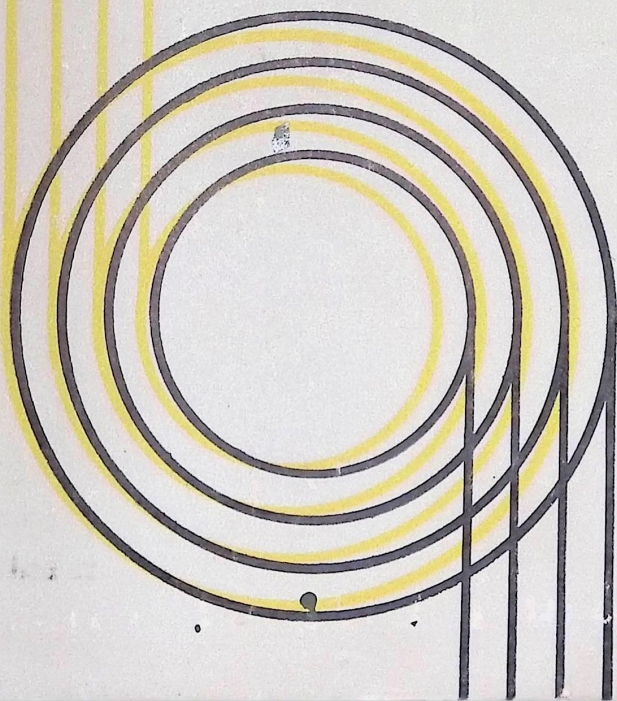
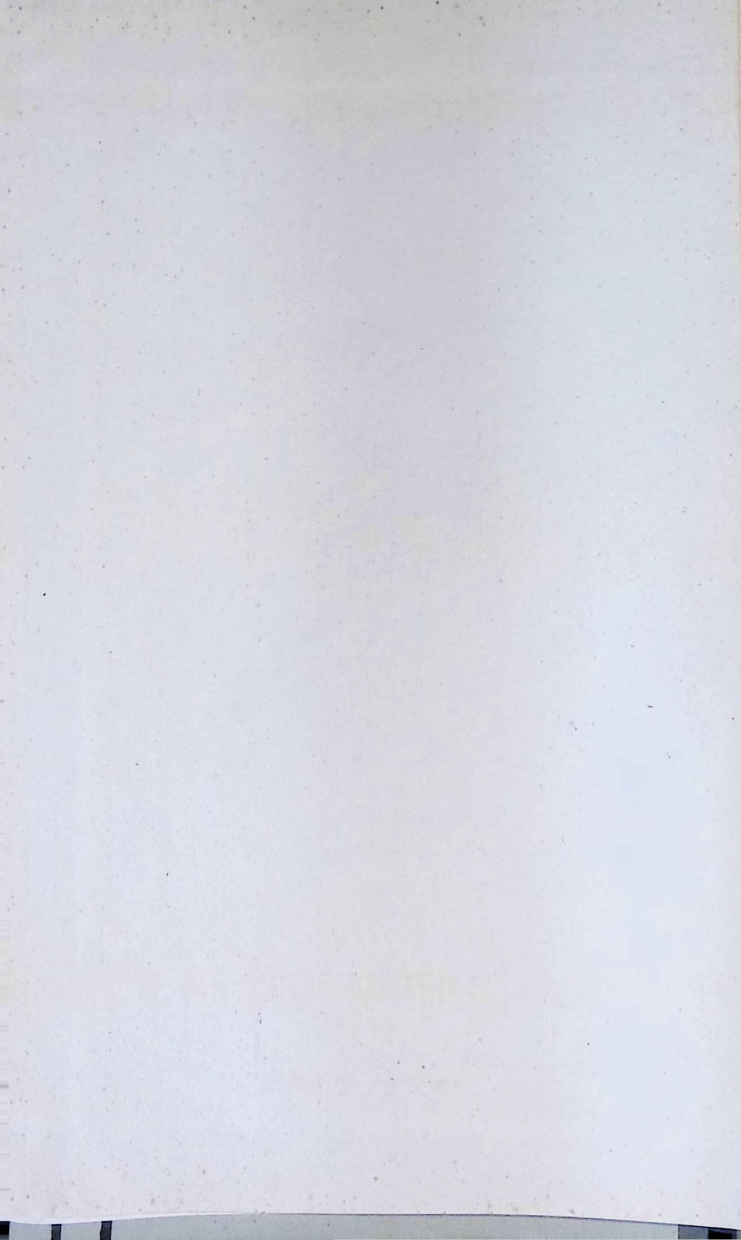


681
С-34

Л. Ю. Симанаускас З. П. Бразайтис

ОСНОВЫ
ПРОЕКТИРОВАНИЯ
МАШИННОЙ
ОБРАБОТКИ
ДАННЫХ



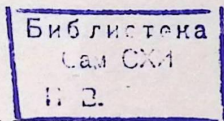


Л. Ю. Симанаускас,
З. П. Бразайтис

681
С-34

ОСНОВЫ ПРОЕКТИРОВАНИЯ МАШИННОЙ ОБРАБОТКИ ДАННЫХ

Допущено Министерством высшего и среднего специального образования СССР в качестве учебника для студентов вузов, обучающихся по специальности «Организация механизированной обработки экономической информации»



МОСКВА «ФИНАНСЫ И СТАТИСТИКА» 1982

С37 Симанаускас Л. Ю., Бразайтис З. П.
Основы проектирования машинной обработки данных:
Учебник. — М.: Финансы и статистика, 1982. — 207 с., ил.
В пер.: 65 к.

В учебнике рассмотрены вопросы применения вычислительной техники в промышленности и других отраслях народного хозяйства. Характеризуются данные, их системы, процессы получения, подготовки и использования данных. Приведены расчеты эффективности систем обработки данных. Проектирование обработки данных решено преимущественно на ЕС ЭВМ.

Для подготовки инженеров-проектировщиков вычислительных установок.

С 0604020101—016 141—82
010(01)—82

ББК 32.97
6Ф7.3

© Издательство «Финансы и статистика», 1982

ПРЕДИСЛОВИЕ

Совершенствование управления народным хозяйством и отдельными его звеньями тесно связано с внедрением новых методов, позволяющих лучше планировать и управлять работой предприятий, объединений, отдельных отраслей и народного хозяйства в целом. К ним относятся методы планирования, математического (линейного, нелинейного, динамического) программирования, математической статистики, исследования операций и др.

Новые методы, позволяющие повысить эффективность хозяйственной деятельности, в свою очередь требуют выполнения более сложных и трудоемких расчетов, вызывают рост объемов перерабатываемой информации. Эту информацию необходимо фиксировать на бумаге, машинных носителях, накапливать, хранить, обновлять и обрабатывать в таких объемах и в такие короткие сроки, что часто эта работа может быть выполнена лишь с применением современных технических средств сбора и обработки информации, и в первую очередь средств вычислительной техники (СВТ).

Опыт показывает, что повышения эффективности управления можно достичь при комплексном решении всех вопросов по совершенствованию планирования, учета и управления работой определенных объектов (предприятий, объединений, отраслей), системной разработке методов и средств реализации мероприятий по созданию и внедрению автоматизированных систем управления (АСУ).

Наиболее важными для организации внедрения средств вычислительной техники являются подсистемы, непосредственно обеспечивающие получение информации, необходимой для органов управления. Это в первую очередь подсистемы информационного, технического, математического и программного обеспечения. Совокупность этих подсистем составляет основу автоматизированной системы обработки данных (АСОД).

Внедрение мероприятий по применению СВТ требует проведения определенного круга работ по их проектированию. Проектирование машинной обработки данных имеет свои теоретические основы и методологию. Оно охватывает некоторую область научных знаний, которая составляет объект настоящего курса. Содержание курса сводится к рассмотрению методов и работ по организации данных, технологических процессов их движения и преобразования, по разработке и внедрению систем такой обработки.

Последовательность изложения материала основана на том, что разработке проектов должно предшествовать выявление осо-

бенностей экономической информации, достоверности данных, а проектирование должно осуществляться с обоснованием предложений по выбору вариантов проектных решений на всех его этапах. Поэтому после ознакомления с основами применения СВТ и проектирования в учебнике излагаются вопросы эффективности применения СВТ, оценки экономической информации и затрат на обработку данных. Проектирование технологии машинной обработки данных рассматривается после вопросов кодирования и размещения данных на носителях, т. е. прежде всего рассматриваются частные вопросы проектирования, а затем — общеорганизационные вопросы, которые именно в такой последовательности решаются при разработке автоматизированных систем машинной обработки данных. Это вызвано желанием обеспечить рациональный путь усвоения студентами материала данного курса, акцентировать их внимание на самые важные вопросы проектирования, приучить студентов выбирать наилучшие проектные решения, их обосновывать с тем, чтобы повысить качество и эффективность разрабатываемых систем.

Учебник предназначен для студентов вузов, обучающихся по специальности «Организация механизированной обработки экономической информации». Материал изложен с учетом того, что студенты уже знакомы с основами алгоритмизации и программирования, техническими средствами обработки информации, программным обеспечением ЭВМ, информационными системами и структурами данных.

ОСНОВЫ ПРИМЕНЕНИЯ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ ДЛЯ ОБРАБОТКИ ДАННЫХ

Глава 1

РАЗВИТИЕ ПРИМЕНЕНИЯ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ

1.1. Этапы развития применения вычислительной техники для механизации планово-учетных работ и автоматизации управления

Средства вычислительной техники создавались и совершенствовались аналогично другим приборам, устройствам и машинам, предназначенным для механизации (автоматизации) труда человека: когда возникала необходимость в них, появлялись и соответствующие технические возможности. На различных этапах развития науки и производства применялись соответствующие подходы, конструктивные решения, элементы. Так, например, первые машины были механическими. Затем появились электромеханические и электронные, менялись их габариты, вид. Но с точки зрения пользователей главными направлениями их совершенствования были: повышение уровня механизации и автоматизации выполнения определенных операций; создание новых средств ввода и вывода данных, запоминающих устройств; дальнейшее развитие возможностей запоминать и хранить данные с их последующей автоматической сортировкой, выборкой и обработкой.

С развитием СВТ менялась и степень централизации их использования: от индивидуального на каждом рабочем месте до централизованного — в специальных подразделениях и организациях (вычислительных установках).

Наиболее простыми по конструкции и применению были механические машины с ручным вводом (набором) исходных данных. Следует отметить, что первая такая машина, суммирующая машина Паскаля (1641 г.), была создана для выполнения расчетов, связанных со сбором налогов, т. е. для обработки экономических данных. Однако наиболее широкое применение для механизации отдельных расчетов нашли машины, предназначенные для выполнения всех четырех арифметических операций (клавишные вычислительные машины). Это прежде всего машины Лейбница и Якобсона (1670 г.), арифмометры (Томаса — 1818 г., Однера — 1880 г.). Затем появились суммирующие и другие машины с устройствами вывода результатов на бумагу (бумажную ленту, закладываемые

непосредственно в машину бланки документов). Позже переход от механических к электромеханическим и электронным клавишным вычислительным машинам позволил во многом улучшить и расширить их эксплуатационные возможности. Поэтому для выполнения отдельных расчетов, обработки первичных документов, проверки правильности таксировки и т. д. они успешно применяются и до сих пор.

Машины с ручным вводом исходных данных почти до конца XIX в. были практически единственным средством обработки данных при выполнении наиболее трудоемких и сложных расчетов. Характеризуя весь этот период (как первый этап применения СВТ), можно выделить следующие его особенности: использование машин, позволяющих выполнять лишь отдельные арифметические операции; ручной ввод исходных данных, набираемых непосредственно на технических устройствах при помощи рычагов, клавишей; возможность использования машин непосредственно на рабочем месте пользователя данных; возможность выполнения расчетов без проведения специальных подготовительных работ.

Значительные изменения в развитии применения СВТ и организации обработки данных способствовали появлению машин с автоматическим вводом исходных данных, фиксируемых на перфокартах, получивших название счетно-аналитических или перфорационных (перфорационных вычислительных) машин (ПВМ). Они с самого начала их появления предназначались для обработки больших массивов данных: первая машина была разработана Г. Голлеритом для обработки данных переписи населения 1890 г. (в США) и названа табулятором. В табуляторе Голлерита были реализованы возможности подсчета количества значений одного или нескольких признаков (подсчета единиц) и определенной группировки перфокарт. Позднее были разработаны и другие машины, предназначенные для обработки перфокарт. Они совершенствовались, развивались принципы, заложенные в первых ПВМ. Табулятор в дальнейшем превратился в суммирующую и счетно-пишущую машину. Пауерсом, Лангфордом, Булем и другими были разработаны также контрольные, сортировальные машины.

С начала XX в. ПВМ стали применяться для механизации учета на промышленных предприятиях: с 1904 г. — в заводской бухгалтерии на сталелитейном заводе в Пенсильвании (США), а через несколько лет — в крупных бухгалтериях Германии. В нашей стране на промышленных предприятиях они впервые были применены в Харькове (1925 г.). Через два года аналогичные машины были установлены в ЦСУ СССР и Наркомате путей сообщения. Еще через три года они уже использовались в Госбанке, Центросоюзе и других организациях. Для эксплуатации этих машин создавались машинносчетные станции (МСС). К 1932 г. в нашей стране таких МСС было создано 12, а за последующие шесть лет — еще 40.

На первых МСС применялись машины Голлерита и Пауерса. С 1935 г. было налажено производство отечественных счетно-ана-

литических машин (на Московском заводе счетно-аналитических машин). Использование ПВМ вполне можно считать вторым этапом применения СВТ. Основными особенностями применения ПВМ являются следующие: использование машин в определенном наборе (комплекте), где каждая из них предназначена для выполнения строго определенной работы (операции или их группы) в общем технологическом процессе обработки данных (подготовки массивов перфокарт с исходными данными, их упорядочения, дополнения другими данными, формирования результатных данных и их выдачи, печати); автоматизация ввода, упорядочения (сортировки) и вывода данных; необходимость специальной настройки отдельных машин комплекта, обрабатывающих данные, зафиксированные на перфокартах (составление схем настройки, коммутации). Все это требует проведения подготовительных работ по построению технологического процесса обработки данных, приспособлению для этого форм документов, разработке специальных условных обозначений (кодов) и т. п., централизации машин, операторов (и другого персонала), работающих на них в определенных подразделениях.

На Московском и Пензенском заводах счетно-аналитических машин, Вильнюсском заводе счетных машин и некоторых других заводах позднее было освоено производство всех основных видов ПВМ, в том числе и узкоспециализированных машин-автоматов: позиционных итоговых перфораторов, перфораторов-репродукторов, раскладочно-подборочных, расшифровочных и других машин. Для выполнения операций умножения и деления разрабатывались различные приставки, умножающие перфораторы. Некоторые из них производятся и эксплуатируются до настоящего времени. Научно-исследовательский институт счетного машиностроения (НИИсчетмаш), Специальное конструкторское бюро вычислительных машин Вильнюсского производственного объединения «Сигма» и другие организации внесли существенный вклад в совершенствование и развитие этих машин, пополнение их комплектов специальными машинами, что помогло значительно улучшить их применение. Так, например, разработанные ими электронные вычислители ЭВ80-3, ЭВ80-3М, ЭВП80-2 («Рута») в сочетании с другими машинами специального назначения, особенно с перфораторами-репродукторами и раскладочно-подборочными машинами, составили удачное сочетание средств для обработки двух или нескольких массивов с выполнением операций таксировки (умножения с использованием массивов постоянной информации — цен, норм и т. д.). Применение этих машин в комплекте с другими ПВМ позволило перейти к следующему этапу — комплексной механизации нормативных и планово-учетных работ.

Комплексность расчетов обеспечивалась применением хранимых на машинных носителях и многократно используемых массивов нормативно-справочной информации, информационной взаимосвязкой всех (в том числе и ранее механизированных) расчетов. На этом этапе развития применения СВТ возникла необходимость

и были созданы предпосылки для механизации расчетов по составлению конструкторских спецификаций, определению потребностей в материальных, трудовых и других ресурсах на изделия и производственную программу, для использования в учете нормативно-справочной и другой информации, заранее зафиксированной на машинных носителях. В дальнейшем этот опыт во многом помог при переходе к применению качественно новых машин, появившихся в 50—60-е годы нашего века. Первые такие машины были разработаны на базе электромагнитных реле — машины Цузе (в Германии, 1936—1941 гг.), МАРК-1 (по проекту Г. Айкена в США, 1939—1944 гг.). Однако вскоре выяснилось, что существенно улучшить характеристики этих машин, расширить их возможности и увеличить быстродействие можно только на базе использования электронных элементов. Первой действующей машиной такого типа была машина ЭНИАК, разработанная под руководством Дж. Маучли и Дж. Эккерта в США (1942—1945 гг.). Широко была известна и машина ЭДВАК, разработанная в США под руководством тех же авторов и Дж. Неймана (1945—1950 гг.).

Важно отметить, что эти машины были предназначены для выполнения расчетов при наличии небольших объемов входных и выходных данных. Поэтому по своим возможностям первые ЭВМ (ЭВМ первого поколения) для экономических расчетов, аналогичных выполняемым с помощью ПЧМ, т. е. связанных с обработкой больших массивов однородной информации, были мало пригодными. Однако благодаря их быстродействию, более высокой надежности и особенно наличию памяти, позволяющей хранить необходимые исходные и другие данные, осуществлять автоматический ввод и др., они быстро привлекли внимание экономистов и организаторов производства. Эти машины начали использоваться для решения некоторых локальных задач, в которых применялись экономико-математические методы, требующие выполнения сложных и трудоемких расчетов при их реализации.

Первые ЭВМ были со штекерным программированием (набираемыми на коммутационной доске программами). К ним относились создаваемые и в нашей стране ЭВМ, получившие название электронных вычислителей. Первой серийной ЭВМ с хранимой в памяти машины программой была машина УНИВАК, которая предназначалась в основном для обработки экономических данных (коммерческой информации). Она представляла интерес и как машина со сравнительно развитой системой контроля правильности передачи и обработки данных. С 1953 г. аналогичные машины стала выпускать фирма ИБМ (США). В 1954 г. начал выпуск машин серии ИБМ-650. Удачная конструкция способствовала их применению не только для инженерно-технических расчетов, но и для решения экономических задач, связанных с обработкой больших массивов информации. Одной из наиболее важных особенностей этих машин, способствовавших их широкому применению для обработки экономических данных, явилось наличие развитой системы периферийных устройств ввода-вывода и устройств памяти. В них

для ввода данных и программ использовались перфокарты и перфоленты, а в качестве внешних запоминающих устройств — блоки с магнитными лентами.

Совершенствование ЭВМ, их устройств ввода и вывода данных, развитие средств программирования способствовало переводу решения на них задач по комплексной механизации планово-учетных работ с использованием массивов нормативной, справочной и другой информации, постоянно хранимой на машинных носителях (в памяти системы). В нашей стране для этого оказались более пригодными ЭВМ второго поколения. Широкое применение нашли ЭВМ «Минск-22», «Минск-32», «Урал-11», «Урал-14», «Рута-110» и другие, появившиеся в 70-х годах. Дальнейшее их развитие и совершенствование применения привело к следующему этапу — автоматизированным системам управления. При создании первых таких систем основное внимание сосредоточивалось на постановке функциональных задач и организации их решения на ЭВМ. Поэтому первые АСУ были лишь машинными системами обработки данных, организованными по единой схеме, в которой процедуры обработки являются взаимообуславливаемыми элементами неразрывного процесса управления, т. е. автоматизированными системами обработки данных.

Начиная с восьмой пятилетки создание АСУ отражается в годовых и пятилетних народнохозяйственных планах. Поэтому часто по пятилеткам и рассматривается их развитие. В первых АСУ, которые создавались и внедрялись в восьмой пятилетке (по стране их было создано 151), в выделении подсистем и задач практически отсутствовали системный подход и классификационное единство. Например, в автоматизированных системах управления предприятиями и производственными объединениями (АСУП), несмотря на то, что подсистемы в основном выделялись по функциональному принципу, среди них были и такие, как «Нормативно-справочное хозяйство», «Бухгалтерский учет». Недостатком являлось и то, что отдельные подсистемы и задачи слабо взаимоувязывались друг с другом, первичная информация предусматривалась отдельно для каждой задачи или их небольших групп. Последнее свидетельствует об отсутствии в этих системах единства в формировании информационной базы системы.

В планах на девятую пятилетку впервые было намечено значительное развитие работ по созданию АСУ. К их осуществлению привлекались специалисты различных специальностей, поэтому во многом пригодился опыт, накопленный при создании и внедрении первых АСУ. На его основе был полнее определен и стабилизирован состав подсистем и задач. В этот период увеличилось количество различных технических средств сбора, подготовки и отображения информации, что позволило лучше организовать обмен данными между их источниками, пользователями и ЭВМ. Однако в организации информационного фонда системы, его формирования и использования наиболее существенные сдвиги были достигнуты лишь в десятой пятилетке, когда стали широко внедряться ЭВМ

Единой системы (ЕС ЭВМ) и другие машины с запоминающими устройствами прямого доступа на магнитных дисках. Характерной особенностью этого периода является переход к созданию и использованию типовых элементов систем, различных программных средств, позволяющих автоматизировать разработку программ для формирования информационных фондов и баз данных, решения функциональных задач управления (формирования результатов).

Большие перспективы развития применения СВТ с широким использованием экономико-математических методов предусмотрены в решениях XXVI съезда КПСС на одиннадцатую пятилетку и на период до 1990 г. Значительное увеличение производства средств вычислительной техники, расширение их ассортимента позволят успешно решать задачи дальнейшего повышения технического уровня, эффективности и качества АСУ, реально осуществить мероприятия по созданию общегосударственной автоматизированной системы сбора, передачи и обработки информации.

1.2. Эволюция работ по проектированию

Внедрению средств вычислительной техники всегда предшествуют подготовительные работы и разработка соответствующей организационно-распорядительной документации. Результаты их выполнения представляются в виде технической документации. Процесс подготовки решений по всем вопросам, связанным с размещением данных на их носителях, организацией движения и преобразования этих данных и разработкой соответствующей документации, принято называть *проектированием машинной обработки данных*.

Содержание работ по проектированию, а также самой проектной документации зависит от сущности проектируемых мероприятий и во многом определяется видом намечаемых к применению носителей данных и технических средств. Наиболее значительно различаются организация и осуществление проектирования применения КВМ, ПВМ и ЭВМ, что в основном определяется степенью механизации (автоматизации) обработки данных и управления.

При подготовке к внедрению клавишных вычислительных машин, с помощью которых можно выполнять лишь арифметические операции и запись данных, т. е. технических средств обработки первичных документов, существующие потоки движения данных и их носителей, а также процессы преобразования данных претерпевают незначительные изменения. Поэтому применение суммирующих, вычислительных, бухгалтерских и фактурных машин обычно не требует проведения больших проектировочных работ. По своему характеру эти работы сравнительно просты, но для их выполнения необходимы специальные знания о возможностях и способах настройки соответствующих машин. Проектируют применение КВМ обычно работники, организующие эксплуатацию или осуществляющие техническое обслуживание соответствующих машин.

При использовании перфорационных вычислительных машин появляются новые носители данных — перфокарты, которые способствуют более полной (чем с помощью КВМ) автоматизации формирования необходимой результатной информации. На них можно выполнять группировку (сортировку) данных, автоматическую печать результатов в виде сводных ведомостей — табуляграмм (отчетов). Здесь также появляются дополнительные технологические операции: кодирование (запись) данных на перфокарты и его контроль, т. е. перфорация и контроль перфорации. Применение таких машин связано с изменением потоков движения данных, разработкой их условных обозначений (кодов), технологических процессов, инструкций и других материалов, обеспечивающих их реализацию.

Примерами проектов применения ПВМ могут служить проекты механизации отдельных участков учетных работ: учета труда и заработной платы, учета материальных ценностей, готовой продукции, ее реализации, составления сводного баланса и др. Разработкой проектов занимались работники машиносчетных станций.

Содержание, объемы и организация проектировочных работ претерпели существенные изменения после того, как в основной комплект ПВМ начали включаться специальные машины-автоматы (перфораторы-репродукторы, раскладочно-подборочные машины, электронные вычислители), с помощью которых можно выполнять дополнительные операции по обработке массивов данных, зафиксированных на перфокартах, такие, как репродукция (получение массивов-копий), дублирование (копирование данных на другие перфокарты, т. е. снятие любого количества копий), объединение, разъединение, совместная обработка с выполнением операций таксировки, агрегирования (суммирования) данных и др.

Проектирование применения полного комплекта ПВМ — это проектирование создания картотек нормативно-справочных данных (норм, нормативов, ценников и др.), технологических процессов формирования и ведения их массивов на перфокартах, механизации сбора оперативных данных, комплексной их обработки. Это позволило расширить круг механизлируемых работ, осуществить комплексную механизацию не только учетных, но и плановых работ, а вместе с тем потребовало создания специальных проектных подразделений в проектно-конструкторских бюро и других организациях.

Кроме индивидуальных проектов существуют типовые, предназначенные для определенных групп предприятий или организаций. Они создаются специализированными проектными организациями. Так, например, научно-исследовательские и проектные организации Госбанка СССР, ЦСУ СССР, других министерств и ведомств разработали типовые проекты по механизации учета в конторах Госбанка, по разработке статистической отчетности и т. д. Наиболее широко известны типовые проекты по механизации основных участков учета на промышленных предприятиях (труда и заработной платы, материалов, основных средств, финансовых операций,

отгрузки и реализации продукции, аналитического и синтетического бухгалтерского учета).

Применение ЭВМ второго поколения привело к созданию автоматизированных систем управления. Сама методология проектирования таких систем с самого начала базировалась на системном подходе. Это способствовало выявлению основных проблем, требующих безотлагательного решения, выполнению проектировочных работ поэтапно, в определенной последовательности. Однако отсутствие необходимой теоретической подготовленности и опыта выполнения таких объемных и сложных по характеру работ, неподготовленность самих объектов и другие причины существенно тормозили проектирование, отрицательно влияли на качество его выполнения. Увеличившиеся объемы работ потребовали большого количества специалистов. Поэтому к проектированию АСУ было привлечено много работников, среди которых были и специалисты, не имеющие достаточного опыта и не учитывающие того, что уже было накоплено, или считавшие, что новые средства (ЭВМ) требуют и совершенно другой методологии проектирования. Это зачастую приводило к тому, что при разработке конкретных проектов по системной организации машинной обработки экономической информации в условиях АСУ не находили достаточного отражения такие сравнительно несложные, но важные вопросы, как кодирование технико-экономической информации, разработка форм документов и т. п. Слабо решались и вопросы организации технологических процессов, особенно вопросы организации подготовки первичных массивов на машинных носителях, формирования и ведения (обновления) массивов нормативно-справочной и другой постоянно хранимой информации. Низкое качество проектов почти всегда приводило к их переработке уже на стадии внедрения с тем, чтобы улучшить их технологичность, обеспечить достоверность данных. Все эти причины, а также индивидуальный подход к созданию каждой системы, в особенности их программного и информационного обеспечения, вели к тому, что системы проектировались долго и большим количеством людей.

Методика проектирования автоматизированных систем управления создавалась вместе с первыми системами, когда каждая проектная организация (подразделение) разрабатывала свои методические материалы. Первой из наиболее общих методик, которая подготовлена в 1971 г., были Общеотраслевые руководящие методические материалы по созданию автоматизированных систем управления предприятиями. В этих материалах определено понятие АСУП как системы управления с применением современных средств обработки данных и экономико-математических методов для регулярного решения основных задач управления производственно-хозяйственной деятельностью предприятия. Четко были выделены стадии создания систем — предпроектная, проектирования и внедрения системы. Определены также номенклатура и содержание разрабатываемой на каждой стадии документации: на предпроектной стадии предусмотрены проведение обследования пред-

приятия, разработка и утверждение технического задания; при проектировании предусмотрены разработка и утверждение технического проекта, а также разработка рабочего проекта. В первых методических материалах определены порядок проведения работ заказчиками и исполнителями, номенклатура и содержание организационно-распорядительной и технической документации. Много внимания уделено организационно-экономическим вопросам управления предприятиями, описанию функциональных подсистем и задач, методике выполнения работ на отдельных стадиях создания АСУП.

С развитием работ по созданию АСУ накапливался определенный опыт. Использование его позволило сократить затраты на создание последующих систем, осуществить типизацию и стандартизацию проектов, отдельных их частей. Это привело к уменьшению разнообразия систем, подсистем и задач, объединению их в определенные типовые группы, к унификации и типизации кодов, форм носителей данных, технологических процессов, алгоритмов и программ.

В 1971—1975 гг. достижения в области выбора автоматизируемых функций, постановки и типизации комплексов задач (участков работ) и их частей в условиях применения ЭВМ второго поколения были обобщены в типовых проектных решениях (ТПР). При их разработке главными элементами типизации принимались задачи (комплексы задач), процедуры их решения, определяемые как программные модули и программы. Основным элементом типизации был принят модуль как независимая часть алгоритма задачи, которая выполняет строго определенные функции обработки данных и контроля, а также имеет стандартные средства связи с другими частями системы. ТПР разработаны главным образом для организации и ведения фонда нормативно-справочной информации (НСИ) с определенной характеристикой форм документов и структуры массивов на перфоносителях и магнитных лентах по основным функциональным подсистемам: управления технической подготовкой производства (дается общая схема организации и ведения фонда НСИ), технико-экономического планирования, оперативного управления основным производством, управления материально-техническим снабжением, бухгалтерского учета, управления сбытом и реализацией продукции. Отдельные ТПР разработаны для проектирования комплексов технических средств и технологических процессов обработки данных в информационно-вычислительных центрах.

В ТПР по каждой задаче дается описание ее организационно-экономической сущности, исходных данных, выходных (результатных) данных, алгоритмов решения, программных модулей в виде блок-схем с указанием последовательности их выполнения, входных и выходных массивов. В отдельные блоки выделяются программы компоновки, сортировки, корректировки массивов, их обработки с получением результатных массивов и табуляграмм.

С появлением ЭВМ третьего поколения главное внимание на-

учно-исследовательских и проектных организаций переключилось на исследование вопросов, связанных с их применением. Основные принципы и методики выполнения работ по созданию АСУ на базе этих машин определены в многочисленных отраслевых и государственных стандартах. Они обобщены в новых отраслевых руководящих методических материалах по созданию АСУП, подготовленных в 1977 г. В них уточнены состав и содержание проектной и организационно-распорядительной документации, порядок выполнения работ. С учетом накопленного опыта предусмотрена возможность совмещения стадий технического и рабочего проектирования, т. е. объединения отдельных технических и рабочего проектов в один технорбочий проект. Также уточнены определения видов обеспечения АСУ (организационного, информационного, технического, программного, лингвистического, правового, математического и эргономического). Детально эти вопросы рассмотрены в последующих разделах учебника, где также изложены и требования соответствующих стандартов, других документов, касающихся определенных проектировочных работ, проектов и материалов, представляемых в них.

1.3. Основные направления совершенствования проектных работ

Проектирование в настоящее время в основном направлено на внедрение ЭВМ третьего поколения, современных средств сбора, передачи, отображения данных и непосредственной связи пользователей с ЭВМ.

Применение ЭВМ с запоминающими устройствами произвольного доступа позволяет не только переводить на машинное решение задачи, сводящиеся к обработке больших массивов данных, но создавать на объектах управления необходимые условия для применения экономико-математических и других методов, требующих выполнения большого количества сложных расчетов. Опыт показывает, что о более широком внедрении методов исследования операций, математической статистики и т. д. можно реально говорить только при переходе от АСУ как систем обработки данных и информационно-справочных систем к информационно-советующим системам управления, в которых основное место занимают задачи, решаемые в реальном масштабе времени, в режиме диалога (ответов на запросы пользователей). Для этого кроме существенных сдвигов по совершенствованию экономико-организационной деятельности необходимо создавать специальные информационные фонды нормативно-справочной и некоторой текущей (оперативной) информации, пригодной для непосредственного использования при решении задач такого типа. В методике проектирования переход к системам более высокого уровня выражается необходимостью организации информационного фонда системы как автоматизированной базы (или банка) данных, широким использованием специально разработанных или генерируемых программных средств, чему во многом способствуют большие сдвиги в совершенствовании

нии программного обеспечения. В таком направлении ведутся интенсивные работы по созданию и совершенствованию общесистемного программного обеспечения, разработке специального программного обеспечения в виде пакетов прикладных программ (ППП) по формированию, ведению и использованию единых баз данных, реализации экономико-математических методов, решению функциональных задач управления и их комплексов. ППП представляет собой комплекс взаимосвязанных прикладных программ со специальной организацией, который обеспечивает значительное повышение производительности труда программистов при создании конкретных программ.

Определенные особенности в проектирование вносит и то, что для сбора и кодирования первичных данных все более широкое применение получают технические средства записи данных непосредственно на магнитные носители. Также внедряются бухгалтерские и другие машины с различными электронными приставками и устройствами автоматической обработки данных и одновременной подготовки документов и машинных носителей и др.

Для повышения эффективности применения СВТ вместе с выявлением новых форм и методов организации данных, их движения и преобразования важно ускорить и сам процесс проектирования, сокращать затраты на разработку и эксплуатацию систем машинной обработки данных. Поэтому проектирование нужно постоянно совершенствовать, применять для его осуществления математические методы и ЭВМ, автоматизировать его.

Глава 2

ЭКОНОМИЧЕСКИЕ ДАННЫЕ И ОСОБЕННОСТИ ИХ ОБРАБОТКИ

2.1. Информационный аспект и информационное обеспечение управления

Управление представляет собой вид целенаправленной деятельности, обеспечивающей согласованность работы, выполняемой для достижения определенных результатов. Когда говорим об управлении, всегда имеем в виду конкретный объект и субъект управления (управляющую и управляемую системы).

Процесс управления конкретной деятельностью состоит в выполнении определенных функций. Они для различных объектов и субъектов управления могут быть совершенно различными, но имеют и много общих черт. Обычно для любого процесса управления характерны цикличность, устойчивость, но вместе с тем и динамичность, строгая последовательность этапов и основных групп операций при реализации соответствующих функций, непрерыв-

ность, требующая постоянства управленческой деятельности, с одной стороны, и дискретности — с другой. Все эти свойства присущи и управлению производством. Например, существует определенная последовательность этапов управления производством, нарушение которой ведет к снижению эффективности производства, иногда и к немалым ошибкам. Так может случиться, если решение принимается без достаточно полного учета реальных ситуаций или без достаточно четкого выяснения целей. С последовательностью этапов связана и дискретность управления производством, так как оно происходит путем выполнения отдельных автономных работ: ставятся цели, оценивается ситуация, анализируются проблемы, разрабатываются и принимаются решения и т. п.

Управление всегда осуществляется с помощью информации. Поэтому наряду с организационной и аналитической работой, выбором вариантов решения важное место занимает непосредственная работа с информацией. Ее можно назвать информационной работой. Более того, на совершенствование управленческой деятельности существенное влияние оказывают рационализация потоков информации, процессов ее движения и преобразования, организация системы информации, улучшение обеспечения информацией управляющих элементов. Поэтому вопросы управленческой и информационной деятельности являются тесно взаимоувязанными. При рассмотрении вопросов совершенствования организации и автоматизации обработки информации отправным пунктом должны служить организация и управление производством, повышение эффективности функционирования предприятия.

Остановимся на некоторых наиболее важных для дальнейшего изучения курса вопросах, которые помогут полнее выявить основные тенденции и направления совершенствования методологических принципов, организационных форм и структур управления для совершенствования системы информации объекта управления с использованием современных технических средств. При этом как раз и выявляются требования к повышению качественных характеристик информации, изменению содержания и сроков формирования (получения) всей необходимой информации, определяются и изменяются связи между результатной и исходной информацией.

Управление как процесс имеет свои статические и динамические характеристики. К первым относятся структура системы управления, организационное закрепление вариантов управленческого труда, распределение функций, установление связей между ними и т. п. Динамичность управления проявляется через его цикличность, т. е. повторяемость функций планирования и учета во времени, постоянное обновление производства и совершенствование методов выполнения соответствующих функций. Эта динамичность становится особенно яркой на современном этапе научно-технического прогресса, когда все более часто обновляется продукция, ускоряются, усложняются и совершенствуются процессы производства, что ведет к частым изменениям функций управления.

Основной элементарной структурной единицей общей системы народного хозяйства, ее производственной ячейкой выступает социалистическое производственное предприятие. Каждое предприятие — самостоятельное хозяйственное звено, характеризующееся организационным, производственно-техническим и экономическим единством. На базе нескольких предприятий, производящих определенный вид продукции или имеющих тесные производственные связи, могут быть организованы производственные объединения.

Перед освоением нового производства (или нового изделия) необходимо провести научно-исследовательские и проектные работы. Некоторые, наиболее простые из них, могут выполняться подразделениями и службами самого предприятия. Однако чаще всего их осуществляют специализированные организации — проектно-конструкторские бюро (ПКБ), специальные конструкторские бюро (СКБ), специальные конструкторско-технологические бюро (СКТБ), отраслевые научно-исследовательские институты (НИИ). Они разрабатывают новые виды изделий, новые, более совершенные (прогрессивные) технологические процессы их производства, определяют направления и тенденции научно-технического прогресса, способы и средства более полного удовлетворения растущих потребностей общества. На базе таких предприятий, проектных и научно-исследовательских организаций создаются научно-производственные объединения.

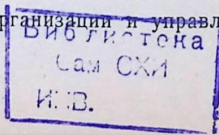
Производственно-хозяйственную деятельность предприятий направляют, координируют и контролируют соответствующие министерства (ведомства). Они обычно это организуют непосредственно через свои подразделения или другие подчиненные организации.

Структуры всех звеньев, распределение функций между ними в каждом конкретном случае должны быть достаточно четко определены. Их нужно выбрать так, чтобы они наилучшим образом способствовали реализации основных целей, конкретно выражаемых через технико-экономические показатели. Но со временем с развитием производства меняются и цели. Структура объекта и организационная система управления с развитием науки и техники, организации производства и управления постоянно меняются и совершенствуются, т. е. имеют динамический характер. Эти важные и актуальные вопросы постоянно находятся в центре внимания партии и правительства, научных и практических работников.

Основные мероприятия, проводимые в последние годы по совершенствованию структуры и организации управления производством, укрупненно можно объединить в следующие группы:

1. Совершенствование экономико-организационной деятельности внутри предприятий и производственных объединений. Эти мероприятия в первую очередь направлены на совершенствование методов и средств планирования и учета, хозяйственного механизма и его воздействия на повышение эффективности производства, на развитие хозрасчета.

2. Совершенствование организации и управления производством.



вом, его подготовкой и развитием. Это связано с тем, что ускорение внедрения достижений научно-технического прогресса все больше требует подготовительных, организационных и других работ, ведет к распределению труда путем специализации предприятий и организаций по отдельным видам производства и группам технологических процессов с завершением производства на головных предприятиях. Основным направлением совершенствования этих работ является создание и укрупнение научно-технических и научно-производственных объединений. Особенность этих работ состоит в том, что объединения должны создаваться не просто путем механического соединения нескольких предприятий, бюро, отраслевых институтов. Объединение непосредственно связано со специализацией и кооперацией производства, интеграцией производства и технологических процессов, перестройкой планирования, оперативного управления и учета, централизацией общего управления подготовкой основного производства, его развитием, вспомогательным производством, децентрализацией некоторых видов производства.

3. Совершенствование отраслевого управления. К этому относятся мероприятия по упразднению некоторых промежуточных звеньев в отраслевых министерствах. Суть таких реорганизаций заключается в перераспределении функций производственно-хозяйственной деятельности путем представления большей самостоятельности предприятиям, что, с одной стороны, повышает их ответственность за выполнение планов, а с другой — повышает ответственность отраслевых министерств за производственно-хозяйственную деятельность и развитие отрасли в целом, за рост эффективности производства, ускорение внедрения новейших научно-технических достижений.

Для того чтобы лучше представить себе главные вопросы, которые возникают при организации информационной деятельности в современных условиях, необходимо выявить основные виды информации, ее источники и пользователей (потребителей), потоки, возникающие при осуществлении управленческой деятельности. Отметим наиболее важные моменты:

1. Работы по подготовке производства завершаются выдачей соответствующей информации, необходимой для обеспечения производства определенного изделия на конкретном предприятии. Так, например, в результате проведения научных исследований разрабатываются решения по структуре, средствам и способам реализации конструкторских замыслов и технологии производства изделий, которые представляются в виде соответствующим образом оформленных и утвержденных проектных документов. Выполнение конструкторских работ — это процесс, завершающийся выдачей конструкторской, а технологических — технологической документации. Работы по определению норм материальных, трудовых и других затрат на составные части изделия и само изделие, других норм и нормативов, необходимых для планирования, учета и организации производства, также заканчиваются разработкой и ут-

верждением соответствующих документов (ведомостей, таблиц и т. п.).

Все работы по формированию документации при подготовке производства могут выполняться на одном или нескольких предприятиях или объединениях, и в первую очередь в проектно-конструкторских (технологических) и научно-исследовательских организациях. В другом случае информацию предприятия получают в готовом виде. К такой информации следует отнести патенты, лицензии, готовую конструкторскую, технологическую и другую документацию. Эта информация при организации или совершенствовании производства нового изделия приобретает у ее изготовителей, поэтому информацию о новом (усовершенствованном) изделии можно определить как информационные ресурсы, которые нужны для освоения нового производства или нового его варианта. Эти ресурсы во многом специфичны по сравнению с материальными, трудовыми и другими, необходимыми для тех же целей. Специфичность заключается в том, что они по своей природе нематериальны и поэтому физически не меняются, не расходуются, не изнашиваются и т. п. Кроме того, они играют первенствующую роль среди других видов производственных ресурсов как регламентирующие потребности в последних.

2. Процесс функционирования предприятий (объединений) и отраслей народного хозяйства требует постоянного повышения научно-технического уровня производства. Поэтому необходимо постоянно выявлять и внедрять новейшие (последние) достижения науки и техники, осваивать производство наиболее совершенных и имеющих большой спрос изделий. Все это делает более динамичным функционирование производства. Вместе с тем постоянное уточнение перспектив развития, подготовка (а при необходимости и перестройка) производства позволяют обеспечить непрерывность работы предприятий, их долгосрочное функционирование. При осуществлении такой деятельности явно выделяется несколько автономных ее видов, выполняемых в строго определенной последовательности: подготовка производства, его обеспечение ресурсами, производство, реализация продукции, развитие производства. Если эти виды деятельности представить как вершины графа, дуги которого указывают направление движения информации, то в таком графе можно выделить два контура: «подготовка производства — развитие производства» и «подготовка производства — его обеспечение ресурсами — производство — реализация продукции — развитие производства». По первому контуру движется информация, связанная с совершенствованием и развитием производства, а по второму — с обеспечением изготовления и реализации продукции. Общность подготовительных этапов в обоих контурах указывает на то, что подготовка производства является одновременно и внедрением новейших достижений и местом, где создается необходимая конструкторская, технологическая и нормативная документация. Общность же конечных этапов (развитие) показывает, что при выявлении основных тенденций и направлений развития и

внедрения достижений научно-технического прогресса нужны данные о производстве и реализации.

То, что под общим управлением должно осуществляться не только функционирование, но и развитие производства, выдвигает новые требования к информации объекта управления. Это относится, во-первых, к объемам и содержанию информации (нужно собирать и обрабатывать информацию о новейших достижениях науки, техники, технологии, организации управления и т. п.) и, во-вторых, к совершенствованию обеспечения производства нужной динамической информацией (постоянное ее обновление с внесением возникающих изменений и дополнений в ее содержание, структуру всей хранимой информации или ее частей, способы и методы ее организации, формирования и преобразования). Все это ведет к расширению информационной базы объекта управления, необходимости обеспечить ее долгосрочное функционирование при достаточно частом ее изменении (обновлении).

3. Управление деятельностью предприятий (объединений) в информационном аспекте проявляется в использовании информации

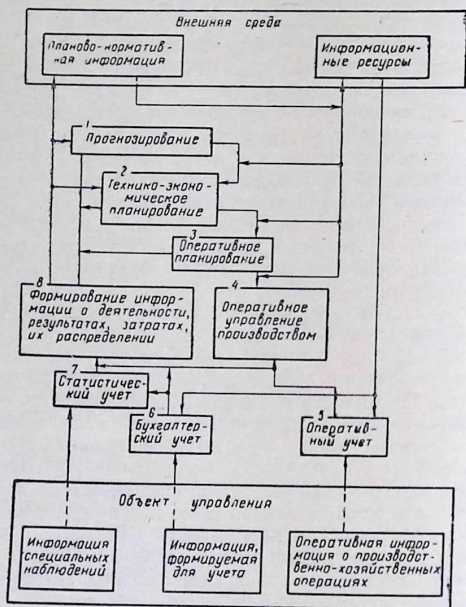


Рис. 2.1. Информационные потоки при осуществлении функций экономико-организационной деятельности

вышестоящих и других организаций (информации внешней среды), информационных ресурсов и учетной (фактической) информации для выполнения функций экономико-организационной деятельности. Кроме того, определенная информация должна быть оформлена, например, для осуществления расчетных операций с работниками предприятия и другими предприятиями и организациями, для передачи другим (в том числе и вышестоящим) организациям данных о деятельности, ее результатах, затратах. Все это ведет к образованию определенных потоков информации.

Потоки информации и их взаимосвязи с функциями экономико-организационной деятельности обобщенно можно представить так, как это показано на рис. 2.1. Все эти функции способствуют осуществлению управления. Кроме них обычно выделяются более общие функции, включающие в себя реализации некоторых (или всех) функций, указанных в блоках 1—8: контроль (блоки 5—7); принятие решений (блоки 1—4); анализ (блоки 8, 1—4, 5—7); регулирование (блоки 3—5); организация (блоки 1—8).

В результате анализа информации, связанной с выполнением этих функций, можно выявить основные потоки информации, ее виды, учитывая при этом однородность требований, предъявляемых к соответствующей информации, используемой при осуществлении отдельных функций и их групп (сочетаний). Например, это позволяет выявить различия между требованиями к информации, необходимой для функций учета, контроля и анализа, с одной стороны, и планирования, регулирования и управления как процессов принятия решений — с другой.

2.2. Техничко-экономическая информация и данные

Понятие информации широко рассматривается марксистско-ленинской философией, где она трактуется как отображение объективного мира, его причинных отношений и связей. Известен и другой подход к информации — как к характеристике неопределенности, зависящей от числа возможных исходов и вероятности каждого из них. Такая теория информации разработана К. Шенноном и Н. Винером. В ней основным понятием является неопределенность, мерой которой служит энтропия, трактующаяся как мера уменьшения неопределенности при совершении какого-то события. Однако при рассмотрении таких специфических процессов, как управление подготовкой производства, производством и экономической деятельностью, приведенные определения информации мало что дают. Здесь на информацию следует смотреть как на сведения о конкретных объектах управления, представляющие интерес при выполнении определенных управленческих функций. Эти сведения могут передаваться, храниться и обрабатываться. Аналогичный подход применяется и тогда, когда рассматриваются вопросы совершенствования организационно-экономической работы с использованием современных экономико-математических методов и

средств вычислительной техники. Поэтому понятие информации широко используется и в документации АСУ.

При проектировании машинной обработки информации важно хорошо уяснить, каковы ее виды и каким образом они подлежат передаче, хранению и обработке, на какие группы и подгруппы их можно разбивать, т. е. классифицировать. Такие классификации помогают выбирать наиболее рациональные способы размещения информации на машинных носителях и строить процессы ее обработки. Это, в свою очередь, способствует повышению качества проектных решений и эффективности управления. При этом важную роль играет правильный подбор классификационных признаков и их иерархии.

Если исходить из того, что информация — это представленные сведения, возникающие при осуществлении определенного вида (сферы) деятельности, то в результате наиболее общей ее классификации можно выделить научную, техническую, экономическую и социальную информацию. Основную долю информации, обрабатываемой с помощью средств вычислительной техники, составляют техническая и экономическая, часто рассматриваемые как общая технико-экономическая информация.

Информация может рассматриваться по отношению к конкретным функциям, с осуществлением которых она связана. Если функции относятся не к внутренней деятельности, т. е. выполняются во внешней среде, то с ними связанная информация (необходимая для них и получаемая в результате их выполнения) по отношению к объекту управления — внешняя. Если функции выполняются в рассматриваемом объекте управления, то с ними связанная информация — внутренняя. Она также может подразделяться по функциям экономико-организационной деятельности на директивную, плановую, статистическую, учетную (фактическую) и т. д.

По стадиям возникновения информация обычно делится на первичную, возникающую при выполнении функций производственной или экономико-организационной деятельности и их операций, и производную. Последняя может быть использована для дальнейших расчетов, т. е. для получения другой производной информации. Такая информация — промежуточная. Информация, которая формируется для выполнения функций организационно-экономической деятельности предприятий и других объектов управления, является результатной. Она представляет собой основной продукт, получаемый при обработке первичной и промежуточной информации.

Часто информация группируется и по другим наиболее важным качественным ее характеристикам: достоверности, полноте (достаточности) для выполнения определенных операций и функций, истинности, полезности, ценности и др. Однако разбиение информации на такие группы часто всеобъемлющего интереса не представляет, а служит лишь указанием аспекта ее рассмотрения или степени удовлетворения некоторым требованиям.

Всю техническую, экономическую и другую информацию объекта управления можно объединить в единую систему, которую при-

нято называть системой информации объекта. Основные свойства таких систем изучены М. Королевым (как экономические информационные системы) [22]. При этом каждая такая система рассматривается как состоящая из одной или нескольких подсистем, их частей (подсистем более низкого уровня) и т. д., до основного элемента таких систем — показателя.

Информация может быть представлена в самых различных формах, зависящих от ее содержания, назначения, применяемых носителей, средств обработки и других факторов. Та разновидность информации, которая связана с управлением деятельностью объекта, наиболее часто представляется в дискретной форме в виде сообщений, показателей и фиксируется в виде определенных последовательностей сигналов (импульсов) или символов.

Информацию, представленную в определенном формализованном виде, позволяющем ее передавать, хранить на различных (в том числе и машинных) носителях и обрабатывать при помощи некоторого процесса, принято называть *данными*.

При рассмотрении вопросов организации применения средств вычислительной техники для автоматизации управления производством понятия информации и данных в определенном смысле часто используются как синонимы. Такое положение в большей степени сложилось исторически: с одной стороны, основная терминология разрабатывалась и совершенствовалась одновременно с развитием работ по применению КВМ, ПВМ и ЭВМ для механизации и автоматизации планово-учетных работ предприятий, отраслей, с созданием и внедрением АСУ, где использовалось понятие информации, а с другой — применялся опыт (и литература) других стран, где в основном используется термин «данные» (например, в описании алгоритмических языков и банков данных). Поэтому и в нашей литературе, а также на практике часто используются два понятия: организация машинной (механизированной, автоматизированной) обработки экономической информации и создание и эксплуатация систем машинной обработки данных.

Понятие информации все же чаще применяется тогда, когда рассматриваются вопросы возникновения и использования определенных сведений для выполнения соответствующих функций, когда речь идет об их качественных характеристиках и эффективности использования.

Об определенных сведениях как о данных принято говорить тогда, когда дело касается их представления на носителях, в памяти машин и устройств, организации их машинной обработки. При этом имеются в виду способы размещения данных на используемых носителях и их непосредственное фиксирование (запись), т. е. данные рассматриваются на логическом и физическом уровнях.

Физически данные записываются при помощи определенных символов или в графическом виде (в виде определенных кривых или фигур). Однако при проектировании их размещения и процессов обработки наименьшей структурной единицей данных являются

ся слова, качественные и количественные реквизиты (признаки и основания показателей), представляемые в виде определенных сочетаний символов, последовательностей цифровых, буквенных или других знаков. Последовательности реквизитов (слов) образуют физические записи, объединяемые в массивы (файлы). Для их размещения необходимо разрабатывать соответствующие формы и макеты. Элементы данных и их сочетания находятся в определенных взаимосвязях, которые учитываются при разработке макетов размещения данных. Разновидности структур данных и способы их размещения на различных машинных носителях подробно рассматриваются при изучении информационных систем и структур данных. Они оказывают существенное влияние на содержание и последовательность работ по проектированию. Эти вопросы в данном аспекте изложены в третьем разделе учебника.

2.3. Этапы движения и преобразования данных

Потоки данных как сведений (информации), возникающих при осуществлении функций экономико-организационной деятельности объекта управления (см. рис. 2.1), всегда берут начало в конкретных источниках, а заканчиваются у определенных пользователей. Для различных объектов управления они могут существенно отличаться. В источниках формируются первичные данные, которые должны быть записаны (зафиксированы) на определенные носители (например, на бумагу). При организации машинной обработки данных они должны переноситься на промежуточные машинные носители, вводиться в память ЭВМ и обрабатываться с целью получения результатных данных (результатной информации), которые передаются их пользователям для выполнения ими соответствующих функций. Такие потоки данных формализованно можно представить при помощи специальных графов, имеющих вершины четырех видов: источников данных, носителей данных, работ (операций, процедур) по переносу данных на другие носители или преобразованию, пользователей данных. Они показывают движение данных через все носители и поэтому представляют собой схемы потоков данных в условиях машинной обработки.

В обобщенном виде схему потоков данных можно представить, как показано на рис. 2.2. Эта схема построена с использованием графических обозначений и по правилам, предусмотренным ГОСТ 19427—74 и ГОСТ 19428—74. В ней выделены наиболее часто встречающиеся варианты потоков данных и укрупненные группы процедур, выполняемых над ними. Схема наглядно иллюстрирует последовательность перезаписи данных с одних носителей на другие и их преобразования. Однако сами процедуры различаются по своему характеру, содержанию, объемам одновременно передаваемых или преобразуемых данных, их носителям и по другим специфическим особенностям. Таким образом, они имеют не только определенную последовательность, но и этапность. При рассмотрении вопросов организации машинной обработки данных можно

выделить четыре наиболее общих, но специфических по отношению друг к другу этапа движения и преобразования данных: первичный, подготовительный, основной и заключительный.

Первичный этап — это формирование и запись на определенные носители первичных данных, т. е. сведений, источниками (местом возникновения) которых являются подразделения и службы объекта управления или другие предприятия и организации (внешняя среда), выполняющие функции, связанные с подготовкой или осуществлением производства, экономико - организационной деятельностью. При организации машинной обработки данных целесообразно выделить два вида их источников:

1) источники, в которых специально создаются научно-технические, нормативно - справочные, плановые и другие данные, необходимые для подготовки, планирования и управления производством, в том числе и те, которые составляют информационные ресурсы (научно - исследовательские, проектные организации, конструкторские, плановые и другие отделы);

2) источники, в которых возникают оперативные фактические (учетные) данные о процессах выполнения производственных, хозяйственных, финансовых и других операций (производственные подразделения, склады, бухгалтерия, финансовый и другие отделы).

Данные (информация) в источниках первого вида, как правило, подготавливаются путем выполнения специального комплекса работ (как определенных функций) и записываются в оформленных документах, которые утверждаются соответствующими должност-

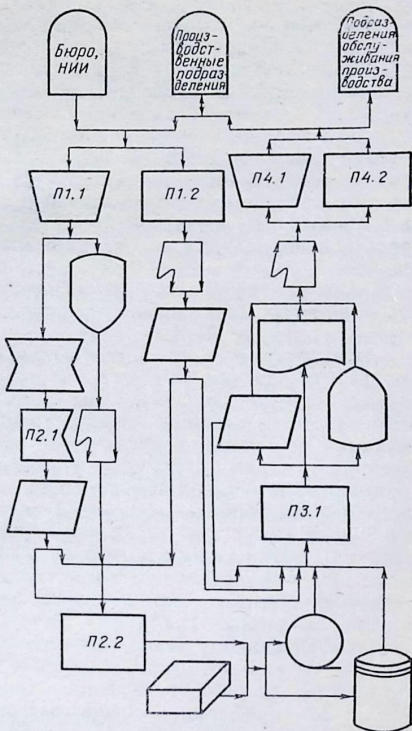


Рис. 2.2. Поток движения и преобразования данных

ными лицами (этим актом соответствующие нормы, цены и т. п. принимаются для пользования). При формировании данных могут выполняться определенные расчеты или экспериментальные работы. Таким образом, для этого также могут применяться современные экономико-математические методы и средства вычислительной техники, причем данные могут записываться сразу на машинные носители.

Данные, возникающие в источниках второго вида, также часто фиксируются на бумаге, т. е. представляются в виде первичных документов. Документы подписываются лицами, которые подтверждают адекватность записанных данных реальной действительности. Важной особенностью является то, что оперативные (учетные) данные характеризуют действительно проделанную работу, выполненные операции, их фактические параметры. От точности и своевременности их фиксирования и обработки во многом зависит эффективность управления (в первую очередь оперативно-го). Это связано с автоматизацией сбора и регистрации оперативных данных, их передачей по каналам связи, с использованием средств непосредственного ввода в ЭВМ и вывода результатов на дисплей.

Как видно из рис. 2.2, первичный этап завершается подготовкой первичных документов (процедура П1.1) или машинных носителей (процедура П1.2).

Подготовительный этап необходим тогда, когда первичные данные фиксируются на носителях, не пригодных для автоматического чтения и ввода данных для их непосредственной обработки на ЭВМ. Так как наиболее распространенными носителями данных являются первичные документы, то выполнение этого этапа часто сводится к переносу (перекодированию) данных с документов на промежуточные машинные носители, т. е. к их записи и проверке правильности выполнения этого переноса. На этом же этапе данные также могут автоматически перезаписываться с одних носителей на другие непосредственно или с передачей их по каналам связи. Это осуществляется тогда, когда необходимо формировать постоянно хранимые или каким-то образом взаимоувязанные совокупности данных (например, базу данных АСУ). На рис. 2.2 к подготовительному этапу относятся процедуры П2.1 и П2.2. При выполнении данного этапа сложной и важной проблемой является обеспечение достоверности записи данных на новый (другой) носитель. Эта проблема особенно остро встает тогда, когда перезапись осуществляет оператор при помощи перфоратора или устройства, записывающего данные непосредственно на магнитные носители. Поэтому при проектировании следует искать возможности автоматизировать подготовительный этап.

Основной этап — это процесс формирования результатных данных и записи (представления) их в таком виде, в котором было бы удобно эти данные воспринять пользователям при осуществлении ими экономико-организационной и другой деятельности. Существенная особенность данного этапа — работа с файлами (массива-

ми) данных, записанных на машинных носителях. Этот процесс обычно сводится к выполнению сравнительно небольшого числа различных процедур, которые на рис. 2.2 условно объединены в одну группу (процедура ПЗ.1). Такими процедурами могут быть: обработка одного или нескольких файлов (или специально организованных баз данных) путем выполнения операций поиска, сортировки, объединения, разъединения и арифметической обработки записей. Результаты работы основного этапа — это результаты решения функциональных задач. Они чаще всего печатаются на бумаге в виде табуляграмм, но могут также представляться пользователям в виде сообщений, показываемых на определенных технических устройствах (экране, табло и т. п.) и записываться на машинные носители.

Заключительный этап — это обеспечение полученными результатами лиц, осуществляющих экономико-организационную деятельность объекта управления, и использование ими этих результатов для выполнения соответствующих функций. Если результаты печатаются на бумаге, т. е. выдаются в виде табуляграмм, то в число процедур этого этапа входит размножение табуляграмм, передача их пользователям, которые на основании этих полученных данных непосредственно или после выполнения дополнительных операций принимают соответствующие решения или выполняют другие функции. При этом следует отметить, что проектирование машинной обработки следует начинать с выявления потребностей пользователей в результатных данных, их требований к срокам получения и качественным характеристикам данных, т. е. с определения того, что должны иметь на заключительном этапе.

Результатные данные могут быть использованы для принятия оперативных решений и последующего воздействия на происходящие процессы (на рис. 2.2 процедура П4.1). Тогда на их основании могут формироваться другие первичные данные. Кроме того, часть результатных данных используется для оценки результатов деятельности предприятия, учета, отчетности и т. п. (процедура П4.2 на рис. 2.2).

Глава 3

СУЩНОСТЬ РАБОТ ПО ПРОЕКТИРОВАНИЮ МАШИННОЙ ОБРАБОТКИ ДАННЫХ

3.1. Типовой состав проектировочных работ

Из общей характеристики проектирования машинной обработки данных как процесса (см. 1.2) видно, что оно всегда сводится к выполнению определенных работ, подготовке соответствующих решений, а их результаты оформляются в виде разрабатываемой документации. Объектами проектирования являются данные, их

потоки, способы организации и размещения данных на отдельных носителях, их совокупностях, работа людей и технических средств, направленная на сбор, передачу и обработку данных.

Проектирование может быть типовым и индивидуальным. Типовое проектирование осуществляется на основе опыта, накопленного при разработке индивидуальных проектов, как процесс его обобщения для групп предприятий, отраслей или видов работ. В каждом конкретном случае оно, как правило, связано с множеством специфических особенностей и может различаться по степени охвата функций (объектов) управления, выполняемым работам и разрабатываемой документацией. Такие проекты обычно разрабатываются высококвалифицированными специалистами. Таким образом, как целью настоящего курса является изучение основ проектирования, то в дальнейшем будем рассматривать преимущественно индивидуальное проектирование. Вместе с тем будем иметь в виду, что необходимо максимально использовать типовые и другие ранее разработанные проекты, их части, отдельные проектные решения. Содержание и объемы работ по индивидуальному проектированию зависят от вида работ (функций), которые автоматизируются, степени их автоматизации, средств вычислительной техники, экономико-математических методов, средств программного обеспечения и др., а также от использования готовых проектов и проектных решений. Проектирование может осуществляться в несколько стадий.

Наиболее простыми по содержанию и менее трудоемкими являются работы по проектированию применения клавишных вычислительных машин. В этом случае данные записываются на документах, и проектирование в основном сводится к пересмотру форм первичных документов и их перестройке таким образом, чтобы операторам, работающим на КВМ, было удобно читать исходные данные и записывать результаты, а также к разработке форм результатных документов, схем настройки машин, инструкций операторам по формированию соответствующих документов. При этом изменяются документооборот и работы по формированию документов с таким расчетом, чтобы арифметическая обработка данных осуществлялась в тех местах, где лучше всего эксплуатировать вычислительные машины. Такими подразделениями обычно служат машиносчетные бюро. Поэтому в процессе проектирования применения КВМ наиболее типичными можно считать следующие работы: обследование существующих документов и документооборота (пути их движения с места формирования до использования); выбор технических средств; проектирование форм документов и документооборота; разработку рабочих инструкций и схем настройки машин на выполнение работ.

При использовании для обработки данных перфорационных вычислительных машин появляются другие носители данных — перфокарты. В связи с этим состав работ по проектированию применения ПВМ (по сравнению с проектированием для КВМ) дополняется разработкой сокращенных условных обозначений данных,

переносимых на перфокарты (кодов технико-экономической информации), макетов размещения данных на перфокартах, технологических процессов перфорации первичных данных, обработки массивов перфокарт, формирования и выдачи результатных данных в виде табуляграмм. При этом меняется и содержание работ по созданию и разработке форм документов, меняется документооборот, разрабатываются иные инструкции и указания по настройке машин. Проектирование применения ПВМ — сравнительно трудоемкий процесс, выполняемый в несколько стадий: проводится подготовка к проектированию, обследуются предприятия и механизмируемые участки работ (объект), обрабатываются и анализируются материалы обследования, разрабатываются эскизный, а затем рабочий проекты, осуществляется внедрение разработанного проекта.

Подготовка к проектированию — это выбор или уточнение работ (их участков), подлежащих механизации, форм организации самого проектирования (бригадной, индивидуальной), составов исполнителей отдельных стадий работ, распределение обязанностей между исполнителями, определение степени участия работников предприятия в проектировании.

Обследование объекта — это ознакомление с производственно-хозяйственным процессом, изучение действующих систем выполнения соответствующих планово-учетных работ, изучение содержания и анализ форм первичных и разрабатываемых документов, сроков их поступления и объемов, определение содержания и объемов работ, выполняемых вручную и подлежащих выполнению на машинах.

Эскизный проект является, по существу, техническим заданием на проектирование и содержит все основные проектные решения, а именно: описание действующей системы выполнения работ, выбор применяемых машин, перечень мероприятий по подготовке объекта к внедрению проекта, систему классификации (кодирования) номенклатур, описание технологических процессов обработки данных, ориентировочный расчет экономического эффекта. Таким образом, в эскизном проекте представляются все решения, которые согласовываются с соответствующими подразделениями предприятия и детализируются в рабочем проекте.

Рабочий проект представляет собой набор проектных документов, в которых уточняется и детализируется эскизный проект. Основные уточненные решения эскизного проекта представляются в первой, принципиальной части рабочего проекта (основные положения), где могут быть указаны и конкретные коды, а также дается назначение каждого документа (описывается их движение и порядок использования). Во второй, рабочей части проекта приводятся конкретные решения по обработке данных, т. е. схема движения данных и детализированные технологические процессы всех работ, указания о последовательности и сроках их выполнения, технологические карты и инструкции на каждую работу и ее операции. Эта часть также должна содержать расчеты объемов работ, график прохождения их во времени, определение затрат на

их выполнение базовым (ручным) и механизированным способом расчет экономического эффекта.

Внедрение разработанного проекта обычно сводится к проведению предварительных работ и мероприятий по подготовке к внедрению, непосредственному внедрению, подведению итогов внедрения, уточнению отдельных проектных решений и проекта в целом, уточнению расчетов по объемам работ, срокам поступления и выдачи документации, уточнению экономической эффективности.

При проектировании комплексной механизации планово-учетных работ предусматривается механизация составления конструкторских и конструкторско-технологических спецификаций, выполнения нормативных, плановых и учетных работ с использованием массивов постоянно хранимых на перфокартах нормативно-справочных и других данных. В таких случаях возникают дополнительные работы по созданию и ведению на машинных носителях картотек применимости составных частей изделий, материальных, трудовых и других нормативов, ценников и т. п.

Если электронные вычислительные машины применяются для тех же работ, как и ПВМ, т. е. для решения отдельных задач и их комплексов, то при проектировании необходимо выполнить все те же самые работы. Кроме того, появляется новая область работ — разработка алгоритмов и программ для ЭВМ. Основная доля работ по проектированию применения ЭВМ, а также современных средств сбора, передачи и отображения данных связана с созданием и внедрением автоматизированных систем управления.

Рассмотрим сущность и общее содержание работ по проектированию применения ЭВМ на примере работ по созданию автоматизированных систем управления промышленными предприятиями и производственными объединениями.

Состав, содержание, порядок выполнения работ по созданию АСУП и оформлению проектной документации в основном регламентированы общепромышленными руководящими методическими материалами по созданию автоматизированных систем управления предприятиями и производственными объединениями (ОРММ) [29]. В них также приведены методики, стандарты и другие методические материалы, определяющие правила выполнения работ и оформления соответствующих документов.

Сущность создания АСУП заключается в комплексном выполнении следующих работ, позволяющих улучшить показатели производственно-хозяйственной деятельности:

- 1) совершенствование организационных структур, методов управления и осуществления экономической деятельности;
- 2) совершенствование процессов принятия решений путем внедрения экономико-математических методов и в первую очередь методов оптимизации;
- 3) создание и внедрение систем машинной обработки данных.

Выделенные группы работ представляют собой целые направления деятельности по совершенствованию управления, но они находятся в тесной взаимосвязи, так как от содержания одних за-

висит содержание других. Поэтому в ОРММ они рассматриваются вместе, на основании этого определены содержание, стадии и последовательность всех работ, а также порядок их выполнения. Отметим положения, наиболее важные при проектировании систем машинной обработки данных:

1. Работы по созданию АСУП разбиваются на три группы, выполняемые последовательно и представляющие отдельные стадии: предпроектную, разработки проектов (проектирования) и ввода АСУП в эксплуатацию. Разработка проектов, в свою очередь, обычно осуществляется в две стадии, и разрабатываются технический и рабочий проекты. Однако это может быть осуществлено в одну стадию. Тогда разрабатывается общий технорбочий проект. Последний вариант рекомендуется применять тогда, когда на практике в других системах уже были разработаны прототипы основных решений.

2. АСУП, как правило, разрабатывают специализированные научно-исследовательские и проектные организации на основе договоров, заключенных с организациями, для которых эти системы создаются. В документации АСУП принято первых называть разработчиками, вторых — заказчиками. Ввиду специфических особенностей каждого объекта, для которого создается система, большинство работ по обследованию объекта, подготовке его к внедрению разработчик поручает выполнять самому заказчику. При этом в ОРММ работы между заказчиком и разработчиком достаточно четко распределены, а также определены обязанности и права каждого из них.

Заказчик при создании АСУП обязан: заключить хозяйственные договоры на создание системы, приобретение технических средств; подготовить задание на строительство или реконструкцию помещений; совместно с разработчиком выполнить работы предпроектной стадии; в необходимые сроки подготовить помещения, приобрести и установить технические средства; разработать и осуществить мероприятия по совершенствованию организации управления и производства; выполнить в намечаемые сроки мероприятия по подготовке объекта к внедрению АСУП.

В плане работ по подготовке объекта к внедрению много места занимают такие работы, которые непосредственно связаны с организацией системы обработки информации: разработка и внедрение локальных классификаторов технико-экономической информации, внедрение общесоюзных классификаторов для кодирования входной и выходной информации, унифицированных форм документов, установленных унифицированной системой документации (УСД), Единой системой конструкторской документации (ЕСКД) и Единой системой технологической документации (ЕСТД).

Основная цель разработчика — это создание автоматизированной системы обработки данных. На предпроектной стадии он участвует в обследовании объекта, обработке материалов обследования, определении расчетов (задач, их комплексов или функций), которые необходимо выполнить машинным способом, в разработке

предложений по изменению информационной базы, определении экономической эффективности.

На основании утвержденного технического задания разрабатываются технический и рабочий проекты. В техническом проекте наиболее существенными являются документы, в которых детализируются решения по комплексу технических средств, описываются выбранные средства программного обеспечения, решения по организации информационного обеспечения, системы классификации и кодирования. В него также включаются решения по организации информационной базы, программы и рабочая документация по формированию и ведению массивов, содержащих нормативно-справочные данные.

Рабочий проект представляет собой рабочую документацию, необходимую для нормального выполнения автоматизируемых работ. Его основную часть составляют программная документация и технологические инструкции по обработке данных. В инструкциях содержатся сведения о выполнении операций сбора, регистрации, передачи и обработки данных.

Все готовые программы и соответствующие инструкции, введенные на контрольных примерах, разработчик сдает заказчику, который после проверки их на тех же примерах составляет акты приема системы в опытную эксплуатацию, проводит ее испытание в реальных условиях. Для этого он подготавливает свой контрольный пример. Принятые программы должны пройти опытную эксплуатацию. После этого программы принимаются в промышленную эксплуатацию. Аналогично принимаются все инструкции и другая документация рабочего проекта.

Конкретный состав проектировочных работ, осуществляемых при подготовке к внедрению различных мероприятий, может существенно меняться. Различными могут быть объемы работ и даже организация их проведения. Однако в методическом аспекте они выполняются аналогично. Поэтому содержанию и методикам проектирования обычно и уделяется главное внимание. Более детально эти вопросы рассматриваются в третьем и четвертом разделах учебника. Опыт показывает, что незнание или недооценка этих специальных вопросов снижает качество разрабатываемых проектов и приводит к последующей их доработке или даже переработке. Это часто случается и из-за недостаточного качества разрабатываемой проектной документации. Однако существуют и более глубокие причины, важное место среди которых занимает качество выработки проектных решений. В каждом отдельном случае практически можно строить несколько (множество) вариантов возможных решений, а потом выбрать наилучшее (оптимальное) по заданному критерию.

Разновидность решаемых при проектировании вопросов не позволяет дать одного способа и подхода к их решению. Даже трудно построить общую модель и выбрать единый критерий оценки различных решений. Так, при выборе автоматизируемых функций (задач или их комплексов) можно исходить из того, насколько это меро-

приятие позволит предприятию увеличить прибыль или улучшить другие показатели своей деятельности. Выбор рационального комплекса технических средств, систем условных обозначений (систем кодирования) технико-экономической информации, макетов и способов размещения данных, вариантов технологических процессов позволяет минимизировать затраты на обработку. Средства программного обеспечения можно выбирать так, чтобы минимизировать затраты на проектирование. При этом различные критерии оценки могут быть противоречивыми. Например, минимизация затрат на проектирование может вызвать их увеличение при эксплуатации. Кроме того, оценки вариантов выбираемых проектных решений могут осуществляться и не в денежном выражении. Поэтому перед тем, как изучить содержание и методики выполнения конкретных проектировочных работ, необходимо более четко уяснить и представить общие принципы оценки эффективности принимаемых решений, оценки информации, а также выявления и определения затрат. Это дает возможность обоснованно выбирать технические и программные средства и принимать решения, способствующие разработке качественных проектов.

3.2. Особенности проектирования машинной обработки данных при создании АСУ

Проектирование машинной обработки данных при создании АСУ представляет собой автономный вид деятельности по выполнению работ, требующих от исполнителей специальных знаний. Поэтому проектирование потоков данных, их размещение на различных носителях и в памяти системы, процессов их движения и преобразования в условиях АСУ вполне естественно выделить в виде автономной группы работ, тем более, что часто при создании АСУ основное внимание уделяется лишь перечисленным вопросам, а вопросы совершенствования организации управления самим объектом решаются отдельно. Но, с другой стороны, проектирование в этих условиях имеет некоторую специфику, так как оно проводится одновременно с другими работами — по совершенствованию организационно-экономической и другой деятельности. Эта специфичность проявляется в следующем:

1. Проектирование машинной обработки данных осуществляется в условиях некоторой неопределенности в том смысле, что оно выполняется одновременно с изменением (перестройкой) потоков информации на объекте управления (предприятии), уточнением структуры и функций управления, изменением потребностей в информации. В таких случаях решения по выбору автоматизируемых функций (задач) должны приниматься на основе проведенного перед этим глубокого и всестороннего изучения не только существующего положения, но и направлений совершенствования этих работ, обобщений последних достижений. В связи с этим повышается роль этапа обследования объекта. Подобное совмещение работ по совершенствованию управления и автоматизации обра-

ботки данных имеет свои достоинства, которые заключаются в том, что в таких условиях можно создать более стройную систему обеспечения информацией ее пользователей, осуществить рационализацию потоков информации на всех этапах ее движения и преобразования. Вместе с тем при совместном выполнении этих работ появляются дополнительные трудности, вызванные некоторой неопределенностью в самом процессе проектирования и итеративностью выбора окончательных проектных решений. Суть заключается в том, что по предлагаемому варианту совершенствования структуры и организации деятельности объекта управления строится система машинной обработки данных; она анализируется и в результате такого анализа вводятся необходимые изменения, которые способствуют улучшению организации движения и использования информации.

2. В условиях АСУ машинная обработка данных всегда имеет системный характер: проектируются постановка и организация решения комплексов информационно взаимосвязанных задач. Это дает возможность проектировать единые информационные фонды, осуществлять унификацию и типизацию схем размещения массивов данных на машинных носителях, процессов формирования массивов и их обработки при решении функциональных задач. Унификации также поддаются имена (идентификаторы) реквизитов, обозначения показателей, системы классификации и кодирования технико-экономической информации, формы документов и т. п.

3. Круг автоматизируемых функций, комплексов и отдельных задач автоматизированных систем управления предприятиями (объединениями), отраслями, а частично и других (например, межотраслевых) АСУ все более четко выявляется и стабилизируется. Это создает условия для применения типовых проектных решений, пакетов прикладных программ и других средств специального программного обеспечения.

4. Машинная обработка данных в условиях АСУ в основном сводится к обработке больших массивов, используемых для решения задач, которые по сложности алгоритмов могут существенно различаться. В решении большинства задач преобладают такие сравнительно простые операции, как поиск, выборка, сортировка, выполнение арифметических действий и выдача результатов в виде многострочных табуляграмм с накапливаемыми итогами нескольких степеней. На основе той же информации решаются и такие задачи, которые требуют применения методов математического программирования и других, реализовать которые можно на заранее известных алгоритмах. Поэтому при проектировании АСУ можно использовать уже имеющиеся типовые программы и программные средства их генерации. В первую очередь это относится к задачам, для решения которых используются данные, формируемые в едином информационном фонде, т. е. к информационно тесно взаимосвязанным задачам, в особенности к задачам учета и составления отчетностей.

5. В системе, как правило, встречаются задачи различной пе-

риодичности решения, в том числе и такие, которые необходимо решать в реальном масштабе времени или по запросам, а также редким заказам. Поэтому задачи можно разбивать на такие группы (или объединять в такие комплексы), к которым они относятся не только по принципу информационных связей, но и по различным другим признакам: периодичности и срокам решения, срокам и способу получения исходных данных, передачи и представления результатов и т. п. При организации решения комплексов таких задач необходимо наиболее полно учитывать многоаспектную группировку задач, сохраняя при этом достаточную системность.

6. В АСУ широко используются единые базы данных и системы управления ими, которые необходимо генерировать с учетом структурных особенностей данных, принимаемых методов доступа и поиска, требований многоцелевого использования хранимой информации. В связи с этим при проектировании системы необходимо провести тщательный анализ существующих способов и средств организации баз данных с тем, чтобы выбрать наиболее рациональный вариант размещения данных в памяти системы, выполнения процедур их формирования, ведения и использования.

Глава 4

ОБЩИЕ ПРИНЦИПЫ ОПРЕДЕЛЕНИЯ ЭФФЕКТИВНОСТИ

4.1. Основы оценки мероприятий по применению вычислительной техники

Средства вычислительной техники принадлежат к числу основных средств современного этапа научно-технического прогресса, позволяющих механизировать и автоматизировать управленческий труд. С их помощью можно выполнять практически все трудоемкие расчеты по обработке данных, лучше обеспечивать информацией лиц, принимающих решения, и т. д. Это создает условия управленческому и другому персоналу шире проявлять свои умственные способности, повышать качество работы.

Внедрение вычислительной техники можно представить как реализацию мероприятия, дающего некоторый эффект. Решение о целесообразности внедрения такого мероприятия должно быть принято после выявления того, какой оно даст полезный результат производственной или управленческой деятельности, т. е. какой даст эффект.

Наиболее общей мерой автоматизации управления может быть та, которая используется для оценки уровня автоматизации производства (оно может быть ручным, механизированным, автоматизированным или автоматическим). Однако такие общие качественные оценки дают слишком обобщенное представление о предполагаемых результатах внедрения проектируемого мероприятия.

Более конкретные оценки уровня автоматизации управленческого труда могут быть получены на основании показателей его охвата автоматизацией, которые можно определить для управленческого персонала (штатной численности), функциональных задач, объемов управленческих работ и т. п.

Уровень автоматизации труда управленческого персонала (K_u) определяется отношением количества управленческих работников, автоматизированно выполняющих функции обработки данных, к общей численности управленческих работников. При этом числитель определяется путем подсчета количества работников, необходимых для выполнения тех работ, которые автоматизируются (при

внедрении мероприятия эти работники относятся к условно высвобождаемым — $Ч_a$). Общее количество работников управления включает в себя работников аппарата управления ($Ч_{уп}$), персонал, обслуживающий средства автоматизации ($Ч_c$), и тех работников, которые условно высвобождаются ($Ч_a$), т. е.

$$K_{\text{ч}} = \frac{Ч_a}{Ч_{\text{уп}} + Ч_c + Ч_a}.$$

Уровень автоматизации управленческого труда как показатель охвата задач ($K_{\text{за}}$) определяется отношением количества (N_a) функциональных задач (комплексов задач или управленческих функций), решаемых с помощью средств вычислительной техники, к общему числу ($N_{\text{ст}}$) задач, решение которых можно или целесообразно автоматизировать для данного предприятия, т. е.

$$K_{\text{за}} = \frac{N_a}{N_{\text{ст}}}.$$

Такая оценка используется при определении научно-технического уровня АСУП. Однако функциональные задачи часто существенно различаются по трудоемкости решения, охвату автоматизируемых функций и другим параметрам. Поэтому данный показатель уровня автоматизации имеет ограниченное применение и может использоваться лишь для оценки степени автоматизации аналогичных функций на различных объектах управления.

Не только количественные, но и качественные оценки можно получить, измеряя уровень автоматизации по отношению к объемам управленческих работ, выражаемым, например, в трудовых затратах. Если эти затраты определяются для однородных работ или операций (процедур) технологического процесса обработки данных, групп функциональных задач, данный показатель рассчитывается по формуле

$$K_{\text{т}} = \frac{T_a}{T_a + T_p},$$

где T_a — трудоемкость автоматизируемых работ;

T_p — трудоемкость неавтоматизируемых работ.

Если же затраты определяются для групп различных функциональных задач, расчеты выполняются по формуле

$$K'_{\text{т}} = \frac{\sum_{i=1}^m (T_i I_i - T_{oi})}{\sum_{i=1}^m (T_i I_i + T_{pi})},$$

где T_i — трудоемкость решения i -й задачи (комплекса задач) вручную (неавтоматизированно);

I_i — индекс изменения трудоемкости решения i -й задачи в связи с автоматизацией;

T_{oi} — трудоемкость части i -й задачи, решаемой вручную в условиях автоматизации;

T_{pi} — трудоемкость подготовки исходных данных и обслуживания технических средств при автоматизированном решении i -й задачи;

m — число задач, решаемых в исследуемой системе.

Уровень использования отдельных технических средств и производительности их комплексов для выполнения вариантов технологических процессов обработки данных можно определить, используя коэффициент автоматизации по объему данных — отношение объема данных, обрабатываемых с помощью технических средств, к их общему объему в рассматриваемой части системы. Если в рассматриваемых вариантах изменяются объемы информации, то это должно учитываться с помощью соответствующих индексов. В этом случае численное значение коэффициента автоматизации определяется по формуле

$$K_a = \frac{\sum_{i=1}^m A_{oi}}{\sum_{i=1}^m A_{pi} \cdot I_i},$$

где A_{oi} — объем данных по i -й задаче, обрабатываемых с помощью технических средств;

A_{pi} — общий объем данных, обрабатываемых при решении i -й задачи вручную;

I_i — индекс изменения объема данных по i -й задаче при автоматизированном решении.

Приведенные и другие коэффициенты дают лишь общие и довольно грубые оценки уровня автоматизации управленческих работ, так как при их определении используются укрупненные данные. Например, принимается, что все выполняемые работы, решаемые задачи и вся обрабатываемая информация одинаковы по сложности, ценности и другим показателям.

Окончательное решение о целесообразности внедрения конкретного мероприятия должно быть принято на основе более точных показателей, определяющих его эффект. Прежде всего, для общественного производства важен не столько достигнутый уровень автоматизации, сколько получаемый эффект, определяемый как любая общественная потребительная стоимость, созданная трудом, изделие или услуга, удовлетворяющие своими свойствами потребности человека.

Эффект может быть социальным, техническим и экономическим. Первый в наиболее общем виде характеризуется степенью удовлетворения потребностей, влиянием на жизненный уровень населения. Он имеет и более конкретные выражения, например в виде показателей роста продукции высшего качества, степени удовлетворения потребностей народного хозяйства в определенной продукции отрасли, показателей объемов производства некоторой важнейшей техники.

Технические показатели эффекта уже давно используются при

оценке отдельных элементов и узлов технических устройств и машин. Они находят все более широкое распространение и при оценке систем машинной обработки данных и автоматизированных систем управления. С их помощью определяется степень прогрессивности технических параметров средств вычислительной техники, уровень автоматизации программирования, использования стандартных и типовых программных средств (пакетов прикладных

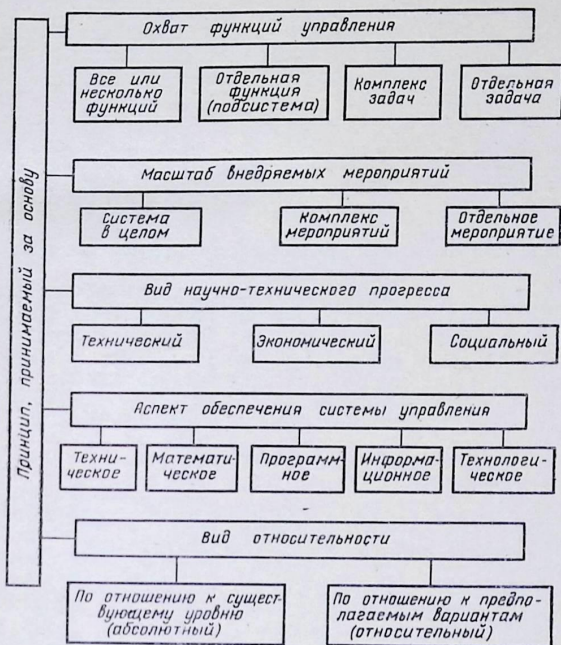


Рис. 4.1. Схема классификации показателей эффекта, получаемого от внедрения средств вычислительной техники

программ и др.), уровень организации информационного фонда и т. д. Такие показатели могут использоваться, например, для оценки качества конструкторской и другой документации, что позволяет более точно определить эффект, получаемый от качественного проведения работ на этапе подготовки объекта к внедрению СВТ. Следует отметить, что технические показатели до сих пор используются недостаточно. Они могут и должны играть важную роль при выборе вариантов проектных решений по информа-

ционному и программному обеспечению, а также технологическим процессам обработки данных.

Экономический эффект в большей степени, чем другие виды эффекта, соответствует тем конкретным целям, которые ставятся перед социалистическим государственным предприятием. Экономический эффект, создаваемый благодаря внедрению СВТ и новых методов планирования и управления, характеризуется улучшением технико-экономических показателей деятельности предприятия. Он показывает величину достигнутого чистого результата, т. е. превышение результата, получаемого по одному варианту, по сравнению с результатом другого варианта и может определяться для всего нового мероприятия или отдельных его частей (при выборе различных вариантов).

Для оценки мероприятия по внедрению вычислительной техники приходится пользоваться несколькими показателями его эффективности. Для более четкого выбора показателей может служить схема классификации, представленная на рис. 4.1. На ее основе в каждом конкретном случае можно выбрать показатели в таком сочетании, чтобы они наиболее полно характеризовали эффект внедрения разработки или функционирования объекта в целом, а также его отдельных частей.

Внедряемые мероприятия могут охватывать различные по масштабу и функциям управления части объекта, на котором они внедряются. Следовательно, возможны и различные варианты этих мероприятий, вложения в них средств. В таких случаях возникает необходимость анализировать и сравнивать возможные варианты и выбирать наиболее эффективные.

4.2. Показатели экономической эффективности и принципы их определения

Эффект показывает только одну — полезную сторону внедряемых мероприятий. Проведение каждого мероприятия связано с определенными затратами, которые для различных мероприятий или их вариантов могут существенно различаться. Поэтому необходима сопоставимость, соизмерение получаемого полезного результата (эффекта) с произведенными затратами. Такой оценочной мерой уровня внедряемых мероприятий служит эффективность. Она, будучи производной категорией от категории эффекта, является более сложной и комплексной. Эффективность в общем случае отражает степень соответствия разрабатываемого проекта своему назначению, его техническое, технологическое и организационное совершенство и экономическую целесообразность. Эффективность производства, по словам К. Маркса, означает не что иное, как возможность и необходимость «...достигать определенного результата при наименьших затратах»¹. Экономическая эффективность определяется отношением получаемого годового экономиче-

¹ Маркс К., Энгельс Ф. Соч. 2-е изд., т. 25, ч. I, с. 96.

ского эффекта от внедрения (осуществления) данного мероприятия к затратам, определившим возможность получения данного эффекта. Тем самым она отражает уровень производительности общественного труда и является конечным критерием и мерилом научно-технического прогресса.

Для АСУП коэффициент экономической эффективности (эффективности капитальных вложений) рассчитывается по формуле

$$E_p = \frac{П^A}{K_K^A},$$

где $П^A$ — годовой прирост прибыли (годовая экономия), получаемой от внедрения системы;

K_K^A — капитальные вложения на создание системы.

Этот коэффициент, а также его обратное значение, т. е. срок окупаемости капитальных вложений (T), и величина годового прироста прибыли ($П^A$) — основные показатели экономической эффективности создания и функционирования АСУП. Срок окупаемости капитальных вложений рассчитывается по формуле

$$T = \frac{K_K^A}{П^A} = \frac{1}{E_p}.$$

Годовая экономия $Э_{год}$ (прирост прибыли — $П^A$) определяется по формуле

$$Э_{год} = П^A = \left(\frac{A_2 - A_1}{A_1} \cdot П_1 \right) + \left(\frac{C_1 - C_2}{100} \cdot A_2 \right),$$

где A_1, A_2 — годовой объем реализуемой продукции до и после внедрения системы;

$П_1$ — прибыль от реализации продукции до внедрения системы;

C_1, C_2 — затраты на 1 рубль реализуемой продукции до и после внедрения системы, определяемые отношением себестоимости к соответствующему годовому объему реализуемой продукции.

В данной формуле каждое слагаемое, указанное в скобках, имеет самостоятельное значение: первое показывает годовой прирост прибыли за счет роста объема реализуемой продукции, а второе — рост прибыли за счет снижения издержек производства.

При определении затрат на 1 рубль реализуемой продукции после внедрения системы (C_2) в годовую себестоимость реализуемой продукции включаются затраты, связанные с эксплуатацией системы.

Общая народнохозяйственная эффективность может определяться как для всей системы, так и при внедрении любого мероприятия по применению средств вычислительной техники. При создании системы обработки данных необходимо провести детальные расчеты по определению экономической целесообразности затрат с точки зрения предприятия, на котором эта система внедряет-

ся, а также с позиций вычислительной установки, осуществляющей обработку данных.

Эффективность внедряемого мероприятия получается от улучшения организационной или производственно-хозяйственной деятельности и от сокращения затрат на обработку данных, т. е. от совершенствования управления и совершенствования обработки данных. Первая часть эффективности — улучшение технико-экономических показателей работы — может быть определена аналогично общей народнохозяйственной эффективности. Что касается второй части эффективности, то она рассчитывается путем сравнения затрат по существующему и предлагаемому вариантам. Для систем машинной обработки данных первая часть иногда называется косвенной, а вторая — прямой эффективностью. При рассмотрении вопроса целесообразности внедрения мероприятия необходимо оценивать общую эффективность. Переоценка роли одной части эффективности и недооценка другой ведет к искажению действительной выгоды от внедрения вычислительной техники. Так, например, об эффективности мероприятия предприятие в основном судит лишь по тому, насколько оно уменьшает трудоемкость обработки данных и высвобождает людей. С другой стороны, показатель экономии затрат на обработку данных служит для оценки различных предлагаемых вариантов.

Выбор варианта осуществляется на базе сравнения определенных показателей эффективности старого (базового) и нового (разрабатываемого) вариантов или двух (нескольких) предлагаемых вариантов. Лучшим считается вариант, обеспечивающий оптимальные значения показателей с учетом выбранных критериев.

Для определения эффективности применения СВТ как мероприятия по внедрению новой техники можно рассчитывать годовой экономический эффект:

$$\mathcal{E} = \mathcal{E}_{\text{год}} - E_{\text{н}} \cdot K_{\text{д}}^{\Lambda},$$

где $E_{\text{н}}$ — нормативный коэффициент экономической эффективности капитальных вложений в данной отрасли;

$K_{\text{д}}^{\Lambda}$ — капитальные вложения, необходимые для создания и внедрения мероприятия (системы).

Если рассматриваются два или несколько вариантов реализации мероприятия, то годовой экономический эффект может быть определен как разность сумм приведенных затрат по сравниваемым вариантам. Они должны рассматриваться в строго сопоставимых условиях сравнения по всем признакам. Но эти затраты представляют собой сумму двух видов затрат, приведенных к одинаковой размерности в соответствии с нормативом экономической эффективности: текущих (себестоимости) и всех единовременных, которые необходимы для осуществления мероприятия. Поэтому они должны быть полностью сопоставимы. Их структура достаточно полно определена лишь для автоматизированных систем управления в целом. Менее разработаны вопросы определения затрат и эффективности для отдельных составляющих АСУ: систем обра-

ботки данных, подсистем, комплексов задач, отдельных задач, технологических процессов, алгоритмов. Это связано с трудностями, возникающими как при выявлении затрат, определяемых прямыми расчетами, так и при распределении общих затрат с выбором базы для такого распределения. При проведении конкретных расчетов по определению приведенных затрат (текущих и единовременных) необходимо руководствоваться следующими принципами.

К единовременным затратам должны быть отнесены не только капитальные вложения, но и затраты, связанные с проведением мероприятий и работ по созданию необходимого программного и информационного обеспечения, составляющего основные средства вычислительного предприятия, по проведению проектных и других подготовительных работ. Если единовременные затраты по сравниваемым вариантам относятся не к одному и тому же периоду времени, то при их сопоставлении следует учитывать это время как лаг. В случаях, когда единовременные затраты связаны с осуществлением нескольких мероприятий и невозможно выделить доли этих затрат на конкретно рассматриваемое мероприятие, их необходимо определить косвенно, согласно выбранной базе. Как правило, при организации машинной обработки информации основную долю капитальных вложений составляют затраты на средства обработки данных, и поэтому за базу можно принять машинное время на обработку или решение задач для рассматриваемых вариантов.

Текущие затраты наиболее полно и точно определяются путем оценки времени, необходимого для выполнения операций технологического процесса, объединенных в группы с учетом стоимостных характеристик различных СВТ; при оценке должны учитываться только те работы и операции, которые нужны для непосредственной реализации известного технологического процесса (на основании разработанных технологических процессов формирования файлов первичной информации, процессов и алгоритмов их обработки при решении задач), исключая те операции, затраты на выполнение которых относятся к единовременным; затраты по отдельным операциям должны учитываться с использованием одинакового уровня оценок.

Оценку затрат по операциям технологического процесса можно производить двумя способами: используя установленные единые цены на машинно-час работы вычислительной техники и определяя себестоимость выполнения операций. Первый способ лучше подходит для предприятий, когда они решают вопросы минимизации оплаты за услуги вычислительных установок. Показатели себестоимости используются при выборе лучшего варианта технологического процесса обработки данных на вычислительных установках.

Эффективность непосредственно связана с качеством, которое также может быть выражено при помощи определенных показателей. Среди них можно выделить временные, стоимостные, технические и потребительские. Конкретные группы показателей качест-

Варианты определения понятия ценности информации

Вариант	Что характеризует ценность информации (формулы определения ее ценности)	Какими показателями выражается ценность информации
1	Ценно то, что может быть использовано, что полученное сообщение новее	Неопределенность, непредвиденность, оригинальность
2	Общее информационное значение единицы информации по отношению к конкретным функциональным задачам: $V_i = (Z_i \alpha + P_i \beta + U_i \gamma + A_i \omega + R_i \delta) \cdot S_i \cdot D_i$ где $\alpha, \beta, \gamma, \omega, \delta$ — весовые коэффициенты; i — номер задачи.	Значимость (Z_i), полезность (P_i), употребляемость (U_i), периодичность или частота использования (A_i), ранг (R_i), своевременность (S_i), достоверность (D_i)
3	Безразмерная величина, характеризующая степень влияния данного сообщения на взаимодействие системы управления и управляемого объекта: $V_i = f(T_i, D_i, Z_i, P_i, U_i, R_i), -1 \leq V_i \leq 1$	Заданное время получения информации (T_i), D_i, Z_i, P_i, U_i, R_i
4	Прибыль, получаемая за определенный период от применения информации, польза от нее, эффективность, которой обладает информация в управлении: $V = \left\{ \frac{1}{F} \sum_{i=1}^m f(P, T_a, Y, n) - g(P_p, Y) \right\} - h(P, F, T_a)$ где T_a — время задержки, необходимое для получения информации; F — интервал времени, в котором используется информация; P_p — вероятность принятия правильного решения без использования информации; Y — число решений, осуществляемых в каждый момент времени; n — прибыль, получаемая от принятия правильного решения; $f(P, T_a, Y, n)$ — прибыль, получаемая от использования информации; $g(P_p, Y)$ — прибыль, получаемая при решении задачи без использования информации; $h(P, F, T_a)$ — средняя стоимость информации.	Долговечность истинности информации в текущих условиях, долговечность истинности оперативно-производственной информации, возможность прогнозирования информации (Π), действительность

щими элементами или путем ее обработки с помощью СВТ. Для предприятия как объекта управления основное значение имеет определение потребностей первого вида. Это связано с тем, что от того, как осуществляется или будет осуществляться реализация основных целей, насколько полно раскрыты и рационально используются все ресурсы, насколько учтены все особенности и условия, зависит эффективность всей деятельности.

Потребности второго вида вытекают из необходимости полу-

чить определенную результатную информацию. Когда информацию обрабатывают управляющие элементы, потребности этих двух видов трудноотделимы. Самостоятельное значение они приобретают тогда, когда необходимо выполнить большие объемы работ по обработке информации, т. е. когда их целесообразно автоматизировать (при создании АСОД). Поэтому при проектировании систем прежде всего следует выявить те особенности, которые возникают при удовлетворении информационной потребности первого вида (потребности в информации, необходимой для непосредственной организации деятельности объекта управления), определить требования к качеству и срокам получения этой информации. После этого нужно перейти к выявлению потребностей в исходной информации.

При определении потребностей в информации как в ресурсах предприятия, а также ее ценности прежде всего возникает вопрос: кому конкретно она нужна и как она может удовлетворять возникающие потребности? При поиске ответа на данный вопрос следует обратить внимание на то, что требования к информации и ее качеству должны рассматриваться с позиций того пользователя, которому она непосредственно должна помочь выполнять соответствующие функции, информационно взаимосвязанные друг с другом. Это, в свою очередь, ведет, с одной стороны, к измерениям полезности информации, используемой различными лицами (управляющими элементами), и с другой — к определению затрат на ее получение. Для количественного измерения ценности (V_y) можно использовать методологию, заложенную в основу определения экономической эффективности АСУ, т. е. прежде всего выявить, насколько используемая информация помогает улучшить показатели работы предприятия, увеличить его прибыль, а затем определить затраты на ее получение и их сопоставить. В связи с этим ценность информации для системы (объекта управления) целесообразно рассматривать не только в отношении субъекта управления (пользователей информации), цели, но и стоимости получения, т. е. в сочетании ее полезности и стоимости.

Ценность информации, используемой для выполнения функций организационно-экономической деятельности, можно определить по следующей формуле:

$$V_y = \begin{cases} C; C \gg \max_{i \in I} C_i, & \text{если информация обязательна;} \\ \max_{i \in I} (P_i - C_i), & \text{если возможны различные варианты} \\ & \text{степени ее полезности,} \end{cases}$$

где P_i — польза (эффект или дополнительный эффект), получаемая от использования информации по сравнению с отсутствием этой информации;

C_i — стоимость информации (по i -му варианту);

I — множество вариантов с различными P_i и C_i ($i \in I$);

C — достаточно большое положительное число, всегда намного больше максимальной стоимости информации.

Эта формула представляет собой обобщение различных вари-

антов оценки информации и определяет некоторый порядок выбора лучшего варианта информации. В первой ее части учитывается специфика той доли информации, которая выступает как обязательная. К такой информации относятся, например, фактические (учетные) данные. При этом она должна поступать и формироваться в конкретные, заранее определенные сроки (периоды), с определенной степенью точности и достоверности. Потребности в такой информации вызваны необходимостью знать, какие другие (кроме информационных) ресурсы были использованы, что и в каком количестве произведено, кому и какие суммы необходимо перечислить и т. п. Формирование и использование этой информации регламентируются соответствующими методиками, нормативными актами, поэтому она всегда является «очень ценной» (для нее $V_y = C$), а затраты на ее получение составляют необходимые издержки производства, поэтому они должны включаться в себестоимость продукции предприятия.

Ценность практически должна определяться только для той части информации, которая может иметь различную полезность (и от этого меняется получаемый эффект), т. е. такой информации, которая может оказывать различное влияние на производство и его результаты (прибыль), определяет вариантность поведения. При этом учитывается тот факт, что на предприятии каждому управляющему элементу, по существу, важна не абсолютная (внутренняя) ценность информации как таковой или ценность для него как индивида, а ценность для всего предприятия (ценность информации по отношению к объекту управления). Например, изобретение может быть ценным, позволяющим сделать существенный сдвиг по улучшению свойств выпускаемых изделий, но отсутствие в рассматриваемый момент возможностей его практической реализации превращает его в бесценную для данного предприятия информацию. С другой стороны, то, что ценно и полезно для отдельного лица, что уменьшает неопределенность или позволяет узнать новое (то, чего оно не знало), может оказаться таким лишь потому, что это лицо до того знало меньше, чем должно было знать.

Так как ценность управленческой информации можно выразить разностью между пользой, полученной от ее применения, и стоимостью получения этой информации, при определении содержания результатной информации и способов ее получения выбор должен пасть на те варианты, сочетание которых при других равных условиях позволит получить наибольшую разность между этими оценками.

Для выбора нужной информации необходимо ее анализировать в различных аспектах, позволяющих наиболее полно выявить ее полезность. Практически таких аспектов немного, поэтому можно выделить наиболее характерные случаи, при которых следует рассматривать лишь общие потребности в информации и когда необходимо определить ее ценность. Для этого можно использовать общую схему анализа, представленную в табл. 5.2.

Многоаспектный анализ информации и ее ценности

Аспект оценки	Ситуации, в которых ценность определяется	Ситуации, в которых ценность не определяется
Время получения	Информация обязательна в момент ее использования. При не-обеспечении этого: а) с протеканием времени теряется ее ценность; б) от срочности получения информации зависит качество результатов	Информация обязательна в момент ее использования. При ее отсутствии: а) функция не может быть выполнена; б) фиксируется, что никакие операции не произведены
Достоверность	От степени достоверности зависит качество (или количество) результатов	Предъявляются конкретные одновариантные требования
Вариантность для принятия решений	Возможны варианты, дающие результаты различного количества или качества (различную эффективность или прибыль)	Информация характеризует процесс (операции) однозначно — что сделано, получено, сколько необходимо сделать и т. п.
Полнота	Можно получить неполную информацию, но от этого зависит и окончательный результат (различная степень полноты — различный результат)	Нужна вся информация

Определение полезности информации и ее ценности по отношению к объекту управления по существу касается только результатной информации (выхода системы информации объекта). Имея эти оценки и другие характеристики результатной информации, можно выдвигать соответствующие требования к исходной информации, необходимой для ее формирования. Здесь проявляется второй вид потребностей (потребностей в исходной информации). Исходная (первичная и промежуточная) информация, используемая для формирования выхода, т. е. результатной информации, с последней взаимоувязана через математические модели задач и процедуры их реализации. Поэтому ценность результатной информации переносится в некотором аспекте и на исходную.

Однако ценность исходной информации рассматривается и в другом аспекте — по отношению к задаче (функции), для решения (или выполнения) которой она используется. Для такой оценки можно использовать показатели, представленные в табл. 5.1 (варианты 2 и 3). Кроме того, при определении ценности исходной информации необходимо учитывать степень ее преобразования для получения результатов и используемые для этого процедуры (процедуры реализации математических моделей задач), а также информационные связи между задачами. Процедуры могут быть представлены в виде технологических процессов формирования массивов и алгоритмов их обработки.

Вопросы оценки информации непосредственно связаны с организацией системы информации, ее размещением на различных носителях, проектированием (организацией) системы машинной обработки. Более детально вопросы построения таких систем и проектирования технологии обработки информации изложены в последующих главах. Предварительно необходимо выделить некоторые наиболее важные требования и особенности, относящиеся к фактору ценности, учитываемому при создании АСОД.

Процесс определения потребностей в информации (определения выхода системы) является итеративным процессом в том смысле, что потребности после их предварительного определения должны уточняться (критически анализироваться) с учетом возможностей и стоимости получения информации, ее ценности в каждом конкретном случае использования.

Информация, передаваемая пользователю в виде сообщений, может быть различной по форме представления, применяемым носителям, своему содержанию (точнее, структуре) на каждом конкретном носителе. Поэтому при организации систем обработки данных вопросы ценности информации следует рассматривать с учетом особенностей формирования конкретных сообщений и размещения информации на носителях. При решении этих вопросов прежде всего необходимо выявить удобства восприятия информации ее пользователями. Как показывает практика проектирования АСОД, восприятию содержания информации часто уделяется явно недостаточное внимание. В решении данного вопроса вырисовываются два направления:

- 1) выбирается такая форма представления информации, чтобы ее пользователю она была наиболее доступна, воспринималась достаточно четко;

- 2) детально определяется информация, которую целесообразно включать в передаваемые пользователям сообщения (табуляграммы, документы и др.). Она распределяется (в виде показателей) по отдельным сообщениям, выбираются форма представления информации и состав реквизитов и показателей с учетом полезности информации и затрат на ее формирование. Тем самым можно более четко определить ценность конкретной представляемой пользователям информации. Например, если носителем информации является бумага (информация представляется в виде табуляграмм), то ее содержание целесообразно определить в следующем порядке:

распределить информацию по документам, образуя автономные подмножества результатной информации;

определить целесообразную избыточность данных как некоторое пересечение выделенных подмножеств, способствующее обеспечению достоверности информации и создающее определенные удобства для пользователей;

выявить затраты на формирование информации всех подмножеств, минимизировать их, определить ценность информации (со-

поставить затраты с полезностью или эффектом от использования информации);

уточнить содержание информационных подмножеств с учетом полезности и затрат на формирование данных.

5.2. Точность и достоверность данных

Процесс машинной обработки данных можно представить в виде потока, на вход которого поступают исходные данные, а на выход подаются те же или преобразованные (результатные) данные. Последние передаются пользователям для выполнения соответствующих функций организационно-экономической деятельности. Потребности в результатных данных и их основные характеристики определяют пользователи этих данных. Поэтому можно принять, что в АСОД требования к выходу (результатным данным) являются заданными. Среди таких задаваемых характеристик данных важное место занимают их точность и достоверность.

Понятие точности уже давно широко используется в технике. Например, в чертежах деталей все размеры указываются с пределами допускаемых отклонений. Это вполне понятно, так как не всегда требуется одинаковая точность, а ее повышение связано с дополнительными затратами. Аналогичное положение практически существует и при определении конкретных числовых значений данных (оснований показателей). Например, заработная плата насчитывается с точностью до копейки, различные нормы определяются с точностью до 3—5 знаков после запятой и т. п. Поэтому при проектировании машинной обработки данных необходимо выяснить, какая степень точности формируемых в системе результатных данных удовлетворяет пользователей, и разработать процессы формирования этих результатов так, чтобы обеспечить эту точность. При этом важно помнить о том, что точность данных теряется лишь при их округлении, сложении (вычитании) данных различной точности, их умножении и делении. Поэтому при разработке алгоритмов и программ в промежуточных данных следует предусмотреть после запятой на 1—2 знака больше, чем требуется в результатах. Более детальное рассмотрение этих вопросов относится к области теории погрешностей, которая особенно широко применяется в математике и физике.

Другой характеристикой оценки АСОД является достоверность данных. Под достоверностью данных понимается их соответствие той действительности, которую они представляют, т. е. уровень адекватного отображения протекающих процессов или явлений.

Для соблюдения требований по обеспечению достоверности при проектировании АСОД необходимо четко проанализировать возможные причины искажения данных, выбрать методы выявления этих искажений в процессах обработки данных, предусмотреть меры по устранению возникающих ошибок и обеспечению требуемой достоверности. При решении этих вопросов в первую очередь необходимо определить возможные места возникновения искажений.

Данные недостоверными могут быть уже на входе системы или могут терять достоверность при их обработке в самой системе. В свою очередь, в процессе обработки достоверность данных может уменьшаться при их передаче, хранении, перекодировании на другие носители и преобразовании по заданным алгоритмам. Таким образом, искажения могут возникать на любой операции технологического процесса формирования и использования базы данных системы.

Потеря достоверности данных в АСОД может произойти из-за искажения или изменения символов, в которых они закодированы, при записи на определенные носители, при передаче или перезаписи на другой носитель, из-за частичной или полной потери данных при хранении и ошибок, возникающих при преобразовании данных. Для различных видов искажения используются свои методы их выявления.

Ошибки в кодировании данных возникают при их записи, перезаписи и передаче по каналам связи. Для обнаружения таких ошибок, т. е. обеспечения достоверности данных на уровне символов (для синтаксического контроля), строятся специальные коды, позволяющие автоматически обнаруживать эти ошибки: запрещающие, корректирующие и др. (см. 7.2). Для этого также используются счетные методы контроля: метод контрольных сумм (по строкам, столбцам, группам данных), метод двойного (повторного) счета, балансовый контроль и др.

Во избежание потери данных при хранении принимаются специальные меры. Во-первых, должна быть обеспечена защита самих машинных носителей от механических повреждений и физических воздействий (например, размагничивания). Для этого машинные носители с данными должны храниться в специальных местах с соблюдением установленных правил. Во-вторых, нужно предусмотреть меры предосторожности от уничтожения данных. Эти средства должны защищать данные от их стирания при обработке, от внесения в них изменений, не санкционированных лицами, ответственными за сохранность данных и поддержание их в активном состоянии. Для этого следует организовать хранение нескольких поколений или копий массивов, предусмотреть специальные программные средства защиты данных.

Преобразование данных в АСОД обычно сводится к формированию массивов и табуляграмм. Поэтому для выявления возможных ошибок и искажений данных необходимо прежде всего определить все машинные процедуры, выполняемые над массивами данных и их элементами. Среди таких процедур преобладают следующие: ввод, сортировка, объединение и разъединение массивов, формирование промежуточных и результатных массивов, формирование и печатание табуляграмм. При этом непосредственная реализация применяемых экономико-математических методов осуществляется путем выполнения процедур формирования результатных данных, когда производится арифметическая и логическая обработка исходных данных. Для обеспечения достоверности по-

лучаемых результатов используются схемные и программные методы контроля. Первые заложены в конструкции машин. Их примером может служить контроль ввода данных в ЭВМ, когда каждый вводимый символ проверяется на нечетность (на представление его нечетным количеством двоичных знаков или импульсов).

Суть программных методов контроля заключается в том, что при разработке алгоритмов предусматриваются дополнительные операции по проверке имеющихся арифметических, логических или других связей между определенными данными, по получению различными способами нескольких (чаще всего — двух) результатов и их сопоставлению и т. п.

Первый вид проверок позволяет осуществить семантический контроль. Методы этого контроля основываются на использовании смыслового значения данных. Важность таких методов заключается в том, что при контроле можно обнаружить искажения данных, допущенные на любом этапе их движения и преобразования, в том числе и на первичном. При помощи таких средств также можно обнаружить ошибки в первичных данных, приписки и другие преднамеренные искажения, недобросовестность выполнения работ по подготовке данных и т. п. Такие меры могут помочь выполнять на предприятии функции контроля и ревизии хозяйственной и организационной деятельности. Суть семантического контроля заключается в том, что проверяются смысловые значения данных, попадание их в определенный допустимый интервал, их согласованность, непротиворечивость и т. д.

Наличие различных методов и средств обеспечения достоверности данных на выходе системы позволяет выявить ошибки в исходных данных, но полностью все же не обеспечивает их достоверности. В немашинных системах первичные данные для обеспечения их достоверности фиксируются в документах. Основной особенностью документов является то, что достоверность записанных в них данных подтверждается подписями соответствующих компетентных лиц. При проставлении подписи дополнительно проверяется соответствие данных действительности. Однако использование на предприятиях и других объектах большого количества различных форм первичных документов тормозит внедрение средств автоматизации сбора и кодирования данных на машинные носители, а ведь любая автоматизация может находить практическую поддержку лишь тогда, когда она проводится с достаточным обеспечением необходимой достоверности данных. Достоверность данных можно реализовать различными способами. Так, например, можно создать специальные технические устройства, на которые нельзя оказать нежелательные воздействия в отношении достоверности данных; можно организовать дополнительный контроль правильности этих данных. Практически для таких целей одновременно с автоматической фиксацией и кодированием исходных данных часто формируются и первичные документы, которые позволяют проверить и подтвердить достоверность этих данных.

Рассматриваемые вопросы обеспечения достоверности резуль-

татных данных играют важную роль при проектировании АСОД. Контроль достоверности данных в каждом конкретном случае должен осуществляться на основе определения наиболее реально возможных способов контроля по каждой операции технологического процесса и выбора такого их сочетания, которое обеспечило бы требуемую достоверность при наименьших затратах на обработку данных и выполнение операций контроля. Поэтому перед изучением методов оценки затрат на обработку данных рассмотрим вопросы оценки достоверности данных.

5.3. Оценка и повышение достоверности данных на выходе АСОД

Искажения данных в системах их машинной обработки могут возникать в результате ошибок, причинами которых могут быть:

- 1) плохая работа персонала, обслуживающего систему (операторов);
- 2) сбои при работе технических средств;
- 3) неверная работа алгоритмов и программ;
- 4) некорректность общего технологического процесса обработки данных;
- 5) неточность в применяемой проектной документации.

Перед началом выполнения каждой операции технологического процесса проектная документация, инструкции, технические и программные средства должны быть соответствующим образом проверены. Такие мероприятия предусматриваются в проектной документации — общих инструкциях по выполнению операций, конкретных рабочих инструкциях. В алгоритмах должны быть предусмотрены точки выхода для проверки промежуточных результатов, позволяющие обнаруживать все искажения, отклонения от нормальной работы, которые не предусматривались при проектировании. При работе программ практически могут появляться только случайные искажения данных. Их могут вызывать сбои, различные случайные факторы или они могут появляться при возникновении ситуаций, которые не были заранее предусмотрены и поэтому предварительно не проверены. Также могут быть обнаружены ошибки, совершенные при проектировании. Таким образом, ошибки при обработке данных в АСОД по своей природе имеют случайный характер, а мера достоверности данных является вероятностной. Поэтому достоверность при выполнении i -й операции технологического процесса может быть выражена вероятностью того, что данные не будут искажены, т. е. $p_i = 1 - q_i$, где q_i — вероятность появления ошибок.

Практика показывает, что абсолютной достоверности достичь почти невозможно. Поэтому для конкретных групп данных необходимо установить предельные реально достигаемые нижние ее уровни. Например, для задач планирования и оперативного управления вероятность появления ошибок (искажения данных) считается удовлетворительной, если она колеблется в пределах 10^{-5} — 10^{-6} , а для задач учета — 10^{-6} — 10^{-7} .

Общая достоверность технологического процесса обработки данных P — это достоверность результатов. При этом можно воспользоваться следующими формулами:

если достоверности отдельных операций равны p_i , а операции являются взаимонезависимыми, то она равна произведению этих достоверностей:

$$P = \prod_{i=1}^n p_i,$$

где n — количество операций технологического процесса;

если в технологическом процессе на k -й операции осуществляется контроль ($k \leq n$), обнаруженная ошибка исправляется, и если при повторной проверке она уже не обнаруживается, то вероятность достоверности данных в этой части процесса можно определить так:

$$P = \prod_{i=1}^k p_i / \left(P_k + \prod_{i=1}^k p_i \right),$$

где P_k — вероятность необнаружения искажения после первого исправления.

При проектировании технологических процессов обработки данных и разработке алгоритмов варианты операций и их контроля следует выбирать так, чтобы они обеспечили требуемую достоверность результатов. Известно несколько таких способов, среди которых наибольшее распространение получили следующие:

1. Вводится информационная или программная избыточность. При этом на ЭВМ приходится выполнять дополнительные операции. Это может значительно увеличить время и стоимость работ по формированию результатов, хотя такой метод и позволяет осуществить наиболее глубокий и полный контроль.

2. Контроль децентрализуется, т. е. по возможности переносится в места формирования и предварительной обработки данных. Такой подход оправдывает себя в отношении задач, связанных с оперативным управлением и регулированием производства.

3. Автоматизируются операции сбора, передачи и записи данных на машинные носители. При повышении степени автоматизации увеличивается достоверность данных. Однако автоматизация связана с дополнительными расходами и определенными организационными трудностями (см. 5.2).

4. Разрабатываются меры по повышению достоверности выполнения операций. Например, можно предусмотреть дополнительные организационные или инженерно-психологические мероприятия, совершенствовать программные средства и т. д.

Таким образом, при проектировании АСОД следует решать задачи выбора методов, а также средств выполнения и проверки правильности всех операций технологических процессов, обеспечивающих необходимую достоверность и минимизирующих затраты на обработку данных. В частных случаях минимизации могут под-

вергаться время обработки данных, себестоимость или стоимость выполняемых работ, а максимизации — достоверность результатов данных и др.

Глава 6

ОЦЕНКА ЗАТРАТ НА ОБРАБОТКУ ДАННЫХ

6.1. Принципы определения затрат на обработку данных

В системах машинной обработки данных преобладает централизованная форма эксплуатации СВТ на вычислительных установках. Разновидность концентрируемых в одном месте средств вычислительной техники, форм организации технологических процессов обработки данных, источников и способов покрытия затрат на эксплуатацию этой техники, форм оплаты труда и другие причины привели к большому разнообразию вычислительных установок. Они различаются также по назначению и самостоятельности. Так, например, на промышленных предприятиях до сих пор функционируют машиносчетные станции, создаются отделы АСУ, в состав которых вычислительные установки входят как автономные подразделения. Однако с расширением применения электронных вычислительных машин основной формой и подразделением по эксплуатации вычислительных машин и выполнению вычислений становятся информационно-вычислительные центры (ИВЦ). Они могут обслуживать только отдельные предприятия, т. е. быть индивидуальными, или предоставлять услуги нескольким предприятиям и организациям. В последнем случае создаются кустовые ИВЦ или вычислительные центры коллективного пользования (ВЦКП).

Такое укрупнение позволяет создавать вычислительные центры с несколькими ЭВМ, которые могут функционировать как самостоятельные предприятия. На ИВЦ также распространены права и обязанности социалистического государственного предприятия, предусмотренные соответствующим положением, намечаются основные показатели, осуществляется народнохозяйственное планирование их деятельности. Поэтому о вычислительных установках вполне обоснованно можно говорить как о предприятиях по обработке информации — вычислительных предприятиях. Это позволяет использовать опыт, накопленный в различных отраслях народного хозяйства, по выявлению затрат на заработную плату (основную и дополнительную), амортизацию СВТ, материалы, электроэнергию и других затрат, определяемых прямым счетом, по выявлению и распределению косвенных затрат на конкретные работы (предоставляемые услуги по обработке данных). К последним следует отнести также те затраты, которые связаны с созданием и поддержанием в активном состоянии программных средств и массивов данных.

Программы в процессе производства на ИВЦ многократно используются для решения задач, но не видоизменяют своей натуральной формы и непосредственно не входят в создаваемый продукт. В связи с этим стоимость программных средств должна переноситься на продукцию, т. е. включаться в издержки производства определенными частями, по мере их постепенного износа. Средства программного обеспечения имеют такую же форму представления, как и данные (не вещественную, а только зафиксированную на определенных физических носителях), и постепенная утрата их стоимости в основном связана с моральным старением. Моральное старение происходит вследствие создания более совершенных в техническом и экономическом отношении средств того же или аналогичного типа и назначения.

Эти свойства программного обеспечения во многом схожи со свойствами средств труда, которые включаются в основные фонды предприятия и выступают как производственные фонды. Поэтому создание, движение и использование средств программного обеспечения должны найти соответствующее отражение в планировании и учете деятельности предприятия и оценке затрат на выполнение работ (изготовление продукции). Для этого вполне естественно применять принципы и методы, разработанные для основных фондов предприятия.

Особое место в процессе производства (в смысле влияния на себестоимость отдельных работ) занимают данные. Чтобы определить себестоимость конкретных работ, необходимо по возможности более точно учитывать затраты на формирование и ведение информационного фонда, используемого при решении комплексов задач. Для этого обратимся к схеме процессов формирования информационного фонда, представленной на рис. 12.1 (см. 12.1), где выделено формирование файлов трех групп (архивы ВхОИ, НСИ и базы данных). Файлы оперативной информации ВхОИ накапливаются, хранятся и используются только в определенном учетном периоде для формирования оперативных, учетных или плановых сообщений (табуляграмм). Таким образом, они полностью потребляются в течение одного производственного цикла, и поэтому затраты на них полностью включаются в себестоимость продуктов, в создании которых они участвуют.

В файлах второй и третьей групп содержится нормативная, справочная и некоторая другая информация. Они создаются не для отдельных задач, а как единые системы информации. Поэтому затраты на них должны распределяться по таким же принципам, как и затраты на основные фонды предприятия и программное обеспечение.

Косвенные затраты на отдельные работы распределяются пропорционально определенным прямым затратам, выбранным в качестве базы для такого распределения (например, пропорционально сумме основной и дополнительной заработной платы). Поэтому прежде всего должны выявляться прямые затраты. Для их определения весь технологический процесс обработки данных необходи-

мо представить в виде последовательности отдельных операций, выполняемых на различных рабочих местах с использованием определенных средств. Операции могут выполняться строго последовательно, а некоторые из них могут частично перекрываться. Для последовательных операций определяется время их выполнения, а для перекрывающихся — дополнительно и степень перекрытия, т. е. коэффициент, указывающий долю одновременно выполняемой части работ по отношению к одной из этих операций. С помощью временных оценок и коэффициентов перекрытия операций по времени определяется общее время на реализацию технологического процесса. Прямые денежные затраты на отдельные операции и весь технологический процесс определяются после оценки времени выполнения каждой отдельной операции. Рассмотрим эти два вопроса более подробно.

6.2. Определение трудовых затрат

Операции технологического процесса могут быть исполнительными и контрольными. Они могут выполняться последовательно или параллельно.

Для определения времени на отдельную операцию необходимо знать объем работ и норму выработки. Отдельно рассмотрим особенности выявления затрат на выполнение операций с применением клавишных, перфорационных и электронных вычислительных машин.

При обработке данных на КВМ обычно рассчитывается время, которое требуется для выполнения основных (арифметических) операций. Все другие затраты оцениваются путем умножения прямых затрат на соответствующий коэффициент оценки полных затрат.

При обработке данных на ПВМ затраты (трудоемкость) определяются по каждой операции (в нормо-часах) и рассчитываются по формуле

$$T = \frac{D}{N},$$

где D — объем работ;

N — часовая норма выработки.

Для оценки объемов работ, выполняемых на различных машинах, применяются соответствующие единицы измерения: для перфораторов — документострока или перфокарта, для сортировок — строка, для табуляторов — картоход и т. д.

При обработке данных на ЭВМ определить время ее работы гораздо сложнее. Время решения задач зависит от многих факторов: характеристик самих задач (их алгоритмов), применяемых программных средств, параметров технических средств, организации вычислительного процесса.

Для определения времени выполнения отдельных процедур машинной обработки информации и всего процесса в целом использу-

ются различные методы. Их можно разделить на две обобщенные группы: аналитические и статистические. Первые опираются на теорию алгоритмов и методы оценки их сложности. К ним относятся методы определения сложности вычислений на машинах Тьюринга, по которым можно получить оценки сложности и продолжительности вычислений на основании разработанных алгоритмов. Такой алгоритмический подход представляет особый интерес на этапе технического проектирования систем машинной обработки данных, когда уже известны общие схемы алгоритмов. В этих схемах выделяются операции, для которых можно рассчитать временные характеристики, выявить все ответвления и повторения. Так как изменение последовательности выполнения операций в алгоритмах носит вероятностный характер, то следует определять вероятность выполнения каждого оператора, входящего в разветвляющиеся и циклические участки алгоритма. После этого для каждого обобщенного оператора (процедуры) с учетом объемов обрабатываемой информации можно рассчитать время его реализации, а затем и время реализации всего алгоритма.

Для практических расчетов более пригодными являются статистические методы, основанные на выявлении таких процедур, для каждой из которых можно определить время выполнения, а также их последовательность.

Общую оценку времени (T) на решение задачи можно получить, если выделить следующие процедуры: ввода исходных данных (время ввода — $T_{\text{вв}}$), их обработки при решении задачи (T_0) и вывода результатов ($T_{\text{выв}}$). Тогда расчет производится по формуле

$$T = K_1 T_{\text{вв}} + K_2 T_0 + K_3 T_{\text{выв}},$$

где K_1, K_2, K_3 — коэффициенты совмещения операций ввода, обработки и вывода данных по отношению к операции, выбранной в качестве базовой (чаще всего — к операции ввода или обработки).

В общем случае при решении i -й задачи время ввода исходных данных с помощью j -го устройства можно определить по формуле

$$T_{\text{вв}}^{ij} = A_{\text{вв}}^i \left(\frac{1}{V_{\text{вв}}^j} + \frac{T_3^j}{P_m^i \cdot K_m^i} \right),$$

где $A_{\text{вв}}^i$ — общий объем вводимых данных;

$V_{\text{вв}}^j$ — скорость ввода;

T_3^j — время замены носителя данных (колоды перфокарт, ленты, диска);

K_m^i — коэффициент использования объема носителя данных;

P_m^i — объем носителя данных.

Время на обработку данных при решении i -й задачи определяется по формуле

$$T_0^i = \frac{A_0^i \cdot a^i}{V_{\text{пр}}} K_K K_A,$$

где A_0^i — общий объем обрабатываемых данных (показателей);

a^i — среднее количество элементарных операций обработки одного показателя;

$V_{\text{пр}}$ — средняя скорость работы процессора;

$K_{\text{к}}$ — удельное время удлинения расчетов на каждый показатель за счет программных методов контроля;

$K_{\text{а}}$ — коэффициент увеличения расчетов в связи с автоматизацией программирования.

Время вывода результатов рассчитывается по формуле

$$T_{\text{выв}}^{ij} = \frac{A_{\text{выв}}^i}{V^j},$$

где $A_{\text{выв}}^i$ — объем выводимых данных при решении i -й задачи;

V^j — средняя скорость работы j -го устройства вывода.

Наиболее точно можно определить время на ввод и вывод информации. Здесь определенные трудности возникают лишь при установлении степени уменьшения технических скоростей соответствующих устройств. Так, например, эксплуатационная скорость ввода массива перфокарт для его обработки практически на 10—30% ниже технической и не превышает 900 перфокарт в минуту (для устройства ЕС-6016), а часто может быть и ниже — 700—800 перфокарт в минуту. Скорость ввода еще меньше при обработке больших массивов. Тогда целесообразно учитывать коэффициент использования устройства, который в среднем достигает значения 0,9.

Техническая скорость устройства вывода на печать в конструкторской документации определена в 380—420 строк в минуту. Для практических расчетов более удобно перевести строки в количество печатаемых листов. Так как на каждом стандартном листе печатаются 68 строк, техническая скорость будет 6 листов в минуту. Однако практически эксплуатационная скорость печатающего устройства в среднем составляет лишь 4 листа в минуту.

Время обработки данных при решении задач (время работы процессора) зависит от многих факторов и колеблется в больших пределах. На это время влияют: характер самой задачи (алгоритм ее решения), требования к достоверности данных, степень использования хранимых на машинных (магнитных) носителях нормативно-справочных и оперативных данных, характеристики устройств ввода всей исходной информации и вывода промежуточных и результативных данных, надежность и рациональность организации выполнения процедур, качество и экономичность используемых программных средств. Как показывает опыт, при обработке информации для обеспечения ее достоверности, т. е. для выполнения контрольных операций, времени тратится больше, чем непосредственно на расчеты. Это время можно определить путем анализа алгоритмов, а при наличии готовых программ время решения задачи можно установить путем хронометража. Для получения средних значений расчеты необходимо проводить несколько раз, используя различные исходные данные.

Для более точного определения потребностей в машинном времени процесс решения задачи необходимо детализировать. Напри-

мер, можно выделить следующие группы процедур: арифметические (вычисления промежуточных и окончательных результатов по формулам); проверки логических условий, представленных в виде уравнений или неравенств; формирования записей файлов на машинных носителях, строк данных и итогов табуляграмм; сортировки (одного файла, нескольких файлов, их объединения и разъединения).

Так, например, для укрупненной оценки времени выделенных процедур, выполняемых на малых ЭВМ, можно воспользоваться следующими формулами:

$$T_{\text{вв}}^i = \frac{A_{\text{вв}}^i}{V_{\text{вв}}} \approx \frac{A_{\text{вв}}^i}{270}; \quad T_{\text{выв}}^i = \frac{A_{\text{выв}}^i}{V_{\text{выв}}} \approx \frac{A_{\text{выв}}^i}{4};$$

$$T^i = T_{\text{вв}}^i (1 - K_c) + T_o^i + T_{\text{выв}}^i (1 - K'_c) \approx 0,25 T_{\text{вв}}^i + T_o^i + 0,1 T_{\text{выв}}^i,$$

где $V_{\text{вв}}$, $V_{\text{выв}}$ — эксплуатационные скорости ввода (с перфокарт) и вывода (на печать);

K_c , K'_c — коэффициенты совместимости работы устройств ввода (с перфокарт) и вывода (на печать) с работой процессора.

Для более точной и полной оценки времени на ввод, вывод и обработку данных необходимы дополнительные сведения о том, какая исходная информация хранится в базе данных, как организованы входные файлы, какими порциями считываются их записи, какая часть файлов нормативно-справочной информации используется и др.

При использовании пакетов прикладных программ основные процедуры и возможные методы их реализации для пользователей оказываются предусмотренными заранее. Так как конкретная программа разрабатывается по заказываемым характеристикам, то для определения времени решения задачи естественно воспользоваться следующим приемом: при апробации разработанного ППП выявляются факторы, определяющие время выполнения расчетов по генерируемым программам, их влияние на данный показатель; даются экспертные оценки времени работы получаемых программ.

Из изложенного следует, что проблемы определения временных затрат на решение различных задач являются сложными и в каждом конкретном случае имеют немало специфических особенностей. Именно поэтому для их решения сначала весь процесс необходимо разбить на явно автономные части или этапы и для каждого из них найти методы, позволяющие определить затраты наиболее просто, но с достаточной точностью.

6.3. Определение затрат в денежном выражении и пути их снижения

Затраты, измеряемые временем работы технических средств или их обслуживания, служат мерой объемов работ, выполняемых на отдельных машинах. Для обработки данных на ИВЦ и форми-

рования выхода АСОД используются машины, которые являются несопоставимыми по стоимости, затратам на обслуживание, эффективности и т. п. В таких случаях измерение затрат труда в натуральных единицах (часах или нормо-часах) не дает полного представления о действительных затратах на работы, выполняемые на различных машинах. Поэтому затраты на решение задач на ИВЦ определяются в денежном выражении, что позволяет их суммировать и сопоставлять.

Расчеты могут производиться для различных целей: определения затрат на выполнение производственного плана ИВЦ или общей оценки его деятельности за рассматриваемый период (например, год); оценки затрат на сравниваемые варианты машинного решения задачи, комплекса задач или выполнения заказов.

Как уже отмечалось, проектирование машинной обработки данных носит вариантный характер, т. е. представляет собой процесс построения и оценки различных наиболее рациональных вариантов проектных решений и выбора из них наилучшего по определенному критерию. Для технологических процессов обработки данных критерием их эффективности обычно является минимум денежных затрат или максимум эффекта.

Варианты можно сравнивать лишь тогда, когда затраты на их реализацию сопоставимы по отдельным элементам, т. е. когда затраты определены по одинаковой методике и с одинаковой степенью детализации соответствующих составляющих. При этом для целей сравнения можно учитывать все затраты или только те, которые меняются наиболее существенно в других вариантах (других проектах), т. е. можно использовать частичные или полные оценки. Чаще всего для сравнения вариантов процессов выполнения работ принимаются следующие оценки:

1. Оценка затрат живого труда на выполнение основной работы (операций технологического процесса). В такие затраты, необходимые для выполнения всех операций, входят основная и дополнительная заработная плата с отчислениями на социальное страхование, определяемые по формуле

$$C_a = \left[\sum_{i=1}^n T_i u_i (1 + K_d) \right] (1 + K_c),$$

где T_i — время, необходимое на выполнение i -й операции;

u_i — часовая тарифная ставка на i -ю операцию;

K_d — коэффициент, учитывающий дополнительную заработную плату;

K_c — коэффициент, учитывающий отчисления на социальное страхование;

n — общее количество операций, необходимых для выполнения работы.

Такой упрощенный расчет удобно проводить в случаях, когда сравниваемые варианты процессов существенно отличаются только по затратам времени на обработку данных, а по другим статьям

затрат и стоимости используемых средств вычислительной техники они примерно одинаковы.

2. Себестоимость обработки данных на ИВЦ. На практике наиболее часто до сих пор применяется метод определения прямых затрат. Суть его заключается в том, что выделяются только заработная плата (C_z) и амортизационные отчисления (C_a). Все остальные затраты относятся к накладным (C_n) и распределяются пропорционально основной и дополнительной заработной плате. Общие затраты и их составляющие определяются по формулам:

$$C_c = C_z + C_a + C_n;$$

$$C_a = \sum_{i=1}^n T_i a_i; \quad C_n = \left[\sum_{i=1}^n T_i u_i (1 + K_d) \right] K_n,$$

где a_i — часовая амортизация машины, используемой для выполнения i -й операции;

K_n — коэффициент, учитывающий накладные расходы.

Если при реализации рассматриваемых вариантов существенно могут меняться некоторые другие затраты, относящиеся к накладным, то их следует выделить отдельно. Такими могут быть затраты на основные и вспомогательные материалы (C_m), затраты на электроэнергию на производственные нужды ($C_{эл}$), оплата специальной организации за техническое обслуживание вычислительной техники ($C_{ор}$), затраты на хранение данных на машинных носителях и их обновление (C_x) и др. Тогда себестоимость и дополнительно выделенные составляющие определяются по формулам:

$$C_c = C_z + C_a + C_m + C_{эл} + C_{ор} + C_x + C'_n;$$

$$C_m = \sum_{j=1}^m K_{mj} P_j; \quad C_{эл} = \left(\sum_{k=1}^p M_{эk} T_{фk} K_{эk} \right) P_э,$$

где C'_n — прочие накладные расходы;

K_{mj} — необходимое количество j -го материала ($j=1 \div m$);

P_j — цена единицы j -го материала с учетом транспортных и складских расходов;

$M_{эk}$ — суммарная мощность k -го оборудования ($k=1 \div p$);

$T_{фk}$ — годовой фонд времени работы k -го оборудования;

$K_{эk}$ — коэффициент использования мощности k -го оборудования;

$P_э$ — цена одного киловатт-часа электроэнергии.

Оплата специальным организациям за техническое обслуживание производится по заключенным договорам и специальным расчетам. Затраты на хранение данных могут определяться исходя из их объемов и затрат на хранение единицы данных (массива на ленте, диске, перфокартах).

Отметим также, что расчеты затрат на решение задач с использованием различных ЭВМ могут производиться по формуле

$$C_{за} = \sum_{l=1}^l T_{эл} C_{эл} + C_{по},$$

где $T_{эл}$ — время работы l -й ЭВМ при решении задачи (в нормо-часах), когда $l = 1 \div r$;

$C_{эл}$ — себестоимость одного машино-часа l -й ЭВМ;

$C_{по}$ — себестоимость работы по подготовке данных для непосредственного решения задачи на ЭВМ.

3. Затраты предприятия (заказчика) на обработку данных. Стоимость работ для заказчика определяется на основании потребности машинного времени в нормо-часах по каждой операции технологического процесса. Для подготовительных операций, выполняемых на КВМ и ПВМ, должны использоваться единые нормы времени (выработки), утвержденные постановлением Государственного комитета Совета Министров СССР по труду и социальным вопросам № 46 от 17 февраля 1977 г. Объем вычислительных работ рассчитывается по расценкам прейскурантов № У-1-94 (для КВМ и ПВМ) и № У-01 (для услуг ИВЦ).

Для определения затрат предприятия на обработку данных с использованием единых тарифов составляются расчетные калькуляции на единицу работы. Единицей калькуляции могут служить обрабатываемый документ, документострока, задача с конкретным объемом исходных данных или результатов и т. п.

Если предприятие обрабатывает данные на собственном ИВЦ, то текущие затраты на обработку данных вычисляются по формуле

$$C_{\tau} = C_c + C_{пс} + C_{пер},$$

где $C_{пс}$ — затраты на обработку данных (в том числе полученных результатов) на самом предприятии;

$C_{пер}$ — затраты на передачу данных в ИВЦ.

4. Полные, т. е. текущие и единовременные затраты. Их сравнение показывает экономию приведенных затрат, которые определяются по формуле

$$C_{при} = C_{\tau} + E_{пвт} \cdot C_e,$$

где C_e — единовременные затраты, необходимые для внедрения АСОД;

$E_{пвт}$ — нормативный коэффициент эффективности капитальных (единовременных) вложений.

Единовременные затраты состоят из предпроизводственных расходов и затрат на оборудование и строительно-монтажные работы, т. е. в них входят затраты:

а) на разработку проектов (постановок задач, технологических процессов их решений, инструкций);

б) на создание массивов данных длительного пользования (нормативных, справочных, баз данных);

в) на отладку и экспериментальную проверку проектов с использованием соответствующих машин;

г) на приобретение и модернизацию оборудования;

д) на дополнительные производственные мощности;

е) по ликвидации технических средств;

ж) в результате изменения величины оборотных средств.

Кроме того, из них исключается остаточная стоимость высвобождаемого оборудования.

Расходы на обработку данных всегда увеличивают общие затраты на управление. Внедрение и использование СВТ будет тем эффективнее, чем меньше придется расходов на передачу, хранение и преобразование данных. Поэтому как при проектировании, так и в процессе эксплуатации АСОД необходимо искать пути снижения трудовых и денежных затрат.

Как уже отмечалось, наиболее полно расходы на систему машинной обработки данных отражаются в приведенных затратах, так как в них входят как текущие, так и единовременные расходы. Поэтому первым (но не самым простым) направлением снижения затрат на АСОД является уменьшение расходов по их составным элементам. Для этого прежде всего необходимо выявить те элементы затрат, которые оказывают основное влияние на общие результаты, и предусмотреть мероприятия по их снижению. Однако для снижения различных затрат имеются далеко не одинаковые возможности. Например, капитальные вложения часто определяются более широкими, чем рассматриваемые в конкретном случае, целями, и их снижение может уменьшить эффективность других мероприятий. Поэтому основными источниками экономии затрат следует считать работы по проектированию рациональной структуры информационной базы, рациональных технологических процессов и алгоритмов обработки данных.

Сокращению затрат на проектирование способствует его автоматизация, применение готовых или типовых проектов, отдельных проектных решений. Реализация других мероприятий по снижению затрат на обработку данных непосредственно связана с выбором наилучших вариантов конкретных проектных решений, осуществляемым пересмотром всех возможных вариантов или путем выбора наилучшего из них с применением математических методов оптимизации. Более подробно эти вопросы рассматриваются при изучении конкретных работ по проектированию в третьем и четвертом разделах учебника.

Глава 7

СИСТЕМЫ КЛАССИФИКАЦИИ И КОДИРОВАНИЯ

7.1. Классификация экономических данных

Классификация в общем смысле представляет собой научный метод исследования природы объектов и явлений с использованием определенных правил их систематизации. Совокупность правил и результатов распределения заданного множества объектов на подмножества в соответствии с установленными признаками их сходства или различия принято называть *системой классификации*, а сам процесс распределения объектов по данной системе — *классифицированием*. Классификация — неременный этап предварительной подготовки экономических данных к машинной обработке, а также предпосылка рациональной организации информационного фонда и моделирования информационных процессов. Ее можно характеризовать как составную часть информационного обеспечения любой АСУ, относящуюся к языковым средствам экономического управления.

Условием проведения любой классификации является наличие множества взаимосвязанных объектов и характеризующих их свойств, которые принимаются за *основание деления* и представляют собой набор *классификационных признаков*. Если в процессе деления определенных объектов учитываются естественные признаки, исходящие из самой сущности этих объектов, то в этом случае классификация отражает существующие в действительности их качественные особенности и представляет собой естественную, содержательную систематизацию данного множества. Для конкретных областей практического применения могут использоваться искусственные (формальные) классификационные признаки, имеющие вспомогательное значение для удобства упорядоченного изучения объектов и накладывающие определенные ограничения на возможности их деления. В связи с этим различают естественную и искусственную классификации. Часто их еще называют соответственно содержательной и формальной. Наиболее совершенным способом систематизации является естественная классификация. При любой классификации должны соблюдаться следующие основные требования:

полнота охвата объектов определенного множества;

непересечение выделяемых групп объектов;
возможность включения новых объектов;
лаконичность, четкость и ясность классификационных признаков;

сохранность установленного классификационного признака на каждом уровне классификации.

Часть объектов, объединяемых по одному или нескольким признакам, представляет собой *классификационную группировку*. Результат очередного распределения объектов одной классификационной группировки принято называть *ступенью классификации*, а совокупность классификационных группировок, расположенных на одних и тех же ступенях, — *уровнем классификации*. Одно и то же множество можно классифицировать по нескольким основаниям, что обусловлено различными аспектами использования данных. Выбор критериев классификации в каждом конкретном случае определяется целями решения как отдельных задач управления экономическим объектом, так и их комплексов.

В зависимости от объединения выбранных оснований классификации и порядка деления объектов различают иерархическую и многоаспектную системы классификации.

Иерархическая система предусматривает деление множества объектов с установлением многоуровневых отношений между классификационными группировками. Такую систему можно представить в виде графа типа дерева. Каждому объекту в данном случае на каждой ступени отводится место только в одной классификационной группировке, т. е. классификационные группировки не повторяются и каждая из них делится только по одному и тому же основанию классификации. Таким образом, логически стройная иерархическая классификация обладает жесткой структурой, фиксированным составом и порядком следования классификационных признаков. В связи с этим она не позволяет производить произвольные группировки объектов по таким сочетаниям признаков, которые не были предусмотрены иерархической классификацией. Тем самым использование данной системы классификации нередко вызывает трудности при решении задач управления, которые связаны именно с непредусмотренными сочетаниями признаков используемых показателей. Для преодоления этих трудностей в иерархические классификации необходимо заранее включить определенную резервную емкость, т. е. дополнительные классификационные группировки внутри каждого уровня, а иногда увеличить и количество уровней классификации (глубину системы).

Указанные недостатки отсутствуют в *многоаспектной системе* классификации, при которой множество объектов делится одновременно по всем основаниям классификации, т. е. параллельно по нескольким независимым признакам. Таким образом, каждый объект может быть отнесен к различным независимым классификационным группировкам в зависимости от нужного аспекта классификации, который определяется соответствующим признаком или их сочетанием.

Для экономической информации наиболее пригодными являются две формы многоаспектной классификации: дескрипторная и фасетная.

Дескрипторная классификация основывается на применении метода координатного (ассоциативного) индексирования, который заключается в отражении классифицируемых объектов и их признаков (свойств) с помощью ключевых слов специального формализованного языка. Каждому классификационному признаку в данном случае соответствует нормализованное ключевое слово, называемое *дескриптором*. Дескрипторы, точно и однозначно выражающие определенные признаки объекта, определяются путем нормализации ключевых слов-синонимов, имеющих одинаковый смысл при разном написании, и слов-омонимов, имеющих разный смысл при одинаковом написании, с целью выбора нормализованного слова.

Дескрипторы включаются в *тезаурусы* (с латинского языка — сокровище), т. е. в систематизированные списки объектов, их признаков (свойств) с ключевыми словами; в них фиксируются постоянные, не зависящие от контекста классификационные отношения между этими признаками, имеющие ассоциативный, синонимический или родо-видовой характер. Ассоциативные дескрипторы отражают объекты, которые обладают определенными общими признаками. Синонимические дескрипторы относятся к объектам с полностью совпадающими признаками, родовые выражают наиболее существенные признаки, которые присущи нескольким объектам, относящимся к одному роду, а видовые дескрипторы, наоборот, характеризуют свойства объектов данного вида, входящего в состав определенного рода.

На основании разработанных тезаурусов объекты соответствующего множества подвергаются координатному (ассоциативному) индексированию с использованием дескрипторов, соответствующих нужному произвольному сочетанию признаков. Именно в процессе такого индексирования, который сводится к подборке нужных ключевых слов, и реализуется дескрипторная классификация экономической информации.

Фасетная классификация осуществляется с использованием параллельных независимых фасетов, представляющих собой некоторые наборы признаков, по которым выделяют независимые классификационные группировки. Отдельные признаки внутри фасетов при этом часто образуют иерархическую структуру. Классификационные группировки из разных фасетов позволяют производить деление объектов по любому возможному сочетанию признаков, при этом признаки, используемые в различных фасетах, не повторяются. Порядок следования фасетов для конкретной классификации задается структурной формулой и зависит от требований, вытекающих из условий решения определенных задач управления. Фасетная классификация наиболее рациональна и удобна для сложной логической обработки данных, хотя она характеризуется и большой сложностью построения. Кроме того, ограниченность

набора фасетов не позволяет использовать произвольные сочетания классификационных признаков, как при дескрипторной классификации, но все же обладает значительно большей гибкостью по сравнению со стандартными иерархическими основаниями классификации.

Выбор той или иной системы классификации зависит от цели проводимых исследований экономических данных и должен производиться с учетом удовлетворения требований всего комплекса задач управления, межклассификационных связей, простоты разработки и ведения классификаторов, отображающих порядок деления множества объектов и позволяющих производить кодирование номенклатур.

7.2. Кодирование номенклатур

Номенклатура представляет собой перечень (список) имен всех объектов какой-либо совокупности, равнозначных в отношении друг к другу и упорядоченных по определенным классификационным признакам.

Цель кодирования номенклатур экономических данных (показателей), отражающихся в соответствующих классификаторах, состоит в том, чтобы представить информацию в компактной и удобной форме. Основой для кодирования информации является ее классификация. *Процесс кодирования* — это присвоение условных обозначений различным объектам определенной номенклатуры по установленным правилам на базе принятого для этого алфавита.

К кодированию, как известно, привело стремление к сокращению описаний хозяйственных процессов и явлений. По мере развития вычислительной техники необходимость в кодировании информации неуклонно возрастала, особенно с совершенствованием перфорационных и развитием электронных вычислительных машин. Это обусловлено прежде всего такими особенностями экономической информации, как высокий удельный вес ее представления в алфавитном виде, большой объем, преобладание логических операций обработки. Использование кодовых обозначений позволяет сократить затраты на составление документов, перенос данных на машинные носители, передачу их по каналам связи, ввод в ЭВМ и машинную обработку при решении задач, вывод результатной информации.

Коды отдельным объектам номенклатур, как уже отмечалось, присваиваются по определенным правилам, совокупность которых представляет собой *систему кодирования*. В зависимости от основы кодирования различают регистрационные и классификационные системы кодирования. Если конкретная номенклатура формируется без предварительной классификации ее объектов по определенным признакам, то им присваиваются *регистрационные коды*, которые строятся с использованием порядковой или серийной системы кодирования, а также системы повторения. *Классификационные коды* присваиваются объектам, предварительно систематизиро-

рованным по определенным классификационным признакам. Они строятся по разрядной (позиционной) или комбинированной системе кодирования.

Порядковая система применяется для кодирования однопризначных, стабильных и простых номенклатур. Она предусматривает присвоение объектам цифровых номеров в порядке их расположения в номенклатуре из натурального ряда чисел без пропуска номеров. Достоинствами данной системы следует считать простоту построения кодов, их малую значность, плотность записи. Вместе с тем порядковые коды не предусматривают группировки объектов по признакам, что затрудняет автоматизацию получения итогов по однозначным признакам при решении задач на ЭВМ. Кроме того, с появлением новых объектов данной номенклатуры происходит нарушение принятой классификации. Это ограничивает применение порядковой системы кодирования рамками стабильных номенклатур. К таким номенклатурам можно отнести, например, категории работающих, отрасли народного хозяйства, некоторые уровни административного деления территории страны и др.

Серийная система служит для кодирования аналогичных простых номенклатур и предполагает присвоение серий номеров объектам, выделенным в отдельные группы по какому-либо признаку. В пределах каждой серии объектам присваиваются номера по порядку. При этом в каждую серию обязательно включаются резервные коды, которые могут присваиваться новым объектам данной номенклатуры. Резерв может быть предусмотрен как в начале, так и в конце каждой серии. Серийная система наиболее экономична по количеству разрядов и вместе с тем содержит необходимый резерв свободных номеров для новых объектов. Это позволяет избежать нарушения классификации с расширением номенклатуры, а автоматическое получение итогов по однозначным классификационным признакам может быть обеспечено с использованием дополнительных сведений о границах каждой серии номеров. Данная система применяется для кодирования относительно стабильных номенклатур, таких, как структурные подразделения предприятия, операции движения материалов, виды оплат и удержаний и др.

Система повторения использует буквенные или цифровые обозначения, непосредственно характеризующие кодируемый объект (например, числовое значение номиналов, вес, размер объекта, разрядность) или ассоциативно связанные с ним некоторые данные, такие, как место расположения, адрес, относительное смещение и т. д. Коды повторения в чистом виде используются редко, но они могут входить в комбинированные коды.

Наиболее простым и наглядным примером использования системы повторения в сочетании с порядковой системой может служить построение кодов единиц измерения. Код единицы измерения состоит из 3 разрядов: старший разряд — вид измерения и его основная единица, кодируемые по порядковой системе (например, 1 — длина, основная единица — метр; 2 — площадь, квадратный

метр; 3 — объем, кубический метр; 4 — масса, килограмм; 5 — штучное количество, штука; ...; 9 — денежная сумма, рубль); второй разряд, принимающий значение нуля или единицы, определяет направление смещения от основной единицы данного вида измерения (0 — в сторону уменьшения, 1 — в сторону увеличения); младший разряд формируется по системе повторения и равняется количеству десятичных разрядов величины смещения от основной единицы измерения. Оно выражается в виде показателя степени основной единицы измерения (без знака). Таким образом, одна из характеристик самой кодируемой единицы измерения непосредственно входит в ее код.

Коды конкретных единиц измерения могут принимать такие значения: 100 или 110 — метр, 101 — дециметр, 102 — сантиметр, 103 — миллиметр, 113 — километр, 116 — тысяча километров; 900 или 910 — рубль, 902 — копейка, 903 — одна десятая копейки, 913 — тысяча рублей, 916 — миллион рублей, 919 — миллиард рублей и т. п. Использование таких кодов позволяет автоматически, программным путем, без дополнительных таблиц расшифровки всех значений кодов выравнивать положение десятичной точки в количественных и суммовых показателях, выраженных в разных единицах измерения.

Рассмотрим еще один возможный вариант использования системы повторения для ассоциативного кодирования работающих, т. е. присвоения им табельных номеров, совпадающих со значением соответствующих адресов записей с данными о каждом работающем на магнитном диске. Такие табельные номера имеют следующую структуру (разряды в скобках могут отсутствовать):

$$\frac{xxx}{Ц} \quad \frac{[x]x}{Д} \quad \frac{[x]}{С} \quad \frac{[xx]x}{Н},$$

где Ц, Д, С обозначают соответственно адреса цилиндра, дорожки и сектора пакета магнитных дисков, а Н — порядковый номер записи.

Ассоциативные табельные номера наиболее пригодны для крупных предприятий и производственных объединений. Их также можно использовать для оперативного управления кадрами на уровне министерств и ведомств. В данном случае может возникнуть необходимость добавления номера тома (пакета дисков), если файл справочных данных о работающих занимает больше одного тома. Увеличение количества разрядов таких табельных номеров окупается возможностью ускоренного поиска справочных данных о любом работающем на всех трех уровнях управления (предприятие — объединение — министерство). Добавление новых записей в ассоциативный файл справочных данных при наметившейся в последние годы тенденции значительного сокращения текучести кадров не требует большого резерва дисковой памяти, тем более, что свободные табельные номера уволенных работающих могут быть использованы уже в следующем календарном году, а переход ра-

ботающего на другое предприятие или объединение того же министерства не влечет за собой изменения его табельного номера. В конечном же итоге такой метод кодирования может обеспечить переход к республиканским и единому общесоюзному кодификаторам работающих в процессе постепенного наращивания общегосударственной системы учета и планирования народного хозяйства.

Разрядная (позиционная) система применяется для кодирования сложных многопризначных номенклатур: каждому классификационному признаку отводится определенное число разрядов (позиций), которое зависит от количества объектов в соответствующей классификационной группировке. Такая система кодирования соответствует иерархической классификации, т. е. классификационные группировки по младшим признакам кодируются в зависимости от кода группировки, выделенной по более старшему признаку. Разрядная система обеспечивает стройность и логичность кодов, четкое выделение каждого классификационного признака, удобство для машинной обработки информации, но вместе с тем требует увеличения разрядности кода. Кроме того, позиционные коды часто характеризуются большой сложностью построения и отсутствием необходимой гибкости при их структурном преобразовании.

По разрядной системе могут быть построены коды таких сложных номенклатур, как материальные ценности, направления производственных затрат, готовая продукция и др. Эта система особенно пригодна для двухпризначных номенклатур, например причин и виновников брака, простоев оборудования, видов изменений и др. (коды таких номенклатур иногда называют матричными или шахматными).

Для кодирования больших многопризначных номенклатур, которые одновременно характеризуются и соподчиненностью, и независимостью отдельных классификационных признаков, используется *комбинированная система*, базирующаяся на различных сочетаниях принципов кодирования по всем рассмотренным системам. Комбинированные коды при всей их стройности и логичности построения имеют наиболее рациональную структуру, обладают достаточной гибкостью, могут применяться как для иерархических, так и для многоаспектных номенклатур, построенных по принципам дескрипторной или фасетной классификации.

Независимо от используемой системы кодирования коды номенклатур должны быть прежде всего ориентированы на эффективную машинную обработку информации, иметь по возможности минимальную длину, обладать достаточной избыточностью и гибкостью. Они должны служить не только для экономии памяти ЭВМ и связанного с этим ускорения обработки данных, но и для повышения уровня автоматизации процессов обработки, к тому же должны быть удобными для пользователей.

При построении кодов следует также предусмотреть возможность автоматического обнаружения ошибок кодирования с помощью ЭВМ. С этой целью коды объектов дополняются контрольным разрядом, определяемым по установленному алгоритму из

значений отдельных разрядов кода. Такие расширенные условные обозначения получили название *кодов обнаружения ошибок*. Для расчета контрольного разряда наиболее широко используются методы контроля по модулю 11, 13, а также 9 и 10. Их сущность заключается в том, что вначале определяется сумма S разрядов кода (или сумма произведений P_i , полученных при умножении каждого разряда p_i на его вес или порядковый номер n_i в коде, т. е. $P_i = p_i n_i$), затем она делится на выбранный модуль m и полученный остаток d от деления вычитается из того же модуля. Полученная разность $k = m - d$ (или ее младший разряд) в виде контрольного разряда приписывается к коду справа. При вводе таких кодов в ЭВМ программным путем выполняются необходимые контрольные вычисления. Примеры построения кодов обнаружения ошибок по модулю $m = 11$ приведены в табл. 7.1.

Таблица 7.1

Построение кодов обнаружения ошибок ($m=11$)

Код (p_i)	Вес (n_i)	P_i	S	d	k	Код обнаружения ошибок (p_i, k)
507	—	—	12	1	0	5070
1364	—	—	14	3	8	13648
17082	—	—	18	7	4	170824
507	3,2,1	15,0,7	22	0	1	5071
1364	4,3,2,1	4,9,12,4	29	7	4	13644
17082	5,4,3,2,1	5,28,0,16,2	51	7	4	170824
507	4,2,1	20,0,7	27	5	6	5076
1364	8,4,2,1	8,12,12,4	36	3	8	13648
17082	16,8,4,2,1	16,56,0,16,2	90	2	9	170829

Автоматический контроль информации может осуществляться также с применением таких кодов, при построении которых заранее запрещается пользоваться некоторыми символами (цифрами) или определенными их сочетаниями (кодowymi словами).

Рассмотренные системы кодирования можно считать традиционными. С их использованием практически строятся только цифровые коды, которые являются наиболее компактными, так как в памяти ЭВМ они могут представляться в упакованном формате. Так, закодированные реквизиты можно обрабатывать на всех вычислительных машинах, в том числе и ЭВМ. На современном этапе развития систем машинной обработки данных и АСУ должны совершенствоваться традиционные и разрабатываться новые, более рациональные системы и методы кодирования, отвечающие требованиям, вытекающим из принципов интегрированной обработки данных и необходимости в автоматизации самого процесса кодирования. Одной из основных предпосылок повышения эффективности АСУ является создание единой системы классификации и кодирования данных.

7.3. Единые системы классификации и кодирования

Важное место в информационном обеспечении АСУ различных уровней занимает Единая система классификации и кодирования данных (ЕСКК), включающая в себя единую методологическую основу классификации и кодирования в разных органах управления, комплекс общесоюзных классификаторов технико-экономической информации (ОК ТЭИ) и автоматизированную систему ведения общесоюзных классификаторов (АСВОК ТЭИ). ЕСКК прежде всего обеспечивает обмен данными между различными уровнями управления народным хозяйством, сопоставление технико-экономических показателей, формируемых в разных органах управления, дает возможность агрегировать информацию и максимально автоматизировать ее сбор, накопление, обработку, хранение, поиск и выдачу.

Единую методологическую основу ЕСКК составляют документы, определяющие структуру, принятые в системе принципы и методы классификации и кодирования, применяемую терминологию, характеристику всех классификаторов и их взаимодействие, порядок пользования и внесения изменений. Совокупность классификаторов делится на три группы:

- общесоюзные, используемые во всех отраслях и на всех уровнях управления;

- отраслевые, используемые в пределах определенной отрасли;

- локальные, используемые в пределах одного объединения, предприятия или их группы.

В ЕСКК должна обеспечиваться полная совместимость всех этих групп на различных уровнях управления. При создании, например, отраслевых классификаторов накладываются ограничения на структуру и значность кодов, предусмотренных в общесоюзных классификаторах. Таким образом, ОК ТЭИ обеспечивают информационную совместимость автоматизированных систем управления различного назначения. В систему ОК ТЭИ в настоящее время включено более 20 общесоюзных классификаторов таких номенклатур, как продукция (классификатор ОКП), оборудование (ОКО), материалы и комплектующие изделия (ОКМКИ), профессии рабочих (ОКПР), должности служащих (ОКДС), виды и разряды работ (ОКВРР), административно-территориальные единицы (СОАТО), предприятия и организации (ОКПО), отрасли народного хозяйства (ОКОНХ), органы государственного управления (СООГУ) и др. Большинство ОК ТЭИ предусматривают два или три раздела описания каждой позиции классифицируемой номенклатуры: идентификации, наименования и фасетных классификационных признаков. Так, например, схема разделения объектов классификации по системе ОКП с примерами построения групповых и индивидуальных кодовых обозначений продукции представлена в табл. 7.2.

Для ведения классификаторов организована система специальных служб, распределенных по следующим четырем уровням:

Кодирование продукции по системе ОКП

Раздел классификатора	Наименование классификационного подразделения	Код	Наименование объекта классификации
Высшие классификационные группировки	Класс	38	Оборудование металлообрабатывающее и деревообрабатывающее
	Подкласс	38 1	Станки металлорежущие
	Группа	38 1 1	Станки токарной группы
	Подгруппа	38 11 3	Станки токарно-револьверные
	Вид	38 113 1	Станки токарно-револьверные с вертикальной осью револьверной головки
Серийно-порядковая часть	Типоразмер	38 1131 3	Станки токарно-револьверные с вертикальной осью револьверной головки с наибольшим диаметром обработки 16 мм
		38 1131 3 1	Станки токарно-револьверные с вертикальной осью револьверной головки с наибольшим диаметром обработки 16 мм нормальной точности
	Первичная группировка	38 1131 31 11	Модель 1Р316
		38 1131 31 12	Модель 1Е316

1) Госплан СССР и Госстандарт СССР, обеспечивающие организационное и методическое единство системы классификаторов;

2) отраслевые головные службы ведения классификаторов, осуществляющие методическое руководство в отрасли, координацию и контроль за выполнением работ по созданию и ведению ЕСКК, которые ведут эталонные экземпляры классификаторов и выпускают бюллетени изменений;

3) службы ведения классификаторов, осуществляющие классификацию и кодирование объектов в пределах отведенных им классов или отдельных группировок, а также ведение контрольных экземпляров отдельных видов общесоюзных и отраслевых классификаторов; они выпускают извещения об изменениях;

4) исполнительные службы ведения классификаторов на предприятиях, составляющие контрольные экземпляры всех видов используемых классификаторов, организующие внедрение ЕСКК на предприятиях, обеспечивающие службы предприятия нужными кодами и запрашивающие необходимые группировки и коды у соответствующих ведущих служб ведения классификаторов.

Успешное внедрение ОК ТЭИ в практику в значительной степени зависит от достоверности и полноты содержащейся в них информации. Поэтому поддержание классификаторов в актуальном и достоверном состоянии приобретает очень важное значение. Учитывая большой объем информации ОК ТЭИ, динамику и объ-

ем изменений по ним, оперативное обеспечение необходимой и достоверной информацией, содержащейся в классификаторах, становится возможным только в условиях АСВОК ТЭИ. Такой системой охвачены основные ОК ТЭИ, используемые прежде всего в АСПР и АСГС. Так, например, в рамках АСГС с 1977 г. функционирует АСВОК ТЭИ, обеспечивающая ведение вышеупомянутых классификаторов СОАТО, ОКПО, ОКОНХ и СООГУ. При этом она выполняет методическое руководство и контроль за ведением этих классификаторов в различных звеньях системы, сбор информации об изменениях и дополнениях в ОК ТЭИ, классификацию и кодирование новых объектов и группировок, внесение изменений и дополнений в классификаторы, периодическое оповещение абонентов системы об этих изменениях и дополнениях, обслуживание абонентов по разовым запросам и другие функции. Данная система имеет трехуровневую иерархическую организационную структуру и включает в себя совокупность подсистем, организованных на союзном, республиканском и областном уровнях АСГС, а также взаимодействует с соответствующими системами других отраслевых и территориальных органов управления.

Постоянное совершенствование системы ОК ТЭИ и АСВОК ТЭИ, расширение ЕСКК является неременным условием создания Общегосударственной системы сбора, передачи и обработки информации для учета, планирования и управления народным хозяйством (ОГАС) на базе функционирующих и разрабатываемых автоматизированных систем управления.

7.4. Методика выполнения работ по классификации и кодированию данных

Разработка кодов — это сложный процесс проектирования, включающий в себя разработку задания на построение классификаторов, подготовку методических материалов по их проектированию, сбор и подготовку исходных данных и, наконец, непосредственное составление, апробирование, корректировку и утверждение конкретных кодов.

Проектирование кодов производится в определенной последовательности:

- определяется номенклатура кодируемых признаков;
 - составляются списки (картотеки) наименований объектов каждой классификационной группировки;
 - систематизируется каждая номенклатура с учетом принятой системы классификации;
 - выбирается наиболее рациональная система кодирования, т. е. определяются правила записи кода;
 - каждому элементу номенклатуры присваивается кодовое обозначение (код);
 - окончательно оформляются закодированные номенклатуры в виде специальных кодовых книг (кодификаторов).
- Прежде всего должны быть изучены особенности, определена

номенклатура кодируемых объектов, ее динамичность, глубина классификации. При этом признаки классификации каждой номенклатуры должны быть согласованы с алгоритмами обработки кодируемых данных и увязаны с классификаторами других номенклатур с тем, чтобы обеспечить единство их использования на всех участках обработки данных.

Выбор системы кодирования осуществляется на основе построения различных возможных вариантов и оценки их эффективности. Такая оценка производится по критериям технической эффективности. Основными критериями могут служить: объем памяти ЭВМ, необходимой для размещения кодификатора (или длина кодовых обозначений), емкость кодификатора, степень избыточности, заполненности и т. п. Избыточность кода, например, появляется тогда, когда по системе классификации номенклатур для удобств пользователей или для обеспечения достоверности данных при их передаче и обработке некоторые возможные значения кодов не используются или запрещаются. При этом учитываются также удобство кода, возможность декодирования и другие качественные критерии. Очень важно выбрать наиболее простую систему кодирования и вместе с тем предусмотреть возможность автоматизации построения и ведения кодификаторов с использованием ЭВМ. Количественные оценки характеристик кодов являются предметом теории информации и кодирования. Рассмотрим некоторые из них.

Емкость классификатора при порядковой и серийной системе кодирования (E_{nc}) определяется по формуле

$$E_{nc} = m^{D_{nc}},$$

где m — основание кода;

D_{nc} — его длина.

Длина кода, в свою очередь, может быть определена по формуле

$$D_{nc} = \log_m E_{nc}.$$

Емкость классификатора при иерархическом поразрядно-последовательном кодировании (E_n) определяется по формуле

$$E_n = m(m^{D_n} - 1)/(m - 1),$$

а если при этом объекты кодируются с выделением блоков по n разрядов, то емкость классификатора (E_{nb}) рассчитывается по формуле

$$E_{nb} = m^n (m^{D_{nb}} - 1)/(m^n - 1).$$

Для фасетных классификаторов емкость можно определить по формулам:

$E_{\phi} = m \cdot D_{\phi}$, если объекты кодируются без выделения блоков;

$E_{\phi b} = \frac{D_{\phi b}}{n} \cdot m^n$, если в кодах выделяются блоки по n разрядов.

Для того чтобы определить, сколько информации об объектах классификатора приходится на один разряд кода, можно воспользоваться показателем степени информативности кодовых обозначений, рассчитываемым по формуле

$$C_n = \Pi/D,$$

где Π — количество признаков, отражающихся в рассматриваемом коде.

Степень заполненности классификатора (C_3) выражается отношением количества объектов (K), кодируемых по данному классификатору, к максимально возможному числу таких объектов ($K_{\max}$) при установленной длине кода, т. е.:

$$C_3 = K/K_{\max}.$$

Степень же избыточности — общая ($C_{\text{изб}}$) и полезная ($C_{\text{изб}}^n$) — может быть определена по формулам:

$$C_{\text{изб}} = D_{\min}/D;$$

$$C_{\text{изб}}^n = \log_m K_v / \log_m K,$$

где $D_{\min}$ — минимальная длина кода, необходимая для кодирования заданного количества объектов (K);

K_v — возможное количество объектов данного классификатора.

Также может оцениваться долговечность классификатора, которая характеризуется временем ($B_{\text{дл}}$), в течение которого классификатор можно пополнять новыми объектами без изменения длины кода:

$$B_{\text{дл}} = (m^D - K)/\Delta K,$$

где ΔK — количество объектов, дополнительно включаемых в классификатор в течение года.

Оценка различных возможных вариантов построения классификаторов с использованием перечисленных и других их характеристик позволяет определить принципы и выработать методику проектирования кодов, наиболее полно отвечающую требованиям их рациональности.

После выбора системы кодирования осуществляются сбор и подготовка исходных данных, необходимых для непосредственного составления систематизированного свода наименований классификационных группировок и отдельных элементов номенклатур, а также их кодовых обозначений. На данном этапе производятся проектирование кодов, наиболее полно отвечающую требованиям их рациональности, их идентификация и последовательное заполнение позиций кодификатора, который иногда может строиться в нескольких вариантах, отличающихся своими лингвистическими и терминологическими особенностями. Кроме того, кодификаторы, как правило, представляются в эталонной и рабочей форме. Эталонный кодификатор, в котором классификационные группировки располагаются в порядке возрастания их кодовых обозначений,

служит для его последующего ведения и разностороннего использования.

Рабочие кодификаторы составляются в целях их применения для анализа результатов решения конкретных задач управления. Они должны быть согласованы с конкретными результатными документами по форме и содержанию.

Одновременно с непосредственным составлением кодификаторов разрабатывается система их ведения, предназначенная для актуализации кодификаторов, их корректировки в централизованном порядке с накоплением информации об изменениях, а также для оперативного извещения пользователей о внесенных дополнениях и изменениях. При этом проектируются такие средства информационного и программного обеспечения системы ведения кодификаторов, как эталонные, корректирующие, архивные, рабочие и справочные массивы, программы процедур актуализации, корректировки, обработки запросов пользователей, поиска и выдачи соответствующих ответов.

Составленные классификаторы проходят экспериментальную проверку, в результате чего определяется их корректность, приспособленность к решению конкретных задач управления в условиях машинной обработки данных; при этом могут быть выявлены определенные недостатки, которые устраняются путем внесения соответствующих коррективов. После апробирования и корректировки кодификаторы утверждаются и вводятся в действие.

7.5. Автоматизация кодирования

Для повышения эффективности проектирования кодификаторов и их использования при решении конкретных задач управления очень важно создание не только автоматизированных систем ведения кодификаторов (см. 7.3, 7.4), но и специальных программных средств, автоматизирующих сам процесс их построения, подготовки к применению в соответствии с определенными целями пользователей. Можно выделить локальные задачи проектирования и использования кодификаторов, которые могут быть решены с помощью ЭВМ:

- присвоение кодов отдельным элементам и классификационным группировкам номенклатуры;

- расчет контрольных разрядов для кодов обнаружения ошибок; определение резерва кодов по каждой классификационной группировке и всей номенклатуре в целом;

- рационализация структур массивов кодификаторов, хранимых во внешней памяти ЭВМ;

- выдача кодификаторов по запросам пользователей во всех необходимых разрезах, т. е. формирование рабочих кодификаторов для конкретных задач.

Наиболее трудоемким является непосредственное кодирование элементов сложных номенклатур. Кроме того, от успешного решения проблемы построения кодов в значительной степени зависит эффективность обработки данных. Только четкое построение сис-

темы кодов позволяет автоматизировать как сам процесс кодирования, так и поиск нужных элементов в списках номенклатур, используя наиболее рациональные методы с применением математического аппарата.

Рассмотрим возможный вариант автоматизации кодирования и поиска элементов сложных номенклатур, основанный на методике, суть которой сводится к следующему. Любую сложную номенклатуру можно представить как систему многоуровневых ветвей вхождения одних элементов в другие. В такой системе в один элемент более высокого уровня входят несколько элементов следующего, более низкого уровня. Очевидно, что в одной ветви того же уровня, а тем более разных уровней, количество элементов неодинаково. Первая задача состоит в том, чтобы это количество элементов во всех ветвях номенклатуры было одинаковым, т. е. максимально возможным (с учетом возможного роста номенклатуры). Другими словами, в номенклатуру необходимо включить фиктивные элементы, дополняющие все ветви до установленного максимального значения. Используя такую дополненную систему ветвей вхождения одних элементов в другие, можно присвоить им коды, состоящие из трех частей:

K — буквенное обозначение всей номенклатуры (один или два разряда);

i — номер уровня, к которому элемент относится (практически требуется один разряд, так как количество уровней даже в самых сложных номенклатурах не превышает десяти, причем самый низкий уровень — нулевой);

j — порядковый номер элемента внутри уровня (его значение меняется от 0 до $(n_i - 1)$, где n_i — количество элементов в одной ветви, а i может принимать значения от 0 до m , т. е. количество уровней рассматриваемой номенклатуры равно $(m + 1)$).

Таким образом, полный код элемента K_{ij} представляет собой идентификатор переменной длины.

Если порядковый номер j элемента уровня i разделить на постоянное количество элементов в одной ветви n_i , то целая часть частного p покажет порядковый номер элемента более высокого уровня $(i + 1)$, в который данный элемент входит, а увеличенный на единицу остаток от деления r указывает на порядковый номер входящего элемента внутри данной ветви уровня i . Поэтому уже в самой системе порядковых номеров элементов заложены правила вхождения одних элементов в другие и принцип организации автоматического кодирования и поиска элементов в списке номенклатуры.

Блок-схема алгоритма кодирования элементов сложных номенклатур по рассматриваемой методике представлена на рис. 7.1. В алгоритме предусмотрен порционный ввод записей массива упорядоченной номенклатуры (по возрастанию признаков R_i всех уровней от m до 1). Каждая запись такого массива представляет собой полный список значений всех классификационных признаков отдельного элемента нулевого уровня номенклатуры. В результате

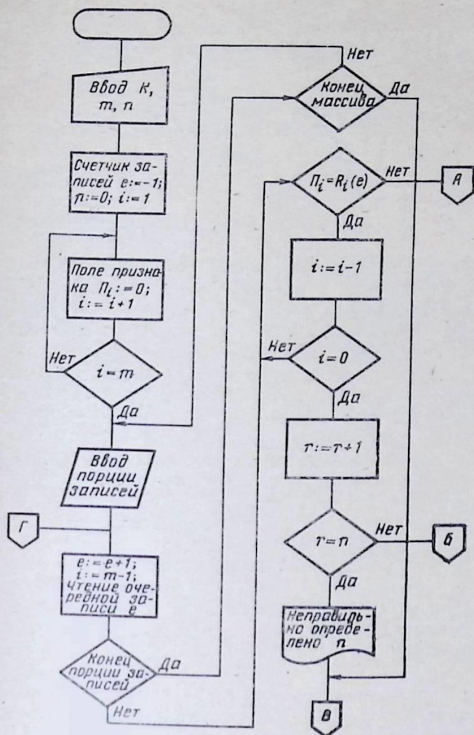


Рис. 7.1. Блок-схема алгоритма автоматического кодирования (начало)

обработки всех записей исходного массива они дополняются кодовыми обозначениями K_{ij} , и составляется кодификатор, также формируемый в виде массива и записываемый на внешний носитель (магнитный диск, ленту и т. п.). В алгоритме предусмотрен контроль правильности определения значения n для конкретной номенклатуры и расчета порядкового номера p . Если в результате проверки устанавливается невозможность продолжения автоматического кодирования, об этом сообщается оператору ЭВМ.

Присвоенные по представленному алгоритму коды элементов сложных номенклатур позволяют путем простого расчета осуществлять переход из одного уровня номенклатуры в другой:

$$Kij \rightarrow K(i-1)(jn) \leftarrow K(i-1)(jn+n-1);$$

$$Kij \leftarrow K(i+1)p, (n-1)\text{-й в ветви.}$$

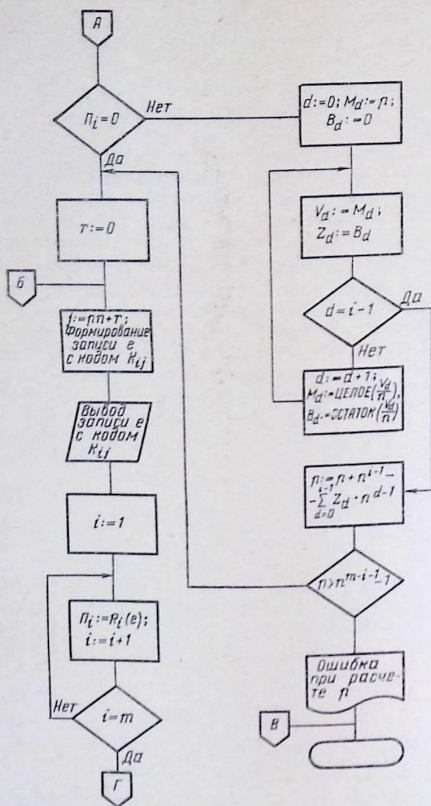


Рис. 7.1. Блок-схема алгоритма автоматического кодирования (конец)

Так, например, для условной номенклатуры М с параметрами $n=25$ и $m=6$ по коду элемента М4164 легко определяем, что в данный элемент входят элементы более низкого уровня, которым присвоены коды М34100, М34101, ..., М34124; в свою очередь, элемент с кодом М2102504 входит пятым по порядку в элемент более высокого уровня, имеющий кодовое обозначение М34100.

Создание системы задач автоматизации построения кодов и кодификаторов обеспечивает высокий уровень автоматизации про-

ектирования информационного обеспечения машинной обработки данных в целом и позволяет эффективно использовать продукт проектирования при решении задач управления.

Глава 8

СИСТЕМЫ ОПИСАНИЯ ДАННЫХ

8.1. Идентификация данных и языки их описания

Рассмотренные в гл. 7 коды экономической информации входят в различные образования информационного фонда (элементы, сегменты, записи, файлы, документы и др.) и непосредственно участвуют в машинной обработке в широком смысле этого слова. Иначе говоря, они применяются для организации системы данных *на физическом уровне*, для физического представления данных, т. е. как они записываются в памяти ЭВМ и на машинных носителях информации. Представление этих же данных в их логических отношениях для нужд пользователей, проектировщиков и прикладных программистов, разрабатывающих средства программного обеспечения обработки данных, возможно при их соответствующем описании *на логическом уровне*. Для логического представления данных, описания структур различных образований информационного фонда используются их *идентификаторы*. Целью идентификации является упрощение описания данных и их отношений, введение однозначного толкования всех структурных элементов данных, входящих в информационную систему, обеспечение возможности использования идентификаторов при построении алгоритмов и программ обработки в качестве имен данных, а также в реализации диалога между пользователем и ЭВМ.

Для идентификации данных могут использоваться различные средства условного обозначения их элементов, составных частей и структур, однако чаще всего при построении систем машинной обработки разрабатываются специальные *языки описания данных* (ЯОД), в общем случае включающие в себя собственный алфавит и правила построения языковых конструкций с использованием данного алфавита. Разработка языка описания данных для АСУ экономическими объектами осуществляется, как правило, с ориентировкой на основную единицу экономической информации — показатель. Для идентификации показателей обычно применяются простые информационные языки с позиционным синтаксисом. При этом различаются тезаурусные языки, а также языки с неполной или развитой грамматикой.

Для создания наиболее простого тезаурусного языка прежде всего необходимо точно определить значение всех употребляемых для выражения понятий показателей слов и составить их полный список. Затем следует произвести классификацию этих слов, рас-

крывающую существующие объективные связи между ними, выделить их основные представители — дескрипторы, т. е. слова и словосочетания с их условными сокращенными обозначениями, которые чаще всего снабжаются определенными индексами, отражающими линейные, иерархические, родо-видовые и другие связи дескрипторов. Идентификаторы реквизитов и показателей строятся с использованием тезаурусных обозначений дескрипторов, которые могут располагаться в любой последовательности или с учетом грамматики языка, например по установленному шаблону.

Языки с неполной грамматикой разрабатываются для описания показателей несложной структуры, когда имеется возможность установить жесткие правила их записи. В этом случае на основе анализа понятий, выражающих экономические показатели системы данных, строится универсальная конструкция высказывания, состоящая из определенного набора признаков позиций (фасетов). Для этого составляется полный список слов конструкции, который затем разбивается на фасетные списки, каждый из которых является словарем для определенной позиции конструкции. Позиция слова в конструкции определяет его роль в высказывании. Номер определенной позиции конструкции совпадает с номером соответствующего фасетного списка, а отдельным словам таких списков также присваиваются номера. В идентификатор показателя включаются пары номеров: номер позиции и номер слова в списке, соответствующем данной позиции. Вместо цифровых номеров здесь

Таблица 8.1

Фрагменты полного списка слов,
используемых для построения ЯОД
с неполной грамматикой

Фасетные позиции (списки)	Номера слов	Слова (лексическое накопление)	Фасетные позиции (списки)	Номера слов	Слова (лексическое накопление)
П1	01	директивный	П3	02	предприятие
	02	плановый		03	цех
	03	учетный		04	участок
П2	04	фактический	П4	15	работник
	09	количество		61	заработная плата
	10	сумма		82	деталь
	11	время		04	год
	17	коэффициент		05	полугодие
	43	средний		06	квартал
	59	относительный		07	месяц

могут использоваться и алфавитно-цифровые обозначения, тем более, что идентификаторы, как правило, начинаются с буквы. Так, например, пользуясь фрагментами полного списка слов, представленными в табл. 8.1, можно построить следующие идентификаторы конкретных показателей:

П102П209П243П302П315П404 — плановая среднесписочная численность (количество) работников по предприятию на год;

П104П361П315П407 — фактическая сумма месячной заработной платы работника;

П102П209П382П304П406 — плановый квартальный выпуск (количество) деталей по участку и т. п.

Языки с развитой грамматикой предусматривают точное отображение смысловых отношений между словами, выражающими показатели всех видов. Для построения такого языка анализируются полные записи всех показателей системы данных, в результате чего составляются список основ слов (без окончаний) показателей и грамматическая схема, общая для всех показателей. В укрупненном виде грамматическую схему показателя можно представить так:

ОпПоОс — ОпДоОс — ОпОбОс,

где Оп — определение, По — подлежащее, До — дополнение, Об — обстоятельство, Ос — определение, относящееся к предшествующему существительному.

Такая схема детализируется в зависимости от грамматических функций членов предложения и категории слов, в результате чего составляется грамматико-семантический шаблон записи показателя. В данном случае в идентификатор любого показателя включаются номера (обозначения) позиций в шаблоне, построенном по детальной грамматической схеме, и номера (обозначения) слов, указанные в списке основ показателей. Идентификацию показателей с использованием такого языка можно производить автоматически с помощью ЭВМ по специальной программе, построенной на базе универсальной модели порождения показателя.

Идентификаторы показателей и их отдельных реквизитов используются в схеме их взаимосвязи, информационных и математических моделях отдельных задач управления, алгоритмах решения этих задач, а иногда и непосредственно в исходных программных модулях.

8.2. Системная идентификация реквизитов и показателей

Системный подход к проектированию языка описания данных, главным образом служащего для идентификации показателей, заключается в полном охвате всех показателей, используемых в данной системе, применении единых классификационных принципов систематизации их реквизитов, использовании единых условных обозначений реквизитов для идентификации всех показателей системы данных, а также общих для всех показателей правил расположения этих обозначений.

Такой подход обеспечивает единство идентификации, возможность определения логических отношений между показателями, расширения системы идентификации и создает предпосылки для автоматизации ее построения и использования в различных целях.

В процессе анализа данных прежде всего следует выделить директивные, плановые, нормативные, учетные, статистические и

аналитические показатели и определить перечень их реквизитов-оснований. Затем эти показатели отдельно группируются с учетом семантических свойств (например, выделяются группы оснований количественных, суммовых, временных, относительных, условных показателей). В каждой группе производится дальнейшее упорядочение реквизитов-оснований по сходству характеризуемых объектов и явлений. Так, например, среди оснований количественных показателей могут быть выделены такие подгруппы: количество в движении (приход, расход, остаток), количество во времени (на начало, на конец периода, на (за) период), общее и единичное количество (выработка всех рабочих, выработка одного рабочего, средняя выработка одного рабочего, среднесписочная численность работающих и т. д.). Наконец, упорядоченные групповые списки реквизитов-оснований в результате их синтеза объединяются в единый перечень оснований показателей, на базе которого производится их идентификация. Условные обозначения каждой позиции такого перечня обычно строятся с использованием индексируемых прописных букв, присваиваемых определенной группе реквизитов-оснований, а иногда выражаются отдельными буквами русского, латинского и греческого алфавитов (при небольшой номенклатуре показателей).

Таблица 8.2

Идентификация реквизитов-оснований показателей

Реквизиты-основания	Обозначения (варианты)		
	1	2	3
Приход (поступление) в количественном выражении	A	K ¹	K1
Расход (отпуск, передача) в количественном выражении	B	K ²	K2
Остаток на начало периода	B	K ³	K3
Остаток на конец периода	Г	K ⁴	K4
Среднесписочная численность основных рабочих	M	K ⁸	K8
Фонд заработной платы	N	C ⁵	C5
Трудоемкость производственной программы	T	T ⁴	T4
Фонд рабочего времени одного рабочего	S	T ¹⁹	T19
Коэффициент выполнения норм выработки	K	P ³	P3

Индексируемые идентификаторы реквизитов — оснований чаще всего имеют верхние цифровые индексы, однако при автоматизированном их построении с помощью ЭВМ, а также в целях их возможного использования при составлении исходных программных модулей все символы идентификаторов следует располагать линейно, на одном уровне. В данном случае используются только прописные буквы, причем в зависимости от применяемых ЭВМ и систем программирования только латинские или одновременно

русские и латинские. Кроме того, идентификаторы оснований при их включении в идентификаторы показателей дополняются условным обозначением, указывающим их принадлежность к директивному, плановому, нормативному, учетному, статистическому или аналитическому показателю. В табл. 8.2 представлены фрагменты перечня реквизитов-оснований с тремя рассмотренными нами возможными вариантами их обозначений.

Для построения идентификаторов реквизитов-признаков показателей следует провести их системный анализ с использованием структурной схемы экономических процессов. При этом выделяются независимые группы однотипных признаков, такие, например, как уровень управления, период и контур управления, характеристика работающих, вспомогательные признаки и т. д. Более детальная систематизация реквизитов-признаков внутри каждой группы приводит к построению идентифицируемых подгрупп, признакам которых могут быть присвоены как уникальные буквенные обозначения, так и индексируемые обозначения (по аналогии с идентификацией реквизитов-оснований). При этом составляется общий список подгрупп реквизитов-признаков и полные упорядоченные перечни признаков всех подгрупп. В табл. 8.3 и 8.4 представлены примеры такого списка и некоторых перечней с возможными вариантами обозначения реквизитов.

Таблица 8.3

Идентификация подгрупп реквизитов-признаков

Группы и подгруппы признаков	Обозначения (варианты)		
	1	2	3
Уровень управления:			
вертикальный	a—ж	v	V
горизонтальный	з—п	w	W
Период управления (интервал времени)	р—я	h	H
Контур управления:			
труд и его оплата	— — — a—п	a	A
средства труда	— — — р—я	z	Z
объект труда	i—s	m	M
продукт труда:			
структура изделия	a'—к'	e	E
планируемая и производимая продукция	л'—ф'	s	S
процесс производства	α—λ	g	G
брак и технологические отходы	σ—Ω	b	B
денежные средства	— — — i—s	c	C
Характеристика работающих	a—п	r	R
Вспомогательные признаки	— — — с—я	u	U

В идентификатор конкретного показателя вслед за обозначением его основания и разреза включаются условные обозначения всех признаков, как правило, располагаемых в порядке их стар-

Идентификация некоторых реквизитов-признаков

Признаки	Обозначения (варианты)		
	1	2	3
1	2	3	4
Вертикальный уровень управления:			
министерство	а	V ₁	V1
объединение	б	V ₂	V2
предприятие	в	V ₃	V3
подразделение (цех, отдел, склад)	г	V ₄	V4
участок	д	V ₅	V5
бригада	е	V ₆	V6
рабочее место	ж	V ₇	V7
Горизонтальный уровень управления:			
исполнитель	з	W ₁	W1
производитель	и	W ₂	W2
поставщик	к	W ₃	W3
держатель фондов материалов	л	W ₄	W4
получатель	м	W ₅	W5
потребитель (заказчик)	н	W ₆	W6
держатель фондов готовой продукции	о	W ₇	W7
Период управления (интервал времени):			
пятилетка	Р	h ₁	H1
год	с	h ₂	H2
полугодие	т	h ₃	H3
квартал	у	h ₄	H4
месяц	ф	h ₅	H5
декада (неделя, пятидневка)	х	h ₆	H6
сутки (день)	ц	h ₇	H7
смена	ч	h ₈	H8
час	ш	h ₉	H9
Труд и его оплата:			
работающий (табельный номер)	а	a ₁	A1
категория труда	б	a ₂	A2
категория работающего	в	a ₃	A3
профессия	г	a ₄	A4
разряд	д	a ₅	A5
форма оплаты	е	a ₆	A6
тарифная ставка	ж	a ₇	A7
вид оплаты	з	a ₈	A8
вид удержания	и	a ₉	A9
группа налогоплательщиков	к	a ₁₀	A10
причина доплаты	л	a ₁₁	A11
виновник доплаты	м	a ₁₂	A12
причина простоя	н	a ₁₃	A13
виновник простоя	о	a ₁₄	A14
вид пособия	п	a ₁₅	A15
вид заболевания	р	a ₁₆	A16

шинства. Приведем несколько примеров обозначения показателей с использованием индексируемых идентификаторов (согласно табл. 8.2 — 8.4):

$K_{h_1 v_1 a_1 a_2 a_3}^{8p}$ или K8PH4V4A4A5A6

— плановая численность основных рабочих на квартал в разрезе цехов по профессиям, разрядам и формам оплаты труда;

$T_{h_2 v_1 a_1}^{19p}$ или T19PH2V4A4

— плановый годовой фонд рабочего времени одного рабочего по цеху и профессии;

$C_{h_5 v_5}^{5f}$ или C5FH5V3

— фактический месячный фонд заработной платы по предприятию.

8.3. Описание данных в условиях применения единых информационных фондов и баз данных

Единые информационные фонды и базы данных к языкам описания данных предъявляют особые требования. Для описания логической (а иногда и физической) организации данных используются языки, служащие не столько для идентификации реквизитов, показателей, записей, файлов и т. п., сколько для объявления используемых структур данных и отношений между различными типами образований информационного фонда. В таких языках, как правило, дополнительно предусматривается возможность описания длины элементов данных, диапазона их допустимых значений, повторяемости вхождения элементов, измерения файлов и их размеров, порядка следования информационных образований, обеспечения секретности данных и т. д. Чаще всего такие языки ориентированы или на программистов, или на пользователей. Они применяются в системах управления базами данных.

При разработке языков описания данных, ориентированных на программистов, прежде всего необходимо исходить из их независимости от конкретных ЭВМ и необходимости в установлении их связи с наиболее распространенными языками программирования. Поэтому такие языки часто строятся на базе используемых систем программирования (Кобол, Фортран, PL/I), что позволяет в стандартизованном виде описывать структуры данных, которые, с одной стороны, не зависят от способов физического хранения и организации данных и, с другой — соответствуют всем ограничениям, установленным для определенного языка программирования. Таким примером может служить язык описания данных CODASYL [20], построенный на базе раздела данных Кобола, который был разработан американской ассоциацией по языкам систем обработки данных.

При построении языка описания данных, ориентированного на пользователей, должны учитываться способы представления ин-

формации в файлах базы данных. Как известно, данные могут представляться в виде деревьев, сетевых и списковых структур а также двумерных таблиц. При этом следует обратить внимание на то, что нетабличная структура базы данных, которой присущи определенные ограничения, в общем случае препятствует ее росту и изменениям, а в конечном счете рост такой базы может привести к нарушению логического представления данных и к непригодности разработанных прикладных программ. Поэтому наиболее рациональными и перспективными представляются реляционные базы данных, построенные из табличных файлов элементов данных, в которых отражаются отношения между данными в самой простой и естественной форме. Любое другое представление данных может быть с некоторой избыточностью сведено к плоским двумерным (табличным) файлам. Процесс перехода к табличному представлению для каждого отношения между данными в базе называется *нормализацией*. Для описания таблиц, наборов их элементов, а также операций над ними разрабатываются специальные математические обозначения. Для таких баз следует использовать и специальный язык манипулирования данными (например, язык Кодда), в котором может быть выражен диалог пользователя с ЭВМ. При этом обязательно предусматриваются операции разрезания и склеивания табличных файлов, что позволяет для одних пользователей создавать таблицы меньшей размерности, а для других — объединять таблицы, в результате чего устанавливаются новые отношения между данными. Таким образом, каждый пользователь имеет возможность рассматривать базы данных как переменный во времени набор нормализованных отношений данных.

Также необходимо отметить, что для описания логической организации единых информационных фондов и баз данных обычно используются два специальных языка: для общего логического представления данных и для представления данных на уровне отдельных задач управления, реализуемых определенными прикладными программами ЭВМ. Система управления базами данных использует общие логические описания данных для получения логических записей из физических. Из этих логических записей по описаниям данных на уровне отдельных задач формируются записи, запрашиваемые конкретными прикладными программами.

При построении языков описания данных, а также языков манипулирования данными, включенными в единый информационный фонд, также используются соответствующие грамматики, которые можно разделить на четыре класса (согласно классификации американского лингвиста Н. Хомского):

- грамматики естественных языков;

- контекстно-зависимые грамматики;

- контекстно-независимые грамматики;

- грамматики с конечным числом состояний автоматного языка.

В настоящее время наиболее распространены языки, порожаемые грамматиками двух последних классов, обладающих наи-

более простым разбором (синтаксическим анализом) конструкций языка, хотя они не всегда обеспечивают эффективное управление базами данных при минимальных затратах на их эксплуатацию (к ним относятся затраты на изучение языка, на составление обращений к базам данных, на их трансляцию). Если для программистов, как уже отмечалось, очень важную роль играет богатство изобразительных средств языка, обеспечивающее использование всех заложенных в базе данных возможностей, то для пользователей системы более существенным выступает высокий уровень языка, обеспечивающий лаконичность описания обращений для типовых условий использования базы данных. Логическая структура грамматики такого языка обычно максимально соответствует системе понятий предметной области, на которую ориентирована база данных.

Накапливаемый опыт внедрения и эксплуатации баз данных в нашей стране, дальнейшие научные исследования в области управления базами данных будут способствовать разработке единых (универсальных) языков описания данных, удовлетворяющих запросам и программистов, и пользователей практически вне зависимости от особенностей предметной области, для которой создается база данных.

Глава 9

ФОРМЫ И МАКЕТЫ РАЗМЕЩЕНИЯ ДАННЫХ

9.1. Носители данных, их классификация и основные характеристики

Данные при их движении и преобразовании всегда размещаются на определенных носителях. Каждый носитель данных — это материальный объект, содержащий сведения, доступные или только для человека, или только для ЭВМ, или, наконец, одновременно и для человека и для машины. В связи с этим все носители мы можем подразделить на человекочитаемые, машиночитаемые (машинные) и комбинированные.

Материальная основа носителей (их физическое строение) предопределяет размещение данных на них, средства их подготовки, роль и место использования, методы организации данных. Поэтому ее можно считать одним из основных классификационных признаков носителей данных, по которому в условиях современных систем машинной обработки они подразделяются на бумажные, магнитные, фотохромные, индикационные (экранные) и жесткие. Носители данных могут отличаться по своему конструктивному исполнению, их роли в различных этапах процессов движения и преобразования данных, способу записи данных на них и т. п. Классификации, которые наиболее существенны при орга-

низации машинной обработки данных, отражены на схеме, представленной на рис. 9.1. Рассмотрим более подробно каждый тип носителей в отдельности.

Бумажные носители являются наиболее простыми, традиционными, но среди них можно выделить совершенно разные группы.

В первую очередь следует выделить первичные документы и табуляграммы (носители результатных данных). Последние составляются на машинах и иногда называются машинограммами. По конструктивному исполнению различаются бланочные, рулонные, карточные и ленточные бумажные носители. По способу записи они могут быть рукописными, печатными и перфорируемыми (перфорационными), а по способу формирования — ручной и автоматической подготовки.

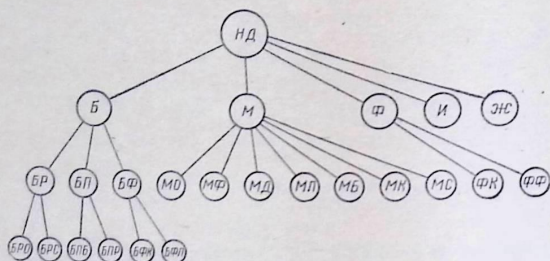


Рис. 9.1. Классификация носителей данных:

НД — носители данных; Б — бумажные; БР — рукописные (БРО — обыкновенные, БРС — специализированные); БП — печатные (БПБ — бланочные, БПР — рулонные); БФ — перфорационные (БФК — карточные, БФЛ — ленточные); М — магнитные (МО — оперативные, МФ — буферные, МД — дисковые, МЛ — ленточные, МБ — барабанные, МК — карточные, МС — сферические); Ф — фотохромные (ФК — кассетные, ФФ — карточные); И — индикационные (экранные); Ж — жетонные

К носителям ручной подготовки прежде всего относятся рукописные (обыкновенные или специализированные) документы, предназначенные для регистрации первичных данных, отражающих определенные количественные и качественные характеристики объекта управления, результаты выполнения хозяйственных и других операций, и имеющие юридическое значение. В отличие от обыкновенных специализированные документы, заполняемые вручную с использованием специальных шрифтов, графических меток и т. п., являются комбинированными носителями данных, т. е. информация с них в ЭВМ вводится непосредственно с помощью специальных устройств, без подготовки промежуточных машинных носителей, которые необходимы при использовании обыкновенных документов, заполняемых чернилами или шариковой ручкой и читаемых только человеком.

К печатным носителям (по своей форме они могут быть бланчными или рулонными) относятся первичные документы с исходными данными (они могут заполняться на пишущих и счетно-табличных машинах или типографским способом) и выводимые табуляграммы с результатами обработки данных на ПВМ или ЭВМ. Результаты могут выводиться на печать не только в виде цифровых, алфавитных или алфавитно-цифровых данных, но и как различных графический материал (схемы, графы, чертежи, рисунки и т. д.). Следует отметить, что параллельно с получением печатных носителей как исходных, так и результатных данных часто подготавливаются машинные носители, предназначенные для автоматического ввода данных в ЭВМ при последующей их обработке.

К перфорационным носителям данных относятся перфокарты и перфоленты, которые наиболее часто используются в виде промежуточных носителей. На них вручную, путем набора на клавиатуре переносятся данные, заранее зафиксированные в первичных документах (или — для перфолент — одновременно с заполнением документов). Перфокарты и перфоленты также могут использоваться как носители результатных данных, т. е. на них могут автоматически выводиться результаты, получаемые при обработке данных на вычислительных машинах.

В целом бумажные носители данных и в условиях АСУ продолжают оставаться основными носителями информации. Прежде всего это относится к первичным документам, отличающимся большим разнообразием по форме, назначению, способам размещения и заполнения данных, их составу, месту возникновения и другим признакам. По этим же признакам можно классифицировать и табуляграммы, которые в основном предназначаются для использования в виде носителей результатных данных (основные табуляграммы). Также составляют табуляграммы, имеющие технологическое назначение: контрольные (для проверки правильности ввода и выполнения операций обработки данных, тестирования программ и т. п.) и вспомогательные (описи хранимых и корректируемых массивов данных, распечатки программ и др.).

Широкое применение перфокарт для обработки экономических данных связано с их пригодностью для постепенного накапливания данных, их хранения в течение длительного периода времени; данные на перфокартах можно многократно использовать, сравнительно легко вносить в них изменения и дополнения, осуществлять объединение, выборку и группировку массивов на специальных машинах, а ввод данных в ЭВМ производить в любой желаемой последовательности. Вместе с тем перфокарты имеют ограниченную информационную емкость, обладают необратимостью записи данных и небольшой скоростью ввода в ЭВМ, они легко деформируются в процессе обработки.

Перфоленты имеют практически неограниченную информационную емкость, обладают значительно большей скоростью ввода по сравнению с картами, но вместе с тем им свойственна низкая физическая прочность, невозможность предварительного упорядоче-

ния данных до ввода в ЭВМ и сложность внесения изменений. Широкое использование перфолент связано с применением средств дистанционной передачи данных на расстояние и специальных приставок подготовки, агрегируемых с пишущими и другими машинами (устройствами).

Основными машинными носителями данных, обрабатываемых на ЭВМ, в современных АСУ и АСОД являются магнитные носители, предназначенные для фиксации данных, закодированных в двоичной системе счисления. Среди них, прежде всего, следует выделить три группы, относящиеся к различным видам машинной памяти: основной (оперативное запоминающее устройство), местной (автономное буферное запоминающее устройство) и внешней (накопители на магнитных дисках, лентах, барабанах, картах, «пузырях»).

Данные, находящиеся в основной (оперативной) памяти ЭВМ, подвергаются непосредственной обработке. Они туда вводятся и выводятся по необходимости использования в процессе решения определенных задач. Оперативная память ЭВМ третьего поколения построена на интегральных схемах (поэтому в данном смысле она чисто условно отнесена к магнитным носителям) и имеет байтовую структуру, т. е. запись и считывание информации могут производиться побайтно (по 8 бит) или словами и блоками, объединяющими группы байт. Оперативная память характеризуется высоким быстродействием и надежностью, но имеет ограниченный объем (от десятков до сотен тысяч, а в наиболее мощных ЭВМ — несколько миллионов байт).

Некоторые периферийные устройства имеют собственные запоминающие блоки, получившие название буферной памяти. По технико-эксплуатационным характеристикам и принципу организации процедуры ввода данных буферную память терминальных устройств можно считать машинным носителем данных, используемым для ввода их в ЭВМ, наравне с перфокартами и перфолентами, так как в этом случае данные вводятся путем последовательного ручного их набора на клавиатуре устройства по соответствующему макету (по аналогии с макетом перфорации).

В качестве устройств внешней памяти ЭВМ преимущественно используются накопители на магнитных дисках и лентах, несколько реже — барабаны и карточные устройства. Магнитные диски широко применяются благодаря их высоким техническим характеристикам и прежде всего возможности прямого (произвольного) доступа к данным, располагаемым на концентрических дорожках дисков. В состав Единой системы ЭВМ входят различные накопители на стационарных или сменных дисках, покрытых магнитным сплавом, которые образуют пакеты определенной емкости (объем памяти одного пакета колеблется в пределах от 1 до 30 млн. байт). Запись и считывание информации производятся посредством электромагнитных преобразователей — аэродинамических головок, плавающих на воздушной подушке на расстоянии нескольких микронов от рабочей поверхности диска. Данные на-

носятся в виде блоков, которые могут занимать от одного до нескольких тысяч байт.

В современных системах машинной обработки данных основными носителями последовательного типа являются накопители на магнитных лентах. Высокое быстродействие, большая информационная емкость, достоверность хранения данных отличают внешнюю память на магнитных лентах от других подобных устройств, но вместе с тем следует отметить, что при работе с лентами нет возможности произвольно выбирать данные, записанные на ленте, а также оперировать несколькими файлами одновременно за один прогон ленты. Стандартная магнитная лента шириной 12,7 мм, применяемая в устройствах памяти ЭВМ Единой системы, имеет 9 дорожек, из которых 8 предназначены для записи данных, а одна дорожка служит для контроля записываемых кодов по четности единичных разрядов. Информация записывается блоками (зонами) с плотностью записи по 8 или 32 бита на 1 мм. Стандартная длина одной ленты — 750 м.

Если раньше основная часть исходных данных, подлежащих вводу в ЭВМ, предварительно переносилась на магнитные диски и ленты с перфокарт или перфолент, то в современных системах обработки данных все шире используются устройства, предназначенные для непосредственной магнитной записи данных на дисках и лентах, что позволяет сократить затраты на подготовку машинных носителей и ввод данных, обеспечивает более простую технологию обработки данных. Однако главным образом внешние запоминающие устройства данного типа используются для вывода промежуточных и результатных данных, для хранения файлов больших объемов, многократно используемых в решении задач управления экономическим объектом. Они применяются для организации единых информационных фондов и баз данных, а также служат для хранения средств программного обеспечения ЭВМ и АСУ.

Магнитные барабаны, имеющие круговые дорожки для записи данных, относятся к запоминающим устройствам прямого доступа. Емкость накопителей на барабанах диаметром 45 см, имеющих 800 рабочих дорожек, составляет 6 млн. байт. Более широкое применение магнитных барабанов сдерживается главным образом невозможностью их быстрой замены и вследствие этого относительно небольшим объемом доступной хранимой информации, а также сравнительно высокой стоимостью.

В качестве машинных носителей данных также могут использоваться магнитные карты, обладающие по сравнению с перфокартами значительно большей плотностью записи (до 100 бит на 1 см) и высокой скоростью считывания данных, примерно в 40 раз превышающей скорость ввода с перфокарт. Вместе с тем применение магнитных карт связано с определенными трудностями их подготовки и хранения, а также значительными затратами из-за их сравнительно высокой стоимости.

В последние годы за рубежом, в частности в США, стали при-

меняться сферические магнитные носители данных, условно называемые «пузырями», относящиеся к внешним запоминающим устройствам прямого доступа, которые по сравнению с магнитными дисками обладают значительно большей плотностью записи, более высоким быстродействием и исключительно высокой надежностью из-за отсутствия механических элементов.

Для хранения различной документальной информации все шире используются фотохромные носители данных, представляющие собой кассетные микрофильмы и карточные микрофиши. Информационная емкость микрофильма шириной 3,5 см и длиной 1 м составляет, например, около 20 млн. байт. Запоминающие устройства, использующие микрофильмы и микрофиши, особенно эффективно могут применяться в информационно-поисковых системах для хранения технологической и проектно-конструкторской документации, нормативно-технических данных, стандартов и т. д.

В системах обработки данных коллективного пользования в условиях телеобработки информации все шире используются видеотерминальные устройства (дисплеи), предназначенные для оперативного ввода и вывода данных. Носителем данных становится экран электронно-лучевой трубки, на котором индицируется как алфавитно-цифровая, так и графическая информация. Система представления знаков дисплея основывается на телевизионном растре. Информационная емкость видеотерминальных устройств, входящих в состав ЭВМ Единой системы, колеблется в пределах от 256 до 2048 знаков. Дисплеи, непосредственно подключаемые к ЭВМ или входящие в состав абонентских пунктов пользователей, обеспечивают диалоговый режим обработки данных.

Иногда для оперативной фиксации определенных постоянных данных удобно пользоваться жетонными носителями (пластмассовыми, металлическими и т. п.) с пробивками, которые считываются специальными устройствами. Примером таких носителей могут быть пропуска работающих, используемые в автоматизированных системах табельного учета, оснащенных комплексом технических средств автоматической проходной предприятия. Информация с жетонных носителей чаще всего переносится на перфоленты или другие промежуточные носители для последующего ввода в ЭВМ.

9.2. Принципы размещения данных

Для размещения данных первостепенное значение имеет выбор типа носителя, отвечающего противоречивым требованиям максимального использования его информационного объема и минимизации времени поиска (выборки) данных, а также стоимости подготовки и хранения информации.

Проектирование размещения данных на носителях вне зависимости от их типа связано с определением содержания и формы документов, с разработкой физической структуры массивов (файлов) данных и макетов, указывающих порядок размещения и некоторые характеристики элементов записей (обозначения, длину,

границы изменения и др.). При этом прежде всего определяется общий состав исходных показателей, используемых в решении задач управления, содержание выходных форм результатной информации и на этой основе устанавливаются перечень и содержание массивов данных, включаемых в единый информационный фонд или используемых для отдельных комплексов решаемых задач. Для каждого вида данных, характеризующих или определенный объект управления (материальные и трудовые ресурсы, оборудование, готовая продукция и т. д.), или соответствующую функцию управления (планирование, учет и др.), определяются формы и макеты представления данных на всех этапах процесса движения и обработки (первичном, подготовительном, основном и заключительном) по каждому типу носителя. Таким образом, в согласованном порядке проектируются (или выбираются) формы первичных документов, макеты машинных носителей для первичных, промежуточных и результатных данных, а также для базовых файлов (массивов единого информационного фонда), формы вывода результатной информации для конкретных пользователей. Основными целями при проектировании и выборе носителей данных являются: минимизация трудовых и стоимостных затрат на подготовку данных и контроль; минимизация объема памяти и стоимости хранения данных; минимизация времени поиска необходимой информации; обеспечение требуемой достоверности данных, простоты использования и возможностей постоянного обновления и наращивания массивов. По перечисленным критериям производится оценка сравниваемых типов носителей. Важным этапом проектирования размещения данных является определение структурного состава формируемых записей на элементарном уровне, т. е. на уровне реквизитов. При этом необходимо исходить из их роли в организации процесса обработки данных. Прежде всего следует выделить специальные реквизиты (такие, как идентификаторы массивов, пользователей, порядковые номера носителей, пачек документов, типы записей и т. д.), справочные и группировочные признаки, количественно-суммовые реквизиты, текстовую информацию. На носителе согласно общим принципам последовательности размещения вначале располагаются специальные и справочно-группировочные реквизиты, постоянные для всех записей массива и определенной группы записей (например, для пачки документов и каждого отдельного документа), затем следуют переменные реквизиты — справочные и группировочные признаки и, наконец, количественно-суммовые основания показателей. При этом последовательность данных на машинном носителе строго соответствует их размещению в первичном документе или получаемом информационном сообщении. На машинный носитель не переносятся данные, имеющиеся в массивах условно-постоянной информации, если они по тем или иным причинам и содержатся в соответствующих первичных документах (это также касается текстовой информации).

Записи одного массива могут быть фиксированного (F), переменного (V) или неопределенного (U) формата в зависимости от

их длины. В формате F все логические записи и их блоки (физические записи) имеют постоянную (фиксированную) длину. Формат предполагает, что переменная длина каждой записи (как логической, так и физической) указывается в начале самой записи. Разновидностью данного формата являются так называемые сегментированные массивы, в которых несколько физических записей представляют собой одну логическую запись. Записи формата L могут иметь переменную, заранее не известную длину и поэтому отделяются друг от друга специальным служебным символом (разделителем записей), который не может встречаться внутри записей. Выбор формата записей зависит от типа носителя и принятой формы представления значений реквизитов: по номинальной или фактической значности. При записи данных по номинальной форме каждому реквизиту отводится постоянное количество разрядов в соответствии со значностью разработанных кодов признаков и наиболее часто встречающейся или максимально возможной значностью оснований показателей. При фактической форме записи отбрасываются все незначащие нули слева как для реквизитов-признаков, так и для оснований. В данном случае, как правило, имеем дело с массивами форматов V или U . При этом внутри каждой записи реквизиты также отделяются друг от друга собственным специальным служебным символом (разделителем элементов записи).

При проектировании размещения данных на носителе необходимо предусмотреть возможность их автоматического контроля. Чаще всего это связано с включением в макеты размещения дополнительных данных (например, контрольных сумм по каждой логической и физической записи, в целом по массиву и др.). Порядок их получения и размещения на носителе зависит от принятых методов обеспечения достоверности данных, их контроля.

Проектирование размещения данных, осуществляемое с учетом рассмотренных общих принципов и требований, должно опираться также на специфику соответствующего носителя и используемых технических средств. Кроме того, свои требования к размещению данных в основной, местной и внешней памяти ЭВМ предъявляют средства системного и прикладного программного обеспечения, которые применяются в конкретной системе машинной обработки данных.

9.3. Методика проектирования форм первичных документов и табуляграмм

Проектирование первичных документов связано с решением трех основных вопросов: выбором необходимых форм единой унифицированной системы документации, построением уникальных форм документов в соответствии с требованиями государственных стандарта или эксплуатационными характеристиками используемых технических средств, организацией рационального документооборота.

Создание унифицированных форм документов представляет собой комплексную проблему и основано на следующих принципах:

1) сокращении количества форм первичных документов с целью построения рационального документооборота и упрощения их обработки на ЭВМ;

2) использовании единых форм документов на различных уровнях управления (министерство, объединение, предприятие);

3) использовании минимального количества данных, вводимых для решения функциональных задач, а также исключении из документов расчетных показателей, которые могут быть получены при машинной обработке, и данных, имеющихся в системе;

4) однократном вводе данных в ЭВМ и многократном их использовании;

5) единой терминологии для всей системы документации, обеспечивающей простоту и удобство при работе пользователя с документами.

Унифицированные формы документов утверждаются или регистрируются Госстандартом СССР. Они разрабатываются и утверждаются ЦСУ СССР, Министерством финансов СССР, другими министерствами и ведомствами на основе типовых форм. Каждой утвержденной форме документации присваивается номер по Общесоюзному классификатору управленческой документации, а реквизиты, включаемые в документ, должны кодироваться в соответствии с общесоюзными или локальными классификаторами технико-экономической информации.

Выбор унифицированных форм первичных документов при проектировании систем машинной обработки данных основывается, прежде всего, на определении состава реквизитов, который необходим для каждой функциональной задачи системы управления с учетом их информационной взаимосвязи, а также на выполнении требований, предъявляемых в связи с применением конкретных средств вычислительной техники.

Несмотря на большое количество разнообразных унифицированных форм первичных документов, для конкретных систем обработки данных разрабатываются уникальные (собственные) формы внутренней документации, удовлетворяющие определенным специфическим условиям пользователей. Проектирование таких форм осуществляется строго в соответствии с требованиями стандартов.

После определения состава реквизитов конкретной формы проектируемого документа производится их размещение на бланке, разбиваемом на определенные зоны с учетом их группировки, логической зависимости и согласованности с макетом их размещения на машинном носителе. Зоны документа могут иметь линейную, анкетную или табличную формы. В линейной форме каждому реквизиту выделяются две клетки: для его наименования и для значения. Анкетная форма предусматривает размещение реквизитов (их наименований и значений) в вертикальной последовательности. Основная зона документа, как правило, представляет собой таблицу с графами по вертикали и строками по горизонтали.

В одном документе чаще всего используется сочетание различных форм построения его зон. Для уменьшения количества документов и сокращения затрат на их заполнение и подготовку машинных носителей целесообразно применять многострочные формы. При этом размещение реквизитов производится в определенной последовательности, которая первоначально отражается в эскизной форме документа.

Реквизиты, переносимые на машинный носитель, по возможности располагаются в одной части документа и выделяются утолщенной линией. Из них отбираются справочно-группировочные признаки, постоянные для данного документа, и размещаются в правом верхнем углу бланка, в общей части документа. В первичные документы часто вводятся специальные графы и строки для записи контрольных сумм, предназначенных для контроля заполнения документов, переноса данных на машинный носитель и ввода их в ЭВМ. В них также предусматриваются подписи лиц, ответственных за содержание данных, записанных на документе.

По эскизу документа проектируется макет размещения данных на машинный носитель, производится их согласование и уточнение содержания документа с учетом других форм взаимосвязанных первичных документов, после чего определяется формат бланка уточненной формы проектируемого документа в соответствии с требованиями стандартов. Если первичный документ предназначен для заполнения на пишущих или счетно-табличных машинах, обязательно учитываются конструктивные особенности этих машин. Свои особенности присущи документам, данные с которых передаются по телеграфу (в них вводятся служебные знаки телеграфирования: начало и конец цифровых данных, возврат каретки, перевод строки). Производятся расчеты размеров бланка документа и выбор стандартного формата. Ширина бланка (в мм) определяется по формуле

$$Ш = П_{\text{л}} + П_{\text{п}} + ш(z + z) + 0,5L_{\text{т}} + 2,2L_{\text{у}},$$

где $П_{\text{л}}$ и $П_{\text{п}}$ — ширина левого и правого чистых полей;

$ш$ — ширина одного знака (символа);

z — максимальное или наиболее вероятное количество знаков (символов) по строке;

z — количество граф (столбцов) документа;

$L_{\text{т}}$ и $L_{\text{у}}$ — количество соответственно тонких и утолщенных вертикальных линий.

Аналогично рассчитывается длина бланка:

$$Д = П_{\text{в}} + П_{\text{н}} + П_{\text{оф}} + вc + 0,5L'_{\text{т}} + 2,2L'_{\text{у}},$$

где $П_{\text{в}}$ и $П_{\text{н}}$ — высота верхнего и нижнего чистых полей бланка;

$П_{\text{оф}}$ — величина поля оформления документа;

$в$ — высота одной строки документа (с учетом высоты знаков);

$с$ — количество строк;

L'_T и L'_y — количество тонких и утолщенных горизонтальных линий.

По полученным значениям ширины и длины ($Ш$ и $Д$) подбирается подходящий формат по стандарту с возможной корректировкой ширины и длины отдельных столбцов (граф) и строк документа. Чаще всего используются следующие стандартные форматы (в мм): 420×297 , 297×210 , 210×148 . По уточненному эскизу подготавливается образец документа для утверждения его формы и передачи в типографию.

При проектировании уникальных документов, предназначенных для автоматического считывания с них информации с помощью специальных устройств, дополнительно учитываются эксплуатационные возможности этих устройств. В нашей стране разрабатывались и производились различные читающие автоматы, использующие нормализованные или стилизованные шрифты, графические метки. В наиболее массовом порядке использовались специальные документы с графическими метками, предназначенные для Всесоюзных переписей населения 1970 и 1979 гг., информация с которых перезаписывалась на магнитную ленту или непосредственно вводилась в ЭВМ с помощью читающих автоматов типа «Бланк».

Организация документооборота, т. е. прохождения документов в процессе их составления и обработки, должна быть максимально простой и отвечать следующим основным принципам:

1) своевременной подготовке документов с широким использованием средств механизации и автоматизации сбора, регистрации и передачи информации;

2) оптимизации путей их прохождения и рациональной обработке;

3) взаимосвязи документов и полноте охвата всех процессов производственной деятельности конкретного объекта управления.

Для рациональной организации документооборота разрабатываются схемы потоков информации по каждому подразделению и объекту управления в целом, анализируется взаимосвязь между различными потоками с использованием математических методов. Так, например, широко известен метод анализа и оптимизации внутренних документопотоков с помощью транспортной модели линейного программирования. При этом в качестве критерия оптимальности выбирается суммарная кратность передач документов по маршрутам их движения.

Проектирование форм вывода результатной информации (табуляграмм) имеет много общего с проектированием первичных документов, особенно в части определения состава и порядка размещения реквизитов, и в то же время характеризуется некоторыми специфическими требованиями, определяемыми самой организацией машинной обработки данных и эксплуатационными возможностями используемых технических средств.

Работы по проектированию форм вывода результатной информации выполняются в такой последовательности:

1) определяется перечень результатных данных;

2) результатные показатели распределяются по табуляграммам и составляется список выводимых форм;

3) каждой форме присваиваются наименование и кодовое обозначение (номер);

4) устанавливаются периодичность составления и срок получения каждой табуляграммы;

5) определяются место использования табуляграммы и необходимое количество экземпляров;

6) разрабатывается эскиз табуляграммы;

7) определяется состав массивов (файлов) данных, необходимых для получения каждой формы;

8) устанавливаются последовательность составления выходных форм и общий технологический процесс, подготавливаются контрольные примеры табуляграмм, они обсуждаются и уточняются;

9) утверждаются уточненные формы вывода результатной информации.

Особенно важной является разработка эскиза табуляграммы. Для каждой формы вывода результатной информации прежде всего определяется содержание трех ее зон: заголовка или титульного листа табуляграммы, выводимого однократно или частично повторяемого на каждом листе; наименований граф и их нумерации, выводимых на каждом листе (заголовка граф); основной предметной (информационной) части, составляющей тело табуляграммы. В последнюю зону включаются строки двух видов: детальные и итоговые. Детальные строки содержат данные, имеющиеся в записях входных массивов. Их размещение по строке выполняется аналогично размещению реквизитов-признаков в первичных документах. Итоговые строки содержат количественные и суммовые итоги по определенным признакам, результаты выполняемых расчетов. Их размещение зависит от количества контролируемых признаков, предусматриваемых условий вывода промежуточных результатов (например, по каждому листу) и других факторов, определяемых пользователем. В итоговые строки часто дополнительно включают текстовые комментарии (например, «итого по ...», «всего» и т. п.).

При определении содержания и формы каждого выходного документа учитываются цели, для которых он предназначен, сфера и особенности его использования, периодичность получения и др. Очень важно установить взаимосвязь различных результатных показателей, чтобы определить возможность их объединения в одну табуляграмму или параллельного составления двух или нескольких однородных табуляграмм. При этом следует учитывать не только технические возможности устройств вывода информации, но и особенности предусматриваемых программных средств подготовки и вывода данных (использование генераторов отчетов; способы формирования титульного листа и заголовка выходного документа; предварительная запись выводимой информации на магнитных носителях и двусторонняя печать; использование диа-

логового режима, терминальных устройств; графическое представление данных и т. д.).

Некоторые наиболее важные вопросы организации вывода результатов более подробно освещены в гл. 13 при рассмотрении технологических процессов формирования и вывода результатных данных.

9.4. Методика и особенности проектирования макетов размещения данных на машинных носителях

На машинных носителях размещаются базовые, первичные, промежуточные и результатные массивы (файлы), представляющие собой определенным образом организованные и поименованные группы логически связанных между собой данных. При размещении массивов на том или ином носителе необходимо стремиться обеспечить экономное использование носителя и эффективную обработку массивов данных. В Единой системе ЭВМ в качестве машинных носителей преимущественно используются перфокарты, перфоленты, пакеты магнитных дисков, магнитные ленты и магнитные барабаны. Массивы, располагаемые на томах последовательного или прямого доступа, подлежат макетированию, т. е. установлению определенного состава и порядка размещения реквизитов с выделением для каждого из них соответствующей части (группы) определенным образом кодируемых элементарных позиций информационной емкости носителя. При этом определяется структура как физических, так и логических записей.

Традиционные макеты перфорации карт, применявшиеся для обработки данных на перфорационных вычислительных машинах, используются и в условиях эксплуатации современных ЭВМ. Для них характерно то, что они предусматривают перфорацию каждой документостроки на отдельной карте с повторением общих для всех строк документа реквизитов в каждой записи массива, имеющей фиксированную длину и чаще всего занимающей не все 80 колонок перфокарты. Такое макетирование легко согласовывается с формой первичных документов, упрощает описание структуры записи в программах пользователя, особенно если используются языки высокого уровня, а также согласование различных макетов друг с другом, но вместе с тем имеет существенные недостатки. В таких макетах практически не предусматривается блокирование логических записей на одной перфокарте (если такая возможность существует). Чаще можно встретить объединение физических записей в одну логическую, размещаемую на нескольких картах. Практически не используются записи переменной и неопределенной длины, хотя для таких массивов возможно применение форматных карт, описывающих структуру записей, причем длину записи можно менять внутри одного и того же массива. Форматное описание структуры записи можно включить в каждую перфокарту, используя при этом битовую строку, нулевые позиции которой указывают отсутствующие элементы стандартной

(полной) структуры записи. Это прежде всего подходит для корректирующих массивов, частично меняющих содержание корректируемых записей. Также можно использовать шаблоны, в которых указывается длина каждого элемента записи.

На перфокарты могут выводиться данные промежуточных и результатных массивов, а также копии файлов базы данных для их восстановления. При макетировании таких массивов следует предусмотреть вывод данных без перекодирования, т. е. по 120 или 160 байт на каждую перфокарту, при этом вывод всего массива можно производить сплошной битовой строкой, разве что с добавлением нумерации перфокарт. К тому же это касается не только алфавитно-цифровой, но и чисто цифровой информации, которая должна пробиваться в упакованном формате, т. е. с двойной экономией карточной памяти. Получаемые таким образом не условные, а действительные копии файлов могут использоваться для частичного или полного их восстановления на магнитных носителях, при этом адреса сбойных участков конкретного файла легко переводятся в соответствующие порядковые номера перфокарт его копии. В этом смысле перфокарты имеют определенное преимущество по сравнению с перфолентой.

Запись (кодирование) данных на перфоленту может производиться ручным способом (их перенос с первичных документов) или автоматически (на перфопроставках, устройствах вывода ЭВМ). Информация наносится в виде непрерывной последовательности символов, каждый из которых кодируется определенной комбинацией пробивок в одной строке. В ЕС ЭВМ предусмотрено использование 5-, 6-, 7- и 8-дорожечных перфолент. Макетирование перфолент имеет свои особенности. Для организации массивов данных на перфоленте используются два формата записей — *F* и *U*. Записи в формате *F* наносятся на ленту друг за другом и не разделяются между собой, а записи в формате *U* отделяются специальным кодом конца блока, роль которого может выполнять любой символ, не встречающийся внутри записей. Блоки могут иметь различную длину. В каждую физическую запись (блок) кроме значений ее реквизитов, которые располагаются в согласованной с соответствующей формой документа последовательности и представляются в любом коде (в том числе и внутримашинном), могут быть включены различные служебные признаки. Так, например, часто используются специальные символы, отделяющие друг от друга документы, их строки, отдельные реквизиты (при их записи по фактической значности) и т. д. В массив могут входить разнотипные блоки. Кроме записей, содержащих информацию документострок, выделяются записи со справочно-группировочными реквизитами, постоянными на каждый документ и их группы (пачки). Тем самым постоянные реквизиты пробиваются один раз и не повторяются для каждой строки документа, как это чаще всего делается при использовании перфокарт. Поэтому макетом перфорации предусматривается пробивка признака типа блока; блоки могут иметь свои порядковые номера.

Массив данных может размещаться на нескольких бобинах перфоленты, а на одной бобине, в свою очередь, можно разместить несколько массивов. В последнем случае массивы отделяются друг от друга участками перфоленты без пробивок, длина которых не менее 25 см.

В современных системах машинной обработки данных перфоленты, как и перфокарты, могут применяться не только для фиксации оперативных первичных данных, подготовленных к вводу, временного хранения некоторых массивов, организации копий файлов данных, но и выполнять роль низшего уровня системной внешней памяти. Поэтому они находят применение и при использовании магнитных носителей данных.

Все более широкое применение получают устройства подготовки данных на магнитных лентах (ЕС-9001, ЕС-9002). С их помощью первичные массивы формируются по предварительно подготовленной программе, предусматривающей набор и размещение данных по шаблону, а также некоторые операции контроля. Такая программа вводится с клавиатуры устройства или с магнитной ленты. Данные вводятся с клавиатуры. При этом производится их программный контроль, они формируются и накапливаются в буферной памяти и отдельными блоками записываются на ленту. Затем производится автоматический контроль правильности записи информации и формирование межблочного промежутка. После окончания ввода всего массива данных устройство переводится в режим верификации, при котором после повторного ввода данных с клавиатуры происходит их сравнение с данными, считанными с магнитной ленты. В случае несоответствия производится корректировка данных на ленте. Также имеется возможность осуществить поиск блока данных на ленте по ключу, вводимому с клавиатуры, и печать данных, считанных с магнитной ленты.

Также используется автоматизированная система подготовки данных на магнитных лентах СПД-9000, разработанная ВНИПИОАСУ на базе малой машины третьего поколения серии АСВТ М-6000. Помимо своего прямого назначения она может быть использована и для выполнения некоторых операций предварительной обработки данных (проверки, сортировки, слияния и др.), а также для диспетчеризации работ по подготовке данных. В систему может входить до 32 клавишных пультов операторов, одновременно осуществляющих ручной ввод данных для первоначального формирования массивов или контроля правильности и корректировки ранее записанных на магнитных лентах массивов.

В ЕС ЭВМ, как уже отмечалось, входит большое разнообразие накопителей данных прямого доступа — различные пакеты магнитных дисков и магнитные барабаны. По используемым форматам физических записей и организации данных все типы томов прямого доступа идентичны, поэтому мы рассмотрим лишь организацию данных на пакетах магнитных дисков для распространенных накопителей типа ЕС-5061. Данный пакет состоит из 11 дисков. Запоминающая (рабочая) поверхность каждого из них пред-

ставляет собой совокупность 203 дорожек, на которые записывается информация. Группа дорожек, доступных при фиксированном положении механизма записи и считывания данных (блока магнитных головок), называется цилиндром. Цилиндр рассматриваемого пакета содержит 20 дорожек, так как на наружных поверхностях нижнего и верхнего дисков запись не производится. Цилиндры нумеруются от 0 до 202, а дорожки внутри цилиндра — от 0 до 19. Данные на дорожку записываются в виде блоков, также нумеруемых с нуля. Таким образом, физический адрес блока задается пятибайтовым двоичным числом вида CCHNR, где CC — номер цилиндра (2 байта), NN — номер дорожки (2 байта), R — номер блока на дорожке (1 байт). Каждый блок записывается в виде двух или трех отдельных полей, разделенных промежутками: поля идентификатора, поля ключа и поля данных. Блоки могут быть с ключами и без них, т. е. поле ключа может отсутствовать. Форматы дорожки с блоками без ключей и с ключами изображены на рис. 9.2 и 9.3.

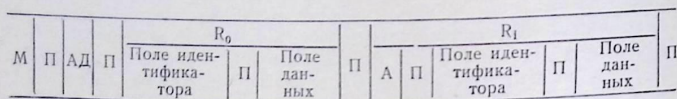


Рис. 9.2. Формат дорожки с блоками без ключей

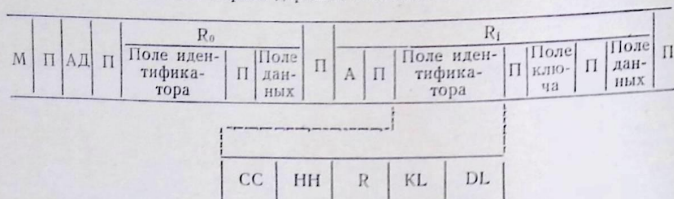


Рис. 9.3. Формат дорожки с блоками, содержащими ключи

Начальная точка записи на всех дорожках пакета указывается специальной меткой М начала оборота, которая опознается при каждом полном обороте дисков. Через определенный промежуток П записывается собственный адрес дорожки АД, в котором кроме номеров цилиндра и дорожки отражается и ее состояние. Далее следует блок R₀, который обычно используется операционной системой для хранения информации о заполненности дорожки данными. Вслед за блоком R₀ идет адресный маркер А, представляющий собой специальное поле, идентифицирующее начало блока. Адресный маркер предшествует каждому блоку, за исключением блока R₀. Идентификатор каждого блока содержит его физический адрес CCHNR, а также информацию о длине ключа (KL) и данных (DL). В поле ключа записывается некоторый признак бло-

ка, который может использоваться в качестве аргумента при поиске блока, если неизвестен его физический адрес. Длина ключа может составлять от 1 до 255 байт (нулевая длина ключа в поле идентификатора означает его отсутствие). Максимальная длина поля данных равна 65 535 байтам, хотя обычно длина блока не превышает длины одной дорожки (3625 байт), или, если в устройстве управления накопителями реализована возможность сегментирования блоков, она не превосходит длины цилиндра. Массивы данных на одном томе прямого доступа разделяются специальным блоком, представляющим собой однобайтовый фиктивный блок с нулевой длиной ключа и данных.

Дорожки пакета делятся на основные (цилиндры 0—199) и запасные (цилиндры 200—202). Если при проверке поверхностей тома, которая выполняется по специальной программе инициализации, в группе основных обнаруживается дефектная дорожка, то она заменяется запасной. Автоматический переход с дефектной дорожки на запасную выполняется с помощью блоков R₀, которые содержат адреса соответственно заменяющей запасной и заменяемой основной дорожек. В процессе инициализации формируются метка и оглавление тома.

Метка тома записывается на нулевой дорожке нулевого цилиндра. Она представляет собой блок, содержащий идентификатор, ключ длиной 4 байта и собственно метку длиной 80 байт. Вслед за данной меткой на этой же дорожке может располагаться еще до 7 дополнительных меток такого же формата.

Оглавление тома — это набор данных, который содержит метки всех массивов данных, находящихся в текущий момент на данном томе. Оглавление также содержит информацию о свободных полях тома. Оно состоит из блоков, имеющих 4-байтовое поле ключа и 96-байтовое поле данных. Каждый блок представляет собой определенную метку, имеющую какой-либо один из 7 возможных форматов. В процессе инициализации первой в оглавлении тома формируется метка 4-го формата, т. е. метка самого оглавления. В ней содержится информация об устройстве, на котором установлен пакет, о размере самого оглавления, о запасных дорожках и др. Далее следует метка 5-го формата с информацией о свободных полях на томе (указываются физические адреса нижней и верхней границы каждого свободного поля). В одной метке может быть описано до 26 непрерывных свободных полей. Если их число больше, формируется еще одна метка 5-го формата. Все метки данного формата связываются в цепочку. Так как при инициализации на томе находится только само оглавление тома, а оставшая его часть свободна, то сразу за меткой 5-го формата в оглавлении следуют метки нулевого формата, которые не содержат никакой полезной информации и записываются только с целью форматирования оглавления тома. В дальнейшем при использовании данного тома на место меток нулевого формата в момент выделения памяти под массивы данных помещаются метки, описывающие эти массивы.

Массивы данных могут описываться метками 1, 2 и 3-го форматов. Метка 1-го формата имеется в оглавлении для каждого массива. Она используется управляющей программой операционной системы для определения физических границ полей тома, выделенных под определенный массив данных. В одной метке данного формата могут быть описаны границы только трех непрерывных полей тома. Если массиву выделяется большее число полей, то в оглавлении тома формируются метки 3-го формата, описывающие последующие поля тома. Одна метка 3-го формата может содержать описания до 13 полей. Метки 1-го и 3-го форматов связываются между собой в цепочку.

Для индексно-последовательных массивов кроме метки 1-го формата формируется метка 2-го формата, в которой описывается физическое расположение различных частей этих массивов (основного поля, поля индексов и поля переполнения). Метка 2-го формата связывается с меткой 1-го формата.

При одновременном обращении к нескольким массивам данных, расположенным на одном томе, для минимизации перемещений магнитных головок эти массивы можно разместить на нескольких цилиндрах так, чтобы на каждом цилиндре находилась определенная порция записей каждого массива. В этом случае в оглавлении тома формируется метка 6-го формата, описывающая ту совокупность распределенных цилиндров, которая выделена под несколько массивов данных. Метки 6-го формата образуют цепочку с меткой 4-го формата.

Кроме рассмотренных меток массивов данных могут формироваться метки пользователя начала и конца массива. Они записываются только на первой дорожке, отведенной для конкретного массива данных. Каждый массив может иметь до 8 меток пользователя начала и конца массива. Такие метки представляют собой блоки, содержащие 4-байтовое поле ключа и 80-байтовое поле данных. Для их обработки используются специальные программы пользователя.

При проектировании размещения информации на магнитных носителях прямого доступа очень важно не только определить содержание и структуру поля данных записи массива, но и выбрать признак для выполнения роли ключа. В этом смысле организация конкретного массива непосредственно зависит от предусматриваемого способа доступа к нему. Операционная система ЕС ЭВМ (ОС ЕС) располагает четырьмя типами организации массивов данных: последовательной, библиотечной (относительной), индексно-последовательной и прямой.

Последовательная организация массивов на магнитных носителях принципиально не отличается от их размещения на перфокартах и перфолентах. Записи размещаются на носителе в той последовательности, в которой они поступают.

При библиотечной организации массив состоит из поименованных разделов (внутри которых данные организованы последовательно) и оглавления, в котором содержатся имена всех разделов

и их относительные адреса. Обращение к любому разделу производится непосредственно по его имени. Если каждая логическая запись массива снабжается своим относительным номером, можно говорить об относительной организации массива (только для ОС ЕС).

Индексно-последовательный массив данных представляет собой набор упорядоченных по возрастанию ключа логических записей формата *F* или *V*. Логические записи, каждая из которых имеет уникальный ключ, могут объединяться в блоки. В данном случае в поле ключа блока помещается ключ последней логической записи блока. Работа с индексно-последовательными массивами предполагает поиск записей по ключу. Для сужения области поиска создаются индексы (указатели): индекс дорожек на каждом цилиндре основного поля массива, индекс цилиндров и до трех уровней главных индексов. Индекс дорожек размещается на нулевой дорожке цилиндра и содержит физический адрес первой записи каждой дорожки данного цилиндра и ключ последней записи на дорожке. На каждом цилиндре может быть предусмотрено поле переполнения, в которое помещаются записи при расширении массива. Индекс цилиндров содержит физический адрес индекса дорожек и ключ последней записи данного цилиндра. Главные индексы строятся в случае, если массив данных занимает большое количество цилиндров и индексы цилиндров сами заполняют несколько цилиндров.

Прямая организация массивов характеризуется наличием связи между ключом блока данных и его адресом. В этом случае ключ блока совпадает с его физическим адресом или с помощью специального алгоритма по ключу определяется адрес блока. При этом необходимо предусмотреть поле переполнения массива для записей, имеющих одинаковые адреса.

На томах прямого доступа массивы данных чаще всего записываются автоматически с помощью системных программ перезаписи данных с других носителей или специальных программ пользователей, но могут использоваться и автономные устройства подготовки данных, предусматривающие ручной набор информации на клавиатуре с предварительным накоплением ее в буферной памяти устройства. В последнем случае применима последовательная организация массивов.

Последовательную организацию имеют также массивы данных, размещаемые на магнитных лентах. В ЕС ЭВМ используются 9-дорожечные магнитные ленты, на которые данные записываются в двоичном коде. Информация начинается с точки загрузки, которая отмечается специальным маркером начала ленты. Информационные блоки, минимальная длина которых равна 18 байтам, разделяются межблочными промежутками, а массивы — специальным блоком, называемым ленточной маркой. Также имеется маркер конца ленты.

В простейшем случае массив данных представляет собой совокупность блоков, расположенных между двумя ленточными мар-

ками или между точкой загрузки и ленточной маркой. Для удобства работы с магнитными лентами рекомендуется использовать стандартные метки массивов. Они представляют собой физические записи длиной 80 байт. При этом информация на ленте имеет следующую структуру: группа меток, идентифицирующая том; массивы данных, каждый из которых включает группу меток начала массива, собственно данные (информационные блоки) и группу меток конца массива (или конца тома).

Группа меток, идентифицирующая том, записывается в начале каждой магнитной ленты и может содержать от одной до восьми меток, с помощью которых автоматически контролируется правильность установки тома на соответствующем накопителе.

Группа меток начала массива данных может содержать от одной до восьми системных меток, а также до восьми пользовательских меток. Чаще всего используются две метки: первая — для контроля правильности подвода запрошенного программой массива данных, а вторая содержит некоторые характеристики массива (формат записи, длину блока, длину логической записи и др.). Группа меток начала массива отделяется от первого информационного блока ленточной маркой.

Группа меток конца массива (конца тома) также может содержать от одной до восьми системных и до восьми пользовательских меток, но чаще всего используются две метки, содержание которых аналогично содержанию соответствующих меток начала массива, что позволяет обрабатывать массив в обратном направлении от конца к началу. После группы меток конца массива следует одна ленточная марка, если вслед за этим массивом размещен следующий, или две ленточные марки, если это последний массив на данном томе. Группа меток конца тома записывается в случае, если массив данных не умещается на одной ленте и его продолжение следует на другой ленте.

Метки пользователя могут обрабатываться только специальными программами пользователя. Возможно также использование так называемых нестандартных меток, для обработки которых должны быть подготовлены специальные программы, встраиваемые в управляющую программу ОС ЕС.

Магнитные ленты, которые предназначены для работы со стандартными метками, предварительно инициализируются. Инициализация тома состоит в записи на нем стандартной метки тома, фиктивной метки начала первого массива данных и ленточной марки.

Структура файлов, размещаемых на машинных носителях, чаще всего представляется в виде таблиц. В общей части таких таблиц указываются наименование и код файла, способ организации, источник формирования, его использование, объем. В содержательной части раскрывается структура файла (его записей), перечисляются реквизиты (поля), указываются их условные обозначения, размеры, форма представления, а при необходимости — их источники и способы получения.

10.1. Принципы организации и размещения данных в памяти системы

Информационный фонд — основа информационного обеспечения АСОД и АСУ. Он представляет собой совокупность всех данных, подлежащих накоплению, хранению, поиску, преобразованию и выдаче в установленном порядке, а также используемых для организации общения человека с ЭВМ. Существуют два подхода к созданию информационного фонда — локальный и интегрированный. При локальной структуре информационного фонда создаются массивы (файлы) данных, предназначенные для решения отдельных задач управления или их комплексов. Файлы при этом обычно бывают простыми последовательными наборами данных, полностью зависящими от программного обеспечения, используемого для их обработки, т. е. конкретных прикладных программ, при помощи которых к этим данным обращаются их пользователи.

Многие файлы локального информационного фонда формируются отдельно для каждого расчета, решения отдельных задач. Поэтому те же данные, используемые в различных расчетах, обычно представляются в другой форме. Это ведет к тому, что в фонде создается большая избыточность данных, возникают определенные трудности при их обновлении, так как практически при большом объеме фонда очень трудно сохранять все файлы на одном и том же уровне обновления. В связи с этим пользователи часто могут обнаруживать явные противоречия в данных и испытывать определенное недоверие к ЭВМ. Кроме того, при постоянном изменении и расширении фонда локальных файлов возникает необходимость в модификации прикладных программ, требующей значительных затрат. Указанные недостатки локальных информационных фондов, в основном использовавшихся в системах обработки данных на базе ЭВМ второго поколения (примерно до конца седьмого десятилетия), полностью или частично отсутствуют в единых (интегрированных) информационных фондах.

Единые фонды также могут иметь файловую структуру, но при этом данные, включаемые в определенный файл, могут использоваться для нескольких приложений, т. е. они имеют многоцелевое назначение. Этим, прежде всего, минимизируется избыточность данных и тем самым повышается их достоверность, исключается возможность несоответствия получаемых результатов решения информационно связанных друг с другом задач, упрощается процесс обновления данных.

Если для создания локальных файлов, физическая и логическая структуры которых полностью совпадают, преимущественно используются перфоносители и магнитные ленты, то единые информационные фонды файловой структуры чаще размещаются на носителях прямого доступа.

Разделение физической и логической структуры данных позволило обеспечить их независимость, а также независимость данных от прикладных программ, что создало возможность перейти к новому уровню организации единых информационных фондов, получившему название базы данных. Это понятие возникло в конце 60-х годов, а практические реализации самой концепции баз данных связаны с внедрением ЭВМ третьего поколения. Базу данных можно определить как управляемый единый информационный фонд, т. е. в нее входят не только соответствующим образом организованные и логически связанные данные, но и система их описания, а также средства, поддерживающие установленные информационные связи.

Возможны два пути разработки баз данных: проектирование оригинальных баз данных; использование типовых баз данных и систем их генерации и интеграции.

Выбор конкретного пути зависит от специфики предметной области, для которой база данных разрабатывается. Для базы данных (или нескольких баз) создается система управления, представляющая собой комплекс программных средств формирования, ведения и использования базы данных. Организационно-техническая система, включающая в себя базу данных, систему управления ею, а также специальные языки, программные и технические средства, в совокупности обеспечивающие многоаспектное использование информации и прямую связь с пользователем, определяется как автоматизированный банк данных (АБД).

Существуют три отдельных представления организации данных в условиях АБД:

общее логическое представление баз данных (на языке описания данных);

логическое файловое представление данных для отдельных прикладных программ пользователей;

физическое представление данных (непосредственное размещение их на запоминающих устройствах, машинных носителях).

К организации и размещению данных в памяти системы предъявляются некоторые общие требования вне зависимости от используемого способа организации информационного фонда или базы данных. К таким общим требованиям относятся:

1) физическая и логическая независимость данных, исключая необходимость перезаписи прикладных программ при изменении устройств памяти ЭВМ, способов физического хранения данных, при добавлении новых элементов данных и реорганизации их общих логических структур;

2) одноразовый ввод и обеспечение многократного использования данных;

3) обеспечение требуемой скорости доступа и поиска данных;

4) простота внесения изменений;

5) обеспечение целостности и достоверности при минимальной избыточности данных;

6) наличие интерфейса для связи с пользователями, средств

защиты и обеспечения секретности данных (интерфейс представляет собой комплекс технических и программных средств, обеспечивающих связь между отдельными элементами АСОД).

Независимость данных обеспечивается тем, что различные логические файлы для конкретных прикладных программ могут быть получены из одной и той же физической совокупности данных на уровне элементов данных, а не на уровне записей. Таким образом, физическое представление данных в памяти системы может сильно отличаться от того представления, которое используется в прикладных программах. Преобразование одного представления в другое осуществляется системным программным обеспечением базы данных или, если это возможно, аппаратурными и микропрограммными средствами ЭВМ.

Целью организации базы данных является исключение избыточных данных там, где это выгодно, и контроль за теми противоречиями, которые возникают из-за избыточности данных. При этом обеспечивается многократное использование одних и тех же данных, между элементами которых устанавливаются необходимые взаимосвязи и разрабатываются специальные средства их соответствующего обслуживания.

В большинстве современных систем машинной обработки данных используются методы доступа и поиска данных, ориентированные на заранее predetermined типы запросов пользователей. Скорость поиска в базе данных в решающей степени зависит от физической организации данных, определяющейся в основном конкретным содержанием регламентированных обращений к базе данных. Однако развитие систем обработки постоянно выдвигает новые требования, заключающиеся прежде всего в обеспечении обработки таких запросов и формирования таких ответов, которые заранее не планируются. Поэтому проблема обеспечения требуемой скорости доступа и поиска данных как при регламентированных, так и в особенности при случайных запросах может быть успешно решена только путем организации многоуровневой памяти системы. Наиболее четко можно выделить три уровня памяти:

оперативное запоминающее устройство;

накопители на магнитных дисках и барабанах (прямой доступ);

накопители на магнитных лентах (последовательный доступ).

Как уже отмечалось в 9.4, низший уровень системной памяти могут представлять и перфоносители.

Данные между различными уровнями памяти можно перемещать с помощью страниц фиксированной длины или отдельными блоками поэтапно. При этом часто используемые элементы данных запоминаются на устройствах с быстрым доступом, а редко используемые — на устройствах с медленным доступом. По аналогичному принципу их можно распределять внутри одного и того же уровня памяти, тем самым увеличивая скорость поиска. Так, например, наиболее часто используемые данные следует запоминать на внутренних цилиндрах диска (со среднего — в сторону уменьшения и увеличения), а наименее часто — на внешних ци-

линдрах (с нулевого — в сторону увеличения и с последнего — в сторону уменьшения).

Способы доступа и методы поиска в базе данных непосредственно связаны с системой внесения изменений в базу. Эта система должна быть предельно простой вне зависимости от того, какая структура данных (древовидная, списковая, сетевая или нормализованная) положена в основу информационного фонда. При ее проектировании следует исходить из того, что для более динамичных данных проще и удобнее использовать нормализованную форму представления (т. е. бинарные отношения между элементами данных) и учитывать то, что обновление данных на определенной ступени может привести к потере первоначальной оптимальности размещения данных (т. е. их следует периодически подвергать реорганизации). Часто претерпевают изменения не только сами элементы данных, но и отношения между ними. Это влечет перестройку системы указателей, отражающих связи между данными. Обычно используются три типа указателей: машинный адрес, относительный адрес или символический идентификатор. Очевидно, что для интенсивно изменяющихся файлов, имеющих много указателей, предпочтительнее применять символические указатели, чем адресные.

При проектировании информационного фонда необходимо предусмотреть средства контроля за его целостностью, т. е. за тем, чтобы элементы данных и связи между ними не разрушались при возникновении ошибок и различного рода случайных сбоев. Система должна сама без потерь восстанавливать данные, периодически производить проверку соответствия их значений некоторым определенным условиям. Также необходимо использовать контроль за достоверностью данных, предотвращающий наличие различных версий одних и тех же элементов данных на различных стадиях их обновления. С этой целью в базах данных предусматривается хранение на определенный период времени оперативно вводимых изменений (корректирующих данных) с возможной пересылкой в архив той части информационного фонда, которая подвергалась корректировке.

Для первоначальной генерации информационного фонда, его последующего ведения, частичных реорганизаций и обновлений, поиска данных, включаемых в формируемые ответы на запросы пользователей, необходимо разработать интерфейс, обеспечивающий возможность просто и эффективно осуществлять обращение к данным, легко формулировать цели такого обращения вне зависимости от всей сложности размещения и адресации данных. Такой интерфейс предусматривает использование по крайней мере двух языков высокого уровня:

для логического представления данных в прикладных программах;

для запросов пользователей (без применения прикладных программ), формируемых при ведении и использовании данных.

Одновременно в системе могут быть предусмотрены средства

защиты данных от некомпетентного обращения к ним, особенно с целью их ошибочного обновления или уничтожения, а также невозможности несанкционированного доступа к данным. Для этого разрабатывается система идентификации пользователей, а в программное обеспечение включается процедура контроля за正确ностью и возможностью обращения к данным.

Следует отметить, что принципы создания единых информационных фондов, баз данных и АБД, а также формы их конкретной реализации находятся в постоянном развитии. Разработка баз данных на современном этапе представляет собой главное направление деятельности в области обработки данных.

10.2. Формирование и ведение единого информационного фонда

Информационный фонд системы обработки данных можно считать единым, если в него входят файлы определенным образом связанных между собой данных, необходимые одновременно для нескольких приложений и исключающие необоснованное дублирование данных. Связи между данными единого фонда могут проявляться в смысловом и технологическом отношении. Цели создания таких фондов заключаются в уменьшении объема хранимой информации, упрощении технологии обработки данных (повышении степени ее автоматизации), обеспечении высокой надежности функционирования системы при многоаспектном использовании одних и тех же данных.

Первые шаги на пути к созданию единых информационных фондов связаны с организацией на предприятиях централизованных нормативно-справочных хозяйств. Высокий уровень постоянства, большой удельный вес (до 60—70% общего объема управленческой информации), многократность использования нормативно-справочной информации — вот те основные факторы, которые определяют возможность и необходимость создания единой нормативной базы любой системы управления. В такую базу включаются основные (первичные) и производные (вторичные) нормативные массивы¹ (нормы расхода, трудоемкости и т. д.), а также справочные массивы (маршрутные, структурные, призначные, табличные и др.). По многим совпадающим характеристикам к нормативно-справочной части информационного фонда можно отнести также массивы плановых данных. Для первоначального формирования этой части фонда характерна высокая требовательность к достоверности данных (должна быть разработана надежная система их контроля), а также сравнительно большой удельный вес расчетных операций, выполняемых как для получения производных (вторичных) нормативов и плановых показателей, так и для обеспечения полного соответствия всех взаимосвязанных нормативно-справочных и плановых данных. То же относится к подготовке

¹ Здесь и далее термин «массив» используется в тех ситуациях, когда речь идет о локальных информационных фондах, а «файл» — о единых базах данных.

корректирующих данных и в особенности к процедуре обновления нормативно-справочной базы.

Условно-постоянная информация, в том числе нормативно-справочная, должна распределяться по массивам с учетом частоты использования как для решения отдельных задач управления, так и для всех расчетов, выполняемых в АСУ. При этом следует формировать меньшее число общих массивов, используемых для нескольких задач. Применение общих массивов целесообразно даже в случае, если в такие массивы необходимо включить некоторые реквизиты, предназначенные специально лишь для отдельных задач, при условии, что основная часть реквизитов используется во многих задачах. Тогда система таких массивов непременно будет рациональной. В связи с тем, что нормативно-справочные массивы в прикладных программах обрабатываются совместно с соответствующими массивами переменной информации, важную роль играет правильный выбор ключевых признаков условно-постоянных массивов, с помощью которых при обработке подбираются их соответствующие записи. Согласованность ключевых признаков, правильное определение их иерархии — непременное условие эффективности обработки данных.

В единый информационный фонд кроме условно-постоянных включаются также файлы переменной информации. Переменная часть фонда может быть организована в виде оперативных, накапливаемых, промежуточных хранимых и результатных файлов. Оперативные файлы используются только для ввода, а результатные — для вывода данных. Накапливаемые и промежуточные хранимые файлы могут быть одновременно входными и выходными, причем накапливаемые файлы постоянно присутствуют в информационном фонде, а промежуточные хранятся в нем регулярно, но временно.

Степень интеграции данных в этих файлах различная. Наиболее интегрированными являются накапливаемые файлы, которые обновляются аналогично условно-постоянным массивам, но, как правило, не частично, а полностью. Связь между рассматриваемыми видами переменных файлов показана на рис. 10.1.

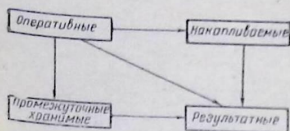


Рис. 10.1. Связь между переменными файлами информационного фонда

Оперативные файлы с первичной информацией используются как для первоначального формирования, так и для последующего ведения (поддержания в актуальном состоянии) накапливаемых файлов, а также могут участвовать в формировании промежуточных и результатных файлов. Для последних они, как правило, используются в режиме поиска (запрос — ответ) без выполнения каких-либо расчетов. Промежуточные файлы целесообразно фор-

мировать в случаях, когда необходимо проводить определенные расчеты с использованием лишь оперативных файлов, если результаты данных расчетов непосредственно не используются, но нужны для последующего формирования выходных данных (автономно в режиме поиска или совместно с другими файлами в режиме расчетов). Что же касается накапливаемых файлов, то здесь необходимо отметить большое разнообразие различных процедур их формирования и обновления (собственно накопление данных по определенным признакам, вычисление различных показателей с применением экономико-математических методов, перегруппировка показателей и др.).

Если оперативные, промежуточные и накапливаемые файлы формируются (а последние и обновляются) в заранее установленном порядке, т. е. регламентированно, то определенная часть результатных файлов может формироваться и произвольно. Это обстоятельство учитывается при разработке программного обеспечения ведения информационного фонда. В него прежде всего включаются программные модули, реализующие регламентированные и регулярные (своего рода внутренние) процедуры ведения фонда, а затем разрабатываются средства для возможного внешнего подключения прикладных программ пользователя по заранее не предусмотренным обращениям к данным, т. е. обеспечиваются возможности расширения использования проектируемого информационного фонда.

Рассмотренный нами наиболее простой вариант организации единого информационного фонда не всегда рационален. В нем остаются следы того этапа применения вычислительной техники, для которого было характерно использование локальных массивов данных. Более высокий уровень организации фонда связан с созданием унифицированных, максимально интегрированных наборов данных, для накопления, обновления и использования которых можно построить в определенной степени стандартизированные операционные системы в виде пакетов прикладных программ. Такие наборы данных также могут быть организованы по файловой структуре, но в более компактном, удобном с точки зрения обслуживания и использования для различных приложений виде, с большими возможностями модификации и расширения информационного фонда.

Информационный фонд такого типа, как уже отмечалось, может создаваться с использованием уже разработанных типовых структур, предполагающих применение конкретной системы управления базой данных, или разрабатываться как оригинальный набор данных, для формирования и ведения которого параллельно проектируются программные средства. В основу организации таких фондов (как типовых, так и оригинальных) закладываются различные принципы. Для примера рассмотрим один из возможных вариантов построения единого информационного фонда.

Фонд состоит из двух частей: главного (генерального, центрального) файла, содержащего накапливаемые и корректируемые

данные и допускающего произвольный доступ к ним, и служебного (организационного) файла, содержащего сведения о составе и структуре накопленных в главном файле данных. Наименьшей единицей данных является элемент (поле), характеризующийся именем, форматом и значением. Имена и форматы устанавливаются пользователем путем описания данных, а значения присваиваются в процессе накопления и обновления данных в файле. Каждому объекту, представляемому в информационном фонде, соответствует определенная совокупность элементов, отражающих различные его характеристики. Все элементы, относящиеся к одному объекту, представляют собой запись данных, а совокупность записей, описывающих множество однородных объектов и имеющих одинаковый фиксированный набор элементов, — файл данных. Таким образом, главный файл фонда может содержать несколько внутренних файлов, определенным образом связанных друг с другом.

Для однозначной идентификации каждого объекта вводится ключевой признак, которым может быть любой уникальный для данного типа объектов элемент. Если в главный файл входит несколько внутренних файлов, то ключевой признак должен быть уникальным во всем объеме главного файла.

Запись любого файла состоит из ключевого признака и совокупности элементов, которые располагаются в ней в возрастающем порядке символических имен. Каждый элемент здесь представляется именем и значением. В каждую запись может включаться дополнительный отличительный признак — номер записи, а также обозначение вида записи, позволяющее различать записи различных форматов. Следует отметить, что на запоминающем устройстве записи целесообразно размещать в сжатом виде, т. е. с исключением из состава записи пустых элементов, имеющих значение пробела или нуля, а также с представлением значений элементов по фактической значности (числовые значения — без незначащих нулей, текстовые — без завершающих пробелов). В связи с этим файлы могут иметь записи переменной или неопределенной длины. В то же время занесение файлов на магнитный диск производится блоками постоянной длины, т. е. количество записей, размещаемых в одном блоке, неодинаково и неопределенно.

Служебный файл предназначен для управления главным файлом. Он, как правило, состоит из трех частей: описания элементов, списка полных названий элементов и описания ключевых признаков. В описании элементов содержатся сведения об их форматах, которые нужны для определения структуры записей главного файла. Для каждого элемента в нем указываются символическое имя и длина значения, а также другие сведения специального назначения. Полные названия элементов являются синонимами их символических имен и предназначаются для облегчения оперирования элементами вне информационного фонда. Отдельный список полных названий может не составляться, если они включаются в описание элементов. Описание ключевых признаков строится по принципу индексно-последовательного метода организации данных.

Оно служит для облегчения поиска данных по значениям ключевых признаков. Практически такое описание представляет собой набор индексов, организуемых на одном или нескольких уровнях. В записи индекса блока указываются, например, относительный номер блока и значение ключевого признака последней записи главного файла, размещенной в данном блоке. Кроме того, следует предусмотреть возможность обращения к таким записям главного файла, ключевые признаки которых неизвестны. Для этого необходимо предварительно построить индексную таблицу для других элементов записи, которые должны выполнять роль ключей прямого доступа, т. е. использоваться для установли связи между значением элемента и адресом соответствующей записи. Следует отметить, что изменение значений ключей прямого доступа требует много времени для исправления индексных таблиц, поэтому принятие решения о включении элемента в список ключей прямого доступа должно быть обосновано тщательным анализом соотношения количества операций поиска по данному элементу и операций изменения его значения. При этом следует учитывать и то обстоятельство, что в информационном фонде всегда имеется возможность производить поиск записей главного файла по значению любого элемента, не являющегося ключом, путем последовательного просмотра всех записей файла.

Работы по созданию информационного фонда рассматриваемого типа начинаются составлением и вводом описания элементов, после чего можно приступить к формированию главного файла в соответствии с предварительно подготовленным описанием структуры его записей. Для этого используются входные данные, как правило, зафиксированные на перфоносителях или магнитной ленте. К ним предъявляются следующие требования:

входной файл должен быть построен из записей жесткого формата, максимальную длину которого определяет тип применяемого носителя;

элементы, входящие в состав записи, должны иметь жесткий формат и всегда занимать одно и то же место в записи;

каждая запись должна содержать обозначение вида записи и значение ключевого признака.

Одновременно с формированием главного файла, а также при последующих его изменениях соответственно создается или корректируется описание ключевых признаков, в том числе и таблица индексов поиска по ключам прямого доступа.

С разработкой программных средств управления данными определенного информационного фонда может быть построена сравнительно простая база данных, которую целесообразно создавать для небольших АСУ при ограниченном количестве объектов, представляемых в главном файле. Для более сложных систем обработки данных информационный фонд такого типа становится слишком громоздким (особенно в части служебного файла), что приводит к значительным затратам на ведение и использование фонда, к неэффективному применению средств вычислительной техники.

Более мощные и гибкие базы данных могут быть построены для таких информационных фондов, которые имеют логическую структуру, позволяющую с большей точностью и полнотой отражать множество различных характеристик объектов, тесно связанных друг с другом в самых разнообразных аспектах. Практически самое широкое распространение получили информационные фонды с иерархической структурой данных, в которых используются различные методы реализации информационных связей, формирования, обновления и прикладного использования данных.

Как известно, все типы структур могут быть представлены в виде деревьев или сетей, которые, в свою очередь, практически всегда могут быть разбиты на группы нормализованных (двумерных) файлов с некоторыми избыточными элементами данных. Таким образом, имеется возможность создания реляционных информационных фондов, скажем, на уровне третьей нормальной формы, которая соответствует установлению элементарных отношений (зависимостей) между элементарными данными [28]. Вопросы организации таких фондов непосредственно связаны с проблемой создания баз данных и АБД, поэтому они рассматриваются в следующем параграфе. Следует подчеркнуть, что единые информационные фонды широкого назначения, являющиеся основой мощных баз данных АСУ, как правило, создаются только с использованием готовых систем управления такими базами. При разработке таких фондов должен решаться вопрос выбора наиболее подходящей системы и настройки ее с учетом конкретных условий пользователей, которые могут быть выражены в различных способах построения логических структур данных и их отображения на физическом уровне, методах индексации и адресации, организации ведения базы данных и т. п.

10.3. Принципы организации единых баз данных

База данных как интегрированная система информационного обеспечения АСУ, включающая в себя собственно информационный фонд, а также программную систему формирования и организации информационных связей этого фонда, должна отвечать следующим основным требованиям:

- 1) интегрированного хранения и централизованного (коллективного) использования данных;
- 2) независимости данных и прикладных программ;
- 3) неизбыточности данных;
- 4) создания и поддержания связей между данными;
- 5) наличия средств обслуживания базы данных.

Главная особенность базы данных — это способность организовывать файлы специальной структуры, в которых связи между данными реализуются внутри записей и между записями. Организация связи внутри записей — это используемый способ физического объединения данных по некоторому ключевому признаку в одну запись. Связь между записями реализуется при помощи адресных

ссылок, которые могут объединять записи, физически расположенные в разных местах одного файла, а также в разных файлах.

Для создания баз данных, как уже отмечалось, используются готовые системы управления ими, обладающие высокой степенью универсальности. Основные свойства таких систем — их адаптивность, инвариантность, модифицируемость и расширяемость. Под адаптивностью понимается возможность приспособления к конкретным условиям пользователя, к изменениям информационного фонда с минимальными изменениями в самой системе управления. Инвариантность обеспечивает возможность выбора определенных элементов системы управления базой данных из нескольких возможных вариантов, а модифицируемость — гибкость системы и дополнительные удобства для пользователей, позволяющие им разрабатывать новые версии отдельных элементов системы. В таких системах обязательно предполагается, что для более полного удовлетворения нужд пользователей они будут развиваться дальше и постепенно дополняться новыми элементами. В этом заключается их расширяемость.

Рассмотрим организацию системы данных наиболее широко используемых баз, имеющих иерархическую структуру. Такие интегрированные базы строятся из информационных файлов двух типов — главных и связующих.

В простейшем виде главный файл представляет собой набор записей, имеющих общий идентификатор и снабженных уникальными идентификаторами, однозначно определяющими каждую запись в данном файле. В запись входит совокупность полей данных, логически связанных с определенным идентификатором (ключом записи). Для главных файлов должно соблюдаться условие единообразия ключевых полей (одинаковые размеры и расположение), а форматы различных записей одного и того же файла могут быть различны. Таким образом, имеется возможность в рамках одного файла объединять записи, представляющие различные информационные объекты. Однако обычно в базу данных включается не один, а несколько главных файлов. При этом исключается дублирование данных. Вместе с тем различные информационные объекты могут иметь одинаковые характеристики. Для отнесения таких характеристик к определенным объектам организуются связующие файлы, являющиеся, во-первых, средством организации связей между данными внутри главных и связующих файлов и, во-вторых, выполняющие функции логического расширения главных файлов. Связующие файлы чаще всего имеют списковую структуру, т. е. представляют собой совокупность групп (списков) записей фиксированного формата, начало которых указывается в соответствующих записях главных файлов, а каждая запись группы содержит адресную ссылку на следующую запись той же группы. Таким образом, все записи одной группы связываются в цепочку.

Как видим, главный файл может быть связан с подчиненным ему связующим файлом, и эти связи используются при расширении записей главного файла. Тогда группа записей связующего файла,

общая для нескольких записей главного файла, подчиняется этому главному. Если же несколько групп записей связующего файла должны быть отнесены к одной записи главного файла, то подчиненным является главный файл. Наконец, для пары файлов (главного и связующего) можно устанавливать перекрестные связи, когда одни записи главного файла связаны с другими его записями. В данном случае оба файла по отношению друг к другу являются подчиненными. При этом необходимо соблюдать следующие условия: никакая запись главного файла, которой подчиняется определенная группа записей связующего файла, не может быть, в свою очередь, подчинена этой же группе записей. Следует отметить, что один главный файл может связываться практически с любым количеством связующих файлов и наоборот. Именно поэтому связующие файлы кроме функции простого логического расширения записей главных файлов могут выполнять свою основную функцию — устанавливать связь между записями нескольких главных файлов. Так, в случае двух главных и одного связующего файлов все они могут быть взаимоподчиненными или, например, связующий файл подчиняется первому главному, а второй главный — связующему файлу. Связи между файлами определяют пути поиска данных. Очевидно, что поиск на каждом участке такого пути направлен к подчиненному файлу.

Данные в главных файлах, формируемых на магнитных дисках, как правило, организуются с использованием одного из двух методов: метода доступа с управляемой последовательностью и метода доступа с отдельным индексом.

Первый метод, во многом схожий со стандартным индексно-последовательным методом доступа, учитывает особенности использования главных файлов в условиях существования взаимосвязей между данными. Кроме индексов трех уровней (главного, цилиндров и дорожек) для каждого файла формируется специальная управляющая запись, помещаемая за индексом дорожек первого цилиндра данных. Минимальная длина этой записи — 80 байт. В нее включаются данные, необходимые для управления последовательностью поиска (отсюда и само название метода): дисковые адреса индексов файла данных, области переполнения, количество записей, количество обращений к файлу и др. Управляющие записи позволяют ускорить процесс поиска нужной информации. Этот метод целесообразно использовать для организации небольших и не часто корректируемых файлов, так как значительное обновление файлов и особенно добавление новых записей в некоторой степени дезорганизует их, значительно увеличивая время доступа к данным. Реорганизация таких файлов представляет длительный и трудоемкий процесс, особенно при большом объеме реорганизуемого файла.

Второй метод доступа также основан на системе индексов трех уровней, но в отличие от первого на самом низшем уровне используется не индекс дорожек, а отдельные индексные ссылки на каждую запись данных. В связи с этим наряду с файлом дан-

ных формируется файл ссылок, каждая запись которого соответствует одной записи файла данных, а в систему индексов файла входят только индексы цилиндров и дорожек. Файл ссылок в этом случае организуется методом с управляемой последовательностью. Файлы с отдельным индексом нужно подвергать реорганизации значительно реже, чем файлы с управляемой последовательностью, но вместе с тем они требуют большего объема памяти, а поиск данных в них осуществляется медленнее. В связи с этим такой метод целесообразно использовать для организации особо крупных или очень часто корректируемых главных файлов.

Связующие файлы записываются на магнитных дисках последовательным методом с оставлением на каждом цилиндре свободных полей для добавления новых записей. Группы записей, относящиеся к одной записи главного файла, связываются между собой адресными цепочками, т. е. каждая запись содержит адрес предыдущей или следующей записи в группе. В связи с тем, что такой группе записей может быть подчинена запись главного файла или, наоборот, сама эта запись может ей подчиняться, в связующих файлах организуются соответственно подчиненные или первичные цепочки. Каждая запись связующего файла входит одновременно в цепочки обоих видов и поэтому содержит по две адресные ссылки. Первичные цепочки строятся по заданным полям управления в последовательности, определяемой пользователем. Подчиненные цепочки формируются в виде списка с обратной очередностью. Поэтому структура связующего файла такова, что все записи первичной цепочки физически размещаются рядом, практически на одном и том же цилиндре, а записи подчиненной цепочки разбросаны и могут располагаться на разных цилиндрах. Учитывая данное обстоятельство, при проектировании главных и связующих файлов необходимо правильно определить частоты входа (с целью прослеживания цепочек) в связующий файл со стороны обоих главных файлов и назначить подчиненным тот из них, через который предполагается реже обращаться к связующему файлу. Это позволяет ускорить наиболее частый поиск данных по направлению: первичный главный файл → первичная цепочка подчиненного ему связующего файла → главный файл, подчиненный связующему.

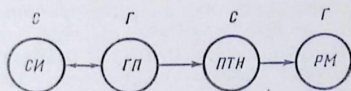


Рис. 10.2. Структура информационного фонда системы СИОД1 и СИОД2 (С — связующие файлы, Г — главные файлы)

Рассмотренная нами организация базы данных использована в системах интеграции и обработки данных СИОД1 и СИОД2, стандартная структура информационного фонда которых представлена на рис. 10.2 (стрелки указывают направление в сторону подчинения). В состав фонда входят следующие файлы:

ГП — главный предметный файл, содержащий данные об изделиях и материалах;

СИ — связующий файл состава изделий (спецификации изделий);

ПТН — связующий файл пооперационных трудовых нормативов (описание технологических маршрутов обработки изделий);

РМ — главный файл рабочих мест (описание оборудования рабочих мест).

Такая структура базы данных может быть использована для ведения нормативно-справочной информации на промышленных предприятиях. Стандартная система управления выполняет обслуживание и реорганизацию информационного фонда, функции которой сводятся к следующему:

- 1) добавлению и исключению отдельных записей и первичных цепочек;
- 2) замене отдельных записей связующих файлов;
- 3) дублированию первичных цепочек;
- 4) установлению и разрыву связей между соподчиненными записями;
- 5) пометке записей главных файлов;
- 6) восстановлению соответствия между логической и физической последовательностью записей в главных файлах с одновременной корректировкой адресных ссылок в связующих файлах и наоборот (реорганизация).

В данной системе также предусмотрены средства прикладного программирования, позволяющие пользователям разрабатывать уникальные программы по поиску и корректировке данных, а также по обслуживанию файлов. Имеется возможность подключать программы пользователя, написанные на языках Ассемблера, Кобола и PL/I.

Представителями более высокого класса иерархических баз данных универсального назначения являются системы ОКА и КАМА. Они предусматривают централизованное формирование, накопление и обновление данных, а система КАМА включает в себя и средства телеобработки информации с использованием дисплеев. Информация, включаемая в базу данных, описывается на информационном языке предкоординатного типа, в котором все возможные информационные высказывания идентифицируются с помощью составных кодовых обозначений. При этом используется специальный тезаурус, позволяющий зафиксировать в его терминах наиболее существенные отношения между данными. Структура ведения диалога предусматривает последовательный вызов на экран дисплея информации, содержащей альтернативные варианты поиска, при этом наименования данных представляются в виде укрупненных элементов информационного языка, т. е. смысловых фасетов.

Несколько иначе организуется база данных с использованием широко применяемой системы ИНЭС. Информационный фонд в ней представляет собой древовидный граф. Ограничения на размеры графа и глубину иерархии накладываются пользователем. В ходе перебора данных можно двигаться по дереву не только вниз, но

и вверх, влево, вправо. Данные описываются на логическом и физическом уровне с формированием дерева описания данных и дерева данных. Для исключения дублирования описания данных имеются средства описания по шаблону. Система обеспечивает ввод, накопление, интегрированное хранение, реструктуризацию, обновление, выборку и вывод данных на печать и на экран дисплея. Взаимодействие пользователя с базой данных осуществляется с использованием языка запросов. Формализованные запросы могут каталогизироваться в специальной библиотеке, а при необходимости пользователь с помощью дисплея может выбирать из них нужные и получать на них ответы.

Для обслуживания пользователей в диалоговом режиме перспективными являются базы данных, построенные на основе инвертированных файлов. В общем случае инвертированный файл — это такой файл, который хранит идентификаторы объектов, связанные с конкретным значением определенного их признака. Примером такой базы данных может быть система TDMS (или усовершенствованный вариант CDMS), разработанная американской фирмой SDC. В совокупности файлов данной системы, предназначенных для хранения данных иерархической структуры, входят словарь элементов данных, индекс таблицы соответствия, список экземпляров, справочник, таблица элементов данных и таблица значений. Поиск значений нужных элементов данных осуществляется путем перехода от одного инвертированного файла к другому в такой последовательности, как эти файлы перечислены выше.

Для организации баз данных в перспективе наряду с другими методами, вероятнее всего, будут использоваться структуры бинарных отношений, размещаемые в ассоциативной памяти системы. Как известно, схему базы данных любой сложности можно представить с помощью только бинарных отношений, а ответы на сложные вопросы можно формировать таким образом, чтобы в каждый момент времени обрабатывалась только одна пара элементов данных. В ассоциативной памяти, в которой размещение записи и ее выборка определяются не машинным адресом, а ее содержанием, признак поиска служит входом, а выход на каждую запись файла содержит по одному биту, характеризующему идентичность данной записи признаку поиска (бит, равный единице, означает совпадение записи с признаком поиска, а равный нулю — несовпадение). При этом можно проводить сравнение записей файла только по некоторым выбранным битам, т. е. с маскированием признака поиска. Такие системы могут обеспечить надежное функционирование баз данных в режиме реального времени.

Перспективное направление организации АБД представляет собой концепция построения орбитальной структуры базы данных. Суть ее сводится к следующему. В связи с тем, что логическая структура базы данных должна отражать не только содержательный аспект ее элементов, но и их состояние, любой простой или сложный элемент базы данных (показатель, их группу, запись,

файл) можно представить как динамическую точку, движущуюся по кольцевой орбите в определенном направлении. На такой орбите каждый элемент проходит три зоны: формирования или корректировки (ФК); активного состояния использования (АИ); пассивного состояния хранения (ПХ).

При этом чередование переходов из одной зоны в другую может происходить по часовой стрелке (ФК → АИ → ПХ) или против часовой стрелки (ФК → ПХ → АИ). Такое представление структуры базы данных, несомненно, носит идеализированный характер.

Для характеристики состояния идеализированного элемента базы данных достаточно указать: 1) орбитальный радиус; 2) направление движения; 3) три точки перехода в зоны ФК, АИ и ПХ относительно некоторой начальной точки.

Начальные точки всех орбит должны лежать на одном радиусе системы — начальном радиусе базы, если при этом известна постоянная для всех элементов базы данных радиальная скорость. Система управления базой данных по этим характеристикам должна определить требуемое состояние каждого конкретного элемента в каждый фиксированный момент времени и обеспечить это состояние. Очевидно, что многие элементы базы данных могут иметь одинаковые характеристики состояния. Такая идеализированная орбитальная структура базы данных позволяет системно выполнять физическое перемещение ее элементов по уровням иерархической памяти системы и организовать их комплексное использование системными и прикладными программами в зонах ФК и АИ.

Рассмотрим основные цели обращения к элементам базы данных в зоне ФК. Формирование нового элемента может осуществляться путем включения его извне, т. е. путем ввода, или же оно сводится к его получению расчетным путем с использованием других элементов базы, одновременно находящихся в зоне АИ. Из принятого принципа орбитального движения элементов базы данных вытекает, что все ранее включенные в базу элементы на каждом обороте подвергаются корректировке, которая сводится к одной из трех процедур:

- 1) удаления в режиме ликвидации или переноса в архив данных для временного хранения;
- 2) собственно корректировки (добавления, замены, удаления);
- 3) сохранения прежнего значения элементов с возможным контролем его соответствия другим, связанным с ним элементам базы данных.

Для преобразования данных (выполнения расчетов) и формирования результатов могут использоваться элементы, одновременно находящиеся в зоне АИ. При этом пользователи системы могут обращаться к активным элементам с целью поиска нужных данных по соответствующим запросам или для решения определенных задач путем выполнения соответствующих прикладных программ.

Элементы базы данных, находящиеся в зоне пассивного состояния хранения (ПХ), физически могут быть перенесены на носитель более низкого уровня. Так, например, если элементы в

состоянии ФК и АИ хранятся на магнитных дисках, то при переходе в состояние ПХ они могут перезаписываться на магнитную ленту.

Орбитальные кольца с одинаковым радиусом, имеющие различные точки перехода из одной зоны в другую, т. е. зонально сдвинутые по отношению друг к другу, образуют логический цилиндр базы данных, все элементы которого характеризуются одинаковой линейной скоростью. Кольца же, имеющие различный радиус, могут располагаться в одной плоскости, образуя логическую секцию базы данных. При этом в одну секцию целесообразно объединять орбиты элементов, имеющих совпадающие точки перехода из одной зоны в другую, т. е. характеризующихся такими точками переходов, которые располагаются на одних и тех же радиусах секции.

Рассмотрим соотношение между логической и физической организацией орбитальной базы данных. Если посетителем данных, находящихся в зонах ФК и АИ, являются пакеты магнитных дисков, то можно представить, что все участки зон ФК и АИ одной логической секции базы данных на физическом уровне соответствуют цилиндру пакета магнитных дисков (в общем случае несколько секций может быть записано на одном распределенном цилиндре или, наоборот, одна секция может располагаться на нескольких, например, одноименных цилиндрах различных пакетов магнитных дисков). Это позволяет одновременно обращаться ко всем элементам базы данных, находящимся в активном состоянии, без перемещения магнитных головок считывания и записи. Таким образом, путем исключения зоны ПХ из логической структуры секций базы данных и сжатия активных участков орбитальных колец устанавливается относительное соответствие между логической и физической структурой базы данных. Исключенные зоны ПХ, как уже отмечалось, могут быть представлены элементами базы данных, записанными на магнитных лентах. При этом все зоны ПХ одного логического цилиндра (а не одной секции) целесообразно формировать в виде отдельного файла, так как в данном случае такие файлы характеризуются одинаковым периодом обращения к ним (с целью записи на магнитную ленту или считывания при переходе элементов в другое состояние).

Для повышения динамичности орбитальной структуры базы данных с учетом того, что некоторые элементы за один период обращения могут менять как линейную скорость, так и направление движения, в системе управления базой данных необходимо предусмотреть возможность перехода элементов из одной орбиты в другую. Переход может осуществляться как по вертикали (внутри секции), так и по горизонтали (внутри цилиндра), при этом для упрощения его организации целесообразно ввести ограничение перехода только на ближайшие соседние орбиты. Такие «блуждающие» элементы, которые могут двигаться не только по плоскостной, но и по трехмерной пространственной траектории, позволяют в каждый фиксированный момент времени обеспечить

нужное состояние базы данных, что не всегда возможно в случае идеализированной структуры орбитальной базы. Однако всегда следует стремиться минимизировать количество таких элементов, так как они вносят дополнительные трудности в организацию базы данных. Прежде всего эти трудности связаны с согласованием характеристик соседних орбитальных колец, из определенных участков которых формируется траектория движения «блуждающих» элементов.

Для ЭВМ четвертого поколения — многопроцессорных вычислительных комплексов «Эльбрус» база данных может быть построена из отдельных файлов, каждый из которых содержит данные и атрибуты файла. В систему файлов вводится адресная информация или межфайловые ссылки, что позволяет обеспечить контекстную защиту данных. Все ссылки необходимо сосредоточить в заголовках файлов, которые собираются в один общий файл, единый для всех магнитных барабанов и каждого семейства накопителей на магнитных дисках. В любом заголовке содержится специальный переключатель, состоящий из межфайловских ссылок, каждая из которых — это относительный адрес заголовка файла в файле заголовков. Таким образом, любой объект в системе файлов может быть достигнут только через ссылки. Очевидно, что на любой объект может быть сколько угодно ссылок из различных переключателей файлов и справочников, на основе которых с использованием механизма входа в базу данных создается система формирования контекстов для конкретных пользователей.

ВЫБОР И ПРОЕКТИРОВАНИЕ ПРОЦЕССОВ
ОБРАБОТКИ ДАННЫХ

Глава 11

ТИПОВЫЕ ПРОЦЕССЫ И ОПЕРАЦИИ МАШИННОЙ ОБРАБОТКИ ДАННЫХ

11.1. Общая характеристика типовых
технологических процессов и операций

Движение и преобразование данных осуществляются путем выполнения различных операций. Каждая операция — это действие или их комплекс, которые выполняются над данными или их носителями на одном рабочем месте, устройстве (комплексе устройств) и в рассматриваемом аспекте берется как элементарная часть более общего процесса. Так, в процессах обработки первичных документов на клавишных вычислительных машинах такими элементарными действиями (операциями) могут быть арифметические операции над реквизитами и запись полученных итогов (суммирования, таксировки и др.), группировка документов, составление на счетно-табличных машинах сводных и других документов. При обработке данных на перфорационных вычислительных машинах могут быть выделены такие операции, как регистрация первичных данных, их передача, приемка для машинной обработки, запись на машинный носитель, составление результатных сводок (табуляграмм) и их выпуск. Применению электронных вычислительных машин обычно сопутствует использование средств сбора, передачи и отображения данных, поэтому в технологическом процессе появляются операции, выполняемые с помощью этих средств. По своему содержанию и способу выполнения различаются разнообразные операции обработки данных на ЭВМ.

Совокупности операций принятого варианта технической реализации конкретной работы (процесса), выполняемых в установленной последовательности, образуют соответствующие технологические процессы обработки данных. Понятия операции и технологического процесса являются в определенном смысле условными, так как вопрос о том, что назвать технологическим процессом и какие выделить в нем операции, во многом решается с учетом не только объективных, но и субъективных факторов.

Количество, характер и содержание выделяемых операций технологического процесса в каждом конкретном случае могут существенно различаться. Их выделение в первую очередь связано

с оценкой таких факторов, как вид и форма представления исходных и формируемых данных, разнообразие и сложность экономико-математических моделей и алгоритмов их реализации, возможности и эксплуатационные характеристики применяемых средств вычислительной техники, организация работы на ИВЦ, требования по достоверности данных, используемые методы и средства обеспечения этой достоверности и др. При проектировании технологических процессов в виде последовательности операций обработки данных обычно выделяются те части общего процесса, для выполнения которых возможно и целесообразно разработать отдельные инструкции, программы или другие указания. Например, такие инструкции разрабатываются для перфорации (переноса) данных с первичных документов на перфокарты (перфолисты) или записи на магнитные носители при помощи клавишных устройств (перфораторов, счетно-табличных машин, агрегированных с перфоприспособками, и др.).

Каждая операция, предназначенная для выполнения определенной работы (части процесса), может иметь различные варианты ее реализации. Поэтому операция технологического процесса характеризуется ее назначением (выполняемыми функциями в процессе обработки данных), методами и средствами реализации, а также качественными и количественными параметрами ее выполнения и получаемыми результатами (затратами на единицу исходных, обрабатываемых или получаемых данных, возможностью обеспечения требуемой точности и достоверности и т. п.). На основании этих характеристик можно разработать рациональные варианты технологических процессов обработки данных. При этом целесообразно использовать такой подход: принять, что имеется определенное множество операций технологического процесса с соответствующими их характеристиками и задание из этих операций построить технологический процесс, позволяющий из исходных данных получить желаемые результаты. Очевидно, что таких технологических процессов может быть построено много: для тех же целей служат различные операции и их сочетания. Поэтому перед решением поставленной задачи необходимо заранее ограничить количество всевозможных операций, т. е. построение вариантов нужно начать с выбора сравнительно небольшого количества операций и их сочетаний и выделения их как основных представителей и типовых элементов технологических процессов.

Типизацию операций технологических процессов обработки данных можно успешно осуществить только на основе четко построенной системы их классификации. Такая система должна помочь наиболее полно определить основные характеристики каждой операции, роль и место ее в общем процессе обработки данных. Сопоставление характеристик различных вариантов позволяет выбрать наиболее пригодные для технологического процесса операции, определить аналогичные характеристики всего построенного варианта этого процесса как типового. Для этих целей классификацию операций необходимо проводить по следующим наиболее

важным признакам: степени автоматизации, применяемым средствам вычислительной техники, этапам процессов движения и преобразования данных, применяемым носителям данных, методам и средствам обеспечения достоверности и др.

По степени автоматизации различаются ручные, машинно-ручные, механизированные, автоматизированные, полуавтоматические и автоматические операции. Ручной способ, несмотря на недостатки, остается пока основным способом обработки документов на клавишных вычислительных машинах и средствах подготовки промежуточных машинных носителей (формирования первичных массивов данных). При формировании массивов чтение исходных данных с первичных документов, их перфорация (запись) на машинный носитель и контроль правильности записи осуществляются человеком. При подготовке массивов на перфокартах используются перфораторы и контрольные, при их подготовке на магнитных носителях — специальные средства записи данных и контрольно-считывающие устройства. Скорость записи данных на машинные носители при этом небольшая, она составляет обычно 1—3 знака в секунду.

По существу, ручными (точнее, машинно-ручными) являются почти все операции, выполняемые на клавишных и перфорационных вычислительных машинах основного комплекта. Эти машины обычно имеют самостоятельное значение в том смысле, что на них производится вся обработка данных, подготовка массивов на машинных носителях (перфокартах), а также составление необходимых ведомостей (формирование и выдача результатных данных) в виде табуляграмм. Кроме того, КВМ и ПВМ могут применяться и при подготовке данных к последующей их обработке на ЭВМ. Например, вместе с ЭВМ довольно широко используются бухгалтерские, фактурные, пишущие и другие машины с перфопроставками, клавишные перфораторы, контрольные, сортировальные машины. Для составления контрольных табуляграмм могут также применяться табуляторы.

Ручной способ выполнения работ (и особенно подготовка массивов первичных данных) имеет существенные недостатки, в первую очередь невысокую достоверность и сравнительно небольшую скорость работы. Это требует их совершенствования, т. е. автоматизации. Автоматизация подготовительного этапа обработки данных осуществляется в следующих основных направлениях: предварительная запись постоянных признаков на машинные носители, что позволяет эту операцию выполнить в период, когда машины не загружены работой и можно тщательно проверить правильность такой записи; автоматическое изготовление машинных носителей одновременно с подготовкой первичных документов (с использованием бухгалтерских, фактурных, контрольно-регистрающих и других машин, агрегированных с ленточными и другими устройствами); автоматическое изготовление машинных носителей в процессе поступления данных по каналам связи или непосредственно от датчиков и контрольно-измерительных приборов.

Типизация операций и технологических процессов связана уменьшением множества возможных вариантов и оставлением только тех, которые удовлетворяют определенным заданным условиям и практически наиболее просто могут быть реализованы при обработке конкретной информации. Например, типовые операции следует предусмотреть только для наиболее часто используемых носителей данных: документов на бумаге, жетонов, устройств, фиксирующих сообщения, перфокарт, перфолент, магнитных лент и дисков. Среди первичных документов можно выделить те, которые подготавливаются отдельно и вместе с машинными носителями, а результатные документы разделить на табуляграммы, составляемые на различных технических средствах: бухгалтерских, фактурных машинах, табуляторах, ЭВМ (их печатающих устройствах), на копировальной или множительной аппаратуре, устройствах отображения данных. Выделение таких разновидностей документов связано с особенностями способов и форм представления данных и существующими различиями в их формировании при использовании соответствующих технических средств.

Операции могут быть исполнительными и контрольными, причем последние выполняются автономно или совмещаются с соответствующими исполнительными операциями. С учетом всех рассмотренных факторов в качестве типовых ручных и машинно-ручных операций обычно выделяются следующие:

1) операции над носителями данных (приемка, сдача заказчику, регистрация, комплектование, передача на хранение, хранение, возвращение);

2) подготовка первичных документов (кодирование и запись, арифметическая обработка, контроль арифметической обработки, подсчет контрольных сумм, регистрация контрольных сумм);

3) запись на машинный носитель (перфорация, верификация, визуальный контроль, исправление (перебивка), составление развернутой контрольной табуляграммы, составление итоговой контрольной табуляграммы, перезапись, ввод в ЭВМ, контроль ввода);

4) операции по обеспечению достоверности и сохранности данных на носителях (формирование и запись контрольных сумм, выяснение ошибок, визуальная проверка: построчное сличение двух документов, ленты суммирующей машины и документов или табуляграммы, двух табуляграмм, документов и перфокарт);

5) проверка и выпуск табуляграмм, составленных на ЭВМ, автономных устройствах печати, ПВМ (табуляторе), КВМ (бухгалтерской или фактурной машине).

На рис. 11.1 представлены фрагменты технологических процессов, характеризующие типовое начало и окончание обработки данных (здесь же конкретизированы машинные носители информации).

Технологический процесс обработки данных можно изобразить в виде связного ориентированного графа, вершинам которого поставлены в соответствие отдельные операции или носители данных. В таком графе вершины могут быть расположены последо-

вательно или параллельно. Поэтому и процессы обработки данных на ЭВМ можно определить как последовательности операций движения и преобразования данных, записанных на машинных носителях. Они представляют собой в определенном смысле автономные группы элементарных операций и процедур соответствующих алгоритмов, а каждый алгоритм — детализацию отдельной операции технологического процесса или его части. Построение таких про-

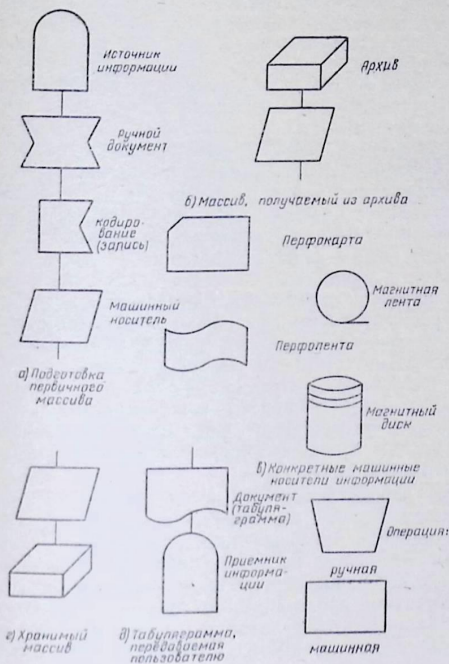


Рис. 11.1. Начало и окончание процессов обработки и носители данных

цессов, выделение и выбор операций, а также основные пути их реализации непосредственно связаны с разработкой и использованием генераторов ввода-вывода, пакетов прикладных программ и других средств программного обеспечения. При этом следует отметить, что они должны рассматриваться как процессы, включающие два этапа: 1) формирование и ведение файлов данных на машинных носителях; 2) формирование и выдача результатов данных.

Выделение типовых операций обработки данных на ЭВМ, т. е. автоматических операций, на современном этапе связано с типизацией и стандартизацией средств программного обеспечения, с применением средств автоматизации разработки (генерации) программ, широким использованием ранее разработанных системных и прикладных программ, их пакетов. Более подробно эти вопросы рассмотрены в 11.3, а также в гл. 12, 13.

Типовые технологические процессы наиболее широко применяются при проектировании использования клавишных и перфорационных вычислительных машин, а также при подготовке массивов данных к их обработке на ЭВМ. Анализ возможных вариантов таких технологических процессов представлен в [8]. Варианты технологических процессов объединяются в группы с учетом разновидностей используемых машин, приставок к ним, степени механизации и автоматизации сбора, предварительной обработки исходных данных и формирования результатов, использования массивов нормативно-справочных данных. Например, для КВМ и средств регистрации (записи) данных в [8] выделяются группы технологических процессов с учетом применения следующих машин: бухгалтерских и фактурных — для составления сводок, выписки счетов-фактур и других документов — с перфоленточными приставками и без них; пишущих (организационных) автоматов; кассовых аппаратов с перфоленточными приставками; устройств сбора первичных данных; читающих автоматов; устройств передачи данных по каналам связи. Для ПВМ выделяются группы технологических процессов с учетом использования следующих машин: основного комплекта ПВМ; итоговых перфораторов; приставок к табулятору; электронных вычислительных перфораторов; считывающих автоматов; любых ПВМ с использованием массивов постоянных (нормативно-справочных) данных.

Выделение рассмотренных технологических процессов способствует унификации работ по обработке данных в машинно-счетных бюро, на станциях и в других подразделениях, механизации и автоматизации отдельных расчетов. Такие технологические процессы в большинстве случаев пригодны и для комплексной механизации планово-учетных работ, но их применение в АСОД ограничено. Основная причина этого заключается в том, что они разработаны как отдельные звенья без учета их использования в системах, т. е. в условиях единых баз данных и организации потоков заданий; в них недостаточно отражены процессы по записи данных непосредственно на магнитные носители и по обработке на ЭВМ третьего поколения с использованием современных средств программного обеспечения; отдельные операции и процессы недостаточно увязаны с обеспечением достоверности результатов и затратами на их выполнение, что не позволяет экономически обосновать выбор наиболее рационального варианта. Поэтому основное направление работ по типизации технологических процессов заключается в установлении перечисленных недостатков.

Типизации подлежат не только сами технологические процес-

сы, но и документация по их описанию, графические схемы их представления. Последние, как правило, строятся для каждой работы по формированию первичных, промежуточных или результатов данных.

Разнообразие операций обработки данных в АСОД порождает большое количество технологических процессов. Однако во многих случаях в них можно выделить определенные аналогичные части. Обычно выделяются следующие последовательности операций, составляющие общий типовой технологический процесс выполнения работ на всех этапах движения и преобразования данных: сбор и запись (фиксирование) первичных данных на объекте управления; подготовка первичных документов (а также данных, непосред-

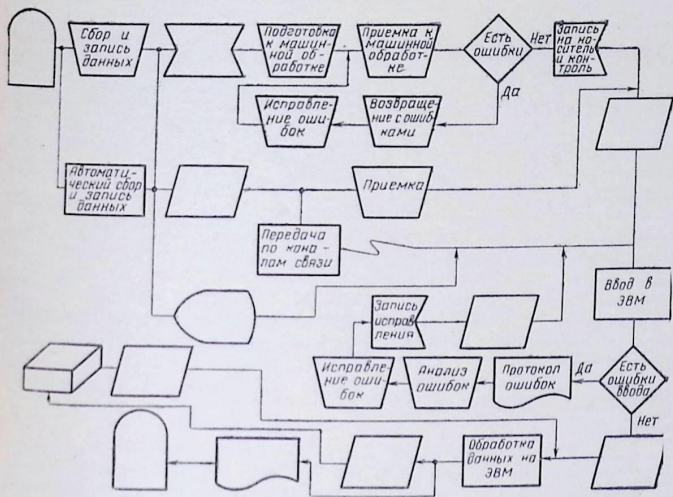


Рис. 11.2. Схема технологического процесса движения и преобразования данных

венно зафиксированных на машинных носителях) к передаче на машинную обработку и их передача, включая передачу по каналам связи; прием и учет данных, включая запросы от абонентов на выдачу справок и результатов расчетов; контроль достоверности принятой информации и устранение обнаруженных ошибок; запись данных на машинные носители и проверка правильности записи (кодирования); ввод данных с машиночитаемых носителей, включая запросы, поступающие по каналам связи и с местных терминалов, проверка правильности их ввода; обработка данных на ЭВМ и вывод результатов на печать, машинные носители, передача по каналам связи и непосредственно на терминалы; проверка

и оформление выходных документов (размножение, брошюровка, переплет, перенос содержимого документов на микрофильмы и т. п.); передача подготовленных выходных документов их пользователям, а массивов данных — на хранение. Графически такой технологический процесс представлен на рис. 11.2. В нем детализованы лишь операции контроля принимаемых документов и ввода данных в ЭВМ. Аналогичный контроль должен осуществляться и при записи данных на машинные носители. Основные виды такого контроля более детально описаны в 11.2.

11.2. Выделение операций технологических процессов

Рассмотренные характеристики операций технологических процессов отражают лишь их потребительские свойства, т. е. пригодность для выполнения конкретной работы вообще. Реализация выбранного варианта технологического процесса всегда связана с определенными затратами времени и денежных средств, а также с необходимостью обеспечения требуемой достоверности результатов. Поэтому для каждого варианта необходимо определить эти характеристики. Для типовых технологических процессов они приобретают особое значение, так как такие процессы предназначены для широкого использования.

Для выявления затрат на весь технологический процесс необходимо оценивать их для каждой операции. Поэтому практически операции следует выбирать с одновременной оценкой затрат на их выполнение, включать в процесс более элементарные операции, затраты на которые можно определить непосредственно.

Рассмотрим принципы выделения технологических операций и оценки временных затрат на их выполнение с применением КВМ, ПВМ и ЭВМ.

Как известно, при использовании КВМ и ПВМ выполняются ручные (машинно-ручные), а для ЭВМ — машинные (автоматические) операции.

Чисто ручной является работа, выполняемая при приемке и проверке первичных документов, контроле и выдаче составленных табуляграмм, ведении соответствующих журналов, при передвижении информации в вычислительной установке, подготовке и завершении работ на вычислительных машинах. Такие работы на отдельные операции разбиваются (или объединяются) с учетом единства факторов времени, пространства, исполнителей и объектов, над которыми они выполняются. Затраты времени в этом случае во многом зависят от субъективных возможностей исполнителей, поэтому для сравнительной оценки времени на выполнение таких операций обычно используются средние величины.

Ручные операции могут быть нормируемыми и ненормируемыми. Затраты времени на выполнение последних определяются как часть (процент) времени выполнения первых. Так, например, используются усредненные нормы выполнения арифметических операций. Такими нормами могут быть: на суммирование — 600 дей-

ствий в час, умножение — 210, деление — 170, а на все вспомогательные операции трудоемкость определяется с учетом оценки основных операций, например умножая полученное время на коэффициент 2.

Операторами часто выполняется ручная запись данных. Это составная операция, в которую входит подготовка к ее выполнению (взятие ручки, поднесение ее к документу и т. п.) и сама запись знаков. Если на запись одного из n знаков тратится a времени, а на подготовку — b , то общее время записи T_3 определяется по формуле

$$T_3 = a \cdot n + b.$$

Практически установлено, что $a = 0,6$ с, $b = 1,2$ с.

При контроле правильности записи информации на машинный носитель много времени тратится на нахождение ошибок путем сличения информации на документе и машинном носителе (контрольной табуляграмме). Эти затраты пропорциональны длине сравниваемых слов. На сравнение двух знаков в среднем тратится около 0,4 с. При сравнении данных на перфоленте с соответствующей информацией на документе, кроме того, требуется около 1,2 с на перевод каждой комбинации пробивок перфоленты в цифры (буквы). Тогда для определения общих затрат на поиск искаженной информации и исправление в массиве, записанном на перфоленте, необходимо время $T_{\text{ин}}$:

$$T_{\text{ин}} = \frac{n(c + ad)}{a} \left(1 - \frac{1}{2^m} \right),$$

где n — количество документострок в массиве;

a — среднее количество документострок в документе;

c — затраты времени на перелистывание одного документа ($c = 1,2$ с);

d — затраты времени на сравнение информации с документострокой по отличительному (ключевому) слову;

m — количество поисков в массиве.

Машинно-ручные операции выполняются при обработке документов, переносе информации на машинные носители, составлении документов на бухгалтерских и фактурных машинах. К ним относятся большинство операций, выполняемых на КВМ и ПВМ. Номенклатура таких операций в первую очередь зависит от возможностей самих КВМ и ПВМ, которые определены в технической документации этих машин. В ней указаны и минимальные затраты времени на выполнение операций (через технические скорости соответствующих машин). Однако практически номенклатура этих операций расширяется, так как появляются дополнительные действия, связанные с подготовкой к выполнению основной операции, ее контролем, завершением и т. п. Для целей проектирования часто применяются нормы не на такие отдельные операции, а на виды работ. Так, Всесоюзным государственным проектно-технологическим институтом по механизации учета и вычислительных ра-

и оформление выходных документов (размножение, брошюровка, переплет, перенос содержимого документов на микрофильмы и т. п.); передача подготовленных выходных документов их пользователям, а массивов данных — на хранение. Графически такой технологический процесс представлен на рис. 11.2. В нем детализованы лишь операции контроля принимаемых документов и ввода данных в ЭВМ. Аналогичный контроль должен осуществляться и при записи данных на машинные носители. Основные виды такого контроля более детально описаны в 11.2.

11.2. Выделение операций технологических процессов

Рассмотренные характеристики операций технологических процессов отражают лишь их потребительские свойства, т. е. пригодность для выполнения конкретной работы вообще. Реализация выбранного варианта технологического процесса всегда связана с определенными затратами времени и денежных средств, а также с необходимостью обеспечения требуемой достоверности результатов. Поэтому для каждого варианта необходимо определить эти характеристики. Для типовых технологических процессов они приобретают особое значение, так как такие процессы предназначены для широкого использования.

Для выявления затрат на весь технологический процесс необходимо оценивать их для каждой операции. Поэтому практически операции следует выбирать с одновременной оценкой затрат на их выполнение, включать в процесс более элементарные операции, затраты на которые можно определить непосредственно.

Рассмотрим принципы выделения технологических операций и оценки временных затрат на их выполнение с применением КВМ, ПВМ и ЭВМ.

Как известно, при использовании КВМ и ПВМ выполняются ручные (машинно-ручные), а для ЭВМ — машинные (автоматические) операции.

Чисто ручной является работа, выполняемая при приемке и проверке первичных документов, контроле и выдаче составленных табуляграмм, ведении соответствующих журналов, при передвижении информации в вычислительной установке, подготовке и завершении работ на вычислительных машинах. Такие работы на отдельные операции разбиваются (или объединяются) с учетом единства факторов времени, пространства, исполнителей и объектов, над которыми они выполняются. Затраты времени в этом случае во многом зависят от субъективных возможностей исполнителей, поэтому для сравнительной оценки времени на выполнение таких операций обычно используются средние величины.

Ручные операции могут быть нормируемыми и ненормируемыми. Затраты времени на выполнение последних определяются как часть (процент) времени выполнения первых. Так, например, используются усредненные нормы выполнения арифметических операций. Такими нормами могут быть: на суммирование — 600 дей-

ствий в час, умножение — 210, деление — 170, а на все вспомогательные операции трудоемкость определяется с учетом оценки основных операций, например умножая полученное время на коэффициент 2.

Операторами часто выполняется ручная запись данных. Это составная операция, в которую входит подготовка к ее выполнению (взятие ручки, поднесение ее к документу и т. п.) и сама запись знаков. Если на запись одного из n знаков тратится a времени, а на подготовку — b , то общее время записи T_3 определяется по формуле

$$T_3 = a \cdot n + b.$$

Практически установлено, что $a=0,6$ с, $b=1,2$ с.

При контроле правильности записи информации на машинный носитель много времени тратится на нахождение ошибок путем сличения информации на документе и машинном носителе (контрольной табуляграмме). Эти затраты пропорциональны длине сравниваемых слов. На сравнение двух знаков в среднем тратится около 0,4 с. При сравнении данных на перфоленте с соответствующей информацией на документе, кроме того, требуется около 1,2 с на перевод каждой комбинации пробивок перфоленты в цифры (буквы). Тогда для определения общих затрат на поиск искаженной информации и исправление в массиве, записанном на перфоленте, необходимо время $T_{\text{пн}}$:

$$T_{\text{пн}} = \frac{n(c + ad)}{a} \left(1 - \frac{1}{2^m}\right),$$

где n — количество документострок в массиве;

a — среднее количество документострок в документе;

c — затраты времени на перелистывание одного документа ($c=1,2$ с);

d — затраты времени на сравнение информации с документострокой по отличительному (ключевому) слову;

m — количество поисков в массиве.

Машинно-ручные операции выполняются при обработке документов, переносе информации на машинные носители, составлении документов на бухгалтерских и фактурных машинах. К ним относятся большинство операций, выполняемых на КВМ и ПВМ. Номенклатура таких операций в первую очередь зависит от возможностей самих КВМ и ПВМ, которые определены в технической документации этих машин. В ней указаны и минимальные затраты времени на выполнение операций (через технические скорости соответствующих машин). Однако практически номенклатура этих операций расширяется, так как появляются дополнительные действия, связанные с подготовкой к выполнению основной операции, ее контролем, завершением и т. п. Для целей проектирования часто применяются нормы не на такие отдельные операции, а на виды работ. Так, Всесоюзным государственным проектно-технологическим институтом по механизации учета и вычислительных ра-

бот ЦСУ СССР под методическим руководством Центрального бюро нормативов по труду при Научно-исследовательском институте труда Государственного комитета Совета Министров СССР по труду и социальным вопросам разработан и этим комитетом утвержден (17 февраля 1977 г.) «Сборник норм времени (выработки) на виды работ, выполняемых на КВМ, ПВМ и устройствах подготовки данных на технических носителях для ввода в ЭВМ». В этих единых нормах учтено время на выполнение подготовительно-заключительных работ, а также на обслуживание рабочего места, отдых и личные надобности, которое составляет 14% оперативного времени. К таким операциям также относится проверка работы машины.

Единые нормы детализированы по признакам работ и условиям их применения. Для целей проектирования и его автоматизации эти нормы можно представить в виде определенных коэффициентов или графических функций.

Снижение достоверности данных может происходить из-за ошибок, возникающих на любой операции технологического процесса, однако основная их доля связана с работой персонала, осуществляющего ручные и машинно-ручные операции сбора, записи, перезаписи и обработки данных. Эти ошибки могут быть как объективными, так и субъективными.

Достоверность данных может быть повышена организационными мерами, учитывающими заинтересованность персонала и устанавливающими ответственность за ошибки. Однако для большинства случаев необходимы дополнительные контрольные операции. Все более широкое распространение получают аппаратные и алгоритмические (программные) средства контроля. Методы такого контроля достаточно полно рассмотрены в [30]. Поэтому мы здесь остановимся лишь на некоторых основных моментах, которые играют важную роль при построении типовых процессов обработки данных на ЭВМ.

Предварительная обработка данных, как известно, позволяет устранить ошибки независимо от их причин и источников возникновения и, таким образом, служит средством сквозного контроля случайных ошибок. Поэтому такая обработка должна рассматриваться как обязательный и важный этап машинной обработки данных, реализация которого главным образом основывается на использовании специального программного обеспечения помехоустойчивого ввода данных. Некоторые функции предварительной обработки данных, проводимой с одновременным выявлением и устранением ошибок в исходных данных, реализуются в процессорных системах подготовки данных аппаратно. Прогресс в развитии устройств подготовки и ввода данных позволяет предполагать, что аппаратные средства обеспечения достоверности данных в будущем будут играть все большую роль.

Программный контроль на современном этапе развития систем машинной обработки данных является главенствующим. Могут быть выделены такие его разновидности:

1. Контроль реквизита, при котором используется избыточность, имеющаяся только в отдельных элементах реквизита или в реквизите в целом. Чаще всего применяются следующие формы этого контроля: контроль предельных значений (по установленному диапазону); контроль по принадлежности и шаблону (проверка разрядности); контроль целочисленности значений числовых реквизитов; проверка условных обозначений реквизитов с использованием самоконтролирующихся кодов. Практически такие формы контроля применяются в различных сочетаниях, т. е. контроль носит комбинированный характер.

2. Контроль совокупности реквизитов документов, использующий синтаксические и семантические связи между реквизитами одного и того же документа, а также их групп. Можно выделить два вида таких связей: арифметические и логические.

Арифметические связи выражаются операциями отношения ($=$, $\neq$, $>$, $\geq$, $<$, $\leq$), а также действиями сложения (например, контрольное суммирование), вычитания, умножения (например, вычисление процента), деления. Логические связи могут существовать между отдельными реквизитами и определяемыми ими показателями. Они позволяют выявить непротиворечивость данных, логическую полноту и последовательность реквизитов в документах.

11.3. Особенности системной обработки данных и программного обеспечения современных средств вычислительной техники

Системный подход к проектированию технологии обработки данных базируется на единстве технологического, информационного, математического и программного обеспечения. Определяющую роль при этом играет системное программное обеспечение (СПО) современных средств вычислительной техники. Оно представляет собой комплекс программ, языков, описаний и инструкций, обеспечивающий взаимодействие всех элементов системы в процессе ее функционирования. СПО предназначено для организации вычислительных процессов и управления ими, эффективного использования аппаратных средств (ресурсов) ЭВМ, автоматизации процедур подготовки и отладки прикладных программ. В его состав входят: операционная система, система программирования, система контроля и диагностики, а также комплексы программ, осуществляющих различные функции по управлению информационным фондом, обеспечению сохранности программных и информационных файлов, организации связи с оператором и т. д.

При реализации технологического процесса совместно с системным программным обеспечением применяется комплекс прикладных программ пользователя, а также функционально настроенных специальных пакетов прикладных программ. Прикладное программное обеспечение должно обладать большой гибкостью и адаптируемостью к различным условиям применения и вместе с тем быть эффективным средством комплексной автоматизации об-

работки данных. Эти требования, вытекающие из системности обработки данных, могут быть удовлетворены путем типизации построения программ, универсализации их входных и выходных форматов, разработки высокоорганизованной системы программной генерации. В этом отношении большие возможности открывают ЭВМ третьего и четвертого поколения, программное обеспечение которых построено с использованием принципов модульности и генерируемости и образует открытые системы. Так, программное обеспечение АСУ и АСОД, базирующихся на вычислительных средствах ЕС ЭВМ, концентрируется вокруг двух типов операционных систем — ДОС ЕС и ОС ЕС. Дисковая операционная система ДОС ЕС обеспечивает: пакетную обработку заданий в мультипрограммном режиме с одновременным выполнением до трех программ; статическое распределение внешних устройств и оперативной памяти; совмещение операций ввода-вывода с работой процессора; использование библиотек программных модулей, широкого набора языков программирования, пакетов прикладных программ; телеобработку данных; протоколирование хода работы вычислительной системы; подключение других различных вспомогательных аппаратных и программных средств. Большими возможностями обладает ОС ЕС, которая может обеспечить функционирование не только однопроцессорных, но и многопроцессорных комплексов, различных технических средств ЕС ЭВМ. При этом работа системы может быть организована как в пакетном режиме (с фиксированным и переменным числом задач — максимум до 15), так и в режиме с разделением времени, а также в реальном масштабе времени. ОС ЕС обеспечивает динамическое распределение памяти, открывает большие возможности для эффективной организации обработки данных, но вместе с тем для ее реализации требуются более мощные технические средства.

Модульность построения компонентов операционных систем дает возможность выбрать необходимый для конкретных условий вариант и сгенерировать его с учетом индивидуальных требований пользователя. Сгенерированный вариант операционной системы накладывает определенные ограничения на прикладное программирование, которое заключается в адаптации применяемых пакетов прикладных программ к условиям проектируемой системы, создании специфичных (оригинальных) модулей прикладного назначения, разработке системы генерации рабочих программ с использованием этих модулей и включении их в единый комплекс, обладающий развитыми адаптационными свойствами и высокой эффективностью функционирования.

Прикладное программное обеспечение современных систем обработки данных, предназначенное для реализации функциональных задач управления определенным объектом, состоит из двух основных частей: системы пакетов прикладных программ, ориентированных на автоматизацию определенных функций управления, реализацию экономико-математических методов для решения различных, в том числе оптимизационных, задач, а также на форми-

рование и ведение информационных фондов; системы стандартизованных прикладных программ, ориентированных на отдельные задачи управления или определенные типовые процедуры обработки данных. Структурно в каждый пакет прикладных программ входят средства генерации и настройки пакета, ведущая программа (в пакетах простой структуры она отсутствует), программные модули тела пакета, сервисные средства обслуживания (к их числу можно отнести, например, входной язык пакета). Стандартизованные оригинальные программы обычно объединяются в специальную библиотеку, которая снабжается средствами формирования рабочих программ и обслуживания программного фонда. Для составления программ пользователей в нашей стране широкое применение получили унифицированные алгоритмические языки: Кобол, PL/I, Фортран, а также РПГ и др. Для программирования наиболее ответственных процедур, отдельных фрагментов задач часто используется язык Ассемблера, позволяющий наиболее эффективно строить рабочие программы (с точки зрения их эксплуатации). Благодаря стандартным средствам операционных систем в рабочие программы можно включать блоки, написанные на различных языках программирования. Сегментация программ вызывается, однако, не только этим обстоятельством: как готовые пакеты прикладных программ, так и создаваемые оригинальные программы должны обладать возможностью настраивания. В этом заключается их гибкость и независимость от структуры используемых данных. В процессе функционирования системы прикладное программное обеспечение может не только наращиваться количественно, но и претерпевать качественные изменения без нарушения принятых общесистемных принципов обработки данных. Сборка рабочих программ с использованием различных модулей позволяет создавать различные версии таких программ, модифицировать их по мере необходимости.

На современном этапе развития АСУ все более широкое распространение получают генерирующие системы составления прикладных программ для решения типовых задач управления, в том числе и оптимизационных. Для таких программ, используемых в модульных системах обработки данных, характерна типовость их получения за три этапа, выполняемых в такой последовательности:

- 1) описание условий задачи в виде определенного отношения, связывающего исходные данные с результатом решения;
- 2) построение конструктивного доказательства, что существует решение данной задачи;
- 3) перевод полученного доказательства в форму программы.

Такие системы генерации программного обеспечения можно отнести к средствам автоматизации проектирования АСУ и АСОД. Более подробно их организация рассматривается в 16.3. Здесь лишь отметим, что в состав таких систем входят: информационная база генерации, входные языки, ведущая программа (общая для всех генерируемых пакетов), трансляторы с входных языков, библиотека модулей, вспомогательные программы.

ФОРМИРОВАНИЕ И ВЕДЕНИЕ ИНФОРМАЦИОННОГО ФОНДА И БАЗЫ ДАННЫХ

12.1. Особенности технологии машинных процедур формирования файлов данных

Под формированием файлов данных мы понимаем их первоначальное создание и включение в единый информационный фонд системы. В технологическом аспекте процесс формирования включает следующие этапы:

- 1) сбор и регистрацию данных;
- 2) предварительную обработку данных;
- 3) запись данных на машинный носитель и подготовку для ввода в ЭВМ с соответствующим контролем;
- 4) ввод данных в ЭВМ с машинных носителей или передачу их по каналам связи с выполнением определенных контрольных функций;
- 5) загрузку файлов (ввод в базу данных).

Технологический процесс в каждом конкретном случае характеризуется определенными особенностями и зависит, прежде всего, от применяемых технических средств, методов предварительной обработки данных, типа и структуры самих данных, организации информационного фонда.

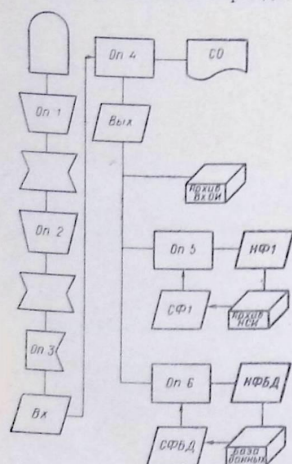


Рис. 12.1. Общая схема процессов формирования файлов данных информационного фонда

Рассмотрим традиционный и до сих пор наиболее распространенный технологический процесс, при котором исходные данные (условно-постоянная или переменная информация) фиксируются на входных документах. Его общая схема представлена на рис. 12.1. Для записи данных (первая операция — Оп.1) могут быть использованы счетно-табличные или пишущие машины. При заполнении документов выполняются операции предварительной арифметической обработки (Оп.2). К таким операциям прежде всего относится подсчет контрольных сумм (по каждой строке или столбцу, по каждому документу или их пачке), используемых для контроля подготовки машинных носителей и ввода данных в ЭВМ. Также могут выполняться некоторые расчеты (например, таксировка

при составлении счетов на фактурных машинах), если предусмотрено включение в документ не только первичных, но и расчетных данных. К предварительной обработке данных следует отнести и операцию логической обработки документов, т. е. комплектование их в пачки по определенным признакам. Заполненные и укомплектованные документы передаются для переноса данных с них на машинные носители, если они не были получены автоматически на перфоприспособках одновременно с заполнением документов. Исходные массивы данных могут размещаться на перфокартах, перфолентах, магнитных лентах и дисках. Запись их на машинные носители осуществляется в данном случае вручную путем набора документированных данных на клавиатуре соответствующего устройства согласно разработанному макету размещения (Оп. 3). К предварительной обработке информации на данном этапе относятся операции контроля подготовки машинных носителей, если при этом используется так называемый счетный метод (контроль по контрольным суммам) или балансовый метод (контроль равенства между показателями при их определенной математической зависимости). При выполнении такого контроля производится распечатка протокола ошибок, по которому уточняются данные соответствующих документов и вносятся изменения на машинные носители. В случае верификации данных производится повторный набор их на клавиатуре и сравнение с содержимым соответствующего носителя. Выявленные при этом ошибки данных исправляются. Исправление для перфокарт сводится к перебивке новой карты и замене ею ошибочной, а для перфолент, магнитных лент и дисков — к переключению с режима дублирования на режим ручной записи и повторному нанесению исправленных данных. При подготовке машинных носителей используются следующие технические средства:

карточные перфораторы и контрольные;

ленточные перфораторы или перфоприспособки к счетно-табличным машинам;

устройства подготовки данных на магнитных лентах (типа ЕС-9002 или СПД-9000, использование которых рассмотрено в 9.4);

устройства подготовки данных на магнитных дисках (например, типа Р810, разработанные СКБ вычислительных машин Вильнюсского производственно-технического объединения «Сигма»).

Для передачи данных по каналам связи чаще всего используются телетайпы.

Наиболее ответственной и сложной процедурой формирования файлов данных является ввод информации с машинных носителей (Оп. 4). В случае использования перфокарт вводу, как правило, предшествует упорядочение (сортировка) массива по определенным признакам. Для ввода данных и загрузки файлов в память системы (на магнитные ленты или диски) подготавливается программа ЭВМ. При этом имеется возможность выбора одного из нескольких вариантов:

- 1) составляется оригинальная программа с использованием физической системы ввода-вывода данных (на уровне Ассемблера);
- 2) оригинальная программа подготавливается на базе логической системы ввода-вывода (на уровне Ассемблера, Кобола, Фортрана, PL/1 и других процедурных или проблемных языков программирования);
- 3) используется универсальный генератор ввода-вывода данных;
- 4) настраивается пакет специальных программ перезаписи файлов с одних носителей на другие;
- 5) генерируется система управления базой данных, выбранная для организации информационного фонда;
- 6) используется принятая система телеобработки данных.

В процессе ввода данные могут подвергаться преобразованию с целью получения требуемого формата или структуры. Форматирование может носить различный характер: от изменения формы представления данных, расстановки отдельных элементов до содержательного преобразования, связанного с выполнением определенных арифметических и логических операций. Наиболее сложными представляются такие процедуры ввода, в которых исполь-

зуется несколько входных массивов, объединяемых в один выходной файл. Слиянию массивов предшествует их упорядочение по общим признакам (логические операции сравнения, поиска, выборки).

Оригинальные программы ввода разрабатываются в случаях, когда возможности готовых средств, имеющихся в системном и прикладном программном обеспечении, не отвечают требованиям, вытекающим из особенностей структурной организации файлов и предусмотренных методов контроля ввода. Реализация методов контроля в таких программах в основном и представляет наибольшие трудности.

Стандартный процесс ввода данных, который может быть реализован с

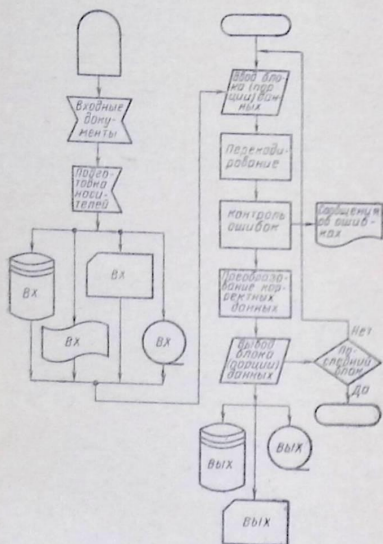


Рис. 12.2. Процесс ввода данных (ВХ — входные файлы, ВЫХ — выходные файлы)

использованием универсального генератора ввода-вывода для ЕС ЭВМ, представлен на рис. 12.2. Здесь предусмотрен блочный (порционный) ввод данных с машинных носителей; они могут представляться в любой системе кодирования, поэтому подвергаются перекодированию во внутренний код ЕС ЭВМ (ДКОИ). Генератор предусматривает несколько видов контроля вводимых данных:

- 1) наличие разделителей записей;
- 2) соответствие определенного элемента записей одному из ключей, представленных в описании данных;
- 3) правильность длины записи;
- 4) соответствие формата каждого элемента записи его описанию;
- 5) наличие элементов в записи;
- 6) правильность длины элементов записи;
- 7) соответствие элемента диапазону его изменения;
- 8) равенство контрольной суммы по каждой записи;
- 9) соответствие иерархии элементов в записи, заданной в описании;
- 10) равенство контрольной суммы по блоку (порции) записей;
- 11) порядок расположения записей;
- 12) порядок вхождения записей (контроль по структурным уровням файла).

При обнаружении ошибочных записей, которые не включаются в выходной файл, на печать выводятся сообщения (на рис. 12.1 — СО). После соответствующего анализа ошибок подготавливается массив корректур, который используется для добавления исправленных записей в ранее созданный выходной файл.

Записи выходных файлов komponуются из элементов входных записей; кроме того, имеется возможность в выходные записи включать такие элементы, которые получаются в результате обработки данных по подключаемым к генератору программным модулям пользователя. Пользователем также могут устанавливаться условия формирования выходных записей.

Выходные файлы, получаемые в результате выполнения программы ввода, являются исходными для формирования информационного фонда на машинных носителях. В этом фонде могут находиться файлы трех видов. Используя терминологию, принятую в методических материалах по созданию АСУП (ОРММ АСУП), их можно называть файлами входной оперативной информации (на рис. 12.1 — ВхОИ), нормативно-справочной информации (НСИ) и базы данных (БД).

В файлах входной оперативной информации обычно хранятся данные, которые накапливаются и используются в одном планово-учетном периоде. Такие файлы в процессе их хранения не корректируются (в них не вносятся изменения), непосредственно не дополняются данными, хранящимися в других файлах. В течение определенного периода они могут пополняться новыми данными.

Файлы нормативно-справочной информации широко применяются при использовании ЭВМ второго и третьего поколения. Они

являются необходимыми и при наличии единых баз данных, например для хранения различной справочной информации, кодификаторов и др. При формировании таких файлов или внесении изменений в них (Оп. 5) каждый новый файл (НФ1) получается из файла вводимых корректирующих данных (Вх) и старого хранимого файла (СФ1).

Третью группу составляют файлы, непосредственно относимые к базе данных. Они формируются и изменяются (загружаются и ведется база данных) с помощью средств специального программного обеспечения (Оп. 6).

В современных АСУ и АСОД и в первую очередь тех, которые базируются на вычислительных центрах коллективного пользования, применяются технические средства передачи данных по каналам связи. Они широко используются и в системах с распределенными базами данных, которые особенно эффективны в условиях крупных производственных и научно-производственных объединений. Практика показывает, что только около половины нормативно-технической и другой документации используется во всех или нескольких подразделениях объединения (остальная часть — лишь в отдельных подразделениях). Поэтому в системе обработки данных целесообразно организовать информационный фонд двух уровней: центральный фонд данных и местные фонды. Для обеспечения обмена информацией в такой системе должна быть организована сеть каналов связи между местными пунктами (МП) и центральным пунктом (ЦП).

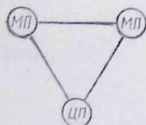


Рис. 12.3. Каналы связи (метод треугольника)

Под линией связи мы понимаем средства, обеспечивающие непосредственную двустороннюю передачу информации между двумя пунктами. Каналом связи будем называть одну или несколько линий связи, обеспечивающих передачу данных между двумя пунктами как непосредственно, так и через другие пункты. Так, например, в систему связи, показанную на рис. 12.3, входят три линии связи, а оба МП связаны с ЦП двумя каналами (с одной и двумя линиями).

В сети, показанной в левой части рис. 12.4, а, имеются 9 линий связи, а каждый МП связан с ЦП десятью каналами связи (2 канала включают по одной линии, 4 — по две и 4 — по три линии).

В зависимости от объемов и интенсивности потоков данных, пропускных возможностей и надежности каналов для организации системы связи между центральным и местными пунктами могут быть использованы четыре метода.

Метод треугольника. По этому методу организуются связи между двумя местными пунктами и центральным пунктом (рис. 12.3). В треугольники включаются все возможные пары местных пунктов, причем каждая из них с ЦП образует локальный треугольник, т. е. линии ранее созданных треугольников не могут быть использованы для организации нового треугольника. Для большей на-

глядности на рис. 12.4, а представлены сети, организованные методом треугольника и включающие по 3 (слева) и 4 (справа) местных пункта.

Линейный метод. В этом случае каждый местный пункт имеет по одной линии связи со всеми другими пунктами системы (ЦП и МП). Такие сети для 3 и 4 местных пунктов показаны на рис. 12.4, б.

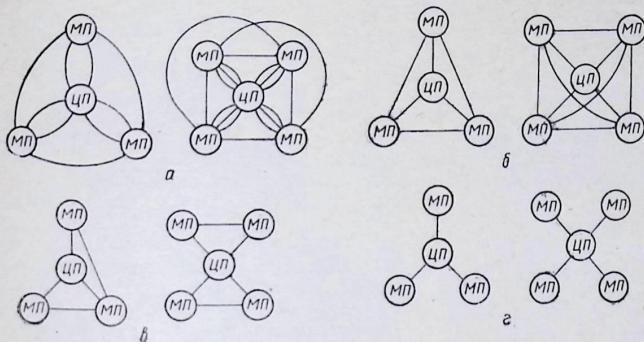


Рис. 12.4. Каналы связи для 3 и 4 МП

Метод соседа. По этому методу каждый местный пункт имеет одну линию связи с центральным пунктом, а также линию связи с одним соседним (как правило, ближайшим) местным пунктом. Как видно из схем, представленных на рис. 12.4, в, в этом случае каждый МП имеет по 2 или 3 линии (в зависимости от четности общего числа местных пунктов).

Метод центра. Используя данный метод, каждый МП отдельной линией связывается только с ЦП, так, как это показано на рис. 12.4, г.

При выборе одного из рассмотренных методов организации сети связи между МП и ЦП необходимо провести их сравнительный анализ и установить, как влияет увеличение общего количества линий связи на рост количества различных каналов связи (в разрезе их типов).

Для обмена данными в системе распределенных баз могут использоваться три способа: автоматизированный, полуавтоматический и автоматический. При автоматизированном способе данные передаются через оборудование абонентских пунктов, используемое в автономном режиме. Данные в канал связи вводятся с перфокарты, а из канала данные выводятся также на перфокарту. При полуавтоматическом способе обмен осуществляется с использованием перфокарты только на абонентских пунктах, т. е. информация

вводится непосредственно в ЭВМ. Автоматический способ предполагает наличие ЭВМ и на абонентских пунктах, и в центре. При этом отправителем и получателем данных является ЭВМ.

Создание файлов и загрузка их в информационный фонд зависят от его структурной организации и принимаемых методов поиска. Эти две стороны информационно-поискового процесса очень тесно связаны друг с другом и взаимообусловлены. Общеизвестно, что при достаточно больших объемах информационного фонда в условиях применения современных ЭВМ наиболее целесообразно использовать индексно-последовательный метод организации файлов данных, а иногда и метод произвольной выборки. Однако реализация функции цели поиска чаще всего связана не с одним, а с несколькими файлами данных, имеющими иерархическую структуру. В таких случаях возникают определенные трудности при организации увязки последовательных шагов поиска при переходе от одного файла к другому. Они сводятся главным образом к тому, что на каждом очередном шаге поиска должен

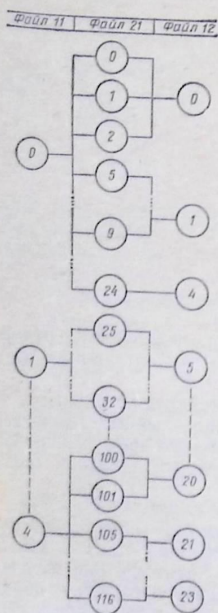


Рис. 12.5. Фрагмент информационного фонда

быть сформирован ключ следующего шага поиска. Наиболее часто встречающееся использование адресных ссылок для этих целей не всегда рационально, к тому же это требует довольно сложных средств программного обеспечения формирования и ведения информационного фонда. Рассмотрим методику организации информационной связи иерархической системы файлов данных с использованием уникальных цифровых идентификаторов записей, позволяющих простым расчетным путем осуществить переход к связанным записям других файлов. На основе такой методики, которая вносит определенные особенности в организацию машинных процедур формирования файлов, могут создаваться взаимосвязанные файлы с использованием специальных программ, обладающих высокой степенью универсальности. Сущность этой методики заключается в следующем.

Для двух непосредственно связанных файлов данных устанавливается постоянное соотношение подчиненных записей ($1:n$), т. е. принимается условие, что к одной записи ведущего файла (файла более высокого уровня) относятся n записей связанного с ним файла более низкого уровня, причем число n для всех записей ведущего файла является постоянным, а его значение для пары кон-

кретных файлов устанавливается по максимально возможному количеству записей более низкого уровня, относящихся к одной записи более высокого уровня. Фактически это означает, что группы записей более низкого уровня, связанные с одной записью ведущего файла, дополняются фиктивными записями до установленного значения n . Всем записям таких файлов (включая фиктивные) присваиваются их порядковые номера $N = pn + i$, где p и i принимают значение от 0 до $(n - 1)$. Отсюда следует, что если порядковый номер N записи более низкого уровня разделить на принятое значение параметра n , то целая часть частного p покажет порядковый номер записи более высокого уровня, к которой данная запись относится, а остаток от деления i указывает на порядковый номер этой записи внутри группы записей, относящихся к p -й записи более высокого уровня. С другой стороны, порядковый номер записи более высокого уровня, умноженный на то же значение n , указывает порядковый номер первой записи группы более низкого уровня, связанной с данной записью. Порядковые номера других записей этой группы устанавливаются путем увеличения значения параметра i от 0 до $(n - 1)$. Как видно из самого принципа присваивания записям их порядковых номеров, используя такие номера, можно осуществить поиск нужных записей с переходом как с более высокого уровня на более низкий, так и наоборот. Присвоение же этих номеров записям всех уровней иерархического информационного фонда можно произвести программным путем по алгоритму, представленному в 7.5. Отметим также, что порядковый номер записи, дополненный номером уровня, может служить уникальным ключом поиска в едином файле данных, объединяющем записи всех уровней.

Изложенная методика организации информационной связи пригодна не только для файлов с простой иерархической подчиненностью, но и в условиях комбинированной связи, когда записи более низкого уровня подчинены не одному, а двум или нескольким ведущим файлам. Рассмотрим конкретный пример использования описанной методики для формирования справочного файла данных о сопротивлениях (резисторах), необходимого для конструкторской, технологической и технической подготовки производства любого приборостроительного предприятия. Сопротивления, как известно, характеризуются определенными значениями номинала (первый файл первого уровня с номером 11), их мощности (второй файл первого уровня с номером 12) и другими техническими и экономическими параметрами (файл второго уровня с номером 21). Значения этих параметров могут повторяться, с одной стороны, для одинаковых значений номинала при разных значениях мощности, с другой — для тех же значений мощности, но при разных значениях номинала. Именно поэтому в иерархический информационный фонд в данном случае следует включить два ведущих файла (11 и 12) и один подчиненный (21). Фрагмент такого фонда можно представить так, как это показано на рис. 12.5. Здесь принято (чисто условно), что сопротивления с одинаковым

значением номинала максимально могут характеризоваться 25 различными наборами параметров ($n=25$), а одному значению мощности может соответствовать до 5 различных комбинаций значений этих же параметров ($m=5$). На рис. 12.5 в кружочках указаны порядковые номера записей соответствующих файлов. Для организации процедуры поиска по всем возможным направлениям (файл $11 \rightarrow$ файл 21, $21 \rightarrow 11$, $12 \rightarrow 21$, $21 \rightarrow 12$, $11 \rightarrow 12$, $12 \rightarrow 11$) используются следующие формулы определения порядковых номеров искоемых записей:

$$N_{21} = N_{11} \cdot n + i \text{ при } i = 0 \div (n - 1);$$

$$N_{11} = \lfloor N_{21}/n \rfloor;$$

$$N_{21} = N_{12} \cdot m + j \text{ при } j = 0 \div (m - 1);$$

$$N_{12} = \lfloor N_{21}/m \rfloor;$$

$$N_{12} = (N_{11} \cdot n/m) + k \text{ при } k = 0 \div (n/m) - 1;$$

$$N_{11} = \lfloor N_{12} \cdot m/n \rfloor.$$

Сформированный таким образом информационный фонд данных позволяет производить поиск по запросам свободной формы, т. е. вне зависимости от того, с какого файла начинается процедура поиска, и даже при неполном охвате всех файлов сформулированным запросом. На первом шаге процедуры поиска всегда выбирается одна запись по ключу поиска, указанному в запросе (для рассмотренного примера им могут быть значения номинала, мощности или некоторых параметров), а затем осуществляется поиск уже по порядковым номерам записей, определенным по порядковому номеру первой выбранной записи. При этом ответ может быть сформирован как из отдельных записей всех трех файлов, так и из их групп (в виде определенного списка).

Если для формирования и ведения информационного фонда используются готовые системы управления базами данных, а также пакеты прикладных программ, то после генерации и настройки этих программных средств подготавливаются входные файлы согласно тем требованиям, которые предъявляются принятой системой с учетом разработанной пользователем структуры базы данных. Для загрузки файлов (записи данных в информационный фонд) могут использоваться системные программы с добавлением нужных модулей пользователя в предусмотренных точках входа. Системы управления базами данных обеспечивают генерацию системных программ, ввод данных в информационный фонд, корректировку структуры базы данных, копирование и восстановление информационного фонда, сбор и обработку статистических данных об использовании файлов базы данных, распечатку файлов, поиск в информационном фонде по запросам от программ пользователя, формирование ответов на запросы с терминалов, ведение словарей базы данных, интерфейс с программами пользователя (через специальные макрокоманды).

12.2. Оценка и выбор вариантов технологических процессов предварительной обработки данных

Предварительная обработка данных предшествует непосредственной загрузке файлов в единый информационный фонд системы или сопровождает ее. Рассмотрим более подробно основные операции предварительной обработки, отмеченные в 12.1.

Прежде всего необходимо подчеркнуть, что в современных системах машинной обработки данных практически полностью можно исключить трудоемкие ручные арифметические операции, т. е. отказаться от выполнения их при заполнении первичных документов (вручную или с использованием КВМ) и передать эту работу ЭВМ. Для этого могут использоваться массивы (файлы) нормативно-справочной информации. Тогда на этапе сбора и регистрации данных остается рассчитывать лишь контрольные суммы. Для наиболее массовых многострочных документов, как уже отмечалось, целесообразно рассчитывать построчные контрольные суммы, а также контрольную сумму по каждому документу. Для контроля порционного ввода данных кроме этой суммы можно использовать и контрольную сумму на каждую пачку документов.

Устранение ошибок при сборе и регистрации информации, а также переносе ее на машинные носители обеспечивается рациональными формами заполняемых документов, последующим контролем вносимой информации, двойной перфорацией или верификацией, использованием методов сравнения дважды введенной информации в ЭВМ или подсчетом контрольных сумм и сравнением их с введенными в ЭВМ контрольными суммами.

Обеспечение достоверности информации при передаче ее по каналам связи базируется на использовании методов перезапроса, избыточных и самокорректирующихся кодов.

Упорядочению (сортировке) подлежат входные массивы на перфокартах, магнитных лентах и дисках, если при первоначальном формировании или обновлении конкретного файла информационного фонда требуется загружать его записи в определенной последовательности (в порядке возрастания или убывания значений ключевых признаков). Для сортировки массивов, записанных на магнитных лентах и дисках, как правило, используются специальные генераторы сортировки и слияния, включенные в системное программное обеспечение ЭВМ.

Операции формирования данных в процессе ввода сводятся не только к перекомпоновке отдельных элементов записей, но и к арифметической и логической обработке с целью получения требуемых выходных данных. Для таких операций подготавливаются специальные программы пользователя.

Наиболее важной операцией предварительной обработки информации с точки зрения обеспечения достоверности данных является контроль ввода файлов в систему. Именно на этом этапе должны быть исключены ошибки данных, вводимых в информационный фонд. Хотя программы ввода — это, по существу, программы преобразования формата вводимых машинных документов в

формат файлов системы, однако около 80% объема этих программ занимают процедуры контроля достоверности вводимой информации и процедуры исправления обнаруженных ошибок.

Наиболее часто используемыми методами контроля при вводе данных являются: сравнение двух независимо подготовленных машинных носителей; сравнение подсчитываемых сумм с введенными контрольными суммами; контроль пределов значений вводимых данных; контроль полноты и последовательности вводимых записей, документов; контроль

правомерности наличия символов и цифр во вводимой порции информации по шаблону; балансовый контроль (определение строк и столбцов матрицы вводимых данных, нарушающих баланс) и др. Программы ввода необходимо строить так, чтобы процесс устранения ошибок требовал минимального количества повторно подготавливаемой и вводимой информации. Важно также, чтобы исправленные данные не задерживались, немедленно направлялись на обработку и не зависели от процедуры исправления ошибок.

Выбор методов и средств контроля подготовки машинных носителей и ввода данных зависит от важности информации, ее структурной организации, объема, стабильности и других характеристик. Особые требования предъявляются к достоверности условно-постоянной информации. Нормативно-справочные, плановые данные длительного пользования должны обладать высоким уровнем достоверности и надежности. Поэтому при формировании и обновлении их файлов необходимо провести лексический, синтаксический и семантический анализ.

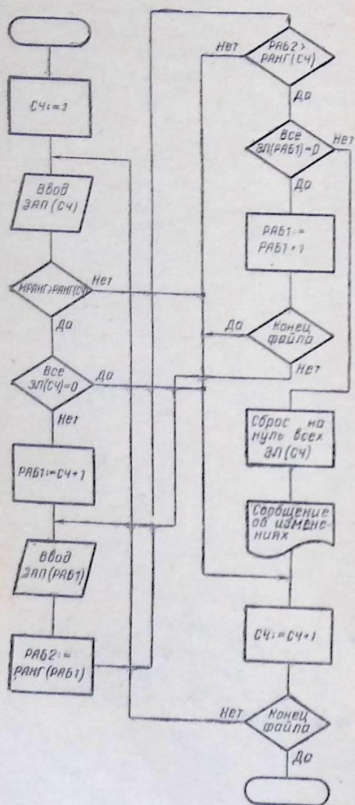


Рис. 12.6. Алгоритм контроля корректности входной информации

Система обработки данных может функционировать нормально только после загрузки всех нужных данных, т. е. после обеспечения полноты информации. Однако бывают случаи, когда требуется провести вариантный расчет до того, как будет обеспечена вся полнота информации. Для этого предусматриваются специальные действия, например обход процедур, использующих недостающую информацию, или переход к специальным процедурам, работающим с неполной информацией. Если такие действия не предусмотрены, то для обеспечения фиктивной полноты информации необходимо установить приблизительные значения недостающих элементов данных, в качестве которых целесообразно выбирать групповые элементы самых высоких рангов без их детализации.

Для примера рассмотрим алгоритм, реализующий такие возможности. Он обеспечивает выполнение следующих функций: отказ от ввода групповых элементов ранга $(n-1)$, если имеются входящие в указанную группу элементы более низкого ранга n ; автоматическое исправление приблизительных (фиктивных) элементов; корректность данных обеспечивается включением в неполный информационный фонд соответствующих данных прошлых периодов. Если при контроле обнаруживается некорректность в какой-то группе элементов, то она должна быть приведена к корректному виду, т. е. если в одной группе заданы элементы как n -го, так и $(n-1)$ -го ранга, то элементы $(n-1)$ -го ранга сбрасываются на нуль. Соответствующее сообщение о внесенном исправлении выдается на печать. Блок-схема рассматриваемого алгоритма представлена на рис. 12.6. Здесь приняты следующие обозначения: МРАНГ — входной параметр, обозначающий максимально возможное значение ранга контролируемой информации; СЧ — счетчик, содержащий номер текущей записи, для которой производится проверка; РАНГ — рабочее поле, содержащее ранг элемента проверяемой записи; РАВ1, РАВ2 — рабочие поля; ЗАП — поле текущей записи; ЭЛ — поле очередного элемента текущей записи.

Данный алгоритм предназначен для контроля полноты и корректности файлов, размещаемых в оперативной памяти ЭВМ или в накопителе с произвольным доступом.

При широком развитии работ по созданию АСУ, их совершенствованию возникает необходимость перевести на индустриальную основу и сам процесс проектирования. Все более остро ощущается необходимость в автоматизации построения и выбора технологических процессов предварительной обработки информации, осуществляемой с целью формирования файлов информационного фонда на машинных носителях.

Разработка общей схемы автоматизированного построения технологических процессов и формирования необходимой документации по последовательности работ и кругу решаемых проблем имеет много общего с проектированием функциональных задач АСУ, их комплексов. Здесь также требуется выяснить сущность автоматизируемых работ, построить математические и информа-

онные модели, создать необходимые информационные файлы и т. д.

Автоматизированно можно разрабатывать проектные решения по выбору наилучших вариантов технологических процессов с задачей необходимой проектной документации. Для этого должна быть проведена большая подготовительная работа. К ней относятся: изучение соответствующих технических и программных средств, тенденций их развития и опыта применения для рассматриваемых работ, возможного различия вариантов технологических процессов; исследование методов обеспечения достоверности информации; определение основных характеристик операций и процессов, нужных для их оценки и выбора; изучение существующей технологической документации и возможных направлений ее унификации.

Задача построения и выбора технологических процессов предварительной обработки информации является многовариантной, так как возможны различные средства, способы и последовательность выполнения операций, их контроля. Оценка выбора чаще всего производится по необходимым для выполнения процесса временным или денежным затратам, достоверности информации на выходе системы.

При построении математической модели выбора прежде всего выделяются операции технологических процессов и определяются те их основные характеристики, по которым в дальнейшем можно выбирать (строить) возможные варианты технологических процессов, давать их количественные оценки. Такими характеристиками могут быть: основные составляющие выполняемой работы (элементарные операции или их совокупности), используемые в них технические средства, требуемые ресурсы и условия выполнения, достигаемая достоверность, время выполнения (для определенного объема работы), себестоимость, приведенные затраты, стоимость по установленным ценам за машинное время и т. п.

После этого можно формулировать задачу оптимизации с функцией цели и ограничениями. В общем виде задачу выбора оптимального по заданному критерию варианта технологического процесса предварительной обработки данных можно сформулировать следующим образом. Известно m различных элементарных технологических операций $P_1, P_2, \dots, P_m$ с соответствующими характеристиками $p_{11}, p_{12}, \dots, p_{1n}$. Необходимо построить такой общий технологический процесс $P_0 = \{P_{01}, P_{02}, \dots, P_{0n}\}$, который позволил бы выполнить соответствующую работу и при этом удовлетворял бы заданным ограничениям и минимизировал (максимизировал) целевую функцию.

Минимизировать можно затраты на выполнение всего процесса, определяемые как сумма затрат на отдельные операции, общее время выполнения процесса (работы), а максимизировать — достоверность информации, определяемую как произведение достоверностей, обеспечиваемых отдельными последовательно выполняемыми операциями, и т. п. Поэтому целевая функция может быть представлена в виде

$$L \rightarrow \text{extr} \sum_{i \in I} P_{ij_0}, \text{ или } L \rightarrow \text{extr} \prod_{i \in I} P_{ij_0},$$

где I — множество операций, которые входят в вариант технологического процесса;

j_0 — выбранная (оптимизируемая) характеристика оценки операций и всего технологического процесса.

Ограничениями задачи могут быть уровень необходимой или допустимой достоверности результатов выполнения всего технологического процесса, объемно-временные требования (непревышение заданного времени) и т. п.

Такую общую задачу решить трудно, поэтому в каждом конкретном случае необходимо формулировать частные задачи, которые могут решаться с помощью известных методов математического программирования и имеющихся программ (пакетов прикладных программ).

При создании АСУ также ставится задача выбора носителей и технических средств вывода результатных данных. При этом можно использовать следующие носители информации: ручной документ, документ, подготавливаемый одновременно с машинным носителем, машиночитаемый первичный документ, дуаль-карту, перфокарту, перфоленту, магнитную ленту, магнитный диск, специальное техническое устройство. Вывод же результатов может производиться на счетно-табличной машине, табуляторе, АЦПУ или специальном техническом устройстве. Не исключена и другая, более сложная постановка данной задачи.

Как уже отмечалось, технологические процессы оцениваются по времени, стоимости (себестоимости, приведенным затратам) и достоверности информации на выходе. Для первых двух оценок важно определить способ и средства оценки каждой операции, выявить последовательность операций процесса. Третья оценка является более сложной и требует глубоких предварительных исследований или экспериментальных работ. Однако для сравнительно несложных и часто встречающихся вариантов можно использовать уже накопленный опыт и построить простые модели их оценки.

Можно сформулировать задачу автоматизации формирования и ведения технической документации. Для ее решения нужна следующая информационная база: кодификаторы технологических операций и реквизитов технико-экономической информации, иногда — описания технологических операций. Используя эту информацию, а также заказ на вариант технологического процесса с указанием исходных и результатных данных, можно сформулировать задачу автоматизации разработки инструкционных и операционных карт технологического процесса.

Также может быть сформулирована задача, дополнительно включающая выбор оптимального варианта технологического процесса. Для этого в информационную базу следует дополнительно ввести описания функций, выполняемых каждым процессом, и оценки, по которым можно выбрать наилучший вариант. Практи-

чески при построении модели такой задачи нет надобности рассматривать все возможные технические средства и варианты их использования в технологических процессах. Всегда по различным причинам можно предварительно исключить неприменимые на данном объекте варианты и тем самым значительно упростить математическую модель задачи.

12.3. Принципы организации и технологические процессы внесения изменений

Ведение созданного информационного фонда предполагает поддержание его в актуальном состоянии. Сущность этой работы мало зависит от уровня системности организации информационного фонда. Поэтому ограничимся рассмотрением ее содержания на примере информационной базы, построенной в виде набора отдельных массивов.

Массивы (файлы) условно-постоянных, а также переменных (накапливаемых) данных должны адекватно отражать состояние тех объектов, характеристики которых они содержат. Для этого в информационный фонд систематически вносятся изменения. Обновление хранимых массивов данных связано с выполнением следующих процедур:

- корректировки массивов условно-постоянной информации;
- аннулирования старых массивов оперативных данных и загрузки новых;
- внесения изменений в накапливаемые массивы переменной информации.

Все виды обновления массивов могут выполняться либо периодически, т. е. по заранее установленному регламенту, либо по мере поступления изменений (в оперативном режиме). Это зависит от характера решаемых задач, в которых используются актуальные массивы информационного фонда, и установленного режима работы всей системы обработки данных.

Наиболее ответственной и сложной является процедура корректировки массивов условно-постоянной (нормативно-справочной, плановой и др.) информации. Для ее осуществления используются два массива — корректируемый и корректирующий, а при корректировке сводных нормативных массивов — корректируемый и тот корректируемый исходный массив, на базе которого рассчитываются новые сводные нормы и нормативы.

Корректирующий массив может включать в себя записи трех типов: новые (добавляемые) записи, ключевые признаки удаляемых записей и полностью или частично заменяемые записи. Тип записи чаще всего указывается в самой записи. Корректировка сводится к поочередному сравнению ключевых признаков записей корректируемого и корректирующего массивов. Блок-схема алгоритма, реализующего данную процедуру, представлена на рис. 12.7. Здесь приняты следующие обозначения: K1 — корректируемый массив; K2 — корректирующий массив; K3 — формируемый (скор-

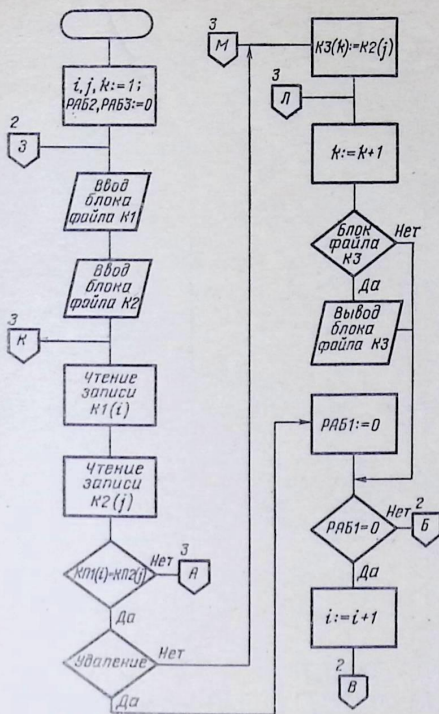


Рис. 12.7. Алгоритм корректировки файлов условно-постоянной информации (начало) (над блоками межстраничных соединителей указаны порядковые номера схем алгоритма)

ректированный) массив; i, j, k — индексы записей соответственно файлов $K1$, $K2$ и $K3$; РАБ1 — рабочее поле для контроля режима корректировки (0 — удаление записи, 1 — добавление новой записи, 2 — перезапись без изменений, неконтролируемые значения — полная или частичная замена записи); РАБ2, РАБ3 — рабочие поля для контроля конца массивов $K1$ и $K2$; КП1, КП2 — ключевые признаки записей массивов $K1$ и $K2$.

Обеспечение достоверности данных во время исправления и обновления информационного фонда может быть организовано так, как это показано на рис. 12.8. Здесь предусматривается хранение трех поколений магнитных лент с массивами условно-посто-

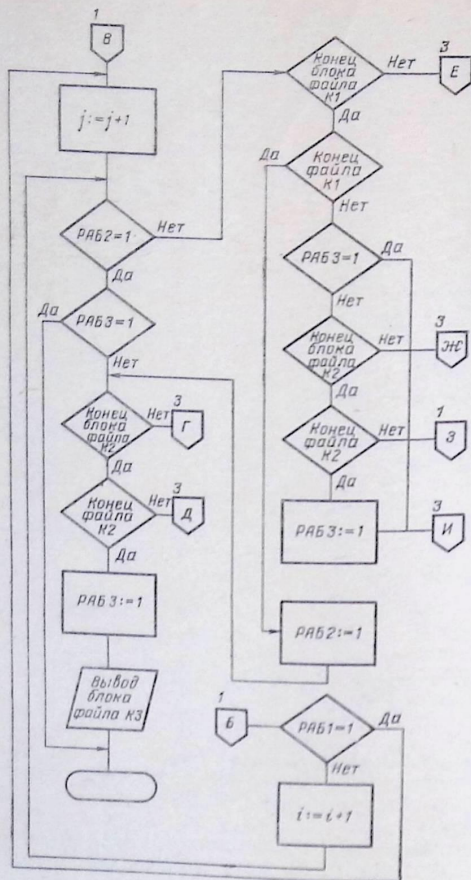


Рис. 12.7 (продолжение)

янной информации («ДЕД», «ОТЕЦ» и «СЫН») с одновременным хранением корректирующих массивов (например, на перфокартах). В системах обработки данных часто используется метод хранения лишь двух поколений массивов условно-постоянной информации, при этом такие массивы обязательно дублируются, например, на магнитных дисках и лентах.

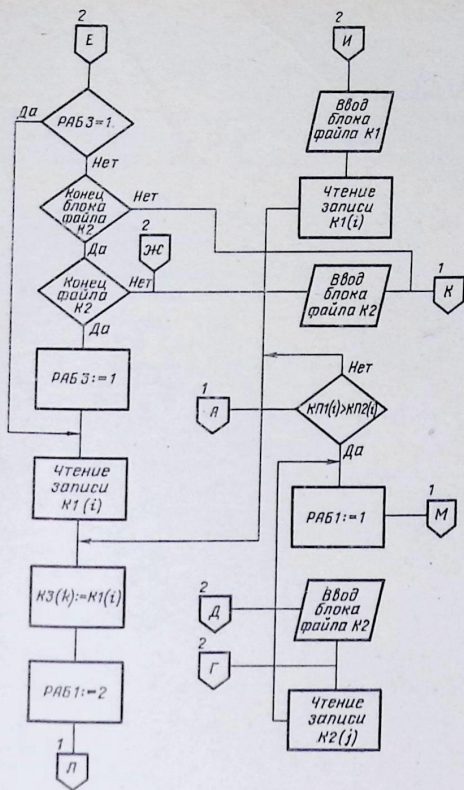


Рис. 12.7 (окончание)

Система внесения изменений в условно-постоянные массивы информационного фонда должна обеспечить абсолютную достоверность данных, поэтому в нее включаются различные процедуры контроля корректирующих массивов, распечатки (полной или фрагментарной) обновленных массивов, средства восстановления информационного фонда.

Анулирование старых оперативных массивов и загрузка новых — более простая процедура по сравнению с внесением изменений в условно-постоянные массивы. Она ничем не отличается от первоначальной загрузки массивов в информационный фонд. Наи-

более важным при этом является вопрос структурных изменений информационного фонда, если пользователю требуется его реорганизация. Оперативные массивы загружаются в определенные периоды времени в зависимости от частоты решения прикладных задач, в которых они используются.

Если оперативные данные накапливаются, то мы имеем дело с обновлением накапливаемых массивов. При этом используются два режима внесения изменений: количественное наращивание

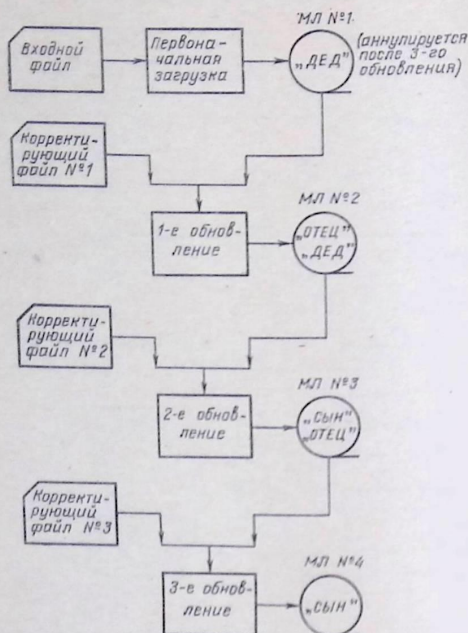


Рис. 12.8. Обеспечение достоверности данных при обновлении информационного фонда

записей накапливаемого массива (добавление новых записей) и арифметическая обработка записей старого накапливаемого и оперативного массивов с получением нового накапливаемого массива качественно другого содержания. Однако вне зависимости от режима накопления данная процедура практически мало чем отличается от рассмотренной нами корректировки условно-постоянных массивов. Некоторые особенности имеет внесение изменений в такие накапливаемые массивы, для формирования и ведения ко-

торых используются два или несколько оперативных массивов, тем более если последние загружаются в информационный фонд с различной периодичностью. Для таких накапливаемых массивов, как правило, приходится выполнять многократное частичное внесение изменений.

Глава 13

ФОРМИРОВАНИЕ И ВЫДАЧА РЕЗУЛЬТАТНЫХ ДАННЫХ

13.1. Технологические процессы обработки и формирования файлов результатных данных

Результатные данные формируются в процессе обработки оперативных, накапливаемых и промежуточных массивов информационного фонда и файлов базы данных. По форме представления они могут быть машинными, печатными и индицируемыми. На машинные носители записываются такие результатные данные, которые выводятся на печать или на экран терминального устройства после получения всех необходимых по конкретной задаче данных, а также те, которые используются при решении других задач. На печать или на экран выводятся результаты решения регламентированных задач и ответы на запросы (в режиме диалога).

Технология обработки файлов базы данных или информационного фонда и формирования результатов зависит от конкретной решаемой задачи. Можно выделить следующие типы задач, решаемых в системах обработки данных:

- 1) поиск элементарного ответа (в виде отдельной записи или ее элементов) на запрос пользователя и его выдачу;
- 2) поиск группового ответа (в виде группы записей или их частей) по запросу с выдачей его пользователю;
- 3) сложный поиск в нескольких файлах базы данных и формирование комбинированных ответов на запросы пользователей (с соответствующей арифметической обработкой или без нее);
- 4) математическую обработку всех записей одного файла;
- 5) математическую обработку нескольких файлов с подбором соответствующих записей по ключевым признакам;
- 6) логическое преобразование файлов данных с целью получения новых структурных элементов записей;
- 7) составные задачи, включающие в себя различные комбинации перечисленных процедур.

Таким образом, получение результатных данных всегда связано с поиском (выборочным или сплошным) и некоторыми преобразованиями информации. Организация поиска зависит от принятых в системе методов размещения файлов в базе данных, их структурного построения, способов организации информационной связи. Непосредственные преобразования данных определяются

содержанием конкретных прикладных задач управления. Получаемые в результате выполнения определенных преобразований выходные данные подвергаются редактированию, которое заключается в образовании форматов данных, пригодных для их вывода, и компоновке записей (строк) согласно принятому шаблону вывода. Иногда результатные данные перед их выводом подвергаются упорядочению (сортировке).

Типовость отдельных процедур обработки и формирования результатов данных позволяет создавать универсальные программные модули, из которых могут собираться различные рабочие программы для решения прикладных задач пользователя. При этом используются модули системного и прикладного программного обеспечения. Например, для решения задач, включающих в себя процедуры ввода, преобразования и вывода данных, часто используются генераторами ввода-вывода и генераторами отчетов. Для подготовки рабочих программ используется специальный входной язык генератора, а на языке принятой системы программирования составляются подключаемые модули пользователя. Исходная программа генератора записывается на специальных бланках 4—5 типов, среди которых можно выделить следующие:

- 1) описание входных данных (указывается характеристика всех входных файлов: имя, носитель, длина блока, длина записи и т. д.; описываются их состав, структура, иерархия составных элементов, их параметры, виды контроля и параметры, необходимые для их реализации);
- 2) описание выходных файлов (указываются способы организации файлов, тип и имя устройств, длина блока и записи, другие характеристики);
- 3) описание выходных данных (перечисляются все формируемые в выходные файлы записи, описываются их состав и структура, параметры отдельных элементов);
- 4) описание блока пользователя или вычислений (указываются точки входа блоков пользователя, их местоположение в библиотеке исходных или объектных модулей; могут описываться контролируемые ключевые признаки входных данных, по которым требуется накапливать определенные итоги, и т. д.).

Задачи, решаемые в АСУ в регламентированном режиме, характеризуются строгими требованиями к оформлению выводимых на печать результатных данных. Это прежде всего относится к текстовой (заголовочной) части выводимых табуляграмм. Наиболее рациональным является использование специального модуля вывода заголовка, выбирающего нужный текст из условно-постоянного справочного файла информационного фонда (по имени табуляграммы, типу листа и строки). В таком модуле предусматривается возможность оперативного включения в формируемую строку некоторых переменных данных (определенных группировочных признаков, даты, нумерации листов и др.).

Как известно, в современных АСУ и АСОД основная масса результатных данных, несмотря на все более широкое применение

терминальных устройств, выводится на печать. Эффективность же работы системы в определенной степени зависит от уровня прямых материальных затрат на выполнение расчетов. Одним из путей снижения затрат является экономия печатной бумаги как самого массового носителя информации. С этой целью следует не только разрабатывать рациональную структуру отдельных табуляграмм, предусматривать параллельный вывод двух или нескольких форм на одном листе, но и использовать двустороннюю печать. Для этого в программное обеспечение системы нужно включить специальный модуль, организующий вывод результатов постранично на лицевой и оборотной стороне бумаги. Прежде всего в основных программных модулях, формирующих поля печати, необходимо предусмотреть блок накопления в оперативной памяти строк одной стандартной страницы (например, по 68 строк со 127 символами в каждой) и ее вывода на магнитный носитель информации (диск, ленту). Для временного запоминания такой страницы потребуется 8636 байт внешней памяти. Наиболее рациональной является такая организация запоминания выводимых на печать данных, при которой на магнитном носителе формируются два файла: первый — для нечетных, второй — для четных страниц. Принципиальная схема фрагмента организации таких файлов в основном программном модуле показана на рис. 13.1. Универсальность рассматриваемой процедуры иногда может вызвать необходимость включения пустых (фиктивных) строк как в первый, так и во второй файлы, причем в последнем могут встречаться даже пустые страницы, если пользователю нужно получить группировочные табуляграммы по различным значениям признаков на отдельных листах. В этом случае первая страница новой группы обязательно печатается на лицевой стороне нового листа, т. е. последняя информационная строка предыдущей группы может оказаться не только на оборотной, но и на лицевой стороне предыдущего листа. Поэтому заполнение поля печати пустыми строками также необходимо предусмотреть в основном программном модуле в зависимости от изменения контролируемых группировочных признаков.

Подготовленные таким образом два файла по специальной программе могут выводиться на печать. При этом на лицевой сто-

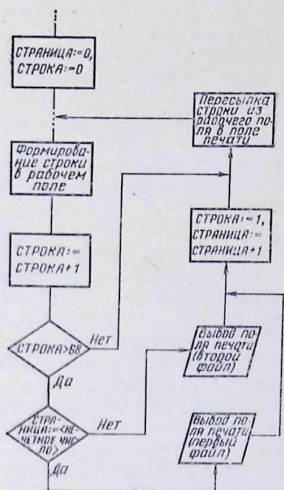


Рис. 13.1. Фрагмент алгоритма формирования выходных файлов печати

роне печатаются нечетные страницы (первый файл), а затем (после соответствующей установки бумаги) на оборотной стороне тех же листов печатаются четные страницы (второй файл). Такую универсальную программу печати можно построить с использованием стандартных модулей перезаписи файлов, входящих в состав системного программного обеспечения ЭВМ. Однако в ней следует предусмотреть возможность не только сплошного, но и порционно-го вывода страниц из обоих массивов, с тем чтобы при желании можно было отпечатать любое количество страниц, начиная с любой. Такая необходимость возникает при сбойных ситуациях: обрыве бумаги, выходе из строя печатающего устройства. Именно в таких случаях вывод одной табуляграммы осуществляется по частям, с повторной печатью страниц, при выводе которых произошли сбои.

Следует также отметить, что записанные на магнитный носитель результатные данные могут выводиться не только на печать, но и на терминальные устройства. Поэтому основные программные модули, формирующие результаты во внешней памяти ЭВМ и не связанные с печатью, могут успешно использоваться и в условиях телеобработки данных с применением дисплеев, абонентских пунктов и других средств.

Рассмотренная нами система организации программного обеспечения технологического процесса вывода результатов на печать может быть улучшена в части экономии внешней памяти ЭВМ, если в специальной программе печати будут предусмотрены параметры адаптации к структуре конкретных табуляграмм. Такими параметрами являются количество строк на одной странице и количество символов в строке, которые в рассмотренной системе нами принимались постоянными и максимально возможными. Программирование предварительного ввода этих параметров и настройки модуля по их значениям не представляет особых затруднений.

При проектировании технологического процесса особое внимание следует уделить вопросам обеспечения условий контроля и точности решения прикладных задач, а также вывода результатных данных. Так, например, в программных модулях, реализующих некоторые математические расчеты, предусматривается их контроль методом двойного счета, по усеченному алгоритму и т. д. Формируемые для вывода записи результатных данных контролируются по формату, а также по отклонениям, предельным значениям, по балансовому и перекрестному методу, по дополнительным связям. В системе должен быть предусмотрен контроль за продолжительностью решения задачи, позволяющий в случае чрезмерного увеличения времени решения сделать вывод о неправильной работе ЭВМ или заикливании программы. Процедуры всевозможного контроля выбираются в зависимости от уровня контрольных функций применяемых готовых системных и прикладных программ, специфики каждой задачи, наличия в ней сложных участков обработки, степени важности формируемых результатных данных.

13.2. Особенности организации обработки данных при решении отдельных задач и их комплексов

При проектировании АСУ большое значение имеет выделение в них отдельных подсистем, их типизация, определение состава задач. Расчленение многофункциональных систем управления на подсистемы более низкого уровня можно осуществить, исходя из конкретных условий и сложившихся принципов обработки данных. Признаками выделения подсистем можно принять функции управления, виды ресурсов или стадии производственного процесса. Однако такая разнородная декомпозиция системы не дает четкого представления о ней в целом, тем более, что опорным пунктом выделения подсистем и задач управления являются существующие, а не проектируемые информационные потоки. В связи с этим часто возникают значительные трудности постановки новых задач по тем экономическим расчетам, которые при существующей системе управления не производятся. При их отнесении к определенной подсистеме часто нарушаются принятые принципы классификации. Во избежание этого можно разработать типовой перечень задач, подлежащих решению на любом предприятии. Тогда в конкретных условиях производства возможно только сокращение или расширение такого предлагаемого перечня. Однако при этом содержательный аспект, глубина решения конкретных задач остаются в стороне и как будто никакого отношения к составу подсистем не имеют.

Применяется и другой подход к определению подсистем и отдельных задач, базирующийся на организации групп однородных структурных блоков АСУ, представляемых в трехмерном пространстве в виде куба, в котором можно выделить подсистемы по каждому из трех постоянных признаков: функциям управления, видам ресурсов и видам (стадиям) производства. При этом определяются три уровня декомпозиции АСУ, т. е. организуются комбинированные подсистемы и задачи, включающие в себя различные сочетания типовых значений перечисленных признаков. Такая функциональная структура системы управления, безусловно, является идеализированной. Реальная же структура вырабатывается при обследовании конкретного объекта путем определения идеализированных подсистем и блоков однородных задач, которые реализуются на данном объекте, исключения нереализуемых и добавления новых задач, выявленных в процессе обследования. Таким образом, структурное построение конкретной АСУ сводится в данном случае к корректировке идеализированной структуры АСУ, т. е. практически также опирается на типовой перечень подсистем и задач, хотя подготовка такого перечня имеет более четкую логическую основу, чем при первом рассмотренном нами подходе.

С учетом положительных и отрицательных сторон существующих наиболее распространенных направлений структурного построения систем управления наиболее целесообразным представляется такой метод типизации подсистем и состава задач, при котором:

предварительный анализ информационной системы производится с использованием некоторых элементов идеализированной структуры АСУ (построение и исследование информационных моделей на уровне документов и показателей);

на базе информационной модели системы управления на уровне показателей формулируются элементарные задачи управления, они объединяются в задачи (блоки, комплексы) более высокого уровня с учетом факторов, влияющих на вход и выход задач и условия их реализации.

Под задачей в самом общем виде понимается высказывание: «Дано A , требуется определить B по алгоритму $P(A)=B$ ». Элементарной будем считать такую задачу, в которой B является экономическим показателем. Таким образом, информационные связи между показателями, отраженные в модели системы (например, матричной), одновременно характеризуют взаимосвязь всех элементарных задач. Для построения такой модели (после детального изучения показателей, используемых при существующей системе управления) следует применить подход, главным образом сводящийся к тому, что после определения основных реквизитов-признаков и реквизитов-оснований составляется полный список показателей, которые теоретически могут быть получены в условиях функционирования АСУ. Для практического расчета и использования в целях управления из такого списка выбираются те показатели, которые действительно необходимы при достигнутом организационном уровне управления. Полный список показателей может быть получен автоматическим образом.

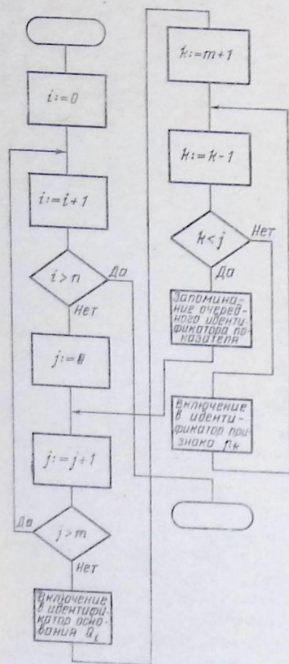


Рис. 13.2. Формирующая часть алгоритма генерации списка показателей (без операций ввода-вывода данных)

чески на ЭВМ. Формирующая часть алгоритма генерации такого списка представлена на рис. 13.2. Здесь Q_i обозначает основания ($i=1 \div n$), а p_k — признаки показателей ($k=1 \div m$).

Такой подход к структурному построению системы управления от ее низшего звена (элементарных задач) в сочетании с принципом составления полного перечня всех теоретически возможных и

выбора практически необходимых для управления показателей дает большие преимущества. Прежде всего они выражаются в том, что имеется возможность объединить элементарные задачи в более крупные блоки и комплексы на основе единого информационного, математического и программного обеспечения системы. Тем самым исключаются трудности, возникающие в связи с необходимой привязкой задач к заранее определенным подсистемам. Задачи более высокого уровня, состоящие из целого ряда элементарных задач, строятся при этом с учетом требований минимизации используемой информации, исключения ее дублирования, эффективной последовательности решения, а также максимального использования всех ресурсов ЭВМ и других средств вычислительной техники.

Мультипрограммный режим решения задач, реализованный в ЕС ЭВМ, позволяет организовать такие потоки задач, выполнение которых значительно повышает эффективность систем обработки данных. Как известно, любую задачу обработки можно разбить на три последовательно выполняемых (или циклически повторяемых) этапа: ввод входных данных, их обработку процессором (счет), вывод полученных результатов. Так как в большинстве экономических задач процедуры ввода и вывода (чаще всего на печать) выполняются медленно, то они могут быть совмещены с вычислительной работой процессора. Пример такого совмещения для работы, состоящей из 5 циклов, представлен на рис. 13.3. В варианте А предусматривается одновременное выполнение (за один цикл) ввода B_i для одной задачи, счета C_j — для другой и печати P_k — для третьей (на схеме медленные процедуры отмечены кружками). В варианте Б вместо процедуры печати для каждой задачи предусмотрены запись выходных данных на магнитный диск D_i (с использованием программных модулей, рассмотренных в 13.1) и последующий вывод этих данных на печать по специальной программе (при этом выполняются две процедуры: $\bar{D}_i$ — чтение дискового файла, P_i — двусторонняя печать). Как видно из представленной схемы, рассматриваемые 5 циклов процесса решения при варианте А охватывают 7 задач, а при варианте Б — 9 задач, причем одновременно выполняются соответственно 3 и 5 программ. Таким образом, процедуры счета, не связанные с печатью, позво-

	1	2	3	4	5
А	B_3	C_3	P_3	B_6	C_6
	P_1	B_4	C_4	P_4	B_7
	C_2	P_2	B_5	C_5	P_5
Б	B_5	C_5	D_5	$\bar{D}_5$	P_5
	P_1	B_6	C_6	D_6	$\bar{D}_6$
	$\bar{D}_2$	P_2	B_7	C_7	D_7
	D_3	$\bar{D}_3$	P_3	B_8	C_8
	C_4	D_4	$\bar{D}_4$	P_4	B_9

Рис. 13.3. Циклические процессы решения комплексов задач (2 варианта)

ляют более рационально загрузить ЭВМ, эффективнее использовать ее ресурсы.

Организация комплексного решения (комплексирование) задач осуществляется на основе их информационных связей, периодичности решения, а также в зависимости от требуемых ресурсов ЭВМ (как аппаратных, так и программных). Для определения очередности решения комплексированных задач им присваиваются соответствующие приоритеты. Так, например, в условиях ДОС ЕС ЭВМ в комплекс первого уровня могут включаться пять задач (трех различных приоритетов), три из которых являются активными, а две ждут подключения к входу системы в случае окончания или прерывания решения более приоритетных задач. Ана-

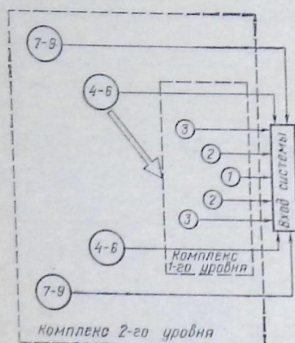


Рис. 13.4. Очередность активизации комплексированных задач (цифры указывают их приоритет)

логично могут быть организованы комплексы второго уровня, включающие по 5 комплексов первого уровня, и т. д. Схема активизации таких комплексов показана на рис. 13.4. Очевидно, что в таких стандартизованных комплексах могут быть и фиктивные задачи.

Последовательность решения задач определяют такие факторы, как информационная взаимосвязь между задачами, сроки выдачи результатов решения, общность использования единого информационного фонда и др. При определении последовательности решения большое значение имеют рациональность размещения файлов исходных и промежуточных данных на машинных носителях, упорядоченность используемых файлов (при последовательной их органи-

зации) и др. Поэтому при формировании потока задач всегда возможен выбор с целью минимизации общего времени решения.

При оперативном управлении работой ИВЦ довольно часто приходится изменять заранее предусмотренную последовательность решения задач, оперативно включать в поток поступающие новые, более важные задачи. Для этого можно использовать методы теории расписаний. При помощи таких методов оцениваются рациональность расписания решения, эффективность упорядочения задач, определяются классы эквивалентных заявок (заказов) на решение задач и отношения предпочтения между ними. При этом используются следующие критерии: максимальная длительность ожидания результатов с момента поступления заявки; среднее или максимальное запаздывание завершения обслуживания заявки сверх установленного срока или вероятность такого запаздывания; потери от ожидания или запаздывания.

В теории расписаний рассматриваются статические и дина-

ческие задачи: когда заявки поступают строго периодически, а расписания составляются каждый раз по всей их совокупности и когда на упорядочение задач влияют результаты реализации предыдущего расписания. Эти методы можно применять лишь при наличии таких характеристик для каждой задачи, как момент поступления, длительность обслуживания, допустимый интервал между поступлением и окончанием обслуживания (при превышении которого появляются потери), штраф за единицу времени ожидания. С методами теории расписаний более подробно можно ознакомиться в [25].

13.3. Основные характеристики различных режимов обращения к системе

К системе машинной обработки данных можно обращаться в различных режимах и с различными целями. От состава реализованных режимов функционирования (пакетного, с разделением времени, в реальном масштабе времени) зависят возможности обращения к системе. Можно выделить следующие основные виды обращения пользователей:

- 1) организацию и ведение информационного фонда (первоначальное формирование, обновление — регулярное в пакетном режиме и произвольное в режиме диалога, реорганизацию);
- 2) решение детерминированных по информационному содержанию и регулярных или случайных по времени задач;
- 3) обработку стандартных запросов в режиме диалога;
- 4) обработку произвольных запросов (как внутренних, так и внешних).

Первый вид обращений достаточно широко освещен в гл. 12. Реализация обращений второго вида в основном сводится к поиску и выборке нужных данных (для этого могут создаваться специальные рабочие файлы, если их использование предусмотрено программой задачи), их преобразованию и выводу последних. Регулярные задачи могут решаться под управлением специальной диспетчерской программы без вмешательства пользователя. Для этого создается система учета состояния программ и файлов данных, запоминания истории функционирования всей системы обработки данных.

В современных условиях для оперативного управления определенным объектом особый интерес представляет обработка данных по запросам пользователей в режиме диалога. Диалоговая система чаще всего строится по одному из двух вариантов (или комбинированно):

- 1) когда цель обращения пользователя к системе формулируется на основе его альтернативных ответов на запросы системы, при этом инициатором диалога является ЭВМ;
- 2) когда полностью сформулированный запрос пользователя

проверяется на корректность с использованием специально разработанных тезаурусов ведущих и подчиненных слов.

Для примера рассмотрим построение диалоговой системы первого типа, предназначенной для оперативного учета движения материальных ценностей на складах промышленных предприятий. После инициативного обращения пользователя к системе она должна последовательно выдавать ему запросы с возможными альтернативными ответами, с тем чтобы четко и однозначно определить

цель обращения, а затем выполнить те функции по обработке данных, которые были сформулированы в выбранных ответах пользователя. Такую систему можно представить в виде графа, показанного на рис. 13.5. Вершины графа представляют собой программные модули, формирующие определенные запросы и выполняющие соответствующие процедуры обработки данных, а его дуги соответствуют альтернативным ответам пользователя на запросы системы (пунктирные линии указывают рекурсивные пути, выбираемые автоматически, без вмешательства пользователя). В процессе диалога встреченные на пути от А до В программные модули последовательно включаются в рабочую программу ЭВМ, они настраиваются согласно ответам пользователя и выполняются. Тем самым реализуется функция цели обращения пользователя к системе. Вершины и дуги представленного графа имеют следующее содержание:

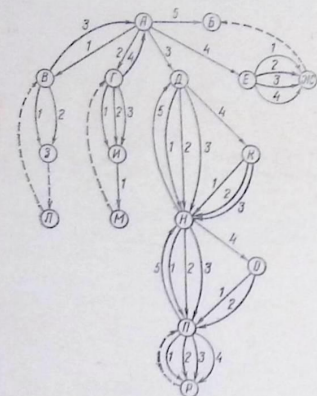


Рис. 13.5. Организация диалоговой системы оперативного учета движения материальных ценностей

А — режим работы: 1 — формирование базы данных (файла движения и остатков материалов или файла цен), 2 — корректировка базы данных (под корректировкой понимается внесение изменений в файл движения и остатков материалов в результате выполнения определенной учетной операции), 3 — поиск данных (переход к процедурам выдачи необходимых оперативных сводок по учету движения материалов), 4 — конец периода (переход к подведению итогов движения материалов за определенный планово-учетный период), 5 — конец работы;

Б — конец работы (прекращение связи с системой);

В — формирование базы данных: 1 — файл движения и остатков материалов, 2 — файл цен, 3 — конец формирования;

Г — операция корректировки: 1 — приход, 2 — расход, 3 — сторно, 4 — конец;

Д — уровень поиска и код: 1 — класс материалов (код — X

... X), 2 — группа (X ... X), 3 — подгруппа (X ... X), 4 — материал (X ... X);

Е — период: 1 — день, 2 — месяц, 3 — квартал, 4 — год;

Ж — конец периода (подведение итогов движения материалов за отчетный период с соответствующей корректировкой файла движения и остатков материалов);

З — установка носителя для формирования файла;

И — ввод данных по операции: 1 — данные;

К — показатель: 1 — количество, 2 — сумма, 3 — количество и сумма;

Л — формирование файла (его загрузка в информационный фонд);

М — корректировка файла;

Н — сводка: 1 — по приходу, 2 — по расходу, 3 — по остаткам, 4 — по обороту (приход и расход), 5 — конец;

О — охват: 1 — с остатком, 2 — без остатка;

П — период: 1 — за день, 2 — с начала месяца, 3 — с начала квартала, 4 — с начала года, 5 — конец;

Р — поиск (выборка нужных данных из файла движения и остатков материалов, формирование и выдача оперативной сводки пользователю).

Так, например, если в процессе диалога был пройден путь, который можно описать схемой (здесь в скобках указаны номера ответов пользователя) [А(2) — Г(1) — И(1) — М — Г(4) — А(5) — Б], т. е. по ответам (2,1,1,4,5) сформирован набор модулей (А, Г, И, М, Г, А, Б), то будет реализована процедура корректировки файла движения и остатков с учетом прихода определенного материала по данным, введенным пользователем с клавиатуры терминального устройства. Результатом выполнения программных модулей, собираемых в рабочую программу по схеме диалога [А(3) — Д(4, X...X) — К(3) — Н(4) — О(1) — П(2) — Р — ...], является получение оборотной ведомости указанного пользователем материала с начала месяца в количественном и суммовом выражении.

Как видно из рассмотренной методики организации диалогового режима оперативного учета движения материальных ценностей, ее применение значительно повышает оперативность учета, делает его более гибким и отвечает высоким требованиям современного уровня управления производством.

Диалоговые системы, строящиеся по второму варианту, предполагают использование тезаурусов входного языка запросов, ориентированного на конкретную предметную область. Основным является тезаурус ведущих (ключевых) слов. Именно в нем первоначально производится поиск слова, содержащегося в запросе пользователя. Структуру каждой записи такого тезауруса можно представить следующим образом: ведущее слово, его синонимы, условное обозначение, коды тезаурусов подчиненных слов, связанных с данным ведущим словом. Отметим, что тезаурусы подчиненных слов могут иметь иерархическую структуру. После нахождения ведущего слова (непосредственно или по синониму) запоми-

нается его условное обозначение и продолжается поиск слов, имеющих в запросе, в подчиненных тезаурусах. Если в запросе отсутствует определенное подчиненное слово, то система выдает пользователю уточняющий запрос с альтернативными ответами, формируемыми из соответствующего тезауруса подчиненных слов. Сформированный таким образом запрос пользователя (с уточнением или без него) принимает определенную стандартную форму, представляющую собой набор условных обозначений тезаурусных слов, которые соответствуют определенным процедурам (блокам) обработки данных. Они должны быть выполнены при формировании и выдаче ответа на запрос пользователя.

Некоторые особенности организации диалогового режима обработки данных присущи системам, использующим распределенные базы данных (см. 12.1). В таких базах все местные и центральный фонды данных должны иметь динамическую структуру. Для этого необходимо обеспечить их сопряжение, т. е. средства «переливания» информации, а также возможность обращения любого пользователя к любому фонду данных распределенной базы. Такая оперативная информационная связь может быть организована при соответствующем режиме работы системы во время обслуживания запросов пользователей. Для этого отдельные организации и предприятия, т. е. хранители местных фондов данных (будем называть их местными пунктами), должны быть обеспечены не только ЭВМ соответствующей мощности, но и необходимой сетью терминальных (дисплейных) устройств и графопостроителей. Все ЭВМ местных пунктов должны иметь каналную связь с цент-

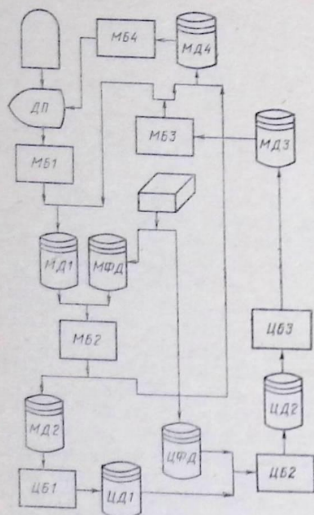


Рис. 13.6. Организация диалоговой системы в условиях распределенной базы данных

ральным пунктом (центром), хранящим центральный фонд данных и обеспечивающим обращение с одного местного пункта к фонду данных другого местного пункта, а также поддерживающим связь с внешней средой. Мощность ЭВМ центрального пункта необходимо подобрать с таким расчетом, чтобы обслуживание всех местных запросов пользователей происходило как можно оперативнее.

Исходя из изложенного, а также учитывая то, что запросы пользователей могут иметь не только стандартную, но и свобод-

ную или полусвободную форму, принципиальную функциональную схему диалоговой системы можно представить так, как это показано на рис. 13.6. В схеме приняты следующие обозначения:

ДП — дисплей;

МБ1 — МБ4 — блоки обработки данных в местном пункте;

ЦБ1 — ЦБ3 — блоки обработки данных в центральном пункте;

МД1 — МД4 — системы оперативных файлов, организуемых в местном пункте;

ЦД1 — ЦД2 — системы оперативных файлов, организуемых в центральном пункте;

МФД, ЦФД — соответственно местный и центральный фонды данных.

Как видно из схемы, в каждом местном пункте запросы пользователей, вводимые с помощью дисплея (ДП), принимаются, анализируются и стандартизируются (блок МБ1), в результате чего формируются файлы корректных и некорректных запросов, включаемые в систему МД1. Далее по корректным запросам в местном фонде данных организуется поиск ответов (блок МБ2), а по некорректным запросам через блоки МБ2, МБ4 и файлы систем МД2, МД4 пользователям передаются соответствующие сообщения. Если поиск ответов заканчивается успешно, то они через специальный файл системы МД4 распределяются по адресам пользователей (блок МБ4). В противном случае, когда нужных данных в местном фонде нет, в системе МД2 формируется файл необслуженных запросов, который передается в центральный пункт. Необслуженные запросы пользователей всех местных пунктов принимаются в централизованном порядке (блок ЦБ1) и записываются в систему файлов ЦД1. По ним производится поиск ответов в центральном фонде данных (блок ЦБ2). Найденные ответы через специальный файл системы ЦД2 распределительным блоком ЦБ3 передаются соответствующим местным пунктам, где формируется файл системы МД3, который через блок МБ3 подключается к файлу ответов системы МД4. Если в результате поиска в ЦФД все же остаются необслуженные запросы, то необходимо перейти к последовательному обращению к местным фондам данных (за исключением того местного пункта, откуда поступил необслуженный запрос) до нахождения ответа. Для этого в системе ЦД2 создается специальный файл необслуженных запросов, который через распределительный блок ЦБ3 разбивается на специальные файлы системы МД3 и передается соответствующим местным пунктам, где через блок МБ3 производится подключение необслуженных запросов пользователей других местных пунктов к файлу корректных запросов системы МД1. В блоке МБ2 найденные ответы на запросы пользователей других местных пунктов включаются в систему файлов МД2 и через центральный пункт направляются по нужному адресу (это должно быть предусмотрено в блоках ЦБ1, ЦБ2). Процедура поиска таких ответов может циклически повторяться при обращении к местным фондам данных других пунктов через центральный пункт, а после обращения ко всем

МФД, если ответ не найден, формируется специальное сообщение пользователю (блок ЦБ3).

Необходимость в «переливании» информации из местных фондов в ЦФД возникает в случае, если ответ на запрос пользователя одного местного пункта формируется в результате поиска в фонде данных другого местного пункта, т. е. когда определенные данные, хранившиеся только в одном местном пункте, оказываются нужными и для пользователей других местных пунктов. Для организации такого «переливания», которое практически сводится к расширению центрального фонда и сужению соответствующего МФД, в блоке МБ2 необходимо предусмотреть формирование в системе МД2 специального корректирующего файла местного фонда с данными по тем ответам, которые передаются пользователям других местных пунктов. По такому файлу производится удаление соответствующих данных местного фонда, передаваемых в ЦФД. Одновременно корректируется (пополняется) ЦФД; это должно быть предусмотрено в блоках ЦБ1, ЦБ2. Здесь используется накапливаемый корректирующий файл, формируемый в системе ЦД1. Однако следует отметить, что рассмотренное нами условие «переливания» данных практически не всегда приемлемо, так как единичные, редкие обращения из одного местного пункта в другой в этом случае могут вызвать чрезмерное, необоснованное расширение ЦФД. В связи с этим в рассматриваемую систему необходимо включить средства учета обращения и принятия решений по «переливанию» данных с учетом определенных критериев.

ПРОЕКТИРОВАНИЕ СИСТЕМ ОБРАБОТКИ ДАННЫХ

Глава 14

ПЛАНИРОВАНИЕ И ОРГАНИЗАЦИЯ РАБОТ ПО ПРОЕКТИРОВАНИЮ

14.1. Планирование работ по проектированию и внедрению проектов

Планирование создания систем обработки данных проводится по следующим основным направлениям: составление планов создания систем как мероприятий по внедрению вычислительной техники; составление плана-графика разработки и внедрения АСОД; разработка оперативно-календарных и других планов подготовки объекта к внедрению разработанных проектов.

Планирование внедрения вычислительной техники сводится к выполнению определенных работ министерствами (ведомствами), внедряющими и создающими (или обеспечивающими создание) системы, по составлению долгосрочных, среднесрочных и краткосрочных программ и планов. Составление таких планов реализации мероприятий по проектированию, подготовке к внедрению разработанных проектов и их непосредственному внедрению — это составная часть общего процесса планирования деятельности предприятий и министерств.

Последующее планирование (составление планов на более короткие сроки) опирается на программы и планы создания систем, и в первую очередь АСУ, на пятилетку. В них определяются основные задания по внедрению вычислительной техники в народное хозяйство.

Пятилетние планы детализируются при проведении среднесрочного планирования, выполняемого перед наступлением каждого года. В них конкретизируются те части работ, которые требуется выполнить в плановом году и которые обеспечены необходимыми ресурсами. Работы по выполнению наиболее ответственных мероприятий, как правило, включают в планы важнейших работ по новой технике отрасли, предприятий и организаций. Планируемые организационно-технические мероприятия должны найти соответствующее отражение в техпромфинплане предприятия, на котором эти мероприятия намечается внедрить.

В конкретизируемых планах должны быть определены номенклатура работ, объемы и сроки выполнения каждой из них. Основной для такой детализации служат планы-графики проектирова-

ния и внедрения мероприятий (АСУ, АСОД и др.), разрабатываемые как составная часть (или приложение) технического задания. В них указываются сроки и необходимые ресурсы, дается распределение работ между исполнителем и заказчиком.

Среднесрочные планы составляются таким образом, чтобы обеспечить внедрение разрабатываемого мероприятия в установленные сроки. Сроки выполнения конкретных планов определяются с точностью, не превышающей половины месяца. Планируемыми работами обычно являются процессы создания отдельных документов проекта или наиболее важных их разделов. При этом заказчик планирует работы на таком уровне, как они определены в планах мероприятий по подготовке объекта к внедрению.

Краткосрочное планирование — это разработка квартальных и месячных планов. Сроки завершения отдельных работ в таких планах устанавливаются с точностью до календарных суток.

Планы выполнения работ могут быть представлены в виде линейных или сетевых графиков. В линейных графиках все отдельные работы по разработке определенных документов перечисляются в виде списка. Здесь для каждой работы указываются ее объем и срок выполнения. Такая форма представления является наиболее простой и наглядной, но вместе с тем слишком статической. Более совершенной и лишенной указанного недостатка формой представления планов служит сетевой график. Здесь акцент делается не на работы, а на их завершение (на события). Сетевые графики позволяют выявлять возможные и необходимые сроки завершения работ, раскрывать резервы ресурсов, которые можно создать путем переноса некоторых работ на более поздние сроки. Кроме того, они дают возможность рационально решать следующие вопросы: определить последовательность выполнения и завершения работ; перераспределить ресурсы на отдельные работы так, чтобы максимально сократился срок завершения всех работ; планировать потребности в ресурсах и стоимость работ по отдельным этапам. Такой план служит важным инструментом для контроля хода выполнения работ. Преимуществом является и то, что для определения основных параметров сетевых графиков можно использовать ЭВМ. Принципы их построения и расчета параметров на ЭВМ общезвестны.

Методическое руководство работами обычно осуществляет головная организация отрасли — институт, проектное бюро и т. п. Каждая такая организация наряду с разработкой перспективных планов и другими общими работами обеспечивает всестороннюю совместимость АСОД, создаваемых в отрасли, осуществляет деятельность, направленную на повышение качества, снижение затрат и сокращение сроков создания и внедрения систем. Для этого разрабатываются необходимые инструктивные материалы, типовые проектные решения и пакеты прикладных программ с учетом специфики отрасли.

Непосредственную организацию работ выполняют министерство и предприятие (объединение) как заказчик. Они в пределах своей

компетенции отвечают за своевременное планирование и финансирование работ, разработку, согласование и утверждение проектов, своевременное выполнение мероприятий по подготовке объекта к внедрению этих проектов.

14.2. Организация обследования объекта и внедрения проектов

Внедрение систем обработки данных обычно связано с изменением документооборота, информационных потоков, применением новых методов выполнения некоторых функций. Эти изменения возникают по двум причинам: совершенствуется система информации в соответствии с требованиями и возможностями внедряемых методов и средств; уменьшается разнообразие, внесенное самими работниками по различным субъективным или даже объективным соображениям. Разнообразие может возникнуть в результате того, что лица, выполняющие соответствующие функции, вносят свои оригинальные методы и способы выполнения определенных работ или общепринятая методика упрощается, модифицируется и приспособляется для конкретных частных нужд на данном предприятии. Кроме того, много функций не имеют формализованного описания, обычно представляемого в виде инструкций, т. е. отсутствуют точные предписания по их выполнению. Для некоторых функций такие материалы в исчерпывающем виде вряд ли вообще могут быть разработаны. Поэтому к созданию АСОД необходимо активно привлекать соответствующих работников предприятия или даже создавать их специальные рабочие группы. Они при обследовании объекта, разработке основных решений, включаемых в техническое задание, при подготовке объекта к внедрению и внедрении проекта в большинстве случаев должны играть решающую роль. Рассмотрим основные виды работ на предприятии, разрабатываемом АСУ и являющемся ее заказчиком.

Создание АСУ, как правило, предполагает нечто большее, чем организация машинной обработки данных и, прежде всего, совершенствование системы управления. Однако основная часть работ на предприятии все же относится к организации обработки данных. Наиболее детально эти работы раскрыты в руководящих методических материалах по созданию АСУП (ОРММ АСУП). Работы по организации и координации создания систем обычно возлагаются на предприятие (объединение), для которого эта система создается. Оно должно обеспечить организационно-техническую подготовку объекта, включая приобретение всего необходимого оборудования, подготовку помещений и т. п. При выполнении работ разрабатывается организационно-распорядительная документация, составляются и реализуются планы оргтехмероприятий, заказываются и вводятся в эксплуатацию технические средства, подготавливаются кодификаторы технико-экономической информации, внедряются унифицированные единые формы конструкторской, технологической и другой документации и т. д. Отдельно следует вы-

делить работы по обследованию объекта и организации внедрения разработанного проекта.

Перед обследованием объекта проводятся соответствующие подготовительные работы. К ним относятся: подготовка и издание распоряжения о проведении обследования, разработка и согласование программы и методики, назначение руководителей, создание и подготовка рабочих групп для выполнения наиболее сложных и трудоемких работ, подготовка анкет для опроса работников, выполняющих конкретные функции, подготовка методики обработки материалов обследования. Качество выполнения этих работ во многом зависит от правильности и точности подбора вопросов, на которые необходимо получить ответы.

При обследовании заказчик должен организовать и осуществить работы по изучению организационной структуры объекта управления, выявлению автоматизируемых функций (комплекса задач), изучению потоков входной и выходной информации и соответствующего документооборота, источников и пользователей данных, по изучению методов решения задач, обработки данных и использования вычислительной техники. Они обычно выполняются отдельными работниками или их группами, поэтому требуют четкой организации их проведения и определенности в том, какие материалы и в каком виде каждый из исполнителей должен представить.

Обработка, анализ и окончательное оформление результатов обследования обычно осуществляются при активном участии разработчиков и в организационном отношении особых трудностей не представляют. При этом следует предусмотреть разработку и оформление функциональной схемы объекта, схем организационной структуры, обмена информацией при существующей системе, произвести оценку объемно-временных характеристик потоков информации. Заказчик должен составлять и реализовать планы по подготовке объекта к внедрению разрабатываемых мероприятий, включая подготовку и создание информационной базы, рассмотрение и утверждение проектной документации. Для этого необходимо широко использовать имеющиеся на предприятии нормативные акты, методически составленные сводные документы (планы, отчеты и т. п.).

Работы по подготовке объекта к внедрению в основном сводятся к приобретению и пуску в эксплуатацию технических средств, подготовке информационной базы и специалистов для работы в новых условиях и проведению организационных мероприятий по обеспечению функционирования системы.

На стадии разработки технического проекта предприятие должно обеспечить разработчика дополнительными исходными данными, необходимыми для проектирования, издать приказы об организации рабочих групп по подготовке информационной базы и проведению других подготовительных работ, организовать обеспечение их выполнения, рассмотрение, согласование и утверждение технического проекта.

На стадии рабочего проектирования заказчик должен завершить формирование информационной базы (подготовить технические и программные средства для ее формирования и ведения, организовать их приемку и внедрение).

На стадии внедрения работы должны быть организованы так, чтобы заказчик в установленные сроки завершил выполнение организационно-технических мероприятий подготовки предприятия к внедрению проекта, организовал опытную эксплуатацию, проведение необходимых испытаний и сдачу системы в промышленную эксплуатацию. Номенклатура этих работ соответствует содержанию разрабатываемой документации.

14.3. Организация проектирования

Проектная или научно-исследовательская организация (разработчик) свою основную деятельность направляет на выполнение проектных работ и разработку соответствующей документации. На предпроектной стадии разработчик лишь принимает участие в работе головной организации отрасли и предприятия. Аналогично разработчик активного участия не принимает и на стадии внедрения, так как все организационные и другие подготовительные работы, которые необходимо выполнить при подготовке к испытанию разработанных программ, процессов и т. д., а также при опытной эксплуатации, входят в обязанности заказчика.

Непосредственную работу по проектированию выполняет разработчик. Эта работа, по существу, сводится к созданию необходимой проектной документации.

Проектирование как процесс целенаправленной деятельности различных групп специалистов требует определенной организации. Организация проектирования машинной обработки данных при создании АСУД достаточно четко определена общепромышленными руководящими методическими материалами, а также разрабатываемыми (при необходимости) для каждой отрасли своими методическими материалами.

Для организации работ должен быть назначен главный конструктор проекта (руководитель темы), который несет ответственность за качество проектирования и соблюдение сроков завершения работ. Его функции, а также права и обязанности определяются действующими положениями и инструкциями.

Непосредственное проектирование выполняют специальные подразделения проектной организации. Большие объемы и различия в содержании работ требуют создания соответствующих специализированных отделов, что позволяет обеспечить не только разделение труда, но и проведение определенной методической работы по унификации и стандартизации отдельных составных частей проектов, разрабатываемых в различных отделах. Такая специализация особенно оправдывается при выполнении работ по постановке задач, при разработке информационного и программного обеспечения. В отделах, обычно специализирующихся по обеспечению

вающим подсистемам, могут быть выделены отдельные группы, выполняющие определенные специфические работы. Например, в отделе постановок задач могут быть выделены группы по видам экономико-математических или информационных моделей задач (по постановке оптимизационных задач, задач прямых расчетов и т. п.). В отделе информационного обеспечения могут быть созданы группы по разработке языков описания данных, построению информационных моделей и структур баз данных, по типизации и выбору средств программного обеспечения баз данных и т. д.

Для проектирования могут также создаваться смешанные группы. При таком методе организации работ можно установить более тесные контакты между разработчиками различных видов обеспечения системы. Так, например, постоянные контакты очень важно поддерживать между постановщиками задач и программистами при разработке оригинального программного обеспечения. Этот способ распространен в организациях, в которых объемы работ сравнительно невелики, например в проектно-конструкторских бюро отраслевых министерств.

Выбор структуры подразделений и распределение функций между ними зависят от профиля организации, выполняющей проектировочные работы, от объемов работ и, конечно, от степени развития в данной организации методических и научно-исследовательских работ.

Проект конкретной АСУ может разрабатывать одна или несколько организаций, однако в любом случае необходимо составлять общие планы выполнения всех этапов работ и конкретные планы на каждый этап, завершающийся оформлением или утверждением соответствующей документации.

При наличии нескольких разработчиков особенно много внимания уделяется разработке методических материалов по содержанию и форме представления конкретной проектной документации. Например, при проектировании информационной базы следует использовать единые системы кодирования не только для файлов и документов, но и для их составных частей, в первую очередь реквизитов и показателей. Это позволяет избежать дополнительных трудностей при объединении отдельных частей проектов в единое целое, дает возможность применять ЭВМ в процессе проектирования.

Глава 15

РАЗРАБОТКА И ДОКУМЕНТИРОВАНИЕ ПРОЕКТОВ

15.1. Состав организационно-распорядительной документации

Результаты проектирования представляются в соответствующим образом оформленных предпроектных и технических документах. В отдельных случаях, когда необходимо построить или реконструировать здания и помещения, также разрабатывается проектно-сметная документация. Кроме того, для выполнения подго-

товительных, проектировочных и других работ, а также разработки указанной документации как основного продукта проектирования создается организационно-распорядительная документация, в которой определяются организационно-правовые и хозяйственные взаимоотношения между сторонами, принимающими участие в создании и внедрении системы. Соответствующие документы разрабатываются на каждой стадии создания АСОД и АСУ, поэтому обычно выделяется документация предпроектной стадии, технического проекта, рабочего проекта (технорабочего проекта) и внедрения.

На предпроектной стадии оформляются следующие организационно-распорядительные документы: договоры на создание системы, разработку предпроектных документов, а также поставку технических средств и обучение (переподготовку) персонала вычислительного центра; приказ об организации работ по подготовке объекта к внедрению системы; планы-графики проведения занятий с персоналом (пользователями информации), разработки классификаторов, приведения конструкторско-технологической документации в соответствие с ЕСКД и ЕСТД.

На стадии подготовки технического и рабочего проектов разрабатывается аналогичная организационно-распорядительная документация (как и на предпроектной) и, кроме того, дополнительные приказы, договоры, планы-графики, а также составляются протоколы. Последние оформляются после согласования постановок задач, принятия зачетов по результатам обучения системам программного обеспечения, рассмотрения технического проекта на научно-технических советах разработчика и заказчика. Кроме того, составляются акты, которыми фиксируется готовность нормативно-справочной документации, проверка на контрольных примерах и приемка в опытную эксплуатацию комплексов задач, выполнение организационно-технических мероприятий по подготовке предприятия к внедрению системы.

При внедрении системы оформляется организационно-распорядительная документация двух типов, связанная с завершением работ по подготовке предприятия к внедрению системы и по вводу ее в эксплуатацию. В группу документов первого вида входят акты завершения работ по оформлению всей необходимой документации, ее заполнению, переносу данных на машинные носители, загрузке информационной базы, ее проверке, а также проверке программ и инструкций на контрольных примерах и в реальных условиях, а при необходимости — и протоколы согласования отступлений от проекта. Документы второго вида — это приказы и планы-графики проведения совместных работ заказчика и разработчика по опытной эксплуатации технических средств и комплексов задач, акты приемки всей системы в эксплуатацию.

15.2. Обработка материалов обследования объекта и подготовка задания на проектирование

На предпроектной стадии создания системы необходимо выбрать круг задач, подлежащих автоматизации, обосновать целесо-

образность и наметить пути реализации соответствующих мероприятий. Работа выполняется путем обработки и анализа материалов обследования объекта с широким изучением литературных источников и опыта, накопленного при реализации аналогичных мероприятий на других объектах. При этом необходимо: выявить возможности обследуемого объекта улучшить соответствующие виды деятельности; провести оценку пригодности, эффективности и качества существующих типовых проектов, проектных решений и других средств по основным видам обеспечения системы (математическому, информационному, программному и т. д.); определить состав мероприятий, обеспечивающих реализацию предлагаемых разработок.

Среди работ, проводимых на предпроектной стадии создания системы, в методическом плане наиболее сложными и важными являются те, которые связаны с обработкой материалов обследования объекта. От качества и полноты этих работ во многом зависит качество всего внедряемого мероприятия. Важность предпроектной стадии подтверждают те теоретические и методические работы, которые широко проводятся ведущими научно-исследовательскими и проектными организациями по внедрению вычислительной техники в народное хозяйство.

Обработка материалов обследования без использования вычислительной техники чаще всего сводится к разработке таблиц, графиков и схем, характеризующих уровень качества функционирования изучаемого объекта, имеющиеся резервы для улучшения работ по конкретным направлениям. Такая обработка обычно проводится путем использования методов и методик, применяемых для анализа хозяйственной и другой деятельности объекта управления за год или другой период времени.

Большие объемы работ по обработке материалов обследования, сжатые сроки, в которые обычно необходимо их выполнить, а также характер самих работ неизбежно требуют их автоматизации. Наиболее широкое развитие получили следующие направления автоматизации обработки материалов, собранных при обследовании объекта: обработка анкет обследования с использованием методов математической статистики; построение и машинный анализ информационных моделей, представляющих собой схемы пространственно-временных связей в системе управления.

Анкетный метод обследования широко применяется не только при создании АСОД, но и при проведении различных социологических и других исследований. Для обработки материалов таких исследований с целью выявления различных возможностей взаимосвязей между исследуемыми объектами или событиями разработаны специальные программные средства в виде стандартных программ и пакетов прикладных программ.

Применение анкетного метода прежде всего связано с разработкой первичного документа — анкеты опроса. Остальная работа, по существу, является организационной и технической. При этом необходимо отметить, что качество всей работы зависит от

того, какие вопросы включаются в анкету, насколько они являются конкретными и понятными тем, кто будет на них отвечать. Кроме того, форма анкеты должна быть приспособлена к машинной обработке, т. е. построена с учетом требований машинной обработки (см. 9.3).

На основании материалов, получаемых при обследовании объекта, выявляются объемы информации и ее потоки, циркулирующие на объекте управления, внутри его подразделений, между ними, а также между объектом управления и внешней средой. Детальное изучение этих потоков способствует выявлению состава и структуры информации, мест ее возникновения, пользователей, качественных и количественных требований к первичной и результирующей информации; это позволяет определить возможности упрощения структуры управления, выбрать более рациональные маршруты движения информации.

Результаты обследования информации объекта формализовано могут быть представлены в виде информационных моделей. При построении таких моделей можно использовать различную степень обобщения, абстракции и различные способы их описания. На объект управления вообще можно смотреть как на физическую систему. Тогда ее можно описать при помощи физических моделей. Такие модели могут быть разработаны на предпроектной стадии. На их основе можно определить объемы, а также места возникновения и использования информации. Для выявления информационных связей между подразделениями объекта управления, носителями информации и т. п. необходимо разрабатывать информационные модели (объект управления представить как информационную систему).

Информационные модели могут быть представлены в виде таблиц, специальных матриц, графиков, математических выражений, определяющих связи между исходными и результирующими структурными единицами информации, подразделениями объекта управления.

Вопросы построения информационных моделей широко рассматриваются в литературе. Наиболее распространенными являются матричные модели [31]. Они строятся на основе схем связей различных уровней. Поэтому они и отражают информационные связи между документами, показателями, подразделениями или работами (задачами). Общая информационная модель имеет четыре основных и два дополнительных блока (раздела). Наибольший интерес при определении информационных связей представляет первый раздел модели — первый квадрант. Его элементы характеризуют непосредственные связи между исследуемыми элементами, а их совокупность может быть представлена в виде матрицы прямых связей, по которой можно построить ориентированный граф. По отношению к данному графу матрица прямых связей представляет собой матрицу смежности, элементы которой соответствуют его вершинам. Поэтому методы анализа информационных связей между исследуемыми элементами можно основывать на использо-

вании свойств графов и матриц смежности. Суть использования матрицы смежности рассмотрим на примере анализа расчетов, выполняемых при решении некоторой функциональной задачи АСУ.

Связи между информационными элементами задачи можно представить, используя специальные матрицы. Если все элементы системы пронумеровать по порядку, то матрицу прямых связей между ними можно записать в виде квадратной матрицы S :

$$S = \begin{pmatrix} 0 & S_{12} & \dots & S_{1n} \\ S_{21} & 0 & \dots & S_{2n} \\ \dots & \dots & \dots & \dots \\ S_{n1} & S_{n2} & \dots & 0 \end{pmatrix},$$

где n — число элементов системы;

$$S_{ij} = \begin{cases} 1, & \text{если элемент } j \text{ непосредственно зависит от элемента } i; \\ 0, & \text{если элемент } j \text{ непосредственно не зависит от элемента } i. \end{cases}$$

При составлении такой матрицы принимается, что каждый элемент может быть связан с любым другим элементом, кроме самого себя, и что между любыми двумя элементами отсутствует обратная связь (например, когда от элемента a_1 зависит элемент a_2 , от него — элемент a_3 и т. д., а от элемента a_i — элемент a_1).

Используя матрицу S , можно определить свойства связей элементов и правила исследования:

1. При изменении порядка нумерации (записи) элементов матрица S не меняется, только необходимо элементы перенумеровать так, чтобы их новые номера соответствовали тем же строкам и столбцам.

2. Если в столбце j все его элементы равны нулю, то соответствующий его номеру элемент не зависит ни от одного другого элемента и является исходным.

3. Если в строке i все ее элементы равны нулю, то от соответствующего номеру этой строки элемента не зависит никакой другой элемент, и тогда возможны два случая:

если в соответствующем столбце j ($j=i$) имеется хотя бы один ненулевой элемент, то соответствующий этому номеру элемент является результирующим;

если в столбце j ($j=i$) все элементы равны нулю, т. е. элемент, соответствующий номеру строки i , не связан ни с одним другим элементом, то он является изолированным и автономным (автономной подсистемой системы, описываемой данной матрицей смежности).

4. Так как при определении матрицы S было принято, что внутри задачи обратные связи отсутствуют, то ее можно привести к треугольному виду, когда под диагональю (слева) располагаются нулевые элементы. Тогда элемент $(i+1)$ будет зависеть только от элементов, имеющих номера от 1 до i (это достигается путем соответствующего изменения номеров элементов системы).

5. Если перенумеровать элементы таким образом, чтобы все ну-

левые столбцы оказались бы впереди, тогда матрицу S можно разбить на четыре части:

$$S = \begin{pmatrix} 0 & S_{12} \\ 0 & S_{22} \end{pmatrix},$$

где S_{12} — матрица порядка $m \times (n - m)$, отражающая связи между исходными и результатными элементами (m — число исходных элементов, n — число всех элементов);

S_{22} — матрица порядка $(n - m) \times (n - m)$, указывающая связи между результатными элементами.

Так как матрица S отражает непосредственные связи между элементами системы, их можно назвать связями первого порядка. Связи порядка i будет отражать матрица S^i :

$$S^i = \begin{pmatrix} 0 & S_{12} \cdot S_{22}^{i-1} \\ 0 & S_{22}^i \end{pmatrix}.$$

Полные связи между элементами укажет матрица Z :

$$Z = \sum_{i=1}^k S^i,$$

где k — наибольший возможный порядок связей между элементами.

Если при исследовании информационных связей необходимо установить только наличие связей любого порядка (от 1 до k), а число этих связей определить не требуется, то при построении матрицы S^i операцию умножения матрицы S_{12} на матрицу S_{22}^{i-1} можно заменить на логическую операцию конъюнкции, а при построении матрицы Z — суммирование на логическую операцию дизъюнкции.

6. Если строки и столбцы матрицы S можно соответственно переставить так, чтобы она приняла вид

$$S = \begin{pmatrix} S'_{11} & 0 \\ 0 & S'_{22} \end{pmatrix},$$

тогда среди представленных ее элементов можно выделить две независимые между собой части, которым соответствуют матрицы структуры S'_{11} и S'_{22} соответственно порядка $(m' \times m')$ и $(n - m') \times (n - m')$, где m' — количество элементов в первой части. Аналогично можно выделить и большее число автономных частей (при наличии таковых).

Теоретические вопросы построения информационных моделей, методы и средства их исследования разработаны достаточно полно. К существенным причинам, тормозящим их широкое практическое использование, следует отнести трудности, связанные с подготовкой исходных данных, т. к. при этом необходимо провести глубокий анализ их полноты, выявить и устранить обратные информационные связи, привести эти данные к форме, пригодной для их непосредственного применения.

Результаты работ, проводимых на предпроектной стадии, оформляются в виде двух предпроектных документов: технико-экономического обоснования системы (как приложения) и технического задания на ее создание.

Технико-экономическое обоснование состоит из следующих разделов: исходных положений, характеристики и технико-экономических данных объекта; обоснования цели создания системы; обоснования задач, их комплексов и средств системы; перечня организационно-технических мероприятий; выводов и предложений. В нем представляются материалы, полученные в процессе проведения обследования объекта, и их обобщение: характеристики организационной и производственной структуры объекта управления, его основные технико-экономические показатели; перечень функций, для выполнения которых необходимо организовать обработку информации; характеристики и направления совершенствования методов планирования, оперативно-диспетчерского управления, учета и анализа; основные критерии эффективности для оценки вариантов автоматизации управления; перечень и обоснование автоматизируемых функций, применяемых типовых проектных решений и других разработок по организационному, техническому, информационному, математическому и программному обеспечению и др.

Техническое задание как основной утверждаемый документ состоит из следующих разделов: основания для разработки; общих положений; основных требований; материалов, используемых для разработки; стадий разработки и внедрения; организации работ и исполнителей; расчета экономической эффективности (а для АСУ — и научно-технического уровня). В нем должны быть представлены описания основных целей, функционального назначения создаваемой системы (разрабатываемых мероприятий), а также состав и перечень автоматизируемых задач (комплексов задач) с их укрупненными характеристиками и указанием влияния автоматизации их решения на повышение эффективности производства, состав и структура применяемых технических и программных средств (в том числе ТПР и ППП), описание информационного обеспечения, структуры баз и банков данных. В техническом задании также определяются стадии разработки с детализацией работ по ним, организация и финансирование их проведения и т. д.

15.3. Разработка технического и рабочего проектов

Процесс разработки технического и рабочего проектов по существу сводится к созданию соответствующей проектной и организационно-распорядительной документации. Рассмотрим ее состав и содержание на примере соответствующих документов, определенных для АСУП.

Документы, разрабатываемые на стадии технического проектирования, можно разделить на общесистемные и документы комп-

лексов задач (отдельных задач). К первым относятся: пояснительная записка, решения по комплексу технических средств, описания организационной структуры управления, организации информационной базы, систем классификации и кодирования технико-экономической информации, расчет экономической эффективности.

Пояснительная записка — это наиболее важный документ технического проекта, в котором представляются основные характеристики объекта (технико-экономические показатели, краткое описание существующей организационной структуры управления, цель создания и область применения системы), а также основные технические решения по всем видам обеспечения и их краткая характеристика. В документе также указываются основание для разработки, перечень организаций, разрабатывающих систему, сроки выполнения основных стадий и этапов всех работ. В других общесистемных документах детализируется то, что представлено в пояснительной записке. Так, при описании решений по отдельным видам обеспечения системы могут быть приведены необходимые расчеты, обоснования по выбору вариантов соответствующих проектных решений, основная характеристика используемых готовых средств (информационного, технического и программного обеспечения). На все документы составляется ведомость («Ведомость документов технического проекта»).

Отдельно следует выделить документ расчета экономической эффективности. В нем представляются основные показатели эффективности, исходные данные, необходимые для их определения, и проведенные расчеты. Документ разрабатывается по утвержденной методике. Для определения экономической эффективности АСУП используется методика, изложенная в [35].

Документ «Описание постановки и решения задачи (комплекса задач)» разрабатывается для каждой отдельной задачи или комплекса задач. Он обычно состоит из нескольких разделов: отдельно выделяется организационно-экономическая сущность задач, математические модели, описание входных, выходных и промежуточных данных, а также алгоритмы решения (представляются как подразделы общего раздела «Алгоритм решения»); в виде приложения к документу представляется контрольный пример решения.

В описании организационно-экономической сущности должны быть в неформализованном виде указаны все сведения, которые необходимы экономистам и организаторам производства при рассмотрении задачи или использовании результатов решения. В данном разделе конкретизируются: цель автоматизации решения комплексов задач (функций управления), ее экономическая целесообразность; перечень структурных подразделений и основных функций их информационной деятельности, связанной с подготовкой исходных данных и использованием результатов решения; описание вновь вводимых функций; предлагаемые методы и варианты решения; методические и другие материалы, на основании которых построены экономико-математические модели задач, содержание расчетов с выявлением специфических ситуаций; организация

представления оперативных данных, получения результатов с указанием периодичности решения; требования к точности и надежности результатов с описанием методов проверки их достоверности и полноты и пр.

На основании материалов, представленных в разделе «Алгоритм решения», разрабатываются оригинальные или подбираются имеющиеся средства программного и информационного обеспечения, производится «стыковка» по работе специалистов по постановке задач, организации информационной базы и программированию, а также экономистов и организаторов производства с ИВЦ. Поэтому в данном разделе должны быть представлены экономикоматематические модели задач, структура массивов (файлов) данных, формы кодограмм сообщений, макеты видеogramм, общие схемы процессов обработки данных, решения задач и алгоритмы реализации отдельных процедур.

Для того чтобы последующую разработку программ могли осуществить другие лица, очень важно детально описать структуру файлов, методы организации данных, составить блок-схемы алгоритмов, для которых требуется подготовить оригинальные программы, тщательно их проверить на контрольных примерах. Практика показывает, что такая обязательная проверка позволяет существенно повысить качество материалов по постановке задач и алгоритмам их решения.

На стадии рабочего проектирования составляются программные документы, технологические инструкции по обработке данных, инструкции по работе в условиях функционирования разрабатываемой системы (внедряемого мероприятия).

Программные документы оформляются по государственным стандартам, объединенным в Единую систему программной документации (ЕСПД). Каждый программный документ состоит из следующих условных частей: титульной, информационной, основной и регистрации изменений. ЕСПД предусматривает разработку таких программных документов: спецификации, текста программы, описания программы, формуляра, общего описания, ведомости эксплуатационных документов, руководства системного программиста, руководства программиста, руководства оператора, описания языка.

Технологические инструкции разрабатываются для следующих операций и работ: сбора, регистрации, предварительной обработки, контроля и выдачи данных; по созданию, ведению и изменению первичной, нормативно-справочной и другой документации; по передаче информации по каналам связи, приему в ИВЦ, переносу на машинные носители, подготовке массивов для их ввода, обработке на ЭВМ; по созданию и обслуживанию файлов информационных фондов (баз данных), архивов перфотек и магнитотек.

В должностных инструкциях устанавливаются права и обязанности для работников, которые обслуживают систему, формируют исходные данные и пользуются результатами, полученными машинным способом.

16.1. Основные направления типизации и стандартизации проектирования

Проектирование машинной обработки данных, содержание и методы выполнения которого рассмотрены в предыдущих главах, представляет собой сложный процесс, обычно требующий больших трудовых и денежных затрат. Поэтому с расширением масштабов внедрения СВТ все более важными становятся вопросы поиска путей сокращения этих затрат. Затраты на создание АСОД и АСУ можно сократить, если использовать ранее разработанные проекты или их части. Так можно поступить лишь в случаях, когда имеют место идентичные ситуации и с помощью ранее выполненных разработок можно достичь желаемых результатов, что практически бывает редко. Поэтому более эффективным может быть путь, когда уже при разработке проектов или отдельных их частей учитываются особенности их возможного использования для определенной группы объектов: подразделений и служб предприятия, группы предприятий, отрасли, всего народного хозяйства. Такие проектные решения могут быть приняты едиными или использоваться как типовые, требующие лишь некоторых дополнений, уточнений или привязки. Содержание, область применения и другие характеристики этих решений могут быть установлены в документах, оформляемых в виде соответствующих стандартов предприятия, отраслевых, государственных (ГОСТ), или в руководящих технических материалах (РТМ). Такими стандартизованными типовыми материалами являются, например, унифицированные системы плановой, учетной, отчетной, организационно-распорядительной, финансовой, конструкторской и технологической документации. Типизации также подвергаются системы классификации и кодирования технико-экономической информации (см. 7.3).

В связи с развитием АСУ проводятся работы по стандартизации терминов, определений, содержания и организации проектирования, структуры и содержания проектной документации, графического представления материалов в ней, классификации АСУ и их элементов. Современному этапу характерно стремление к совершенствованию структуры АСУ и процессов их разработки, базирующемуся на накопленном положительном (и отрицательном) опыте с ярко выраженной тенденцией расширения типизации и стандартизации.

Объектами типизации становятся все более сложные элементы систем: кроме кодов технико-экономической информации, форм первичных документов, макетов машинных носителей, способов размещения данных, операций технологического процесса и отдельных программ в настоящее время типизации и стандартизации подвергаются задачи, их комплексы, функции управления, процес-

сы подготовки и ведения файлов информационного фонда системы, процедуры решения задач. Применение ЭВМ третьего поколения, имеющих запоминающие устройства прямого доступа сравнительно большого объема, развитые общесистемные программные средства, позволяет использовать более совершенные методы организации файлов, дает возможность полнее осуществить системный подход к организации части информационного обеспечения системы в виде базы данных с реализацией современных концепций банков данных. Это непосредственно связано с возможностью перехода к автоматизации проектирования. Основными направлениями совершенствования проектирования АСОД и АСУ обобщенно следует считать: 1) типизацию структурных элементов АСУ; 2) разработку баз и банков данных; 3) разработку дополнительного и модернизацию общего и специального программного обеспечения; 4) разработку типовых операций и процессов обработки данных; 5) разработку методов и средств по автоматизированной обработке данных обследования объекта управления.

Первое направление ведет к уменьшению неоправданного разнообразия систем, включению их в определенную типовую группу. Другие направления требуют разработки и использования специальных способов представления данных и средств программного обеспечения, прикладных программ, а также типовых технологических процессов. При этом типизация и автоматизация проектирования должны предусматривать реализацию следующих основных принципов:

1) обеспечения решаемых задач исходными данными на основе общей системы хранения информации, построенной таким образом, чтобы она была независимой от количества и содержания решаемых задач;

2) построения единых схем обмена данными между системой и пользователями;

3) использования единых форм производственно-хозяйственных документов, их приспособленности для людей и ЭВМ;

4) обеспечения универсальности способов информационного отображения производственно-хозяйственной деятельности;

5) применения средств и методов проектирования, основанных на принципах блочных и модульных систем.

По степени охвата автоматизируемых функций управления все широко известные методы автоматизации проектирования обычно делятся на элементные, подсистемные, объектные и модельные.

Суть *элементного* метода заключается в том, что заранее разрабатывается определенное множество элементарных типовых частей системы, а затем из них строится проектируемая система. Такие элементы включаются в специальную библиотеку, которая при необходимости может пополняться новыми элементами. Каждый элемент в библиотеке может быть представлен несколькими вариантами, отличающимися применяемыми средствами, методами реализации, временными или другими характеристиками. Примером реализации такого метода служат типовые проектные решения

АСУП, разработанные в 1971—1975 гг. на основе достижений, имевшихся в то время в области постановки и типизации задач и их частей (программных модулей) в условиях применения ЭВМ второго поколения. Они не утратили своей актуальности и по сей день. При разработке ТПР была создана проектная документация по экономическому, программному и техническому обеспечению АСУП, позволяющая осуществлять проектирование систем по методу агрегатирования ее с оригинальной проектной документацией, отражающей специфику объекта.

Все ТПР подразделены на классы. Класс «Техника», например, определяет состав, размещение и порядок использования технических средств, предназначенных для решения задач АСУП.

Класс «Задача» охватывает комплекс алгоритмов и программ, реализующих определенные функции управления. В зависимости от реализуемых функций ТПР данного класса подразделены на подклассы (функциональные подсистемы). Подсистемы составлены из комплексов задач (по модульному принципу). При этом предусматривается разработка модулей алгоритмов (экономическое обеспечение) и модулей программ (программное обеспечение).

Каждый автономный программный модуль характеризуется завершенностью и постоянством алгоритма. Для работы модуля требуется не только обрабатываемая, но и управляющая информация, применяемая для настраивания модуля на конкретную работу (конкретный процесс обработки данных). Эта информация задается при помощи параметров, которые в конкретных условиях могут изменяться в определенных пределах.

Программные модули могут охватывать выделенные типовые процедуры или стандартные блоки обработки информации: перепись данных с одного машинного носителя на другой, например с перфокарт на магнитную ленту, и т. п.; дополнение содержания одного файла данными других и результатами обработки данных того же самого файла; формирование нового файла и запись на магнитную ленту или другой машинный носитель; сортировка файла, находящегося в памяти ЭВМ; формирование и печатание таблиц; объединение нескольких файлов для дальнейшей совместной обработки с получением промежуточных файлов, разбивка файла на несколько частей (по определенному признаку) и др.

Управляющая программа используется для ввода управляющих операторов (параметров) и формирования в оперативной памяти таблиц рабочих модулей. Описание структуры входных и результатных данных при помощи управляющих операторов позволяет автоматически настраивать программы решения задачи по одной из возможных структур этих данных. Так же производится автоматический вызов рабочих модулей. Массив управляющих операторов организуется по общесистемным принципам программного обеспечения ЭВМ.

Следует отметить, что серьезной проблемой в использовании программных модулей является информационная связь между ни-

ми (создание внутренней информационной базы) и их связь с исходными и результатными документами (создание внешней информационной базы).

Информационная связь между подсистемами реализуется через файлы промежуточной и нормативно-справочной информации. Она организуется при помощи системы алгоритмов формирования и ведения файлов нормативно-справочных данных. Для этого используются стандартные методы подготовки и записи на машинные носители нормативно-справочной информации, унифицированные структуры массивов и пр.

Подсистемный метод автоматизированного проектирования представляет более высокую степень интеграции типовых элементов систем. Их декомпозиция осуществляется на уровне функциональных подсистем, каждая из которых представляется в виде пакета прикладных программ. Пакеты обычно строятся как средства реализации некоторых экономико-математических методов, а не как объединение задач, решаемых в рамках определенных подразделений предприятия. Это позволяет разрабатывать пакеты программ, пригодные для более широкого их использования. Однако такой подход имеет и некоторые недостатки, так как модели, заложенные в пакетах, оказываются неполными или недостаточно гибкими, что требует дополнительных (оригинальных) разработок.

Пакеты прикладных программ по сравнению с другими средствами программного обеспечения отличаются своей адаптивностью, т. е. возможностью их настройки на конкретные особенности объекта, для которого они применяются. При помощи средств ППП можно генерировать свои программы согласно заданным характеристикам.

Программы генерации позволяют разрабатывать специальные структуры пакетов на уровне агрегатной компоновки или на уровне настраиваемых модулей. В первом случае сгенерированный пакет имеет модульную структуру и состоит из ряда самостоятельных процедур, после каждой из которых возможен переход к одной или нескольким последующим процедурам. Каждая процедура (программный модуль) реализует один способ или метод вычисления. При использовании настраиваемых модулей можно по необходимости менять их параметры.

Важным свойством пакета является его автономность, т. е. законченность формы и возможность применения как для разработки всей системы, так и для перевода на машинное решение отдельных задач управления или их комплексов. Пользователь в составе пакета получает тексты программ на машинном носителе, а также необходимую техническую и эксплуатационную документацию. Поэтому пакет прикладных программ можно определить как средство генерации составляющих АСУП, т. е. автоматизированного (машинного) формирования системы, подсистемы или их элементов из совокупности типовых модулей, причем каждый пакет состоит из набора алгоритмов, программ, реализующих стандартные процедуры, блоков генерации и совокупности инструкций

и методических указаний, определяющих процесс задания конфигурации генерируемых элементов системы и порядок их формирования.

Основная особенность *объектного* метода автоматизации проектирования — создание всей системы для определенной группы предприятий. При этом она, как правило, имеет блочную структуру. Блоки могут обеспечить автоматическую настраиваемость на варианты расчетов определенных показателей, формируемые документы, форматы данных, варианты средств вычислительной техники.

Модельное проектирование применяется при наличии общей информационной модели системы, аппарата ее совершенствования и развития, когда можно использовать некоторые общие параметры настройки и др. Следует отметить, что такие методы лишь начинают развиваться и находятся на стадии научных исследований и экспериментов.

Для автоматизации проектирования большой интерес представляет группа пакетов, предназначенных для выполнения общих работ по созданию систем: обработки материалов обследования объекта с целью построения рациональной структуры информационного фонда, его частей, выбора технических средств, вариантов технологических процессов обработки данных, последовательности выполнения процедур при решении задач. Теоретической основой при разработке таких пакетов, как правило, служат модели межотраслевого баланса, методы теории графов и матриц.

Автоматизации работ по проектированию систем способствуют и включенные в операционную систему ЭВМ третьего поколения средства автоматического построения блок-схем алгоритмов транслируемых программ, записанных на языке любого уровня. Это ускоряет обнаружение ошибок (в особенности логического характера) на стадии отладки отдельных программных модулей. К тому же с использованием автоматического построения блок-схем алгоритмов корректных программ ускоряется и сам процесс оформления программной документации.

16.2. Типизация и автоматизация проектирования информационного фонда и базы данных

Создание единых информационных фондов и баз данных — наиболее ответственный и сложный этап проектирования АСОД и АСУ. Независимо от специфических условий и требований к конкретным информационным фондам можно определить общую последовательность их разработки, в которой можно выделить следующие этапы: выявление содержания требуемой выходной информации; выявление источников и содержания возможной входной информации; определение методов и формул расчета; построение экономико-математических моделей; построение схемы процессов выполнения необходимых расчетов; моделирование и исследование возможных вариантов с целью выбора из них наиболее рациональ-

ных (оптимальных) вариантов структуры информационного фонда, базы (баз) данных, технологических процессов и алгоритмов формирования соответствующих файлов; выбор необходимых средств технического и программного обеспечения.

Как показывает практический опыт, информация и информационные процессы обычно легко поддаются формализации, т. е. их представлению в виде таблиц, схем, математических моделей, построение которых требует типизации и способствует автоматизации проектирования.

Методы типизации и автоматизации проектирования информационных фондов весьма специфичны и отличаются от аналогичных методов, применяемых в других случаях (например, при проектировании ЭВМ и других технических средств), тем, что здесь типизации и автоматизации подвергаются как сам объект (его организация), так и процесс его проектирования. Для автоматизированного построения информационного фонда могут быть использованы существующие типовые проектные решения. Автоматизация же процесса проектирования сводится к выбору наилучших вариантов проектных решений и разработке необходимой документации. Опыт показывает, что с использованием только готовых компонентов реальную систему практически невозможно построить, не говоря уже о том, что и они должны привязываться к конкретным условиям. Таким образом, возникает проблема аналитического моделирования информационного обеспечения, т. е. построения модели единого информационного фонда и базы данных на детализированном (элементарном) уровне. Создание такой модели для имитации функционирования информационной базы современных больших систем представляется почти нереальной задачей. Для этого нет совершенных программных и технических средств, позволяющих осуществить такое моделирование. Поэтому практически для автоматизации проектирования информационных фондов и баз данных можно использовать лишь частные модели, т. е. модели определенных частей системы. Рассмотрим некоторые возможные варианты типизации и автоматизации структурного проектирования информационного фонда.

Как известно, составные единицы информации, используемые для решения задач АСУ, функционально связаны друг с другом. Такие связи выражаются определенными формулами расчетов (математических моделей). Поэтому можно построить систему проектирования структуры информационного фонда на основании формул расчетов, записанных на определенном проблемно-ориентированном языке. При реализации такой системы на первом этапе предусматривается выполнение лексического, синтаксического и семантического анализа формул, а затем разрабатывается рациональная структура информационного фонда. Если критерием рациональности принять частоту встречаемости показателей в формулах, то рациональной будет иерархическая структура, в которой чаще используемые показатели займут более высокие уровни иерархии. Таким образом построенная структура типа дерева по-

звolyет минимизировать среднее время выбора показателей из фонда для множества задач АСУ. При этом также максимально ликвидируется дублирование показателей. Математическим аппаратом работы такой системы проектирования служит теория графов.

Известна методика автоматизации построения структур баз данных на основе запросов, предъявляемых к системе. Ею можно пользоваться лишь при небольшом количестве пользователей, так как построение структуры по многим фиксированным запросам большого количества пользователей может оказаться нерациональным. В последнем случае при конструировании, например, сетевой базы данных лучше учитывать лишь множество ее элементов и функциональные зависимости между ними, т. е. схему строить безотносительно к запросам, основываясь на требованиях непротиворечивости и минимальной избыточности данных. При этом непротиворечивость данных здесь заключается в согласованности связей между ними. Один из способов преодоления противоречий в данных и сокращения их избыточности — это приведение реляционной модели базы данных к третьей нормальной форме (см. 10.3). Реляционная модель базы данных, как известно, представляет собой преобразованную структурную модель, являющуюся отображением базы данных в виде совокупности меняющихся во времени нормализованных отношений ее элементов. Во избежание несогласованности таких отношений строится граф межреляционных зависимостей, представляющий собой схему допустимых связей между элементами базы данных. С использованием этого графа отдельные элементы базы данных объединяются в наборы (записи), в которых поддерживается определенное состояние связей между данными, а также определяются пути доступа к ним.

Согласно рассмотренной методике построения сетевой базы данных проектирование ее сводится к следующим основным этапам:

- 1) определению множества элементов;
- 2) определению структуры функциональных зависимостей;
- 3) построению реляционной модели базы данных в третьей нормальной форме;
- 4) построению графа межреляционных зависимостей;
- 5) определению путей доступа и описанию структуры базы данных на языке описания данных в соответствии с графом межреляционных зависимостей.

Для файловой организации информационного фонда немаловажное значение имеет вопрос структурной интеграции и дифференциации данных. Здесь под интеграцией понимается процесс объединения элементов данных (признаков и оснований показателей) в одну запись файла, а под дифференциацией — расчленение составных записей файла на определенные части.

Данные могут вводиться в базу данных в виде входных файлов, структура которых определяется форматом входных документов. Каждая входная запись одного файла, как правило, характеризуется уникальным значением ключевого признака. Записи же двух

или нескольких различных первичных входных файлов часто могут отличаться лишь набором реквизитов-оснований, и поэтому их призначные части могут быть объединены. Именно интеграция таких данных позволяет исключить дублирование информации, уменьшить объемы файлов, включаемых в информационный фонд, формировать основные и рабочие файлы с наименьшими затратами машинного времени. Поэтому эффективность интеграции данных в наибольшей степени проявляется в результате рационализации внутренней структуры файлов. Вместе с тем интеграция данных может оказывать отрицательное влияние на эффективность непосредственного решения самих задач системы обработки данных или АСУ, так как для выполнения отдельных прикладных программ приходится использовать интегрированные файлы с относительной избыточностью информации. Кроме того, в некоторых случаях объединенные или входные файлы даже по своему объему не отвечают требованиям, вытекающим из основного критерия рациональности базы данных — минимизации объема требуемой памяти ЭВМ. Поэтому наряду с вопросами интеграции данных при создании систем обработки информации следует рассмотреть и решить проблему целесообразности дифференциации данных, причем решить ее, как уже отмечалось, с точки зрения не только рациональности структуры файлов, но и эффективности решения прикладных задач пользователей системы, т. е. использования информационного фонда. Рассмотрим один из методов выбора интегрированной или дифференцированной структуры файлов данных.

В общем случае структуру записи интегрированного входного файла можно представить в виде последовательности:

$$(A, B, C, \dots, R, \dots, Z,)$$

где $A, B, C, \dots$ — реквизиты-признаки;

R — ключевой признак, имеющий в каждой записи уникальное значение;

Z — реквизиты-основания (в частном случае один реквизит-основание).

Такой входной файл можно характеризовать следующими параметрами:

n — количество реквизитов в каждой записи (включая уникальный ключ и основание);

m_i — количество различных значений i -го реквизита ($i = 1 \div n$, т. е. $m_1 = m_A, m_2 = m_B, \dots, m_n = m_Z$, причем $m_R = m_Z$;

s_i — разрядность i -го реквизита записи.

Используя принятые обозначения, можно записать формулу определения объема входного файла данных:

$$Q_{\text{вх}} = m_R \sum_{i=1}^n s_i. \quad (16.1)$$

При дифференциации такого файла прежде всего организуется один основной файл, структура записи которого имеет вид (R, Z) .

Кроме того, из входного файла формируются дополнительные файлы, в каждый из которых включаются все значения ключевого признака и значения одного из остальных реквизитов-признаков, т. е. записи дополнительных файлов имеют следующую структуру: (A, R) , (B, R) , (C, R) и т. д. Отметим, что повторяющиеся здесь значения признаков $A, B, C, \dots$ выносятся в подзаголовки групп записей с различными значениями ключевого признака R , т. е. одинаковые значения этих признаков в каждой записи не повторяются. Как видно из принятого принципа дифференциации входного файла, объемы основного ($Q_{\text{осн}}$) и всех дополнительных файлов ($Q_{\text{доп}}$) могут быть определены по формулам:

$$Q_{\text{осн}} = m_R (s_R + s_Z), \quad (16.2)$$

$$Q_{\text{доп}} = (n - 2) m_R s_R + \sum_{i=1}^{n-2} m_i s_i. \quad (16.3)$$

Исходя из критерия экономии общего требуемого объема памяти ЭВМ, можно сформулировать условие дифференциации входного файла:

$$Q_{\text{вх}} > Q_{\text{осн}} + Q_{\text{доп}}. \quad (16.4)$$

Используя формулы (16.1), (16.2) и (16.3), это же условие можно выразить через параметры входного файла:

$$m_R \sum_{i=1}^n s_i > (n - 2) m_R s_R + \sum_{i=1}^n m_i s_i. \quad (16.5)$$

В табл. 16.1 приведены исходные и расчетные данные для выбора интегрированной или дифференцированной структуры двух конкретных файлов (при $n = 5$).

Таблица 16.1

Расчет объемов интегрированных и дифференцированных файлов данных

	Параметры	Реквизиты					$Q_{\text{вх}}$	$Q_{\text{осн}} + Q_{\text{доп}}$
		A	B	C	R	Z		
A	s_i	3	4	2	4	3	480	712
	m_i	4	6	3	30	30		
B	s_i	3	4	4	3	12	7800	7338
	m_i	6	20	10	300	300		

Как видно из приведенных данных, дифференциация первого файла (часть А табл. 16.1) является нецелесообразной, а дифференциация второго файла (часть Б табл. 16.1) дает экономию памяти почти на 6%. Все же выбор того или другого варианта ор-

ганизации файлов по рассмотренному нами общему условию дифференциации не может быть окончательным, особенно в случаях, когда это условие не удовлетворяется. Здесь дополнительно следует учитывать то обстоятельство, что при решении отдельных задач вместе с основным файлом используются не все дополнительные файлы, а только их определенная часть. Поэтому экономия объема памяти, необходимой для информационного обеспечения конкретных прикладных программ, в условиях дифференциации данных может быть достигнута даже при нарушении общего условия дифференциации (16.5). В связи с этим для окончательного выбора внутренней структуры файлов необходимо проанализировать использование данных в разрезе отдельных прикладных программ пользователей системы. Принятие решения по выбору структуры файлов, а также сам процесс дифференциации можно автоматизировать, используя специальные программы ЭВМ, разработка которых не представляет особых трудностей.

16.3. Автоматизация построения специального программного обеспечения для решения задач и их комплексов

Одним из основных направлений автоматизации проектирования систем машинной обработки данных является создание средств автоматизированного построения программного обеспечения, т. е. программирования задач управления в широком смысле этого слова. Как известно, в современных системах часто используются готовые пакеты прикладных программ, но это не освобождает разработчиков систем от создания специального программного обеспечения, подготовки оригинальных рабочих программ.

Подготовка рабочих программ ЭВМ для решения конкретных задач АСУ с использованием современных системных средств программирования включает в себя три этапа:

- составление программ (исходных модулей) на языке программирования;

- трансляцию исходных модулей с получением переместимых (объектных) программ;

- редактирование переместимых модулей с целью получения абсолютных фаз программ, готовых к выполнению.

Два последних этапа выполняются автоматически по системным программам транслятора и редактора связей. Однако редактирование программ заключается не только в привязке отдельных модулей к конкретному месту оперативной памяти ЭВМ, но и в сборке сегментированных рабочих программ из отдельных модулей. В связи с этим при составлении задания на включение отдельных фаз в рабочую программу задачи возникают определенные трудности. Прежде всего они сводятся к тому, что программист должен детально описать структуру собираемой программы и порядок включения в нее редактируемых модулей, т. е. необходимо подготовить подробную инструкцию выполнения процесса редактирования. На данном этапе подготовки программ можно автома-

тизировать процесс сборки, в особенности в условиях использования единых баз данных.

Принципы структурного программирования позволяют разрабатывать микромодули, каждый из которых предназначен для расчета одного типового выходного показателя. Создание банка таких микромодулей, реализующих типовые алгоритмы расчетов, не представляет особых затруднений. Таким образом, каждый модуль может быть идентифицирован списком входных показателей и одним выходным показателем (практически для этого всегда можно использовать дескрипторные условные обозначения показателей согласно принятой системе их идентификации, например по генерируемому списку показателей, рассмотренному в 13.2). Отсюда следует, что идентифицированные программные микромодули могут объединяться в рабочую программу по принципу соответствия их входа и выхода, т. е. процесс редактирования таких микромодулей может происходить автоматически по специальной программе авторедактирования, анализирующей информационные связи задачи и на этой основе выбирающей для включения в конкретную рабочую программу те микромодули, которые эти связи реализуют.

Очевидно, что по одному и тому же микромодулю может быть определен целый ряд выходных показателей, если алгоритмы их расчета полностью совпадают. Таким образом, банк микромодулей по своему объему вовсе не соответствует количеству выходных показателей функциональных задач АСУ, а каждый микромодуль практически описывается набором различных входов и выходов. При этом такой набор с развитием системы может постоянно расширяться.

Процесс авторедактирования по рассматриваемой методике в зависимости от постановки задачи может быть организован по двум направлениям. Он всегда начинается с поиска и выбора того микромодуля, который определяет выходной показатель задачи (в общем случае — нескольких микромодулей, если в задаче предусмотрено получение нескольких выходных показателей). Далее может быть организовано наращивание собираемой рабочей программы другими микромодулями, если задача сформулирована явно, т. е. с указанием входного потока показателей. Подобное направление авторедактирования реализуется без вмешательства пользователя (постановщика задач). При этом каждый шаг наращивания можно представить как процедуру поиска в банке такого микромодуля, выходной показатель которого является входным для уже выбранных микромодулей, и включения его в рабочую программу. Процесс наращивания продолжается до тех пор, пока программа не будет полностью обеспечена входным потоком показателей согласно постановке задачи.

Второе направление реализации процесса авторедактирования связано с неявной постановкой задачи, когда входной поток показателей не определяется, а задаются только выходные данные. При этом, аналогично первому варианту, всегда выполняется первый шаг авторедактирования (он может быть и единственным).

Далее следует предусмотреть возможность уточнения постановки задачи с использованием диалогового режима. Принцип такого уточнения заключается в том, что программа авторедактирования выдает запрос постановщику с указанием требуемого потока входных показателей. Если он удовлетворяет условиям задачи, то после позитивного ответа постановщика процесс авторедактирования заканчивается. В противном случае выполняется следующий шаг наращивания рабочей программы. Следовательно, постановщик получает запросы после каждого шага наращивания и принимает соответствующее решение о необходимости продолжения или прекращения процесса авторедактирования.

Следует отметить, что процесс наращивания рабочей программы может происходить как в прямом, так и в обратном направлении. Обратное наращивание — это исключение микромодулей из собираемой программы. Это может случиться при неправильной постановке задачи, когда пользователь указывает неполный или ошибочный поток входных показателей. Для выполнения корректирующих шагов авторедактирования следует предусмотреть средства запоминания истории сборки рабочей программы, с тем чтобы обратное наращивание при необходимости происходило по той же траектории сборки. Кроме того, в целях ускорения процесса авторедактирования все шаги наращивания вначале следует выполнить фиктивно, т. е. проложить только саму траекторию сборки без фактической сборки, а затем по ней (без корректировки или после нее) реализовать процесс авторедактирования. Очевидно, что обратное наращивание чаще всего встречается при неявной постановке задач.

Таблица 16.2

Авторедактирование при явной постановке задачи

Шаг	1	2	3	4	5	6	7	8
Искомый выход	12	8	10	11	7	8	8	9
Включаемый модуль	Е	Б	Г	Д	А	(Б)	(Б)	В
Передаваемые входы	8 10 11	10 4 11 5 6	11 7 8	7 8 8 9	8 1 8 2 9 3	8 4 9 5 6	9 4 5 6	3 6
Неудовлетворенный входной поток	1 4 2 5 3 6	1 2 3	1 2 3	1 2 3	—	—	—	—
Наращиваемая программа	Е	Е Б	Е Б Г	Е Б Г Д	Е А Б Г Д	Е А Б (Б) Г Д	Е А Б (Б) Г (Б) Д	Е АБ Б (Б) Г (Б) Д В

Для более наглядного представления об авторедактировании программных модулей рассмотрим условный пример. Пусть в банке имеется определенный набор программных микромодулей (А,Б,В,Г,Д,Е,...), характеризующихся такими входами и выходами (показатели здесь обозначены порядковыми номерами): (1,2,3)→А→(7), (4,5,6)→Б→(8), (3,6)→В→(9), (7,8)→Г→(10), (8,9)→Д→(11), (8,10,11)→Е→(12) и т. д. Необходимо получить рабочую программу, определяющую показатель (12), когда во входной поток включаются показатели (1,2,3,4,5,6). Шаги процесса наращивания такой программы характеризует табл. 16.2. В ней показаны все 8 шагов редактирования, хотя 6-й и 7-й шаги, выделенные утолщенной линией, и не вызывают повторного включения в программу микромодуля В. Таким образом, после 8 шагов наращивания (два из них фиктивные) может быть выполнена рабочая программа, в которую входят (в порядке, обратном наращиванию) модули В, Б, А, Д, Г, Е. Именно в этой последовательности они должны быть выполнены в процессе решения поставленной задачи.

Если неявно формулируется задача получения показателя (11) без указания входного потока показателей, то на первом шаге авторедактирования с использованием того же набора микромодулей будет выбран модуль Д со входом (8,9); если это не удовлетворяет пользователя, то могут быть выполнены и последующие шаги наращивания согласно табл. 16.3 (шаг обратного наращивания помечен знаком минуса).

Как видно из представленных данных, рабочая программа в данном случае может быть собрана из одного, двух (в двух вариантах) или трех микромодулей. Постановщик задачи может остановить процесс авторедактирования после любого шага наращивания программы.

Используя рассмотренную методику авторедактирования программных модулей, можно автоматизировать один из самых сложных этапов подготовки рабочих программ. Как уже отмечалось, наиболее эффективно ее можно использовать в условиях единой базы данных, так как в этом случае авторедактирование может быть непосредственно связано с выполнением получаемой рабочей программы, когда переход от сборки программы к решению самой задачи может быть осуществлен без вмешательства пользователя. Это обеспечивается тем, что и база данных, и система авторедактирования пользуются единым идентификационным фондом. Со-

Таблица 16.3

Авторедактирование при неявной постановке задачи

Шаг	1	2	3	—4
Искомый выход	11	8	9	8
Включаемый микро-модуль	Д	Б	В	—Б
Передаваемые входы (требуемый поток)	8 9	9 4 5 5 6 6	4 3 (6)	3 8 6
Наращиваемая программа	Д	Д Б	Д Б В	Д В

единение постановки задач, редактирования и выполнения программ ускоряет процесс решения задач и служит важной предпосылкой реализации режима реального масштаба времени по решению высокоприоритетных оперативных задач управления.

Для практической реализации рассмотренных принципов авторедактирования первостепенное значение имеет упомянутый идентификационный фонд, средствами которого описываются входные и выходные показатели программных модулей и задач, характеризуются и устанавливаются информационные связи между отдельными модулями, а также организуется обращение к базе данных. Такой фонд может быть создан на базе специальных тезаурусов. Отдельные тезаурусы при этом, как правило, составляются для каждой группы реквизитов, характеризующих определенные однородные свойства объектов управления. В них включаются уникальные дескрипторы всех реквизитов, которые используются для построения идентификаторов показателей. Идентификационный фонд иногда целесообразно дополнить таблицами (перечнем) всех идентификаторов показателей с их порядковыми номерами, которыми более эффективно можно оперировать в процессе внутримашинной обработки данных. Не вникая в подробности организации такого фонда, отметим лишь то, что с его помощью можно организовать подготовку заданий по авторедактированию и выполнению программ решения конкретных задач пользователя на естественном языке, без применения дескрипторных условных обозначений входных и выходных показателей. Это значительно упрощает постановку задач управления, а выполнять ее может любой пользователь, не имеющий специальной подготовки.

ЛИТЕРАТУРА

1. Автоматизация проектирования систем управления/Под общ. ред. В. А. Трапезникова. М., Статистика, 1979. 205 с.
2. Алгоритмы и организация решения экономических задач. М., Статистика, вып. 1—12.
3. Алферова З. В. Теория алгоритмов. М., Статистика, 1973. 163 с.
4. Атавмян И. О. и др. Надежность автоматизированных систем управления. М., Высшая школа, 1979. 287 с.
5. Ахо А. и др. Построение и анализ вычислительных алгоритмов. М., Мир, 1979. 536 с.
6. Белоногов Г. Г., Новоселов А. П. Автоматизация процессов накопления, поиска и обобщения информации. М., Наука, 1979. 256 с.
7. Блохин Ю. И. Классификация и кодирование технико-экономической информации. М., Экономика, 1976. 191 с.
8. Ванагс Э. Я. Стандартизация технологических процессов машинной обработки данных. М., Статистика, 1974. 176 с.
9. Ванагс Э. Я., Ванагс И. Я. Автоматизированная система обработки данных административного района. М., Статистика, 1980. 151 с.
10. Вольфенгаген В. Э. и др. Реляционные методы проектирования банков данных. Киев, Высшая школа, 1979. 192 с.
11. Выбор комплекса технических средств АСУП/В. И. Аверин, А. П. Иванов, И. А. Кручинин и др. М., Статистика, 1973. 206 с.
12. Гендель Е. Г., Левин Н. А. Оптимизация технологии обработки информации в АСУ. М., Статистика, 1977. 232 с.
13. Глушков В. М. и др. Обработка информационных массивов в автоматизированных системах управления. Киев, Наукова думка, 1970. 181 с.
14. Евдокимов В. В., Рейнер В. А. Машинный синтез АСУП. М., Статистика, 1980. 222 с.
15. Евсюков К. Н., Колин К. К. Основы проектирования информационно-вычислительных систем. М., Статистика, 1977. 216 с.
16. Зимин Л. Н., Совета Н. Н., Филиппов Б. В. Средства подготовки данных в АСУП. М., Статистика, 1976. 192 с.
17. Зингер И. С., Модин А. А., Коротяев М. Ф. Экономико-организационные основы создания систем обработки данных. М., Статистика, 1978. 157 с.
18. Зингер И. С., Куцык Б. С. Обеспечение достоверности данных в автоматизированных системах управления производством. М., Наука, 1974. 136 с.
19. Иванов А. П. Вычислительные параметры экономических задач. М., Статистика, 1976. 164 с.
20. Информационные системы общего назначения (Аналитический обзор систем управления базами данных). М., Статистика, 1975. 463 с.
21. Карасев О. И. и др. АСУП на базе пакетов прикладных программ. М., Статистика, 1978. 141 с.
22. Королев М. А., Клешко Г. Н., Мишенин А. И. Информационные системы и структуры данных. М., Статистика, 1977. 183 с.
23. Кляков Ю. И., Горьков Б. М. Банки данных для принятия решений. М., Сов. радио, 1980. 208 с.
24. Куценко С. П. Проектирование, организация и планирование районных вычислительных установок. Проблемы теории и практики. М., Статистика, 1974. 280 с.

25. Липаев В. В. Распределение ресурсов в вычислительных системах. М., Статистика, 1979. 241 с.
26. Мамиконов А. Г. Методы разработки автоматизированных систем управления. М., Энергия, 1973. 336 с.
27. Мамиконов А. Г. и др. Модели и методы проектирования информационного обеспечения АСУ. М., Статистика, 1978. 221 с.
28. Мартин Дж. Организация баз данных в вычислительных системах. Пер. с англ. М., Мир, 1980. 664 с.
29. Общепромышленные руководящие методические материалы по созданию автоматизированных систем управления предприятиями и производственными объединениями (АСУП). М., Статистика, 1977. 262 с.
30. Проектирование машинной обработки экономической информации/Г. А. Титоренко, Г. С. Федорова, А. В. Абанина и др. 3-е изд., перераб. и доп. М., Статистика, 1979. 287 с.
31. Модин А. А. и др. Справочник разработчика АСУ. 2-е изд., перераб. и доп. М., Экономика, 1978. 583 с.
32. Соколов Р. В. Экономико-информационное моделирование процессов преобразования информации в АСУП. Изд. Ленингр. ун-та, 1980. 152 с.
33. Тимофеев Б. А., Литвинов В. А. Технология обработки данных в АСУП. Киев, Техника, 1978. 176 с.
34. Экономическая информация (Методологические проблемы). М., Статистика, 1974. 226 с.
35. Методика определения экономической эффективности автоматизированных систем управления предприятиями и производственными объединениями (ГКНТ СССР. Госплан СССР. АН СССР). М., Статистика, 1979. 62 с.

ОГЛАВЛЕНИЕ

Предисловие	3
-----------------------	---

РАЗДЕЛ ПЕРВЫЙ

ОСНОВЫ ПРИМЕНЕНИЯ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ ДЛЯ ОБРАБОТКИ ДАННЫХ

Глава 1. Развитие применения вычислительной техники	5
1.1. Этапы развития применения вычислительной техники для механизации планово-учетных работ и автоматизации управления	5
1.2. Эволюция работ по проектированию	10
1.3. Основные направления совершенствования проектных работ	14
Глава 2. Экономические данные и особенности их обработки	15
2.1. Информационный аспект и информационное обеспечение управления	15
2.2. Техничко-экономическая информация и данные	21
2.3. Этапы движения и преобразования данных	24
Глава 3. Сущность работ по проектированию машинной обработки данных	27
3.1. Типовой состав проектировочных работ	27
3.2. Особенности проектирования машинной обработки данных при создании АСУ	33

РАЗДЕЛ ВТОРОЙ

ЭФФЕКТИВНОСТЬ МАШИННОЙ ОБРАБОТКИ ДАННЫХ

Глава 4. Общие принципы определения эффективности	36
4.1. Основы оценки мероприятий по применению вычислительной техники	36
4.2. Показатели экономической эффективности и принципы их определения	40
Глава 5. Оценка экономической информации	44
5.1. Потребности в информации и ее ценность	44
5.2. Точность и достоверность данных	51
5.3. Оценка и повышение достоверности данных на выходе АСОД	54
Глава 6. Оценка затрат на обработку данных	56
6.1. Принципы определения затрат на обработку данных	56
6.2. Определение трудовых затрат	58
6.3. Определение затрат в денежном выражении и пути их снижения	61

РАЗДЕЛ ТРЕТИЙ

ПРОЕКТИРОВАНИЕ СИСТЕМ ДАННЫХ

Глава 7. Системы классификации и кодирования	66
7.1. Классификация экономических данных	66
7.2. Кодирование номенклатур	69
7.3. Единые системы классификации и кодирования	74
7.4. Методика выполнения работ по классификации и кодированию данных	76
7.5. Автоматизация кодирования	79
Глава 8. Системы описания данных	83
8.1. Идентификация данных и языки их описания	83
8.2. Системная идентификация реквизитов и показателей	85
8.3. Описание данных в условиях применения единых информационных фондов и баз данных	89
Глава 9. Формы и макеты размещения данных	91
9.1. Носители данных, их классификация и основные характеристики	91
9.2. Принципы размещения данных	96
9.3. Методика проектирования форм первичных документов и табуляграмм	98
9.4. Методика и особенности проектирования макетов размещения данных на машинных носителях	103
Глава 10. Информационный фонд и база данных	111
10.1. Принципы организации и размещения данных в памяти системы	111
10.2. Формирование и ведение единого информационного фонда	115
10.3. Принципы организации единых баз данных	120

РАЗДЕЛ ЧЕТВЕРТЫЙ

ВЫБОР И ПРОЕКТИРОВАНИЕ ПРОЦЕССОВ ОБРАБОТКИ ДАННЫХ

Глава 11. Типовые процессы и операции машинной обработки данных	129
11.1. Общая характеристика типовых технологических процессов и операций	129
11.2. Выделение операций технологических процессов	136
11.3. Особенности системной обработки данных и программного обеспечения современных средств вычислительной техники	139
Глава 12. Формирование и ведение информационного фонда и базы данных	142
12.1. Особенности технологии машинных процедур формирования файлов данных	142
12.2. Оценка и выбор вариантов технологических процессов предварительной обработки данных	151
12.3. Принципы организации и технологические процессы внесения изменений	156
Глава 13. Формирование и выдача результатных данных	161
13.1. Технологические процессы обработки и формирования файлов результатных данных	161

13.2. Особенности организации обработки данных при решении отдельных задач и их комплексов	165
13.3. Основные характеристики различных режимов обращения к системе	169

РАЗДЕЛ ПЯТЫЙ

ПРОЕКТИРОВАНИЕ СИСТЕМ ОБРАБОТКИ ДАННЫХ

Глава 14. Планирование и организация работ по проектированию	175
14.1. Планирование работ по проектированию и внедрению проектов	175
14.2. Организация обследования объекта и внедрения проектов	177
14.3. Организация проектирования	179
Глава 15. Разработка и документирование проектов	180
15.1. Состав организационно-распорядительной документации	180
15.2. Обработка материалов обследования объекта и подготовка задания на проектирование	181
15.3. Разработка технического и рабочего проектов	186
Глава 16. Автоматизация проектирования	189
16.1. Основные направления типизации и стандартизации проектирования	189
16.2. Типизация и автоматизация проектирования информационного фонда и базы данных	193
16.3. Автоматизация построения специального программного обеспечения для решения задач и их комплексов	198
Литература	203

Симанаускас Леонас Юозапович,
Бразайтис Зенонас Пранович

ОСНОВЫ ПРОЕКТИРОВАНИЯ МАШИНОЙ
ОБРАБОТКИ ДАННЫХ

Рецензенты С. И. Волков

А. Д. Евдокимов

Редактор Р. И. Кушкина

Мл. редактор Т. М. Кудина

Техн. редактор Л. Г. Челышева

Корректоры Я. Б. Островский, Г. И. Терновская

Худож. редактор Э. А. Смирнов

Переплет художника Е. П. Евсеева

ИБ № 1110

Сдано в набор 20.08.81. Подписано в печать 31.12.81.
А 12458. Формат 60×90^{1/16}. Бум. тип. № 3.
Гарнитура «Литературная». Печать высокая.
П. л. 13,0. Усл. п. л. 13,0. Усл. кр.-отт. 12,31.
Уч.-изд. л. 14,6. Тираж 15000 экз. Заказ 2728. Цена 65 к.

Издательство «Финансы и статистика», Москва, ул. Чер-
нышевского, 7

Великолукская городская типография
управления издательств, полиграфии и книжной торговли
Псковского облисполкома, г. Великие Луки,
ул. Полиграфистов, 78/12

