 13, 16]): именно за счет ограничения «разнообразия», т. е. типизации, и в конечном счете примитивизации систем, было достигнуто упрощение централизованного управления разработками АСУ.

Негативная роль ОРММ была в последующем осознана, особенно в период перехода предприятий на самокупаемость и хозрасчет. Типовые решения и структуры можно использовать лишь на начальных стадиях создания АСУ, а по мере их развития все больше начинают проявляться индивидуальные особенности конкретных предприятий и объединений и связанная с этим индивидуальность АСУ.

В этих условиях для управления разработками АСУ требуются методические материалы, в которых не только определялось бы, *что* нужно делать в процессе разработки АСУ и диктовались бы готовые типовые проектные решения, а давались бы рекомендации о том, *как* нужно принимать решения по выбору структуры АСУ, средств ее реализации в конкретных условиях, необходимо разрабатывать соответствующие методики и модели для принятия решений по управлению разработками АИС и АСУ с учетом конкретных особенностей предприятий и организаций.

Очевидно, что принципы построения и эффективность АСУ существенно зависят от уровня развития информационных технологий, разработки новых технических средств и программных продуктов.

С появлением в середине 1980-х гг. персональных ЭВМ происходит корректировка идеи АСУ от ВЦ и централизации управления – к распределенному вычислительному ресурсу и децентрализации управления. При этом уменьшается нагрузка на централизованные вычислительные ресурсы и верхние уровни управления, что позволяет сосредоточить в них решение крупных долгосрочных стратегических задач.

В то же время для обеспечения эффективного управления крупными предприятиями остается актуальной идея создания *интегрированных АСУ*, а для обеспечения информацией по группам основных функций организационного управления предприятий, корпораций – *корпоративных информационных систем* – КИС.

4.3. Виды и принципы разработки структуры функциональной части АСУ

Виды структур функциональной части АСУ

Первоначально основной формой представления структуры ФЧ АСУ была древовидная иерархическая структура. АСУ делилась на подсистемы (или комплексы задач), подсистемы – на группы задач, а последние, в свою очередь – на отдельные задачи.

В *Общепромышленных руководящих методических материалах по созданию АСУП*⁸³, утвержденных Постановлением Государственного комитета по науке и технике Совета Министров СССР 10 мая 1971 г. № 145, типовая структура ФЧ первоначально включала 7 подсистем (рис. 4.4).

Затем типовая структура ФЧ АСУП была дополнена подсистемой управления качеством, а в последующем постепенно расширялась. Подсистемы в первых структурах ФЧ АСКП, а затем – в типовой структуре ФЧ АСУП были ориентированы на основные

⁸³ *Общепромышленные* руководящие методические материалы по созданию автоматизированных систем управления предприятиями (АСУП). – Минск, 1971. – 376 с.

функциональные подразделения существовавших систем организационного управления, откуда и произошел термин *функциональная подсистема*.

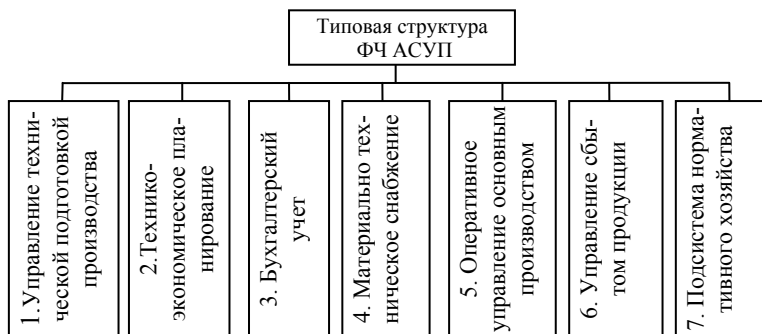


Рис. 4.4. Типовая структура ФЧ АСУП

Число подсистем в АИС первой очереди конкретных предприятий было разным, но обычно не превышало числа подсистем типовой АСУП (7–8), рекомендованной в *Общепромышленных руководящих методических материалах*. Число задач по подсистемам колебалось довольно в широких пределах, в зависимости от принятой детализации задач и конкретных особенностей предприятия.

В 80–90-е гг. XX в. в связи с необходимостью объединения в единую систему разнообразных автоматизированных систем, разрабатывавшихся на предприятии возник и был широко распространен термин «Интегрированные автоматизированные системы управления промышленными предприятиями».

В соответствии с объективными процессами материального производства и управления создавались информационные системы управления следующих основных типов (рис. 4.5):

- автоматизированные системы научных исследований (АСНИ);
- системы автоматизированного проектирования изделий (САПР) и автоматизированные системы технологической подготовки производства (АС ТПП);
- автоматизированные системы общезаводского управления, ориентированные на автоматизацию функций управления объединением (корпорацией, трестом, концерном и др.) (АСУО) и предприятием (АСУП);

- автоматизированные системы для комбинированного организационного и технологического управления (АСУ ОТ);
- автоматизированные системы управления гибкими производствами (АСУ ГПС), включая АСУ отдельными производствами, цехами, участками, гибкими автоматизированными линиями (ГАЛ), гибкими производственными модулями (ГПМ);
- автоматизированные системы управления технологическими процессами (АСУ ТП);
- автоматизированных систем контроля качества и испытания изделий (АСКИ).

Объединить эти подсистемы помогла структура ФЧ, приведена на рис. 4.5 [21].



Рис. 4.5. Структура ФЧ ИАСУ

Однако по мере развития АСУ число подсистем увеличивалось. При возрастании числа подсистем нарушается важное требование к иерархическим структурам – гипотеза Миллера, согласно которой число составляющих одному узлу должно быть 7 ± 2 .

Трудности управления разработками АСУ при существенном увеличении числа подсистем, т. е. числа составляющих на

верхнем уровне иерархической структуры ФЧ, привели к тому, что на некоторых предприятиях стали вводить еще один обобщающий уровень, который был предложен на Волжском автозаводе (история развития структуры ФЧ АСУ ВАЗа приведена на рис. 4.6) и назван *направлениями*.

Для иллюстрации развития структуры ФЧ АСУ выбран пример ВАЗа, на котором создана наиболее развитая АСУ, являющаяся основой функционирования производства.

Направление объединяло несколько подсистем, связанных между собой. В принципе можно было бы укрупнить подсистемы, однако к тому времени в практике разработок АСУ с понятием подсистемы был связан вполне определенный объем работ, поэтому более удобным оказалось ввести новый термин. Пример направлений, выделенных на ВАЗе в 1984 г., приведен на рис. 4.6, в).

В дальнейшем по мере развития АСУ появилась (впервые на Волжском автозаводе) новая форма автоматизированных систем – *организационно-технологические АСУ (АСУ ОТ)*, которые создавались для повышения степени оперативного управления отдельными участками производства. АСУ ОТ сочетают функции управления технологическими процессами (АСУ ТП) и функции организационного управления производством (АСУП), причем большая часть массивов организационного управления формируется на основе массивов управления техпроцессами (что возможно только при практически полной автоматизации технологических процессов). Подсистемы АСУ ОТ возникали либо на основе АСУ ТП, либо – как развитие подсистем АСУП.

Вначале АСУ ОТ рассматривали как одно из направлений. Но иерархический принцип представления имеет в числе своих специфических особенностей стремление к обособленности выделенных ветвей, т. е. в иерархических структурах сильны вертикальные связи и предельно ослаблены горизонтальные взаимосвязи между направлениями, подсистемами, которые после их выделения в самостоятельные начинают развиваться независимо, как бы соперничая друг с другом.

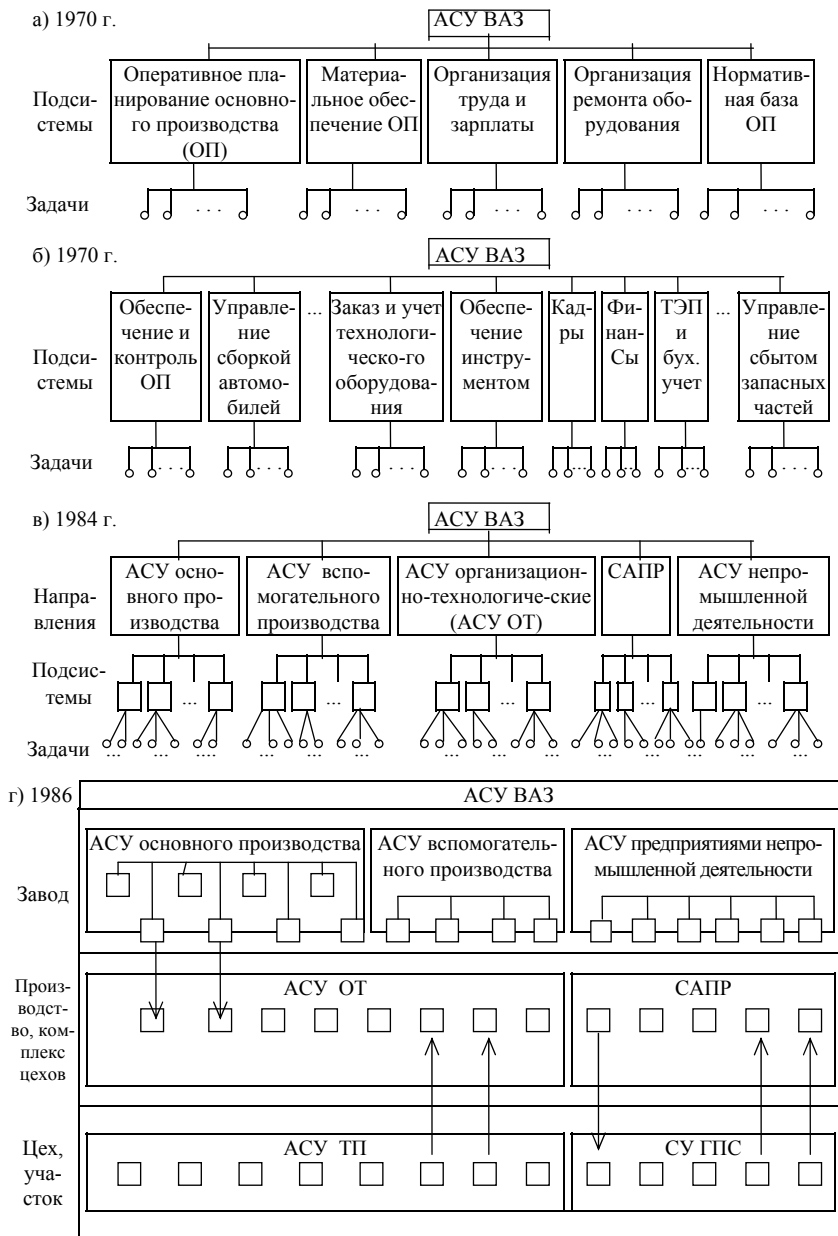


Рис. 4.6. Развитие структуры ФЧ АСУ ВАЗа

Поэтому, если включить АСУ ОТ в структуру ФЧ АСУ как равноценное направление с АСУ ТП и АСУП, то по мере самостоятельного развития этих направлений возможно дублирование работ в АСУ ОТ и АСУ ТП или в АСУ ОТ и АСУП. Для того, чтобы усилить взаимодействие этих трех направлений, их оказалось удобнее поместить в структуре друг под другом.

В этом случае устанавливаются вертикальные согласующие взаимосвязи между подсистемами и задачами этих направлений, и по мере развития подсистем АСУ ТП в направлении усиления функций организационного управления переводить их на уровень АСУ ОТ, а по мере усиления функций управления технологическими процессами в подсистемах АСУП – на уровень организационно-технологических АСУ.

Аналогичные взаимоотношения целесообразно установить между подсистемами АСУП, САПР и СУ ГПС.

Сформированная в результате *многоуровневая структура* ФЧ АСУ ВАЗа приведена на рис. 4.6, *г*.

Аналогично была сформирована и структура АСУ объединения АвтоВАЗ, в которой к приведенным на рис. 4.6, *г* уровням добавился еще уровень объединения, а на нижележащей по отношению к нему страте, наряду с АСУ ВАЗа, были помещены АСУ других предприятий, входящих в объединение, структуры ФЧ АСУ которых формировались аналогично АСУ ВАЗа.

Приведенные примеры структур ФЧ АСУ иллюстрировали развитие структуры по мере развития предприятий и их АСУ: от первоначального представления структур ФЧ в виде древовидных иерархических структур до структур ФЧ АСУ в виде *многоуровневой, стратифицированной*.

Для разработки структур ФЧ АСУ конкретных предприятий применялись подходы и методы системного анализа – методики структуризации целей и функций систем управления, методы их оценки, и выбора на этой основе подсистем и задач, подлежащих автоматизации при разработке соответствующей очереди АСУ.

Принципы и примеры разработки структур ФЧ АСУ

Применение методики, базирующейся на двойственном определении системы. *А. И. Уёмов*, определяя систему через понятия «вещи», «свойства», «отношения», предложил двойственное определение⁸⁴, в одном из которых свойства q_i характеризуют элементы a_i , а в другом – свойства q_j характеризуют связи (отношения) r_j (1.1, е):

$$\begin{aligned} S &\equiv [\{a_i\} \& \{r_j(q_j)\}], \\ \text{def} \quad a_i &\in A \quad r_j \in R \quad q_j \in QR \end{aligned} \tag{4.1}$$
$$\begin{aligned} S &\equiv [\{a_i(q_i)\} \& \{r_j\}]. \\ \text{def} \quad a_i &\in A \quad q_i \in QR \quad r_j \in R \end{aligned}$$

В работах *Б.Д. Кошарского* (см. ссылки в [1, 13]) было показано, что этим определениям соответствует два способа представления системы управления:

п р о ц е д у р н о е – как множества объектов A , на котором реализуются заранее определенные отношения R с фиксированными свойствами Q_R ; при этом, если системообразующее отношение определено во времени, то это представление соответствует структуризации системы по циклу управления: планирование, организация, регулирование, учет и т. п. (набор функций цикла управления изменяется по мере развития экономики и зависит от конкретных условий);

ф а к т о р н о е – как множества объектов A , обладающее заранее определенными свойствами Q_A с фиксированными между ними отношениями R ; при этом могут быть выделены такие составляющие, как основное и вспомогательное производство, основные и оборотные фонды, трудовые ресурсы, материально-техническое обеспечение и другие объекты управления на предприятии (набор их также определяется конкретными условиями).

Б. Д. Кошарский показал, что каждый из этих способов представления системы в отдельности дает неполное описание системы управления, а для выявления системных особенностей конкретного предприятия необходимо один способ описания дополнить другим, двойственным ему, т. е. что только совместное использование процедурного и факторного представлений системы

⁸⁴ *Уёмов А. И.* Системный подход и общая теория систем / А. И. Уёмов. – М.: Мысль, 1978. – 272 с.

позволяет обеспечить конкретизацию и полноту анализа целей и задач организационного управления.

Отметим, что это утверждение о полноте справедливо лишь в рамках принятой концепции системы.

Такое требование на практике реализуется либо путем формирования и анализа матрицы «цикл управления–объект управления» (табл. 4.1), после оценки которой формируются двойственные структуры и осуществляется выбор из них наилучшей; либо путем параллельного формирования двойственных вариантов структуры (рис. 4.7), с использованием взаимно обратной последовательности признаков структуризации.

Таблица 4.1

Матрица «цикл управления–объект управления»

Объект управления (ОбУ) Производственный признак	Цикл управления (ЦУ) Временной признак					
	Прогнозирование	Перспективное планирование	Организация	Текущее планирование	Оперативное управление	Учет Контроль Анализ
	<i>ПП</i>	<i>ПП</i>	<i>ОРГ</i>	<i>ТП</i>	<i>ОУ</i>	<i>УКА</i>
Научно-исследовательская работа (НИР)	–	–	–	–	–	–
Производство основной продукции (ОП)	+	+	+	+	+	+
Вспомогательно-обслуживающее производство (ВП)	–	–	–	+	–	+
Транспорт (Т)	–	–	–	–	–	–
Материально-техническое снабжение (МТС)	–	–	–	+	+	+
Трудовые ресурсы (кадры – К)	–	–	–	–	–	+
Сбыт продукции (СбП)	–	–	–	+	–	+
Финансы (Ф)	–	–	–	–	–	–

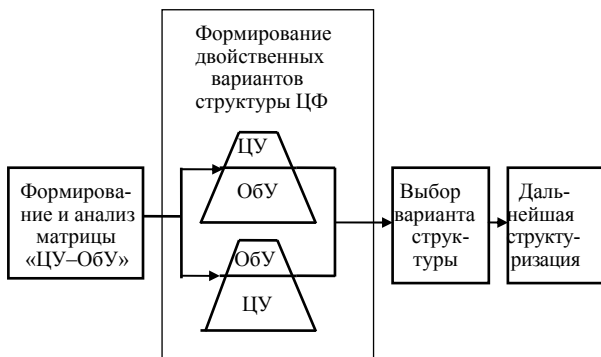


Рис. 4.7. Структура методики, основанной на двойственном определении системы А. И. Уёмова

На рис. 4.8 приведен пример вариантов таких обратных друг другу структур, соответствующих анализу составляющих в табл. 4.1.

В этом примере использовались классификаторы, рекомендованные обычно в справочниках разработчика АСУ, чтобы получить реальную ситуацию, имевшую место при разработке типовой структуры функциональной части (ФЧ) АСУП, которая в течение ряда лет была основой разработки структур ФЧ многих АСУ предприятий.

Полученные структуры анализируют.

Первая структура (рис. 4.8, а) имеет «вырожденные» ветви и не соответствует требованиям равномерности структуризации и др., предъявляемым к структурам целей, рассмотренным, например, в [1, 13, 16]. Вторая структура (рис. 4.8, б) с этой точки зрения лучше, но в ней одна ветвь (первая, связанная с управлением основным производством) «перегружена» по сравнению с остальными. Если разделить ее, выделив в самостоятельные подцели (подсистемы) ТПП, ТЭП и ОУОП (аббревиатуры см. в табл. 4.1), то структура будет удовлетворять требованию равномерности. Эта структура и была положена в основу типовой структуры функциональной части АСУП. В дальнейшем была добавлена подсистема управления качеством продукции (УКачП).

Методика Кошарского–Уёмова нашла широкое применение в различных отраслях при структуризации целей и функций предприятий в процессе разработки структуры функциональной части их автоматизированных систем управления.

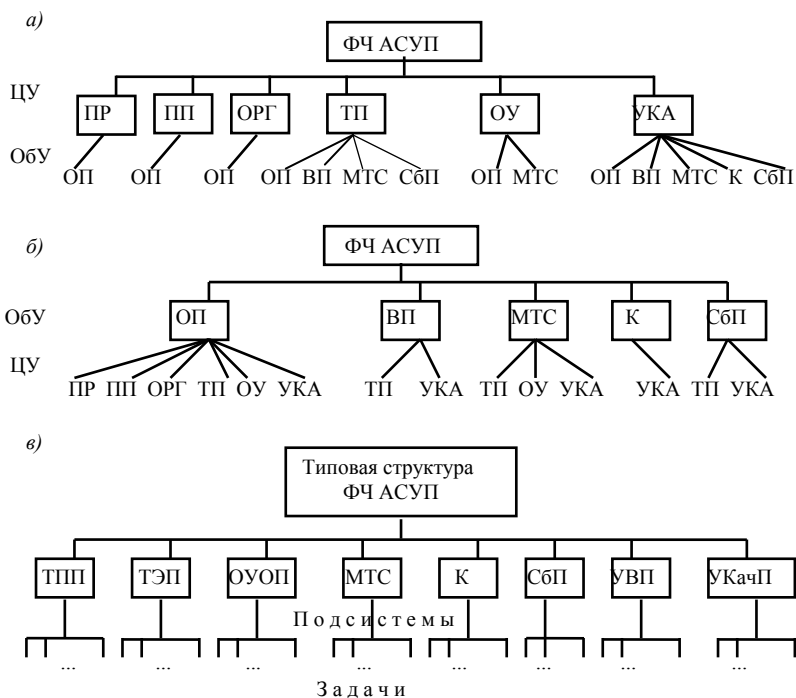


Рис. 4.8. Варианты структур ФЧ АСУП

При этом в ряде случаев подход, положенный в ее основу, оказался столь естественным, что имена первых авторов были забыты, принцип двойственного представления не упоминался, а использовались сразу признаки «цикл управления» и «объект управления».

Методика является удобной для анализа целей и функций действующих предприятий, для которых можно провести обследование существующей системы управления и выявить объекты управления. Однако в ней нет средств для определения новых объектов, новых видов деятельности и функций, связанных с развитием предприятия (внедрением новой техники, технологии и т. п.), что ограничивает применение методики при реконструкции и проектировании новых предприятий.

Применение методики, основанной на концепции системы, учитывающей среду и целеполагание. В основе методики лежит определение системы **В. Н. Сагатовского** ⁸⁵, в котором учитываются понятия цели Z , среды SR и интервала времени ΔT , периода существования системы, влияющего на процесс целеобразования:

$$S \stackrel{def}{=} \langle A, R, Z, SR, \Delta T \rangle. \quad (4.2)$$

На основе этого определения дано обоснование положенной в основу методики философской концепции системы и используемых для ее раскрытия признаков структуризации системных моделей. Разработана и исследована методика группой ученых томских вузов **Ф. И. Перегудовым, В. Н. Сагатовским, В. З. Ямпольским, Л. В. Кочневым** под руководством **Ф. И. Перегудова**. Поэтому кратко ее иногда называют методикой Ф. И. Перегудова – В. Н. Сагатовского.

Структура методики приведена на рис. 4.9.

Основные этапы методики соответствуют уровням структуризации, Методика излагается в терминах и формулировках ее авторов, но с некоторыми дополнениями, предложенными ими позднее.

Уровень 1. Формирование глобальной цели системы.

Цель либо задается вышестоящей организацией, либо воссоздается на основе анализа директивных документов. Цель должна быть ориентирована на конечный продукт, для получения которого существует или создается система.

Конечным продуктом может быть любой результат социальной деятельности: материальная продукция, новый научный результат, научная информация и т. д.

Уровень 2. Декомпозиция по признаку «виды конечного продукта» (ВКП).

Осуществляется в тех случаях, когда система производит разные виды конечного продукта. При наличии большого числа разновидностей продукции классификатор по этому признаку может быть двухуровневым.

⁸⁵ **Основы** системного подхода и их приложение к разработке территориальных АСУ / Под ред. Ф.И. Перегудова. – Томск: Изд-во ТПИ, 1976. – 440 с.

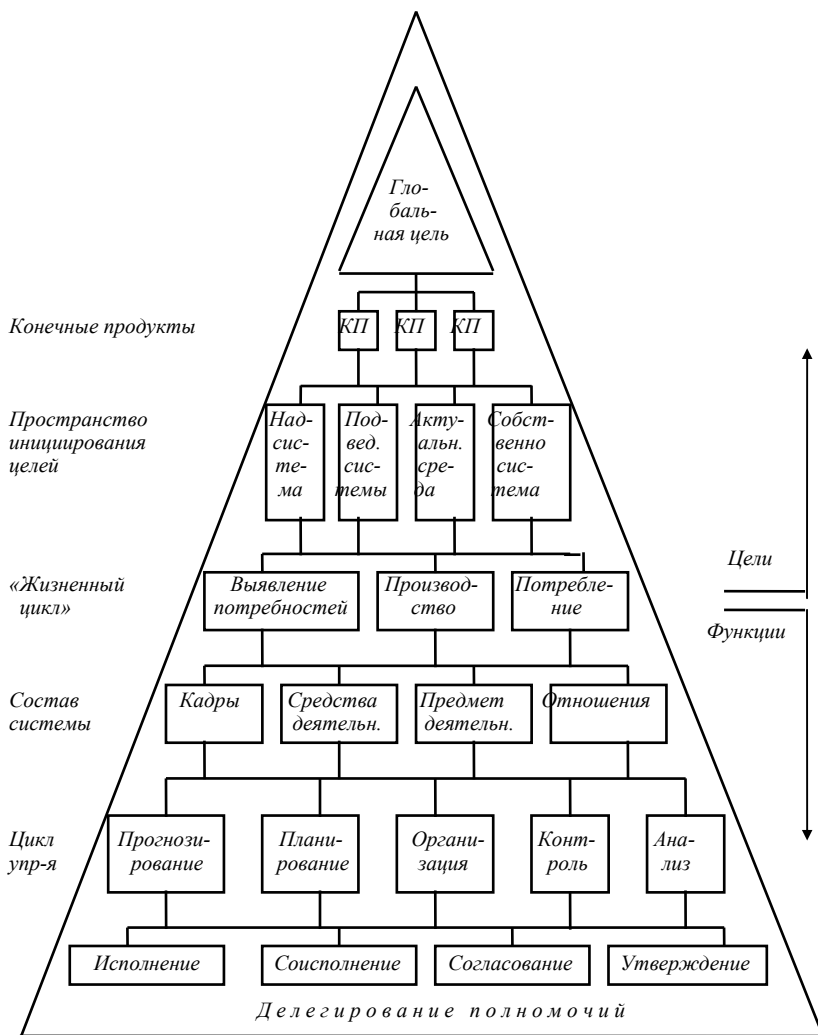


Рис. 4.9. Структура методики, учитывающей взаимодействие системы со средой

Виды конечного продукта зависят от того, для чего строится структура целей. Если речь идет о производстве, то конечным продуктом является вы-

пускаемая продукция, а если структура целей строится для аппарата управления, то это планы, решения и другие нормативно-методические документы, обеспечивающие выпуск соответствующего вид продукции.

Уровень 3. Декомпозиция по признаку «пространство инициирования целей» (ПИЦ).

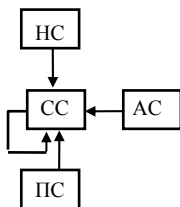


Рис. 4.10. ПИЦ

Формируются подцели исследуемой системы, иницируемые требованиями и потребностями окружающей среды, которые имеют отношение к производству конечного продукта проектируемой или исследуемой системы.

При этом все системы, с которыми взаимодействует исследуемая в процессе производства конечного продукта, делятся на четыре класса (рис. 4.10): *над-система* (НС) или *вышестоящие системы* (ВС),

формулирующие главные требования к конечному продукту (и потребности в нем); *нижележащие* или *подведомственные системы* (ПС), требования которых выступают в основном в качестве ограничений на свойства конечного продукта или потребностей в организации ремонта и других видов обслуживания материально-технической базы для производства конечного продукта, *существенная* или *актуальная среда* (АС), т. е системы ⁸⁶, исследуемая *собственно система* (СС), подцели которой иницируются собственными (внутренними) потребностями, мотивами, программами, постоянно возникающими в развивающейся системе и также трансформирующимися в требования к конечному продукту.

Отметим, что этот признак структуризации базируется на *закономерности коммуникативности* (см., например, [1, 13, 16]).

Уровень 4. Декомпозиция по признаку «жизненный цикл».

Определяются различные подэтапы получения конечных продуктов в зависимости от их видов – от формирования или прогнозирования потребностей в продукте до потребления или поставки заказчику (см. примеры на рис. 4.9).

⁸⁶ В дальнейшем **В. Г. Колосовым** была предложена идея выделения *дружественной* (поставщики, потребители, аналогичные предприятия, с которыми сотрудничает рассматриваемая система), *конкурентной* (предприятия, выпускающие аналогичную продукцию или поставляющие ее на рынок из других стран или регионов) и *безразличной* среды, которая на данном этапе является нейтральной, но которая со временем может стать дружественной или конкурентной.

Начиная с этого уровня декомпозиции обычно становится удобнее оперировать не термином «подцель», а термином «функция», и считать, что «дерево целей» как бы перерастает в «дерево функций».

Уровень 5. Декомпозиция по основным элементам (составу) системы (СС), в результате которой формируются функции, вытекающие из потребностей основных элементов системы, которые объединяются в три основные группы – *кадры К, предмет деятельности ПД и средства деятельности СД*, на практике иногда вызывает затруднения, и его бывает удобно интерпретировать как «объекты деятельности» (основной – ПД, и обеспечивающей – СД, К и др. ресурсы системы).

Уровень 6. Декомпозиция по признаку «Управленческий цикл», классификатор которого, предлагаемый авторами методики, приведен на рис. 4.9.

Уровень 7. Декомпозиция по признаку «Делегирование полномочий», классификатор по которому также приведен на рис. 4.9.

Примечание. Методика развивалась. В первоначальном ее варианте было пять уровней декомпозиции. Позднее добавилось еще два. В зависимости от приложений менялась последовательность признаков: во многих приложениях удобнее на первое место вынести признак «пространство инициирования целей», помогающий уточнить «виды конечного продукта». По-разному можно раскрывать «жизненный цикл» системы.

Методика нашла наиболее широкое применение из всех рассматриваемых: она использовалась при разработке «дерева целей» управления хозяйством области; при формировании структуры функциональной части территориальной АСУ Томской области; структуры ФЧ отраслевой АСУ (ОАСУ) Минвуза РСФСР; при корректировке организационной структуры Минвуза и т. д. (см. ссылки в [1, 13]).

Наибольшее распространение получили первые три признака структуризации (т. е. признаки формирования собственно структуры целей, и особенно признак «пространство инициирования целей», в основе которого лежит закономерность коммуникативности (разделяющая сложную среду на

надсистему, подведомственные системы, актуальную среду и внутреннюю среду – собственно систему, постоянно изменяющуюся в развивающейся системе).

Получаемые с помощью этой методики структуры цели и функций существенно полнее, чем при использовании предыдущей.

Пример структуры приведен на рис. 4.11.



Рис. 4.11. Пример структуры целей и функций, полученных на основе методики, учитывающей взаимодействие системы со средой

Анализ *пространства инициирования целей* заставляет обратить внимание на необходимость определения платежеспособного спроса на продукцию (услуги) предприятия (НС), поддержания в работоспособном состоянии оборудования, организации ремонта,

изготовления специнструмента, спецоснастки и на другие функции, обусловленные потребностями ПС, а также на функции координации и кооперирования с предприятиями и организациями АС и функции СС.

Таким образом, благодаря большей полноте определения системы, данного **В. Н. Сагатовским**, по сравнению с двойственным определением системы **А. И. Уёмов** и раскрытию этого определения в хорошо отработанной совокупности признаков структуризации и примеров классификаторов по этим признакам, рассматриваемая методика обеспечивает бóльшую полноту анализа целей и функций систем управления по сравнению с предыдущей методикой, помогает выявлять новые, ранее не выполнявшиеся на предприятии функции. Методика является хорошим средством анализа целей и функций в условиях изменяющейся среды, развития предприятия (организации), при внедрении в производство и управление различного рода нововведений, техническом перевооружении и реконструкции или проектировании новых предприятий.

В установившемся же режиме функционирования предприятия эта методика может оказаться избыточной. Кроме того, возникают затруднения при определении «видов конечного продукта» четвертой ветви (собственно системы управления), для структуризации которой оказалась удобной методика, основанная на концепции деятельности (см. например, [1, 13, 16]).

Сравнивая методики с точки зрения положенных в их основу концепций, можно дать некоторые рекомендации по их выбору в конкретных условиях.

Так, концепция двойственного определения системы **А. И. Уёмов** ориентирована на описание статики системы, фиксацию уже достигнутых представлений о ней у ЛПП. Она, разумеется, допускает включение новых объектов управления, изменение функций в цикле управления (что и происходит по мере развития методики), однако в ней нет средств, которые помогали бы выявить новые объекты, функции, виды деятельности, такие, как внедрение новой техники, технологии, нововведений в управленческой деятельности. Методику, основанную на концепции системы, учитывающей ее взаимодействие со средой, полезно применять на этапах развития системы, пересмотра производственной и организационной структур,

при проектировании новых предприятий. Она, помогает выявить новые виды деятельности, объекты управления.

Существуют и другие методики структуризации целей и систем, рассмотренные в [1, 13, 16], необходимость в использовании которых определяется конкретной проблемной ситуацией.

Необходимость в использовании методики, базирующейся на концепции деятельности ⁸⁷, возникает в тех случаях, когда исследуемый или создаваемый объект недостаточно изучен, т. е. в случаях постановки новых проблем, структуризации целей развития новых видов деятельности.

Методика *Р. Акоффа* и *Ф. Эмери* ⁸⁸ помогает обеспечить полноту выявления подцелей и функций для системы управления районом, городом, для управления непромышленной деятельностью предприятия (культурно-бытовыми, детскими учреждениями и другими учреждениями социальной сферы).

Таким образом, при выборе и разработке методики структуризации целей и функций системы управления нужно учитывать состояние системы (находится ли она в стабильном состоянии, либо требуется существенный пересмотр ее целей и функций в связи с реконструкцией производства, перестройкой системы управления, изменением принципов организации экономики и т.п., либо система еще мало исследована), характер анализируемого вида деятельности, степень познания объекта (т. е. имеющиеся представления о нем у лиц, формирующих структуру ЦФ), отведенный период времени на проектирование или преобразование системы управления, влияющий на возможность выполнения методики в полном объеме, включая оба этапа, и т. п.

Поэтому целесообразно разрабатывать обобщенную методику формирования и анализа структур целей и функций, которая включала бы несколько методик структуризации, несколько методов оценки структур, предусматривала бы возможность выбора методики получения первоначального варианта (вариантов)

⁸⁷ *Применение* системного анализа на разных уровнях правления в высшей школе: Обзор. информация / Под ред. В.Н. Волковой. – М.: НИИВШ, 1977. – 65 с.

⁸⁸ *Акофф Р. О целеустремленных системах* / Р. Акофф, Ф. Эмери. – М.: Сов. радио, 1974. – 272 с.

структуры и методов оценки компонентов структур ЦФ, наиболее подходящих для соответствующего периода развития предприятия и его системы организационного управления, с учетом характера, объемов и других конкретных особенностей создаваемого предприятия (организации).

Принципы автоматизации формирования структуры целей и функций

Изучение процесса формирования и анализа структур целей и функций управления показало, что это – сложный, итеративный процесс, требующий уточнения признаков структуризации, классификаторов по этим признакам, изменения их последовательности, обсуждения вариантов структуры и внесения изменений в исходные классификаторы. Если структуру параллельно формируют разные специалисты, то даже при использовании одной и той же методики они, как правило, формируют разные варианты структуры, что обусловлено проявлением закономерности эмерджентности (см. [1, 13, 16]) на каждом уровне иерархии. При сопоставлении вариантов структуры, согласовании мнений экспертов нужно обеспечить быструю повторяемость формирования новых, уточненных структур, что весьма трудоемко.

Все это обусловило необходимость поиска путей автоматизации формирования и анализа структур целей и функций, которая позволила бы сократить время на получение структуры, не снижая при этом степени полноты.

Анализ пошагового формирования структуры ЦФ с использованием методик структуризации позволяет частично формализовать этот процесс и подготовить возможность автоматизации. Анализ показал также, что по мере добавления признаков структуризации при продвижении с верхнего уровня на нижний получаемые подцели и функции становятся все более конкретизированными. Это означает, что при помещении рядом составляющих классификаторов смежных – вышестоящего и нижележащего – уровней иерархической структуры и при продвижении по структуре сверху

вниз происходит уточнение содержания функций управления, т. е. как бы появляется новый смысл (вследствие осмысления формируемых сочетаний человеком, их воспринимающим и оценивающим).

При сопоставлении таких конкретизированных подцелей (функций) можно ставить вопрос о том, какие из них в условиях конкретного предприятия являются более значимыми, по каким нужно осуществлять централизованное управление, создавать автоматизированные процедуры поддержки принятия решений и т. п.

Эффект появления нового смысла исследуется в теории множеств (при формировании множеств из «пар», «троек», «*n*-ок» элементов исходных множеств), в математической лингвистике и теории языков (при создании тезаурусов).

Например, в соответствии с тезаурусным или лингвистическим представлением рассматриваемой ситуации можно сформировать списки понятий (классификаторы по признакам структуризации) и, помещая рядом составляющие из разных списков (операция *конкатенации*), получить таким образом все возможные конкретизированные функции, которые должны быть в структуре целей и функций при принятой концепции системы и соответствующей методике структуризации.

Организация такой процедуры базируется на математической операции «размещения с повторениями». Программа для ее реализации должна содержать несколько вложенных циклов (по числу списков). Однако число получаемых размещений с повторениями согласно известной теореме комбинаторики определяется как $R = k_1 * k_2 * \dots * k_n$, где k_1, k_2, \dots, k_n – число элементов в исходных списках (классификаторах), и если число списков (признаков структуризации), например, 7, и в каждом классификаторе содержится хотя бы по 7 составляющих, то $R = 7 * 7$, что много не только для человека, но и для ЭВМ, а если элементов и списков больше, то перебор (хотя число размещений с повторениями меньше, чем число сочетаний) становится практически нереализуемым.

Поэтому решение проблемы автоматизации формирования структуры ЦФ оказалось возможным только в диалоговом ре-

жиме. Опираясь на основную идею «дерева целей» – отсечение малозначимых (несущественных или несуществующих) элементов «дерева целей» на каждом шаге структуризации – можно существенно сократить перебор и осуществить автоматизацию формирования структуры ЦФ.

На основе этой идеи были определены принципы разработки автоматизированных диалоговых процедур анализа целей и функций, базирующиеся на получении размещений с повторениями из исходных списков классификаторов по используемым признакам структуризации, в результате чего возникает известный в теории множеств эффект появления нового смысла, уточнения подцелей и функций.

Автоматизированная процедура реализуется в диалоговом режиме, что позволяет сократить перебор путем отсечения малозначимых или не имеющих смысла размещений на каждом шаге структуризации.

Идея организации такой человеко-машинной процедуры иллюстрируется рис. 4.12.

На рис. 4.12 приводится пример применения методики, базирующейся на двойственном определении системы *А. И. Уёмова*, согласно которой на верхних уровнях иерархической структуры применяются признаки структуризации «объект управления (OU)» – «цикл управления (ZU)», а в качестве третьего признака структуризации используется признак «виды продукции (VP)», классификатор которого включает автомобили (АВТ), запасные части (З/Ч), товары народного потребления (ТНП). Участие человека показано ответами по двоичной систем «+» – «-».

Естественно, в классификаторы вводятся только ключевые слова. Поэтому на рис. 4.12 приведены в качестве примера формулировки подцелей, получаемые на основе отобранных размещений, составленных из ключевых слов. Глобальная цель названа – «Основные направления (ОН) деятельности предприятия».

Разрабатывая процедуру автоматизации формирования структуры ЦФ, желательно предусмотреть возможность изменения последовательности

признаков структуризации при выводе результатов анализа на дисплей или принтер.

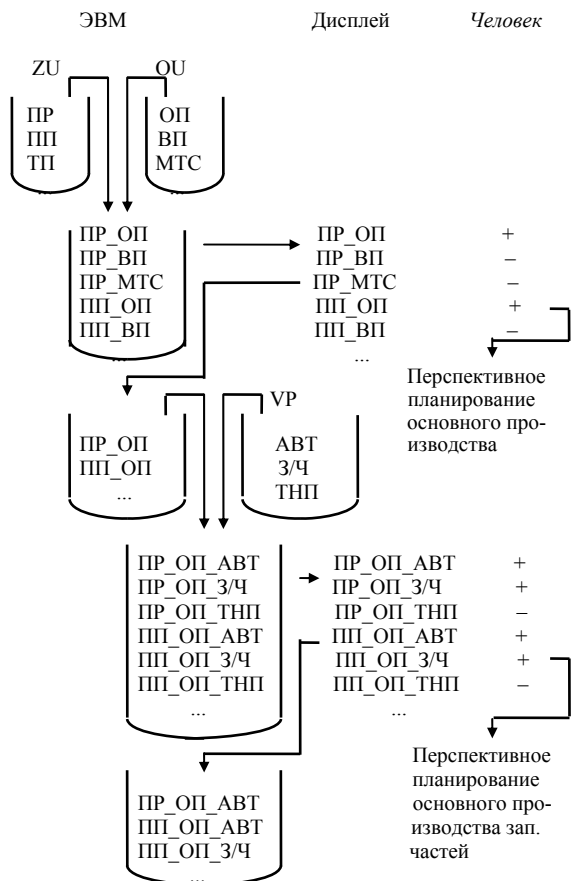


Рис. 4.12. Принципы автоматизации формирования структур ЦФ

Пример двойственных вариантов структуры для двух верхних уровней был приведен на рис. 4.8, а и б.

Появление нового смысла может быть связано не только с операцией конкатенации. Между уровнями могут быть и причинно-следственные

отношения, и отношения подчиненности, включения и т. п., что, например, имеет место при использовании других методик структуризации.

После разработки структуры целей и функций проводится оценка составляющих ее структуры для выявления наиболее значимых. При этом следует параллельно использовать *экспертные* и *косвенные количественные оценки*, которые затем обрабатываются совместно (см., например, [1, 13, 16]).

При экспертной оценке в качестве критериев используется система критериев, аналогичная принятой в методике ПАТТЕРН⁸⁹, но с некоторыми модификациями: в ней предлагаются следующие группы критериев – относительной важности (значимости), взаимосвязанности, экономические оценки (последняя – заменяет группу критериев «состояние-срок»).

Идея косвенных количественных оценок предложена на основе анализа иерархических структур с помощью информационного подхода, из результатов которого следует, что структурированность ветвей иерархической структуры определяет придаваемую им фактическую значимость⁹⁰.

В качестве косвенных количественных оценок могут быть использованы: число подразделений, выполняющих данную функцию, число документов, подготавливаемых для реализации функции, структурированность ветви (например, число тем, на которые разделена проблема), внимание к соответствующим подцелям и функциям в директивных документах, периодической печати, источниках НТИ и т.п. При выборе косвенных количественных оценок следует учитывать «пространство инициирования целей и факторов», т. е. учитывать требования и потребности надсистемы, отраженные в законодательных актах и директивных документах, аналогичных предприятий актуальной среды, интересы подведомственных подразделений, инициативы структурных единиц собственно системы управления

⁸⁹ **Лопухин М. М.** ПАТТЕРН – метод планирования и прогнозирования научных работ / М. М. Лопухин. – М.: Сов. радио, 1971. – 160 с.

⁹⁰ **Волкова В. Н.** Оценка целевых структур при разработке планов в системе высшей и средней специальной школы: Экспресс-информация / В. Н. Волкова, А. И. Коношенко, А. Ф. Марьенко – М.: НИИВШ, 1980. – 19 с.

При обработке результатов оценки применяются не только традиционные методы усреднения, но и выявление *противоречивых мнений* с последующим содержательным анализом этих оценок.

При этом для сопоставления оценок удобно применять графические представления в виде гистограмм.

Для сбора и обработки экспертных и косвенных количественных оценок и для сравнительного информационного анализа структур следует создать группу соответствующих специалистов. В необходимых случаях могут быть использованы методы организации сложных экспертиз, базирующиеся на информационном подходе [1, 3, 13].

На основе полученных результатов оценки исходная структура корректируется: составляющие, получившие наименьшие оценки значимости по сравнению с другими и не получившие при этом высоких оценок связности с высокозначимыми, либо исключаются из структуры ЦФ, либо опускаются на нижележащие уровни иерархии, и, напротив, составляющие, получившие высокие оценки значимости, могут быть перенесены на более высокие уровни иерархической структуры. При такой корректировке могут возникнуть вырожденные ветви, разные варианты новой структуры. Может возникнуть необходимость изменить логику формирования структуры.

В результате получается структура ФЧ АИС, представляющая собой совокупность подсистем, подлежащих автоматизации и задач, на которые делятся подсистемы. При необходимости вначале могут быть выделены группы задач. А при увеличении числа подсистем подсистемы могут быть объединены в направления, а если возникнет необходимость, то формируются многоуровневые структуры, подобно тому, как это было сделано на ВАЗе (рис. 4.6).

4.4. Разработка структуры обеспечивающей части АИС

Разделение АСУ на функциональную и обеспечивающую части, а последней – на информационное обеспечение (ИО), техническое (ТО), организационное (ОргО), программное (ПО) и другие виды обеспечения, позволило привлечь для уточнения соответствующих видов обеспечения специалистов в этих областях.

Такой подход к организации разработок АСУ помог справиться со сложностью системы и ускорить разработку АСУ путем параллельного проведения работ по анализу и выбору структуры отдельных видов обеспечения. Однако, если разрабатывать отдельные проекты ИО, ТО, ОргО и других видов обеспечения, то после разработки этих проектов возникает достаточно сложная задача их согласования, взаимоувязки принятых структур этих видов обеспечения, критериев, учитываемых при их разработке и т. д.

Поэтому на определенном этапе развития работ по созданию АСУ был даже сформулирован специальный принцип – единства ИО, ТО и ОргО как основных видов обеспечения, при разработке структур которых возникали несогласованности – и было рекомендовано проектировать структуру ОЧ АСУ с самого начала как единую с уточнением структур отдельных видов обеспечения в рамках общего проекта.

В качестве примера приводится подход к разработке структуры ОЧ АИС, предложенный в 1970-е годы на основе идеи постепенной формализации модели принятия решений⁹¹.

Задача обоснования структуры ОЧ АСУ была сформулирована следующим образом.

Под структурой ОЧ АСУ понимается сеть информационных служб (Главный информационно-вычислительный центр, локаль-

⁹¹ **Волкова В. Н.** К методике проектирования автоматизированных информационных систем / В. Н. Волкова // Автоматическое управление и вычислительная техника. – М.: Машиностроение, 1975. – Вып. 11. – С. 289–300; **Волкова В. Н.** Постепенная формализация моделей принятия решений / В. Н. Волкова. – СПб.: Изд-во Политехн. ун-та, 2006. – 120 с.

ные вычислительные центры производств, цехов и других подразделений, автоматизированные рабочие места и другие составляющие ОргО) с размещенными в ней массивами хранения информации, документами (ИО), техническими средствами регистрации, хранения, передачи, обработки, представления информации (ТО), программным обеспечением (ПО), методическим обеспечением (инструкциями для пользователей, положениями о подразделениях и т. п.) и другими видами обеспечения.

Задача выбора структуры ОЧ формулируется как задача наилучшего размещения всех этих компонентов с учетом единой согласованной системы критериев и ограничений, обеспечивающей наиболее эффективную реализацию подсистем и задач, включенных в структуру ФЧ АСУ на соответствующем этапе ее развития (т. е. соответствующей очереди АСУ).

В такой постановке задача на первый взгляд кажется практически нерешаемой. Действительно, представить эту задачу классом *хорошо организованных* систем, т. е. создать математическую модель, в которой взаимосвязи между компонентами структуры ОЧ и целями АСУ (задачами, входящими в структуру ФЧ) были бы описаны в виде аналитических зависимостей, практически невозможно.

Более реально представить эту задачу моделью *плохо организованной* системы, т. е. использовать статистические исследования, отражающие основные характеристики потоков информации, и на этой основе предложить структуру ОргО, информационных массивов, требуемые технические средства, т. е. разработать ориентировочный вариант структуры ОЧ. Однако и этот путь – не лучший с точки зрения затрат времени и доказательства правомерности принятого решения.

Такой подход возможен при корректировке структуры ОЧ, когда есть предварительные исследования информационных потоков. В условиях же создания новой АИС получение статистики на основе «ручного» способа сбора информации может дать неверные результаты, поскольку, во-первых, «ручной» способ обычно реализован не так, как в автоматизиро-

ванном варианте, а во-вторых, при автоматизации могут принципиально измениться формы сбора информации.

Наиболее целесообразно отобразить задачу с помощью класса *самоорганизующихся, развивающихся систем* и организовать процесс «выращивания» структуры ОЧ с помощью модели постепенной формализации задачи.

На практике после определения видов обеспечения процесс формирования модели для решения данной задачи, приведенный на рис. 4.13, был несколько упрощен.

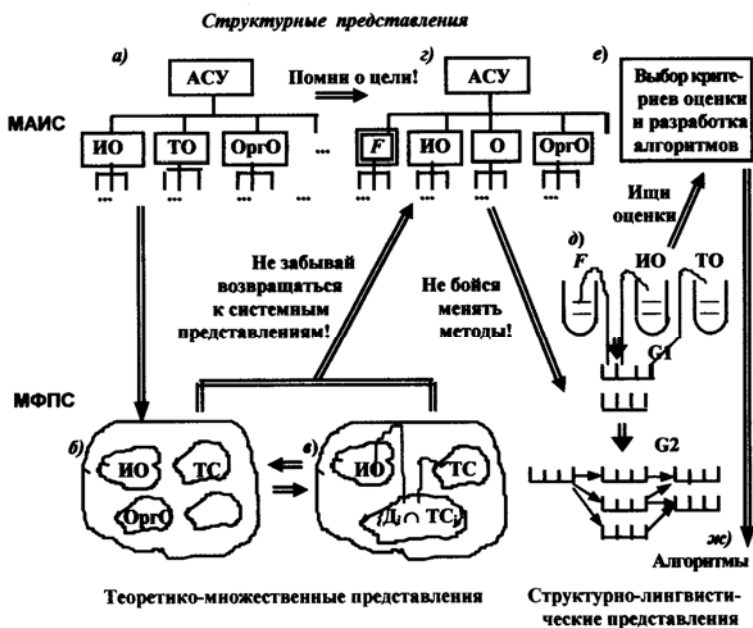


Рис. 4.13. Постепенная формализация модели обоснования структуры ОЧ АИС

Так, при решении задачи формирования структуры ОЧ АИС первый этап может быть определен тем, что ОЧ АИС разделена на виды обеспечения, т. е. сформирована структура-состав ОЧ (рис. 4.13, а).

Затем для уточнения состава видов обеспечения применены теоретико-множественные представления (рис. 4.13, б), помогающие вначале расширить состав видов обеспечения, найти способы объединения элементов из разных подмножеств (рис. 4.13, в).

В практике разработки АСУ этому этапу соответствовал поиск взаимосвязей между компонентами ОЧ с помощью формирования и экспертной оценки матриц типа «ТО–ОргО» или «ИО–ТО». На основе оценки элементов матриц «ТО–ОргО» размещались технические средства в подразделениях предприятия. В матрицах «ИО–ТО» оценивалось, с помощью каких технических средств следует обрабатывать соответствующие документы, что первоначально при ограниченных мощностях ЭВМ иногда было оправдано.

Однако на основе анализа матриц, отображающих некоторые взаимосвязи между компонентами ОЧ, решить задачу формирования структуры ОЧ не удастся. Анализ матриц, взаимоувязка полученных с их помощью взаимосвязей между элементами – весьма трудоемкая работа, задерживающая разработку АИС. Поэтому согласно рассматриваемому подходу, целесообразно *возвратиться к системно-структурным представлениям*, с помощью которых активизируется использование интуиции и опыта ЛПР, перечень множеств анализируется и при необходимости дополняется принципиально важными подмножествами для дальнейшего моделирования.

В частности, в рассматриваемой задаче перечень исходных подмножеств (ИО, ТО, ОргО соответственно) дополнен подмножеством задач, входящих в структуру ФЧ АСУ, и функций F (рис. 4.16, з), которые выполняются при решении задач сбора и первичной обработке информации, входящими в структуру ФЧ АСУ 1-й очереди.

Для дальнейшей реализации идеи комбинирования элементов в поисках вариантов решения задачи (т. е. путей прохождения информации при ее сборе и первичной обработке) могут быть выбраны более удобные и подсказывающие правила формирования вариантов *лингвистические* представления, являющиеся основой разработки языка моделирования путей прохождения информации.

С помощью языка моделирования можно получить многоуровневую модель (рис. 4.13, д), позволяющую формировать варианты информационных потоков. После формирования вариантов следования информации выбраны критерии и способы оценки графо-семиотической модели (рис. 4.13, е). Приняты разные варианты – от экспертной оценки вариантов сбора и первичной

обработки информации (нижний уровень рис. 4.13, д) до поиска алгоритмов последовательного преобразования оценок компонентов предшествующих уровней модели в оценки компонентов последующих уровней.

Структура полученной методики формирования и анализа информационных потоков приведена на рис. 4.14.



Рис. 4.14. Методика обоснования структуры ОЧ АИС

Этапы 1–7 выполняются по каждой из задач ФЧ АСУ, а затем проводится содержательная интерпретация, обобщение результатов моделирования задач ФЧ АСУ и определение на этой основе состава необходимых и достаточных компонентов структуры ОЧ.

Последовательность этапов может быть иная, чем на рис. 4.13 и 4.14. Например, вначале можно применить целевой подход, т. е. проводить активное обследование, отбор элементов модели с помощью предварительно сформированной структуры функций, на основе которой провести опрос сотрудников исследуемого или аналогичных объектов, получить необходимые сведения в документах и т. д. Или можно обратиться к оценкам раньше, чем это сделано на рис. 4.13, – оценить значимость подцелей и функций, что помогло бы ускорить выбор лучшего варианта и т. д. При выполнении этапа 5 полезно обращаться к этапу 3, а при выполнении этапа 7 – к этапу 5, что показано на рис. 4.14.

Приведенный пример демонстрирует, что постепенная формализация становится своего рода «механизмом» развития системы, «выращивания» модели принятия решения. По мере развития

модели происходит смена методов из групп МАИС и МФПС. В результате можно получить разные варианты формализованной модели. В процессе моделирования следует учитывать рекомендации, приведенные на рис. 2.6 и 4.13, типа *«используй то, что знаешь»*, *«не увлекайся перечислением»*, *«не забывай возвращаться к системным представлениям»*, *«помни о цели»*, *«не бойся менять методы»* и т. п. [4].

При «выращивании» модели можно накапливать информацию об объекте, фиксируя все новые компоненты, связи, правила взаимодействия компонент, и, используя их, получать отображения последовательных состояний развивающейся системы, постепенно создавая все более адекватную модель реального, изучаемого или создаваемого объекта. При этом информация может поступать от различных специалистов и накапливаться во времени по мере ее возникновения в процессе развития объекта и наших представлений о нем.

При постановке задачи для принципиально нового объекта или процесса постепенная формализация позволяет обосновать принципы разработки языка автоматизации моделирования и обобщенную формальную модель с пошаговым доказательством ее адекватности на каждом витке моделирования.

Рассмотренный пример иллюстрирует основную суть подхода – применение идеи постепенной формализации для формирования нового алгоритма, получения новой методики системного анализа в тех случаях, когда имеет место большая начальная неопределенность задачи и отсутствуют аналоги для ее решения.

Такого рода подходы и методики применялись достаточно редко. Обычно структура АСУ развивалась на основе аналогий и возникающих текущих потребностей. В то же время по мере усложнения автоматизированных систем предприятий и организаций усложняется и управление ходом их развития, и идея приведенного подхода может оказаться полезной.

На практике может быть применена упрощенная методика разработки и развития АИС и АСУ организации, идея которой приводятся разделе 4.5.

4.5. Разработка методики для проектирования АИС предприятия

Для достаточно крупных предприятий в соответствии с Общеотраслевыми руководящими методическими материалами по созданию АСУП был определен порядок разработки АСУ, приведенный на рис. 4.1 и 4.2. Разрабатывались методики формирования структур функциональной и обеспечивающей частей АСУ, подобные приведенным выше.

Однако в последующем практика показала, что принцип единства ИО, ТО и ОргО как основных видов обеспечения, принятый в рассмотренной методике обоснования обеспечивающей части АСУ, важен при разработке первой очереди АСУ, когда принимаются принципиальные решения по выбору технических средств и структуры организационного обеспечения.

Следующие очереди разрабатываются в условиях выбранных технических средств, и развитие технической базы осуществляется с учетом совместимости с существующими, а главной задачей становится формирование структуры *информационного обеспечения*.

Поэтому в методике разработки структуры ФЧ и ОЧ для предприятия (организации) предусматривается два этапа:

1. Разработка (развитие) структуры функциональной части АСУ.

При выполнении этого этапа разрабатывается структура целей и функций организации (подэтап 1.1), проводится оценка составляющих этой структуры (подэтап 1.2.) с использованием *экспертных* и *косвенных количественных* оценок и на этой основе определяется состав подсистем и задач, подлежащих автоматизации в первую очередь, т.е. сформировать структуру ФЧ АИС.

2. Определение состава информационных массивов, необходимых для реализации подсистем и задач ФЧ АИС.

Эту задачу, как правило, можно выполнить, задав предварительно совокупность информационных массивов (с учетом уже созданных), и сопоставив их с подсистемами и задачами ФЧ АИС, определить значимость массивов для реализации этих задач и функции, которые не обеспечены информацией, и дополнить совокупность массивов новыми.

Пример методики приведен на рис. 4.15.

1. Формирование структуры ФЧ АСУ

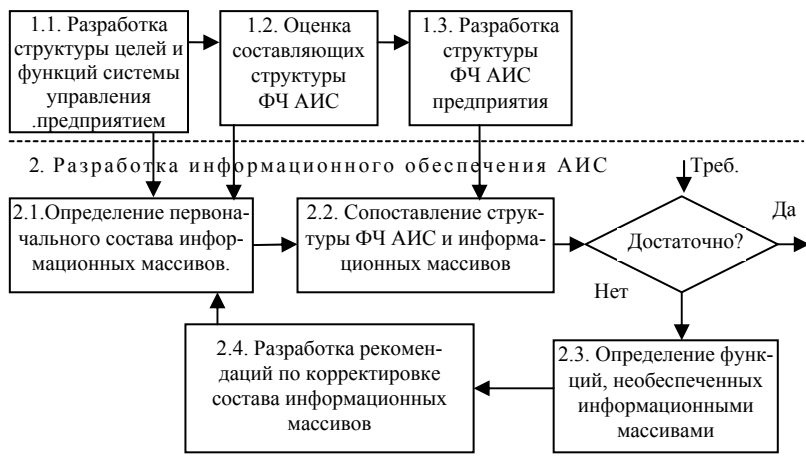


Рис. 4.15. Структура методики проектирования информационной системы

При выполнении подэтапа 1.1 применяются методики структуризации целей систем управления и автоматизированные диалоговые процедуры анализа целей и функций

При выполнении подэтапа 1.2 для оценки структуры целей и функций путем выявления наиболее значимых составляющих предлагается параллельно используются *экспертные оценки* и *косвенные количественные оценки*, которые затем обрабатываются совместно.

При экспертной оценке в качестве критериев используется система критериев, аналогичная принятой в методике ПАТТЕРН⁹², но с некоторыми модификациями: в ней предлагаются следующие группы критериев – относительной важности (значимости), взаимосвязанности, экономические оценки (последняя заменяет группу критериев «состояние-срок»).

⁹² Лопухин М. М. ПАТТЕРН – метод планирования и прогнозирования научных работ / М. М. Лопухин. – М.: Сов. радио, 1971. – 160 с

Идея косвенных количественных оценок предложена ⁹³ на основе анализа иерархических структур с использованием информационного подхода из результатов которого следует, что структурированность ветвей иерархической структуры определяет придаваемую им фактическую значимость.

В качестве косвенных количественных оценок могут быть использованы: число подразделений, выполняющих данную функцию, число документов, подготавливаемых для реализации функции, частота обращения к выполнению функции и информационным массивам, обеспечивающим ее выполнение, внимание к соответствующим подцелям и функциям в директивных документах, периодической печати, источниках НТИ и т.п. При выборе косвенных количественных оценок следует учитывать «пространство инициирования целей и факторов», т. е. учитывать требования и потребности надсистемы, отраженные в законодательных актах и директивных документах аналогичных предприятий актуальной среды, интересы подведомственных подразделений, инициативы структурных единиц собственно системы управления

При обработке результатов оценки применяются не только традиционные методы усреднения, но и выявление *противоречивых мнений* с последующим содержательным анализом этих оценок. При этом для сопоставления оценок удобно применять графические представления в виде гистограмм.

При выполнении подэтапа 1.3 на основе полученных результатов оценки исходная структура корректируется: составляющие, получившие наименьшие оценки значимости по сравнению с другими и не получившие при этом высоких оценок связности с высокозначимыми, либо исключаются из структуры ЦФ, либо опускаются на нижележащие уровни иерархии, и, напротив, составляющие, получившие высокие оценки значимости, могут быть перенесены на более высокие уровни иерархической структуры. При такой корректировке могут возникнуть вырожденные ветви, разные варианты новой структуры целей и функций. В последнем случае следует перейти к подэтапу 1.1 и повторить формирование и оценку структуры ЦФ.

⁹³ Волкова В. Н. Оценка целевых структур при разработке планов в системе высшей и средней специальной школы: Экспресс-информация / В. Н. Волкова, А. И. Коношенко, А. Ф. Марьенко – М.: НИИВШ, 1980. – 19 с.

На основе этого анализа формируется структура ФЧ, в которой компоненты первого сверху уровня структуризации первоначально называли подсистемами, нижележащего по отношению к уровню подсистем – группами задач или задачами. По мере развития АСУ названия менялись, и при необходимости формировались многоуровневые структуры ФЧ (см. структуру ФЧ АСУ ВАЗа). Но для малых и средних предприятий обычно сохранялось представление структуры ФЧ в виде древовидной иерархической структуры.

При разработке информационного обеспечения (этап 2) в совокупности массивов могут быть массивы, обеспечивающие информацией производственный процесс, материально-техническое снабжение и т. п. массивы фактографической информации, и массивы научно-технической информации. После того, как определен окончательный состав массивов, разрабатывается структура информационного обеспечения, т. е. определяется структура входных, архивных, информационно-поисковых массивов, взаимосвязи между массивами и т. п.

Принципы построения и эффективность АСУ существенно зависят от уровня развития информационных технологий. С появлением в середине 1970-х гг. персональных ЭВМ происходит корректировка идеи АСУ: от ВЦ и централизации управления – к распределенному вычислительному ресурсу и децентрализации управления. Для обеспечения информацией по группам основных функций организационного управления предприятиями в настоящее время приобретают готовые программные продукты. Вариант методики для выбора готовых программных продуктов приводится в следующем разделе.

Особенно актуален выбор готовых программных продуктов для предприятий малого и среднего бизнеса, которые не имеют средств для проведения собственных разработок по созданию программного обеспечения для создания необходимых информационных массивов.

4.6. Выбор готовых программных продуктов

Предприятия малого и среднего бизнеса, как правило, не имеют средств для внедрения единой автоматизированной системы, поэтому существует практика выбора имеющихся на рынке программных продуктов, автоматизирующих соответствующие области управления (бухгалтерский учет, материально-техническое обеспечение и т. п.) с последующей интеграцией в единую систему.

При выборе готовой информационной системы (ИС) обычно учитывают технические и стоимостные характеристики: количество рабочих мест (с возможностью одновременного доступа), разделение функций пользователей, скорость функционирования системы, сложность ее администрирования, проработанность механизмов, обеспечивающих права доступа, стоимость ИС и т. п. В то же время важно учитывать такие критерии как полнота автоматизации функций управления и степень использования возможностей приобретаемой готовой ИС.

Поэтому методика должна быть основана на сопоставлении функций системы управления предприятия (или подсистемы АСУ) и функциональных возможностей существующих на рынке программных продуктов и определении *степени автоматизации* функций конкретного предприятия с помощью выбираемого программного средства и *коэффициент использования* этого средства.

Степень автоматизации определяется как отношение количества автоматизированных за счёт конкретного программного продукта функций f_{ia} к общему количеству функций f_{io} организации:

$$A = \frac{\sum f_{ia}}{\sum f_{io}}, \quad (4.4)$$

Степень использования функциональных возможностей автоматизированной системы:

$$k_{и} = \frac{\sum \varphi_{ia}}{\sum \varphi_{is}}, \quad (4.5)$$

где φ_{ia} – число функциональных возможностей программного продукта, используемых в кадровой службе конкретной организации; φ_{is} – общее число функциональных возможностей программного продукта.

Структура методик приведена на рис. 4.16.

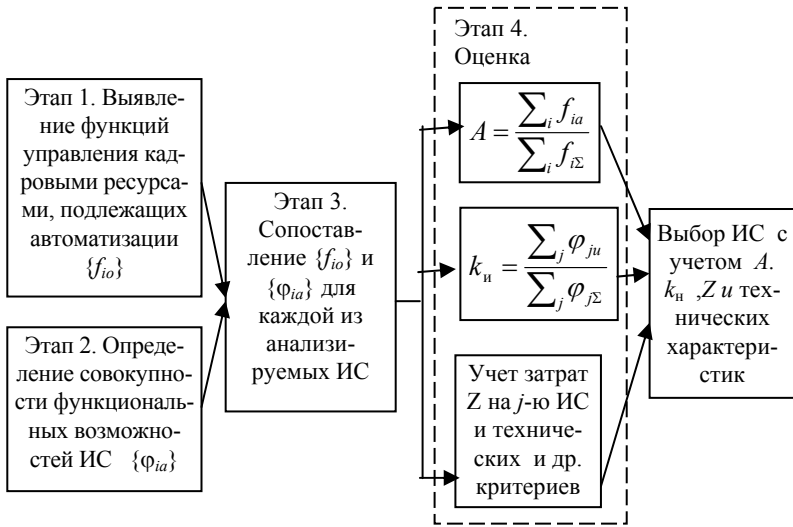


Рис. 4.16. Структура методики выбора ИС

При разработке методики для конкретной организации помимо приведенных критериев степени автоматизации и коэффициента использования функциональных возможностей автоматизированной системы, следует учитывать технические и экономические критерии, значимые для организации и могут быть поставлены оптимизационные задачи или задачи типа оптимальной кластеризации многомерных объектов⁹⁴.

⁹⁴ Мышенков К. С., Комплексный критерий качества и классификации многомерных объектов / К.С. Мышенков, В. И. Карпов, В. В. Гетьман. Свидетельство об официальной регистрации программ для ЭВМ № 200631.

4.7. Интегрированные автоматизированные системы управления

В 80-90-е гг. XX века в связи с необходимостью объединения в единую систему разнообразных автоматизированных систем, разрабатывавшихся на предприятии возник и был широко распространен термин «*Интегрированные автоматизированные системы управления*» (ИАСУ)⁹⁵.

В соответствии с объективными процессами материального производства и управления создавались информационные системы управления следующих основных типов:

- автоматизированные системы научных исследований (АСНИ);
- системы автоматизированного проектирования изделий (САПР) и автоматизированные системы технологической подготовки производства (АС ТПП);
- автоматизированные системы общезаводского управления, ориентированные на автоматизацию функций управления объединением (корпорацией, трестом, концерном и др.) (АСУО) и предприятием (АСУП);
- автоматизированные системы для комбинированного организационного и технологического управления (АСУ ОТ);
- автоматизированные системы управления гибкими производствами (АСУ ГПС), включая АСУ отдельными производствами, цехами, участками, гибкими автоматизированными линиями (ГАЛ), гибкими производственными модулями (ГПМ);
- автоматизированные системы управления технологическими процессами (АСУ ТП);

⁹⁵ При подготовке данного раздела использовалась информация из статьи «Интегрированные автоматизированные системы управления промышленным предприятием» из Справочника [17].»

- автоматизированных систем контроля качества и испытания изделий (АСКИ).
- системы автоматизированного проектирования (САПР).

Требования по повышению уровня организации производства, обеспечению четкой координации действий подразделений предприятия могут быть выполнены путем интеграции всех функций управления в единую интегрированную систему управления (ИАСУ). ИПСУ должна комплексно обеспечивать автоматизацию процессов стратегического планирования, экономического и технического развития предприятия, маркетинговых и научных исследований, проектирования новых видов продукции, управления технической подготовкой производства, текущей производственно-хозяйственной, сбытовой и финансовой деятельностью и, наконец, автоматизацию управления основными и вспомогательными технологическими операциями.

Под *интегрированными автоматизированными системами управления* (ИАСУ) понимаются многоуровневые человеко-машинные системы, охватывающие решение задач управления экономикой предприятия, административной деятельностью, исследованиями конъюнктуры рынка, реализацией товаров, а также задач проектирования изделий, технической подготовки производства, организации и управления технологическими процессами.

Структуры ФЧ ИАСУ может быть представлена подобно рис. 4.5 и 4.6.

4.8. Общегосударственная автоматизированная система (ОГАС)

Как было сказано в гл. 2, в нашей стране работы по автоматизации управления проводились с 1960-х гг. на основе системно-целевого подхода. АСУ создавались для всех уровней и направлений управления страной: государственного, отраслевого, регионального, предприятий и организаций.

Задача построения общегосударственной автоматизированной системы управления была поставлена первым заместителем Председателя Совета Министров (тогда *А.Н. Косыгиным*) в ноябре 1962 года ⁹⁶.

Были создана специальная комиссия под председательством *Б.Н. Малиновского* ⁹⁷, в которую вошли ученые-экономисты, в частности, академик *Н.П. Федоренко*, начальник ЦСУ *В.Н. Старовский*, первый заместитель министра связи *А.И. Сергийчук*, а также другие работники органов управления

К этому времени у нас в стране уже имелась концепция Единой государственной сети вычислительных центров (ЕГСВЦ) для обработки экономической информации, которую выдвинули академик *В.С. Немчинов* и его ученики. А в конце 1950-х гг. *А.И. Китова* в письме главе СССР *Н.С. Хрущеву* (которое он послал в ЦК КПСС 7 января 1959 г.), предложил создать общенациональную компьютерную сеть многоцелевого назначения, предназначенную для планирования и управления экономикой в масштабе всей страны на основе Единой государственной сети вычислительных центров (ЕГСВЦ), отмечают ряд известных отечественных и зарубежных историков науки ⁹⁸, Имеется и ряд публикаций *А.И. Китова*, в которых содержится эта идея ⁹⁹.

В 1963 году председателем Межведомственного научного совета по внедрению вычислительной техники и экономико-математических методов в народное хозяйство СССР при Государственном комитете Совета Министров СССР по науке и тех-

⁹⁶ *Малиновский Б.Н.* История вычислительной техники в лицах / Б.Н. Малиновский. – Киев: Изд-во «КИТ», ПТОО «А.С.К.», 1995. – 384 с..

⁹⁷ Советский и украинский учёный, ветеран вычислительной техники, член-корреспондент НАН Украины, доктор технических наук, профессор. Заслуженный деятель науки и техники Украины

⁹⁸ www.kitov-anatoly.ru

⁹⁹ *Китов А. И.* Автоматизация управленческих работ / А.И. Китов, Ю.И. Черняк // Автоматизация производства и промышленная электроника. Т. 1, М.: Государственное научное издательство «Советская энциклопедия», 1962. С. 26-32..И до.

нике был утвержден **В.М. Глушков.**, который был главным идеологом разработки и создания *Общегосударственной автоматизированной системы учёта и обработки информации* (ОГАС), предназначенной для автоматизированного управления всей экономикой СССР в целом. Для этого им была разработана система алгоритмических алгебр и теория для управления распределёнными базами данных¹⁰⁰.

Конечной целью ОГАС было объединение в единую систему тысяч вычислительных центров (ВЦ), отдельных автоматизированных систем управления предприятиями (АСУП) и автоматизированных систем управления отраслями народного хозяйства (ОАСУ). Несмотря на то, что сам проект ОГАС был заморожен, отдельные его элементы претворялись в жизнь, особенно в области военной промышленности. На отдельных предприятиях создавались системы АСУ-70 и АСУ-80.

По мнению **В.М. Глушкова** наличие планового хозяйства в СССР позволило создать самую эффективную систему управления экономикой, и понимая это, он сделал ставку на ОГАС. Создание ОГАС давало уникальную возможность объединить информационную и телекоммуникационную структуру в стране в единую систему, позволявшую на новом научно-техническом уровне решать вопросы экономики, образования, здравоохранения, экологии, сделать доступными для всех интегральные банки данных и знаний по основным проблемам науки и техники, интегрироваться в международную информационную систему. Реализацию ОГАС могла бы вывести страну на новый уровень развития, соответствующий постиндустриальному обществу.

В публикациях и в интернете обсуждаются политические причины того, что ОГАС не была реализована. Однако, представляется, что проблема была не только в этом. Для создания

¹⁰⁰ **Глушков В. М.** Введение в кибернетику / В. М. Глушков. – Киев: Изд-во АН УССР, 1964.

такой сложной системы **В. М. Глушков** предлагал разные концепции ¹⁰¹.

Первый вариант концепции ОГАС был основан на принципе централизованного управления, в соответствии с которым все создаваемые отраслевые, республиканские, территориальные вычислительные центры подчинялись головной организации – Главному вычислительному центру (ГВЦ) Госплана СССР, созданного при Госплане СССР, но размещенном в самостоятельном здании.

Была предложена трехступенчатая система: Первая ступень ОГАС – пункты сбора и первичной обработки информации, а также вычислительных центров предприятий и некоторых исследовательских организаций. Вторая ступень ОГАС - концентрировать основные вычислительные мощности в нескольких десятках крупных опорных центров с мощностью каждого центра порядка 1-1,5 млн. операций в секунду. Эти центры должны быть расположены в местах наибольшей концентрации потоков экономической информации и обслуживать прилегающую к ним территорию. Третья ступенью ОГАС – головной вычислительный центр, осуществляющий оперативное руководство всей сетью вычислительных центров страны и непосредственно обслуживающий высшие правительственные органы. Помимо базовых крупных вычислительных центров, в сеть должны были включаться порядка 20 тысяч вычислительных центров предприятий и организаций, занимавшихся непосредственной обработкой экономической информации.

¹⁰¹ Автор учебника имела возможность слушать доклады **В.М. Глушкова**, в которых он излагал концепции ОГАС, на заседаниях Ученого совета Все-союзного научно-исследовательского института проблем организации управления (ВНИИПОУ), директором которого был **Д. Жимерин**, а научным руководителем разработки ОГАС **В.М. Глушков**, который работал в Институте кибернетики в Киеве, но приезжал на заседания ученого совета во ВНИИПОУ.

Этот вариант не был принят, поскольку предусматривал усиление роли Госплана, и в нем неясна была роль Совета Министров СССР и Государственного Комитета по науке и технике (ГКНТ). при Совете Министров СССР, Центрального статистического управления (ЦСУ) СССР, руководитель которого пытался предложить свой проект ОГАС, в котором ведущая роль отводилась ЦСУ СССР, уже накопившему значительные массивы экономико-статистической информации.

Второй вариант, предложенный **В. М. Глушковым** – объединение всех создаваемых ВЦ и АСУ. Однако достаточно быстро стало ясно, что АСУ отрасли и АСУ отраслевого министерства – не одно и то же, поскольку последнее всегда в большей мере ориентировано на решение собственных нужд сотрудников министерства. Аналогичная проблема возникает и при разработке территориальных АСУ (ТАСУ): АСУ органов государственного управления также в большей мере ориентированы на собственные нужды сотрудников, обеспечение их информацией в соответствии с их потребностями. Кроме того, такой принцип в условиях СССР, управление в котором было подчинено строго централизованному принципу, не мог быть принят.

Тогда **В. М. Глушков** предложил **третий вариант**, основанный на концепции многоуровневой структуры, сочетающей *территориальный* и *отраслевой* принципы и принципы *вертикальных* и *горизонтальных* связей между автоматизированными системами управления и вычислительными центрами

Вертикальная цепочка связей соответствует установленному порядку подчиненности в структуре управления народным хозяйством: предприятие – объединение – министерство – Госплан – Совет Министров.

Горизонтальные связи устанавливаются между различными предприятиями, учреждениями, организациями, системами, относящимися к разным министерствам и ведомствам, т. е. находящимися в различных цепочках вертикальных связей.

В. М. Глушков считал, что для эффективного планирования необходима «... унифицированность, полнота, объективность и иерархическая организованность информации, возможность быстрого (автоматического) обмена информацией между любыми предприятиями и органами управления любого уровня ... независимо от их ведомственной принадлежности»¹⁰².

Для реализации этой концепции предлагалось «... в ОГАС» должно существовать централизованное (межведомственное) звено, выполняющее функции диспетчеризации и коммуникации сообщений. Его техническую базу составляет система общегосударственных (межведомственных) информационно-вычислительных центров (ОГИВЦ), являющихся одновременно и центрами коммуникации сообщений. Расположение этих центров определяется конфигурацией системы связи»²²⁰.

Такой вариант структуры ОГАС **В. М. Глушков** иллюстрировал рисунком 4.17.

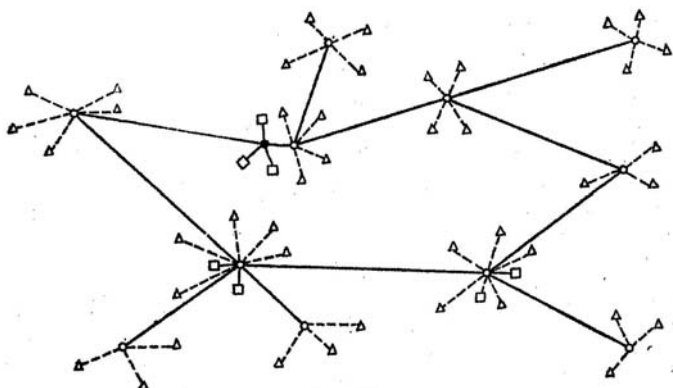


Рис. 4.17. Вариант структуры ОГАС

Для обеспечения взаимодействия между ОГИВЦ предлагалось создать человеко-машинную Общегосударственную диспетчерскую службу (ОГДС), а несколько позднее – Общегосу-

¹⁰² Глушков В.М. Введение в АСУ / В.М. Глушков. – Киев: Техніла, 1972. – С. 291 – 303.:

дарственную систему диспетчерских служб (ОГСДС). А на нижних уровнях предполагалось создавать информационно-диспетчерские пункты (ИДП).

Возникшие при обсуждении этой концепции вопросы о роли существующих общегосударственных органов управления предлагалось решить вначале расположением головной ОГИВЦ в непосредственной близости от места сосредоточения центральных органов управления и соединением последних с головным ОГИВЦ широкополосными каналами связи. Аналогично решить вопрос и в республиках и для территориальных образований создавать даже нескольких ГИВЦ, работающих по принципу вычислительных центров коллективного пользования (ВЦКП). А в последующем все же в ОГАС при общегосударственных и территориальных органах управления было предложено создавать соответствующие АСУ – Автоматизированную систему плановых расчетов (АСПР) при Госплане, АСУ государственной статистик (АСГС), АСУ Госстроя (АСУС), АСУ Госснаба (АСУГ), автоматизированную систему Госкомтруда (АСТЗ), АСУ ГКНТ (АСУНТ), АСФЮ – для обеспечения финансово-юридической деятельности по управлению госбюджетом, и др. АСУ общегосударственных органов управления, разработка которых была поручена соответствующим министерствам и ведомствам. Предлагалось даже включить в эту структуру Государственную автоматизированную систему научно-технической информации (ГАСНТИ). Отраслевые министерства включены на уровне общегосударственных органов управления. Аналогично решалась проблема на уровне республик, областей и др. территориальных образований.

Между АСУ одного уровня предусматривались горизонтальные связи. Во главе ОГАС был поставлен Совет Министров СССР (СМ СССР).

Модель структуры ОГАС приведены на рис. 4.18.

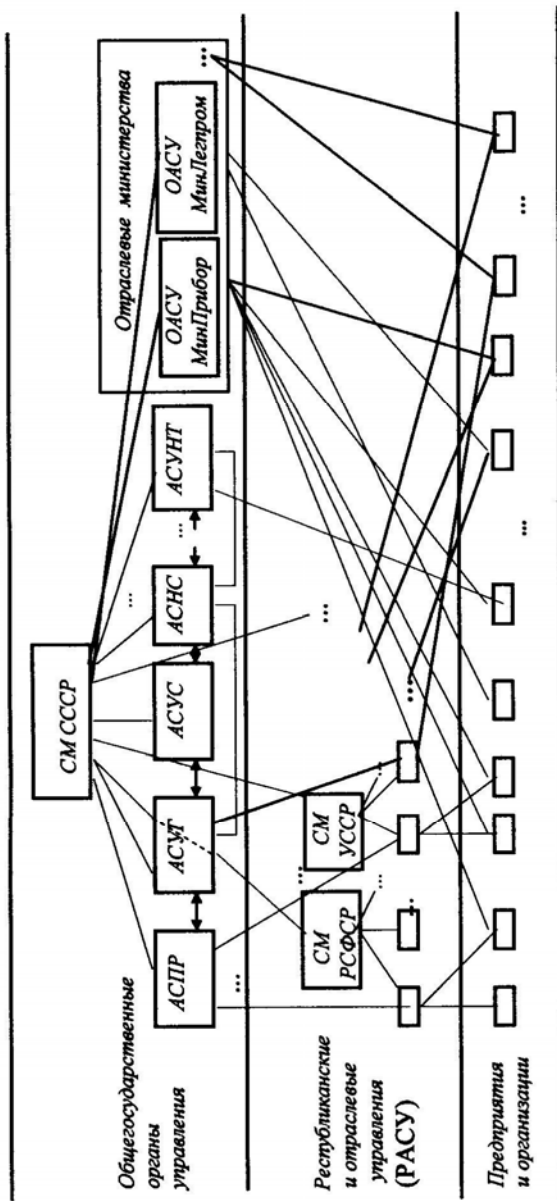


Рис. 4.18. Модель структуры ОГАС

а рис. 4.18 не показан уровень областей и других территориальных объединений, который формируется по принципу республиканских органов управления, но только вместо министерств, в областных органах управления – соответствующие отделы или управления (в настоящее время обычно называемые департаментами). На рис. 4.18 также показан принцип двойного подчинения АСУ предприятий и организаций, которые в соответствии с территориально-отраслевым принципом управления, принятым в стране и в концепции ОНАС, должны были представлять информацию и в ОАСУ, и в ТАСУ,

В таком варианте реализация ОГАС начала создаваться в середине 1960-х гг. и велась до 1990-х гг., когда изменилась экономическая и политическая ситуация в стране.

Структуру, приведенную на рис. 4.18, можно квалифицировать как стратифицированную структуру с вертикальными и горизонтальными связями. Идею диспетчеризации, предложенную В.М. Глушковым, реализовать не удалось. Да и вряд ли это возможно.

Могут быть реализованы принципы эшелонированного управления *М. Месаровича*, рассмотренные, например в [1, 13]. В соответствующих принципах функционирования структуры типа «эшелоны» могут быть установлены информационные и координационные связи (с различной степенью координируемости), которые должны быть регламентированы в соответствующих нормативно-методических документах. Но в тот период работы *М. Месаровича* еще не были переведены и широко известны в нашей стране, и соответствующие документы разработаны не были.

Вопросы для самоконтроля

1. История разработки АИС и АСУ.
2. Принципы разработки АИС и АСУ.
3. Функциональная и обеспечивающая части АИС и АСУ.
4. Виды структур функциональной части АИС и АСУ.
5. Принципы разработки структуры функциональной части АИС.

6. Методика разработки структуры обеспечивающей части АИС на основе постепенной формализации модели принятия решений.
7. Методика выбора готовых программных продуктов, реализующих отдельные функции ИС организации.
8. Методические материалы, регламентирующие разработку АСУ.
9. Понятие о интегрированной АСУ.
- 10 История возникновения идеи ОГАС.
- 11 Структура ОГАС.

Лабораторный практикум

После изучения материала данной главы предусмотрены лабораторные работы:

- Разработка структуры функциональной части АИС (АСУ) с использованием автоматизированных диалоговых процедур формирования и оценки иерархических структур.
- Разработка методики выбора информационных массивов для реализации функций системы управления предприятием (организацией).
- Разработка методики выбора для ИС предприятия готовых программных продуктов специального назначения.

Лабораторные работы выполняются на примере конкретной организации (по выбору студента) и включаются в курсовой проект (курсовую работу).

При выполнении лабораторных работ применяются автоматизированные диалоговые процедуры анализа целей и функций (АДПАЦФ) и реализации моделей организации сложных экспертиз, принципы работы и информационные технологии выполнения которых изложены в дополнительных учебных пособиях¹⁰³.

¹⁰³ **Автоматизированные** диалоговые процедуры анализа целей и функций систем управления: учеб. пособие / Под ред. В. Н. Волковой. – СПб.: Изд-во Политехн. ун-та, 2010. – 72 с.; **Волкова В. Н.** Методы организации сложных экспертиз: учеб. пособие / В. Н. Волкова, А. А. Денисов. – СПб.: Изд-во Политехн. ун-та, 2010. – 128 с.

Глава 5. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПОСТРОЕНИЯ И ПРОЕКТИРОВАНИЯ ДОКУМЕНТАЛЬНЫХ ИНФОРМАЦИОННО-ПОИСКОВЫХ СИСТЕМ НАУЧНО-ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ

5.1. Виды научно-технической информации и структура Государственной системы НТИ

Научно-техническая информация может быть представлена в *документальной* и *фактографической* форме. Но в любой форме она может быть *семантической* (содержательной и понятной для широкого круга пользователей) или *прагматической* (содержательной и существенной для принятия решений в конкретных условиях).

При этом производственная документация, оперативная информация о состоянии производства или рынка и другие виды фактографической информации, используемые в экономике для принятия текущих решений, не относятся к научно-технической информации.

Научно-техническую информацию классифицируют по разным признакам:

- по способу восприятия (визуальная, звуковая, тактильная (осязаемая);
- по форме представления (текстовая, графическая (изобразительная);
- по типу, по назначению; по способу распространения;
- по степени аналитико-синтетической обработки.

Основные виды научно-технической информации приведены в табл. 5.1. [9].

Для сбора, хранения и предоставления научно-технической информации создаются соответствующие хранилища и справочно-информационные фонды.

Таблица 5.1.

Виды научно-технической информации

Основание деления	Наименование признака	Примечание
1. Назначение информации	Массовая	Предназначенная для всех, а не только для узких специалистов
	Специальная	Предназначенная только для специалистов в конкретной области
2. Тип информации	Документальная	Фиксированная в научных документах
	Фактографическая (точнее концептуально-фактографическая)	Идеи и факты, извлеченные из научных документов
3. Способ распространения информации	Опубликованная	Широко распространяемая посредством тиражного размножения научных документов, прошедших официальную регистрацию
	Неопубликованная (и не публикуемая)	Не рассчитанная на широкое распространение и/или не прошедшая официальной апробации
4. Степень аналитико-синтетической переработки	Первичная	Непосредственные результаты научно-исследовательской и опытно-конструкторской работы (монографии, статьи, отчеты)
	Вторичная	Результаты аналитико-синтетической переработки первичной информации, сведения о документах, содержащих научно-техническую информацию
5. Области получения и/или использования информации в отраслях народного хозяйства и других областях человеческой деятельности	Название отрасли науки или области применения	Астрономическая, биологическая, географическая, идеологическая, искусствоведческая, историческая, медицинская, промышленная, сельскохозяйственная, строительная, физическая, химическая, экологическая и т. п.

За рубежом первоначально формировались *центры анализа информации*, постепенно укрупняемые путем объединения родственных центров, а в нашей стране создавались сразу *интегральные комплексы*, ставящие своей задачей преодоление дублирования работ по индексированию, хранению и обработке НТИ и объединяемые Государственной системой НТИ.

Центры анализа информации

По мере осознания роли информации как важнейшего ресурса развития общества и исследования ее основных особенностей, в том числе таких, как тиражируемость и многократность использования, старение (и нередко связанное с этим «загрязнение» (*pollution*) информационного пространства, неаддитивность, некоммуникативность, кумулятивность, зависимость реализуемости и эффективности от степени использования информации и т.п., все больше осознавалась необходимость работ по систематизации, оценке и обобщению информации. Поэтому в 1960-е гг. стали создаваться центры анализа информации.

Термин «*Informations Analysis Center*» (IAC) был предложен *Дж. Симпсоном* (США) в 1964 г. [10]. В 1967 г. при консультативном Комитете по научно-технической информации президента США была создана Секция по центрам анализа информации (*Panel of IAC*). Эквивалентом этому термину в немецком языке служит *Wissensbewertungsstelle* («центр по оценке знаний»), а во французском *centre d'analyse de l'information*.

В практике создания таких центров сложилось 2 направления:

- центры, занимающиеся сбором, индексированием, хранением, поиском и распространением информации;
- центры, занимающиеся содержательной оценкой и интерпретацией научной информации.

К числу центров первого типа относят: научно-технические библиотеки; издательства, торговые ассоциации или профессиональные организации, занимающиеся изданием и распространением научных книг и журналов; информационные центры, занимающиеся подготовкой библиографических и реферативных изданий; центры, занимающиеся сбором, хранением и распро-

странением документальной информации по собственной инициативе или запросам потребителей.

Центры второго типа являются по сути научными учреждениями и организациями, занимающимися оценкой и обобщением информации. Такие центры вначале создавались по отдельным отраслям фундаментальной науки (биология, физика, химия, науки об окружающей среде) или даже научным направлениям.

Вначале центры создавались независимо, часто дублируя работу друг друга. Например, в США к середине 1960-х гг. было создано более 100 центров разного рода, причем иногда по несколько в одной и той же области знаний (см. табл. в [9, с. 329]).

Затем стала наблюдаться тенденция к их объединению. Так:

- * в 1965 г. министерство здравоохранения, образования и социальной политики США объединило 18 ИАС в Информационный центр по ресурсам образования (*Educational Resources Informational Center – ERIC*);

- * в 1966 г. был создан Комитет по научно-техническим данным (*Comitee on Data for Science and Technology – CODATA*), в который вошли представители международных союзов и ряда ведущих в тот период стран – СССР, США, Великобритании, ФРГ, Канады, Японии и др.;

- * в 1970 г. три крупных информационных центра США (*BioScience Information Service of Biological Abstracts – BIOSIS, Chemical Abstracts Service – CAS, Engineering Index Inc. – Ei*) объединились для изучения проблем дублирования обработки одних и тех же публикаций и возможностей устранения или хотя бы сокращения такого дублирования путем кооперации и обмена информационными массивами на машинных носителях.

Примерами такого рода центров в нашей стране являются:

- * созданный в 1952 г. Всесоюзный институт научной и технической информации (ВИНИТИ);

- * созданная в 1965 г. Государственная система стандартных справочных данных (ГСССД), которая должна была обеспечивать не только сбор, обработку и предоставление информации о свойствах веществ и материалов, но и качественную переработку информации, оценку ее достоверности; при этом ГСССД сразу создавалась как разветвленная многоуровневая система с децентрализованным получением и распространением научной информации;

- * центральные научно-исследовательские институты технико-экономической информации (ЦНИИТЭИ), которые создавались для ведущих отраслей промышленности (ЦНИИТЭИ Приборостроения, ЦНИИТЭИ Автопрома и т. п.).

Интегральные информационные системы

Для того, чтобы преодолеть проблемы дублирования работ по индексированию, хранению и обработке НТИ, «загрязнения» информационного пространства была предложена [9] идея создания *интегральных информационных систем*.

Интегральная информационная система (ИИС) – это совокупность методов и средств, позволяющих при *однократном* описании, индексировании и реферировании научных документов обеспечить *многоаспектную* обработку содержащейся в них информации и *многократное* ее использование для удовлетворения разнообразных информационных потребностей [10, ~~Принцип~~]. Принцип одноразового ввода и многократного использования информации, разумеется, реализовывался и в ряде зарубежных центров анализа информации. Но они, как правило, создавались для отдельных отраслей: по физике (в Американском физическом институте и в Институте инженеров-электриков в Великобритании), по ядерной физике и технике (Международное агентство по атомной энергии, Центр информации и документации Евратома), по медицине, химии и химической технологии и т. д.

В СССР замысел ИИС подразумевал не просто создание многоаспектных информационных банков или баз данных, а получение эффекта *целостности*, т. е. новых *эмерджентных* свойств, которых не было у элементов, объединяемых в ИИС.

Получение эффекта целостности базировалось на создании условий для совместной работы ученых и специалистов – как тех, кто обеспечивает подготовку сигнальной информации и реферирование (осуществляемое высококвалифицированными учеными и специалистами, работающими, как правило, на внештатной основе), так и пользователей, которые влияют на возникновение новых системных свойств посредством обратной связи в режиме ИРИ.

Для того, чтобы реализовать этот замысел, была предложена принципиальная схема функционирования ИИС (рис. 5.1), и с 1971 г. в ВИНТИ начала разрабатываться интегральная информационная система АССИСТЕНТ (Автоматизированная Справочно-

информационная СИСТЕма по Науке и Технике), в функции которой входило оказание помощи ученым и специалистам в решении не только научно-исследовательских, но и производственных, организационных, управленческих задач.

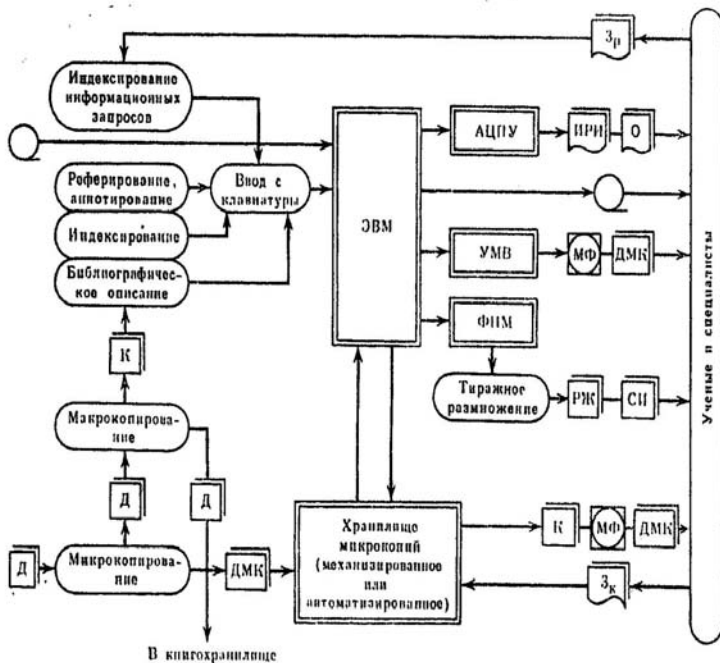


Рис. 5.1. Принципиальная схема функционирования интегральной ИС

Д – документы; ДМК – диамикрокарты; З_р – разовый информационный запрос; З_к – запрос копии документа; ИРИ – избирательное распределение информации; О – ответы на разовые информационные запросы; К – макрокопии документов; МФ – рулонные микрофильмы; РЖ – реферативные журналы; СИ – бюллетени сигнальной информации; УМВ – устройство микрофильмирования; ФНБ – фотонаборная машина

С основными задачами и особенностями этой системы можно познакомиться в [9, с. 361–377], что может быть полезным для развития современных систем информационного обеспечения научных исследований.

Государственная система научно-технической информации

Для более полной реализации замысла целостной интегральной системы в 1970-е гг. в СССР было начато создание *Государственной системы научно-технической информации* (ГСНТИ) и поставлена задача разработки *Единой автоматизированной системы научно-технической информации*, в последующем переименованной в Государственную – ГАСНТИ. Руководил созданием ГСНТИ и ГАСНТИ Государственный комитет Совета Министров СССР по науке и технике.

В настоящее время в стране сохранилась общегосударственная система НТИ, которую условно называют Российской ГСНТИ (РГСНТИ), структура которой приведена на рис. 5.2.

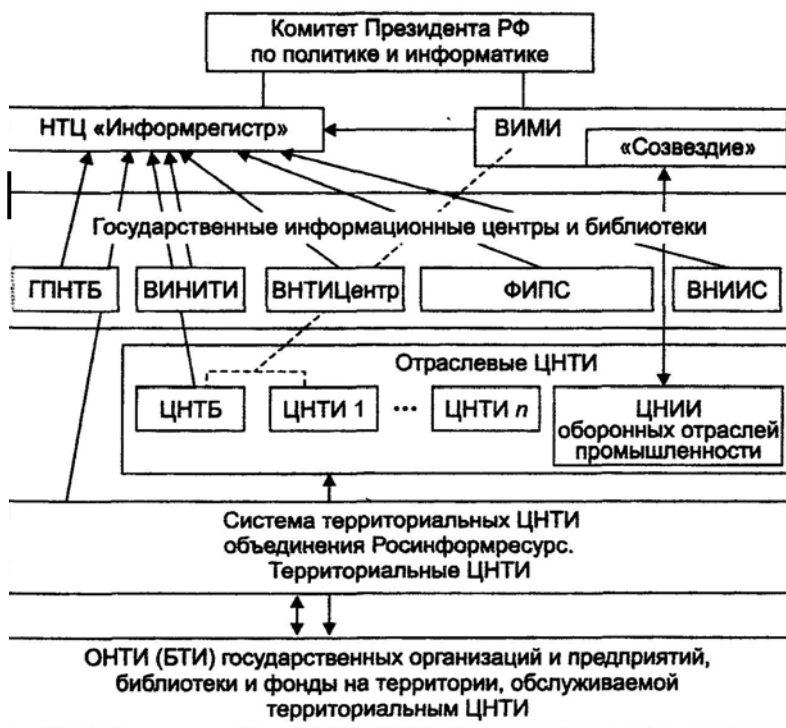


Рис. 5.2. Структура РГСНТИ

Структура РГСНТИ включает [11, 19]:

1. Государственные информационные центры и библиотеки, основными из которых являются:

- Государственная публичная научно-техническая библиотека России (ГПНТБ России) – головной орган научно-технических библиотек, фонды которой составляют отечественные и зарубежные научные издания, неопубликованные переводы, авторефераты диссертаций и др. источники НТИ; Всероссийский институт научной и технической информации (ВИНИТИ), фонды которого содержат отечественные и зарубежные периодические издания, сборники трудов, реферативные журналы и ряд др.;

- Всероссийский научно-технический информационный центр (ВНТИ-Центр), формирующий и сохраняющий документы государственной регистрации НИОКР, отчеты по НИР и ОКР, диссертации, фонды алгоритмов и программ;

- Федеральный институт промышленной собственности (ФИПС) и Информационно-издательский центр Российского агентства по патентам и товарным знакам, накапливающий фонды отечественной и зарубежной патентной литературы, товарные знаки (бывшее НПО «РОСПАТЕНТ»);

- Всероссийский научно-исследовательский институт стандартизации (ВНИИС) и Федеральный фонд государственных стандартов, общероссийских классификаторов технико-экономической информации, международных, региональных стандартов, правил, норм и рекомендаций по стандартизации, национальных стандартов зарубежных стран Федерального агентства по техническому регулированию и метрологии (в прошлом – ВНИИ классификации, терминологии и информации по стандартизации и качеству – ВНИИКИ, разрабатывающий и собирающий стандарты и др. нормативно-техническую документацию);

- НТЦ «Информрегистр» Комитета при Президенте РФ по политике информатизации, занимающийся регистрацией и предоставлением сведений об имеющихся в стране информационных ресурсах на машинных носителях, в том числе библиографических и справочных БД.

2. Система территориальных центров НТИ (ЦНТИ) объединения Росинформресурс Миннауки РФ, включающая 69 центров.

Каждый из территориальных центров НТИ (ЦНТИ) включает научно-технические библиотеки, располагает уникальными фондами опубликованных и неопубликованных информационных изданий, отражающих сведения о научных и технических разработках, материальных и сырьевых ресурсах

своих регионов, об организациях и предприятиях региона, его экономике и т. п. ЦНТИ взаимодействует с отделами НТИ (ОНТИ) государственных организаций и предприятий, частных фирм на территории обслуживаемых регионов, имеющих свои библиотеки и фонды документов, отражающие производственную, коммерческую и иную информацию (всего более 100 тыс. организаций).

3. Система НТИ министерств и ведомств России.

Центральные научно-технические библиотеки (ЦНТБ) и отраслевые ЦНТИ оборонных отраслей промышленности, объединяемые автоматизированной информационной системой «Созвездие» с центральным звеном во Всероссийском научном институте межотраслевой информации (ВИМИ) при Комитете Президента РФ по политике информатизации.

ВИМИ выполняет функции методического и координационного центра для ведущих отраслевых ЦНТИ. Отраслевые ЦНТИ взаимодействуют с информационными подразделениями предприятий и организаций своей отрасли и создают справочно-информационные фонды, содержащие техническую, отраслевую нормативную, технологическую и др. виды документации.

4. Система территориальных ЦНТИ объединения Росинформресурс Миннауки РФ 69 территориальных ЦНТИ.

5. Отделы НТИ (ОНТИ) и бюро технической информации (БТИ) государственных организаций и предприятий, библиотеки и фонды на территории, обслуживаемой территориальным ЦНТИ (более 100 тыс.).

Особое место в системе ГСНТИ занимают Российская книжная палата, регистрирующая и хранящая информацию об опубликованных в РФ произведениях печати и государственной библиографии, и Российский государственный архив научно-технической документации Федеральной архивной службы России, содержащий информацию о документах научно-исследовательских, проектных, конструкторских, технологических организаций и предприятий федерального значения, передаваемых на постоянное хранение. Эти общегосударственные хранилища являются относительно независимыми в общей структуре ГСНТИ.

С начала 90-х гг. XX в. в компьютерных сетях началось предоставление возможности широкой аудитории пользователей знакомиться с электронными вариантами выпускаемых в стране компьютерных изданий.

5.2. Информационные потребности и информационное обслуживание

При работе с информацией (особенно с документальной) важное значение приобретает проблема определения потребителя (исследователя) и удовлетворения его потребностей.

Информационной потребностью (ИП) называют осознанную необходимость в знаниях или, другими словами, осознание недостаточности имеющихся знаний.

При выявлении информационных потребностей следует учитывать: *тип информации* (тематика, виды документов; *качество* и *количество* информации (достоверность, полнота, избыточность, надежность информационных источников); *способы представления* информации потребителю (устно, на экране, оригинал, копия, информационное издание, микрофильм и т. п.); *сроки представления* с учетом минимизации разрыва между появлением, опубликованием информации и доведением ее до потребителя.

Для выявления и обеспечения информационных потребностей необходима система *информационного обслуживания* с соответствующей организацией информационных потоков.

Информационные потребности зависят от различных объективных и субъективных факторов. Различают *субъективные* ИП, которые определяются мнением и требованиями потребителя, его представлениями, знаниями, умением сформулировать запрос и потребности в объективно необходимой информации; и потребности в *объективно необходимой* информации, которая должна быть связана с задачей, проблемой, которую исследует потребитель в соответствии со своими служебными обязанностями, в частности, в системе управления предприятием (организацией).

При выполнении профессиональных обязанностей возникают потребности в текущей и ретроспективной информации, узко- и широкотематической информации, отраслевой, межотраслевой, правовой, нормативной и др. видов информации.

Для более эффективной организации информационного обслуживания необходимо учитывать принципиальную особен-

ность – субъективизм и неадекватность выражения информационных потребностей. Иными словами, информационная потребность субъекта определяется степенью понимания стоящих перед ним задач и знанием возможностей систем информационного обслуживания.

Изучение запросов потребителей информации является первоочередной задачей научно-информационной деятельности. Задача эта является весьма сложной, поскольку потребители, как правило, не умеют четко сформулировать свои потребности, и они уточняются по мере работы с получаемой информацией.

Существующие методы изучения информационных потребностей делят на две группы:

- *косвенные* или *документальные*, базирующиеся на анализе документальных источников (карт обратной связи, запросов потребителей информации и т. п., независимо от того, запросили ли специалисты);

- *прямые*, связанные с участием потребителя информации, изучением задачи, проблемы, объекта, конкретной ситуации, которые позволяют выявить потребность в объективно-необходимой информации конкретных потребителей.

Прямые методы обычно реализуются с помощью анкетирования, интервьюирования и других экспертных методов. С их помощью изучают не столько информационные потребности, сколько представление субъекта об этих потребностях. Использование прямых методов позволяет не только обеспечить информационную поддержку членов коллектива, но и определить, какая именно информация нужна каждому для исполнения своих обязанностей.

В то же время следует иметь в виду, что решения, принимаемые на основе экспертных оценок, основаны на субъективных мнениях экспертов и далеко не всегда позволяют определить истинные информационные потребности.

Применение косвенных методов связано с изучением предметной области и функционально-должностных обязанностей

специалистов. Здесь широко используются методы анализа систем, базирующихся на анализе структуры целей и функций деятельности потребителя информации, а также методы математического моделирования.

В результате формируется информационная модель, отражающая потенциальные информационные потребности и позволяющая организовать информационное обслуживание в процессе управления предметной областью.

Проблема изучения информационных потребностей и запросов потребителей информации может решаться на основе статистических исследований информационных потоков, изучения их закономерностей (см. гл. 7), таких как экспериментально открытый закон рангового распределения *Дж. Ципфа*, теоретически уточняющий его закон *Б. Мандельброта*, закономерности упорядочения параметров документальных информационных потоков *С. Брэдфорда*, *Б. Викери*, закономерность концентрации-рассеяния, исследованная применительно к проблемам определения информационных потребностей *В. И. Горьковой* и т.п.

Перспективным представляется определение запросов потребителей на основе структуризации целей и функций деятельности руководителей (при обслуживании в режиме ДОР) или специалистов (обслуживаемых в режиме ИРИ).

Для решения проблемы обеспечения информационных потребностей в теории научно-информационного поиска предложены различные виды информационного обслуживания: *регламентное обслуживание по стандартным запросам (СЗ)*, *избирательное распределение информации (ИРИ)*, *дифференцированное обслуживание руководителей (ДОР)*, *ретроспективный поиск (РП) по произвольным запросам*.

Доступ к научно-технической информации – вид *информационных услуг, информационного обслуживания*, существовавшего еще до появления электронной вычислительной техники.

Механизация и автоматизация доступа к НТИ началась в 1970-е гг. При этом стали различать библиотечную информацию, ориентированную на широкий круг читателей, и специальную

научно-техническую информацию, которую собирают, формируют справочно-информационные фонды и предоставляют специалистам особые, отделенные от библиотек, подразделения, создаваемые на предприятиях и в организациях – отделы научно-технической информации (ОНТИ) или бюро технической информации (БТИ).

В настоящее время информационное обслуживание (сервис) представляет собой область профессиональной деятельности, обеспечивающей поиск, обработку, оформление и распространение информации.

В процессе развития системы информационных коммуникаций сформировались три вида информационного обслуживания: *документальное, фактографическое и концептографическое.*

Документальное обслуживание общества и различных его институтов, в том числе науки, техники и экономики, традиционно обеспечивалось научно-технической информацией. Сущность его заключается в том, что информационные потребности удовлетворяются путем предоставления документов, сведения из которых потребитель извлекает самостоятельно. Развитие этого вида обслуживания сделало доступным массовому потребителю не только печатные и машиночитаемые документы, но и базы данных различных видов и наименований.

Дальнейшее совершенствование ИТ позволило перейти к *фактографическому* информационному обслуживанию. Его основу составляет предоставление потребителю сведений (данных, фактов) в соответствии с его запросом. Это более сложная форма организации обслуживания. Для формирования и реализации запроса требуется знать характеристики объекта поиска, связи между ними, т. е. необходимо множество сведений, позволяющих сформулировать запрос и понять его.

Различают два вида фактографического обслуживания: информационно-поисковое и информационно-решающее.

В первом случае выполняется лишь поиск информации по заданным характеристикам; во втором – обеспечивается предос-

тавление информации об объекте, получаемой в результате работы алгоритмов преобразования данных.

Необходимость интерпретации, анализа, реферативной обработки полученных сведений (данных), наличие огромных объемов информации привели к возникновению и развитию *концептографического* информационного обслуживания. Оно предполагает предоставление потребителю так называемой ситуативной информации: интерпретации данных, оценки, рекомендации, прогнозы, реферативные обзоры, рецензии, переводы с иностранных языков и т. д. Другими словами, осуществляется предоставление информации, которой в явном виде в документах и базах данных может и не быть. Такая информация может быть получена на основе аналитико-синтетической обработки научно-технической информации или на основе анализа и обработки (на основе моделей принятия решений, моделей прогнозирования и т. п.) фактографической информации. Концептографическое обслуживание можно рассматривать как форму интерпретированного документального и фактографического обслуживания, или как разновидность научной деятельности в области информационного анализа.

Существуют различные *виды* документального информационного обслуживания. Основными можно считать следующие:

- *регламентное обслуживание* в форме подготовки и издания обзорных и реферативных информаций, библиографических указателей и т. п., или обслуживание по стандартным запросам;
- оповещение отдельных специалистов (абонентов) о текущих публикациях, представляющих для них потенциальный интерес путем *избирательного* (адресного) *распределение информации (ИРИ)* по «профилям интересов», сформулированных потребителями, или *дифференцированного обслуживания руководителей (ДОР)* по мере поступления информации, необходимой для принятия управленческих решений;
- *ретроспективный поиск*, т. е. отыскание документов, в которых находятся сведения по конкретным запросам, в массиве

всех накопленных источников информации, что может осуществляться либо в форме подготовки тематических подборок, либо в оперативном режиме.

Наряду с основными предлагались более полные классификации видов обслуживания. В частности, в [9, с. 247] приводится обобщенная классификация по нескольким признакам (табл. 5.2).

Таблица 5.2

Виды информационного обслуживания

Признак классификации	Вид обслуживания
<i>По источнику инициативы</i>	<ul style="list-style-type: none"> • принудительное • по запросам потребителей
<i>По типам документов</i>	<ul style="list-style-type: none"> • обслуживание опубликованными документами и/или их копиями • обслуживание копиями неопубликованных документов (отчетов по НИР, диссертаций, переводов и т. п.)
<i>По направленности или адресности</i>	<ul style="list-style-type: none"> • информационные издания (много адресов) • избирательное распределение информации (один адрес)
<i>По периодичности или срочности</i>	<ul style="list-style-type: none"> • текущее оповещение • ретроспективный поиск
<i>По способу доведения документов до потребителя</i>	<ul style="list-style-type: none"> • непосредственная передача документов или их копий потребителям • двухступенчатое обслуживание (вначале информационное издание-сигнал, а затем – копии заинтересовавших потребителя документов)
<i>По назначению, цели</i>	<ul style="list-style-type: none"> • ознакомление • для практического использования

5.3. Понятие о документальном информационном поиске

Теория информационного поиска первоначально сформировалась при исследовании документальных информационно-поисковых систем (ДИПС) научно-технической информации.

Под *информационным поиском* в таких системах понимается некоторая последовательность операций, выполняемых с целью отыскания документов (статей, научно-технических отчетов, описаний к авторским свидетельствам и патентам, книг и т. д.), содержащих определенную информацию (с последующей выдачей самих документов или их копий), или с целью выдачи фактических данных, представляющих собой ответы на заданные вопросы [9, с. 248].

Массив элементов информации, в котором производится информационный поиск, был условно назван *поисковым массивом* [9].

Процесс поиска документов может быть формализованно описан в терминах теории множеств следующим образом: D – некоторое множество документов или библиотека (поисковый массив); Q – множество информационных запросов; R – отношение, свойство, при наличии которого любому $q \in Q$ ставится в соответствие подмножество $D' \subset D$, называемое ответом на информационный запрос.

Очевидно, что прочитать каждый документ информационной базы, чтобы найти необходимый, практически невозможно. Поэтому на протяжении истории развития информационного поиска разрабатывались и совершенствовались различные методы поиска.

Каждому документу, вводимому в поисковый массив, ставится в соответствие *поисковый образ документа* (ПОД), который представляет собой характеристику, отражающую основное смысловое содержание документа (этим ПОД отличается от кода, присваиваемого информационному элементу в фактографических информационных системах).

В виде такой же краткой характеристики – *поискового предписания* или *поискового образа запроса* (ПОЗ) должен быть сформулирован и информационный запрос. Благодаря этому процедура поиска может быть сведена к простому сопоставлению поисковых образов документов с поисковым предписанием.

Однако такое сопоставление допустимо лишь в случае, если ПОД и ПОЗ описаны в терминах единого языка.

Для реализации процедуры описания и документов и запросов с помощью ПОД и ПОЗ разрабатывают *информационно-поисковые языки* (ИПЯ), которые имеют различные возможности.

Простейшим ПОД может являться заглавие документа, переписанное в форме перечня входящих в него слов. При этом необходимо установить некоторые правила типа включения в ПОД существительных в единственном числе и именительном падеже, глаголов – в неопределенной форме или в виде отглагольных существительных и т. п. правил, общих для формирования ПОД и ПОЗ.

В более развитых вариантах для уточнения содержания документа можно в качестве ПОД использовать аннотацию или реферат документа, также переписанные в виде перечня входящих в них слов с учетом некоторых правил, общих для формирования ПОД и ПОЗ.

В практических условиях реализации ИПС число слов, включаемых в ПОД и ПОЗ, не должно быть слишком большим, поэтому вводится понятие *ключевые слова*, которые являются наиболее значимыми для отображения содержания документа.

Некоторые идеологи информационного поиска (например, Ч. Мидоу [8]) отождествляли понятия *ключевое слово* и *дескриптор*. Однако в дальнейшем эти понятия стали использоваться в разном смысле.

Для отображения ПОД и ПОЗ разрабатываются *информационно-поисковые языки* (ИПЯ).

Для сопоставления ПОД и ПОЗ применяют различные *критерии поиска* или *критерии смыслового соответствия* (КСС).

Точность поиска определяется точностью отображения документов и запросов с помощью их поисковых образов и степенью совпадения ПОД и ПОЗ. Поэтому возможна неполнота вы-

дачи документов, либо, напротив, «поисковый шум», которые представляют собой своеобразную плату за облегчение процедуры информационного поиска.

Для оценки качества поиска вводят специальные критерии. Обобщенные критерии качества поиска – *релевантность* и *пертинентность* [9, с. 282–287].

Под *релевантностью* понимается соответствие выдачи запросу, т. е. релевантность характеризует качество алгоритма поиска.

Под *пертинентностью* понимается соответствие выдачи потребностям лица (или лиц), для которого (которых) осуществляется поиск информации, т. е. пертинентность характеризует смысловыражающие возможности ИПЯ, точность отображения с его помощью информационных потребностей.

Предлагаются и используются различные конкретные критерии количественной оценки релевантности и пертинентности. Разработка комплекса критериев оценки качества информационного поиска – достаточно сложная проблема: состав и количественные характеристики критериев зависят от конкретного назначения и принципов реализации ИПС.

Информационный поиск осуществляется с помощью информационно-поисковой системы, понятие о которой развивалось. Кратко рассмотрим представления об ИПС и ее структуре.

5.4. Информационно-поисковая система и структура ИПС

Структура и функционирование конкретной ИПС зависят от вида и состава информационных источников, от способов реализации информационного поиска. В то же время есть некоторые общие принципы построения и функционирования ИПС, которые кратко рассматриваются в данной главе.

Анализ определений ИПС. Под *информационно-поисковой системой* первоначально понималась некоторая совокупность

или комплекс связанных друг с другом отдельных частей, предназначенных для выявления в каком-либо множестве элементов информации (документов, сведений и т.д.), которые отвечают на информационный запрос, предъявляемый к системе [9].

С учетом приведенного выше описания процесса информационного поиска ИПС (*IPS*) можно определить следующим образом:

$$IPS \underset{def}{\equiv} \langle D, Q, R, D' \rangle, \quad (5.1)$$

$D' \subset D$

где D – некоторое множество документов или библиотека (поисковый массив); Q – множество информационных запросов; D' – ответ на информационный запрос; R – множество отношений, свойств, при наличии которых любому запросу $q_i \in Q$ ставится в соответствие подмножество D' .

Более полно **А. И. Черный** предложил представить ИПС в виде совокупности четырех основных компонентов [17, с. 18]:

$$IPS \underset{def}{\equiv} \langle LS, D, TS, N \rangle, \quad (5.2)$$

где LS – *логико-семантический аппарат* (включающий информационно-поисковые языки – один или более, правила индексирования и критерии выдачи); D – *поисковый массив* (т. е. определенное множество снабженных поисковыми образами документов, в котором отыскиваются необходимые); TS – *технические средства* (т. е. какие-то приспособления или устройства, которые необходимы для записи и хранения поисковых образов, для хранения документов и осуществления процесса сопоставления поисковых образов документов с поисковым предписанием или поисковым образом запроса); N – *люди, взаимодействующие с системой* (т. е. те, кто пользуются данной ИПС и обслуживают ее – осуществляют индексирование документов и информационных запросов, выбирают стратегию поиска, а также выполняют другие интеллектуальные операции, без которых невозможен информационный поиск).

Затем я обеспечения возможности автоматизации процедуры информационного поиска было предложено [9, 17] в ИПС выделить два уровня рассмотрения – *абстрактный* и *конкретный*.

Абстрактной ИПС была названа совокупность ИПЯ (*retrieval language – RL*), правил индексирования (*IND*) и критерия выдачи или критерия смыслового соответствия (*KSS*):

$$IPS \underset{def}{=} \langle RL, IND, KSS \rangle. \quad (5.3)$$

Конкретной ИПС названа практически реализованная система, включающая массив документов *D*, в котором производится информационный поиск, технические средства *TS* реализации ИПС, а также взаимодействующих с ней людей *N*.

Структура функционирования ИПС в таком понимании приведена на рис. 5.3.

В соответствии с рассмотренным выделением в ИПС *абстрактного* и *конкретного* уровней и с учетом особенностей хранения документальной информации (библиотеки, архивы и т.п. хранилища) процедуру информационного поиска документальной информации было предложено разделить на два контура [17]:

1) семантическое осмысление запроса и выдача адресов (шифров, кодов), соответствующих запросу документов; на рис. 5.3 этот контур показан сплошными линиями;

2) отыскание самих документов (вручную или с помощью специализированных технических средств, если ими оборудовано хранилище); на рис. 5.3 – штриховые линии.

Второй контур связан с разработкой специализированных технических средств хранения больших массивов документов и работой по переоборудованию хранилищ, а собственно проблемы информационного поиска решаются в *первом контуре*.

С учетом вышерассмотренного первый контур ИПС представляет собой ее логико-семантический аппарат и состоит из трех основных блоков (рис. 5.4):

- *информационно-поискового языка,*

- системы перевода (индексирования) на этот язык и
- логики, обеспечивающей поиск, которые, в свою очередь, могут быть детализированы и реализованы разными способами.

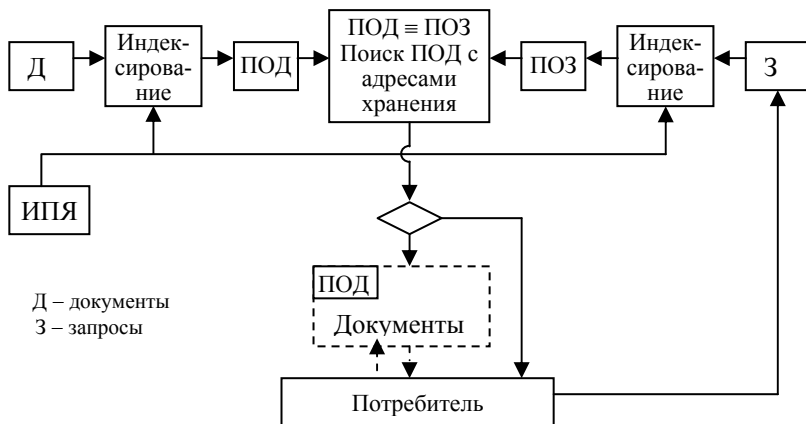


Рис. 5.3. Структура функционирования ИПС



Рис. 5.4. Структура логико-семантического аппарата ИПС

Представление ИПС в виде двух контуров является в настоящее время наиболее распространенным.

В некоторых системах контуры могут быть совмещены.

Напротив, иногда возникает необходимость выделять не два, а большее число контуров, что помогает организовать последовательно углубленный анализ текстов документов.

Такие варианты реализуются, например, в документально-фактографических системах нормативно-правовых и нормативно-методических документов.

В символической форме, принятой выше, *абстрактная ИПС* (1-й контур) представляет собой совокупность ИПЯ (*RL*), правил (системы) индексирования (*IND*) и логики (*LOG*), включающей наряду с критериями смыслового соответствия базисные отношения:

$$IPS \underset{def}{=} \langle RL, IND, LOG \rangle. \quad (5.4)$$

В теории и практике различают ИПС разных видов.

Документальные ИПС (ДИПС) в ответ на вводимые в них информационные запросы выдают оригиналы, копии или адреса хранения документов, содержащих требуемую информацию.

Фактографические ИПС (ФИПС) предназначены для выдачи непосредственно требуемой информации (например, температуры кипения какой-либо жидкости, статистических показателей, содержащихся в соответствующих отчетных документах и т.п.).

При этом существуют фактографические системы двух видов:

1) системы, в которых сразу формируются массивы фактографической информации, параллельно с документальными;

2) системы, в которых массивы фактографической информации формируются на основе массивов документальной информации.

Информационные системы второго вида могут, в свою очередь, формироваться как *документально-фактографические* (ДФИПС и АДФИПС), содержащие массивы двух видов:

- *документальные* и сопряженные с ними массивы *фактографической* информации;
- *информационно-логические* ИПС.

В отличие от документальных, фактографических и документально-фактографических ИПС 1-го вида, которые могут по запросам выдавать только информацию, введенную в них ранее, *информационно-логические системы* представляют собой информационные системы более высокого класса: они должны выдавать не только ранее введенную в них информацию, но и производить, если необходимо, логическую переработку этой информации с целью получения новой информации, которая в явном виде не вводилась в ИС.

Информационно-логическую систему (ИЛС или *ILS*) можно определить как совокупность ИПЯ (*RL*), правил перевода с естественного языка на информационный, т. е. правил индексирования (*IND*) и правил логического вывода (*LV*), которая предназначена для алгоритмического получения новой информации (I_n) [9, 17]:

$$IPS \underset{def}{\equiv} \langle RL, IND, LV, I_n \rangle. \quad (5.5)$$

Развивая представления об информационных системах, способных получать новую информацию, **Ю. И. Шемакин** предлагает понятие *информационно-семантической системы* [18, с. 60]:

$$ISS \underset{def}{\equiv} \langle a, St, tp_{iss}, co, t_i \rangle, \quad (5.6)$$

где a – цель; St – структура; $tp_{iss} \in TP$ – подмножество технологических процессов для данной *ISS*; co – условия; t_i – время.

Входящие в определение (5.6) составляющие могут быть детализированы с учетом конкретной реализации ИПС.

Особенно важно уточнять состав технологических процессов:

$$tp_{iss} \equiv \langle met, re, SemSI \rangle, \quad (5.7)$$

где met – методы; re – средства; $SemSI$ – семантическая переработка семантической информации.

Для организации проектирования информационных систем **Ю. Ф. Тельнов** [14] предлагает определение, в котором учитываются и цели (запросы) Q , и элементы (внешние E_n и внутренние E_i), и отношения R (включая динамические взаимодействия), и функции F (процессы, операции), и период времени T , и

закономерности Z , определяющие структуру системы и ее взаимодействие с внешней средой:

$$S \underset{def}{=} \langle Q, E_i, E_n, T, F, R, Z \rangle. \quad (5.8)$$

Выбор определения ИПС зависит от конкретного объекта, для которого она разрабатывается, назначения и условий разработки и функционирования информационной системы.

5.5. Информационно-поисковый язык

Понятие об информационно-поисковом языке

Применение естественного языка для отображения ПОД и ПОЗ связано со значительными трудностями, обусловленными наличием в языке синонимов, омонимов и т. п. неоднозначностей использования терминов естественного языка. Поэтому на определенном этапе развития теории и практики создания ИПС вместо естественного языка стали применять искусственные информационно-поисковые языки – ИПЯ.

Существуют различные названия и определения специализированного языка, с помощью которого отражают основное содержание документов, вводимых в ИПС.

Информационно-поисковый язык (retrieval language) – «это специализированный искусственный язык, предназначенный для выражения основного содержания документов или информационных запросов с целью отыскания документов в некотором их множестве» [9, с. 259].

Информационно-поисковый язык (ИПЯ) используется для отображения содержания документов *информационно-поисковой системы* в *поисковом образе документа* – ПОД, и запроса – в *поисковом образе запроса* – ПОЗ, или *поисковом предписании*.

Такой язык называли вначале *информационным языком (ИЯ)*, предъявляя к нему требование однозначной записи содержания документа; *языком индексирования (index language)*, определяемым как совокупность или система символов или ин-

дексных терминов и правил их использования для выражения предметного содержания документов; *документальным языком* (*language documentaire*) и т. п. (подробнее с обзором этих терминов можно познакомиться в [9]).

В окончательном варианте понятийного аппарата теории информационного поиска утвердился термин *информационно-поисковый язык* (*retrieval language*).

Обобщая различные представления об информационно-поисковом языке, можно дать следующее определение:

Информационно-поисковый язык (ИПЯ) является формализованной семантической системой, обеспечивающей передачу (запись) содержания документа в объеме, необходимом для целей поиска.

Документ, записанный на этом языке, может быть в принципе и не понят человеком, даже если в записи используются слова естественного языка, поскольку в ИПЯ употребление слов, выражений, отношений между ними стандартизировано определенным образом.

Задачей ИПЯ является перевод содержания документа в *поисковое предписание* или *поисковый образ документа* (при вводе документа в ИПС) и перевод содержания запроса пользователя в *поисковый образ запроса* (*поисковое предписание*).

Первые исследователи в качестве составляющих ИПЯ выделяли: *алфавит* (набор буквенных и цифровых символов); *слова*, формируемые из алфавита с помощью морфологических правил – *морфологии*; *словарь* перевода (в котором каждому слову или осмысленной конструкции естественного языка сопоставлено слово или словосочетание ИПЯ); правила, отражающие взаимоотношения между словами документа, которые в конкретных ИПЯ реализуются, например, с помощью *текстуальных* или *контекстуальных* отношений, или с помощью специальных правил грамматики – *синтаксис*.

Словарь может состоять из *ключевых слов* (*словосочетаний*) или *дескрипторов*. Вначале некоторые авторы (например,

Ч. Мидоу [8] отождествляли эти понятия и понимали под дескриптором все слова, выбранные для включения в словарь.

Однако в дальнейшем термину *дескриптор* стали придавать более сложный смысл: в отличие от *ключевых слов*, выбираемых предварительно из документов массива, для поиска в котором разрабатывается ИПЯ, под *дескриптором* понимается некоторый (выбранный разработчиком ИПЯ), обобщающий термин для отображения группы синонимов или слов, которые для целей поиска в конкретной ИПС можно считать синонимами.

Такие слова объединяют в класс *условной эквивалентности*, обобщаемый соответствующим *дескриптором*, и если в тексте документа или запроса встречается слово из данного класса, то его заменяют в ПОД или ПОЗ дескриптором.

Таким образом, *дескриптор* – специальное понятие, введенное и используемое в теории *информационного поиска* [8, 9, 17].

В современных информационно-поисковых языках под *дескриптором* понимают имя класса *условной эквивалентности* [9, 17].

Класс условной эквивалентности формируется из ключевых слов, связанных *парадигматическими отношениями*.

Парадигматические (базисные) отношения – один из видов семантических отношений, предложенных в теории *информационного поиска* и применяемых при разработке *информационно-поисковых языков*.

Парадигматические отношения представляют собой внетекстовые смысловые отношения между лексическими единицами ИПЯ, которые устанавливаются на основании потребностей информационного поиска.

Роль парадигматических отношений сводится к следующему. Принципиальной особенностью естественного языка является тот факт, что в нем одни и те же события могут быть описаны в разных терминах. Тогда в *поисковом образе документа* – ПОД, и *поисковом образе запроса* – ПОЗ, могут быть использованы разные слова с сохранением смысла документа и запроса.

Кроме того, на практике может оказаться необходимым отыскивать документы, в которых речь идет о более частных понятиях, чем в ПОЗ. Не потерять такие документы может помочь введение парадигматических (базисных) взаимоотношений между дескрипторами ИПЯ.

В широком смысле в состав парадигматических отношений включают отношения синонимии (тождество означаемых при различии означающих), омонимии (тождество означающих при различии означаемых), отношения, основанные на одинаковости основы при различных окончаниях (парадигмы склонения и спряжения).

Однако в более узком смысле при разработке ИПЯ иногда предлагается под *парадигматическими (базисными) отношениями* понимать «*лишь такие отношения между словами (означающими), которые основаны на существовании тех или иных связей между означаемыми*» [9, с. 433].

Разные специалисты предлагают различные способы определения парадигматических связей: по *сходству предметов*, по *принадлежности к одному классу*, *ассоциативные отношения* (ассоциации по *смежности* в пространстве и во времени, по *сходству*, по *контрасту*, отношения *соподчинения*, «*вид-род*», «*причина-следствие*», «*часть-целое*» и т.п.).

При этом допускается произвольное установление отношений в конкретном ИПЯ, с ориентацией на повышение эффективности информационного поиска.

В частности, **Э. С. Бернштейн, Д. Г. Лахути и В. С. Чернявский**¹⁰⁴ использовали при разработке ИПС «Пусто-Непусто» парадигматические отношения, которые определяют как отношения, существующие между словами поискового языка независимо от контекста, называя именно их базисными отношениями (БО), и задавали их списком (включая в тезаурус). БО увеличивают семантическую силу системы, позволяют формулировать запросы в терминах, отличных от терминов, употребляемых в релевантных документах.

¹⁰⁴ **Бернштейн Э.** Вопросы теории поисковых систем / Э. Бернштейн, Д. Лахути, В. Чернявский. – М.: ВНИИЭМ, 1966. – 64 с.

Фиксированные базисные отношения могут быть заданы различными способами: с помощью структуры слов (как в УДК), с помощью системы ссылок, с помощью деревьев дескрипторов и т. п.

Следует иметь в виду, что, стремясь улучшить результаты поиска, можно увеличить «шум», т. е. избыточную выдачу.

В первых работах по теории информационного поиска термин дескриптор использовался иногда [9] как синоним понятия *ключевое слово*

В различных языках эти компоненты ИПЯ используются по-разному. Словарь может иметь достаточно сложную структуру, т. е. представлять собой *тезаурус*, который может включать в себя и алфавит, и слова, и словосочетания, и более сложные конструкции.

Термин *тезаурус* (от греч. «θησαυροζ», «thesauros» – сокровищница, богатство, клад, запас и т.п.) в общем случае характеризует «совокупность научных знаний о явлениях и законах внешнего мира и духовной деятельности людей, накопленную всем человеческим обществом» [9, с. 85]. Этот термин был введен в современную литературу по языкознанию и информатике в 1956 году Кембриджской группой по изучению языков. В то же время термин существовал раньше: в эпоху Возрождения тезаурусами называли энциклопедии. С обзором определений тезауруса и первых тезаурусов можно познакомиться в [9, с. 415–432, 469–505].

Особую роль в формировании тезауруса играют *базисные (парадигматические)* отношения, которые исторически являются элементом *логики ИПС*.

В математической лингвистике и семиотике термин тезаурус используется в более узком смысле, для характеристики конкретного языка, его многоуровневой структуры.

Для этих целей удобно пользоваться одним из принятых в лингвистике определений тезауруса как «*множества смысло-*

выражающих элементов языка с заданными смысловыми отношениями»¹⁰⁵.

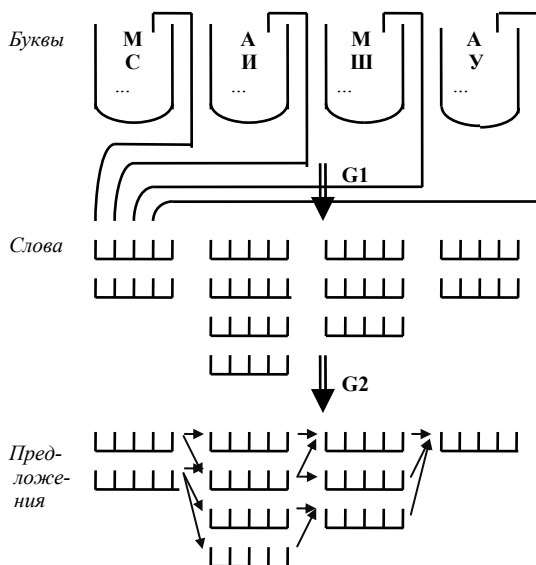


Рис. 5.5. Принципы формирования структуры тезауруса

Это определение иллюстрирует рис. 5.5 на элементарном примере формирования слов из букв и предложений их слов.

Разумеется, в реальных тезаурусах уровни носят иные названия: ключевые слова, дескрипторы, абзацы и иные лингвистические и логические элементы.

При этом между уровнями тезауруса могут существовать различные взаимоотношения – от древовидных иерархических до причинно-следственных.

Таким образом, тезаурус позволяет представить структуру языка в виде уровней (страт) множеств слов, предложений, абзацев и т. п., смысловыражающие элементы каждого из которых фор-

¹⁰⁵ Шрейдер Ю. А. Информация в структурах с отношениями / Ю. А. Шрейдер // Исследования по математической лингвистике, математической логике и информационным языкам: сб. – М.: Наука, 1972. – С. 147–159.

мируются из смысловыражающих элементов предшествующих структурных уровней.

Правила формирования смысловыражающих элементов второго, третьего и последующих уровней в тезаурус не входят. Они образуют грамматику информационно-поискового языка (*G1, G2* и т. д.). В тезаурусе же определяется только вид и наименование уровня, характер и вид смысловыражающих элементов.

Иногда вместо термина *смысловыражающие элементы* используется термин *синтаксические единицы* тезауруса. Однако, это менее удачный термин, так как при формировании элементов нового множества смысловыражающих элементов каждого последующего уровня (при образовании слов из букв, фраз и предложений из слов и т. д.) у элементов вновь образованного множества появляется новый смысл, т. е. как бы проявляется закономерность целостности, и это хорошо отражает термин «смысловыражающий элемент».

Понятие тезауруса стало в первую очередь использоваться при разработке информационно-поисковых языков, но в последующем его стали применять и при создании других искусственных языков – языков моделирования, автоматизации проектирования.

Тезаурус позволяет охарактеризовать язык с точки зрения уровней обобщения, ввести правила их использования при индексировании информации. В теории научно-технической информации [9, 10, 17, 18, 20 и др.] исследуются различные свойства тезауруса.

Можно говорить о *глубине* тезауруса того или иного языка, характеризуемой числом уровней, о видах *уровней обобщения*, и, пользуясь этими понятиями, сравнивать языки, выбирать более подходящий для рассматриваемой задачи или, охарактеризовав структуру языка, организовать процесс его разработки.

В практике создания информационно-поисковых систем наиболее известен словарь-тезаурус «Тезаурус ASTIA»¹⁰⁶.

В системе SMART¹⁰⁷ содержится два вида тезаурусов:

¹⁰⁶ *Vikery B. C.* Thesaurus of ASTIA Descriptors, 2nd Ed //Armed Forces Technical Information Agency. – Arlington, Virginia: Dec. 1962.

- Тезаурус с иерархической структурой понятий.

Дает возможность для любого номера понятия найти их «родителя», «сыновей», «братьев» и множество возможных перекрестных ссылок.

- Словарь синонимов или тезаурус.

Используется для замены значащих слов номерами понятий, каждое из которых представляет класс основ слов, близких по смыслу.

Простейшими тезаурусами являются словари *дескрипторов* при толковании дескриптора как имени класса условной эквивалентности, формируемого на основе *парадигматических отношений*.

Тезаурусы разрабатываются и в отечественных отраслевых системах научно-технической информации (напр., в АСНТИ-геология¹⁰⁸). Термин тезаурус иногда используется в более широком смысле.

Например, **Ю. И. Шемакин** тезаурусом называет сложную систему организации в автоматизированных системах управления и обработки информации разных ее видов (научно-технической, управленческой, представляемой в документальной и фактографической форме)¹⁰⁹.

Морфологию и *синтаксис* удобно объединять единым термином – *грамматика*. Тогда говорят, что ИПЯ состоит из *тезауруса* и *грамматики*, а затем рассматривают смысловыражающие элементы (синтаксические единицы) тезауруса и правила грамматики.

Под *грамматикой* (которую иногда называют *синтактикой*, *синтаксисом*, что сужает понятие грамматики, исключая из него *морфологию*) понимаются правила, с помощью которых форми-

¹⁰⁷ *Сэлтон Г.* Автоматическая обработка, хранение и поиск информации / Г. Сэлтон. – М.: Сов. радио, 1973. – 560 с.

¹⁰⁸ *Автоматизированная* система научно-технической информатики – разработка и эксплуатация / К. И. Володин, Л. Л. Гульницкий, И. Ф. Пожариский и др. – М.: Финансы и статистика, 2004. – 192 с.

¹⁰⁹ *Шемакин Ю. И.* Тезаурус в автоматизированных системах управления и обработка информации / Ю. И. Шемакин. – М.: Воениздат, 1974. – 172 с.

руются смысловыражающие элементы языка. Пользуясь этими правилами, можно «порождать» (формировать) грамматически (синтаксически) правильные конструкции или распознавать их грамматическую правильность.

Простейшими правилами грамматики являются синтагматические (текстуальные) отношения.

При создании и использовании искусственных языков для информационно-логических систем применяют понятия математической лингвистики. В частности, понятия *порождающей* и *распознающей* грамматики.

Под *порождающей* грамматикой понимается совокупность правил, с помощью которых обеспечивается возможность формирования (*порождения*) из первичных элементов (словаря) синтаксически правильных конструкций.

Под *распознающей грамматикой* – правила, с помощью которых обеспечивается возможность распознавания синтаксической правильности предложений, фраз или других фрагментов языка ¹¹⁰.

На базе математической лингвистики развивается *теория формальных грамматик Н. Хомского*.

Классы формальных грамматик Н. Хомского ¹¹¹ считаются основой *теории формальных языков*.

Формальный язык определяется как множество (конечное или бесконечное) предложений (или «цепочек»), каждое из которых имеет конечную длину и построено с помощью некоторых операций

¹¹⁰ *Гладкий А. В.* Элементы математической лингвистики / А. В. Гладкий, И. А. Мельчук. – М.: Наука, 1969. – 88 с.; *Гросс М.* Теория формальных грамматик / М. Гросс, А. Лантен; под ред. А. В. Гладкого. – М.: Мир, 1971. – 294 с.

¹¹¹ *Хомский Н.* Три модели для описания языка / Н. Хомский // Кибернетический сборник. – Вып. 2. – М.: Изд-во ИЛ, 1961; *Хомский Н.* Введение в формальный анализ естественных языков / Н. Хомский, Дж. Миллер // Кибернетический сборник: Новая серия. – Вып. 3. – М.: Мир, 1965.

(правил) из конечного множества элементов (символов), составляющих алфавит языка.

Формальную грамматику определяют в виде четверки множеств

$$G = \langle V_T, V_N, R, A \rangle, \quad (5.9)$$

где V_T – множество *основных* или *терминальных* символов;

V_N – множество *вспомогательных* или *нетерминальных* символов;

R – множество *правил вывода*, или *продукций*, которые могут иметь вид

$$\alpha \rightarrow \beta, \quad (5.10)$$

где $\beta \in (V \cup V_N)$, т. е. β – цепочка конечной длины из терминальных и нетерминальных символов множеств V_T и V_N ; $\alpha \in (V_T \cup V_N)V_N(V_T \cup V_N)$, т. е. α является цепочкой из терминальных и нетерминальных символов, содержащей по крайней мере один нетерминальный символ из V_N ; A – множество аксиом (в грамматиках комбинаторного типа, к которым относятся грамматики **Н. Хомского**, A состоит из одного начального символа S , причем $S \subset V_N$).

Учитывая, что в литературе по формальным грамматикам, как правило, не стремятся к содержательной интерпретации получаемых выводов, а рассматривают лишь формальную сторону процессов порождения и распознавания принадлежности цепочек к соответствующему классу грамматик, приведем содержательный пример порождающей грамматики.

Предположим, дано: $V_T = \langle \epsilon_1, \epsilon_2, n, l \rangle,$
 $V_N = \langle S, P \rangle.$

<i>Порождающая грамматика</i>	<i>Распознающая грамматика</i>	
$\left\{ \begin{array}{l} S \rightarrow SP \\ S \rightarrow \epsilon_1 S \\ S \rightarrow \epsilon_2 S \\ S \rightarrow n \\ P \rightarrow l \end{array} \right.$	$\left\{ \begin{array}{l} SP \rightarrow S \\ \epsilon_1 S \rightarrow S \\ \epsilon_2 S \rightarrow S \\ n \rightarrow S \\ l \rightarrow P \end{array} \right.$	(1) (2') (3') (4') (5')

(5.11)

Применяя правила R левой части (5.11) в приведенной последовательности, получим

$$S \Rightarrow SP \Rightarrow \epsilon_1 SP \Rightarrow \epsilon_1 \epsilon_2 SP \Rightarrow \epsilon_1 \epsilon_2 n P \Rightarrow \epsilon_1 \epsilon_2 n l.$$

(1) (2) (3) (4) (5)

Это – формальная сторона процесса порождения. Для того, чтобы получить интерпретируемое выражение, нужно расшифровать терминальные символы, включенные в V_N , где ϵ_1 – ВСЕ, ϵ_2 – ВОЗРАСТЫ, n – ПОКОРНЫ, l – ЛЮБВИ.

Тогда полученное предложение

« $\epsilon_1 \epsilon_2 n l$ » – «ВСЕ ВОЗРАСТЫ ПОКОРНЫ ЛЮБВИ».

Если изменить последовательность применения правил, то будут получаться другие предложения. Например, если применить правила в последовательности (1) \Rightarrow (3) \Rightarrow (2) \Rightarrow (4) \Rightarrow (5), то получится «ВОЗРАСТЫ ВСЕ ПОКОРНЫ ЛЮБВИ». Если применить не все правила: например, (1) \Rightarrow (2) \Rightarrow (4) \Rightarrow (5), то получим «ВСЕ ПОКОРНЫ ЛЮБВИ».

Если же попытаться получить предложение, как у А. С. Пушкина – «*Любви все возрасты покорны*», то, как бы мы не меняли последовательность правил, получить эту фразу не удастся. Нужно изменить первое правило: вместо $S \rightarrow SP$ включить в R правило $S \rightarrow PS$.

Из примера видно, что вид порождаемых цепочек (предложений) зависит от *вида правил (исчисления)* и от *последовательности их применения (алгоритма)*.

С помощью приведенного примера легко также продемонстрировать тесную связь понятия «*грамматически правильный*» с языком (грамматикой).

Распознающая грамматика для рассматриваемого примера будет содержать как бы «перевернутые» правила – правая часть (5.11), которые должны применяться в обратной последовательности. Пример представления анализа правильности предложения с помощью правил распознающей грамматики приведен на рис. 5.6.

При распознавании правильности предложения если не оговаривать, что предложение (цепочка) грамматически правильно с точки зрения правил данного формального языка, то можно, пользуясь формальной грамматикой в первоначальном виде, получить вывод, что приведенная фраза Пушкина грамматически неправильна с точки зрения правил грамматики (5.11).

Действительно, с точки зрения правил грамматики для построения делового текста, которым соответствуют правила (5.11), другие поэтические строки часто получали бы формальную оценку «грамматически неправильно». И, напротив, если построить грамматику на основе анализа пушкинского стиля, то в деловом тексте получились бы предложения типа «Я решение свое принял правильное» (подобно фразе «Я памятник себе воздвиг нерукотворный»).

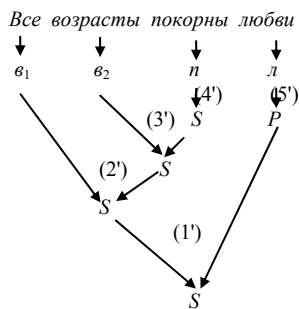


Рис. 5.6

Понятие формальной грамматики используют при создании языка моделирования соответствующего литературного или музыкального произведения – пародий, подражательств или, как иногда принято говорить, произведений соответствующего стиля или класса.

Например, известны работы **Р. Х. Зарипова**¹¹² по моделированию музыкальных произведений в стиле, или в классе, массовых советских песен, моделирование процесса сочинения стихотворных произведений и т. п.

Подобным же образом можно моделировать порождение деловых писем или других документов, имеющих, как правило, не только формализованный стиль, но и формальную структуру. Аналогично можно создавать языки моделирования структур, языки автоматизации проектирования сложных устройств и систем определенного вида (класса).

Основу таких работ составляют идеи, которые можно пояснить с помощью классов грамматик, впервые предложенных **Н. Хомским**¹¹³.

¹¹² **Зарипов Р. Х.** Машинный поиск вариантов при моделировании творческого процесса / Р. Х. Зарипов. – М.: Наука, 1983. – 232 с

¹¹³ **Хомский Н.** Три модели для описания языка / Н. Хомский // Кибернетический сборник. – Вып. 2. – М.: ИЛ, 1961.

Разделение грамматик на классы определяется видом правил вывода R . В зависимости от них можно выделить четыре основных, наиболее часто рассматриваемых класса грамматик (табл. 5.3).

Таблица 5.3

Основные классы грамматик Н. Хомского

Класс	Характеристика
<i>1-й класс. Неукорачивающие (НУ-грамматики)</i>	На правила вывода накладывается только одно требование, чтобы в левой части правила вывода было всегда меньше символов, чем в правой, т. е. чтобы правила были неукорачивающими, не уменьшали число символов в выводимых цепочках. Иногда эти грамматики называют грамматиками типа ноль (<i>нулевого типа</i>) или <i>алгоритмическими</i> .
<i>2-й класс. Нонтекстные, контекстно-связанные</i>	На правила вывода, помимо требований неукорачиваемости, накладывается ограничение, чтобы на каждом шаге изменялся только один символ в контексте, т. е. чтобы $Z1 B Z2 \rightarrow Z1 W Z2$, где B – один нетерминальный символ, W – непустая цепочка символов, т. е. $W \neq \emptyset$. Иногда применяют термин – грамматика <i>непосредственных составляющих</i> (НС-грамматики).
<i>3-й класс. Контекстно-свободная (КС-грамматика) или бесконтекстная</i>	Кроме неукорачиваемости требуется, чтобы правила имели вид $B \rightarrow \beta,$ т. е. α всегда состоит из одного вспомогательного символа.
<i>4-й класс. Автоматные (А-грамматики)</i>	На правила вывода накладывается по сравнению с третьим классом еще одно ограничение, требующее, чтобы в правилах вывода нетерминальный символ всегда стоял справа или слева, т. е. с одной стороны. Если нетерминальный символ стоит слева, т. е. правила имеют вид $A \rightarrow aB$ или $A \rightarrow a$, где $(A, B) \in V_N$, $a \in V_T$, автоматная грамматика является <i>праволинейной</i> ; если нетерминальный символ стоит справа – то автоматную грамматику называют <i>леволинейной</i> .

В полной теории формальных грамматик с правилами типа подстановки есть и промежуточные классы.

В теории формальных грамматик показано, что имеет место следующее соотношение:

$$A \subseteq KC \subseteq HC \subseteq HU. \quad (5.12)$$

Иногда доказывают, что имеет место строгое вхождение

$$A \subset KC \subset HC \subset HY. \quad (5.12,a)$$

При исследовании разных классов формальных грамматик получены результаты, которые позволяют сделать вывод, что по мере уменьшения числа ограничений, накладываемых на правила вывода, т. е. по мере продвижения в (5.11) слева направо, в языке увеличивается возможность отображения смысла, т. е. возможность выражения с помощью формальных правил семантических особенностей проблемной ситуации. Говорят, что формальная система становится более богатой. Однако при этом в языке растет число алгоритмически неразрешимых проблем, т. е. увеличивается число положений, истинность или ложность которых не может быть доказана в рамках формальной системы языка.

Здесь мы сталкиваемся фактически с проблемой Гёделя, которая в теории формальных языков обсуждается обычно в терминах этой теории. А именно: вводится понятие «*операция определена (или не определена) на множестве языков данного класса*»; и считается, что операция определена на множестве языков данного класса, если после применения ее к языкам, входящим в это множество, получается язык, принадлежащий множеству языков этого класса.

Например, если $Y_1 \subset KC$ и $Y_2 \subset KC$, и если $(Y_1 \cup Y_2) \subset KC$, то операция объединения \cup определена на классе KC -языков.

Характеризуя с помощью введенного понятия классы языков, отмечают, что в последовательности (5.11) по мере продвижения слева направо увеличивается число операций, которые не определены на множестве языков данного класса.

Следует оговорить, что дело обстоит не так прямолинейно. Точнее было бы сказать, что для большого числа операций нет доказательств, что они определены на классах HC -языков и HY -языков, т. е. эти доказательства становятся сложнее или вообще (в силу теоремы Гёделя) нереализуемы средствами теории формальных грамматик.

Приведенное упрощенное представление проблемы помогает обратить внимание тех, кто будет заниматься разработкой языков программирования или программных систем, языков моделирования, автоматизации проектирования, на необходимость учета следующей закономерности: *чем большими смысло-*

выражающими возможностями обладает знаковая система, тем в большей мере растет в ней число алгоритмически неразрешимых проблем (т. е. тем менее доказательны в ней формальные процедуры).

При выходе в класс произвольных грамматик, в котором не выполняется даже условие неукорачиваемости, доказать допустимость тех или иных формальных преобразований средствами математической лингвистики практически невозможно, и поэтому в поисках новых средств исследователи обратились к семиотическим представлениям. Здесь можно провести как бы формальную границу между лингвистикой и семиотикой.

При создании ИПЯ с тезаурусом и грамматикой важную роль играют понятия *семантики* и *прагматики*.

Под *семантикой* понимается *содержание значение, смысл* формируемых или распознаваемых конструкций языка; под *прагматикой* – *полезность для данной цели, задачи*.

В естественном языке различить понятия, с помощью которых характеризуются термины «*семантика*» и «*прагматика*», трудно; обычно пояснить различие можно лишь при парном сопоставлении терминов [20 и др.]:

<семантика> :: = <содержание> | <смысл> | <значение>;
<прагматика> :: = <смысл> | <значение> | <полезность>.

Поэтому принято рассматривать эти понятия на примерах. Поясним различие между семантически и прагматически правильными конструкциями языка на следующих легко запоминающихся примерах.

Традиционно для пояснения синтаксической правильности и семантической бессмыслицы используется предложенный **Л. В. Щербой** пример «*Глокая куздра тцето борзданула бокра и курдычет бокрѣнка*» (в котором просто нет ни одного слова естественного языка, имеющего смысл). Но примеры можно найти и в естественной речи.

Предложение «*Муха лукаво всплеснула зубами*» синтаксически правильное, но не имеет смысла в естественном русском языке в обиходном, широком употреблении, т. е. является с точки зрения пользователей русским языком семантически неправильным (исключим пока гипотетическую ситуацию сказки, в которой муха может быть наделена указанными свойствами).

Другое предложение «*Маленькая девочка собирает цветы на лугу*» – синтаксически и семантически правильное. Однако для директора завода (если это луг, а не заводской газон, и – учтем личный фактор – если эта девочка не его дочь) это предложение не несет никакой информации, т. е. прагматически (с точки зрения целей руководителя) является неправильным. Другое дело, если «Иванов (который в данный момент должен находиться на рабочем месте) собирает цветы на лугу». Тогда это предложение было бы и прагматически правильным.

Возвратимся теперь к примеру с мухой. Приведенное предложение, *семантически неправильное*, может в гипотетической ситуации сказки оказаться *прагматически правильным*, что важно иметь в виду при применении лингвистических представлений.

Пользуясь этими правилами, можно лучше отразить смысл документа или запроса в ПОД и ПОЗ, повышая релевантность поиска.

Виды и классификации ИПЯ

В зависимости от используемых компонентов ИПЯ бывают разных видов.

В качестве первоначально использовавшихся ИПЯ **Ч. Мидоу** [8] выделяет следующие (рис. 5.7).

Иерархические классификации.

Например, Десятичная классификация *Дьюи*¹¹⁴, классификация библиотеки Конгресса США, Универсальная Десятичная классификация (УДК)¹¹⁵, которая представляет собой модификацию системы Дьюи, получившую широкое применение.

Иерархические классификации обеспечивают возможность расширения «вниз», т. е. уточнения описания документа, но они отличаются жесткостью, их достаточно трудно изменять.

- *Язык предметных заголовков*

Подобно иерархической классификации использует фиксированное число предметных классов (часто располагаемых по алфавиту), но для его терминов обычно не используется определенный код. Язык позволяет любому документу приписывать более чем один термин, почти не имеет структуры и средств для выражения взаимоотношений между терминами.

¹¹⁴ *Dewey V. Dewey Decimal Classification and Relative Index*, Forest Press, Lake Placid Club / V. Dewey. Essex Country, New York, 1959.

¹¹⁵ *Сборник методических материалов по Универсальной Десятичной Классификации.* – М.: Гос. Комитет Совета Министров СССР по науке и технике, ГПНТБ, 1969. – 485 с.

Предметные заголовки используются, например, в журналах (рубрики), в классификаторах специальностей вузов и ВАК. Этот язык можно считать языком иерархической классификации, но с ослабленной структурой, что облегчает его разработку, но затрудняет изучение и применение.

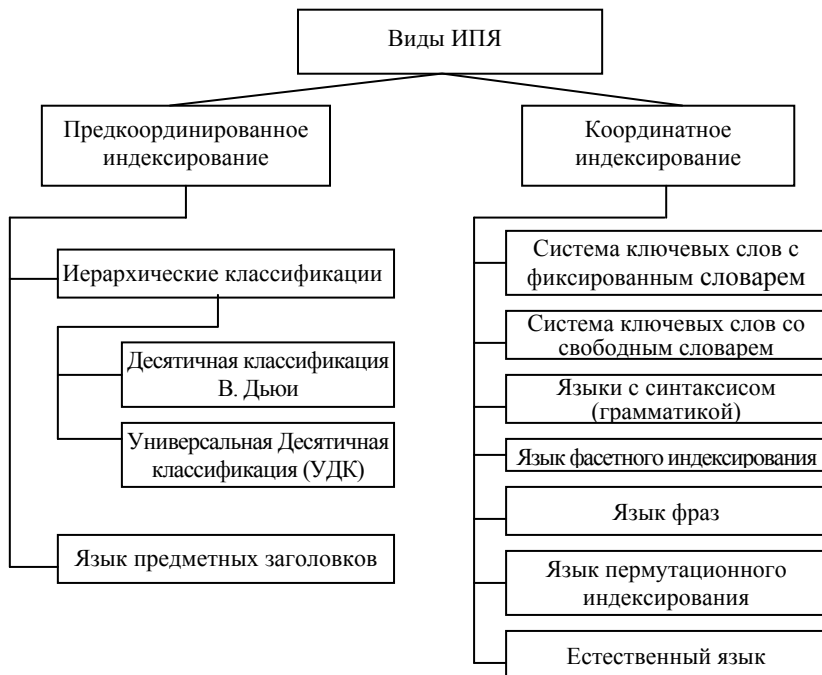


Рис. 5.7. Классификация ИПЯ *Ч. Мидоу*

Рассмотренные языки иногда называют *предкоординированными* (*pre-coordinate*) системами, поскольку семантические комбинации терминов не определены в словарном составе ИПЯ, а составляются его разработчиками. Такие языки неизбежно имеют пробелы, в них трудно отражать новые предметные области (что легко видеть на примере системы УДК).

Поэтому для более совершенного описания содержания документа стало применяться индексирование ключевыми словами – *координатное индексирование*:

- *Система ключевых слов с фиксированным словарем.*

Такие языки могут применяться для узкоспециализированных ИПС с достаточно формализованной (унифицированной) терминологией (напр., система «УниTERM», разработанная *М. Таубе*¹¹⁶). Эта система подобна системе предметных заголовков, но в отличие от нее, во-первых, ключевые слова короче предметных заголовков (обычно это единичные слова, иногда – короткие словосочетания), и во-вторых, объем полного словарного состава существенно больше. Словарный состав здесь, как правило, фиксирован, и отсутствуют средства установления связей между словами (синтаксис), но включение набора слов в ПОД или ПОЗ позволяет пользователю как бы угадывать эти связи, что помогает более полно описать исходный документ или запрос.

- *Система ключевых слов со свободным словарем.*

Такие языки позволяют пользователю выбирать для описания документа любые слова (за исключением союзов и предлогов), руководствуясь их ролью в отображении содержания документа. Это позволяет точнее отобразить содержание документа, но может снизить релевантность поиска, поскольку составители ПОЗ не могут предусмотреть точки зрения составителей ПОД.

- *Языки с синтаксисом (грамматикой).*

В качестве простейшего из ИПЯ этого вида *Ч. Мудой* [8] выделяет язык *помеченных дескрипторов (tagged descriptors)*, с помощью которого отображение смысла в ПОД и ПОЗ осуществляется путем присоединения к основному дескриптору (или ключевому слову) уточняющих дескрипторов (ключевых слов), роль которых состоит в том, чтобы либо классифицировать основной дескриптор как имя собственное, характерный признак или действие, либо объединить в одну группу дескрипторы, относящиеся к одному и тому же предмету документа. В современных поисковых системах Интернет применяются операции AND, OR, отображающие логические операции дизъюнкции и конъюнкции), в простейшем варианте – слова заключают в кавычки.

В дальнейшем были разработаны ИПЯ с грамматикой, содержащей более сложные правила.

- *Язык фасетного индексирования.*

В более развитых в синтаксическом отношении вариантах такого языка различные дескрипторы могут изменять значения друг друга. Простейшим примером такого синтаксиса является запись команды ЭВМ,

¹¹⁶ *Taube M.* Unit Terms in Coordinate Indexing / *M. Taube, C.D. Gull, I.S. Wachtel* // *Amer. Documentations.* –1952. – № 3, 4. – P. 213–218.

состоящей из собственно оператора и адреса хранения информации. В качестве примера можно также привести уточненное описание товара, включая фасон, цвет, цену и т.п. характеристики товара. Такой синтаксис основан на известном в теории множеств положении: в результате помещения рядом элементов разных множеств возникает эффект появления нового смысла. Языки такого вида позволяют частично устранять омонимию с учетом контекста. Различные роли, которые играют дескрипторы в таких языках, называют *фасетами* ¹¹⁷.

Для фасет могут быть использованы дескрипторы из одного и того же словаря. Располагаются фасеты в порядке значимости дескрипторов для отображения содержания индексируемого документа. В отличие от иерархических классификаций фасеты можно располагать в произвольном порядке.

- *Язык фраз.*

В качестве ПОД используются индексирующие фразы. В этом случае контекст ключевых слов позволяет частично снять проблемы семантической неоднозначности.

Трудность здесь состоит в выборе фраз, включаемых в язык. Кроме того, отсутствует возможность расширения ПОД.

- *Язык пермутационного индексирования.*

Пермутационный указатель включает контекст каждого слова, содержащегося в фразе и называется указателем ключевых слов, взятых в контексте, или указателем типа KWIC ¹¹⁸.

Идею такого указателя легче пояснить примером:

Системы индексирования **документов**

Системы **индексирования** документов

Системы индексирования документов

При этом ПОД образует колонка ключевых слов в центре, расположенных в порядке алфавита. Можно считать этот ПОД фразой, но он эффективнее предыдущего с точки зрения смысловыражающих возможностей. Однако пермутационное индексирование трудоемко и неэффективно экономически.

- *Естественный язык.*

Наиболее точно может отразить семантику текста, однако помимо трудоемкости и экономической неэффективности, возникают проблемы

¹¹⁷ *Vikery B. C.* On Retrieval Systems Theory / B. C. Vikery. – London: Butterworth, 1961, 33 p.

¹¹⁸ *Vikery B. C.* General Information Manual, Keyword-in-Context (KWIC) Indexing / B. C. Vikery. New York: IBM Corp., White Plains, 1962.

синонимии, омонимии и др. неоднозначности естественного языка, затрудняющие алгоритмизацию поиска.

Уменьшить неоднозначности языка помогает *словарь*.

В истории развития информационного поиска разрабатывались и применялись разнообразные словари: словарь синонимичных пар; словарь с многократными связями (например, двуязычный словарь); классификационная таблица; словарь с определениями на естественном языке; отрицательный словарь (содержащий запрещенные словосочетания); словарь-тезаурус («Тезаурус ASTIA»¹¹⁹ и др.).

Таким образом, существуют *ИПЯ, использующие ключевые слова; дескрипторные ИПЯ без грамматики и с грамматикой, ИПЯ с отрицательным словарем, ИПЯ с тезаурусом* и т. д.

Существуют и иные классификации ИПЯ.

Так, в [17, с. 31–36] предлагается следующее разделение ИПЯ (рис. 5.8):



Рис. 5.8. Классификация ИПЯ *А.И. Черного*

¹¹⁹ *Vikery B. C.* Thesaurus of ASTIA Descriptors, 2nd Ed. // B. C. Vikery // Armed Forces Technical Information Agency. – Arlington, Virginia: Dec. 1962.

1. Предкоординированные ИПЯ.

1.1. Перечислительные классификации: иерархические, алфавитно-предметные (по Мидоу – язык предметных заголовков).

1.2. Фасетные классификации.

2. Посткоординированные (координатные по Мидоу) ИПЯ.

2.1. Дескрипторные языки (с координацией посредством использования операции логического умножения или пересечения \cap).

2.2. Семантические коды, задающие парадигматические отношения структурами лексических единиц (код Перри – Кента ¹²⁰, RX-коды языка «Бит» ¹²¹).

2.3. Синтагматические языки с развитой системой средств отображения синтагматических отношений (см., например, язык СИНТОЛ ¹²²).

3. *Языки библиографических ссылок.* Предлагались классификации ИПЯ по типу их словарного состава (см. ссылки в [11]), типам языковых единиц, степени их сложности, характеру отношений между этими единицами (виду грамматики), системам индексирования и т. п.

При этом следует иметь в виду, что на практике конкретный ИПЯ нельзя строго отнести к тому или иному классу, поскольку: во-первых, некоторые ИПС могут работать и в режиме без грамматики, и в режиме с грамматикой (например, системы СИНТОЛ ¹²³, СМАРТ ¹²⁴; а, во-вторых, ИПС развиваются, и ос-

¹²⁰ *Perry J. W.* Tools for machine literature searching. Semantic code dictionary. Equipment, Procedures / J. W. Perry and A. Kent. – New York, Interscience Publishers, 1958. – 241 p.

¹²¹ *Информационно-поисковая система «БИТ»* / Э. Ф. Скороходько, Л. Э. Пшеничная, П. Н. Кар-Ялайне и др. – Киев: Наукова думка, 1968. – 219 с.

¹²² *СИНТОЛ* // Сборник переводов по вопросам информационной теории и практики. – М.: ВИНТИ, 1968. – С. 36–47, 50–52, 66–72, 76–80.

¹²³ *СИНТОЛ* // Сборник переводов по вопросам информационной теории и практики. – М.: ВИНТИ, 1968. – С. 36–47, 50–52, 66–72, 76–80.

¹²⁴ *Сэлтон Г.* Автоматическая обработка, хранение и поиск информации / Г. Сэлтон. – М.: Сов. Радио, 1973. – 560 с.

новой является развитие ИПЯ. Поэтому обычно ИПС и ИПЯ описывают рядом характеристик, с тем чтобы пользователь мог выбрать желаемые.

При выборе ИПЯ необходимо оценивать их эффективность. При оценке эффективности ИПЯ используют различные критерии.

В частности, **Ч. Мидоу** предлагает четыре меры: *семантическую силу, многозначность, компактность и стоимость* [8, с. 71–79].

Качественные графические иллюстрации исследований, приводимые в [8], показывают, что оценки семантической силы и многозначности ИПЯ коррелируют друг с другом, вступая в противоречие с компактностью и стоимостью, и что наиболее сложной проблемой является оценка семантической силы ИПЯ.

Проблема оценки ИПЯ, в свою очередь, является составной частью более общей проблемы – оценки *качества информационного поиска*.

5.6. Система индексирования

Процедуру перевода с естественного языка на ИПЯ называют *индексированием*. Результатом такого перевода является ПОД (при вводе документов в ИПС) или ПОЗ (при индексировании запроса пользователя).

Проблема индексирования связана с семантическим анализом текстов документов. Сложность ее связана с тем, что индексирование документов, вводимых в поисковые массивы, и запросов пользователя разнесены во времени.

Для алгоритмизации и автоматизации индексирования необходимо решить проблему выбора для включения в ПОД или ПОЗ наиболее значимых ключевых слов, дескрипторов, фраз (в зависимости от лексических единиц ИПЯ).

Важность можно определить несколькими признаками:

- статистически, т. е. на основе частоты использования термина в документе;

- на основе высказываний автора (его мнения, отраженного в заглавии документа или подзаголовках, выделяемых автором в документе);

- с помощью грамматики, позволяющей отразить взаимосвязи между лексическими единицами, содержащимися в контексте;

- по критериям важности, сформулированным пользователем, для чего при индексировании документов могут быть указаны весовые коэффициенты дескрипторов.

Система индексирования конкретной ИПС в основном определяется возможностями ИПЯ, имеющимися в нем лексически и синтаксическими средствами. Однако есть и некоторые специфические правила и рекомендации, исследование которых позволило выявить некоторые разновидности систем индексирования.

Существуют различные типы систем индексирования:

1. К первому типу относят системы *свободного индексирования*.

При этом способе из индексируемого документа выписываются в ПОД слова или словосочетания, которые отражают содержание индексируемого документа. Кроме этого, элементами ПОД могут быть слова, отсутствующие в этих документах, но отражающие более точно смысл их текстов с точки зрения целей создания ИПС. Выписанные элементы упорядочиваются в алфавитном порядке. Такой упорядоченный набор слов (словосочетаний) представляет собой ПОД при этом типе индексирования. Аналогично – из текста запроса пользователя формируется ПОЗ.

Такой процесс индексирования является принципиально неалгоритмическим, т. е. неавтоматизируемым.

2. При втором методе, который условно называют методом *полу-свободного индексирования*, из документа выписывают слова и словосочетания вначале так же, как и при свободном индексировании.

Однако выписанные элементы сравнивают затем с фиксированным словарем, не найденные в нем – устраняют, а оставшиеся, упорядочиваемые в алфавитном порядке, представляют собой ПОД (или ПОЗ).

3. Третий способ индексирования основан на *статистическом подходе*.

Выбор слов (выражений) исходного текста, подлежащих включению в ПОД, производится на основе статистического анализа текста, при котором его слова рассматриваются как знаки, не имеющие семантических значений. При этом предлагались различные *статистические критерии*, основанные на сопоставлении относительной частоты употребления слова в документе и относительной частоты употребления слова в представительном массиве документов (т. е. в репрезентативной статистической выборке).

Например, в [9] предлагаются следующие количественные критерии:

$$|F - R| > k; \quad \frac{F}{F + R} < k; \quad \frac{F}{R} > k,$$

где F – относительная частота употребления слова в документе; R – относительная частота употребления слова в представительном массиве документов.

Легко видеть, что в основе приведенных соотношений лежит идея, согласно которой информационная значимость слова определяется расхождением частоты его употребления в данном документе и во всем потоке рассматриваемых документов.

Возможны различные подходы к определению расхождения:

- согласно первому вычисляется расхождение между частотой употребления слов в потоке документов данной тематики (монотематический поток) и частотой встречаемости этого слова в многотемном потоке документов (политематический поток);
- второй принцип основан на вычислении расхождения частоты употребления слова в потоке текстов данной тематики и частоты этого же слова в потоке текстов тематики, далекой от данной («противоположной» тематики).

Статистический способ индексирования может быть алгоритмизирован и автоматизирован, и в настоящее время имеются средства автоматизированного статистического анализа текстов.

Однако самостоятельного практического применения в ИПС этот способ не нашел, он используется как вспомогательный в сочетании с семантическим анализом текстов документов.

4. К четвертому типу относят *системы индексирования, контролируемые заданным словарем (тезаурусом)*.

Алгоритм индексирования сводится к тому, что каждое слово текста сравнивается с точностью до основы со словарем, совпавшие слова записываются в ПОД.

В некоторых системах словарь используется как помощник специалисту, занимающемуся индексированием текста. К таким системам относится, например, УДК.

В других системах такой словарь является элементом алгоритма индексирования: слово, одновременно встретившееся в тексте и в словаре, записывается в ПОД. В дескрипторных ИПЯ в ПОД (ПОЗ) записываются не само слово текста, а соответствующий ему дескриптор.

Перспективным представляется индексирование документов с использованием специально разработанных иерархических классификаций, отражающих цели поиска и использования документов.

Такие классификаторы могут использоваться в качестве ИПЯ в информационных системах нормативно-методического обеспечения управления: иерархический классификатор, объединяющий нормативно-методические документы, разрабатывается на основе структуры целей (основных направлений) и функций деятельности предприятия.

Иерархический классификатор ИПЯ может быть основой системы избирательного распределения информации: разрабатывается классификатор потребностей категории работников, пользующихся системой ИРИ.

5.7. Логика ИПС. Критерии смыслового соответствия

Как показано на рис. 5.4, под *логикой* ИПС понимаются *критерии выдачи* или *критерии смыслового соответствия* (обязательный элемент), *базисные (парадигматические)* и *текстуальные (синтагматические)* отношения между словами ИПЯ (базисные и/или текстуальные отношения могут и отсутствовать). *Критерий смыслового соответствия (КСС) или критерий выдачи* позволяет решать вопрос о выдаче или не выдаче того или иного документа, т. е. являются основой алгоритма поиска.

Существуют следующие виды КСС [9, 17]:

* «На полное вхождение» или «на вхождение».

Условием выдачи документов является полное вхождение ПОЗ в ПОД. Иными словами, документ выдается, если множество дескрипторов, образующее ПОЗ ($M_{\text{ПОЗ}}$), полностью входит (рис. 5.9) в множество дескрипторов, содержащихся в ПОД ($M_{\text{ПОД}}$), или совпадает с $M_{\text{ПОД}}$, т. е. $M_{\text{ПОЗ}} \subseteq M_{\text{ПОД}}$.

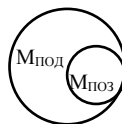


Рис. 5.9

* «На частичное вхождение» ПОЗ в ПОД (пересечение ПОД и ПОЗ).

Документ выдается, если ПОД и ПОЗ совпадают частично, т. е. если часть дескрипторов, содержащихся в $M_{\text{ПОД}}$, совпадает с дескрипторами, входящими в $M_{\text{ПОЗ}}$ (рис. 5.10): $M_{\text{ПОЗ}} \cap M_{\text{ПОД}}$.



Рис. 5.10

* КСС с учетом базисных отношений.

Документ выдается в том случае, если для каждого дескриптора запроса в ПОД встретился либо сам дескриптор, либо дескриптор, связанный с исходным базисными отношениями.

* КСС с учетом текстуальных и базисных отношений.

Различие с предыдущим заключается в том, что сравнение дескрипторов ПОЗ и ПОД должно осуществляться с точностью до совпадения текстуальных отношений, в которые их прообразы вступают соответственно в запросе и документе.

* КСС с учетом весовых коэффициентов информативных слов или дескрипторов.

Каждому информативному слову в запросе приписывается весовой коэффициент (W_i). Весовые коэффициенты в ПОЗ определяются пользователем и нормируются. Сумма всех весовых коэффициентов в запросе должна быть константой ($\sum W_i = \text{Const}$). Выдача эшелонируется в зависимости от суммы весовых коэффициентов слов запроса, совпавших со словами, употребляемыми в документе. Количество эшелонов выдачи, а также соответствующие каждому из них суммы весовых коэффициентов (порог) определяются разработчиком системы в процессе ее отладки.

* КСС с учетом синтаксических отношений.

Вводятся правила грамматики и сопоставляются синтагмы, формируемые из дескрипторов (или ключевых слов) с помощью введенных правил.

5.8. Критерии оценки качества поиска и информационно-поисковых систем

В теории информационного поиска предлагаются и используются различные критерии оценки качества информационно-поисковой системы.

Разработка комплекса критериев оценки качества информационного поиска – достаточно сложная проблема: состав и количественные характеристики критериев зависят от конкретного назначения и принципов реализации ИПС.

Оценка или метод оценки – это алгоритмическая процедура, которая любому оцениваемому объекту изданной области ставит в соответствие некоторый другой объект, называемый значением оценки. Полностью алгоритмическую процедуру оценки называют формальной оценкой.

Различают два типа оценок:

- * оценки-описания, значения которых характеризуют непосредственно систему безотносительно к другим системам;

- * оценки-шкалы, значения которых определяют сравнительные достоинства различных поисковых систем.

От «оценки-описания» требуется, чтобы ее значения позволяли достаточно полно судить о существенных свойствах оцениваемых объектов, например, предсказывать их поведение в тех или иных конкретных условиях. В этом случае «оценка-описание» называется эффективной.

От «оценки-шкалы» требуется, чтобы ее значения упорядочивали множество оцениваемых объектов, например, различных ИПС, не вступая при этом в противоречие с существующими у нас содержательными представлениями о сравнительных достоинствах этих объектов. В этом случае «оценка-шкала» называется здоровой.

Содержательные представления о сравнительных достоинствах систем являются содержательной оценкой. Объективная формальная оценка не должна противоречить содержательной.

Следует иметь в виду, что одна и та же формальная оценка может рассматриваться и как «оценка-шкала», и как «оценка-описание».

Содержательная оценка подразумевает оценку полезности информации для потребителя, для результатов его основной деятельности. При этом оценка эффективности получаемой информации подразумевает оценку ее полезности и затрат на ее получение. Кроме того, для строгой оценки необходимо выделить долю результата, которая получена именно благодаря полученной информации, что крайне затруднено.

С учетом сказанного вместо оценки эффективности поиска ограничиваются оценкой функциональной эффективности.

Оценки поисковых систем делят на два класса, которые называются *внешними* (или функциональными) и *внутренними* оценками.

Внешние, или *функциональные*, оценки основаны на сравнении результатов работы системы с результатами идеального содержательного поиска, осуществляемого экспертом. В теории *информационного поиска* для этого введены понятия *релевантности* и *пертинентности*.

Первоначально в классической теории информационного поиска (см. раздел 5.3) под *релевантностью* понимали соответствие выдачи запросу, т. е. оценку качества алгоритма поиска. Под *пертинентностью* – соответствие выдачи потребностям лица (или лиц), для которого (которых) осуществляется поиск информации, т. е. пертинентность – характеристика смысловыражающих возможностей ИПЯ, точность отображения с его помощью информационных потребностей.

В настоящее время иногда ¹²⁵ термин *релевантность* используют в более широком смысле, и различают релевантность первого рода (формальную релевантность), которая соответствует термину, первоначально введенному в теории информационного поиска, и релевантность второго рода, соответствующую понятию *пертинентности*.

¹²⁵ *Максимович Г. Ю.* Информационные системы / Г. Ю. Максимович, А. Г. Романенко, О. Ф. Самойлюк. – М.: Изд-во Рос. экон. акад., 1999. – 198 с.

Для оценки релевантности используют такие критерии как *полнота*, *точность* поиска, *потери*, *шум*, которые могут быть представлены в виде различных соотношений.

В качестве критериев оценки качества информационного поиска вводится понятие коэффициента корреляции поиска¹²⁴:

$$r = \frac{ad - bc}{\sqrt{(a+c)(b+d)(d+c)(a+b)}}, \quad -1 \leq r \leq 1.$$

Кроме того, к числу показателей функциональной эффективности **Г. Ю. Максимович, А. Г. Романенко, О. Ф. Самойлюк**¹²⁴ относят *оперативность* поиска; *специфичность* поиска – отношение числа невыданных нерелевантных документов (d) к общему числу нерелевантных документов ($d + b$), где b – число выданных нерелевантных документов:

$$C = \frac{d}{d+b}; \quad 0 \leq C \leq 1.$$

На практике при оценке ИПС с большими массивами информации точные измерения числа релевантных и нерелевантных документов в общем массиве или в массивах выданных документов затруднено. Поэтому могут использоваться энтропийные показатели

Энтропийные меры могут быть получены на основе исследования выборки из информационного массива, т.е. могут использоваться вероятностные меры неопределенности исходного массива p_0 , массива выданных p_1 , и массива невыданных p_2 документов, вычисленные на их основе H_0, H_b, H_{nb} и соответствующие меры W .

В частности, предлагается интегральный энтропийный показатель как мера упорядоченности поискового массива документов, являющаяся результатом процесса поиска по заданному запросу:

$$W = \frac{H_0 - H_n}{H_0},$$

где H_0 – допоисковая (априорная) энтропия, H_n – послепоисковая (апостериорная) энтропия.

При этом для измерения априорной и апостериорной энтропии предлагается использовать меры концентрации релевантных документов в общем массиве, в массиве выданных документов и в массиве невыданных документов, в относительных единицах:

$$P_0 = \frac{a + z}{a + g + z + d}; \quad P_1 = \frac{a}{a + g}; \quad P_2 = \frac{z}{z + d},$$

где a – число релевантных документов, выданных в результате поиска; z – число релевантных документов, не выданных в результате поиска; g – число нерелевантных документов, выданных в результате поиска; d – число нерелевантных документов, невыданных в результате поиска.

Энтропийная мера может быть представлена и в логарифмической форме.

Например, для оценки энтропии используются натуральная логарифмическая мера (т. е. неопределенность измеряется в неперах):

неопределенность исходного массива:

$$H_0 = [P_0 \ln P_0 + (1 - P_0) \ln (1 - P_0)];$$

неопределенность массива выданных документов:

$$H_b = [P_1 \ln P_1 + (1 - P_1) \ln (1 - P_1)];$$

неопределенность массива невыданных документов:

$$H_{nb} = [P_2 \ln P_2 + (1 - P_2) \ln (1 - P_2)].$$

При формировании подобных мер могут быть использованы и иные логарифмические шкалы: двоичные логарифмы (биты), восьмеричные логарифмы (байты), десятичные логарифмы.

Внутренние оценки основываются на таких структурных качествах системы как *сложность*, *степень близости к человеческой логике или естественному языку*, *степень алгоритмичности*, на оценке компонентов ИПС, и в частности информационно-поискового языка (ИПЯ) и т. п.

Например, **Ч. Мудой** [8] предлагает оценивать качество информационно-поискового языка, используя следующие критерии: *семантическая сила* (expressiveness), *многозначность* и *компактность* языка, *стоимость выбора термина*.

Семантическая сила – это способность языка идентифицировать объект, различать мелкие особенности объектов, описывать объект с разной степенью детализации.

Подразумеваются потенциальные возможности ИПЯ, а не умение им пользоваться. Самой большой семантической силой обладает естественный язык.

Многозначность означает, что слово или синтаксическая единица тезауруса имеет более чем одно значение (омографы) или, напротив, что некоторое значение может иметь более одного символического представления в словарном составе ИПЯ (синонимия). Кроме того, одинаково звучащие слова могут иметь различные значения (полисемия или омонимия).

Синонимия и омография могут существовать и в синтаксических единицах, состоящих из нескольких слов.

Компактность характеризуют физический размер или длину терминов словаря или поисковых образов, составленных их числа терминов, необходимых для отображения смысла документов и запросов.

Стоимость характеризует цену процесса принятия решения по выбору терминов (ключевых слов, дескрипторов или иных синтаксических единиц) для отображения смысла документа или запроса.

В общую стоимость входят: *стоимость обучения пользованию языка, стоимость составления и совершенствования словаря, затраты, связанные с устранением ошибок*, допущенных при выборе терминов, *затраты времени на индексирование документов и составление ПОЗ*.

Предлагаемые **Ч. Мидоу** оценки не являются независимыми и взаимоисключающими. ИПЯ может быть семантически сильным, но многозначным.

Компактность слов в словарном составе языка не определяет стоимости, т. е. затрат времени, труда на выбор терминов.

ИПЯ характеризуют также *словарным составом* и наличием *грамматики*. При наличии тезауруса ИПЯ можно охарактеризовать глубиной *тезауруса*, т. е. количеством уровней, видов смы-

словыражающих элементов или синтаксических единиц тезауруса. Характеристики ИПЯ являются *внутренними* оценками информационно-поисковой системы, влияющими на оценку качества информационного поиска, по критерию пертинентности.

Количественные оценки релевантности

Формальные оценки релевантности можно получить только для релевантности первого рода, т. е. для релевантности в исходном ее понимании в теории информационного поиска.

Для введения критерия релевантности следует задать процедуру определения меры семантической близости поискового образа документа поисковому образу запроса и некоторое пороговое значение этой меры. Если мера превышает пороговое значение, то документ релевантен запросу.

ПОД и ПОЗ представляют собой множества ключевых слов или дескрипторов в зависимости от вида информационно-поискового языка. Для их сопоставления используют критерии смыслового соответствия, которые определяют на основе совпадения ключевых слов (дескрипторов) в ПОД и ПОЗ.

ПОД и ПОЗ можно представить в виде четких и нечетких множеств.

Для четких множеств вводят нормированную меру релевантности:

$$0 \leq \mu_R(a, b) \leq 1.$$

Критерий смыслового соответствия можно представить в виде:

$$r_a = \langle \mu_R(a, b) \geq \alpha \rangle,$$

где $\mu_R(a, b)$ – функция вычисления меры релевантности (или просто мера релевантности); α – пороговое значение релевантности, такое, что

$$r_a = \begin{cases} 1 & \text{при } \mu(a, b) \geq \alpha \text{ документ выдается} \\ 0 & \text{при } \mu_R(a, b) < \alpha \text{ документ не выдается} \end{cases}$$

Изменяя пороговое значение α , можно организовать выдачу различных совокупностей документов, которую в теории информационного поиска называют *эшелонированной выдачей*. Каждый эшелон соответствует определенной мере семантической близости совокупности документов запросу.

Очевидно, что чем больше пороговое значение α , тем более жесткие условия налагаются на смысловую близость документа запросу. В нормированных мерах при $\alpha = 1$ для выдачи документа требуется полное совпадение ПОД и ПОЗ.

Оценку релевантности можно характеризовать *полнотой* выдачи (или *потерями*), т. е. числом невыданных релевантных документов), и *точностью* (или *шумом*), т. е. числом или процентом «лишних» документов, которые выданы в результате поиска, но не являются релевантными.

Например, в [17] оценки полноты R и точности и T вводятся следующим образом:

$$R_i = \frac{a_i}{n_i}, \quad T_i = \frac{a_i}{m_i}, \quad 0 \leq R_i \leq 1, \quad 0 \leq T_i \leq 1,$$

где a_i – число релевантных документов, формально выданных системой на i -й запрос; m_i – число всех формально выданных на i -запрос системой документов; n_i – число всех релевантных документов, соответствующих запросу.

При этом рекомендуется полноту и точность определять на основе нескольких поисков N по запросу, и определять их средние значения и суммарные относительные оценки:

$$R = \frac{\sum_{i=1}^N \frac{a_i}{n_i}}{N}, \quad T = \frac{\sum_{i=1}^N \frac{a_i}{m_i}}{N} \quad \text{– средние относительные оценки;}$$

$$R = \frac{\sum_{i=1}^N a_i}{\sum_{i=1}^N n_i}, \quad T = \frac{\sum_{i=1}^N a_i}{\sum_{i=1}^N m_i} \quad \text{– суммарные относительные оценки,}$$

где N – число поисков.

Величины $(1-T)$ и $(1-R)$ называются соответственно шумом и потерями.

В работах *А. И. Михайлова, А. И. Черногo и Р. С. Гиляревскогo* [9, с. 306] предлагается наглядная матрица для определения полноты и потерь, точности и шума (табл. 5.4).

Таблица 5.4.

Матрица для определения критериев релевантности

		Релевантны	Нерелевантны		
		A_1	A_2		
Выдано	B_1	a	b	$a+b$	
Не выдано	B_2	c	d	$c+d$	
		$a+c$	$b+d$	$a+b+c+d$	

Полноту поиска измеряют отношением числа выданных релевантных документов (a) к общему числу релевантных документов массива ($a+c$):

$$R = \frac{a}{a+c}, \quad 0 \leq R \leq 1.$$

Точность поиска T – отношение числа выданных релевантных документов (a) к числу общему выданных документов ($a+b$)

$$T = \frac{a}{a+b}, \quad 0 \leq T \leq 1.$$

Соответственно потери L и шум S можно представить следующим образом:

$$L = \frac{c}{a+c} \quad \text{и} \quad S = \frac{b}{a+b}, \quad 0 \leq L \leq 1, \quad 0 \leq S \leq 1.$$

В [17] предлагается, проведя серию экспериментов n по определению полноты и точности поиска, определить среднюю полноту и среднюю точность:

$$P_{cp} = \frac{1}{n} \sum_{i=1}^m \frac{a_i}{a_i + b_i}; \quad T_{cp} = \frac{1}{n} \sum_{i=1}^m \frac{a_i}{a_i + c_i}.$$

Используются и иные способы усреднения (см., напр., [9, 17]).

Например, в связи с оценкой системы СМАРТ Сэлтон¹ вводит нормированную полноту R_N и нормированную точность P_N

$$R_N = \frac{1}{N} \sum_{i=1}^N \frac{n_i}{n},$$

где N – число документов в массиве; n – число всех релевантных документов в массиве; n_i – число релевантных документов, выданных до i -го ранга включительно,

$$P_N = \frac{1}{N} \sum_{i=1}^N \frac{n_i}{i},$$

где i – номер ранга.

В [9] предлагается также, пользуясь табл. 5.4, ввести (с определенной степенью приближения) показатели в терминах теории вероятностей:

$\frac{a}{a+c} = P_{B1}(A_1)$ – условная вероятность выдачи релевантных документов;

$\frac{b}{b+d} = P_{B2}(A_1)$ – условная вероятность выдачи нерелевантных документов;

$\frac{c}{a+c} = P_{B1}(A_2)$ – условная вероятность невыдачи релевантных документов;

$\frac{d}{b+d} = P_{B2}(A_2)$ – условная вероятность невыдачи нерелевантных документов.

$$P_{B1}(A) + P_{B1}(A_2) = 1$$

$$P_{B2}(A) + P_{B2}(A_2) = 1$$

Предлагались и другие способы оценки релевантности (см. обзор в [9])

Со способами применения для оценки релевантности нечетких множеств можно познакомиться в работах **Г. Ю. Максимо-вича, А. Г. Романенко, О. Ф. Самойлюк**¹²⁶.

¹ Сэлтон Г. Автоматическая обработка, хранение и поиск информации / Г. Сэлтон. – М.: Сов. радио, 1973. – 560 с.

Определение полноты системы связано с определением содержательной выдачи на каждый запрос.

Существует несколько способов определения этой выдачи:

* **Сплошной просмотр всего экспериментального массива.**

Достоинством этого способа является надежность, недостатком – трудоемкость.

* **Метод документа-источника («метод Клевердона»).**

Состоит в том, что по некоторым документам массива, выбранным более или менее случайно, составляются запросы с таким расчетом, чтобы каждый документ-источник был релевантен составленному по нему запросу.

* **Метод контрольных документов.**

По запросу, полученному по произвольно выбранному документу-источнику, проводится содержательный поиск путем сплошного просмотра массива, начиная, например, с документа-источника, до нахождения первого релевантного документа, который объявляется контрольным. Значение полноты для системы считается теперь как доля запросов, по которым система выдала контрольный документ в общем количестве запросов.

* **Метод объединения формальных выдач.**

Применяется при сравнении нескольких поисковых систем («оценка-шкала»). Он состоит в том, что по каждому запросу эксперт просматривает только те документы, которые выдавались хотя бы одной из этих поисковых систем. Содержательной выдачей считается совокупность обнаруженных релевантных документов, и относительно нее определяется полнота, которая отличается от истинной полноты каждой из рассматриваемых систем.

Вопрос о представительности массива документов и массива запросов, выбранных для определения формальных оценок, в общем виде не решен. Считается, что более или менее устойчивые оценки (колебания не превышают 5 %) можно получить на массиве в 4000 документов, массив запросов при этом должен быть порядка нескольких сотен.

Таким образом, релевантность (формальная) характеризует свойства средств логико-семантического аппарата информационно-поисковой системы и зависит от возможности отображения ПОД и ПОЗ с помощью информационно-поискового языка, принятых в ИПС алгоритмов поиска и системы индексирования.

¹²⁶ *Максимович Г. Ю.* Информационные системы / Г. Ю. Максимович, А. Г. Романенко, О. Ф. Самойлюк. – М.: Изд-во Рос. экон. акад., 1999. – 198 с.

5.9. Разработка и отладка ИПС

При разработке ИПС обычно выполняются следующие этапы:

1) анализ совокупности документов, представляющих соответствующую научно-техническую область, для которой создается ИПС; выбор ключевых слов (КС), характеризующих содержание каждого документа; КС выбирают из заголовков (или из аннотаций) документов, включаемых в информационный массив (информационную базу) ИПС. Ключевые слова – существительные или отглагольные существительные, взятые в единственном числе и именительном падеже, прилагательные, глаголы в неопределенной форме;

2) формирование из ключевых слов словаря (первой стадии разработки информационно-поискового языка); ИПЯ является основой для формирования ПОД, а в последующем ПОЗ.

3) выбор системы индексирования документов; обычно вначале выбирается индексирование по алфавиту; а в последующем могут быть выбраны системы индексирования по значимости или с использованием статистического словаря.

4) формирование ПОД каждого документа с использованием словаря; первоначально при выборе системы индексирования по алфавиту слова в ПОД располагаются по алфавиту; а в дальнейшем в процессе отладки системы может быть выбран иной способ индексирования.

Примечание. Возможно свободное индексирование, без словаря (система Унитерм); тогда будет отсутствовать п. 2.

При формировании информационного массива в структуру записи (сложного термина – при использовании языка логического программирования¹²⁷) включают номер документа, наименование документа, фактографические сведения (дата его издания, место издания, автор и т. п., в зависимости от характера информации), фасета из КС ПОЗ. ПОЗ может описываться отдельно, тогда в него помимо КС включается номер документа.

¹²⁷ *Анастасиади П. Г.* Турбо-Пролог и его применение для создания информационно-поисковых систем: учеб. пособие / П. Г. Анастасиади, В. Н. Волкова. – СПб.: Изд-во Политехн. ун-та, 2007. – 56 с.

5) формируется ПОЗ или поисковое предписание; при применении ИПЯ – с его использованием;

6) проводится отладка ИПС; в процессе которой оценивается релевантность выдачи и исследуется несколько способов повышения релевантности:

- изменение алгоритма поиска и/или КСС;
- изменение способа индексирования (статистический, с учетом значимости КС для отображения смысла документа или запроса, т. е. с учетом весовых коэффициентов ключевых слов в ПОД и ПОЗ);
- развитие ИПЯ: внесение изменений в словарь путем введения классов условной эквивалентности, дескрипторов или иной способ изменения словаря; введение грамматики.

Изложенная методика разработки и отладки ИПС представляется в виде структурной схемы типа приведенной на рис. 5.11.

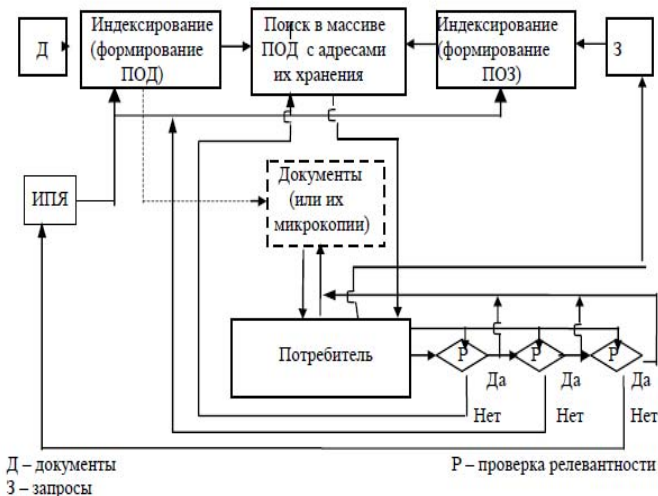


Рис.5.11. Схема отладки ИПС

С применением рассмотренной методики студенты выполняют лабораторную работу с использованием языка логического программирования Турбо-Пролог.

Вопросы для самоконтроля

1. Понятие о документальном информационном поиске и информационно-поисковой системе.
2. Поисковый образ документа и поисковый образ запроса.
3. Документальные информационно-поисковые системы.
4. Структура документальной информационно-поисковой системы.
5. Логико-семантический аппарат.
6. Информационно-поисковый язык.
7. Классификации ИПЯ. Предкоординированные и посткоординированные (кординатные) ИПЯ.
8. Deskрипторные ИПЯ без грамматики и с грамматикой.
9. Ключевое слово. Deskриптор.
10. Понятие о тезаурусе.
11. Система индексирования.
12. Логика ИПС.
13. Критерий смыслового соответствия.
14. Парадигматические и синтагматические отношения.
15. Документальные информационно-поисковые системы научно-технической информации.
16. Структура ГАС НТИ .
17. Виды источников научно-технической информации.
18. Режимы информационного обслуживания.
19. Оценки информационно-поисковых систем.
20. Понятие о релевантности и пертинентности.
21. Понятие об информационно-логических и информационно-семантических системах.

Лабораторный практикум

Разработка и отладка документальной ИПС без грамматики и с грамматикой с использованием методики, изложенной в разделе 5.9, и одного из языков логического программирования (например, языка Турбо-Пролог) включая формирование структуры информационного массива, ИПЯ ключевых слов и ИПЯ deskрипторного типа, алгоритмы (в языке Турбо-Пролог – сложные термы) реализации информационного поиска¹²⁸.

¹²⁸ *Анастасиади П. Г.* Турбо-Пролог и его применение для создания информационно-поисковых систем: учеб. пособие / П. Г. Анастасиади, В. Н. Волкова. – СПб.: Изд-во Политехн. ун-та, 2007. – 56 с.; Доорс Дж. Пролог – язык программирования будущего / Дж. Доорс, А. Р. Рейблейн, С. Вадера / Пер. с англ. А. Н. Волкова. – М.: Финансы и статистика, 1990. – 144 с. (Prolog – Progrzmming for Tomorrow by J. Doors, A. R. Reblein & S. Vadera. – Signa Press *Winstow).

Глава 6. ПРИНЦИПЫ ПОСТРОЕНИЯ И ПРОЕКТИРОВАНИЯ ДОКУМЕНТАЛЬНО-ФАКТОГРАФИЧЕСКИХ ИНФОРМАЦИОННО-ПОИСКОВЫХ СИСТЕМ

6.1. Понятие о документально-фактографических информационно-поисковых системах

Документально-фактографическая информационно-поисковая система (ДФИПС) представляет собой ИПС, предназначенную для выдачи непосредственно требуемой информации (например, температуры кипения конкретной жидкости, статистических показателей и иных конкретных сведений, содержащихся в соответствующих отчетных документах, публикациях и т. п. документальных информационных массивах, и др.).

Существуют ДФИПС двух видов:

- * *документальные* и сопряженные с ними массивы *фактографической* информации, которые формируются параллельно;
- * *информационно-логические системы*, в которых фактографическая информация извлекается из документальной в процессе поиска.

Первоначально разрабатывались документально-фактографические системы, в которых массивы фактографической информации формировались параллельно с массивами документальной информации.

Наиболее известными примерами таких систем являются ДФИПС «Фтор» и «Спектр» [9, 10].

АДФИПС «Фтор». Была разработана в ВИНТИ в 1963 г. и предназначена для оперативного оповещения химиков о новых публикациях по фторорганическим соединениям (документальная ИПС), а также для быстрого поиска сведений о химических соединениях и реакциях (фактографический режим работы ИПС).

Для реализации фактографического режима формируются фактографические массивы, в которых записываются:

- * реакция или соединение, краткие сведения о них, содержащиеся в публикации;

- * дополнительные сведения (механизм и кинетика реакции, спектральные данные, физические и физико-химические свойства и т. д.);

- * выходные данные публикации-источника.

Для описания документов разработан специальный ИПЯ, предназначенный для исчерпывающего описания реакций и свойств фторорганических соединений.

В АДФИПС «Фтор» предусмотрена возможность работы в следующих режимах обслуживания: ИРИ, ретроспективный поиск; издание пермутационных библиографических указателей по тематике ИПС; изготовление копий публикаций-источников по заказам подписчиков.

АДФИПС «Спектр». Разработана в 1970 г. группой сотрудников Новосибирского отделения органической химии и Вычислительного центра Сибирского отделения Академии наук СССР. Первоначально эта АДФИПС предназначалась для идентификации органических соединений по их инфракрасным спектрам.

Для этого разработан специальный вариант кодирования инфракрасного спектра, поскольку АДФИПС предназначалась не только для опознавания веществ по их инфракрасным спектрам, но и для выявления различных корреляционных связей. В АДФИПС предусмотрен режим стандартных запросов. При этом возможен поиск одновременно по 4-м запросам, а в последующем была создана подсистема АДФИПС для электронной спектроскопии, в которой возможен поиск одновременно по 8 запросам.

В последующем стали понимать документально-фактографический поиск как процесс автоматизации извлечения и логического преобразования фактографических данных с целью получения новой информации.

Такие ИС являются одним из направлений развития информационно-логических систем. При разработке таких ИЛС необхо-

димо проводить анализ системы понятий и методов рассуждений, принятых в соответствующей отрасли знаний, разрабатывать более сложные ИПЯ с правилами логического вывода. ИС такого рода базируются на применении методов дискретной математики и искусственного интеллекта.

Документально-фактографические ИПС особого типа составляют основу *Автоматизированных систем нормативно-методического обеспечения управления* (АСНМОУ) предприятием (организацией) и Информационных систем нормативно-правовой документов.

6.2. Автоматизированные системы нормативно-методического обеспечения управления предприятиями и организациями

Автоматизированная система нормативно-методического обеспечения управления (АСНМОУ) предприятием (организацией) должна содержать нормативно-правовые, нормативно-методические, нормативно-технические и организационно-распорядительные документы (НПД, НМД, НТД и ОРД), которые обеспечивают реализацию принятых проектных и управленческих решений в процессе функционирования предприятия (организации).

Пример структуры СНМОУ и АСНМОУ, соответствующей основным функциям, которые должна выполнять СНМОУ, приведен на рис. 6.1¹²⁹.

В качестве методической основы создания АСНМОУ используют идею стратифицированного представления процедуры поиска информации с углублением на каждой ступени анализа документов, содержащихся в АСНМОУ, путем структуризации их текстов (рис. 6.2):

¹²⁹ Чудесова Г. П. Преобразование организационной структуры при изменении формы собственности предприятия / Г. П. Чудесова. – СПб.: Изд-во СПбГТУ, 1995. – 190 с.

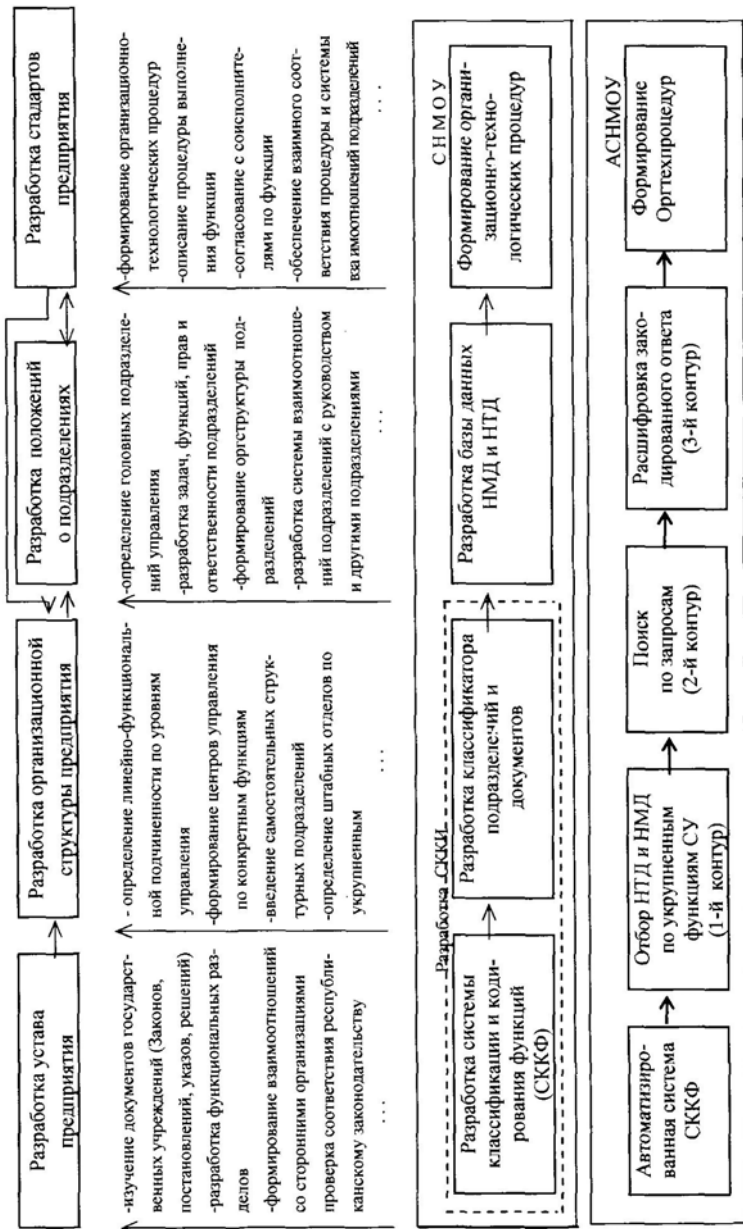


Рис. 6.1. Функции и структура СНМОУ и АСНМОУ

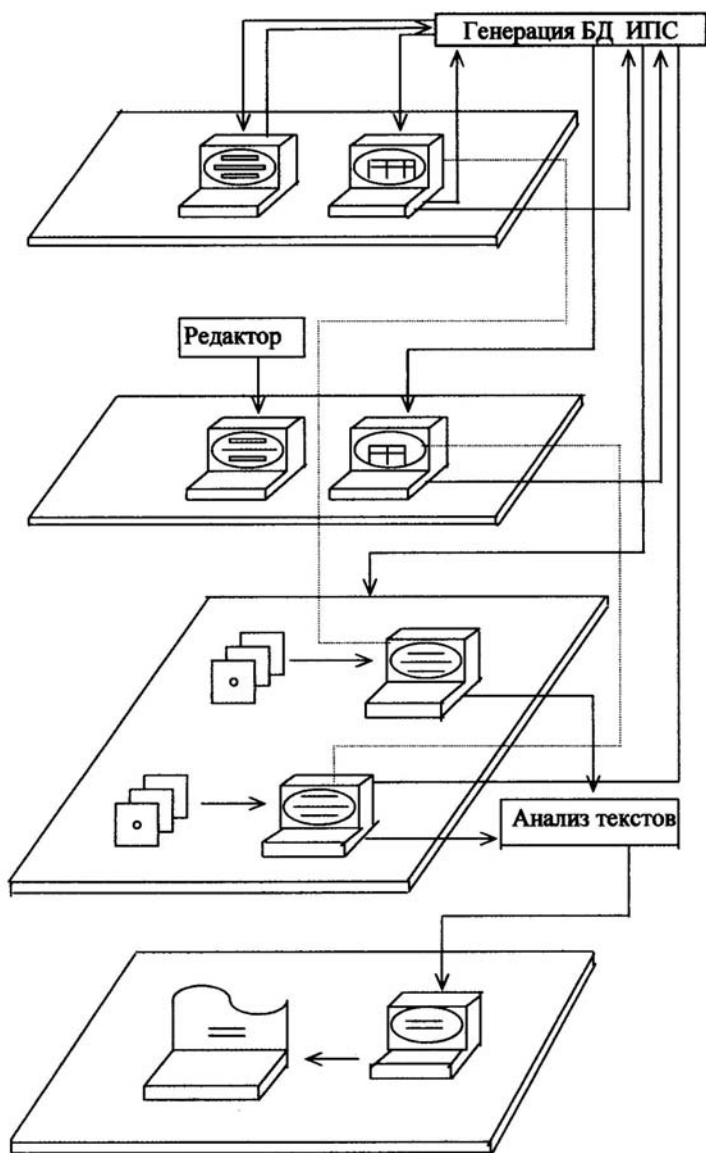


Рис. 6.2. Принципы стратифицированного построения АСНМОУ

- на верхней страте осуществляется поиск документов по функциям управления;
- на второй сверху страте – поиск разделов документов в соответствии с запросом пользователя, взаимоотношений между разделами связанных друг с другом документов;
- на третьей страте – вывод текстов на дисплей или принтер (полного текста или его разделов);
- на четвертой страте (которая реализуется не для всех документов) – аналитико-синтетическая обработка (АСО) текстов документов (что, например, требуется при поиске информации в текстах законов и других НПД).

Виды баз данных определяются спецификой конкретного предприятия (организации).

Например, базы данных можно выделить таким образом: создать базу данных НПД (общегосударственных, региональных), базу данных об НМД, НТД и ОРД органов отраслевого управления и предприятия. Для АСНМОУ предприятия может оказаться целесообразным создать отдельные базы СТП, должностных инструкций, положений о подразделениях и т. п. Базы данных для крупных предприятий и организаций имеют значительный объем. Для организации эффективных процедур поиска и корректировки информации, их нужно структурировать. Исследование особенностей БД АСНМОУ и работы с ними показало, что выбрать наиболее целесообразную жесткую структуру БД практически невозможно, так как, с одной стороны, подразделения предприятия принимают участие в выполнении нескольких укрупненных функций управления, а, с другой стороны, одна и та же функция выполняется несколькими подразделениями.

Кроме того, одна и та же функция регламентируется в документах разного вида – и в положениях о подразделениях, и в СТП, и в ОРД, и т. п. При этом одинаковые или сходные с различной степенью детализации и с несколько модифицированными формулировками функции в разных документах заводи-

рованы по-разному, в соответствии с группированием и индексированием функций, принимаемым при разработке документа его авторами.

Используя терминологию теории научно-технической информации, удобно говорить не о создании баз данных для различных страт, а о создании двух-, трех- или более контурных АДФИПС.

В частности, при создании БД нормативно-правовых документов исследовалось два варианта АДФИПС: двухконтурный – с поиском документов в *первом* контуре и извлечением из них разделов (которые содержат сведения о правовых нормах, представляющих собой фактографическую информацию для пользователя) во *втором*; и трехконтурный с поиском в контурах последовательно в первом – документов, во втором – разделов, и в третьем – статей (содержащих еще более конкретную фактографическую информацию о правовых нормах).

Аналогично могут быть организованы БД нормативно-методических документов.

Например, при создании АДФИПС положений о подразделениях предприятия или БД НМД, НТД и ОРД могут быть образованы следующие контуры: 1) отбор положений или других НМД, НТД, ОРД, соответствующих запросу (т. е. структуризация под запрос – например, по укрупненным функциям, по группам подразделений); 2) поиск в базе отобранных документов по запросам пользователя необходимых ему сведений о функциях, сроках и способах их выполнения и т. п. фактографической информации (поиск может осуществляться по признакам, предусмотренным при разработке АДФИПС); 3) вывод фрагментов документов, отобранных в соответствии с запросами во втором контуре, на дисплей или принтер в удобной для пользователя форме (например, представления разделов положений, содержащих необходимые сведения о функциях, сроках их выполнения, исполнителях и т. п., в виде таблицы).

Пример трехконтурной АДФИПС нормативно-методических документов АСНМОУ (положений о подразделениях организации предприятия), приведен на рис. 6.3¹³⁰.

¹³⁰ Чудесова Г. П. Преобразование организационной структуры при изменении формы собственности предприятия / Г. П. Чудесова. – СПб.: Изд-во СПбГТУ, 1995. – 190 с.

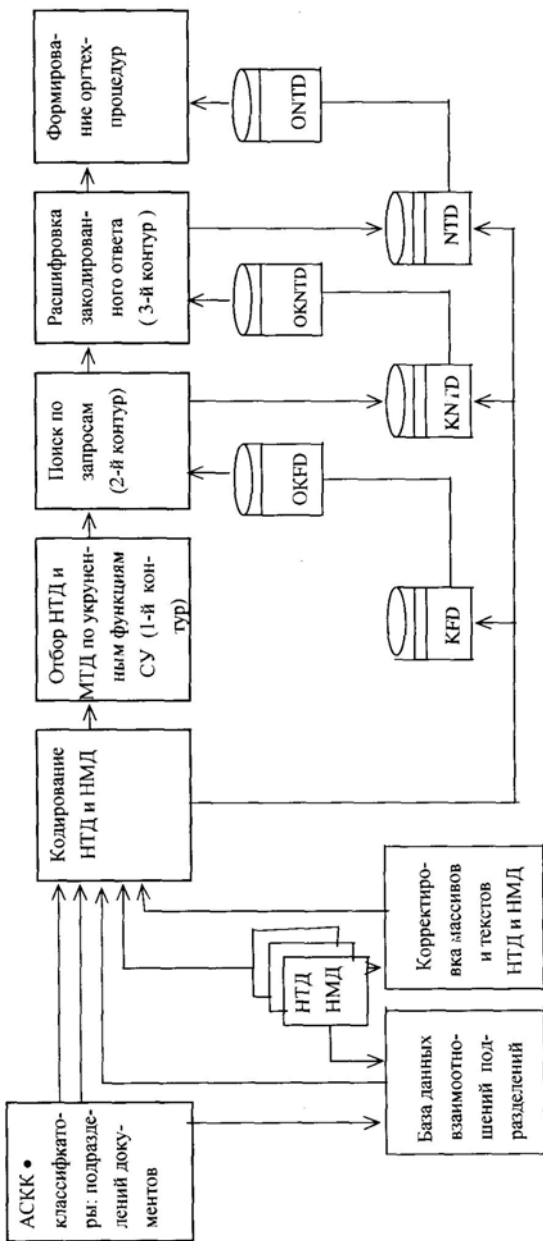


Рис. 6.3. Структура трехконтурной АДПИПС

Структуры БД в этом примере имеют следующий вид (в приводимом примере принята фасетная классификация).

Структура первого контура БД положений о подразделениях:

$$nmo = \langle N1, N2, F1, F2, F3 \rangle,$$

где $N1$ – номер отдела в существующей СНМОУ; $N2$ – наименование отдела; $F1, F2, F3$ – индексы укрупненной функции в структуре целей и функций предприятия, выполнение которой регламентируются в положении.

Структура БД раздела «Взаимосвязь» положений о подразделениях предприятия (второго контура АДФИПС), составляющими которой являются функции положения:

$$nmo = \langle F1, F2, F3, NI, DI, FJ, DK, NP \rangle,$$

где $F1, F2, F3$ – индексы укрупненной функции в структуре целей и функций предприятия, выполнение которой регламентирует подготавливаемый документ; NI – номер отдела-изготовителя документа в существующей СНМОУ; DI – наименование вида исходного документа (или конкретного документа), для подготовки которого выполняется функция; FJ – индекс функции внутри положения; DK – индекс документа, получаемого в результате выполнения функции; NP – номер отдела, в который передается документ для дальнейшей обработки (согласования). Поиск фактографической информации в этом контуре может осуществляться по любому параметру.

Третий контур – вывод соответствующих отобраным кодам текстов.

В приведенном примере АДФИПС предусмотрен также блок формирования и анализа организационно-технологических процедур (ОТП) подготовки и реализации управленческих решений (подробнее характеризуемый, например, в [1, 13, 16]).

Соответственно – для БД нормативно-правовых документов структуры контуров имеют следующий вид.

Структура БД НПД первого контура:

$$npd = \langle N, T, D, M, F1, F2, F3, IF \rangle,$$

где N – номер документа в БД НПД; T – тип документа (закон, постановление и т. п.); D – наименование документа; M – место опубликования или хранения документа; $F1, F2, F3$ – индекс укрупненной функции, регламентируемой документом; IF – имя файла, в котором хранится текст документа (или фрагменты текста).

Структура БД НПД второго контура:

$$npd = \langle N, D, R, F1, F2, F3, F4, F5, F6, IFR \rangle,$$

где N – номер документа в БД НПД; D – наименование документа (закона, постановления, и т. п.); R – наименование раздела документа; $F1, F2, F3, F4, F5, F6$ – индекс функции, регламентируемой разделом; IFR – имя файла, в котором хранится раздел текста документа.

При необходимости может быть сформирован контур для статей НПД или контур аналитико-синтетической обработки информации.

Для ускорения разработки АСНМОУ для конкретных предприятий и организаций разрабатываются системы ее генерации АДФИПС, с примерами которых можно познакомиться в [1, 13].

Приведенные примеры ДФИПС относятся к классу ИПС, в которых фактографический поиск понимается как процесс отыскания уже готовых данных и фактов, извлекаемых из текстов в процессе подготовки и индексирования входных данных и вводимых в массивы ИПС. В этом случае теория и методика разработки таких ИПС практически не отличаются от принципов проектирования чисто документальных ИПС, и нередко ДИПС развивались в ДФИПС.

Вопросы для самоконтроля

1. Документально-фактографические информационные системы. Примеры.
2. Структура системы нормативно-методического обеспечения управления предприятием.

Практикум

Разработка методики и программы формирования вариантов организационно-технологических процедур (ОТП) подготовки и реализации управленческих решений (с применением правил формирования синтагм или правил типа «условного следования за»).

Глава 7. ПЕРСПЕКТИВНЫЕ НАПРАВЛЕНИЯ РАЗВИТИЯ ТЕОРИИ ИНФОРМАЦИОННЫХ КОМПЛЕКСОВ

7.1. Концепция архитектуры предприятия и ее применение для развития теории информационных систем

Предпосылки возникновения понятия «архитектура предприятия»

В настоящее время все больше внимания уделяют новому направлению, которое стали называть *архитектурой предприятия*.

Это направление изначально предназначалось для решения двух основных проблем в сфере информационных технологий (ИТ), выявленных двадцать лет назад.

Первая проблема заключается в постоянном увеличении сложности ИТ-систем, что приводило к увеличению затрат на их построение.

Вторая вызвана тем, что со временем получить реальную отдачу от ИТ-систем становится труднее, т. е. несмотря на всевозрастающую стоимость ИТ-систем, организациям с большим трудом удавалось поддерживать их соответствие требованиям бизнеса.

Очевидно, эти проблемы взаимосвязаны. Чем сложнее система, тем труднее ею управлять. Чем лучше удастся справиться со сложностью системы, тем выше вероятность получения от нее реальной выгоды.

В итоге – высокие затраты, низкая эффективность. Эти проблемы, впервые выявленные двадцать лет назад, постепенно достигли критической точки. Стоимость и сложность ИТ-систем выросли экспоненциально, а реальная польза от них резко снизилась.

Для создания простой однопользовательской информационной системы понятие «архитектора» не нужно. Но в процессе создания корпоративной, распределенной системы представление об архитектуре все более усложняется.

Принципы построения архитектуры предприятия, которые двадцать лет назад представлялись странными и далекими от действительности, в настоящее время становятся все более востребованными.

За прошедший период было разработано множество методологий построения архитектуры предприятия.

История развития архитектур предприятия приведена в следующем подразделе и в Приложении 1.

Краткая история развития концепции архитектуры предприятия

Понятие «архитектура предприятия» появилось в 1987 г. в статье *Дж. А. Захмана* «Структура архитектуры информационных систем», опубликованной в журнале «IBM Systems Journal»¹.

Видение Захмана заключалось в том, что для обеспечения высокой ценности и гибкости бизнеса необходим целостный подход к архитектуре систем, в рамках которого каждая существенная проблема рассматривается с разных точек зрения. Такой подход к созданию архитектуры систем представляет собой то, что Захман изначально называл *архитектурной структурой* информационных систем, а впоследствии – *структурой архитектуры предприятия*.

В последующем Захман внес вклад в разработку архитектуры ИС Министерства обороны США. Эта попытка была предпринята в 1994 г., а сама концепция получила название «Базовая архитектура технического обеспечения для управления информацией» (Technical Architecture Framework for Information Management – TAFIM).

Преимущества, обеспечиваемые такими архитектурами предприятий, как TAFIM, были отмечены Конгрессом США. Конгресс в 1996 г. принял закон, известный как акт Клингера – Коэна², который определил реформу управления информационными технологиями.

¹ *Захман Дж.А.* Структура архитектуры информационных систем / Дж. А. Захман // IBM Systems Journal. – 1987. – Том 26. – № 3.

² Акт Клингера – Коэна от 1996 г. (PL 107–347) (см. THOMAS [Библиотека Конгресса]).

В соответствии с этим законом всем федеральным агентствам было предписано принять меры по повышению эффективности инвестиций в ИТ. Для надзора за выполнением закона был сформирован совет директоров по информационным технологиям, в который вошли директора по информационным технологиям из всех основных правительственных органов.

С этого периода началось активное развитие идеи архитектуры предприятий и организаций.

В апреле 1998 г. Совет директоров по информационным технологиям начал работу над первым крупным проектом – структурой архитектуры федеральной организации (Federal Enterprise Architecture Framework – FEAF). Версия 1.1 данной структуры была выпущена в сентябре 1999 г.

В 1998 г., четыре года спустя после разработки TAFIM и два года спустя после оформления этой методологии в виде акта Клингера–Коэна, методология TAFIM была официально отменена Министерством обороны. Все наработки по TAFIM были преобразованы в открытую группу, а затем – в новый стандарт, известный в настоящее время под названием TOGAF (The Open Group Architectural Framework).

Через некоторое время полномочия совета директоров по информационным технологиям по архитектуре федеральной организации были переданы Административно-бюджетному управлению.

В 2002 г. Административно-бюджетное управление переработало методологию FEAF и переименовало ее в архитектуру федеральной организации (Federal Enterprise Architecture – FEA).

Несмотря на весьма продуктивную деятельность федерального правительства по разработке архитектуры предприятия прогресс практически отсутствовал. В 2004 г., восемь лет спустя после принятия акта Клингера–Коэна, предписывающего использовать эффективные процессы планирования ИТ, Бюджетно-контрольное управление представило отчет о внедрении идеи архитектуры.

В 2003 г. Министерством Обороны США была разработана DoDAF (Department of Defence Architecture Framework), а консорциумом The Open Group – методология TOGAF (The Open Group Architecture Framework – структура архитектуры The Open Group)¹³¹.

В 2005 г., примерно в то же время, когда Административно-бюджетное управление стало доминировать в области разработки архитектуры предприятия в государственном секторе экономики, лидирующей в частном секторе экономики стала другая организация – группа Gartner.

Методология Gartner не является ни таксономией (как модель Захмана), ни процессом (как TOGAF), ни полной методологией (как FEA). Эта методология представляет собой набор практических рекомендаций по построению архитектуры предприятия.

Компания Gartner считает, что архитектура предприятия должна базироваться на системно-целевом подходе. На основе этого подхода в 2006 г сформулированы рекомендации, касающиеся разработки архитектуры в виде последовательности шагов и задач участников, которые, однако, не детализированы до уровня моделей процесса разработки архитектуры, а затем разработана

К 2005 г. компания Gartner стала одной из наиболее влиятельных организаций, занимающихся консалтингом на уровне директоров по информационным технологиям. Однако в области разработки архитектуры предприятия наиболее известной ИТ- и консалтинговой компанией была не Gartner, а Meta Group.

В 2005 г. Компания Meta Group предложила Методику META Group, отличительной особенностью которой является более детальное и формализованное описание именно *процесса* разработки архитектуры и всех его составляющих по управлению ИТ- программами и проектами.

¹³¹ www.opengroup.org.

В 2005 г. компания Gartner приобрела компанию Meta Group, и разработала принципы принятия согласованных решений, объединяющих различающиеся подходы.

История возникновения основных видов архитектур приведена на рис. 7.1. Краткая характеристика архитектур – в табл. 7.1.

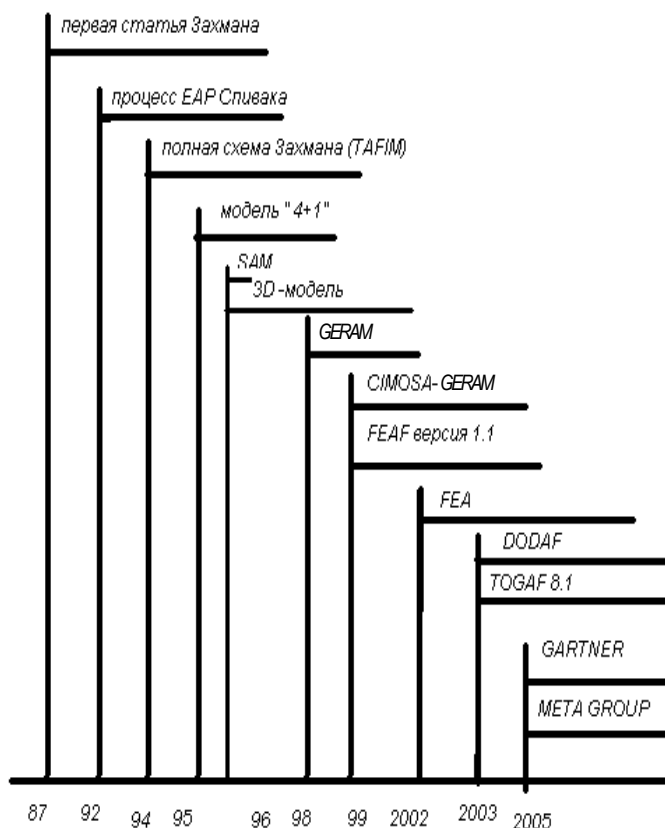


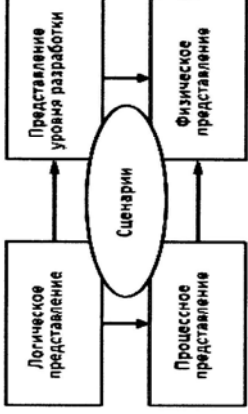
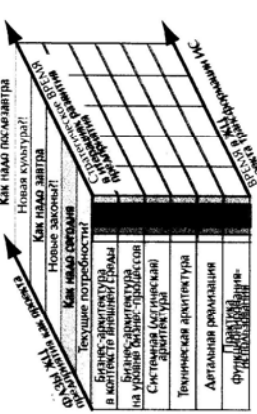
Рис. 7.1. Диаграмма развития методов построения Архитектуры предприятия

Таблица 7.1

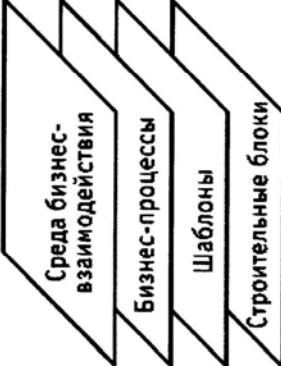
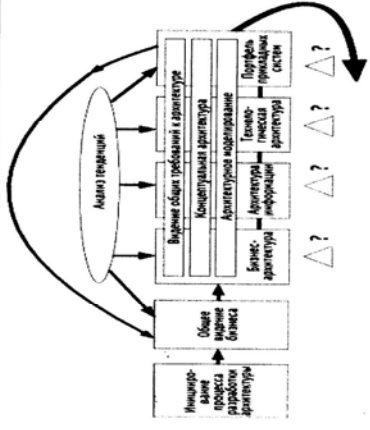
Краткая характеристика концепций архитектуры предприятия

Дата	Автор Название	Основная идея	Иллюстрация
1987	Модель Дж.А. Захмана.	Модель предприятия представляется в виде набора согласованных описаний, которые соотносятся с ячейками формализованной матрицы, в которой учтены существенные для архитектуры аспекты предприятия. По столбцам матрицы разнесены основные аспекты деятельности (объекты – «что», действия – «как», местоположения – «где», люди – «кто», время – «когда» и мотивы – «почему»). По строкам – различные описания системы с точки зрения бизнес-руководителей, менеджеров и разработчиков	
1994	Модель Министерства обороны США ТАFIM (Technical Architecture Framework for Information Management) – База архитектуры технического обеспечения для управления информацией	В матрице учтены существенные для архитектуры аспекты предприятия. Заполнение матрицы происходит сверху вниз.	
1992	Стивен Сливак. Модель EAP (Enterprise Architecture Planning) – Планирование архитектуры предприятия) [7, 18].	Методика EAP обеспечивает взгляд на предприятие с точки зрения его бизнес-функций и требований к информационному обеспечению. Применен сегментный подход – 4 уровня, 7 блоков, определяющих архитектуру и соответствующий план ее реализации. Модель основана на упрощенной матрице Захмана 1987 г.	
1996	Английская консалтинговая компания Systems Advisers Ltd. разработала стратегическую модель архитектуры SAM (Strategic Architecture Model)	SAM – это надстройка над моделью архитектуры предприятия Захмана. Она предоставляет общие структуры для определения архитектуры и механизмы, позволяющие организовать и анализировать информацию об архитектуре. SAM использует нотацию «сфер интересов» для представления целостного набора фактов о предприятии и «отношений», которые связывают эти факты в «полезные» группы. «Сферы интересов» SAM позволяют систематизировать всю информацию, имеющую отношение к определенному предмету.	

Продолжение таблицы 7.1.

Дата	Автор, фирма Название	Основная идея	Иллюстрация
1995	Филипп Круттен. Модель «4+1» (точнее 2The 4+1) View Model of Architecture	<p>Предлагается использовать 5 представлений. Четырьмя основными представлениями в этой методике являются следующие.</p> <p>Логическое представление – отвечает на вопрос: что система должна выполнять в терминах конечных пользователей? Для иллюстрирования могут применяться диаграммы классов (в нотации языка UML).</p> <p>Процессное представление. Учитывает нефункциональные требования к системе, включая производительность и доступность. Описывает вопросы параллельного исполнения и синхронизации процессов.</p> <p>Физическое представление. Описывает размещение программных компонент системы на аппаратных платформах и аспекты, связанные с физическим расположением системы.</p> <p>Представление уровня разработки. Описывает фактическую организацию модулей системы, разделение ее на подсистемы, которые могут разрабатываться независимо.</p> <p>Архитектура системы во многом определяется сценариями, которые объединяют все представления вместе.</p>	
1996–1997	Евгений Зиндер. «3D-модель» предприятия	<p>Введена ось времени, где располагаются интервалы осуществления различных проектов и стадий развития ИС и всего предприятия. В качестве других осей выступает матрица Захмана: 1) ось уровня проектирования и использования ИС; 2) ось раздела обеспечения и аспекта работы ИС; 3) ось времени, в котором развивается предприятие и его ИС.</p> <p>Стратегический и детальный анализ могут рассматриваться и как разные стадии, что демонстрирует принцип адаптации схемы к жизненным циклам разных типов. Характер расположения архитектурных компонентов ИС в этом третьем измерении отличается большим разнообразием, поскольку в реальной жизни многие процессы трансформации предприятия идут параллельно и итерационно.</p>	

Окончание таблицы 7.1

Дата	Автор, фирма Название	Основная идея	Иллюстрация
2005	Компания Gartner. Модель Gartner	<p>Компания Gartner считает, что архитектура предприятия должна начинаться с того, что организация собирается достичь, а не с текущего положения дел. После того как в организации будет сформировано единое представление о ее развитии, можно рассматривать влияние этого представления на архитектуру бизнеса, технологическую архитектуру, информационную архитектуру и архитектуру решений.</p> <p>В данном подходе сформулированы рекомендации, касающиеся разработки архитектуры в виде последовательности шагов и задач участников, которые, однако, не детализированы до уровня моделей процесса разработки архитектуры.</p> <p>Методика описания архитектуры Gartner представляет собой как бы трехмерный куб.</p>	
2005	Компания Meta Group. Методика META Group	<p>Отличительной особенностью методика META является более детальное и формализованное описание процесса разработки архитектуры и всех его составляющих.</p> <p>Архитектура реализуется на практике через процесс управления ИТ-программами и проектами.</p> <p>Предусматривает совместное участие представителей бизнес-подразделений и ИТ в выработке общего понимания набора требований.</p> <p>Организация процесса разработки архитектуры и создание начальной версии архитектуры предприятия, согласно META Group, состоит в прохождении следующих этапов:</p> <ol style="list-style-type: none"> 1. Видение общих требований в архитектуре. 2. Разработке концепции архитектуры. 3. Архитектурное моделирование. <p>Методика предлагает формализованные шаблоны, обеспечивающие разработку основных документов: «Выделение общих требований» и «Принципы концептуальной архитектуры».</p>	

Развитие представлений об архитектуре предприятия

Термин «архитектура» означает множество близких по смыслу понятий, но применяющихся в различных приложениях. Известных формальных определений архитектуры существует несколько сотен ¹³².

В англоязычной литературе – методиках, статьях, стандартах – термину «*архитектура предприятия*» соответствует термин *enterprise architecture*.

Дж. А. Захман в 1987 г. в статье «Структура архитектуры информационных систем» вначале называл свою концепцию *архитектурной структурой* информационных систем, а впоследствии – *структурой архитектуры предприятия*.

В публикациях, в основном в Интернете, содержится большое количество рекомендаций, материалы аналитических компаний, теоретические и практические наработки в области архитектуры ИТ таких организаций, как Совет Директоров по Информационным Технологиям госорганизаций США (CIO Council) и ряда других ¹³³

Поскольку Захман предложил термин «архитектура» применительно к информационной системе, то первоначально было распространено определение в программистском стиле: «*Архитектура системы состоит из нескольких компонент, внешних свойств и интерфейсов, связей и накладываемых ограничений, а также архитектуры этих внутренних компонент*».

Однако вскоре **Е. Зингер** предложил более широкое определение: «*Архитектура предприятия – это вовсе не архитектура информационных систем или технологий, данное понятие охватывает и устройство бизнеса (деятельности по государственному управлению, если это министерство), и базовые техно-*

¹³² См., напр., сайт Института Проектирования Программного Обеспечения Карнеги-Меллона (SEI – Carnegie Mellon Software Engineering Institute) www.sei.cmu.edu/technology/architecture.

¹³³ Enterprise Architecture – Sergey Orlik (Microsoft Platforma 2011). www.slideshare.net/sorlik/enterprise-architecture-serge.

логии (например, станки, банкоматы, технологические процессы и т. д.), и работников всех видов и рангов, и информационные технологии»¹³⁴.

«Архитектурный взгляд» на системы (как ИТ-системы, так и бизнес-системы) определен в стандарте ANSI/IEEE 1471-2000 как *фундаментальная организация системы, состоящая из совокупности компонент, их связей между собой и внешней средой, и принципы, которыми руководствуются при их создании и развитии*».

Формальное определение в стандарте IEEE1471¹³⁵. Института инженеров-электриков и электронщиков для определения архитектуры предлагает метамодель. Этот стандарт определяет такие абстрактные элементы архитектуры, как представления, системы, среды, обоснования, заинтересованные стороны и т. д. В соответствии с этим представлением *«система обладает некоторой архитектурой, которая может быть определенным образом описана с различных точек зрения в зависимости от интереса тех людей (участников), которые рассматривают архитектуру системы. Каждой точке зрения на архитектуру системы соответствует определенное представление, основу которого составляет некоторый набор моделей»*.

Однако этот стандарт не определяет структуру собственно архитектуры предприятия. Например, говорится о том, что необходимо иметь различные представления архитектуры, но при этом не указывается, какие это должны быть представления.

Применительно к архитектуре предприятия формальное описание впервые было сформулировано в стандарте ISO 15704, который был предложен рабочей группой IFAC/IFIP (International Federation of Automatic Control / International Federation for Information Processing). Идея состояла в том, чтобы разработать максимально общую, так называемую эталон-

¹³⁴ **Зиндер Е.** «3D-предприятие» – модель трансформирующейся системы / Евгений Зиндер // http://www.iteam.ru/publications/it/section_53/article_1272.

¹³⁵ **Стандарт** IEEE 1471-2000. Рекомендации IEEE по архитектурному описанию преимущественно программных систем. Подробную информацию об этом стандарте описания архитектуры можно получить по адресу <http://www.enterprise-architecture.info/Images/Documents/IEEE2014712000.pdf>.

ную (reference) модель архитектуры предприятия, которая охватывала бы дополнительно процесс развития предприятия во времени как проект, а также учитывала бы роль человеческого фактора. Тогда архитектуры отдельных подсистем, в том числе ИТ-системы предприятия, могут быть разработаны как специфические уточнения такой общей модели.

Разработанная как приложение к данному стандарту, такая общая эталонная модель архитектуры получила название GERAM (Generalized Enterprise Reference Architecture and Methodology – Обобщенная референсная архитектура и методология предприятия). Фактически GERAM представляет собой абстрактное описание архитектуры общего уровня, которая может быть использована для «привязки» и сравнения между собой различных практических моделей архитектур. В частности, в ее рамках определяются такие понятия, как «продукт», «жизненный цикл», «роль персонала в системе», «моделирование процессов» применительно к задачам описания функционирования предприятия¹³⁶.

В Национальном стандарте Российской Федерации¹³⁷ архитектура определяется как *«описание (модель) основного устройства (структуры) и связей, частей системы (физического или концептуального объекта или сущности)»*.

При этом предлагается различать два типа архитектур, имеющих отношение к интеграции предприятия, а именно:

а) системные архитектуры (называемые иногда архитектурами «типа I»), действие которых распространяется на проекти-

¹³⁶ Базовый стандарт «Архитектура Предприятия – ISO 15704:2000 «Requirements for enterprise-reference architectures and methodologies».

¹³⁷ Национальный стандарт Российской Федерации «Промышленные автоматизированные системы. Требования к стандартным архитектурам и методологиям предприятия. Industrial automation systems. Requirements for enterprise-reference architectures and methodologies» (Дата введения – 2010-01-01).

² Национальный стандарт Российской Федерации «Промышленные автоматизированные системы. Требования к стандартным архитектурам и методологиям предприятия. Industrial automation systems. Requirements for enterprise-reference architectures and methodologies» (Дата введения – 2010-01-01)

рование системы, например, на компьютеризированную, являющуюся частью системы интеграции предприятия;

b) стандартные проекты предприятия (называемые иногда архитектурами «типа 2»), действие которых распространяется на организацию разработки и выполнения проекта, например, интеграцию предприятия или другую программу развития предприятия.

В некоторых работах выделяют такие виды, как *системная архитектура* (архитектура систем – System Architecture) и *программная архитектура* (архитектура программного обеспечения – Software Architecture).

При этом, хотя методика описания и проектирования архитектуры отдельных прикладных систем имеет много общего с подходами к описанию архитектуры предприятия в целом, тем не менее, архитектура программных систем является отдельной областью знаний, которой посвящено большое количество соответствующих публикаций.

Под «программной архитектурой» в зависимости от контекста может пониматься как архитектура взаимодействия приложений в рамках информационной системы предприятия (т. е. *архитектура приложений*), так и *архитектура программных модулей* или *архитектура взаимодействия различных классов в рамках одного приложения*.

Каждая из архитектур, в свою очередь, может рассматриваться с тем или иным уровнем детализации. Так, для *программной* архитектуры выделяют следующие уровни описания архитектуры:

- **концептуальная архитектура** определяет компоненты системы и их назначения, обычно в неформальном виде. Это представление часто используется для обсуждения с нетехническими специалистами, такими как руководство, бизнес-менеджеры и конечные пользователи функциональных характеристик системы (что система должна уметь делать, в основном, с точки зрения конечного пользователя);

- **логическая архитектура** выделяет, прежде всего, вопросы взаимодействия компонент системы, интерфейсы и используемые протоколы. Это представление позволяет эффективно организовать параллельную разработку;

- **физическая реализация**, которая описывает привязку к конкретным узлам размещения, типам оборудования, характеристикам окружения, таким как, например, используемые операционные системы и т. п.

Интересный пример анализа различных аспектов деятельности программного архитектора предлагается в стандарте IEEE 1471-2000¹³⁸.

В соответствии с определениями компании Gartner¹³⁹ архитектура – это:

• *общий план или концепция, используемая для создания системы, такой как здание или информационная система, или «абстрактное описание системы, ее структуры, компонентов и их взаимосвязей»;*

• *семейство руководящих принципов, концепций, правил, шаблонов, интерфейсов и стандартов, используемых при построении совокупности информационных технологий предприятия.*

Первое определение ориентировано на описание существующих и будущих систем, второе – на процесс их построения.

В. И. Галактионов¹⁴⁰ предлагает рассматривать архитектуру предприятия в двух аспектах:

- статическом – в некоторый фиксированный момент времени;
- динамическом – как процесс перехода (миграции) от текущего состояния к некоторому желаемому состоянию в будущем.

Рассматриваемая в статике архитектура предприятия состоит из следующих элементов:

- миссия и стратегия, стратегические цели и задачи;
- бизнес-архитектура;
- системная архитектура.

Рассматриваемая в динамике архитектура предприятия — это логически связанный цельный план действий и скоординированных проектов, необходимых для преобразования сложившейся архитектуры организации

¹³⁸ <http://www.bredemeyer.com>.

¹³⁹ Структура архитектуры предприятия Gartner: развитие, 2005 г. / Грета А Джеймс, Роберт А. Хэндлер, Энн Лапкин и Николас Галл. – 25 октября 2005 г. – Код Gartner: G00130855.

¹⁴⁰ **Галактионов В. И.** Системная архитектура и ее место в архитектуре предприятия / В. И. Галактионов // Директор информационной службы, 2002. – № 5.

к состоянию, определенному как долгосрочная цель, базирующийся на текущих и планируемых бизнес-целях и бизнес-процессах организации.

Проектирование системной архитектуры предполагает разделение системы на **наиболее крупные составные части** и принятие **конструктивных решений**, которые после их принятия *с трудом поддаются изменению*. Если впоследствии оказывается, что нечто изменить легче, чем казалось вначале, это «нечто» легко исключается из «архитектурной» категории.

Статический срез системной архитектуры на определённый момент времени включает:

- архитектуру приложений – функциональный и компонентный состав информационной системы;
- архитектуру данных – способы взаимодействия систем и хранения данных;
- архитектуру оборудования – используемые технические средства/решения.

Другими аспектами системной архитектуры являются:

- способы и планы миграции от текущего состояния архитектуры к целевому;
- способы передачи реализаций между средами;
- стоимость решения, включая капитальные и операционные расходы.

При этом В.И. Галактионов считает, что существует более одного способа описания архитектуры. Степень важности каждого из этих способов. меняется в продолжении жизненного цикла.

А. В. Данилин и **А. И. Слюсаренко** проводят анализ определений архитектуры и предлагают в зависимости от контекста определять архитектуру системы как [5]:

1) многоаспектное описание или план задуманной или развиваемой системы на уровне ее компонентов, детализированное в достаточной мере для руководства ее воплощением, а также принципы и руководящие материалы, определяющие руководство конструированием и развитием системы во времени;

2) структура существующей системы как совокупность ее компонентов и их взаимосвязей.

В варианте второго пункта этого определения архитектура у предприятий была всегда.

Опираясь на методологию Gartner, в соответствии с которой подход к формулировке архитектуры должен основываться на анализе бизнес-процессов и поддерживающих их приложений, авторы [5] делают вывод о том, что *«архитектура предприятия является одним из инструментов организационных изменений всего предприятия в целом с использованием ИТ, и особенно той части организации, которая отвечает за информационные технологии»*, и что *«представление об архитектуре предприятия имеет свои корни в дисциплине, которая получила название “системное мышление”»*.

И дают более подробное разъяснение этой идеи следующим образом.

«Отличительной характеристикой решений, принимаемых в отношении архитектуры, является то, что эти решения должны приниматься с учетом широкой, или системной, перспективы. Любое решение, которое может быть принято локально (например, в рамках подсистемы), не является архитектурным для системы в целом. Это позволяет делать различие между детальным проектированием и принятием решений по поводу практической реализации системы, с одной стороны, и архитектурными решениями – с другой. Первые решения имеют локальное влияние, а вторые – систематическое. Поэтому для проектных решений нужна соответствующая более широкая перспектива, позволяющая учесть системное влияние решений более высокого уровня, что дает возможность достичь желаемого уровня компромиссов и соглашений между составными частями для обеспечения должного уровня качества системы в целом».

При этом в [5] отмечается, что организации испытывают постоянные трудности в синхронизации целей и задач бизнеса и процессов развития своих информационных систем. Существует как бы «облако неопределенности» между определением организацией и обеспечивающей ее ИТ-инфраструктурой своих целей и задач.

Архитектура информационных технологий и архитектура предприятия в целом по мнению авторов [5] является основным механизмом интерпретации и реализации целей организации через адекватные ИТ-инфраструктуру. Это достигается через создание

определенного количества взаимосвязанных архитектурных представлений, которые с использованием различных методик описания архитектуры разбивают архитектуру предприятия на различное количество моделей и определений, относящихся к таким областям, как *бизнес, информация, прикладные системы, технологическая инфраструктура*.

Бизнес-модели описывают стратегию организации, структуры управления, требования, ограничения и правила, а также основные бизнес-процессы, включая взаимосвязи и зависимости между ними. Т. е. *бизнес-архитектура* описывает на уровне предприятия в целом то, как реализуются основные функции организации, включая организационные и функциональные структуры, роли и ответственности, расположение, время, типы файлов и баз данных и других информационных хранилищ.

Архитектура информации определяет ключевые активы, связанные со структурированной и неструктурированной информацией, требующейся для бизнеса, включая расположение, время, типы файлов и баз данных и других информационных хранилищ.

Архитектура прикладных систем описывает те системы, которые и обеспечивают необходимый функционал для реализации логики бизнес-процессов организации.

С точки зрения *технологической архитектуры* важные модели включают описание ИТ-сервисов, которые требуются для реализации перечисленных выше трех других областей архитектуры.

При этом *логические модели* ИТ-сервисов построены в абстрактной, технологически независимой форме и оставляют свободу для оптимального выбора конкретных технологий. Но, в конце концов, архитектура предприятия завершается *физическими моделями*, которые определяются технологиями, аппаратными и программными платформами, выбранными для реализации ИТ-сервисов.

Иерархическое построение архитектуры позволяет облегчить ее восприятие человеком.

В определениях можно выделить, по крайней мере, три различных аспекта:

- иерархия архитектур различных организационных систем;
- соотношения между объективной реальностью и субъективным восприятием;

- соотношения между общесистемной архитектурой и частными архитектурами.

Можно говорить об *архитектуре предприятия в целом, архитектуре уровня отдельных проектов или семейства продуктов, об архитектуре отдельной прикладной системы.*

При этом все виды архитектуры должны использовать сходные средства описания и представления результатов, опираться на методы декомпозиции (структуризации) сложных систем, а архитектура представляет собой некоторую модель реальной системы, которая динамически изменяется, сохраняя соответствие оригиналу.

Профессионалы в области информационных технологий понимают под архитектурой ИТ достаточно большой спектр понятий – структурированное семейство технических руководств, включая концепции, принципы, правила, шаблоны и интерфейсы, а также взаимосвязи между ними, которые используются при создании новых информационных систем и развитии существующих систем. В отличие от них, профессионалы в области бизнеса не рассматривают этот вопрос как вопрос исключительно технологий. Наоборот, они пользуются терминами «бизнес-модели», «бизнес-процессы» и иногда – «бизнес-архитектура».

В процессе эволюции понятия «Архитектура предприятия» этот термин означал все более комплексный и всеобъемлющий подход к описанию и практике использования информационных технологий по обеспечению основной деятельности организации и, как результат, получение все более широкого спектра преимуществ.

В ранних работах ИТ-архитектура понималась в основном как *технологическая* архитектура или архитектура, определяющая *инфраструктуру* информационной системы. Работы по описанию архитектуры были сосредоточены на формировании технологических стандартов и принципов, включая проведение инвентаризации различных технологий, используемых в организации. Такой подход позволяет добиться определенных частных выгод, связанных прежде всего с уменьшением стоимости заку-

пок и эксплуатации информационных систем, уменьшением затрат на разработку приложений и обучение персонала. Однако он был ограниченным, так как не подразумевал ориентацию на решение бизнес-задач.

Следующей ступенью явилось понятие *корпоративной информационно-технологической архитектуры* масштаба предприятия (EWITA – *Enterprise-wide information technology architecture*). Это означало, что усилия по описанию архитектуры предприятия должны включать в себя описание архитектуры информации и архитектуры прикладных систем, а не только технологический уровень. Основное направление работ при этом состояло в совместном использовании общих данных, исключении дублирования бизнес-функций, координации управления пользователями, ресурсами, информационной безопасностью за счет улучшений в управлении портфелем прикладных систем. Корпоративная информационно-технологическая архитектура масштаба предприятия описывает то, как компоненты информационной системы связаны между собой; точно так же бизнес-архитектура описывает то, как элементы бизнеса связаны между собой.

Такой подход обеспечивает более эффективное взаимодействие различных структурных подразделений организации, совместный доступ к информации различных подразделений и внешних организаций (клиентов, партнеров, поставщиков); уменьшение дублирования с точки зрения параллельной реализации близких по функционалу прикладных систем для различных бизнес-подразделений; решение проблем, которые затрагивают интересы нескольких подразделений, например, интеграция и взаимодействие информационных систем.

Важным логическим шагом для эффективного описания существующих в организации процессов и планируемых изменений явилось введение понятия *архитектуры предприятия* (*Enterprise Architecture*), которая объединяет *корпоративную ИТ-архитектуру* масштаба предприятия с *бизнес-архитектурой* и позволяет обеспечить достижение стратегических целей предприятия.

Преимуществами такого включения бизнес-архитектуры в контекст рассмотрения целостной архитектуры предприятия являются большая способность организации к изменениям или динамичность (*agility*) и синхронизация возможностей информационных технологий с бизнес-стратегией, обеспечение вариативности бизнес-стратегии за счет возможности изменений в обеспечивающих процессах и технологических решениях; лучшие перспективы с точки зрения использования возможностей информационных технологий по формированию бизнес-стратегии.

Таким образом, концепция архитектуры предприятия явилась результатом поиска некоторого целостного подхода, который обеспечил бы взгляд на организацию как единую систему.

Введены понятия *архитектуры предприятия*, *архитектуры приложений*, *архитектуры данных*, *технологической архитектуры*, что позволяет кратко охарактеризовать различные многомерные структуры различных страт ИС – страты данных, программных приложений, клиент-серверов, технологической страты, ИС предприятия в целом, т. е. *архитектуры предприятия*.

Концепция архитектуры предприятия представляет собой многомерную модель, описывающую структуру и функции предприятия, позволяет получить детализированное описание информационной системы предприятия, включая цели и задачи; процессы и их организация; данные и используемые технологии.

При этом понятия *структуры* и *архитектуры* используются неоднозначно:

- либо понятие «структура» характеризует строение, расположение, порядок, т. е. конфигурацию ИС – «структура архитектуры» у Захмана;

- либо существует точка зрения ¹, в соответствии с которой под *архитектурой* понимается описание системы с точки зрения конечных пользователей и интерфейсов взаимодействия с внешней средой, т. е. как внешний взгляд на ИС, а *структуру* ИС описывают в виде взаимодействующих между собой подсистем, т. е. внутреннюю конфигурацию ИС. При этом каждая подсистема может

¹ **Бойченко А. В.** Основы открытых информационных систем / А. В. Бойченко, В. К. Кондратьев, Е. Н. Филинов. – М.: Изд. Центр АНО «ЕОАИ», 2004. – 126 с.

быть разделена на составные части в иерархии, вплоть до модулей прикладных программ, принимаемых за неделимые элементы. Таким образом, понятие структуры представляется в форме иерархии (древовидной или стратифицированной), включающей несколько уровней разбиения, и полученные структурные единицы, в свою очередь, могут быть представлены в виде архитектурного описания по отношению к внешним структурным единицам.

7.2. Концепция ситуационных центров

В теории разработки АСУ предусматривалось, что АИС являются первой очередью АСУ. В дальнейшем предпринимались попытки разработки вначале *информационно-справочных* и *информационно-советующих систем* (ИСС).

Затем появился термин *«системы поддержки принятия решений (СППР)»*.

В то же время практика развития информационных систем для предприятий и организаций в основном пока ограничивается разработкой или приобретением программных продуктов, обеспечивающих учетно-отчетную деятельность систем организационного управления предприятиями.

Эти ИС хорошо обеспечивают отчетность предприятий перед вышестоящими органами отраслевого и территориального управления и разнообразными контролирующими организациями.

Если разработаны требования к объему, времени выпуска продукции и др. количественным характеристикам производственных предприятий или обслуживающих организаций, то АИС позволяют определить рассогласование реальных результатов с желаемыми и представить эту информацию лицам, принимающим решение.

Первоначально поставленная задача создания именно АСУ, была реализована в основном на уровне технологических процессов – АСУ ТП, которые разрабатывались для ряда техноло-

гических процессов, для непрерывных производств, иногда – для управления конвейером.

А для системы организационного управления даже промышленными предприятиями реализованы были только АИС, обеспечивающие информацией лиц, принимающих решения, а процессы собственно принятия решений остаются за человеком.

В то же время для ряда прикладных проблем требуются новые подходы к информационно-аналитическому обеспечению управленческой деятельности.

Определенным шагом на пути к повышению степени автоматизации управления предприятиями и организациями являются системы бизнес-аналитики¹⁴¹, которые на основе различных технологий обработки данных обеспечивают кластеризацию систем, группирование и классификацию информации, что в ряде ситуаций является действенной помощью при принятии решений.

В настоящее время развивается форма информационно-аналитических систем, называемых *ситуационными центрами* (СЦ), которые должны представлять собой инструмент поддержки управленческой деятельности.

Концепция ситуационного центра была предложена английским кибернетиком *Стаффордом Биром* в 1970-е гг. Первый ситуационный центр для первых лиц государства был создан под руководством Бира в конце 1970-х гг. В настоящее время их число составляет несколько сотен и продолжает возрастать.

В нашей стране активное развитие ситуационных центров началось в 1990-е гг.

Современный ситуационный центр определяют как организационно-техническую систему, реализующую функции подготовки и поддержки принятия управленческих решений, которая позволяет наиболее полно и оперативно представлять информацию о сложившейся ситуации органам управления, прогнозировать возможные сценарии ее развития, оперативно подготавливать

¹⁴¹ См., напр., *Паклин Н. Б.* Бизнес-аналитика: от данных к знаниям (+ CD) / Н. Б. Паклин, В. И. Орешков. – СПб.: Питер, 2009. – 624 с.

возможные альтернативные варианты управленческих решений и оценивать их последствия.

Создание и распространение ситуационных центров было вызвано высокой динамичностью, сложностью, многоаспектностью, увеличением степени неопределенности задач управления, решаемых в органах государственной власти и крупных корпорациях. Создание таких систем помогает лицам, принимающим решение, в переработке больших объемов информации, в подготовке сценариев принятия решений.

Функционирование ситуационных центров основано на следующих принципах¹⁴²:

- непрерывный мониторинг и моделирование протекающих процессов, прогнозирование сценариев развития ситуаций;
- визуализация управленческих ситуаций и причинно-следственных связей анализируемых событий;
- организация коллективной выработки решений с использованием информационных ресурсов, интеллектуальных информационных технологий и средств отображения информации;
- обеспечение оперативного синтеза альтернативных решений.

Таким образом, ситуационные центры интегрируют в своем составе системы поддержки принятия управленческих решений (Decision Support System) и презентационные центры (Presentation Center).

Современные ситуационные центры позволяют решать такие задачи, как:

- обеспечение информационной поддержки руководителей;
- осуществление непосредственного доступа руководителей к территориально удаленным информационным ресурсам как структурных подразделений, так и других организаций;
- согласование и обеспечение целостности функционирования информационно-коммуникационных систем;

¹⁴² *Информационно-аналитические средства поддержки принятия решений и ситуационные центры* / Н. И. Ильин, Н. Н. Демидов, А. Н. Данчул и др.; под ред. А. Н. Данчула. – М.: Изд-во РАГС, 2006. – 326 с.

- осуществление доступа к информации организаций, взаимодействующих при принятии решения;
- сокращение временных и финансовых затрат, вызванных несовместимостью информационно-телекоммуникационных систем, дублированием подготовки данных, их противоречивостью, затруднениями с доступом, выборкой и передачей информации;
- интеграция информационных систем структурных подразделений в единое информационное пространство.

По характеру поддерживаемых управленческих задач выделяют три класса ситуационных центров (табл. 7.2)¹⁴³.

Таблица 7.2

Классы ситуационных центров

Класс СЦ	Назначение	Примеры
1. Оперативно-диспетчерские центры, решающие в реальном времени задачи оперативного управления сложными организационно-технологическими процессами с многочисленными информационными потоками	Изначально в центрах этого класса основное внимание уделялось презентационной компоненте, с развитием средств вычислительной техники в них не только стали появляться современные средства отображения информации коллективного пользования, но и все большую роль стало играть аналитическое информационно-программное обеспечение принятия решений	<ul style="list-style-type: none"> • Центр управления космическими полетами. • Центры управления перевозками. <p>И т. п.</p>

¹⁴³ Выступление *Данчула А. Н.* / Роль ситуационных центров в системе поддержки принятия решений государственного управления: круглый стол 29.11.2010.

Класс СЦ	Назначение	Примеры
<p>2. Ситуационные центры анализа и управления кризисными ситуациями.</p>	<p>Основное назначение подобных центров заключается в предотвращении кризиса за счет своевременного предоставления лицам, принимающим решения (ЛПР), исчерпывающей информации по текущему состоянию контролируемых объектов и прогнозов возможных сценариев развития событий.</p> <p>В случае же, если кризиса избежать не удалось, такие ситуационные центры становятся, по сути, оперативными штабами по управлению процессами локализации (ликвидации) последствий кризиса</p>	<ul style="list-style-type: none"> • Центр управления кризисными ситуациями Министерства по чрезвычайным ситуациям, созданный на основе решений, отработанных при ликвидации последствий Чернобыльской катастрофы в 1986 г. Этот центр обеспечивает визуализацию текущего и прогнозируемого состояния анализируемой ситуации, отображая имеющиеся силы и средства, предлагаемые рекомендации. На основе этой информации принимаются решения, которые в режиме реального времени доводятся до исполнительных структур и подразделений. <p>Ситуационные центры этого класса имеются в Минатоме, Минприроде, других ведомствах и организациях</p>
<p>3. Информационно-аналитические ситуационные центры</p>	<p>Ориентированы на принятие не только оперативных, но и стратегических решений. К ним относятся многие из созданных в последнее время ситуационных центров для самых разнообразных применений. К этому классу можно отнести и широко известную «ситуационную комнату» президента США</p>	<ul style="list-style-type: none"> • Ситуационный центр в Совете Безопасности (СБ) при Президенте РФ. СБ позволяет осуществлять мониторинг, моделирование последствий, анализ событий, которые происходят в экономике, социальной сфере, в области национальной безопасности, помогая тем самым выработать решения (создан в 1994 г.). • Ситуационный центр в резиденции Президента РФ (введен в строй в феврале 1996 г.).

Класс СЦ	Назначение	Примеры
	<p>В различных источниках указывается, что Президент США обслуживают от 4 до 5 ситуационных центров.</p>	<ul style="list-style-type: none"> • Ситуационные центры полномочных представителей Президента РФ в федеральных округах (пусковые комплексы развернуты в 2004 г). • Информационно-аналитические ситуационные центры в ряде министерств и ведомств Российской Федерации, в администрациях глав свыше десяти субъектов Федерации, наиболее развитым из которых является СЦ ОАО «Российские железные дороги» – Ц РЖД (сдан в постоянную эксплуатацию в 2002 г.)

В настоящее время активно развиваются ситуационные центры третьего класса.

Существуют и другие классификации СЦ: по составу систем ситуационного моделирования (СЦ наблюдения, аналитические, полнофункциональные); по масштабу (степени сложности) решаемой задачи; по размещению (стационарные, мобильные, виртуальные); по степени детерминированности решаемых задач; по целевой направленности (контроля, управления, кризисных ситуаций, обучения); по способу отображения информации (коллективные, индивидуальные); по универсальности (специальные, настраиваемые).

Ситуационный центр представляет собой помещение, предназначенное для оперативного принятия управленческих решений, а также для контроля разного рода объектов и событий. Ситуационные центры используются федеральными и региональными органами власти, органами местного самоуправления, крупными промышленными предприятиями, образовательными учебными учреждениями; оснащаются новейшими средствами коммуникаций: оборудованием для интерактивного представле-

ния информации, системой конференцсвязи, видеоконференцсвязи. Средства, используемые в ситуационных центрах, дают возможность наглядного отображения огромных объемов информации, на основе которой принимаются решения и доводятся до исполнителей.

В связи с необходимостью ознакомления государственных служащих с современными управленческими и информационно-аналитическими технологиями, реализуемыми в ситуационных центрах, в Российской академии государственной службы (РАГС) при Президенте РФ в 2004 г. создан учебно-исследовательский ситуационный центр. В СЦ РАГС ¹⁴⁴ предусматривается проведение занятий в форме лекций-демонстраций, лекций-дискуссий, кейс-стадий, деловых игр. Поддержка практических занятий в СЦ РАГС реализуется в виде учебно-аналитических задач (УАЗ) по решению управленческих проблем. На практических занятиях обеспечивается автоматизированная поддержка коллективного доступа к информационным ресурсам, обмена данными между автоматизированными рабочими местами участников занятий и визуализация результатов их действий на видеостене.

Занятия, проводимые в СЦ РАГС, могут быть посвящены всем или некоторым из следующих укрупненных этапов решения проблемы:

- неформальное обсуждение аналитического доклада;
- формализация проблемной ситуации и формулировка проблемы;
- групповое моделирование проблемной ситуации;
- принятие коллективных решений.

Основное содержание кейс-стадии – групповое моделирование проблемной ситуации; это комплекс действий по установле-

¹⁴⁴ *Данчул А. Н.* Первая очередь ситуационного центра РАГС: Концепция построения и состав // Информационно-аналитические средства поддержки принятия решений и ситуационные центры: Материалы научно-практической конференции, состоявшейся в РАГС 28 – 29 марта 2005 / Под общ. ред. А. Н. Данчула. – М.: изд-во РАГС, 2006. – С. 19–27.

нию возможных тенденций развития проблемной ситуации; возможных последствий ее развития; экспертному прогнозированию возможного развития проблемной ситуации в условиях применения различных стратегических решений. При использовании видеоконференций возможна организация микрогрупп удаленных участников.

В деловых играх групповое моделирование целесообразно дополнять процедурами принятия коллективных решений. Между микрогруппами распределяется несколько различных функциональных ролей участников коллективных действий.

Информационное обеспечение (ИО) СЦ РАГС строится по четырехуровневой схеме:

- 1) локальное ИО учебно-аналитических задач;
- 2) информационное обеспечение СЦ РАГС, включающее адаптированные базы данных статистической информации об экономике России и ее регионов;
- 3) информационное обеспечение общеакадемического уровня, включающее правовые базы данных и ресурсы, доступные с помощью информационно-обучающего портала «Государственное управление и местное самоуправление»;
- 4) информационные ресурсы Интернет.

При выборе технологии реализации информационных процессов (ИП) в СЦ РАГС учитывались их следующие основные характеристики:

- коллективность – участие в ИП группы взаимодействующих в интересах достижения общей цели пользователей;
- интеллектность – вследствие нечеткости исходной постановки и слабой формализуемости решаемой проблемы доля интеллектуальных действий, выполняемых человеком или в редких случаях программами искусственного интеллекта, соизмерима с долей рутинных действий или превышает ее;

- интерактивность – высокая чередуемость действий, выполняемых человеком и программно-техническими средствами, обуславливающая диалоговый характер реализуемых ИП;
- уникальность – отсутствие типовых полномасштабных технологий реализации ИП решения проблем;
- большой объем и разнообразие форм представления используемой информации, в том числе – значительная доля и существенная важность визуальной и аудиоинформации;
- сеансовость – возможность разбиения ИП решения проблемы на сеансы, разделенные определенными или неопределенными промежутками времени;
- высокая доля вспомогательных действий по подготовке сеанса: разработка или выбор сценария сеанса, сбор, отбор и подготовка данных, выбор и подготовка (настройка) необходимой программно-технической среды.

Ориентированность СЦ на решение уникальных, а не типовых задач означает, что он должен рассматриваться как система с развивающимися (за счет включения новых задач) функциями, а создание новых задач – как штатный процесс функционирования. Уникальность также влечет достаточно длительный и ресурсоемкий цикл подготовки к решению задачи. В связи с этим в составе программного обеспечения СЦ РАГС весомую часть составляет инструментальное программное обеспечение, предназначенное для автоматизации разработки этих задач.

Большой объем и разнообразие форм представления информации обеспечивается в СЦ РАГС прежде всего за счет использования видеостены, позволяющей одновременно воспроизводить визуальную и с различных АРМ и видеоисточника. Для полного использования преимуществ, предоставляемых видеостеной, важно чтобы эти формы поддерживались и специальным программным обеспечением.

Решение основных прикладных проблем информационно-аналитической поддержки управленческой деятельности в ситуационных центрах может быть достигнуто в ходе выполнения исследований по следующим направлениям.

1. Разработка теории (методов) создания информационно-аналитических систем как перманентно развивающихся систем решения уникальных слабо-формализованных задач.

2. Разработка методов и средств интеграции разнообразных программных средств и баз данных в условиях постоянного обновления их перечня.

3. Разработка методов и средств поддержки различных классов коллективной (в том числе территориально-распределенной) деятельности.

4. Разработка методов визуализации (создания информационных образов) решаемых задач, организация хранения и поиска визуальной информации.

Развитию идей и форм реализации ситуационных центров способствуют регулярно проводимые на базе РАГС конференции¹⁴⁵ и семинары по проблемам создания ситуационных центров, заседания за «круглым столом» на базе Аналитического центра при Правительстве РФ.

Эти конференции, заседания и экспертные обследования ситуационных центров позволили выявить основные проблемы, которые необходимо решить для развития ситуационных центров, основными из которых являются следующие¹⁴⁶:

- СЦ не встроены в цепочку принятия решения;
- нет понимания роли СЦ как со стороны руководителей, так и со стороны обслуживающего персонала;
- не в полной мере используются имеющиеся ресурсы

¹⁴⁵ *Ситуационные* центры и перспективные информационно-аналитические технологии поддержки принятия решений: Матер. науч.-практич. конф. РАГС, 7-9 апреля 2008 г. – М.: Изд-во РАГС, 2009. – 418 с.; *Ситуационные* центры 2010. Современные информационно-аналитические технологии поддержки принятия решений: Матер. науч.-практич. конференции РАГС, 27–28 апреля 2010 / Под общ. ред. д-ра техн. наук, профессора А. Н. Данчула. – М.: Изд-во РАГС, 2011. – 420 с.

¹⁴⁶ *Ситуационные* центры: нереализованный потенциал / В. И. Ерохин, М. В. Садофьев, Т. В. Еферина, И. М. Алексеева // Стратегическое управление: от идеи до результата. – М.: Аналитический центр при Правительстве РФ, 2011. – С. 84–103.

- основные функции большинства СЦ – пока еще только презентации и учебные мероприятия;
- на уровне органов управления территориями и в надведомственных структурах ЛПР, как правило, не заинтересованы в развитии СЦ;
 - кадры не готовы;
 - слаба методическая подготовка;
 - основными источниками информации являются Интернет и Росстат;
 - редко используется многомерное представление и многофакторный анализ;
 - документальное обеспечение сводится к должностным инструкциям.

На Федеральном уровне СЦ функционируют более эффективно, чем на региональном

В то же время анализ показывает, что развития требуют исследовательские, методические, образовательные, организационные проблемы, проблемы нормативно-методического обеспечения управления функционированием СЦ. Необходимо разрабатывать типовые сценарии, типовую методологию функционирования СЦ. Необходима также пропаганда полезности СЦ для организации коллективного принятия решений. Заимствуя зарубежный опыт, необходимо учитывать, что в России в отличие от многих зарубежных стран, не ситуационная психология, а когнитивное мышление.

Таким образом, создание и практическое использование ситуационных центров как важного средства информационно-аналитических технологий поддержки принятия решений требует развития теоретических и методических основ процессов принятия решений на базе СЦ.

Утверждение Стратегии национальной безопасности и Основ стратегического планирования в Российской Федерации как инструментов Государственного управления дополнительно к прежним функция ситуационных центров добавило задачу создания систе-

мы ситуационных центров, ориентированных главным образом на стратегическое планирование и согласование стратегических целей общества и стратегических целей государства в лице Правительства РФ. Создание системы ситуационных центров является компонентом более крупной задачи обеспечения вертикальных и горизонтальных связей в системе стратегического планирования в РФ, основанной на принципах долгосрочного и среднесрочного программно-целевого управления, решением задачи мониторинга и прогнозирования чрезвычайных ситуаций, создания системы независимой экспертной оценки программно-проектной деятельности органов государственной власти¹⁴⁷.

В следующих разделах рассматриваются подходы к разработке информационных систем организаций, основанные на применении теории систем и системного анализа.

7.3. Подходы к анализу и проектированию систем и их применение для развития теории проектирования информационных систем

На протяжении всей истории развития теории систем предлагались и применялись различные подходы к представлению (отображению), анализу и проектированию систем.

Традиционный подход, применяющийся в математических исследованиях, требует определить элементы (переменные, константы) и связать их соответствующим соотношением (формулой, уравнением, системой уравнений), отображающим принцип взаимодействия элементов.

Когда задачи усложнились, и такое соотношение не удавалось сразу найти, то предлагалось формировать «пространство состояний» элементов и вводить «меры близости» между элементами это-

¹⁴⁷ **Ситуационные** центры: нереализованный потенциал / В. И. Ерохин, М. В. Садофьев, Т. В. Еферица, И. М. Алексеева // Стратегическое управление: от идеи до результата. – М.: Аналитический центр при Правительстве РФ, 2011. – С. 84–103.

го пространства. Такой подход вначале пытались применить для исследования сложных систем.

Предлагалось обследовать систему, выявить все элементы и связи между ними. Этот подход называли иногда «перечислением» системы. При обследовании применялись разные способы: 1) *архивный* (изучение документов и архивов предприятия); 2) *опросный* или *анкетный* (опрос сотрудников, в том числе с помощью специально разработанных вопросников – анкет).

Однако первые же попытки применить такой подход к исследованию систем управления предприятиями и организациями показали, что «перечислить» сложную систему практически невозможно. В истории разработки автоматизированных систем управления был такой случай. Разработчики написали несколько десятков томов обследования системы, а так и не смогли приступить к созданию АСУ, поскольку не могли гарантировать полноты описания. Руководитель разработки вынужден был уволиться, и в последствии стал изучать системный подход и популяризировать его.

Учитывая трудности «перечисления» системы, с самого начала возникновения системных теорий исследователи искали подходы к ее анализу и созданию.

Основными подходами к анализу систем, предлагавшимися разными исследователями, являются следующие:

- в начальный период становления теории систем развивался бихевиористский подход, основанный на исследовании поведения (*behaviour* – поведение) систем; однако этот подход весьма трудоемок и не всегда реализуем;
- американский ученый **М. Месарович**¹⁴⁸ предложил подходы, которые назвал *целенаправленным* и *терминальным* (от *терм* – элементарная частица, интересующая исследователя);
- польский ученый **Р. Куликовски**¹⁴⁹ предложил называть аналогичные подходы *декомпозицией* и *композицией* системы;
- швейцарский астроном, венгр по происхождению **Ф. Цвикки**¹⁵⁰ предложил и развил *морфологический подход*, который помогает искать полезные объединения элементов путем их комбинаций;

¹⁴⁸ **Месарович М.** Общая теория систем и ее математические основы / М. Месарович // Исследования по общей теории систем: сб. переводов / Под ред. В.Н. Садовского и Э.Г. Юдина. М.: Прогресс, 1969.– 520 с.

¹⁴⁹ **Куликовски Р.** Оптимальные и адаптивные процессы в системе автоматического регулирования / Р. Куликовски. – М.: Наука, 1967.

¹⁵⁰ **Zwicky F.** Morfological astronomy / F.Zwicky. – Berlin: Springer-Verlag, 1957. – 299 p.

- американская корпорация *RAND*¹⁵¹ предложила подход к созданию сложных программ и проектов, названный «*деревом целей*»;

- в практике проектирования сложных технических комплексов возникли термины *язык моделирования*, *язык автоматизации проектирования*, применяющиеся для отображения взаимосвязей между компонентами проекта; при разработке языков моделирования применяют *математическую логику* и *математическую лингвистику*, в которой есть удобный термин для описания структуры языка – *тезаурус* (см. гл. 5), и подход называют иногда *лингвистическим* или *тезаурусным*;

- при исследовании и формировании структур были предложены следующие подходы: путем поиска связей между элементами; или, напротив, путем устранения лишних связей (см. [1, 13]).

С учетом рассмотренных подходов в настоящее время на основе обобщения предшествующего опыта сформировалось два основных подхода к отображению систем, первоначально предложенных для формирования структур целей¹⁵².

а) «сверху» – методы *структуризации* или *декомпозиции*, *целевой* или *целенаправленный* подход;

б) «снизу» – подход, который называют *морфологическим* (в широком смысле), *лингвистическим*, *тезаурусным*, *терминальным*, методом «*языка*» системы. С помощью этого подхода определяется «пространство состояний» системы и реализуется поиск взаимосвязей (мер близости) между элементами.

Подход «снизу» можно реализовать, применяя не только комбинаторные приемы (морфологический и т. п.), но и бихевиористский подход, вариант которого при автоматизации моделирования поведения объектов в настоящее время иногда называют *процессным*.

¹⁵¹ *Лопухин М. М.* ПАТТЕРН – метод планирования и прогнозирования научных работ / М. М. Лопухин. – М.: Сов. радио, 1971. – 160 с.

¹⁵² *Теория систем и методы системного анализа в управлении и связи* / В. Н. Волкова, В. А. Воронков, А. А. Денисов и др. М.: Радио и связь, 1983. – 248 с.

Подходы «сверху» и «снизу» называют также *аксиологическим* и *каузальным*, соответственно ¹⁵³.

Аксиологическое представление системы – отображение системы в терминах *целей* и целевых функционалов. Этот термин используют в тех случаях, когда необходимо выбрать подход к отображению системы на начальном этапе моделирования и противопоставить это отображение описанию системы в терминах «перечисления» элементов системы и их непосредственного влияния друг на друга, т. е. *каузального представления*.

Каузальное представление системы – описание системы в терминах влияния одних переменных на другие, без употребления понятий *цели* и *средств* достижения целей. Этот термин происходит от понятия «cause» – причина, т. е. подразумевает причинно-следственные отношения. При каузальном представлении будущее состояние системы определяется предыдущими состояниями и воздействиями среды. Такое представление является развитием отображения системы в виде «пространства состояний», характерного для большинства математических методов моделирования. Применяют каузальное представление в случае предварительного описания системы, когда *цель* сразу не может быть сформулирована, и для отображения системы или проблемной ситуации не может быть применено *аксиологическое представление*.

На практике обычно эти подходы сочетают.

В 1970–1980-е гг. при проектировании организационных структур были предложены три подхода к решению этой проблемы ¹⁵⁴.

- *Нормативно-функциональный* подход направлен на унификацию организационных форм управления в рамках отрасли. Разработка и внедрение типовых организационных структур явились первым шагом на пути внедрения принципов их научно обоснованного построения. Однако ориентация на типовую номенклатуру функций управления и структурных управленческих подразделений не позволяет учесть особенности конкретных предприятий и условия их деятельности.

¹⁵³ *Математика* и кибернетика в экономике: словарь-справочник / Отв. ред. Н. П. Федоренко – М.: Экономика, 1975. – 700 с.

¹⁵⁴ *Мильнер Б. З.* Системный подход к организации управления / Б. З. Мильнер, Л. И. Евенко, В. С. Рапопорт. – М.: Экономика, 1983. – 224 с.

- *Функционально-технологический* подход основан на рационализации потоков информации и технологии ее обработки, на формировании и анализе организационно-технологических процедур подготовки и реализации управленческих решений. Этот подход обеспечивает возможность достаточно полно учесть особенности конкретного предприятия (организации), отличается гибкостью и универсальностью. Вместе с тем, он характеризуется высокой трудоемкостью, использованием стабильной номенклатуры сложившихся функций управления, подчинением оргструктуры схеме документооборота.

- *Системно-целевой* подход заключается в построении структуры целей, определении на ее основе функций управления и их организационном оформлении. Преимущества этого подхода заключаются в возможности учитывать особенности объекта управления и условия его деятельности, изменять и расширять состав функций, проектировать разнообразные организационно-правовые формы предприятий. Трудности в использовании подхода связаны с проблемой перехода от совокупности целей и функций к составу и подчиненности структурных звеньев, обеспечивающих их реализацию.

Применительно к исследованию и разработке ИС эти подходы можно интерпретировать следующим образом.

Обобщающий подход «сверху», называемый целевым, целенаправленным, системно-целевым, основан на структуризации или декомпозиции системы *в пространстве*. Этот подход позволяет расчленив исходную большую неопределенность на более обозримые и выбрать методы их анализа и проектирования, сохраняя целостность представления об исследуемой системе или решаемой проблеме на основе иерархической структуры (древовидной, стратифицированной). Подход применялся при разработке АИС и АСУ для крупных предприятий, при реструктуризации систем организационного управления.

Подход «снизу», основанный на анализе пространства состояний, поиске «мер близости» между компонентами с помощью различных, в том числе статистических методов, морфологического

моделирования отличается большой трудоемкостью. Для анализа пространства состояний в экономике разработаны методы бизнес-аналитики Data-Mining, реализуемые с помощью соответствующих программных продуктов (Deductor¹⁵⁵ т. п.).

В настоящее время для проектировании информационных систем широкое применение нашел подход, основанный на анализе бизнес-процессов, кратко называемый *процессным*.

Процессный подход (который можно считать развитием *функционально-технологического подхода*) основан на структуризации *во времени*, на представлении процессов в форме графов.

При несомненной привлекательности применение функционально-технологического подхода долгое время было практически нереализуемым из-за большой трудоемкости, отсутствия правил и средств автоматизации формирования графов, отображающих процессы в системах.

В 1990-е годы была разработана методология SADT – (Structured Analysis and Design – структурный анализ и проектирование (предложена *Дугласом Россом*¹⁵⁶), представляющая собой совокупность методов, правил и процедур, предназначенных для построения функциональной модели объекта какой-либо предметной области. На ее основе разработаны и стали широко применяться функционально-ориентированные и объектно-ориентированные CASE- и RAD-технологии. Компьютерная реализация методологии SADT получила название IDEF (Icam Definition). Основными структурными моделями являются модели процессов IDEF0 и IDEF3, модель данных IDEF1X¹⁵⁷. Созданы стандарты IDEF и DFD, ориентированные на анализ процессов (в

¹⁵⁵ См., напр., *Паклин Н.* Бизнес-аналитика: от данных к знаниям / Н. Паклин, В. Орешков. – СПб.: Питер, 2009. – 624 с.

¹⁵⁶ *Ross D.* Applications and extension of SADT // IEEE. Computer. – April, 1995.

¹⁵⁷ См., напр., *Черемных С. В.* Структурный анализ систем: IDEF-технологии / С. В. Черемных, И. О. Семенов, В. С. Ручкин. – М.: Финансы и статистика, 2003. – 208 с.

том числе бизнес-процессов), что позволило создать теорию проектирования экономических информационных систем [14, 15]. Для реализации моделей применяются автоматизированные средства – BPWin, ARIS, язык UML (Unified Modeling Language – Унифицированный язык моделирования)¹⁵⁸.

Популярность CASE-методологии и RAD-технологий базируется на разработке принципов и автоматизации формирования процессов, на развитии методов их формирования (на основе анализа «жизненного цикла» производства, обслуживания или других процессов, причинно-следственных связей и т. п.), что и обеспечило развитие *процессного подхода*, преимущества которого заключаются в возможности учитывать особенности конкретного объекта и условий его деятельности.

На основе анализа методик проектирования документальных, фактографических и документально-фактографических информационных систем, приведенных в гл. гл. 4–6, с точки зрения рассмотренных подходов можно сделать следующие выводы.

1. Проектирование АИС и АСУ разных уровней в 1970-е гг. осуществлялась на основе *системно-целевого подхода*, т. е. на основе *структуризации целей и функций* системы управления (*целевой* или *целенаправленный* подход, подход «сверху»), формирования на ее основе структуры *функциональной части* АСУ (АИС) и определения средств для ее реализации – *обеспечивающей части*, т. е. информационного, технического, программного, организационного и других видов обеспечения.

Для управления разработками АСУ были подготовлены и изданы соответствующие общепромышленные руководящие методические материалы, в которых отражались теория и основная терми-

¹⁵⁸ См., напр., **Фаулер М.** UML в кратком изложении: Применение стандартного языка объектного моделирования / Пер. с англ./ М. Фаулер, К. Скотт. – М.: Мир, 1996. – 112 с.

нология, обязательная для использования при представлении отчетных материалов по разработке и внедрению АИС и АСУ.

2. При разработке ГАСНТИ также вначале был применен *системно-целевой подход*, разработана классификация АС НТИ для разных уровней управления – общегосударственном, отраслевом, региональном, предприятий и организаций, и соответствующие нормативные документы. Однако в настоящее время ГАСНТИ развивается в основном только на общегосударственном уровне.

3. Разработка документальных информационно-поисковых систем обычно проводится в виде следующих этапов: а) анализ совокупности документов, представляющих соответствующую научно-техническую область, для которой создается ИПС; б) выбор ключевых слов, характеризующих содержание каждого документа; в) формирование из ключевых слов словаря (первой стадии разработки информационно-поискового языка); г) индексирование документов с использованием словаря; д) поиск; е) проверка релевантности выдачи и отладка ИПС.

Постепенно ИПЯ развиваются, вначале на основе формирования дескрипторного словаря, затем – тезауруса, а при необходимости – и разработки грамматики. Существующие ДИПС (в том числе поисковых систем Интернет) редко развиты до уровня наличия в них многоуровневого тезауруса и грамматики. В них обычно используются простейшие правила грамматики – конъюнкция и дизъюнкция (названные упрощенно – «AND», «OR» и т. п.), а термин «тезаурус» используется в упрощенном варианте, в лучшем случае – это двухуровневый словарь.

С использованием терминологии рассмотренных подходов, можно считать, что разработка ДИПС начинается с анализа «пространства состояния», т. е. следует квалифицировать подход как *терминальный, лингвистический, тезаурусный*, метод «языка» системы, в упрощенном названии – как подход «снизу».

По мере развития ДИПС при разработке систем, реализующих режимы избирательного распределения информации и дифференцированного обслуживания руководителей, применя-

ют *системно-целевой подход*, т. е. начинают разработку с исследования потребностей научных подразделений и руководителей.

4. Для разработки экономических информационных систем создана теория проектирования, основанная на применении *процессного подхода*, анализе бизнес-процессов, представляемых специальными графами.

Идею применения графического представления и анализа информационных потоков (*функционально-технологический подход*) при проектировании информационных систем начинали применять в 1970-е гг.¹⁵⁹. Был разработан язык моделирования информационных потоков, реализованы средствами языка логического программирования РЕФАЛ¹⁶⁰. Однако его применение долгое время было практически нереализуемым из-за большой трудоемкости и отсутствия стандартных автоматизированных процедур (проведенные исследования с применением языка РЕФАЛ были экспериментальными).

Разработка CASE- и RAD-технологий, стандартов IDEFO и DFD, ориентированных на анализ процессов (в том числе бизнес-процессов и соответствующих им информационных потоков) и средств их автоматизации позволила создать широко применяющуюся в настоящее время теорию проектирования информационных систем, основанную на *процессном подходе* [14, 15 и др.].

5. При автоматизации управленческой деятельности предприятий малого и среднего бизнеса, которые, как правило, не имеют

¹⁵⁹ **Волкова В. Н.** Некоторые вопросы автоматизации проектирования АИС: автореф. дис. на соискание ученой степени канд. техн. наук по специальности 05.13.01 – Техническая кибернетика и теория информации / В. Н. Волкова. – М., 1973. – 24 с.; **Волкова В. Н.** К методике проектирования автоматизированных информационных систем / В. Н. Волкова // Автоматизация управления и вычислительная техника: кн. – Вып. 11. – М.: Машиностроение, 1975. – С. 189–300.

¹⁶⁰ **Волкова В. Н.** Автоматизация проектных работ при создании автоматизированных информационных систем / В. Н. Волкова // Типизация и автоматизация процессов проектирования АСУ: матер. Всесоюзн. Семинара. – Душанбе: Изд-во Дониш, 17–19 апреля 1972. – С. 79–87; **Волкова В. Н.** Исследование возможностей сокращения перебора при семиотическом моделировании проектных задач АСУ / В. Н. Волкова, А. Г. Красовский // АСУ и синтез оптимальных систем: сб. тезис. докл. науч.-техн. конф. – М.: НТОРЭС, 1974. – С. 18–20.

средств для внедрения единой автоматизированной системы, существует практика выбора имеющихся на рынке программных продуктов, автоматизирующих соответствующие области управления (бухгалтерский учет, материально-техническое обеспечение и т. п.) с последующей их интеграцией в единую систему. Для выбора этих программных продуктов разрабатывают соответствующие методики (пример приведен в разделе 4.6), основанные на применении *системно-целевого подхода*, т. е. разработке структуры целей и функций организации и выбора программных продуктов, функциональные возможности которых обеспечивают реализацию функций и задач организации.

7.4. Анализ определений системы и их применение для развития теории информационных систем

В теории систем существуют различные определения системы (см. обзоры в [1, 13, 16]).

В первых определениях система рассматривалась как совокупность только *элементов* a_i и *связей* r_j между ними (*Л. фон Бергман* и др.):

$$\begin{aligned}
 S &\stackrel{\text{def}}{=} \langle A, R \rangle, \text{ где } A = \{a_i\}, R = \{r_j\} \\
 S &\stackrel{\text{def}}{=} \langle \{a_i\}, \{r_j\} \rangle, \\
 &\quad a_i \in A \quad r_j \in R \\
 S &\stackrel{\text{def}}{=} \langle \{a_i\} \& \{r_j\} \rangle, \\
 &\quad a_i \in A \quad r_j \in R
 \end{aligned}
 \tag{7.1}$$

Затем – с учетом их *свойств* Q (*А. Холл, А. И. Уёмов* и др.):

$$S \stackrel{\text{def}}{=} \langle A, R, Q \rangle.
 \tag{7.2}$$

Затем в определения стал вводиться *системообразующий критерий* (*М. Г. Гаазе-Рапопорт*), а позднее – и *цель* Z в явном виде:

$$S \stackrel{\text{def}}{=} \langle A, R, Z \rangle,
 \tag{7.3, a}$$

где Z – цель, совокупность или структура целей.

В определении *В. Н. Сагатовского* уточняются условия целеобразования – *среда* SR , *интервал времени* ΔT , т. е. период, в рам-

ках которого будет существовать система и ее цели: система – «конечное множество функциональных элементов и отношений между ними, выделенное из среды в соответствии с определенной целью в рамках определенного временного интервала»¹⁶¹:

$$S \underset{def}{=} \langle A, R, Z, SR, \Delta T \rangle. \quad (7.3, б)$$

При этом понятие цели использовалось в перечислении наряду с элементами, связями, свойствами, т. е. рассмотрение системы все же начиналось с элементов и связей, а не с формулирования целей (цель предполагалась заданной).

В последующем в определении предлагается учитывать наблюдателя N :

$$S \underset{def}{=} \langle A, R, Z, N \rangle. \quad (7.4, а)$$

На необходимость учета взаимодействия между изучаемой системой и исследователем или проектировщиком первоначально указал **У. Р. Эшби**¹⁶² и в качестве обобщающего названия лиц, исследующих или проектирующих систему, ввел термин «наблюдатель».

Первое определение, в которое в явном виде включен наблюдатель, дал **Ю. И. Черняк**: «Система есть отражение в сознании субъекта (исследователя, наблюдателя) свойств объектов и их отношений в решении задачи исследования, познания»¹⁶³:

$$S \underset{def}{=} \langle A, Q_A, R, Z, N \rangle. \quad (7.4, б)$$

Ю. И. Черняк стал учитывать и язык наблюдателя L_N : «Система – отображение на языке наблюдателя объектов, отношений и их свойств в решении задачи исследования, познания»:

$$S \underset{def}{=} \langle A, Q_A, R, Z, N, L_N \rangle. \quad (7.4, в)$$

Анализ приведенных определений системы показывает, что первые определения опирались на подход к исследованию и про-

¹⁶¹ **Основы** системного подхода и их применение при разработке территориальных АСУ / Под ред. Ф.И. Перегудова. – Томск: ТГУ, 1976. – с. 13–14.

¹⁶² **Эшби У.Р.** Введение в кибернетику / У.Р. Эшби. – М.: Ин. лит., 1959. – 432 с.

¹⁶³ **Черняк Ю. И.** Системный анализ в управлении экономикой / Ю. И. Черняк. – М.: Экономика, 1975. – 191 с.

ектированию системы, базирующийся на отображении пространства состояний (*элементов, связей, свойств*) и поиске мер близости на этом пространстве. Включенная в последующем в определения *цель* рассматривалась также как один из компонентов пространства состояний, а «*наблюдатель*» и его «*язык*» являлись субъектом и средством анализа пространства состояний. Этот подход в теории систем **М. Месарович** называет *терминальным*, **Ю. И. Черняк** – *лингвистическим* или *методом «языка» системы*; для краткости в теории систем принят упрощенный термин – подход к исследованию или проектированию системы *от элементов*, т. е. как бы «снизу».

В то же время по мере развития теории систем было осознано, что наиболее предпочтительным при создании информационных систем, особенно для социально-экономических объектов является аксиологический подход – от целей, потребностей (подход «сверху»), т. е. *системно-целевой подход*.

На основе анализа определений и по аналогии с определением, которое было дано в [18] для информационно-семантических систем, предложено более общее определение, в котором реализуется *системно-целевой подход*¹⁶⁴:

$$S \underset{def}{\equiv} \langle Z, STR, TECH, COND, N \rangle, \quad (7.5)$$

где $Z = \{z\}$ – совокупность или структура целей;

$STR = \{STR_{пр}, STR_{орг}, STR_{ИТ}, \dots\}$ – совокупность структур, реализующих цели ($STR_{пр}$ – производственная, $STR_{орг}$ – организационная, $STR_{ИТ}$ – ИТ-структура т.п.);

$TECH = \{meth, means, alg, \dots\}$ – совокупность технологий (методы – *meth*, средства – *means*, алгоритмы – *alg* и т.п.), реализующих систему;

¹⁶⁴ **Волкова В. Н.** Развитие определения системы / В. Н. Волкова // Матер. Международной научно-практической конференции «Системный анализ в проектировании и управлении»: сб. – СПб.: Изд-во СПбГПУ, 2001. – С. 12–14.

$COND = \{\varphi_{ex}, \varphi_{in}\}$ – условия существования системы, т. е. факторы, влияющие на ее создание и функционирование (φ_{ex} – внешние, φ_{in} – внутренние);

N – «наблюдатели» (по *У.Р. Эшби* и *Ю.И. Чепняку*), т. е. лица, принимающие и исполняющие решения, осуществляющие структуризацию целей, корректировку организационной и производственной структуры, осуществляющие выбор методов и средств моделирования и т. п.

В теории информационного поиска для определения информационно-поисковых систем научно-технической информации предлагались определения, примеры которых были приведены в гл. 5. Проанализируем эти определения с точки зрения принципов, положенных в их основу.

Первоначально определения ИПС отображали *процесс информационного поиска* в множестве документов D ; с помощью множества информационных запросов Q и отношений R , при наличии которых любому запросу $q_i \in Q$ ставится в соответствие подмножество D' ответов на информационный запрос [9, 17].

$$IP_{S \underset{def}} \equiv \langle D, Q, R, D' \rangle, \quad (7.6)$$

$D' \subset D$

Затем *А. И. Черный* [17] дал определение, отображающее *принципы построения и функционирования ИПС*

$$IP_{S \underset{def}} \equiv \langle LS, D, TS, N \rangle, \quad (7.7)$$

$$LS \underset{def} \equiv \langle RL, IND, KSS \rangle. \quad (7.8)$$

где LS – *логико-семантический аппарат* (т. е. информационно-поисковый язык RL , правила индексирования IND и критерии смыслового соответствия KSS);

D – *поисковый массив* (т. е. определенное множество снабженных поисковыми образами документов);

TS – *технические средства* (т. е. приспособления или устройства, которые необходимы для записи и хранения поисковых образцов, для хранения документов и осуществления процесса сопос-

тавления поисковых образов документов с поисковым предписанием или поисковым образом запроса);

N – люди, взаимодействующие с системой (т. е. те, кто пользуются данной ИПС и обслуживают ее – осуществляют индексирование документов и информационных запросов, выбирают стратегию поиска и т. д.).

В последующем **Ю. И. Шемакин** дал определение информационно-семантической системы [18, с. 60], в котором учитываются понятия цели, структуры, технологических процессов, методов, средств и условий поиска:

$$ISS \underset{def}{\equiv} \langle a, St, tp_{iss}, co, t_i \rangle, \quad (7.9)$$

где *a* – цель; *St* – структура; $tp_{iss} \in TP$ – подмножество технологических процессов для данной *ISS*; *co* – условия; *t_i* – время.

Входящие в определение (8.8) составляющие могут быть детализированы с учетом конкретной реализации ИПС. Особенно важно уточнять состав технологических процессов:

$$tp_{iss} \equiv \langle met, re, SemSI \rangle,$$

где *met* – методы; *re* – средства; *SemSI* – семантическая переработка семантической информации.

Определение (8.8) позволяет лучше учесть конкретные условия и особенности средств реализации при проектировании современных ИПС.

Для организации проектирования информационных систем **Ф. Тельнов** [14] предложил определение, в котором учитываются и цели, и *элементы*, разделяя их на внешние и внутренние, и *отношения*, и *функции* (*процессы*, *операции*), и период времени *T*, и закономерности, определяющие структуру системы и ее взаимодействие с внешней средой:

$$S \underset{def}{=} \langle G, E_i, E_n, T, F, R, Z \rangle. \quad (7.10)$$

где *G* – цели; *E_n* – внешние элементы; *E_i* – внутренние элементы; *R* – отношения, включая *динамические взаимодействия*; *F* – функции (процессы, операции); *T* – период существования сис-

темы; Z – закономерности, определяющие структуру системы и ее взаимодействие с внешней средой.

Определение **Ю.Ф. Тельнова** (8.9) положено в основу развиваемого им подхода к проектированию информационных систем, базирующегося на анализе бизнес-процессов, т. е. *процессного* подхода.

Обобщенное определение ИПС. С учетом анализа развития определений системы в теории систем и в теории информационного поиска можно дать определение ИС, в котором учитывались бы и *элементы* разного рода, и *цели* (как правило, многоуровневая структура функциональной части АСУ или совокупность информационных потребностей), и *структуры* (структура информационного, программного и др. видов обеспечения в АСУ, структура ИПС), и *отношения* между компонентами, и *технологии* в широком смысле (методы, технические средства, алгоритмы, информационные технологии обработки информации), и *среда* (внешние и внутренние факторы), и *временной интервал*, и *логико-семантический аппарат*, и *люди*, взаимодействующие с системой (проектировщики, пользователи и обслуживающий персонал). При этом с учетом большого числа факторов можно вначале включить в систему укрупненные множества, а затем раскрыть их:

$$S \underset{def}{\equiv} \langle Z, STR, LS, TECH, COND, SR, \Delta T, N \rangle, \quad (7.11)$$

где Z – цели, которые в АИС могут интерпретироваться как структура функциональной части АСУ, а в документальных и документально-фактографических ИС – как потребности;

STR – структуры информационных массивов, т.е. баз данных в фактографических системах; поисковых массивов документов D (определенное множество снабженных поисковыми образами документов) в документальных ИПС, хранилищ информации разного рода (выбор термина диктуется объемами информационных массивов и конкретными условиями);

$LS \equiv \langle RL, IND, KSS \rangle$ – логико-семантический аппарат, включающий информационно-поисковые языки RL , систему индексирования IND и критерии выдачи (или критерии смыслового соответствия KSS);

$TECH$ – технологии в широком смысле, включая TS – технические средства (т.е. приспособления или устройства, которые необходимы для сбора, регистрации, хранения, обработки и представления информации), $meth$ – методы сбора, хранения, обработки информации, включая алгоритмы alg , программные процедуры или пакеты прикладных программ PPP , информационные технологии IT и т.п.;

$COND$ – условия, т.е. внешние φ_{ext} и внутренние φ_{int} факторы, влияющие на создание и функционирование ИС; для их анализа полезно использовать признак «пространство инициирования целей», т.е. выявлять факторы надсистемы и актуальной сред (φ_{ext}), подведомственной и собственно системы (φ_{int});

ΔT – временной интервал создания и функционирования («жизни») информационной системы,

N – люди, взаимодействующие с системой, т.е. те, кто заказывают ИС, проектируют ИС, пользуются данной ИС и обслуживают ее – осуществляют индексирование документов и информационных запросов, выбирают стратегию поиска.

Для обобщающего названия всех, кто имеет отношение к разработке и использованию систем, **У.Р. Эшби** в начале развития кибернетики и теории систем ввел термин «наблюдатель»; а в настоящее время часто используют термин «стейкхолдеры». Буквальный перевод термина «*stakeholders*» – «держатели ставок», т.е. игроки в тотализатор на ипподроме. При переводе на русский язык предлагались варианты «акционеры», «заинтересованные стороны». Но каждый из них неточен: «акционеры» – неполный состав лиц, взаимодействующих с системой, в их числе могут быть и «незаинтересованные». Поэтому был охрaнен термин без перевода, и его часто применяют при исследовании или проектировании именно информационных систем.

Определение (7.11) может интерпретироваться с учетом значения вида информационной системы, условий ее разработки.

Запись определений в формализованном виде помогает сохранять целостное представление о концепции, положенной в основу определения.

Компоненты, входящие в определения, конечно, можно трактовать по-разному. Можно и для элементарной базы данных назвать цель или назначение ее создания, перечислить запросы, отвечающие какой-то совокупности потребностей, определить структуру БД, методы и средства ее реализации, условия использования.

Однако важно понимать, что, разрабатывая информационную систему для предприятий, организаций, следует исходить из их целей, т. е. сформировать структуру ФЧ АИС на основе анализа целей и функций организации и определения наиболее значимых подсистем для автоматизации. А, следовательно, нужно начинать с формулировки концепции проекта системы, формирования определения системы, соответствующего этой концепции, анализа целей и роли всех видов информационных ресурсов для реализации целей, а уж затем выбирать технологии (методы, алгоритмы, средства, в том числе готовые программные продукты) с учетом их соответствия поставленным целям.

Приведенное определение помогает теоретически обосновать концепцию многоуровневой информационной системы, рассматриваемую в следующем разделе.

7.5. Концепция многоуровневой информационной системы¹⁶⁵

По мере развития представлений об автоматизации информационного обеспечения любых организаций становится очевидным, что для функционирования и управления организацией нужен широкий спектр информационного обеспечения, включая разнородную информацию от фактографической до документальной.

Основной особенностью автоматизации информационного обеспечения процессов принятия решений является необходимость интеграции между собой большого количества независи-

¹⁶⁵ Раздел подготовлен совместно с аспиранткой *Ю.А. Голуб.*

мых друг от друга и территориально распределенных информационных массивов.

В этих условиях задача информационного обеспечения процессов принятия решения представляет собой задачу формирования алгоритмов вызова информационных массивов, хранящихся и актуализируемых в различных распределенных и независимых базах данных и информационных системах.

Решение задачи в такой постановке потребовало разработки новой концепции строения и функционирования автоматизированной информационной системы.

Для регулирования распределения финансовых и других ресурсов на разработку и поддержку функционирования разнородных баз данных и информационных систем и алгоритмов сбора информации для обеспечения потребностей лиц, принимающих решения, или других пользователей, необходимо объединить независимые базы данных и алгоритмы в единую систему. Появилось понятие *интегрированной информационной системы*. Предлагались различные способы представления структуры таких систем.

В связи с осознанием огромного влияния информационных ресурсов на социальное и экономическое развитие всего мира, отдельных стран и регионов появилось понятие *информационной инфраструктуры* (ИИ). Термин был введен в зарубежных и отечественных публикациях в конце 1990-х гг.¹⁶⁶.

Термин «*инфраструктура*» (лат. *infra* – под, внизу, *structure* – строение, постройка) появился вначале в строительном деле и означал основание, фундамент, нижнее строение.

В дальнейшем понятие было расширено и в переносном смысле стало означать совокупность внешних по отношению к рассматриваемому объекту сооружений, в том числе не только в строительстве, но и по отношению к производству, к организациям непроизводственной сферы.

¹⁶⁶ *Clinton W. J., Gore A.* Technology for America's Economic Growth, a New Direction to Build / W. J. Clinton, A. Gore // Executive Office of the President. Washington, DC, 1993, 39 pp.

В нашей стране на необходимость введения понятия *информационной инфраструктуры* (ИИ) впервые было обращено внимание в статье **Ф. Широкова** и **В. Дрожжинова** ¹⁶⁷.

Попытки разработки концепции информационной инфраструктуры организации на примере вуза были инициированы в 1990-е гг. студентами **А. А. Ионовым** и **И. В. Релиной** – членами Молодежного академического общества при Научном совете по проблемам высшей школы, подготовки и аттестации научно-педагогических кадров Санкт-Петербургского Отделения Международной академии наук высшей школы ¹⁶⁸.

Для управления развитием ИИ ее предлагалось условно представить в виде 4-х страт (рис. 7.2): *пользовательской, функциональной, информационной, коммуникационной* (информационная супермагистраль).

<p><i>Страта 1. Пользовательская</i></p> <p>Собственно пользователи, включая формирование их потребностей в информации и правила взаимодействия с информационной системой</p>
<p><i>Страта 2. Функциональный слой</i></p> <p>Совокупность услуг, предоставляемых пользователям различными поставщиками информации: доступ к базам данных и библиотекам, стандартам и т. д.</p>
<p><i>Страта 3. Информационный слой</i></p> <p>Информация в телекоммуникационных сетях, базах данных и информационных хранилищах библиотек, отделов научно-технической информации и т. п.</p> <p>Информационные ресурсы: справочники, сайты, статистическая информация, нормативно-правовая информация и пр.</p>
<p><i>Страта 4. Информационная супермагистраль</i></p> <p>Информационные телекоммуникации. Информационные технические средства</p>

Рис. 7.2. Страты информационной инфраструктуры

¹⁶⁷ **Широков Ф.** Три ступени Альберта Гора / Ф. Широков, В. Дрожжинов // Компьютер-Пресс, 1994. – № 9. – С. 43–46. – № 10. – С. 87–95.

¹⁶⁸ **Волкова В. Н.** Информационная инфраструктура как средство управления учебно-научным процессом / В. Н. Волкова, А. А. Ионов, И. В. Релина // Сб. тезисов докладов III Международной научно-технической конференции: Высокие интеллектуальные технологии образования и науки. – СПб.: СПбГТУ, 1996. – С. 185–186.

Для обоснования концепции создания интегрированной информационной системы современного вуза были проанализированы определения систем и подходы к их разработке, предложенные в теории систем.

Определение (7.11), отражающее системно-целевой подход к проектированию систем, может интерпретироваться с учетом назначения, вида информационной системы, условий ее разработки.

Например, стратифицированную структуру, приведенную на рис. 7.3, учитывая сложившиеся концепции и термины в теории разработки АСУ и систем научно-технической информации, можно обосновать следующим образом.

Цели Z в определении (7.11) реализуются в форме *функциональной страты*. Назначение *функциональной страты* (страта 2) – обеспечить удобный доступ к информационным хранилищам – *информационной страте* (страта 3).

Для этого нужно структурировать направления деятельности, цели Z организации, определить взаимосвязи между направлениями деятельности, целями (подцелями) и составляющими информационной страты, как показано на рис. 7.3.

Одной из форм реализации доступа к информации является структура функциональной части АСУ, предназначенной для обеспечения информацией сферы организационного управления. В библиотеках подобную функцию выполняет предметный указатель.

Информационная страта реализует структуру информационного обеспечения – *STR* в определении (7.11). Она объединяет хранилища информации разного рода, создаваемые в организации в разных формах и на различных носителях.

Нижняя страта, названная *коммуникационной*, реализует компоненту *TECH*, т. е. включает техническое, алгоритмической, программное обеспечение информационной системы, в том числе взаимодействие в сети Интранет, если она создана в организации, и взаимодействия между локальными сетями или отдельными компьютерами, технические средства сбора, регистрации, хранения и обработки информации.



Рис. 7.3. Многоуровневая структура информационной системы

Компоненту N отражает *пользовательская* страта. Эта страта должна обеспечить доступ к информации не только руководителям, но и всем сотрудникам организации. Для ее реализации необходимо ставить задачи определения информационных потребностей пользователей.

В системах научно-технической информации ставится задача определения информационных потребностей пользователей и создания систем избирательного распределения информации (ИРИ) и дифференцированно-го обслуживания руководителей (ДОР).

В АСУ для обеспечения фактографической информацией имеются работы по созданию автоматизированных рабочих мест (АРМ) для специалистов соответствующей квалификации, должностных лиц.

Для реализации пользовательской страты необходимо установить взаимосвязи соответствующих групп пользователей.

В научно-исследовательской организации – это научные сотрудники, руководители проектов, руководители организации и т. п., на производственных предприятиях – руководители отделов организационной структуры, в медицинском учреждении – врачи, медицинские сестры, администрация; в вузе – это, например, преподаватели, студенты, сотрудники лабораторий и организационных подразделений аппарата управления; и т. п.

Необходимо определить взаимосвязи компонентов пользовательской страты с существующими направлениями и структурами функциональной страты, обеспечить доступ к этим структурам, в том числе в форме рекламы, информационных писем (например, о конференциях, проводимых мероприятиях, формах доступа к информационным источникам, к средствам издания и размножения информации и т. п.).

В идеале желательно создать соответствующую информационно-поисковую систему для информирования пользователей о существующих функциональных направлениях и хранилищах по запросам.

Пользовательская страта в организациях, как правило, пока не создана, даже для руководящих работников организации. Она складывается в процессе опыта принятия решений.

Стратифицированное представление, приведенное на рис. 7.3, помогает уточнить назначение и взаимодействия различных аспектов реализации компонентов информационной системы, входящих в определение системы (7.11).

В связи с неоднозначным использованием в публикациях термина «инфраструктура», более целесообразно трактовать структуру рис. 7.3 как многоуровневую (стратифицированную) интегрированную информационную систему (МИИС).

Концепция стратифицированной информационной системы в конкретных условиях требует уточнения и развития.

Например, если учесть другие составляющие компоненты технологии *TECH*, входящей в определение (7.11), то между функциональной и информационной, а также между пользовательской и функциональной стратами возможно введение дополнительных страт, содержащих средства (алгоритмы, программные продукты) для обеспечения возможности более полного исследования взаимосвязей между компонентами этих страт.

Если учесть логико-семантический аппарат *LS*, то практическая реализация рассмотренной концепции осуществляется путем создания информационно-поисковых систем, работающих в режимах избирательного распределения информации и ретроспективного поиска по произвольным запросам, на базе ресурсов информационной страты, сети Интернет и локальных вычислительных сетей.

Создание МИИС требует разработки и применения соответствующих методов и автоматизированных процедур для определения состава компонентов каждой из страт, оценки влияния того или иного проекта в сфере МИИС на развитие организации.

Стратифицированное представление помогает решать проблему управления проектами и программами развития ИИ на основе их оценки с целью распределения финансовых, материальных и кадровых ресурсов, для чего разрабатывается методика, учитывающая степень влияния проектов на цели организации.

Для исследования взаимосвязей между стратами необходимо использовать методы и модели системного анализа: методики структуризации целей и функций, методы организации сложных экспертиз.

В частности, для анализа и формирования *функциональной страты* целесообразно использовать методики структуризации целей и функций, методы анализа информационных потребностей, разрабатываемые в теории информационного поиска. Оценку значимости элементов каждой из страт можно проводить с применением методов организации сложных экспертиз [1, 3, 13 и др.] (метода парных сравнений в модификации *Т. Саати*, метода решающих матриц *Г. С. Поспелова*, информационного

подхода *А. А. Денисова*), учитывать разнородные критерии – технические, экономические, социальные.

Например, четырехстратную структуру представленную на этапе предварительных исследований (рис. 7.3) можно сократить до трехстратной (рис. 7.4), пропустив *функциональную страту*, которая практически отсутствует в существующих в настоящее время информационных системах вуза.

Для организации сложной экспертизы можно применить вначале метод решающих матриц, предложенный *Г.С. Поспеловым* как средство стратифицированного расчленения проблемы с большой неопределенностью на подпроблемы и поэтапного получения оценок.

С помощью метода решающих матриц можно оценить относительные информационные потребности, затем – компоненты информационной страты их обеспечивающие, а затем – компоненты нижней страты, т.е. технические и программные средства, используемые для создания информационной страты.

Тогда в модели могут быть реализованы следующие страты (сверху вниз): информационные потребности с относительными оценками $\alpha_1, \alpha_2, \dots, \alpha_j, \dots, \alpha_{n\alpha}$; информационные массивы – $\beta_1, \beta_2, \dots, \beta_i, \dots, \beta_{n\beta}$; составляющие (компоненты) информационной супермагистральной – $\gamma_1, \gamma_2, \dots, \gamma_k, \dots, \gamma_{n\gamma}$. (рис. 7.4).

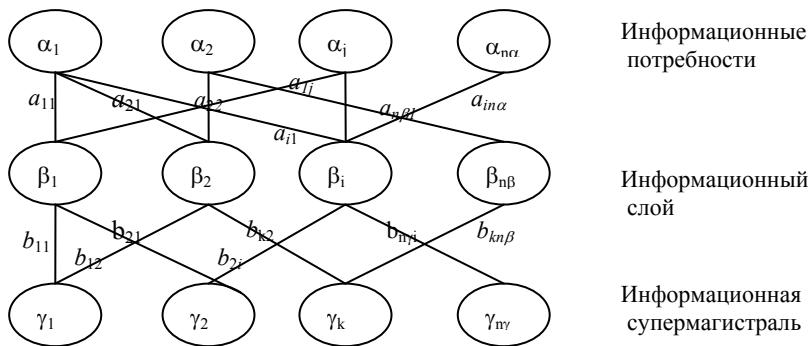


Рис 7.4. Модель решающих матриц

Каждая строка решающей матрицы характеризует относительную значимость компонентов нижележащего уровня для реализации вышестоящего.

Вначале оценивается относительная значимость α_j компонентов верхнего уровня – информационных потребностей.

Оценка проводится методом нормирования. Проверяется выполнение условия нормирования $\sum_{j=1}^{n\alpha} \alpha_j = 1$.

Затем оценивается a_{ij} – относительная значимость компонентами информационного слоя (баз данных, массивов научно-технической информации и т.п.) для реализации информационных потребностей пользователей.

Проверяется выполнение условия нормирования для каждой j :

$$\sum_i a_i = 1.$$

На основе α_j и a_{ij} вычисляется β_i – относительная значимость компонент информационной страты

$$\beta_i = \sum_j a_{ij} \alpha_j.$$

Проверяется выполнение условия нормирования $\sum_{i=1}^{n\beta} \beta_i = 1$.

Далее оценивается b_{ki} – относительная значимость компонентов информационной супермагистрали (т.е. информационных и программных средств) для реализации информационной страты.

Проверяется выполнение условия нормирования для каждой i

$$\sum_k b_k = 1.$$

Вычисляется γ_k – относительные веса компонент информационной супермагистрали для реализации информационной страты:

$$\gamma_k = \sum_i b_{ki} \beta_i$$

Проверяется выполнение условия нормирования $\sum_{k=1}^{n\gamma} \gamma_k = 1$.

Алгоритм реализации метода решающих матриц приведен на рис. 7.5.

При оценке технических и программных средств возможно формирование двух отдельных матриц и оценка значимости их компонентов для реализации информационной страты.

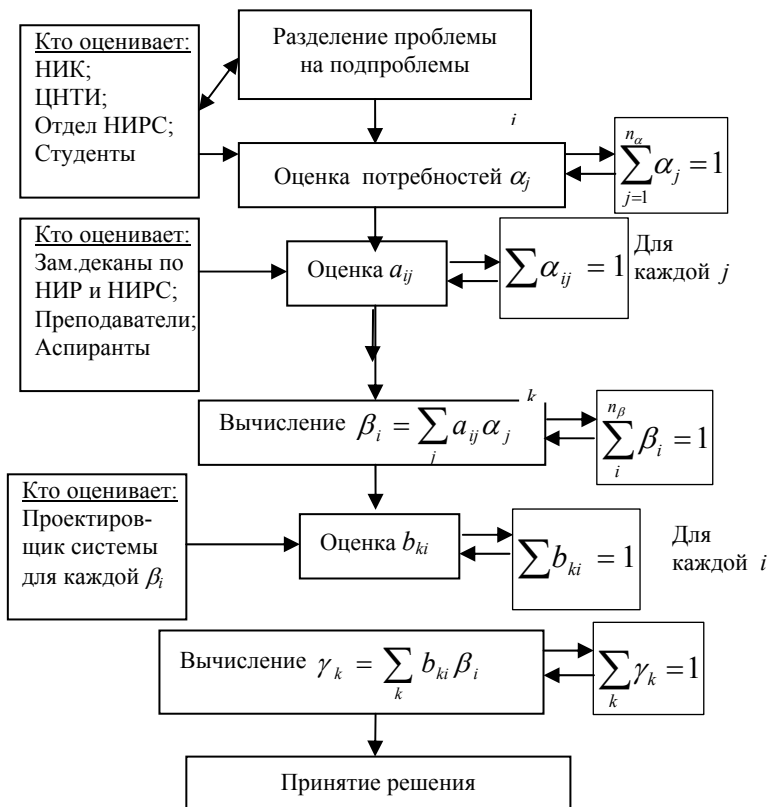


Рис. 7.5. Алгоритм реализации модели решающих матриц

Еще более объективный анализ можно получить с помощью подхода к разработке модели организации сложных экспертиз, базирующегося на использовании методов структуризации и информационных оценок степени целесообразности исследуемых компонентов, что реализуется на основе многоуровневой оценки влияния компонентов каждой нижележащей страты на реализацию компонентов вышестоящей страты.

В основу построения модели положена идея модели решающих матриц, а оценки между уровнями уточняются на основе информационного подхода *А. А. Денисова* [6], позволяющего приводить разнородные критерии (количественные и качественные) к единым информационным единицам, что помогает сопоставлять эти критерии и получать оценки потребности в различных видах услуг, используемых для проведения сравнительного анализа.

В качестве информационной меры, характеризующей значимость компонентов нижележащего уровня на вышестоящий прием принять предложенную *А.А. Денисовым* [3, 6], меру потенциала H (1.18, в), которая позволяет учесть одновременно два критерия p и q .

Тогда в модели, сохраняя страты (сверху вниз): информационные потребности с относительными оценками $\alpha_1, \alpha_2, \dots, \alpha_j, \dots, \alpha_{nn}$; информационные массивы – с оценками $\beta_1, \beta_2, \dots, \beta_b, \dots, \beta_{nb}$; составляющие (компоненты) информационной супермагистрали – с оценками $\gamma_1, \gamma_2, \dots, \gamma_k, \dots, \gamma_{ny}$, вводятся оценок между p' и q между стратами (рис. 7.6).

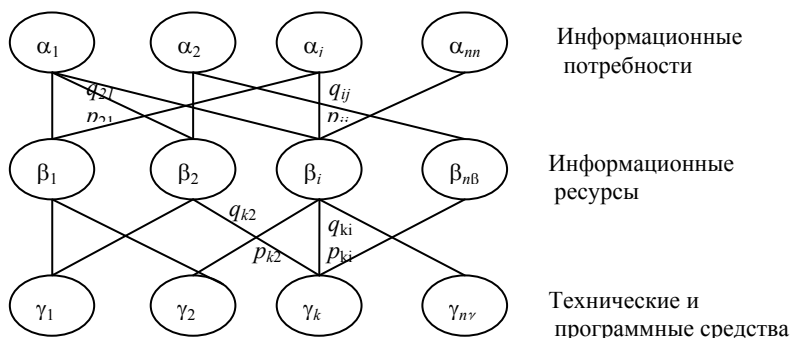


Рис. 7.6. Модель, основанная на информационных оценках

Оценка p' производится единичными экспертами, для которых определяются сферы компетентности.

Получение обобщенной оценки проводится в соответствии с алгоритмом, приведенным на рис. 7.7, что позволяет получить уточненные оценки каждой компоненты.

Компоненты верхней страты могут быть оценены либо методом нормирования, либо для верхней страты информационная мера степени влияния информационных потребностей может быть оценена как степень влияния p' потребностей пользователей на реализацию целей вуза, т.е. без учета q $H_{\alpha_j} = -\log(1 - p_{\alpha_j})$.

Для получения значения α_j необходимо выполнить процедуру нормирования, т.е. $H_{\Sigma_{\alpha}} = \sum H_{\alpha} \alpha_j = \frac{H_{\alpha_j}}{H_{\Sigma_{\alpha}}}$.

На рис. 7.6. и 7.7. приняты следующие обозначения:

$$\alpha_j = \frac{H_{\alpha_j}}{H_{\Sigma_{\alpha}}} \quad - \quad \text{относительная значимость компонентов верхнего уровня – информационных потребностей.}$$

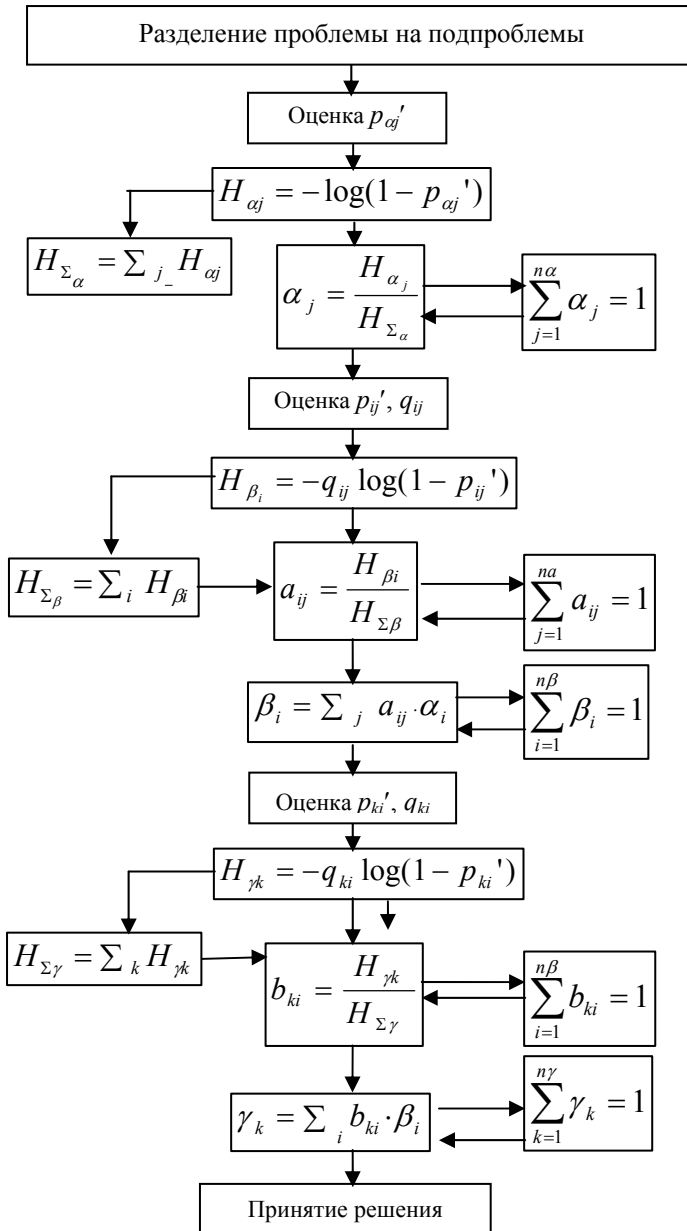


Рис. 7.7. Алгоритм реализации модели, основанной на информационных оценках

H_β – информационная мера степени влияния информационного слоя (информационных баз) на реализацию информационных потребностей.
 $H_{\beta_i} = -q_{ij} \log(1 - p_{ij})$;

p_β – степень влияния компонент информационной страты (информационных баз) на реализацию информационных потребностей;

q_{ij} – вероятность выбора и реализации компонент информационной страты с точки зрения лиц, принимающих решение при их выборе на вышестоящем уровне.

H_γ – информационная мера степени влияния компонент информационной супермагистралей (технических средств) на реализацию информационных баз. $H_{\gamma_{ki}} = -q_{ki} \log(1 - p_{ki})$;

p_γ – степень влияния компонент информационной супермагистралей (технических средств) на реализацию информационных баз;

q_{ki} – относительная значимость компонент информационной супермагистралей (технических средств) на реализацию информационных баз.

β_i – относительная значимость компонент информационной страты
 $\beta_i = \sum a_{ij} \cdot \alpha_j$;

$b_{ki} = \frac{H_{\gamma_{ki}}}{H_{\Sigma \gamma_{ki}}}$ – относительная значимость компонент супермагистралей для реализации компонент информационной страты;

γ_k – относительные веса компонент информационной супермагистралей.
 $\gamma_k = \sum b_{ki} \cdot \beta_i$.

Для реализации рассмотренных методов организации сложных экспертиз разработаны автоматизированные диалоговые процедуры [8].

Экспериментальные исследования рассмотренных моделей проведены на примере разработки пилотного проекта информационной системы для обеспечения информационных потребностей студентов.

Таким образом, предлагаемая концепция развития интегрированной информационной системы основана на стратифицированном представлении ее структуры, в условиях использования территориально распределенных и взаимно независимых информационных баз и временного объединения их на основе разработки заранее исследованных и формализованных алгоритмов и программных процедур, обеспечивающих подбор необходимой информации для удовлетворения запросов потребителей.

При реализации предлагаемой концепции информационная система представляет собой комплекс организационных структур, обеспечивающих функционирование и развитие информационного пространства организации, и средств информационного взаимодействия, т. е. включает совокупность информационных центров, банков данных и знаний, систем связи; обеспечивает доступ потребителей к информационным ресурсам на основе разработанных алгоритмов и нормативного обеспечения, регламентирующего доступ к информации.

Обращение пользователя к информационной системе должно инициировать процесс, с помощью которого обеспечивается поиск необходимой пользователю информации на основе использования территориально распределенных информационных массивов.

Интересно отметить, что многоуровневой структуры информационной системы, полученной на основе определения системы, могут быть поставлены в соответствие компоненты, определяемые концепцией архитектуры предприятия типа TO-GAF (табл. 7.1). Если эти компоненты поместить друг под другом – от компоненты «Пользователи» до компоненты «Технологическая архитектура», то получится многоуровневая структура, приведенная на рис. 7.8.

В этой структуре функциональная страта формируется на основе анализа бизнес-процессов, а между функциональной стратой (архитектурой бизнеса) и информационной стратой (архитектурой данных) помещена страта программных приложений, которую можно интерпретировать как определенный логико-семантический аппарат *LS*.

В рассмотренных примерах реализации концепции системы, лежащей в основе определения системы (7.11), не использовалась пока компонента *COND* – условия функционирования системы с учетом внутренних факторов φ_{int} и внешних факторов φ_{ext} среды, в которой функционирует информационная система.

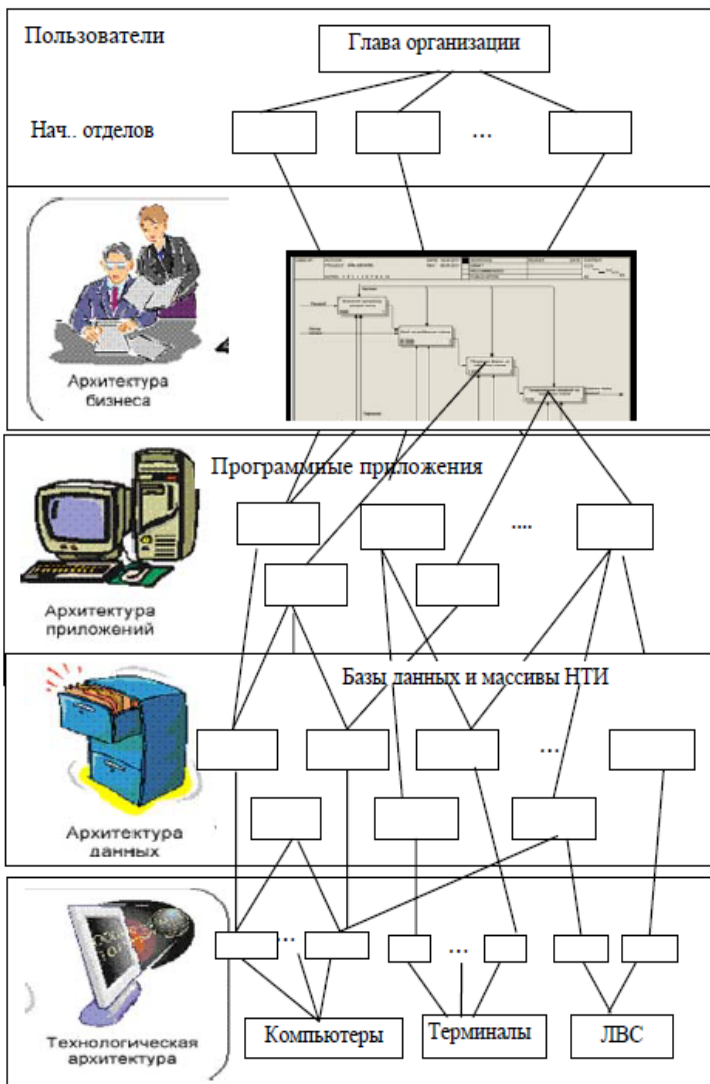


Рис.7.8. Многоуровневая структура ИС, основанная на идее TOGAF

В то же время эту компоненту необходимо учитывать как на этапе проектирования, так и в процессе функционирования ИС, что и подсказывает определение (7.11).

В частности, в процессе функционирования ИС при поиске информации в ответ на запросы пользователей необходимо учитывать, что информация может содержаться не только в массивах собственно ИС, но и во внешних базах данных, содержащихся, например, в сети Internet, локальных вычислительных сетях, в источниках научно-технической информации, доступных по межбиблиотечному абонементу (МБА) и т. п.

7.6. Проблемы и перспективы создания единой системы информации на основе объединения идей ГСНТИ и ОГАС

В настоящее время идея ОГАС практически забыта. Неудачу ее создания объяснили политическими причинами. Хотя, как было показано в гл. 5, существуют и теоретические проблемы.

Идея ГСНТИ, ее функции, сформировавшаяся в конце XX в. определении и регламентированные постановлениями Правительства РФ¹⁶⁹ от 24 июля 1997 г. № 950 и от 10 июля 1998 г. № 736, в основном, сохраняются.

В состав ГСНТИ в соответствии с прошлым опытом, частично закрепленным этими постановлениями, входят десятки федеральных и региональных центров научно-технической информации. Основные работы по развитию ГСНТИ направлены на организацию использования федеральных и региональных информационных ресурсов в области науки и техники, в том числе электронных, описание этих ресурсов, создание баз данных по приоритетным направлениям развития науки, технологий и техники, проведение важнейших научных исследований и разработок в информационной сфере, внедрение общесистемной нормативной базы, стандартов и классификационных систем в области научно-технической информации, гармонизированных с международными, создание телекоммуникационной среды доступа потребителей к информационным ресурсам ГСНТИ.

¹⁶⁹ <http://www.gsnti.ru>.

В то же время теоретические проблемы и нормативно-методические документы, регламентирующие ее функционирования и развития, требуют обсуждения.

В соответствии с Постановлением от 24 июля 1997 г. № 950 г., подписанном Председатель Правительства РФ **В.С. Черномырдиным**, было утверждено Положение о государственной системе научно-технической информации и координация деятельности ГСНТИ была возложена на Министерство науки и технологий Российской Федерации

Однако через год Министерство науки и технологий Российской академия наук Российской Федерации Приказом-распоряжением от 14 октября 1998 г. № 192/15, подписанном Министр науки и технологий **М. П. Кирпичниковым** и Президентом Российской Федерации Российской академии наук **Ю. С. Опиовым** возложило функции головной организации ГСНТИ на ИНИТИ РАН и Миннауки России.

Через месяц Приказом Министерства науки и технологий РФ от 24 ноября 1998 г. №225 был образован Координационный совет государственной системы научно-технической информации Министерства науки и технологий РФ.

В рамках Федеральной целевой программы «Электронная Россия (2002-2010 годы)» была разработана Концепция создания распределенной информационно-аналитической системы ГСНТИ ¹⁷⁰, определяющая комплекса нормативных, организационных и программно-технологических мероприятий, направленных на создание и развитие ГСНТИ России.

В соответствии с этой концепцией основным средством реализации ГСНТИ становится распределенная информационная система (РИС) на базе пяти порталов, объединяющая ГСНТИ (рис. 7.3), осуществляющая мониторинг состояния электронных информационных ресурсов и навигацию в них. Порталы определены, при крупнейших российских центрах научно-

¹⁷⁰ Концепция создания распределенной информационно-аналитической системы ГСНТИ. www2.viniti.ru/index.php.

технической информации - ВИНИТИ, ГПНТБ, ВИМИ, Институт промышленного развития «Информэлектро», Объединение «Росинформресурс». Распределение основных функций между создаваемыми порталами: приведено в табл. 7.9.

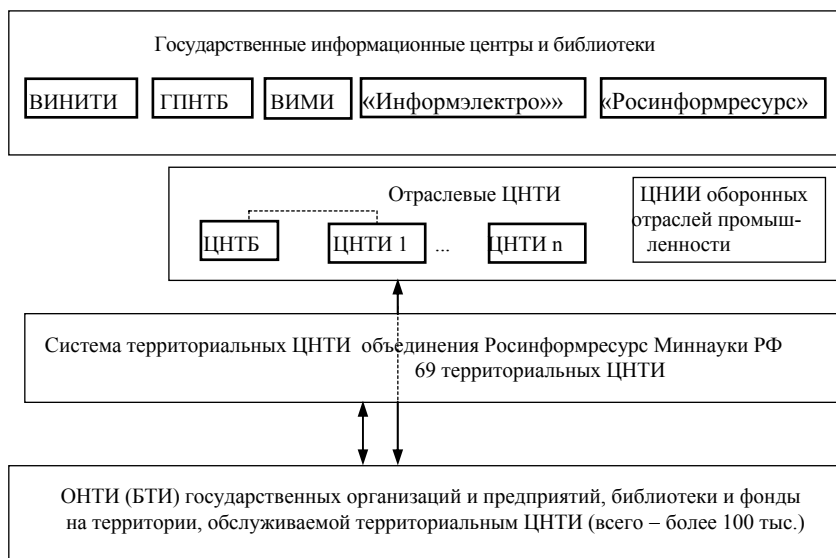


Рис. 7.9. Структура ГСНТИ

Таблица 7.3

Распределение основных функций между создаваемыми порталами

Центры НТИ	Функции
Всероссийский институт научной и технической информации (ВИНИТИ) Российской академии наук и Минпромнауки России	Формирование, ведение и обеспечение доступа к ресурсам федерального информационного фонда реферативной информации по опубликованным документам в области точных, естественных и технических наук, навигация по всем ресурсам ГСНТИ, включая ресурсы ВНИЦ (федеральный фонд по непубликуемым документам - отчетам о НИОКР, диссертациям, алгоритмам и программам), ресурсы ИНИОН РАН и ВНИИКИ взаимодействие с порталами РАН и Минобразования России, РФФИ, а также порталами, обеспечивающими пользователю выход на зарубежные ресурсы НТИ;

Центры НТИ	Функции
Государственная публичная научно-техническая библиотека (ГПНТБ) России Минпромнауки России	Формирование, ведение и обеспечение доступа к федеральному фонду библиографической информации по отечественной и зарубежной научно-технической литературе, ведение автоматизированной системы российского сводного каталога по научно-технической литературе, навигация по ресурсам научно-технических библиотек ГСНТИ и страны
Объединение "РОСИНФОРМРЕСУРС" Минпромнауки России	Формирование, ведение и организация использования региональных информационных фондов, баз и банков данных, как составной части государственных ресурсов научно-технической информации, включая базы данных о результатах научно-технической и инновационной деятельности регионов России; мониторинг патентной, технологической и производственной деятельности предприятий и товарных рынков в регионах; обеспечение доступа заинтересованных пользователей в регионах к базам и банкам данных ГСНТИ, зарубежным и международным базам и банкам данных.
Федеральное государственное унитарное предприятие «Всероссийский институт межотраслевой информации» (ВИМИ) Минпромнауки России	Формирование, ведение и обеспечение доступа к базам и банкам по научно-исследовательским, опытно-конструкторским работам и результатам научно-технической деятельности оборонного комплекса, навигация по аналогичным ресурсам, формируемым другими организациями ГСНТИ и страны
Институт промышленного развития "ИНФОРМЭЛЕКТРО" Минпромнауки России	Формирование, ведение и обеспечение доступа к информационным фондам, базам и банкам данных по результатам научно-технической и производственной деятельности, технико-экономической информации и результатам технико-экономических исследований в гражданских отраслях промышленности; информационно-аналитическое обеспечение по этим вопросам заинтересованных федеральных органов власти и хозяйствующих субъектов

Таким образом, с точки зрения теории систем и структур **М. Месаровича** в рассмотренной концепции предложена распределенная стратифицированная структура, координация в которой осуществляется пятью ведущими центрами НТИ в соответствии с их функциональными особенностями и возможно-

стями.. По замыслу авторов Концепции распределенная информационная система ГСНТИ должна представлять собой целостную технологическую и программную среду взаимодействия организаций по обмену электронными информационными ресурсами или информацией об этих ресурсах на основе единых принципов и открытых общепринятых стандартов. Осуществление этого проекта позволяет создать интегрированную информационную систему, объединяющую федеральные, отраслевые и региональные информационных ресурсы и обеспечивающую в рамках единых стандартов сбор, накопление, обработку, поиск и представление информации в интересах государственных органов власти, субъектов экономики, науки и образования, общества в целом.

Опыт функционирования ГСНТИ последних лет свидетельствует о том, что за прошедшие годы при отсутствии централизованного управления исчезли совсем или изменили характер своей практической деятельности десятки отраслевых информационных институтов и сотни центров научно-технической информации на предприятиях. Возникла проблема получения достоверной информации непосредственно от предприятий, отражающей их состояние, возможности и результаты деятельности. Появление множества частных информационных фирм и агентств, развитие Интернета пока не обеспечивают потребностей как самих участников производства, так и государственных органов управления. Различия в принципах, структуре, составе описания и представления информации, используемой терминологии и поисковых систем приводит к несопоставимости данных, предоставляемых различными производителями информации. Нет единой телекоммуникационной и информационной среды, ориентированной на единые стандарты передачи и обработки информации.

Кроме того, распределенная структура с принятыми принципами координации функций пятью ведущими центрами НТИ решает задачи координации сбора и хранения информации, но не решает проблему организации доступа к информации.

Иными словами возникла ситуация, подобная одному из вариантов структуры ОГАС, который предлагал **В.М. Глушков**. Практическую реализацию варианта создания какой-то специализированной диспетчерской службы или системы диспетчерских служб ОГСДС в этой концепции не удалось осуществить.

В соответствии с закономерностями теории систем в системе важны не только элементы, связи, цели, но и «наблюдатель» по **У.Р. Эшби**, «стейкхолдеры по **Р. Акоффу**, т.е. лица, принимающие решение (ЛПР).

В первоначальной структуре ГСНТИ, приведенной в главе 6, функции централизованной координации возлагались на Комитет при Президенте по политике и информатизации.

Но в последующих постановлениях и приказах ЛПР «опустили» вначале на уровень Министерства по науке и технической политике, а затем – на уровень организации, находящейся внутри системы (ВИНИТИ) и не обладающей необходимыми правами (компетенциями) для принятия решений по доступу ко всем видам информации, которые должна объединять единая информационная система.

В этих условиях было принято решение о коллективном органе по типу стейкхолдеров в теории управления ИТ-проектами, для чего был создан Координационный совет.

Для усиления координации при образовании коалиций **М.Месарович** предлагает эшелонированную структуру, в которой информационные и координационные связи должны быть хотя бы прописаны в соответствующих нормативных документов-договоров.

Разработка таких нормативных документов должно базироваться на предварительном исследовании методик и алгоритмов поиска информации для решения соответствующих задач. Эта идея начинает реализовываться

Например, для организации функционирования распределенной стратифицированной структуры может быть использована идея, предложенная при решении задачи предоставления населению региона различных пособий, компенсаций, социальных вы-

плат, документов и т.д., которая поставлена как задача автоматизации сбора необходимых сведений и документов и предоставления их лицу, принимающему решение, в электронном виде.

Задача может решаться путем создания системы, осуществляющей поддержку принятия решений в режиме инициативного информационного обслуживания. Организация информационного обеспечения процессов принятия управленческих решений и информационных потребностей пользователей при обращении ЛПР или другого лица к информационной системе должно инициировать процесс, с помощью которого обеспечивается поиск необходимой пользователю информации на основе использования территориально распределенных информационных массивов. В этих условиях задача информационного обеспечения процессов принятия решения представляет собой задачу формирования алгоритмов вызова информационных массивов, хранящихся и актуализируемых в различных распределенных и независимых базах данных и информационных системах¹⁷¹.

Для реализации такой информационной системы в условиях распределенных баз данных необходимо на основе полученных алгоритмов разрабатывать соответствующие регламенты.

Вопросы для самоконтроля:

- 1 Новые направления развития теории информационных систем.
2. Понятие об архитектуре информационной системы предприятия.
3. Понятие о ситуационных центрах.
4. Направления развития теории информационных систем на основе исследуемых в теории систем подходов к их проектированию и развитию.
5. Концепция многоуровневой информационной системы, базирующаяся на определении системы, отображающем системно-целевой подход.
6. Перспективы создания единой информационной системы для обеспечения научных исследований.

¹⁷¹ *Тибилова Г.С.* Применение системного анализа в проектировании интегрированной автоматизированной системы «Одно окно» / Г. С. Тибилова // В сб. трудов XII международной научно-практической конференции: «Системный анализ в проектировании и управлении». – СПб.: Изд-во СПбГПУ, 2008. – С.142-144.

ЗАКЛЮЧЕНИЕ

Анализ существующего опыта и теоретических исследований в области информатики и разработки комплексных средств автоматизации различного вида и назначения показывает, что первоначально, в основном, использовался термин «система», было принято говорить о разработке информационных систем, информационно-управляющих систем, интегрированных и интегральных информационных систем и т.п.

Однако в современных условиях, как правило, для создания информационного обеспечения управления организациями и сложными техническим объектами используются разнообразные автоматизированные источники информации, как разрабатываемые самостоятельно, так и создаваемые на основе готовых программных продуктов (обычно называемых также информационными системами, корпоративными информационными системами), а нередко используется и информация из внешних баз и хранилищ данных. Поэтому, удобнее использовать термин «комплексы» – не только применительно к сложным техническим автоматизированным комплексам, но и применительно к информационному обеспечению управления предприятиями и организациями. При этом следует понимать, что термин «комплекс» представляет собой совокупность разного рода средств и источников информации, но комплекс может первоначально и не обладать общесистемными свойствами, в них может и не возникнуть эффект эмерджентности. В то же время, используя термин «комплекс», обычно все же подразумевают, что информационный комплекс должен быть ориентирован на достижение определенных целей, его функционирование должно подчиняться определенным закономерностям.

Для разработки различных информационных систем и комплексов используются нередко специфические принципы, специальная терминология. В то же время есть и общие принципы,

определяемые теорией систем и теориями информационных систем разного вида и назначения.

Поэтому представляется полезным провести краткий анализ современного состояния и перспектив развития теории информационных систем как научного направления, которое может стать основой разработки единой теории, полезной для разработки и реализации информационных комплексов.

Становление любой науки проходит 4 стадии. Интерпретируем состояние реализации этих стадий для становления теории информационных систем как научного направления.

1. Определение места теории информационных систем среди других научных направлений. Систематизация и классификация ИС.

На основе приведенных в гл. гл. 1 и 3 сведений о структуризации информатики, определении места в этой теории и классификации информационных систем можно сделать вывод, что первую стадию теория информационных систем как наука прошла.

В то же время представляется, что принятую на рис. 3.1 классификацию ИС можно в свою очередь структурировать более детально. Развития и постоянной корректировки требует классификация ИС специального назначения, поскольку в настоящее время информационные технологии активно развиваются и на их основе создаются новые программные продукты и платформы.

1. Углубленное изучение информационных систем, их элементов, структуры и принципов функционирования.

В структуре информатики выделены *информационные элементы, информационные процессы и информационные системы*.

Виды, структуры, возможности **информационных элементов** (включая формы регистрации и хранения информации, технические средства сбора, передачи, хранения, обработки, представления информации), в настоящее время достаточно хорошо изучены и представлены в литературе.

Рассматриваются во многих монографиях и учебных пособиях, изучаются в различных дисциплинах – «Вычислительные системы, сети и телекоммуникации», «Операционные системы», «Базы данных», «Программная инженерия», «Алгоритмизация и программирование» и т. п.

Для исследования **информационных процессов** активизировалось применение предложенного в 1970-е гг. функционально-технологического (процессного) подхода, реализуемого на основе предложенной в 1990-е гг. методологии SADT и семейства IDEF-технологий, которые служат полезным инструментом для анализа процессов в информационных системах.

Изучаются специальными дисциплинами – «Информационные системы и технологии», «Проектирование информационных систем» «Моделирование бизнес-процессов».

Кроме того, информационные процессы как процессы возникновения и распространения научно-технической информации исследуются на основе закономерностей информетрии, что рассматривается в гл. 2.

В данном учебном пособии рассмотрены основные сведения об **информационных системах**: их виды и особенности построения и проектирования, теоретические основы разработки АИС и АСУ предприятия (организации), включая принципы формирования структуры функциональной и обеспечивающей частей, принципы выбора готовых программных продуктов (гл. 4); приведены основные сведения о документальных информационно-поисковых системах (ДИПС), включая понятие об информационном поиске и информационно-поисковой системе, информационно-поисковом языке, системе индексирования, логике ИПС, критериях смыслового соответствия, критериях оценки ИПС (гл. 5); дано представление о документально-фактографических информационно-поисковых системах (ДФИПС), о системе нормативно-методического обеспечения управления (СНМОУ) предприятия (организации) и ее автоматизированном варианте (АС-НМОУ) как информационной системе типа ДФИПС (гл. 6).

3. Формирование научной методологии и методик исследования информационных систем и их компонентов, использование методов точных наук, формальных методов и моделей.

К настоящему времени в информатике и теории информационных систем получен ряд методологических и методических результатов:

- в 1970-е гг. для управления разработками АСУ и ГАСНТИ подготовлены Общепромышленные методические материалы, регламентирующие разработку АИС и АСУ, Единый порядок и технические условия разработки и внедрения автоматизированных подсистем обработки, поиска, хранения, выдачи и передачи информации, Техническое задание на совершенствование Единой системы

научно-технической информации в стране, которые можно считать стандартами того периода;

- в 1990-е гг. разработана методология SADT – (Structured Analysis and Design Technique) – структурный анализ и проектирование, представляющая собой совокупность методов, правил и процедур, предназначенных для построения функциональной модели объекта какой-либо предметной области. Компьютерная реализация методологии SADT получила название IDEF (Icam DEFinition или Integrated DEFinition). Для реализации моделей применяются автоматизированные средства – BPWin, ARIS, язык UML (Unified Modeling Language – Унифицированный язык моделирования);

- для управления ИТ-проектами корпоративных информационных систем предприятий разработаны разнообразные многочисленные стандарты ¹⁷²;

- разработаны подходы к проектированию систем и методики проектирования и выбора информационных систем для конкретных организаций (гл. 5 и анализ подходов в разделе 7.3); при разработке ИС применяют расчетные модели оценки их эффективности;

- при создании автоматизированных систем научно-технической информации разработана теория информационного поиска и документальных информационно-поисковых систем (гл. 5);

- при создании систем нормативно-методического обеспечения управления разработан подход, принципы создания многоконтурных автоматизированных систем нормативно-методического обеспечения управления (гл. 6);

- исследованы *закономерности информетрии* (гл. 2), базирующиеся на применении статистических методов и помогающие в исследовании информационных потоков и совершенствовании процесса информационного поиска;

- при разработке структуры обеспечивающей части АИС применен подход постепенной формализации модели, основанный на применении методов дискретной математики для доказательства адекватности формализованных модели;

- разработана концепция многоуровневой информационной системы (раздел 7.5);

¹⁷² **Сатунина А. Е.** Управление проектами корпоративной информационной системы предприятия: учеб. пособие / А. Е. Сатунина, Л. А. Сысоева. – М.: Финансы и статистика, ИНФРА-М, 2009. – 352 с. (см. перечень стандартов на с. 338–340).

- для исследования взаимосвязей между стратами многоуровневой информационной системы предложено использовать формальные модели, основанные на применении методов организации сложных экспертиз, базирующихся на идее метода решающих матриц *Г. С. Поспелова* и информационном подходе к анализу систем *А. А. Денисова* (раздел 7.5).

Таким образом, в теории информационных систем применяются статистические методы, некоторые расчетные модели, методы дискретной математики, модели организации сложных экспертиз, основанные на формализованных методах.

В то же время формальные математические методы и оптимизационные модели в современной информатике и теории информационных систем применяются недостаточно, т. е. эту стадию развития теории информационных систем еще нельзя считать завершенной.

4. Развитие теоретического компонента – выработка понятий, категорий, методологических установок, создание теоретических концепций, формирование категориального аппарата.

В современной литературе существует множество определенных информатики и информационных систем. Это происходит оттого, что данная область знаний относительно новая и соответствующий понятийный аппарат не совсем устоялся.

В процессе разработке АСУ и АИС сформировалась определенная терминология (гл. 4). При создании автоматизированных систем научнотехнической информации разработана теория информационного поиска, теория документальных и документально-фактографических информационно-поисковых систем, введен ряд специальных понятий (ПОД и ПОЗ, парадигматические и синтагматические отношения, критерии смыслового соответствия, критерии оценки – релевантность, пертинентность и т. д. (гл. 5). При разработке специализированных информационных систем применяется терминология, удобная для конкретных приложений.

Таким образом, к настоящему времени сложились комплексы понятий, категорий, методологических установок, теоретических концепций, ориентированных на разработку и эксплуатацию информационных систем разного вида и назначения. Подготовлен словарь-справочник [11], учебники и учебные пособия [2, 11, 14, 15, 21 и др.].

В то же время обобщающие понятия категориального аппарата теории систем, обеспечивающие взаимопонимание специалистов различных областей знаний, в независимо складывающихся теориях информационных систем разных видов трактуются неодинаково.

Например, понятие «**цель**».

Применительно к информационным системам *в технических* приложениях трактуется как требования к проектируемой системе.

В *социально-экономических организациях* *цель* трактуется и как *назначение, миссия* (для внешней среды), и как *цели и функции* организации, и как *потребности* ее сотрудников. Для того, чтобы индивидуальные потребности были подчинены целям и задачам организации, в теории автоматизированных систем управления введено понятие *функциональной части АСУ*, что позволяло проще решить проблему целей и потребностей. Структура ФЧ АСУ формируется на основе анализа целей и функций организации и включает подсистемы и задачи, подлежащие автоматизации в соответствующий период планирования работ по созданию АСУ (для удобства вводилось понятие первой, второй и т. п. очереди АСУ).

В теории *научно-технической информации, документальных ИПС* вместо понятия «цель» вводится понятие «*информационные потребности*», которое первоначально было довольно жестко «привязано» к индивидуальным потребностям, но в последующем стало исследоваться применительно к потребностям организаций.

Различные названия присваивают неделимым **элементам** или **компонентам** системы и **способам их отображения**: в АИС – *базы данных и кодирование*; в документальных ИПС – *информационные массивы и индексирование* с помощью информационно-поискового языка в форме ПОД и ПОЗ.

По разному называют и исследуют **информационные процессы**: *документальные информационные потоки* в документальных системах НТИ, *бизнес-процессы* в АИС и АСУ, *организационно-технологические процедуры* подготовки и реализации управленческих решений в АСНМОУ.

Неоднозначно используется и понятие «**структура**».

Если в системах научно-технической информации сохраняется понятие *структуры* ИПС, то для объединения разнообразных баз данных, используемых для управления предприятием, введено понятие *архитек-*

туры предприятия, архитектуры приложений, архитектуры данных, технологической архитектуры, что позволяет кратко охарактеризовать структуры различных страт многоуровневой ИС – страты данных, программных приложений, клиент-серверов, технологической страты, ИС предприятия в целом (архитектура предприятия). При этом понятия *структуры* и *архитектуры* используются неоднозначно: либо понятие «структура» – конфигурация ИС («структура архитектуры» у Захмана), либо существуют точка зрения, в соответствии с которой под архитектурой понимается описание системы с точки зрения конечных пользователей и интерфейсов взаимодействия с внешней средой, т. е. как внешний взгляд на ИС, а структуру ИС описывают в виде взаимодействующих между собой подсистем, т. е. как внутреннюю конфигурацию ИС. При этом к структура представляется в форме древовидной или стратифицированной иерархии, и полученные структурные единицы, в свою очередь, могут быть представлены в виде архитектурного описания по отношению к внешним структурным единицам.

Благодаря возросшим возможностям вычислительной техники, и особенно в результате появления персональных ЭВМ, первоначально развивающиеся самостоятельно *теории построения и функционирования АИС* и *теории поиска в документальных ИПС* сближаются, происходит взаимное проникновение понятий, принципов и теоретических положений.

В частности, по мере роста объемов и усложнения структуры АИС возникает необходимость в заимствовании понятий теории информационного поиска, разработанной ранее для систем НТИ (в частности, понятия *тезауруса*, *многоконтурности* ИПС и т. п.).

В свою очередь, в решении проблемы разработки ИПЯ с высокими смысловыражающими возможностями может помочь структуризация, которая исходно была основой классификации и кодирования информации в АИС и АСУ, а при разработке систем НТИ могут оказаться полезными понятия *функциональной* и *обеспечивающей* частей, введенные в нормативных документах по разработке АИС и АСУ.

Специалисты различных областей знаний (философы, экономисты, специалисты по разработке автоматизированных и программных средств и т. п.) исследуют возможности создания единой теории, которая стала бы основой создания информационных комплексов различного вида и назначения. Ведутся поиски на-

звания такой науки. Для названия единой науки предлагаются названия «Теория информационных систем», «Теория информационных систем и процессов». Возможно название «Теория информационных комплексов». В то же время с точки зрения автора учебного пособия было бы целесообразно возродить понятие «информатика» в первоначальном определении ее **Ф. Е. Темниковым** как комплексной теории информационных элементов, теории информационных процессов и теории информационных систем (кратко охарактеризованного в гл. 1),

Для развития такой информатики можно использовать предложенные в гл. 7 концепции, базирующиеся на анализе концепции *архитектуры предприятия*, концепции *ситуационных центров*, возможностей применения для развития теории информационных систем подходов, принципов и моделей *теории систем* и *системного анализа*, и в частности, проведен анализ подходов к проектированию систем и возможности их реализации при создании информационных систем; анализ определений системы и обобщенное определение информационной системы, реализующее системно-целевой подход, и на этой основе – *концепция многоуровневой информационной системы*.

Важным направлением развития информатики как теории информационных систем и процессов является развитие *методов семантического анализа информации*, которые исследуются в самостоятельно развивающихся направлениях теории информационных систем – теориях *интеллектуальных систем* и *анализа текстов*. Представляется также перспективным продолжить исследования семантики текстовой информации, начатые в 1960–1970-е гг. на базе математической лингвистики и приостановленные развитием средств сети Интернет, обеспечивающих поиск избыточной информации, но не решающих проблему помощи пользователю в ее семантическом анализе.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. **Волкова В. Н.** Теория систем и системный анализ: учебник / В. Н. Волкова, А. А. Денисов. – М.: Изд-во Юрайт, 2012. – 579 с.
2. **Волкова В. Н.** Теоретические основы информационных систем / В. Н. Волкова. – СПб.: Изд-во Политехн. ун-та, 2012. – 228 с.
3. **Волкова В. Н.** Методы организации сложных экспертиз: учеб. пособие / В. Н. Волкова, А. А. Денисов. – СПб.: Изд-во Политехн. ун-та, 2010. – 128 с.
4. **Горькова В. И.** Информетрия: Количественные методы в научно-технической информации / В. И. Горькова // Итоги науки и техники. Сер. Информатика. Т. 10. – М.: ВИНТИ, 1988. – 328 с.
5. **Данилин А.** Архитектура и Стратегия. «Инь» и «Янь» информационных технологий / А. Данилин, А. Слюсаренко. – М.: Интернет-университет информационных технологий – ИНТУИТ.ру, 2005 // razum.ru/komp/59613-danilin-a-slyusarenko-a.
6. **Денисов А. А.** Современные проблемы системного анализа: учебник. / А. А. Денисов. – 3-е изд., перераб. и доп. – СПб.: Изд-во Политехн. ун-та, 2008. – 304 с.
7. **Козлов В. Н.** Теория информационных систем: учеб. пособие / В. Н. Козлов, В. Ф. Кисоржевский, Б.И. Морозов. – СПб.: Изд-во СПбГПУ, 2003. – 210 с.
8. **Мидоу Ч.** Анализ информационно-поисковых систем / Ч. Мидоу. – М.: Мир, 1970. – 368 с.
9. **Михайлов А. И.** Основы информатики / А. И. Михайлов, А. И. Черный, Р. С. Гиляревский. – М.: Наука, 1968. – 756 с.
10. **Михайлов А. И.** Научные коммуникации и информатика / А. И. Михайлов, А. И. Черный, Р. С. Гиляревский. – М.: Наука, 1976. – 435 с.
11. **Прикладная** информатика: Справочник: учеб. пособие / Под ред. В.Н. Волковой и В.Н. Юрьева. – М.: Финансы и статистика, ИНФРА-М, 2008. – 768 с.
12. **Сешнс Р.** Сравнение четырех ведущих методологий построения Архитектуры / Р. Сешнс // <http://msdn.microsoft.com/ru-ru/library/ee914379.aspx> Май 2007.

13. **Системный** анализ и принятие решений / Словарь-справочник: учеб. пособие / Под ред. В.Н. Волковой и В.Н. Козлова. – М.: Высшая школа, 2004. – 616 с.

14. **Смирнова Г. Н.** Проектирование экономических информационных систем: учебник / Г. Н. Смирнова, А. А. Сорокин, Ю. Ф. Тельнов / Под ред. Ю.Ф. Тельнова. – М.: Финансы и статистика, 2005. – 512 с.

15. **Тельнов Ю. Ф.** Реинжиниринг бизнес-процессов / Ю.Ф. Тельнов. – М.: Финансы и статистика, 2004. – 320 с.

16. **Теория** систем и системный анализ в управлении организациями: Справочник: учеб. пособие / Под ред. В.Н. Волковой и А.А. Емельянова. – М.: Финансы и статистика, ИНФРА-М, 2009. – 848 с.

17. **Черный А. И.** Введение в теорию информационного поиска / А. И. Черный. – М.: Наука, 1975. – 298 с.

18. **Шемакин Ю. И.** Компьютерная семантика / Ю. И. Шемакин, А. А. Романов. – М.: Научно-образовательный центр «Школа Китайгородской», 1995. – 343 с.

19. **Шрайберг Я. Л.** Автоматизированные библиотечно-информационные системы России: состояние, выбор, внедрение, развитие / Я. Л. Шрайберг, В. С. Воройский. – М.: Изд-во «Либерия», 1996. – 271 с.

20. **Шрейдер Ю. А.** Логика знаковых систем / Ю. А. Шрейдер. – М.: Знание, 1974. – 43 с.

21. **Юрьев В. Н.** Информационные системы в экономике: учебник / В. Н. Юрьев, В. Н. Волкова. – СПб.: Изд-во Политехн. ун-та, 2006. – 538 с.