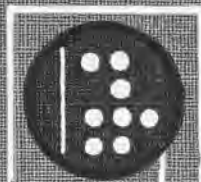


Е.И.Воробьев, А.И.Китов

введение в медицинскую кибернетику



Е. И. ВОРОБЬЕВ, А. И. КИТОВ. *Введение в медицинскую кибернетику*. М., «Медицина», 1977, 288 с.

В связи с широким внедрением методов кибернетики в практику работы больниц, поликлиник, санитарно-эпидемиологических станций, клинических лабораторий и других медицинских учреждений необходима популяризация основных принципов и методов медицинской кибернетики. Разделы медицинской кибернетики введены в курсы подготовки студентов во многих высших медицинских учебных заведениях.

В книге в доступной форме изложены основные сведения по кибернетике и ее применению в медицине и здравоохранении. Описаны задачи и методы кибернетики, строение и принципы работы электронных вычислительных машин (ЭВМ), а также методика программирования и решения задач на этих машинах. На примерах дано представление о математических методах выбора оптимальных решений в медицине и здравоохранении. Достаточно подробно освещены вопросы поиска и обработки на ЭВМ медицинской информации. Рассмотрены конкретные примеры применения ЭВМ в больницах (для управления процессом обслуживания больных, ведения историй болезней и др.), в органах управления здравоохранением. Обсуждены вопросы использования кибернетических методов в сфере управления медицинской наукой и в области медицинской научной информации.

Авторы монографии — доктор медицинских наук Е. И. Воробьев, кардиолог, специалист в области организации здравоохранения и радиационной медицины, и доктор технических наук профессор А. И. Китов — специалист в области вычислительной техники и автоматизированных систем управления.

Книга предназначена медицинским работникам, интересующимся применением ЭВМ и кибернетики в медицине и здравоохранении.

В книге 16 схем, 14 табл., библиография: 28 названий, глоссарий и список сокращений.

В $\frac{50103-249}{039(01)-77}$ 285-77

ПРЕДИСЛОВИЕ

XXV съезд КПСС уделил большое внимание вопросам развития медицины, здравоохранения. В отчетном докладе съезду Л. И. Брежнев подчеркнул, что среди социальных задач, стоящих перед Коммунистической партией, нет более важной, чем забота о здоровье советских людей. «Наши успехи здесь общеизвестны. Но надо видеть и стоящие в этой области проблемы. Они связаны с улучшением организации здравоохранения, расширением сети больниц и поликлиник, ростом производства медицинского оборудования и высокоэффективных лекарств. Они связаны также с дальнейшим развитием медицинской науки, развертыванием широкой борьбы против наиболее опасных заболеваний»¹.

В «Основных направлениях развития народного хозяйства СССР на 1976—1980 гг.» отмечается необходимость усиления исследований в области молекулярной биологии, физиолого-биохимических и иммунологических основ жизнедеятельности человеческого организма с целью ускорения решения важнейших медико-биологических проблем борьбы с сердечно-сосудистыми, онкологическими, эндокринными, вирусными и профессиональными заболеваниями, болезнями нервной системы.

В решении перечисленных проблем большое значение имеет ознакомление широких кругов медицинских работников с достижениями других наук, которые могут быть использованы в медицине и здравоохранении. Одной из таких наук является медицинская кибернетика. Медицинская кибернетика подобно экономической кибернетике, технической кибернетике, военной кибернетике и другим практическим отраслям кибернетики является прикладной наукой, в которой общие концепции кибернетики используются для изучения состояния

¹ Брежнев Л. И. Отчет Центрального Комитета КПСС и очередные задачи партии в области внутренней и внешней политики. М., Политиздат, 1976, с. 49.

здоровья людей и для выработки решений по управлению этим состоянием (по использованию сил и средств, организации здравоохранения).

В данной книге рассмотрены отдельные задачи и случаи применения кибернетики в медицине и здравоохранении, изложено существо самой кибернетики, математических методов обработки информации и выбора оптимальных решений, принципов работы ЭВМ и программирования для них.

С учетом того что важнейшим условием эффективного использования ЭВМ в медицине и здравоохранении является возможность взаимодействия между врачами и ЭВМ на естественном или частично формализованном медицинском языке, в книге достаточно подробно описаны вопросы обработки текстовой информации (глава IV). Изучение этой главы не потребует от читателей каких-либо дополнительных знаний, кроме обычных знаний грамматики русского языка; в то же время знакомство с материалом этой главы позволит врачу понять, как ЭВМ анализирует и опознает фразы естественного языка из вводимых в нее документов и запросов, как формирует ответы также в виде фраз естественного медицинского языка. Эти сведения будут полезны врачам при выполнении наиболее трудной и ответственной части работы, связанной с внедрением ЭВМ и методов кибернетики в медицину: при разработке форм входной и выходной документации для решения задач на ЭВМ и выборе способов представления информации для ее хранения в памяти ЭВМ.

В связи с наметившимся переходом в народном хозяйстве страны на ЭВМ третьего поколения в книге наряду с изложением общих принципов работы ЭВМ описаны особенности процессов решения задач на ЭВМ третьего поколения.

Значительное место в монографии уделено изложению существа математических методов выбора оптимальных решений и исследования сложных систем с помощью ЭВМ, причем акцент сделан на возможности и области применения этих методов и подходы к постановке задач.

В книге нашел отражение зарубежный и отечественный опыт применения ЭВМ в медицине и здравоохранении, а также в области научной информации и прогнозирования науки.

В связи с тем что книга рассчитана в основном на врачей, занятых практической и исследовательской работой в различных медицинских учреждениях, а также на работников организации и управления здравоохранением, авторы стремились при изложении указанных вопросов избегать технических подробностей и математических деталей, а излагать основную суть методов, их возможности и области применения.

Цель книги — помочь читателю оценить возможности применения кибернетики в собственных областях деятельности, наметить выбор подходящих способов представления информации и методов ее обработки и оценить хотя бы ориентировочно состав необходимых для этой цели технических средств.

Освоение любой новой области знания невозможно без изучения новых понятий, а следовательно, и новых терминов. Кибернетика и вычислительная техника особенно богаты специальной терминологией. Чтобы облегчить читателю ее усвоение, в тексте термины сопровождаются подробными пояснениями, а в конце книги дан глоссарий специальных терминов и список сокращений, составленный Н. С. Будько.

ОСНОВНЫЕ СВЕДЕНИЯ ИЗ КИБЕРНЕТИКИ**1. Возникновение кибернетики**

Кибернетика — наука о процессах управления и структуре управляющих систем в машинах, живых организмах и их объединениях. В. М. Глушков дает следующее определение кибернетики: «Кибернетика — наука об общих законах получения, хранения, передачи и преобразования информации в сложных управляющих системах». Основные положения сформулировал в 1948 г. американский ученый Н. Винер в книге «Кибернетика, или управление и связь в животном и машине» (М., 1958). Возникновение кибернетики обусловлено, с одной стороны, потребностями практики, выдвинувшей задачи создания сложных устройств автоматического управления, а с другой — развитием научных дисциплин, изучавших процессы управления в различных физических областях и подготовивших создание общей теории этих процессов. К числу таких наук относятся теория автоматического регулирования и следящих систем, теория электронных программно-управляемых вычислительных машин, статистическая теория передачи сообщений, теория игр и оптимальных решений и т. д., а также комплекс биологических наук, изучающих процессы управления в живой природе (рефлексология, генетика).

В отличие от указанных наук, занимающихся изучением конкретных процессов управления, общая (или теоретическая) кибернетика изучает то общее, что свойственно всем процессам управления независимо от их физической природы, и ставит своей задачей создание единой теории этих процессов.

Для развития современной науки характерны два процесса. Один из этих процессов называется дифференциацией (расщеплением) наук. Сущность его состоит в том, что научные дисциплины, ранее охватывавшие большие области явлений, распадаются на ряд самостоятельных наук, изучающих более узкие области явлений. Дифференциация вызвана постоянным разви-

тием наук, в ходе которого изыскиваются более глубокие методы исследования, обнаруживающие коренные различия между явлениями или процессами, бывшими при прежних грубых методах изучения объектами одной науки. Вторая причина дифференциации заключается в быстром увеличении числа изученных фактов, в росте объема накопленных научных знаний.

Поясним сказанное на следующем примере.

Во времена Ньютона и Эйлера математика существовала как единая наука. В настоящее время математика представляет собой целый комплекс наук, таких, как математический анализ, теория функций действительного переменного, теория функций комплексного переменного, высшая алгебра, векторный анализ, тензорный анализ, высшая геометрия, топология, номография, теория чисел и многие другие. Другой пример. Мы являемся свидетелями дифференциации научной дисциплины терапии на ряд специализированных направлений: кардиология, пульмонология, гастроэнтерология, эндокринология и др.

Однако наряду с дифференциацией происходит и другой процесс, противоположный по своему характеру. Этот процесс называется интеграцией — объединением наук в новые научные области. Сущность интеграции состоит в том, что в результате углубления научных методов и главным образом благодаря сопоставлениям результатов, даваемых различными науками, возникают новые точки зрения на явления, которые прежде считались не имеющими между собой ничего общего. Это позволяет установить общую точку зрения и применить общий метод к изучению таких явлений. Примером интеграции может служить возникновение общей теории колебаний, рассматривающей с единой точки зрения такие явления, как механические колебания, акустические колебания (звук), электромагнитные колебания (свет, радиоволны) и многие другие.

Примером интеграции наук является также формирование в настоящее время новой научной области, получившей название кибернетики.

Кибернетика изучает с единой общей точки зрения вопросы управления и связи в различных системах, которые до сих пор рассматривались в таких далеких друг от друга науках, как математика, логика, биология, психология, физиология, лингвистика, теория связи, тео-

рия автоматического регулирования, теория электронных цифровых программно-управляемых машин.

Слово «кибернетика» в переводе с греческого означает искусство кормчего. Название «кибернетика» для не существовавшей в то время науки было введено французским ученым А. М. Ампером, сделавшим попытку осуществить классификацию наук, и встречается в его книге, вышедшей в свет в 1834 г. Ампер считал, что должна существовать наука об управлении государствами, и именно для нее ввел это название.

Однако кибернетика как область знания возникла лишь 100 лет спустя, в 40-х годах текущего столетия. Толчком к ее развитию послужило появление и быстрое совершенствование новых вычислительных средств — электронных программно-управляемых машин.

При изучении работы электронных программно-управляемых машин была замечена некоторая аналогия, существующая между функциями таких машин и функциями мозга человека. Например, машина обладает свойством хранения чисел, что соответствует функции памяти человеческого мозга. Решение математических задач на электронной программно-управляемой машине осуществляется на основании тех же законов логики, которыми для этих целей пользуется человек.

Дальнейшее изучение этой аналогии показало, что сходство между функциями электронной программно-управляемой машины и функциями мозга действительно существует. В основе этого сходства лежит тот факт, что как электронная программно-управляемая машина, так и мозг представляют собой самоуправляемые системы, работа которых состоит в обмене информацией между отдельными их частями и в преобразовании этой информации.

По своей физической природе, характеру протекающих в них физических процессов, по своим возможностям и надежности в работе программно-управляемая машина и мозг, конечно, совершенно различны, но и машина и мозг являются системами, принимающими, перерабатывающими, хранящими и передающими информацию. Некоторые коллективы людей можно рассматривать тоже как системы, принимающие, перерабатывающие, хранящие и передающие информацию.

Для любых процессов управления характерно наличие организованной системы, состоящей из управляю-

щих и управляемых (исполнительных) органов; взаимодействие данной организованной системы с внешней средой, являющейся источником случайных или систематических помех; осуществление управления на основе приема и передачи информации; наличие цели и алгоритма управления. Понятие цели управления в кибернетике имеет весьма широкий смысл, включающий также естественную целесообразность управляющих систем живой природы, направленную на обеспечение устойчивости биологических видов и приспособление их к изменениям внешней среды. Изучение проблемы естественного причинного возникновения целесообразных управляющих систем живой природы является важной задачей кибернетики, которая должна позволить глубже выяснить соотношения причинности и целесообразности в живой природе.

Кибернетика изучает процессы управления с информационной стороны, отвлекаясь от энергетических или конструктивных характеристик реальных управляющих систем, поэтому кибернетику определяют и как науку о способах восприятия, передачи, хранения, переработки и использования информации в машинах, живых организмах и их объединениях.

В задачу кибернетики входит также систематическое сравнительное изучение структуры и различных физических принципов работы управляющих систем с точки зрения их способности воспринимать и перерабатывать информацию.

Кибернетика по своим методам, является наукой, широко использующей математический аппарат, а также сравнительный подход при изучении различных процессов управления. В качестве основных разделов кибернетики могут быть выделены теория информации, теория методов управления и теория управляющих систем.

2. Информация и сигнал¹

Как было сказано выше, кибернетика есть наука о связи, управлении и контроле в живых организмах и машинах. Но кибернетика изучает лишь одну сторону

¹ Авторы излагают понятие информации в соответствии с тем его толкованием, которое принято в большинстве работ, посвященных этому вопросу. Однако такое толкование не является единственно возможным (прим. ред.).

живого организма или машины, рассматривая их как системы, в которых вырабатывается информация, передается по каналам связи, воспринимается, преобразуется и сохраняется в памяти. Такие системы называются кибернетическими.

В кибернетике понятие информации является весьма широким. Сообщения, публикуемые в газетах, передаваемые по радио, получаемые из разговора с другими людьми, являются информацией. «Сведения» о внешней среде и ее изменениях, получаемые мозгом человека или животного от органов чувств, тоже являются информацией. «Указания», получаемые мускулатурой тела человека или животного от мозга о том, какие движения должны быть произведены, также представляют собой информацию. Исходные данные задачи, подлежащей решению, вводимые в электронную программно-управляемую машину, как и промежуточные или окончательные результаты решения задачи, — также не что иное, как информация. Информация — это приращение «знаний», новые, ранее неизвестные «сведения». Из сообщения, содержание которого уже известно получателю, последний не извлекает информации¹.

Органы или приборы кибернетической системы вырабатывают и передают информацию по назначению в виде сигналов. Среда, служащая для передачи сигналов, называется каналом связи. Например, при передаче телеграмм сигналами служат импульсы электрического тока, а каналом связи — металлический провод с необходимой аппаратурой. При передаче информации путем устной речи сигналами являются звуковые колебания, а каналом связи — воздух².

¹ Из сообщения, содержание которого уже известно получателю, можно извлечь информацию, хотя и не ту, которая выражена этим сообщением. Именно в этом плане следует понимать афоризм: «Ты мне сказал один раз, и я тебе поверил. Ты мне сказал второй раз, и я засомневался. Ты мне сказал третий раз, и я понял, что ты говоришь неправду». Более того, мы иногда можем получить информацию из того, о чем умолчал говорящий (прим. ред.).

² Такие широко употребляемые выражения, как, например, «по каналу связи передается информация», являются чистойшей методикой (иносказанием), такой же, как и выражение «школа вышла на субботник». По каналам связи передаются сигналы, и ничто другое. Только тогда, когда эти сигналы поступят в мозг, и если мозг в состоянии их правильно воспринять, только тогда эти сигналы будут восприняты как информация (прим. ред.).

Процесс передачи информации может быть весьма сложным. Например, при передаче по радио информация сначала передается в виде звуковых сигналов диктором, говорящим перед микрофоном. В микрофоне звуковые сигналы преобразуются в электрические. В передатчике радиостанции электрические сигналы превращаются в электромагнитные колебания — радиоволны. Затем радиоволны опять становятся электрическими сигналами, которые динамическим громкоговорителем или телефоном снова преобразуются в звуковые сигналы.

Наряду с передачей информации часто приходится наблюдать и хранение ее. Мозг человека обладает способностью запоминать информацию. Речь, произносимую оратором, можно записать на магнитофонную пленку. Орган или прибор, служащий для хранения информации, называют запоминающим устройством, или памятью. В памяти информация хранится тоже в виде сигналов.

Чаще всего сигналы, применяющиеся для «запоминания» информации, представляют собой более или менее устойчивые изменения среды, образующей запоминающее устройство. Так, при использовании в качестве запоминающего устройства ферромагнитной пленки сигналами являются группы намагниченных участков этой пленки. Запоминание информации мозгом человека происходит в результате возникновения устойчивых изменений в определенных его клетках или их группах.

Можно сказать, что каналы связи служат для передачи информации в пространстве, а запоминающие устройства для особого вида передачи — во времени: из прошлого в настоящее и из прошлого и настоящего в будущее. Нужно, однако, заметить, что передача информации по каналу сопровождается некоторой потерей времени (от момента посылки сигнала до момента его приема), а запись информации в память или считывание ее из памяти не могут быть произведены без некоторого переноса ее в пространстве (на расстояния, определяемые физическими размерами запоминающего устройства).

Иногда передача информации осуществляется путем транспортировки запоминающего устройства, хранящего эту информацию. Каналом связи при этом является совокупность транспортируемого запоминающего уст-

ройства и оборудования, с помощью которого происходит транспортировка. Примером такой передачи информации может служить почтовая связь.

Возможны также случаи хранения информации путем многократной передачи ее по некоторому каналу связи из одного конца его в другой и обратно (или циркуляции сигналов в замкнутом канале связи). Такой канал связи вместе с устройствами приема и обратной отправки («отражения») сигналов, несущих информацию, представляет собой динамическое запоминающее устройство.

Из сказанного видно, что информация может находиться в состоянии как передачи (по каналу связи), так и хранения (в памяти). Информация может также подвергаться преобразованиям, на чем мы остановимся в дальнейшем.

Сразу же подчеркнем, что передача и прием информации есть лишь один из видов связи между объектами, так как не всякая связь между объектами является обменом информацией (сигналами) и не между всякими объектами существует связь, состоящая в обмене информацией. Такая связь имеется лишь между частями одной и той же кибернетической системы. Обмен информацией между кибернетическими системами происходит только тогда, когда они являются частями более сложной кибернетической системы. Например, при разговоре двух человек по воздуху, как по каналу связи, происходит передача сигналов, которые в итоге мозгом слушающего воспринимаются в виде некоторой информации. Если те же звуки, т. е. те же сигналы, достигают ушей животного (или даже человека, которому неизвестен язык говорящих), то в этом случае передачи информации (по крайней мере той, которой обмениваются говорящие) не происходит. Таким образом, смысл сигналов зависит не только от этих сигналов, но и от того, кем или чем, т. е. какой кибернетической системой, они восприняты.

Мы знаем, что все явления и объекты мира существуют во взаимосвязи. Эта связь бесконечно многообразна. Кибернетика, изучающая лишь один из видов связи — обмен информацией и при том лишь между некоторыми объектами — частями кибернетических систем, не может претендовать и не претендует на то, чтобы быть наукой о всеобщей связи.

Раздел кибернетики, называющейся теорией информации, изучает информацию с количественной стороны.

Теория информации изучает способы преобразования, передачи и хранения информации. Информация передается с помощью сигналов — физических процессов, у которых определенные параметры находятся в однозначном соответствии с передаваемой информацией. Установление такого соответствия называется кодированием.

Хотя для передачи любых сигналов необходима затрата энергии, количество ее в общем случае не связано с количеством, а тем более со смыслом передаваемой информации. В этом заключается одна из принципиальных особенностей процессов управления: управление большими потоками энергии может осуществляться сигналами, требующими для передачи незначительного количества энергии.

Центральным понятием теории информации является мера количества информации, определяемая как изменение степени неопределенности в ожидании некоторого события, о котором говорится в сообщении до и после получения сообщения. Эта мера позволяет измерять количество информации в сообщениях подобно тому, как в физике измеряется количество энергии, при этом смысл и ценность передаваемой информации для получателя не учитываются. Помимо статистического подхода, в последние годы в теории информации наметился семантический подход, целью которого является выработка количественных методов оценки смысла и ценности сообщений. К этому направлению относится изучение общих вопросов природы научной информации, ее дискретного характера; соотношений, определяющих ценность информации, извлекаемой из эксперимента, а также вопросов автоматической выработки понятий, их классификации и опознавания образов.

Исключительный интерес с точки зрения кибернетики представляет изучение созданных природой естественных способов кодирования наследственной информации, обеспечивающих сохранение в ничтожных объемах наследственного вещества огромных количеств информации, содержащих уже в зародышевой клетке признаки взрослого организма.

Метод подсчета количества информации, содержащейся в некотором сообщении, наиболее полно был

впервые изложен американским ученым Клодом Шенноном, который занимался теоретическими вопросами передачи электрических сигналов по проводам. Очень скоро выяснилось, что его формула для подсчета количества информации может с успехом использоваться и в других областях науки и техники.

Метод подсчета количества информации проще всего уяснить из следующих рассуждений.

Рассмотрим систему событий A_1, A_2, \dots, A_n , обладающую тем свойством, что одно и только одно из событий этой системы должно произойти. Пусть p_1, p_2, \dots, p_n — соответственно вероятности¹ наступления событий A_1, A_2, \dots, A_n (очевидно, $p_1 + p_2 + \dots + p_n = 1$). Система событий A_1, A_2, \dots, A_n , рассматриваемых в совокупности с их вероятностями, называется конечной схемой. Такую конечную схему обозначают следующим образом:

$$\begin{array}{l} A_1, A_2, \dots, A_n \\ p_1, p_2, \dots, p_n \end{array} \quad (1.1)$$

Наступление одного из событий конечной схемы называется ее испытанием.

Всякая конечная схема содержит некоторую неопределенность. Известен перечень возможных событий, известна вероятность каждого события, но какое из событий в действительности произойдет — неизвестно.

Сообщение о результате испытания конечной схемы снимает эту неопределенность. Естественно считать, что количество информации, содержащейся в сообщении о результате испытания конечной схемы, равно величине неопределенности, содержащейся в этой конечной схеме. Как показал К. Шеннон, эта неопределенность, а значит, и количество информации в сообщении может быть вычислена из формулы, носящей его имя:

$$H = -(p_1 \log_a p_1 + p_2 \log_a p_2 + \dots + p_n \log_a p_n). \quad (1.2)$$

Здесь a — некоторое положительное число больше единицы. Ниже будет показано, что выбор числа a равноценен выбору единицы измерения.

¹ Вероятностью события называется отношение числа шансов, которые ему благоприятствуют, к общему числу равновозможных шансов (такое определение не претендует на строгость). Например, если в ящике находятся 7 черных и 3 белых шара, то вероятность того, что человек с завязанными глазами вынет из ящика черный (а не белый) шар, составит 0,7.

Кроме того, принято, что если $p=0$, то

$$p \log_a p = 0. \quad (1.3)$$

Знак минус в правой части формулы (1.2) взят потому, что все слагаемые $p_i \log_a p_i$ неположительны. Вероятность p , если не равна нулю, является положительным числом не больше 1, т. е. $\log_a p$ будет отрицательным числом или нулем. Следовательно, и произведение $p \log_a p$ будет отрицательным числом или нулем. Таким образом, этот знак делает правую часть формулы (1.2) положительной.

Рассмотрим внимательно формулу (1.2) и убедимся в целесообразности выбранного способа измерения количества информации.

Предположим, что одно из событий A_1, A_2, \dots, A_n является достоверным. Например, событие A_1 . Тогда его вероятность будет равна единице, а вероятность каждого из остальных событий будет равна нулю. Конечная схема принимает вид:

$$\begin{pmatrix} A_1 & A_2 & \dots & A_n \\ 1 & 0 & \dots & 0 \end{pmatrix} \quad (1.4)$$

Ясно, что такая конечная схема не содержит никакой неопределенности, и сообщение о результате ее испытания не будет содержать никакой информации.

Формула (1.2) дает именно этот результат:

$$H = -(1 \log_a 1 + 0 + \dots + 0) = 0.$$

Очевидно, наибольшую неопределенность содержит конечная схема, все события которой равновероятны. В случае такой конечной схемы до ее испытания труднее всего сделать сколько-нибудь обоснованное предположение о результатах испытания.

Исследуя на максимум величину H как функцию переменных p_1, p_2, \dots, p_n (эти выкладки здесь не приводятся), легко получить, что своего наибольшего значения H достигает при $p_1 = p_2 = \dots = p_n = \frac{1}{n}$, т. е. в случае равной вероятности событий, образующих конечную схему. При этом конечная схема принимает вид:

$$\begin{pmatrix} A_1 & A_2 & \dots & A_n \\ \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \end{pmatrix} \quad (1.5)$$

Сообщение о ее испытании содержит количество информации, равное:

$$H = - \left(\frac{1}{n} \log_a \frac{1}{n} + \frac{1}{n} \log_a \frac{1}{n} + \dots + \frac{1}{n} \log_a \frac{1}{n} \right) = \log_a n. \quad (1.6)$$

Обратим особое внимание на простейший частный случай конечной схемы (1.5), именно на случай, когда $n=2$. При этом схема состоит из двух событий: A_1 и A_2 . Применяя обозначение математической логики, событие A_2 можно записать символом \bar{A}_1 (это читается: не A_1). Конечная схема имеет вид:

$$\begin{pmatrix} A_1 & \bar{A}_1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad (1.7)$$

Сообщение о результате ее испытания содержит количество информации, равное числу:

$$H^* = \log_a 2.$$

Если мы положим $a=2$ (в последней и во всех предыдущих формулах), то будем иметь:

$$H^* = \log_2 2 = 1$$

Тем самым мы выбираем единицу измерения количества информации.

Таким образом, за единицу количества информации принято количество информации, содержащееся в сообщении об испытании конечной схемы, состоящей из двух равновероятных событий. Эта единица получила название б и т.

Выбранная мера количества информации обладает еще тем свойством, что при подсчете с ее помощью количества информации, содержащегося в сообщении об испытании двух независимых между собой конечных схем, получается число, равное сумме количеств информации, содержащихся в сообщениях об испытании каждой из конечных схем в отдельности, т. е., как говорят, мера количества информации обладает свойством аддитивности.

Ограничимся доказательством этого факта для простейшего случая, когда каждая из конечных схем состо-

ит из двух событий. Доказательство в общем случае совершенно аналогично приводимому.

Итак, рассмотрим две конечные схемы:

$$\begin{pmatrix} A_1 & A_2 \\ p_1 & p_2 \end{pmatrix} p_1 + p_2 = 1 \quad (1.8)$$

$$\begin{pmatrix} B_1 & B_2 \\ q_1 & q_2 \end{pmatrix} q_1 + q_2 = 1 \quad (1.9)$$

Сообщения о результатах испытания этих конечных схем содержат соответственно следующие количества информации:

$$H_1 = -(p_1 \log_2 p_1 + p_2 \log_2 p_2). \quad (1.10)$$

$$H_2 = -(q_1 \log_2 q_1 + q_2 \log_2 q_2). \quad (1.11)$$

Очевидно, испытание обеих конечных схем (1.8) и (1.9) эквивалентно испытанию новой конечной схемы:

$$\begin{pmatrix} A_1 \text{ и } B_1 & A_1 \text{ и } B_2 & A_2 \text{ и } B_1 & A_2 \text{ и } B_2 \\ p_1 q_1 & p_1 q_2 & p_2 q_1 & p_2 q_2 \end{pmatrix} \quad (1.12)$$

Каждое событие конечной схемы (1.12) представляет собой сочетание одного из событий, входящих в конечную схему (1.8), с одним из событий, входящих в конечную схему (1.9). Вероятность наступления такого комбинированного события, как известно, равна произведению вероятностей наступления событий, входящих в комбинацию.

Подсчитаем количество информации, содержащееся в сообщении об испытании конечной схемы (1.12):

$$H = -[p_1 q_1 \log_2(p_1 q_1) + p_1 q_2 \log_2(p_1 q_2) + p_2 q_1 \log_2(p_2 q_1) + p_2 q_2 \log_2(p_2 q_2)]. \quad (1.13)$$

Представляя в этом выражении каждый логарифм произведения в виде суммы логарифмов сомножителей и выполняя затем несложные алгебраические преобразования, получаем:

$$H = -[p_1 q_1 \log_2 p_1 + p_1 q_1 \log_2 q_1 + p_1 q_2 \log_2 p_1 + p_1 q_2 \log_2 q_2 + p_2 q_1 \log_2 p_2 + p_2 q_1 \log_2 q_1 + p_2 q_2 \log_2 p_2 + p_2 q_2 \log_2 q_2] = -(q_1 + q_2) (p_1 \log_2 p_1 + p_2 \log_2 p_2) + (p_1 + p_2) (q_1 \log_2 q_1 + q_2 \log_2 q_2).$$

Учитывая условия

$$p_1 + p_2 = 1,$$

$$q_1 + q_2 = 1,$$

а также формулы (1.10) и (1.11), получим:

$$H = H_1 + H_2, \quad (1.14)$$

что и требовалось доказать.

Все сказанное выше убеждает нас в том, что способ измерения количества информации выбран целесообразно.

Приведем два примера на вычисление количества информации.

Во многие игры в качестве составного элемента входит метание кубика, на гранях которого написаны цифры от единицы до шести.

Пусть конечная схема состоит из шести событий, каждое из которых представляет выпадение одного из указанных чисел при метании кубика. Если кубик изготовлен из однородного материала с большой степенью точности, то все события можно считать равновероятными. Конечная схема имеет вид:

$$\left(\begin{array}{cccccc} A_1 & A_2 & A_3 & A_4 & A_5 & A_6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{array} \right).$$

Сообщение о результате метания кубика содержит следующее количество информации:

$$H = \log_2 6 \approx 2,585 \text{ бита.}$$

Предположим теперь, что на одной из граней кубика написана цифра 1, а на пяти остальных — цифра 2. В этом случае конечная схема состоит из двух событий; A_1 (выпадение цифры 1) с вероятностью $1/6$ и A_2 (выпадение цифры 2) с вероятностью $5/6$:

$$\left(\begin{array}{cc} A_1 & A_2 \\ \frac{1}{6} & \frac{5}{6} \end{array} \right).$$

Сообщение о результатах метания кубика содержит меньше информации, чем в первом примере, а именно:

$$\begin{aligned} H &= - \left(\frac{1}{6} \log_2 \frac{1}{6} + \frac{5}{6} \log_2 \frac{5}{6} \right) = \\ &= \log_2 6 - \frac{5}{6} \log_2 5 \approx 0,65 \text{ бита.} \end{aligned}$$

Как уже было сказано, передача информации имеет место только между частями одной и той же кибернетической системы. Такие части соответствуют друг другу: одна передает информацию, предназначенную другой, а последняя воспринимает и «понимает» сигналы, несущие эту информацию. В связи с этим расчленение на части «сообщений» (сигналов), передаваемых внутри кибернетической системы, не всегда допустимо. Части сообщений могут оказаться «непонятными» для приемника информации и не могут содержать информации, тогда как их совокупность является носителем положительного количества информации. Сказанное позволяет объяснить следующий парадокс.

Предположим, что производится двукратное метание монеты, имеющей правильную форму, так что выпадения лицевой и оборотной стороны ее («орла» и «решки») равновероятно. Пусть при этом событий А состоит в том, что результаты обоих метаний одинаковы, а событие В в том, что они различны. Рассматриваемая конечная схема имеет вид:

$$\begin{pmatrix} A & B \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Предположим далее, что сообщение о результате испытания конечной схемы имеет вид: «орел, орел». Такое сообщение полностью снимает неопределенность конечной схемы и несет в себе, согласно ранее сказанному, один бит информации. Разделим это сообщение на две части «орел» и «орел». Очевидно, ни одна из этих частей, взятая в отдельности, не несет в себе никакой информации о результате испытания нашей конечной схемы. Получается (и в этом состоит парадокс), что совокупность двух сообщений содержит один бит информации, а каждое из этих сообщений содержит ноль битов, т. е. $0+0=1$.

Парадокс разрешается следующим образом. Указанные части сообщения для рассматриваемых условий задачи не являются сообщениями, хотя они внешне тождественны с сообщениями, имеющими смысл для некоторых других конечных схем. Сформулированной задаче они не соответствуют и потому действительно не имеют смысла, т. е. не несут информации, тогда как их совокупность является сообщением и несет информацию.

Теория информации применяется при изучении каналов связи кибернетических систем. Многие результаты, полученные методами теории информации, имеют непосредственное практическое применение. Коротко остановимся на некоторых из них.

Количество информации, передаваемой за единицу времени по некоторому каналу связи, называется скоростью передачи информации. Информация передается по каналу связи, как уже говорилось, путем передачи сигналов. Эти сигналы представляют собой некоторые процессы, протекающие в канале связи. Но канал связи не

может быть полностью изолирован от воздействия внешней среды. Посторонние явления, протекающие в каналах связи и искажающие сигналы, получили название шумов, или помех.

Итак, передача информации происходит в условиях наличия шумов, искажающих полезные сигналы. Иногда искажения сигналов достигают такой степени, что искажается информация, содержащаяся в сигналах. Шумы даже могут быть такими сильными, что сигналы становятся для приемника совершенно неразличимыми.

Надежностью передачи называется вероятность приема неискаженной информации. Весьма интересна доказанная в теории информации теорема, гласящая, что с увеличением скорости передачи ее надежность уменьшается и, наоборот, за счет уменьшения скорости передачи можно повысить ее надежность. Простейшим способом повышения надежности передачи информации является многократное повторение передачи одного и того же сообщения. Ясно, что при этом снижается общая скорость передачи информации.

Следует подчеркнуть, что для «понимания» передаваемой информации приемник должен заранее располагать определенными сведениями, «знать», из каких событий состоит конечная схема и как будет закодировано сообщение о результате ее испытания.

Сигналы, хранящиеся в «памяти», также подвержены действиям помех, которые могут исказить их, причем в ряде случаев может исказиться представленная ими информация. Например, записи, произведенные в записной книжке, могут стираться и становиться неразборчивыми.

В случае полного стирания отдельных слов смысл записи может измениться. Судьям и следователям, допрашивающим свидетелей, хорошо известно, как сильно может отличаться то, что запомнил очевидец, от того, что он видел в действительности. Значит, память человека тоже подвержена влиянию помех.

Теория информации показывает методы повышения количества информации, передаваемой в единицу времени по каналу связи. Суть этих методов сводится к учету вероятностей появления в передаваемых сообщениях различных символов, слов, предложений и т. п. При этом наиболее часто встречающиеся символы, слова и т. п. кодируются короткими кодами (имеющими небольшое число цифр или сигналов), а редко встречающиеся символы

кодируются более длинными кодами. Практическое значение имеет следующая теорема, доказанная в теории информации: количество передаваемой в единицу времени информации (скорость передачи информации) по каждому каналу связи имеет максимум, который не может быть превышен никакими способами кодирования информации. Максимум скорости передачи по каналу связи получил название пропускной способности канала связи, так как он представляет собой наибольшее количество информации, которое можно за единицу времени передать по каналу. Эту теорему приходится учитывать, например, при проектировании телефонных линий в больших городах.

При обучении человека разным наукам информация передается извне в его мозг по разным каналам. Во время слушания лекции таких каналов два: слуховой (по которому передается речь лектора) и зрительный (передающий информацию, выписываемую лектором на доске или содержащуюся в демонстрируемых им наглядных пособиях). При прохождении практики в лаборатории или на производстве каналами связи для передачи информации могут быть и другие пути (осязательные, обонятельные, вкусовые и т. д.). При правильном обучении пользуются одновременно несколькими из этих каналов связи.

Методы теории информации применяются при решении задач медицинской машинной диагностики для оценки роли (информативности) различных симптомов-признаков и их комбинаций.

3. Преобразование информации и алгоритмы

Преобразование информации осуществляется путем изменения несущих или хранящих ее сигналов.

Мы уже знаем, что существуют преобразования сигналов, не изменяющие содержания представленной ими информации (перекодировка информации). Но возможны преобразования сигналов, сопровождающиеся закономерным изменением содержания отвечающей им информации. Такие преобразования сигналов и информации называют содержательными. Примером содержательного преобразования информации является преобразование, выполняемое при решении математических или логических задач.

Предположим, что нам задана следующая информация.

Один гражданин купил 5 м сукна и 3 м сатина, заплатив за покупку 53 р. Другой гражданин за 4 м сукна и 12 м сатина заплатил 52 р. Обозначая буквой x цену 1 м сукна, а буквой y цену 1 м сатина, мы можем написать:

$$5x + 3y = 53$$

$$4x + 12y = 52$$

Эти два уравнения содержат ту же самую информацию, которая прежде была заложена в словесную формулировку задачи. Переход от слов к алгебраическим уравнениям является перекодировкой, а не преобразованием информации.

Умножим первое уравнение на 4 и вычтем из него второе уравнение. Получим: $16x = 160$.

Разделив обе части этого уравнения на 16, получим: $x = 10$.

Подставляя значение x в одно из исходных уравнений, находим: $y = 1$.

В результате последних преобразований, произведенных нами над уравнениями (над сигналами), мы получили: $x = 10$, $y = 1$. Эти сигналы обозначают информацию: 1 м сукна стоит 10 р., а 1 м сатина — 1 р.

Мы видим, что исходная информация оказалась преобразованной. Для преобразования информации был необходим запас сведений из алгебры, т. е. для того, чтобы выполнить содержательное преобразование информации, нужен некоторый предварительный ее запас. При этом нужны не только правила выполнения алгебраических действий над уравнениями и числами, но и сведения о том, какие действия и в какой последовательности надо произвести для получения искомого результата. Другими словами, необходима информация, определяющая преобразование информации. Таким образом, мы установили, что, кроме преобразуемой, существует еще преобразующая информация. Конечно, преобразующая информация не является каким-то особым видом информации. Деление информации на преобразуемую и преобразующую отражает лишь функции, выполняемые ими в заданных условиях. В других условиях, например, когда мы узнаем, что известный нам метод решения некоторой задачи должен быть определенным образом изменен, информация, которая первоначально именовалась преобразующей, станет в этих условиях преобразуемой.

Описанное выше решение задачи выполняется мозгом человека в соответствии с правилами алгебры и требованиями, содержащимися в преобразующей информации, которая ему заранее известна. Мышление человека,

являющееся при всей своей сложности также преобразованием информации, происходит в соответствии с законами логики и тоже с помощью преобразующей информации, запасы которой в мозге человека огромны. Существуют машины и устройства, позволяющие производить некоторые преобразования информации с получением тех же результатов, которые при переработке такой же информации получил бы человек. Это вовсе не означает, что в машинах производятся те же преобразования информации, которые выполняет мозг человека. Преобразующая информация может быть весьма разнообразной по своему виду. Иногда она может быть задана в виде таблицы соответствия между сигналами, несущими исходную информацию, и сигналами, отвечающими преобразованной информацией. При этом преобразование информации сводится к выбору результата из таблицы на основании поступающих исходных данных. В частном случае такая таблица может быть составлена в процессе эксперимента, например, таблица для получения информации о силе сопротивления воздуха движению самолета по информации о скорости движения самолета относительно воздуха и об атмосферном давлении. Необходимое преобразование информации может осуществляться также путем вычислений по специально подобранным формулам, заменяющим таблицы и т. п.

В управляющих системах информацию принято различать по назначению, по той роли, которую она выполняет в этих системах. Информацию, поступающую от чувствительного органа (устройства), называют осведомляющей информацией, а информацию, передаваемую управляющим органом (устройством) в управляемые органы (устройства), — управляющей информацией. При этом говорят, что в управляющем органе осведомляющая информация преобразуется в управляющую. Так, под влиянием внешних воздействий органы чувств человека вырабатывают осведомляющую информацию, которая в виде сигналов передается по центроостремительным нервам в центральную нервную систему. Например, если человек случайно прикоснулся рукой к горячему предмету, происходит описанный только что процесс. Мозг преобразует эту информацию в управляющую, которая передается по центробежным нервам мышцам руки. Под воздействием управляющей информации мышцы сокращаются и выводят руку из соприкосновения с горячим пред-

метом. На этом работа мозга не прекращается. Из мозга передается управляющая информация мышцам глаз. Глаза направляются на предмет, вызвавший ощущение боли. Из глаз в мозг поступает новая осведомляющая информация. Одновременно поступает информация и от других органов чувств. Эта информация преобразуется, в частности, и логически, причем используется и запасы информации, хранящиеся в памяти. Если окажется, что ожог вызван загоревшимся предметом, то мозг благодаря преобразованию информации поймет, что случился пожар, и пошлет в мышцы управляющую информацию, под воздействием которой они придут в действие, и будут приняты меры к тушению пожара.

Аналогичные процессы преобразования информации происходят и в автоматах, построенных человеком.

Управляющий орган представляет собой преобразователь информации. Он подвергает осведомляющую информацию содержательному преобразованию, а затем производит перекодировку полученных результатов таким образом, чтобы сигналы, несущие управляющую информацию, были способны воздействовать на управляемый орган.

Передача информации в виде сигналов всегда сопровождается некоторым расходом энергии. Количество энергии, необходимое для передачи информации, не зависит от содержания и количества информации, а прежде всего определяется чувствительностью элемента, принимающего сигналы (информацию), свойствами каналов связи и характером сигналов, используемых для передачи информации.

В тех случаях, когда передача информации производится с целью управления большими количествами энергии (например, для управления атомной энергией), количество энергии, расходуемой на передачу управляющей информации, во много раз меньше количества энергии, которой управляет эта информация. Это обстоятельство с выгодой используется при автоматическом управлении или управлении на расстоянии агрегатами, работа которых связана с расходом больших количеств энергии.

Приборы или органы, осуществляющие преобразование информации, как и каналы связи и запоминающие устройства, подвержены влиянию внешних воздействий, посторонних по отношению к системе, в состав которой

они входят. Эти воздействия (помехи) могут вызывать сбои в работе преобразователей информации, приводящие к искажениям выдаваемой ими информации.

Математика в поисках способов решения своих задач издавна занималась разработкой алгоритмов, при этом под алгоритмом понималось четкое предписание, определяющее вычислительный процесс, вполне понятное и не допускающее никакого произвола для его исполнителя. Человек, даже не знающий той области математики, для решения задач которой разработан алгоритм, но понимающий содержащиеся в нем указания и точно их выполняющий, получает искомое решение задачи. Можно сказать, что наличие алгоритма позволяет решать задачи механически.

Алгоритму должны быть присущи следующие три свойства:

1. Определенность, состоящая в четкости образующих его указаний, их полной понятности для исполнителя и полной однозначности.

2. Массовость, заключающаяся в его применимости не к одной единственной системе, а к некоторому множеству систем исходных данных. Это значит, что алгоритм должен позволять решать любую задачу из некоторого класса задач, а не какую-нибудь единственную конкретную задачу.

3. Результативность, т. е. определяемый алгоритмом процесс должен привести от исходных данных задачи к результату, который является именно решением задачи.

При этом математика требует, чтобы алгоритм был потенциально осуществим, т. е. чтобы для любой допустимой системы исходных данных он приводил к искомому результату после выполнения конечного (хотя бы и очень большого) числа «шагов» (операций). Сами операции, производимые при выполнении алгоритма, должны быть весьма простыми. Может оказаться, что алгоритм при некоторых исходных данных не приводит к определенному результату или не заканчивается в конечный срок. В таком случае говорят, что алгоритм не применим к данному классу задач (а может быть, применим к более узкому классу задач).

В качестве примера алгоритма можно привести известный алгоритм Эвклида, разработанный для решения задач следующего вида. Даны два натуральных числа a и b (т. е. два положительных целых числа). Требуется найти их общий наибольший делитель.

Алгоритм Эвклида может быть сформулирован в виде предписания, состоящего из пяти указаний.

Указание I. Проверить, справедливо ли неравенство $a > b$. Если да, перейти к указанию IV, если нет — к указанию II.

Указание II. Проверить, справедливо ли неравенство $b > a$. Если да, перейти к указанию V, если нет — к указанию III.

Указание III. Принять a за искомый ответ. Прекратить процесс.

Указание IV. Отнять b от a и результат принять за a . Вернуться к указанию I.

Указание V. Отнять a от b и результат принять за b . Вернуться к указанию I.

Например, если требуется найти общий наибольший делитель чисел 12 и 18, то считаем $a=12$, $b=18$ и приступаем к выполнению алгоритма.

Выполняя указание I, мы устанавливаем, что неравенство $a > b$ не справедливо, и переходим к указанию II. При выполнении указания II мы обнаруживаем, что неравенство $b > a$ справедливо, и переходим к указанию V. После выполнения этого указания получаем $a=12$, $b=18 - 12=6$ и переходим снова к указанию I. Указание I теперь требует перехода к указанию IV, в силу которого получаем $a=12-6=6$; $b=6$ и снова возвращаемся к указанию I. Последнее направляет нас к указанию II (так как $a=b$, а не $a > b$), которое направляет нас к указанию III, гласящему (в данном случае), что общий наибольший делитель чисел 12 и 18 равен 6. Это же указание предписывает нам теперь остановиться, прекратить дальнейшее выполнение алгоритма.

Алгоритм Эвклида может выполнять человек, не знающий, что такое общий наибольший делитель, но понимающий смысл указаний, образующих алгоритм, и умеющий их выполнять.

Известно, что существуют машины, которые могут выполнять математические операции (например, арифмометр может выполнять сложение, вычитание, умножение и деление чисел). Каждое предписание алгоритма требует выполнения некоторой простой операции. Если такие операции будет выполнять некоторое устройство (машина, которая, кроме операций над информацией, может делать переходы от одного указания к другому), входящее в состав кибернетической системы, то, очевидно, при соответствующем выборе алгоритма можно получить необходимое преобразование информации. По этим соображениям понятие алгоритма стало широко использоваться также и в кибернетике.

В кибернетике алгоритм — это информация, определяющая содержательное преобразование некоторого допустимого множества исходных сообщений (сигналов) в искомые результаты (тоже сигналы или состояния некоторого устройства).

Таким образом, алгоритм перерабатывает информацию либо в новую информацию, как это происходит в преобразователе информации, либо в действия некоторого устройства, как это имеет место в управляемых органах (устройствах). Результатами могут быть либо изменения состояния, если состояния являются положениями в пространстве, или сами состояния, если состояния представляют собой различные виды движения. Частный, простейший вид алгоритма — применение таблицы, т. е. осуществление заранее заданного соответствия.

В кибернетике свойство массовости алгоритма заключается в его применимости к любому сообщению, которое в данной кибернетической системе может быть передано преобразователю информации. Поскольку решение математической задачи является частным случаем преобразования информации, понятие алгоритма в математике следует рассматривать как частный случай этого понятия в кибернетике.

Так как алгоритм является информацией и, следовательно, представляется в виде совокупности сигналов, он сам может подвергаться преобразованиям. Эти вопросы изучает теория алгоритмов.

Особый интерес представляют два вида преобразований алгоритмов: содержательное преобразование в процессе их выполнения (алгоритм, преобразуя информацию, попутно преобразует и самого себя) и эквивалентные преобразования, при осуществлении которых алгоритм превращается в новый, который, однако, перерабатывает исходную информацию в те же результаты, что и первоначальный алгоритм. Первый вид преобразований позволяет делать алгоритмы особо компактными. При втором виде преобразований можно одни алгоритмы заменять другими, которые оказываются более простыми, более быстродействующими или надежными. Производя эквивалентные преобразования алгоритмов, являющихся описаниями преобразователей информации, мы получаем описания новых преобразователей информации, эквивалентных исходным.

Кроме описанных вопросов, теория алгоритмов изучает также способы построения из заданных алгоритмов более сложных, т. е. способы конструирования алгоритмов.

В сложных кибернетических системах возникает, передается, накапливается в памяти и подвергается преоб-

разованиям информация о различных последовательностях событий, свойствах предметов и др. Анализ сложных систем, когда информация передается в виде дискретных сигналов (такие кибернетические системы называются дискретными), показал, что в них существует некоторый набор элементарных сигналов, каждый из которых либо соответствует элементарному событию, либо описывает возможный вид связи между элементарными событиями. Сигналы, несущие информацию о сложных событиях, строятся по определенным правилам из элементарных сигналов или из сигналов, которые уже построены из элементарных сигналов. Другими словами, в каждой дискретной кибернетической системе существует свой язык, служащий для представления информации. Такой язык называется информационным языком, а элементарные его компоненты — морфемами.

Морфемой называется минимальная значащая часть высказывания или основная единица языковой системы, обладающая свойствами значимости, предельности и повторяемости. Морфемы могут состоять из еще более простых сигналов, которые не имеют уже самостоятельного значения и называются символами информационного языка. В реальных дискретных кибернетических системах как число символов, так и число морфем конечно. В тех случаях, когда морфемы получаются из символов путем объединения их в последовательности, символы называются буквами. Совокупность правил построения сообщений из морфем и из уже построенных сообщений называется синтаксисом языка. Соответствие между сообщениями и событиями, т. е. информация, передаваемая сигналами, представляет собой семантику информационного языка.

В простейших случаях информационный язык состоит из нескольких морфем и имеет «пустой» синтаксис, т. е. в языке не предусмотрено образование из морфем более сложных сообщений.

Кибернетические системы могут состоять из подсистем, в каждой из которых действует свой информационный язык. В этом случае при передаче информации из одной подсистемы в другую производится переход от одного информационного языка к другому, представляющий собой перекодирование информации. Между информационными и естественными разговорными языками существует некоторая аналогия, однако есть и коренные

различия, обусловленные разницей между естественными природными явлениями и кибернетическими системами, учитывающими только одну сторону подобных явлений.

4. Кибернетические системы

К кибернетическим системам относят любые самоуправляемые и саморегулирующиеся системы. Это могут быть искусственные системы, созданные человеком (например, холодильник, поддерживающий нужную температуру в камере; автоматический рулевой на корабле), и естественные системы (например, организмы животных), в которых температура тела или состав крови автоматически поддерживаются постоянными, несмотря на различные возмущающие воздействия. Саморегулирующимися могут быть и некоторые коллективы людей и объединения животных. Простейшей системой такого рода является система из определенного вида животных (например, зайцев или карасей) и хищников, питающихся этими животными (соответственно волков или щук): чрезмерное размножение хищников приводит к уменьшению источника их питания, а это вызывает уменьшение численности хищников; в свою очередь уменьшение численности хищников вызывает рост популяции соответствующих животных. Такая же связь существует и в более сложных, многоступенчатых системах: растительность — гусеницы; гусеницы — насекомоядные птицы; насекомоядные птицы — хищные птицы.

Простейшим видом кибернетических систем являются самоуправляемые системы.

В общем случае можно считать, что самоуправляемая система состоит из трех основных элементов: управляемого объекта, воздействующего на внешнюю среду; чувствительного элемента, на который влияют воздействия внешней среды и изменения состояния управляемого объекта; и регулятора, или управляющего элемента. Регулятор по возможности должен быть изолированным от непосредственных воздействий внешней среды. Нередко в состав самоуправляемых систем входит еще четвертый элемент — запоминающий (память). Упрощенная схема самоуправляемой системы приведена на схеме 1.

Чувствительный элемент соединен с регулятором каналом связи 1. По этому каналу связи осведомляющая информация, выработанная чувствительным элементом,

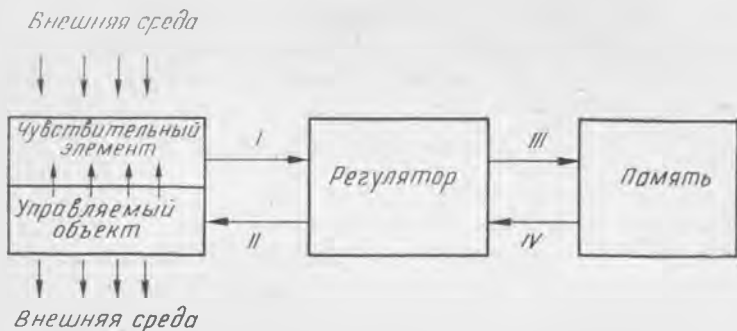


Схема 1. Упрощенная схема самоуправляемой системы.

поступает в регулятор. Здесь происходит ее преобразование в управляющую информацию, которая по каналу связи II передается управляемому объекту. По каналу связи III регулятор может пересылать в память как осведомляющую, так и соответствующую ей управляющую информацию, а по каналу связи IV вызывать ранее запасенную информацию для повторного использования (что убыстряет его работу) или для совместного использования с вновь полученной осведомляющей информацией (накопление «опыта» кибернетической системой). По каналу IV происходит также пересылка из памяти в регулятор преобразующей информации, которая требуется для переработки осведомляющей информации в управляющую.

Необходимо отметить, что в самоуправляемых системах между управляемым объектом и регулятором имеется так называемая обратная связь, обратное воздействие управляемого объекта на регулятор. Благодаря этому чувствительный элемент воспринимает изменения состояния управляемого объекта, возникающие под влиянием управляющей информации, и посылает в регулятор осведомляющую информацию об этих изменениях.

В наиболее простых случаях чувствительный элемент и управляемый объект могут быть объединены в одно целое.

В любой кибернетической системе чувствительный орган (или многие чувствительные органы) под влиянием внешних воздействий и воздействий на него других частей системы вырабатывает осведомляющую информацию. Точно так же в любой самоуправляемой системе возникшая в регуляторе управляющая информация уп-

равляемыми органами преобразуется в «действия» системы, влияющие на внешнюю среду. Это возможно благодаря тому, что управляющая информация (как и всякая другая) физически существует в виде определенных сигналов; эти сигналы и вызывают необходимые физические изменения в управляемых органах. Вид этих изменений зависит от характера сигналов, т. е. от информации, которую они несут.

Наиболее интересную и практически важную разновидность самоуправляемых систем представляют собой саморегулирующиеся системы. В этих системах, помимо описанных выше процессов, происходит выработка преобразующей информации. В дискретных системах этот процесс представляет собой видоизменение или выработку новых алгоритмов. Здесь имеется в виду не самоизменение алгоритма, которое может происходить в процессе его выполнения и заложено в нем самом, а изменение, вызываемое внешними причинами. Каким образом протекает процесс выработки преобразующей информации в естественных кибернетических системах, например в организмах животных, выяснить еще не удалось, хотя работы в этом направлении начаты еще до возникновения кибернетики (например, труды И. П. Павлова) и ведутся весьма интенсивно. В искусственных самоорганизующихся системах для этой цели применяют различные генераторы случайных сигналов. Такие сигналы поступают в регулятор, участвуют в выработке управляющей информации и, если состояние системы в результате этого улучшается (о чем регулятор «узнает» благодаря наличию в системе обратной связи), они запоминаются и применяются в дальнейшем. Для оценки состояния системы в состав ее преобразующей информации включают алгоритм, который также может подвергаться изменениям.

Изучение самоуправляемых и саморегулирующихся систем интересно потому, что проливает свет на ряд вопросов физиологии, биологии, психологии, без изучения которых невозможно построить сколько-нибудь сложные автоматы.

Исследуя вопросы построения автоматов, выполняющих отдельные психические функции человека, кибернетика встала перед проблемой анализа умственных действий человека с тем, чтобы разложить их на определенные элементарные преобразования информации.

Возник вопрос, какие виды умственной работы, выполняемой человеком, могут быть возложены на автомат. Полного ответа на него еще дать нельзя, однако с уверенностью можно утверждать, что всякую умственную деятельность человека, которую удалось описать в виде алгоритма, можно переложить на машину. Для этого требуется алгоритм представить на языке данной машины и ввести его в ее память. Требование математики, заключающееся в потенциальной осуществимости алгоритма, должно быть при этом заменено требованием реальной осуществимости алгоритма за допустимое время. Этот вопрос решается рассмотрением в совокупности алгоритма и кибернетической машины, предназначенной для его выполнения. Если найдена кибернетическая машина, осуществляющая алгоритм за допустимое время, то эту машину, а также выполняемый ею алгоритм называют эффективным.

Создание саморегулирующихся кибернетических систем позволяет предположить, что возможности воспроизведения с помощью машин некоторых умственных функций человека не ограничиваются функциями математических расчетов и технических вычислений.

В результате усилий, направленных на создание кибернетических машин, выполняющих те или иные умственные функции человека, возникла упоминавшаяся выше важная область кибернетики, изучающая вопросы построения вычислительных и управляющих алгоритмов, — теория алгоритмов.

Весьма важной является задача сравнительного анализа процессов переработки информации, происходящих в различных естественных и искусственных организованных системах. Кибернетика выделяет следующие основные классы информационных процессов: мышление и рефлекторная деятельность живых организмов; изменение наследственной информации в процессе эволюции биологических видов; переработка информации в автоматических системах: переработка информации в экономических и административных системах; переработка информации в процессе развития науки. Выяснение общих закономерностей этих процессов составляет одну из основных задач кибернетики.

Теория управляющих систем изучает структуру и принципы построения управляющих систем. В общем случае управляющей системой может быть назван любой

физический объект, осуществляющий целенаправленную переработку информации (нервная система животного, система автоматического управления движением самолета, электронная программно-управляемая вычислительная машина, система управления отраслью промышленности и др.).

Кибернетика изучает абстрактные управляющие системы, представленные в виде математических схем (моделей), сохраняющих основные информационные свойства соответствующих классов реальных систем. В рамках кибернетики возникла специальная математическая дисциплина — теория автоматов, изучающая класс дискретных управляющих систем, моделирующих работу нейронных сетей. Большое теоретическое значение имеет выяснение на этой основе механизмов мышления и структуры мозга, обеспечивающих возможность восприятия и переработки огромных количеств информации в органах малого объема с ничтожной затратой энергии и с исключительно высокой надежностью.

В кибернетике выделяются два общих принципа построения управляющих систем: обратной связи и многоступенчатости (иерархичности) управления. Принцип обратной связи дает возможность управляющей системе постоянно учитывать фактическое состояние всех органов и реальных воздействий внешней среды. Многоступенчатая схема управления обеспечивает экономичность и устойчивость управляющих систем.

Сочетание принципов обратной связи и иерархичности управления придает управляющим системам свойство «ультраустойчивости», позволяющее системе автоматически находить оптимальные положения при любых непредвиденных изменениях внешней обстановки. Эти принципы обеспечивают не только устойчивость, но и приспособляемость управляющих систем к изменившимся условиям и являются основой, в частности, эволюции биологических видов.

Эти принципы, реализованные в нервной системе животных, дают возможность развития, обучения и приобретения опыта живыми организмами в процессе их жизни. Постепенная выработка условных рефлексов и их непрерывное усложнение являются ни чем иным, как повышением уровней управления в нервной системе данного животного. Принципы обратной связи и иерархичности управления используются также при построении слож-

ных управляющих систем в технике и организации процессов управления в общественной жизни.

При изучении управляющих систем возникают два рода вопросов:

1) анализ структуры управляющей системы и определение реализуемого ею алгоритма;

2) синтез из заданных элементов системы, обеспечивающей выполнение заданного алгоритма.

Общими требованиями при этом являются требования заданного быстродействия, точности работы, минимального количества элементов, надежности функционирования, минимума затрачиваемой энергии. Весьма плодотворно при исследовании структуры различных управляющих систем моделирование этих систем на ЭВМ универсального назначения.

Кибернетика занимается также вопросами общей теории организации систем. Для количественной оценки степени организации систем вводится числовая мера, характеризующая количеством информации, которую требуется ввести в систему для обеспечения перехода системы из начального беспорядочного состояния в требуемое организованное состояние.

Особый интерес представляют самоорганизующиеся системы, способные самостоятельно переходить из произвольных начальных состояний в устойчивые состояния, соответствующие характеру внешних воздействий. В общем случае состояние таких систем изменяется под влиянием внешних воздействий произвольным образом. Однако эти системы обладают упомянутыми выше регулирующими механизмами высших уровней, которые осуществляют запоминание и целенаправленный отбор наиболее устойчивых состояний, соответствующих характеру внешних воздействий.

Свойство самоорганизации может проявляться только у систем определенной степени сложности, в частности с избыточностью структурных элементов и случайным характером связей между ними, изменяющихся в результате взаимодействия системы с внешней средой. К таким системам относятся, например, сети нейронов мозга, некоторые типы колоний живых организмов, а также некоторые типы сложных экономических или административных объединений.

Основные разделы кибернетики по содержанию очень тесно связаны друг с другом. Процессы самообучения и

самоорганизации систем близки к процессам опознавания и классификации информации и т. д. Отдельные разделы кибернетики образуют единую теорию процессов управления, находящуюся еще в стадии становления.

Помимо методологического значения, кибернетика имеет большое практическое значение как теория, обеспечивающая единый подход к изучению искусственно созданных и естественных управляющих систем, а также рациональное построение различных систем сбора, передачи, хранения и обработки информации. Методы теории информации позволяют оценивать экономичность и надежность различных способов кодирования информации и проектировать каналы связи с высокой пропускной способностью и помехоустойчивостью и эффективные запоминающие устройства. Математические методы нахождения оптимальных решений используются в области экономического планирования и анализа. Они помогают правильно использовать ресурсы и производственные возможности и осуществлять оптимальное управление как отдельными предприятиями, так и отраслями хозяйства.

Кибернетика является теоретической основой автоматизации технологических процессов. Комплексная автоматизация на основе самообучающихся систем позволяет обеспечить достижение наивыгоднейших режимов управления, что особенно важно для сложных производств. Необходимой предпосылкой такой автоматизации является наличие для данного производственного процесса детального математического описания (математической модели), которое вводится в ЭВМ, управляющую процессом, в виде программы ее работы. В эту машину поступает информация о ходе процесса от различных измерительных устройств и датчиков, и машина на основе имеющейся математической модели процесса рассчитывает его дальнейший ход при тех или иных командах управления. Так как подобное моделирование и прогнозирование протекает значительно быстрее реального процесса, то имеется возможность путем расчета и сравнения ряда вариантов выбирать наивыгоднейший режим управления. Оценка и выбор вариантов могут производиться как самой машиной полностью автоматически, так и с помощью человека — оператора. Важную роль при этом играет проблема оптимального взаимодействия человека — оператора и управляющей машины.

Большое практическое значение имеет выработанный кибернетикой единый подход к анализу и описанию (алгоритмизации) различных процессов управления и переработки информации путем последовательного расчленения этих процессов на элементарные акты, представляющие собой альтернативные выборы («да» или «нет»).

Систематическое применение этого метода позволяет формализовать все более сложные процессы умственной деятельности, что является первым необходимым этапом в их последующей автоматизации.

Для повышения эффективности научной работы весьма перспективна проблема информационного симбиоза машины и человека, т. е. непосредственного взаимодействия человека и ЭВМ в процессе творческого мышления при решении научных задач.

Под непосредственным влиянием кибернетики возникла новая инженерная наука — бионика, изучающая управляющие системы и чувствительные органы живых организмов с целью использования их принципов действия и организации для создания соответствующих технических устройств. Получаемые при этом результаты находят практическое применение, в том числе, например, в области протезирования. Разработка подобных систем в свою очередь позволяет подойти к более глубокому пониманию процессов, происходящих в управляющих системах живой природы.

Кибернетика широко применяется в медицинской диагностике, лечении, планировании и управлении здравоохранения.

ПРИНЦИПЫ УСТРОЙСТВА ЭВМ И ЭЛЕМЕНТЫ ПРОГРАММИРОВАНИЯ

1. Устройство ЭВМ

Общая классификация средств вычислительной техники показана на схеме 2.

Собственно ЭВМ, составляющие техническую базу кибернетики, на указанной схеме обозначены как вычислительные машины цифровые электронные. Они включают три основных вида машин: универсального применения, управляющие машины и машины специализированные, рассчитанные на решение какой-либо одной задачи. Естественно, что данная классификация (как, впрочем, и любая классификация) является условной. Сейчас существует такое большое число различных видов и подвидов ЭВМ, что их трудно уложить в какую-то единую классификационную схему. Рассматриваемые ЭВМ относятся, как уже сказано, к классу цифровых машин. Это значит, что они оперируют с информацией, представленной в цифровом кодированном виде, например в общеизвестной десятичной системе счисления.

Из схемы классификации видно, что существует второй класс вычислительных машин, называемых аналоговыми. Название это связано с тем, что большинство машин этого класса работает по принципу аналогии — подобия между изучаемыми явлениями различной физической природы. Известно, например, что процессы распространения тепла, движения жидкостей и газов, распространения акустических волн и т. д. имеют много аналогий. Это значит, что математические уравнения, описывающие какой-либо один процесс, могут подходить и для описания другого процесса (из другой области явлений), если только в этих уравнениях придать другой физический смысл переменным величинам и коэффициентам уравнений. Отсюда следует, что для изучения процессов одной физической природы может быть построена модель другой физической природы.



Схема 2. Классификация вычислительных машин.

Наиболее удобны для воспроизведения и измерения в лабораторных и в производственных условиях процессы движения электрического тока, поэтому они используются для изучения других явлений. С помощью специальных электрических схем можно выполнять непосредственно различные арифметические и вообще математические действия (сложение, вычитание, умножение, деление, извлечение квадратного корня, дифференцирование, интегрирование, получение логарифмов, тригонометрических функций и т. д.). В этих схемах участвующие величины представляются в виде непрерывных значений электрических параметров (напряжения, силы тока, частоты и т. д.). В связи с этим аналоговые ЭВМ называются также машинами непрерывного действия в отличие от цифровых машин, оперирующих с дискретными кодированными значениями величин — с числами. На схеме 2 показаны и некоторые другие вспомогательные классы средств вычислительной техники (аналоговые механические устройства — например, логарифмическая линейка; цифровые настольные счетные машинки, перфорационные машины).

ЭВМ имеют две основные области применения: переработка информации для человека и по заданию человека; переработка информации, автоматически поступающей из внешних источников (без участия человека) и выдача результатов (опять автоматически, без участия человека) для управления другими машинами.

Машины первой группы называются вычислительными машинами, так как применяются для вычислений, т. е. для решения научных или инженерных задач. Машины второй группы получили название управляющих, так как служат для автоматизации процессов управления различными станками, энергетическими установками, транс-

портными средствами и т. п. Существуют также ЭВМ, в той или иной мере сочетающие эти функции. Некоторые управляющие машины наряду с автоматическим управлением объектами выдают результаты и в числовой форме для анализа их человеком. Есть также вычислительные машины, которые получают исходные данные для расчетов автоматически от датчиков, связанных с различными измерительными устройствами, а результаты выдают на печать или на экраны для восприятия человеком.

Перфорационные машины в настоящее время (в связи с широким применением ЭВМ) играют вспомогательную роль. Они используются в основном для решения простых бухгалтерских и статистических задач массового характера в тех случаях, когда отсутствуют ЭВМ. Типовой комплект перфорационных машин состоит из основной машины-табулятора, осуществляющего арифметические действия и печать таблиц; машины для сортировки данных и устройства для подготовки данных для ввода в указанные машины. Это устройство называется перфоратором, а работница, работающая на нем, — перфораторщицей.

Используются два основных способа ввода данных в вычислительные машины:

— перфорация, т. е. пробивка отверстий на стандартных листах картона (перфокартах) или на стандартных бумажных листах (перфолентах);

— непосредственный ввод в ЭВМ (запись на магнитный носитель информации либо ввод в оперативную память с клавиатуры дисплея — см. ниже).

В перфорационные машины данные вводятся только путем перфорации и только с помощью перфокарт. В электронные машины данные вводятся обоими способами. При перфорации определенные комбинации отверстий на фиксированных позициях означают определенные буквы или цифры. Работа перфораторщицы сходна с работой машинистки, только здесь при нажатии соответствующей клавиши пробиваются отверстия в перфокарте (или перфоленте). Затем пробитые карты вводятся в так называемое читающее устройство, которое прощупывает их поверхность и посылает электрические сигналы в соответствии с наличием пробивок. Определенные комбинации пробивок в каждой колонке перфокарты соответствуют определенным цифрам или буквам.

В ЭВМ используется ряд способов представления чисел, или, как принято говорить, ряд систем счисления. Для записи исходных данных и для выдачи результатов расчетов применяется обычная десятичная система. Внутри машины числа представляются в так называемой двоичной системе, в которой основанием системы служит двойка и, таким образом, все числа выражаются сочетанием нулей и единиц. Напомним, что в математике принято считать, что всякое число в нулевой степени равно единице. Имея это в виду, можно записать, например, число 29 в виде суммы $2 \cdot 10^1 + 9 \cdot 10^0$.

Подобное разложение по степеням основания можно получить для любого числа и в двоичной системе. Например, для числа 29 двоичное разложение будет иметь вид: $1 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0$. Записывая только последовательность коэффициентов при степенях основания, получим двоичное представление числа 29 в виде: 11101. В качестве переходной системы между десятичной и двоичной используется так называемая двоично-десятичная система, в которой основанием системы является 10, а каждая цифра десятичного числа представляется в двоичной системе. Например, то же число 29 в двоично-десятичной системе будет иметь вид: 00101001. Здесь первые четыре позиции представляют двойку, а следующие четыре позиции — девятку. Переход от десятичной системы к двоично-десятичной осуществляется во время перфорации: нажимается клавиша с соответствующей цифрой, а пробиваются сразу нужные комбинации из четырех двоичных позиций. Дальнейший переход к чисто двоичной системе производится уже внутри машины путем специальных расчетов. При выдаче результатов из машины осуществляется обратный процесс преобразований. Таким образом, видно, что применение только двух двоичных цифр никак не ограничивает возможностей представления в ЭВМ различных данных. Буквы, например, представляются 5-, 6- или 8-разрядными двоичными комбинациями — кодами; каждой букве соответствует определенная двоичная комбинация. Из букв строятся слова, фразы, предложения и таким образом машина может хранить и обрабатывать не только числовые данные, но и текстовую информацию, например, наименования медикаментов, фамилии пациентов, формулировки диагнозов, назначений и т. д. Применение двоичной системы в ЭВМ обусловлено техническими соображениями (про-

стога выполнения арифметических действий, удобство запоминания и др.).

На схеме 3 приведена укрупненная структура ЭВМ. Основными частями этой машины являются:

1. Оперативное и внешнее запоминающие устройства (ОЗУ и ВЗУ), которые служат для приема, запоминания и выдачи информации. ОЗУ выдает информацию в любом порядке и с высокой скоростью, но обладает ограниченным объемом. ВЗУ выдает данные последовательно (с меньшей скоростью), но имеет большую емкость.

Оперативное запоминающее устройство состоит из отдельных ячеек, в каждую из которых может быть помещено одно число. Все ячейки запоминающего устройства машины перенумерованы подряд и каждой ячейке присвоен постоянный номер, называемый адресом ячейки.

Ячейки строятся из элементарных частей, называемых двоичными разрядами. Для кодирования десятичных цифр нужно иметь не меньше четырех разрядов. В машинах третьего поколения применяются, как правило, очень короткие ячейки (по 8 двоичных разрядов), рассчитанные на хранение элементарных единиц данных. Такая элементарная единица данных из 8 двоичных разрядов называется байтом. Одним байтом можно выразить десятичные цифры или одну букву.

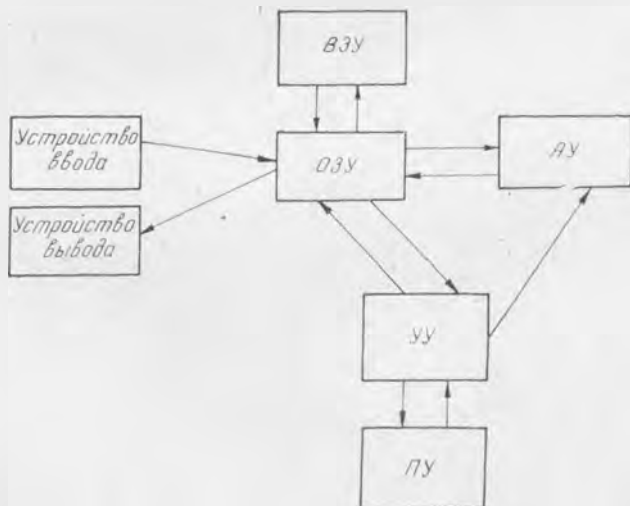


Схема 3. Укрупненная структура ЭВМ.

ВЗУ служит как бы резервуаром большой емкости, а ОЗУ — рабочим резервуаром: во время работы ЭВМ происходит передача данных между этими устройствами, так что в ОЗУ всегда хранятся данные, связанные с текущими вычислениями.

2. Арифметическое устройство (АУ), выполняющее отдельные операции над числами. В это устройство поступают из оперативного запоминающего устройства два исходных числа и из устройства управления (см. ниже) — команда, какую операцию над ними следует выполнить. Оно выполняет эту операцию, например складывает два заданных числа, и выдает результат в оперативное запоминающее устройство. В следующий такт работы ЭВМ в АУ поступают два других числа и новая команда; оно выполняет эту команду и т. д.

3. Устройство управления (УУ), управляющее выполнением программ решения задачи; УУ выбирает из ОЗУ поочередно команды программы, расшифровывает их и выдает всем остальным устройствам ЭВМ сигналы, обеспечивающие реализацию данной команды.

Машинная команда — это условный числовой код (выраженный также двоичным числом), заставляющий машину выполнять определенную операцию. Последовательность команд составляет программу работы машины. Программа решения задачи составляется заранее и служит для решения определенной задачи (при различных исходных данных). Исходную программу составляет человек — программист, как правило, в некотором условном, символическом виде. Перед решением задачи программа вводится в ЭВМ и преобразуется в последовательность машинных команд.

Программа работы машины, представленная в виде последовательности чисел, хранится в запоминающем устройстве машины. Характерной особенностью сложных вычислительных процессов является необходимость изменять порядок расчетов в зависимости от того, какие получаются промежуточные данные. Это осуществляется с помощью специальных команд ЭВМ, называемых командами условного и безусловного переходов.

Команды программы выполняются в том порядке, как они записаны и введены в ОЗУ, до тех пор, пока не встретится команда условного перехода или команда безусловного перехода. Команда условного перехода осуществляет проверку заранее заданного условия и переход к

определенной команде программы в зависимости от результата этой проверки («да» или «нет»).

Команда безусловного перехода осуществляет переход в любое заранее заданное место программы без проверки каких-либо условий.

Работу этих команд хорошо иллюстрирует описанный выше алгоритм Эвклида. В нем указания I и II содержали команду условного перехода: в зависимости от того, выполняется или не выполняется приведенное там неравенство, для продолжения вычислений надо было обращаться в одно или другое место алгоритма. Указания IV и V содержали команду безусловного перехода: независимо от результата последнего действия следовало переходить к указанию I.

Помимо перечисленных частей ЭВМ, на схеме показаны пульт управления (ПУ) и устройства ввода и вывода данных (УВВ). ПУ служит для пуска и остановки машины и контроля ее работы. Ввод данных производится, как уже говорилось, с помощью перфокарт и перфолент. Результаты выдаются из ЭВМ с помощью печатающих устройств.

Широкое применение в современных ЭВМ получили дисплеи, т. е. устройства ввода информации с клавиатуры и вывода информации на экраны. Они состоят из экрана телевизионного типа и клавиатуры пишущих машинок.

Дисплеи бывают двух основных типов: алфавитно-цифровые и графические. Алфавитно-цифровые дисплеи предназначены для ввода и вывода (отображения) текстовой алфавитно-цифровой информации. Графические дисплеи обеспечивают ввод и отображение, помимо алфавитно-цифровой информации, также различных геометрических фигур — графиков, схем, рисунков, чертежей и т. д. Оба вида дисплеев обеспечивают оперативное взаимодействие между человеком и ЭВМ в реальном масштабе времени.

Генерация графических изображений на экране дисплея, как и выдача текстовой информации, осуществляется по программам, выполняемым ЭВМ, к которой подключены дисплеи. Как правило, дисплеи имеют собственную оперативную память небольшого объема (1300—4000 ячеек) для временного хранения вводимой или отображаемой информации. Для ввода данных дисплей снабжается клавиатурой. Дисплей, предназначенный только

для отображения информации, может не иметь клавиатуры.

Важной деталью дисплея является так называемое световое перо — световой указатель, который позволяет указывать любую точку на экране. Таким путем можно выделять отдельные буквы или цифры текстового изображения или отдельные точки на графических изображениях, удалять их или наносить новые.

К одной ЭВМ может подключаться много дисплеев — несколько десятков и даже сотен, причем они могут располагаться на любом расстоянии от ЭВМ. При необходимости это позволяет пользоваться вычислительной машиной лицам, даже удаленным от нее на большие расстояния. Дисплеи, расположенные вдали от ЭВМ и используемые для взаимодействия с ЭВМ (для ввода в ЭВМ исходных данных и получения результатов решения задач), называются также терминалами (т. е. оконечными устройствами).

Запоминающие устройства ЭВМ строятся на различных технических принципах: магнитная запись, электростатическая запись, голография и др. В настоящее время наибольшее применение получили магнитные запоминающие устройства различных видов. Оперативные запоминающие устройства чаще всего строятся в виде кубов, состоящих из сотен тысяч маленьких магнитных сердечников (колечек). Каждый сердечник служит для запоминания одной двоичной цифры (нуля или единицы). Устройства с такими сердечниками обладают высоким быстродействием (запись или считывание одного числа производится за 1 мкс), высокой надежностью, малыми габаритами и малым потреблением энергии. Во внешних запоминающих устройствах используются магнитные ленты (подобные лентам магнитофонов), магнитные барабаны, магнитные диски. Особенно большую роль в современных машинах играют запоминающие устройства на магнитных дисках. Подробнее эти устройства будут рассмотрены в дальнейшем. Емкость внешних магнитных запоминающих устройств измеряется сотнями миллионов чисел, что позволяет их использовать для хранения различных архивов, каталогов и т. д.

Принято условно различать ЭВМ четырех поколений. Машины первого поколения строились на электронных лампах. Для машин второго поколения характерно ши-

рокое применение полупроводниковых элементов (транзисторов и диодов). В настоящее время производятся машины третьего поколения, в которых применяются элементы на так называемых интегральных схемах, т. е. весьма малогабаритных блоках, содержащих в малом объеме сложные схемы.

Применение, в каждом следующем поколении новых элементов привело к резкому уменьшению габаритов машин, повышению их надежности и быстродействия. Уже существуют машины с быстродействием в десятки миллионов операций в секунду. Совершенствовалась и структура ЭВМ, или, как говорят, их архитектура. ЭВМ третьего поколения строятся по модульному принципу, при котором имеется возможность изменять состав оборудования машины в зависимости от объема задач, которые предполагается решать в том или ином вычислительном центре: можно к машине присоединить дополнительные блоки памяти, устройства ввода и вывода информации и другие устройства.

Машины третьего отделения могут обслуживать заказчиков на расстоянии, используя телефонную или телеграфную связь. Одновременно каждая машина может решать несколько задач.

В настоящее время уже создаются ЭВМ четвертого поколения, которые будут использовать новые физические принципы — голографию, лазерную технику и обладать еще большими информационными и вычислительными возможностями. Упростится процесс общения рядовых потребителей с этими машинами: задания для них можно будет задавать на обычном языке (по телефону, телеграфу, почтой) и результаты получать сразу на экране телевизионного типа. Сейчас в СССР и за рубежом выпускаются и находятся в эксплуатации в основном ЭВМ третьего поколения.

В табл. 1 приведены основные характеристики наиболее распространенных ЭВМ единой серии: ЕС-1022, ЕС-1033 и ЕС-1050. Эта система машин построена на интегральных схемах с применением прогрессивных конструктивно-технологических решений и процессов, что обеспечивает высокую надежность и возможность массового промышленного выпуска ЭВМ.

Помимо ЭВМ серии ЕС, в нашей стране выпускаются ЭВМ третьего поколения серии АСВТ (агрегатированные средства вычислительной техники).

Таблица 1

Основные характеристики некоторых ЭВМ третьего поколения

Тип ЭВМ	Быстродействие, тыс. операций в секунду	Емкость байты			Примечания
		ОЗУ	МД	МЛ	
ЕС-1022	80	128—512 тыс.	15—60 млн.	80—160 млн.	<ol style="list-style-type: none"> 1. Возможно подключение до 128 медленных и 2 быстрых каналов обмена данными 2. Предназначена для ВЦ учреждений и предприятий
ЕС-1033	200	256 или 512 тыс.	15—60 млн.	До 160 млн.	<ol style="list-style-type: none"> 1. Могут решаться одновременно 15 задач 2. Возможно подключение до 256 медленных и 3 быстрых каналов обмена данными 3. Предназначена для крупных ВЦ, отраслевых АСУ
ЕС-1050	450	До 1024 тыс.	60—120 млн.	Практически неограниченна	<ol style="list-style-type: none"> 1. Может решаться одновременно несколько десятков задач 2. Возможно подключение 256 медленных и 6 быстрых каналов обмена данными 3. Предназначена для крупных ВЦ и АСУ

Важным свойством ЭВМ серий ЕС и АСВТ является программная совместимость разных моделей. Это значит, что разные модели имеют единую систему команд и могут выполнять одни и те же программы без каких-либо переделок. Ограничением в возможности выполнения какой-либо программы на данной модели ЭВМ может оказаться только ее мощность (быстродействие и емкость запоминающих устройств). При этом выполнение сложных программ на малых моделях этой серии может потребовать слишком много времени, т. е. практически оказаться нереальным.

Очень важно, что все модели комплектуются внешними или периферийными устройствами из единого стандартного набора (для ввода и вывода перфокарт и перфолент, для печати данных, для отображения данных на экранах, для хранения данных на магнитных лентах и дисках, для связи с абонентами по телефонным и телеграфным каналам связи). Эта стандартизация обеспечивает возможность расширения состава внешних устройств в зависимости от конкретных нужд потребителей. Для единой системы ЭВМ предусмотрено единое математическое обеспечение; потребитель вместе с ЭВМ получает набор готовых программ для решения типовых задач. Сюда входят прежде всего программы так называемой операционной системы, облегчающие управление работой машины (вводам и выводам данных, очередностью решения задач и т. д.), а также средства автоматизации программирования (набор языков программирования разного назначения — экономика, инженерное дело, научные расчеты и соответствующие средства для автоматического перевода исходных программ, записанных на одном из этих языков, в машинные программы). Кроме того, с ЭВМ поставляются библиотеки программ типовых, наиболее часто встречающихся задач (решение уравнений, статистическая обработка экспериментальных данных, расчет оптимальных планов и т. д.).

Рассмотрим в качестве примера несколько подробнее характеристики типовой ЭВМ третьего поколения — ЕС-1033. Эта ЭВМ средней производительности: имеются ЭВМ менее мощные и более мощные по сравнению с ней. Она широко предназначается для решения большого круга научных, экономических и информационных задач. Собственно вычислительная часть машины, называемая процессором и включающая ОЗУ, АУ и УУ, имеет следующие характеристики. Длина основного машинного слова равна 4 байтам, т. е. 32 двоичным разрядам. В составе машины может быть до восьми накопителей на магнитных дисках емкостью по 7,25 млн. байтов. Ввод данных в ЭВМ может производиться либо с перфоленты со скоростью 90 тыс. строк в минуту, либо с перфокарт со скоростью 500 карт в минуту.

Ввод небольших количеств данных осуществляется с помощью клавиатуры электрифицированной пишущей машинки (которая обычно стоит на пульте управления

ЭВМ) со скоростью 600 знаков в минуту. Вывод данных производится с помощью ленточного перфоратора со скоростью 9000 строк в минуту, с помощью карточного перфоратора со скоростью 100 карт в минуту, с помощью электрифицированной пишущей машинки со скоростью 600 знаков в минуту.

Основной вывод результатов решения задач производится с помощью алфавитно-цифрового печатающего устройства (АЦПУ), выдающего данные на широкую бумажную ленту со скоростью 770 строк в минуту (в строке предусматривается до 128 знаков, включая пробелы).

ЭВМ ЕС-1033 обладает возможностями работы в реальном масштабе времени (т. е. принимает данные и выдает ответы решения задач в процессе самих вычислений, без задержек) одновременно со многими абонентами-заказчиками. Для этого она имеет сопряжения с двумя видами каналов связи (обычно именуемых просто каналами): мультиплексным и селекторным.

Мультиплексный канал служит для подключения к системе ЕС-1035 медленнодействующих вводных и выводных устройств; скорость передачи данных в мультиплексном режиме составляет 40 000 символов в секунду; количество подканалов 256.

Селекторный канал служит для подключения к системе ЕС-1035 быстродействующих вводных и выводных устройств; он работает со скоростью передачи данных 800 килобайт в секунду.

Математическое обеспечение ЕС-1035 включает операционную систему, обеспечивающую управление режимами работы ЭВМ, регулирование порядка прохождения задач, управление внешними устройствами, в том числе управление работой накопителей на магнитных дисках, а также разветвленную систему автоматизации программирования задач. Последняя включает ряд алгоритмических языков (универсальный язык программирования ПЛ/1, язык для выдачи различных отчетов РПГ, машинно-ориентировочный язык Ассемблера и др.) с соответствующими трансляторами.

Помимо перечисленных ЭВМ, широкое применение в медицинских учреждениях находит ЭВМ третьего поколения М6000. Указанная ЭВМ относится к числу малых ЭВМ и предназначается для использования в качестве вспомогательной ЭВМ в крупных ВЦ, в качестве управляющей ЭВМ при автоматизации технологических про-

цессов и при моделировании сложных динамических систем, а также в отдельных научно-исследовательских и лечебных учреждениях. Информация хранится в 18-рядных двоичных ячейках, в которых размещаются по 2 символа-байта с контрольными разрядами. Скорость выполнения операций до 200 тыс. в секунду. Емкость оперативной памяти от 8192 до 65 736 байтов. Операции с повышенной точностью выполняются по подпрограммам. К ЭВМ может подключаться дополнительное оборудование для ввода и вывода данных без прерывания работы ЭВМ. Таким путем могут вводиться различные физиологические параметры непосредственно во время измерения. Программное обеспечение включает алгоритмические языки ФОРТРАН, сокращенный АЛГОЛ-60, язык диалогового типа БЭЙСИК, машинно-ориентированный язык мнемокод. Подробнее алгоритмические языки будут рассмотрены ниже.

Данная ЭВМ строится по модульному (наращиваемому) принципу, имеет очень небольшие габариты, малую потребляемую мощность и может обслуживаться 1—2 человеками.

Кроме перечисленных ЭВМ, в составе ВЦ учреждений используются счетно-перфорационные (СПМ) и настольные счетно-клавишные машины (ВКМ). Обычно СПМ используются в виде следующего типового комплекса: перфоратор алфавитно-цифровой ПА-80-2/1М; контрольный алфавитно-цифровой КА-80-2/1М; сортировка С-80-7; табулятор алфавитно-цифровой ТА-80; электронная вычислительная приставка к табулятору ВП-3 (для выполнения дополнительных операций); перфоратор итоговый позиционный ПИ-80У; перфоратор-репродуктор ПР-80У; расшифровочная машина.

СПМ применяются для механизации отдельных видов вычислительных работ, связанных с учетом материальных ценностей, бухгалтерскими расчетами, статистической обработкой информации.

В заключение обзора технических средств обработки информации следует упомянуть о некотором вспомогательном оборудовании, обычно используемом совместно с ЭВМ.

Телеграфный аппарат СТА-2М — это ленточный буквопечатающий аппарат, работающий на стандартной телеграфной ленте шириной 10 мм и оборудованный приставками автоматизации — трансмиттерной и реперфо-

раторной для нанесения алфавитно-цифровой информации в виде отверстий на перфорированную ленту шириной 17,5 мм. Предназначен для перевода на перфоленду данных с различных первичных документов. Средняя скорость обработки данных 40—100 знаков в минуту.

Абонентский телеграфный аппарат Т-63 представляет собой рулонный буквопечатающий аппарат, оборудованный автоматизирующими узлами: реперфоратором, транзиттером, автоответчиком, автостопом, звонковой сигнализацией. Этот аппарат печатает алфавитно-цифровую информацию на стандартной рулонной бумаге шириной 215 мм. Предназначен для работы с абонируемым каналом связи, для передачи информации в двух режимах: прямая передача телеграмм с пульта (текст и в виде перфоленды) и через транзиттер (только в виде перфоленды). Скорость прямой передачи такая же, как и СТА-2М; скорость передачи с транзиттера в 10 раз большая. Может быть использован автономно для подготовки данных (записи данных на перфоленду).

Настольные вычислительные многоклавишные машинки (тип «Искра II» или «Элка») предназначены для ручного выполнения вычислительных работ небольшого объема.

Перспективы развития технических средств медицинской кибернетики. ЭВМ, средства связи и другие средства сбора и обработки медицинских данных в общем случае должны обеспечивать одновременную работу в реальном масштабе времени со многими абонентами, удаленными на расстояние от нескольких сотен метров до нескольких сотен километров. Для того чтобы можно было отказаться от ведения бумажной документации (истории болезни, рецепты и т. д.) и заменить ее хранением данных в памяти ЭВМ, эти машины должны быть высоконадежными и обеспечивать многодневную круглосуточную работу. Для этого в технических средствах должно широко использоваться дублирование и резервирование аппаратуры. Емкость и надежность запоминающих устройств ЭВМ значительно увеличиваются. В составе ЭВМ будут запоминающие устройства емкостью до 1 млрд. слов с временем обращения в несколько миллисекунд — для хранения данных (полных историй болезней) пациентов, находящихся на обслуживании в больницах или поликлиниках, а также для хранения различных словарей, справочников и т. д. Кроме того, ЭВМ

должны обладать запоминающими устройствами емкостью в несколько миллионов слов с временем обращения 5—10 мкс — для хранения текущих данных и программ, находящихся в работе.

В состав ВЦ медицинского назначения должна входить развитая система периферийных устройств, обеспечивающих непосредственное обращение к ЭВМ (ввод или получение данных) со стороны врачей и среднего медицинского персонала без использования промежуточных этапов перфорации, кодирования и т. д. Выдача данных должна осуществляться с помощью дисплеев (терминалов) преимущественно в наглядной форме (графики, рисунки, схемы, таблицы, диаграммы и т. д.) с временем задержки (после ввода запроса) не более 1 с для врачей и не более 1 мин для административного и хозяйственного персонала. Время выдачи печатных документов оперативного характера не должно превышать нескольких минут.

Для исключения очередей в пользовании терминалами необходимо иметь их достаточное количество. По-видимому, для больницы на 1000 коек необходимы 30—40 терминалов, для поликлиники — 5—10 терминалов, для санитарно-эпидемиологической станции — 3—5 терминалов.

Желательно, чтобы обращение к терминалам производилось не с помощью клавиатуры, а с помощью голоса или путем показа световым указателем (пером).

Для внутренней связи в больнице с ВЦ целесообразно использовать видеотелефоны. Мощные ЭВМ в кустовом медицинском ВЦ должны дополняться малыми ЭВМ-спутниками, предназначенными для приема и обработки сообщений, переключения каналов связи, формирования, хранения и выдачи данных на устройства наглядного отображения, а также для специальных задач (обработка лабораторных анализов, электрокардиограмм, рентгеновских снимков и т. д.).

В состав периферийных устройств должны входить устройства, воспринимающие магнитные отметки, печатный (и даже рукописный) текст, голосовую речь, а также типовые устройства ввода перфокарт и перфолент.

Особое внимание при создании медицинских ВЦ должно обращать на автоматический контроль первичной информации, при этом должно проверяться ее соответствие ранее накопленным данным, отсутствие логиче-

ских и формальных ошибок (например, нахождение в допустимых пределах данных лабораторных анализов, выполнение заданных соотношений между разными позициями сообщений и т. д.).

Для обеспечения сохранности врачебной тайны необходимо обеспечить автоматическую идентификацию лиц, обращающихся к тем или иным данным в памяти ЭВМ. Для этой цели могут быть использованы условные коды, присвоенные постоянно каждому лицу, имеющему доступ к ЭВМ. Коды могут либо набираться на клавиатуре терминалов перед началом работы, либо вводиться с помощью специальных жетонов.

Программы работы ЭВМ с экранными терминалами должны предусматривать по крайней мере три диалоговых режима работы:

1) обучающий режим, когда ЭВМ выдает данные небольшого объема с подробным пояснением их назначения и порядка работы с экранным терминалом;

2) общий эксплуатационный режим, при котором разнообразные данные выдаются с минимальными пояснениями;

3) специализированный, при котором выдаются большие объемы однотипных данных для специалистов, без дополнительных инструкций по их использованию.

Применение диалогового режима обеспечит оперативное взаимодействие врачей и других медицинских работников с автоматизированными информационными системами.

Вполне реально в ближайшем будущем использование для связи человека с ЭВМ не только зрительного канала, но и слухового — голосового канала.

2. Элементы программирования

Для составления программы решения какой-либо задачи сначала составляют блок-схему вычислительного процесса, показывающую в наглядной графической форме весь ход процесса. Отдельные более или менее самостоятельные участки вычислений обозначают блоками в виде прямоугольников, внутри которых указывают либо наименование, либо содержание вычислений. Процедуры сравнения различных величин и выбора одного из двух возможных направлений продолжения вычислений обозначают ромбами, внутри которых указывают условия вы-

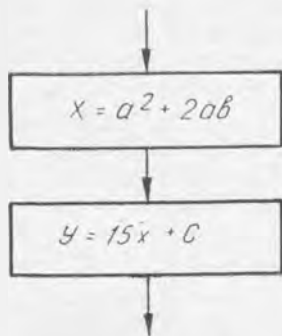


Схема 4. Пример вычислительных блоков.

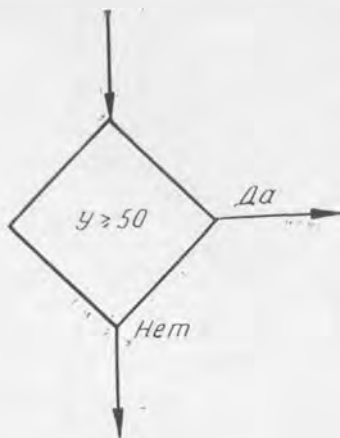


Схема 5. Пример логического блока.

бора (схемы 4, 5). Существуют другие блоки для ввода и вывода данных, для указания способов хранения данных и других действий. Отдельные блоки соединяют стрелками, показывающими порядок выполнения этих блоков. Все блоки обычно нумеруют, каждый блок имеет свой номер. Прямоугольники и ромбы являются основными видами блоков, которые достаточны для составления первой (эскизной) блок-схемы решения задачи. Подобные блок-схемы могут (и должны) составляться авторами задачи (медиками, экономистами, инженерами и т. д.), не являющимися программистами. Полезен также следующий прием, с помощью которого можно показывать разрывы в блок-схеме. Стрелку, которая должна соединить два удаленных друг от друга блока, делают с разрывом: отходящий конец заканчивается кружком, где пишут номер блока, к которому должна была подойти стрелка, а у этого блока рисуют входящий конец стрелки, начинающийся кружком, содержащим номер исходящего блока.

По этим блок-схемам программисты могут составлять более детальные блок-схемы машинных алгоритмов с указанием способов ввода и вывода данных, видов используемых накопителей, этапов ручной обработки и т. д. При решении сравнительно простых задач программисты могут обходиться без детальных блок-схем и состав-

лять программы, ориентируясь на эскизные блок-схемы. В конечном счете весь процесс вычислений, т. е. каждый блок, должен быть разбит на последовательность элементарных операций, и для каждой из них должна быть составлена своя команда. В первые годы применения ЭВМ этот процесс выполнялся в основном вручную, людьми — программистами. На составление и выверку программ тратилось много времени. Сейчас широко применяется автоматизация программирования, при которой программист записывает процесс решения задачи на некотором формальном, символическом языке, а сама машина по специальной программе, называемой транслятором, переводит программу на машинный язык, т. е. в последовательность команд. При этом сокращается трудоемкость программирования и резко уменьшается число ошибок. Пользуясь символическими языками, после соответствующего обучения, программы для ЭВМ может писать на алгоритмическом языке и специалист, не являющийся специалистом по программированию.

Сейчас одно из основных направлений в развитии ЭВМ заключается в разработке удобных символических языков, которые бы позволяли составлять задания для машин специалистам различных профилей.

В настоящее время используются два основных вида символических языков: машинно-ориентированные языки (язык Ассемблера, автокод) и алгоритмические языки.

Язык Ассемблера (это название принято для ЕС ЭВМ) называется машинно-ориентированным потому, что используются команды (операторы), по форме напоминающие машинные команды, в которых в символическом (буквенном) виде представляются коды машинных операций и величины, участвующие в операциях.

Вообще программирование задач для ЭВМ — очень трудоемкое дело даже при использовании современных средств автоматизации программирования. Особенно длительным является программирование, когда применяются машинные языки или язык Ассемблера, так как при этом весь вычислительный процесс приходится расчленять на элементарные операции, касающиеся отдельных величин. При этом велика вероятность появления ошибок, которые затем приходится находить и исправлять в процессе отладки программ. Недостатком программ, написанных на машинном языке или языке типа Ассемб-

лера, является также то, что эти программы применимы только для той ЭВМ (того типа), для которой они написаны. При переходе на ЭВМ другого типа программы необходимо переписывать заново на языке новой ЭВМ.

Эти недостатки привели к тому, что в настоящее время для программирования различных конкретных задач применяют в основном алгоритмические языки (АЛГОЛ, ФОРТРАН, КОБОЛ, ПЛ-1 и др.). Для научных и инженерных задач применяют языки АЛГОЛ и ФОРТРАН, а для программирования экономических задач широко используют язык КОБОЛ (и его различные модификации АЛГЭМ, АЛГЭК и др.). Универсальный язык программирования ПЛ-1 (Programming Language-1) применяется для программирования математических, экономических и информационно-логических задач. Этот язык имеет в своем составе средства для обработки текстов (строк символов), для обработки составных величин, для организации ф а й л о в¹ различного типа. В алгоритмических языках для указания арифметических действий используются обычные знаки операций (+, —, ×, /) и круглые скобки, указания логических операций — знаки этих операций ∨, ∧, →, ≡, ⊃ (см. раздел 4-й третьей главы) и знаки операций отношения (> — больше, < — меньше, ≥ — больше или равно, ≤ — меньше или равно, = — равно, ≠ — не равно). Для указания последовательности вычислений и выбора направления вычислений в зависимости от условий используются стандартные словесные символы: **если, то, иначе, перейти к**. Символ **перейти к** служит для указания переходов в программе от одних о п е р а т о р о в к другим. Тот оператор, к которому требуется перейти, должен иметь перед собой метку (имя или номер), и эта метка указывается после символа **перейти к**. Например, возможен такой участок программы на алгоритмическом языке:

Если температура ≥ 37 , то перейти к анализ
иначе перейти к расчет.

Здесь слово «температура» обозначает некоторую переменную величину, а слова «анализ» и «расчет» — метки некоторых операторов программы.

¹ Файл — набор данных одинаковой структуры, объединенных общим назначением, которому присвоено имя. Так, файлом будет, например, введенная в ЭВМ совокупность историй болезней для всех лечащихся в данном медицинском учреждении. Название этого файла — истории болезней.

Применение ЭВМ в медицине ведется в основном по пути создания интегрированных систем сбора и обработки различных медицинских данных или автоматизированных систем управления медицинскими учреждениями. И в том и в другом случае решаются не изолированные задачи, а комплекс взаимосвязанных задач, использующих общие входные, выходные или промежуточные данные. При этом важнейшим вопросом является организация информационной базы.

Под информационной базой (базой данных) принято понимать совокупность документов, машинных массивов информации, классификаторов, словарей, а также инструкций, определяющих порядок ввода, кодирования, формирования, контроля, выдачи и использования всех видов информации в системе.

Элементарную единицу информации, имеющую определенный смысл и не допускающую ее расчленение на более мелкие составные части, принято называть величиной или показателем. Каждая величина состоит из данных двух видов — признаков и оснований. Например, в величине «масса 73 кг» признаков два — «масса» и «килограмм» и одно основание — «73». В некоторых величинах может отсутствовать основание. Так, величина «заболевание — ангина» состоит только из признаков.

Помимо простых величин, существуют сложные величины, состоящие из нескольких простых. Так, дата — сложная величина, имеющая три признака, год, месяц, число, и три основания и соответствующие им числовые значения. Систематизированную совокупность величин, характеризующих один объект (акт), называют записью. Примерами записей являются истории болезней, рефераты статей, накладные на отпуск товаров и т. п. Среди величин, входящих в запись, выделяют одну или несколько величин, присущих только данному объекту. Такие величины называют ключами, или идентификаторами записей. Они служат для опознания и отыскания обозначаемых ими записей. Совокупность однотипных записей, характеризующих однородные объекты, называется массивом.

При программировании составляют описания величин, записей и массивов, используемых при решении задачи. Указывают тип величины, количество разрядов для значения, пределы изменения, состав и порядок наличия в

записи, число записей в массиве. Для описания величин в алгоритмических языках используются стандартные слова (**целый, вещественный, массив, список** и др.). Если, например, массив целых чисел представляет собой матрицу A размером 10×20 элементов, то описание этого массива может выглядеть так:

Целый массив A [1 : 10, 1 : 20].

Здесь выделенные слова представляют собой условные символы, показывающие, что некоторая прямоугольная таблица чисел, имеющая наименование A , содержит только целые числа. Число строк в этой таблице 10, а число столбцов 20. Каждый элемент этой таблицы обозначается A_{ij} , где i — номер строки, а j — номер столбца; i меняется от 1 до 10, а j — от 1 до 20.

Все программное обеспечение (называемое часто математическим обеспечением) с точки зрения выполняемых функций принято делить на три основные части:

1) программы для непосредственного решения производственных или научных задач. Это так называемые пользовательские или прикладные программы;

2) средства для автоматизации процессов программирования, включающие алгоритмические языки разной сложности и назначения, и программы перевода с этих языков на машинные языки;

3) операционные и вспомогательные программы, предназначенные для управления порядком решения задач на ЭВМ, совместной работой ЭВМ и периферийных устройств или работой нескольких ЭВМ, а также для ведения информационных массивов, библиотек программ. Часто в состав операционных систем включают и средства автоматизации программирования, и, кроме того, типовые программы, используемые при обработке информации (программы сортировки, перезаписи массивов, печати документов и др.).

С точки зрения порядка разработки программ все программное обеспечение обычно делят на две части:

1) общее программное обеспечение, поставляемое вместе с машиной и используемое при решении любых задач на ЭВМ независимо от специфики ее использования. Сюда относятся все программы операционной системы, понимаемой в широком смысле. Существует два типа операционных систем: **д и с к о в а я о п е р а ц и о н н**

ная система (ДОС) и общая операционная система (ОС). По сравнению с ДОС ОС обладает более широкими возможностями; она обеспечивает дополнительно к ДОС возможность работы ЭВМ в реальном масштабе времени и обмен данными между ЭВМ и большим количеством терминалов. В связи с этим она сложнее и требует большего объема памяти, чем ДОС;

2) специальное программное обеспечение, создаваемое для каждой конкретной области применения ЭВМ и учитывающее ее специфику.

Такие программы (и алгоритмы), несмотря на специфичность, часто оказываются пригодными для решения аналогичных задач в других областях. По этой причине в стране организована специальная служба государственного и отраслевого учета, хранения и выдачи по запросам разработанных алгоритмов или программ. Поэтому начиная применение ЭВМ в новой области (например, для обработки историй болезней или для расчетов дозы лучевой терапии), нужно использовать имеющееся общее программное обеспечение, поставляемое с машиной, а также специальное программное обеспечение, содержащееся в государственном или отраслевых фондах алгоритмов и программ, и только при невозможности их использования создавать новые программы.

Как показывает опыт, при разработке сложных программ чрезвычайно важно обеспечить с самого начала работы правильное взаимопонимание и четкое взаимодействие между автором и будущим потребителем результатов задачи, и программистом, занимающимся ее разработкой.

Желательно, чтобы описание задачи и все особенности последующего ее использования заказчик (врач, организатор здравоохранения, экономист, плановик и т. д.) представлял программисту в письменном виде. Последующие изменения и уточнения постановок задач, требования к формам документов, которые должны вырабатываться ЭВМ, требования к точности выходных данных и т. п. также должны оформляться письменно. Это позволит во многих случаях избежать недоразумений и исключить непроизводительные затраты труда программистов.

Серьезное внимание должно уделяться оформлению окончательной документации на программы, которая

должна включать описание для каждой задачи (программы) форматов входных и выходных документов и показателей с указанием видов величин (числовые, буквенные) и их разрядности. Должны быть описаны существо и ограничения методов обработки и способов контроля вычислений и составлены подробные инструкции операторам по подготовке информации и по работе на ЭВМ. Для эксплуатации программ и их передачи в другие организации необходимы подробные блок-схемы и распечатки программ с пояснениями, обеспечивающими возможность модификации программ для изменившихся условий их использования.

Главным принципом составления сложных программ является принцип модульной структуры. Сложная программа строится из автономных кусков, называемых модулями и выполняющих законченные функции преобразования информации. Модули связываются между собой с помощью специальной управляющей программы — диспетчера, которая включает их в работу в нужной последовательности.

Для сложных программ, состоящих из ряда автономных программных модулей, принято название «пакет пользовательских программ».

Модульный принцип построения сложных программ обеспечивает гибкость и устойчивость работы всей сложной системы программного обеспечения АСУ. При этом имеется возможность одновременного и независимого составления различных пользовательских программ и рабочих подпрограмм и простота их замены и корректировки, а также упрощается процесс отладки сложной системы программ.

Важным моментом в построении любой достаточно сложной медицинской программы является обеспечение контроля за первичными (входными) данными и контроля за правильностью вычислений. Нужно принять в качестве одного из принципов построения любых медицинских автоматизированных информационных систем следующее положение: везде, где ЭВМ может быть применена для контроля и исправления ошибок, допускаемых людьми (врачами, сестрами, перфораторщицами и т. п.) при подготовке, записи и вводе информации, она должна использоваться для этих целей и в общую программу (комплекс программ) должны включаться соответствующие участки.

3. Особенности программирования для ЭВМ третьего поколения

Рассмотрим кратко основные особенности построения программного обеспечения ЭВМ третьего поколения, имеющие значение при разработке АСУ. Во-первых, на организацию информационных массивов и всех вычислительных процессов в АСУ значительно влияет наличие в ЭВМ третьего поколения запоминающих устройств на магнитных дисках — устройств большой емкости с прямым доступом к информации. В таких устройствах удобно размещать библиотеки программ, словари, справочные таблицы, таблицы кодирования и другие часто используемые данные. При этом резко упрощается и ускоряется процесс поиска и переписи этих данных в оперативную память по сравнению со случаем, когда подобные данные постоянно хранятся на магнитных лентах. Подробнее основные сведения о структуре запоминающих устройств на магнитных дисках и способах расположения информации на них будут приведены ниже. Во-вторых, большое влияние на организацию вычислительных процессов и методику программирования задач АСУ оказывает наличие в ЭВМ третьего поколения развитых операционных систем, о которых говорилось выше. Напомним, что операционной системой в общем случае принято называть программу (или комплекс программ), постоянно находящуюся в оперативной памяти ЭВМ и обеспечивающую управление процессом решения одной или нескольких задач.

Операционные системы начали применяться уже в машинах второго поколения (например, Минск-32), но в машинах третьего поколения они получили достаточно полное развитие.

Основной функцией операционных систем является управление порядком решения различных задач и порядком выполнения различных этапов одной задачи в режиме так называемой пакетной обработки. При пакетной обработке в ЭВМ вводятся для решения сразу несколько задач вместе с необходимыми исходными данными. Вся эта информация размещается во внешних запоминающих устройствах и поочередно вызывается операционной системой в оперативную память для исполнения. Операционная система может учитывать и заранее установленную приоритетность (важность) задач. Опе-

рациональная система следит за тем, чтобы вместе с вызовом какой-нибудь программы для исполнения ей подавались бы в нужные рабочие области памяти и соответствующие исходные данные, а результаты решения своевременно выдавались на печать или во внешний накопитель, освобождая место для следующих программ и данных. Операционная система обеспечивает оптимальную загрузку всего оборудования ЭВМ и ее эффективное использование за счет совмещения во времени решения различных задач или разных этапов одной задачи. При этом может, например, оказаться так, что для одной задачи ЭВМ производит вычисления, для другой задачи выдает результаты на печать, а для третьей задачи производит ввод данных с внешнего накопителя.

Операционная система обеспечивает реализацию сложных стандартных процедур, связанных с обращениями к внешним устройствам — накопителям, устройствам ввода и вывода, каналам связи и др. Благодаря этому программист освобождается от необходимости каждый раз подробно программировать эти процедуры; он указывает в своей программе только укрупненные команды (так называемые макрокоманды или экстракоды), а реализацию этих команд обеспечивает операционная система с помощью стандартных подпрограмм, находящихся в ее составе. При этом имеется возможность постоянного пополнения и эффективного использования библиотек стандартных подпрограмм. Особенно развита эта функция в ЭВМ третьего поколения в связи с возможностью хранить такие библиотеки на магнитных дисках. Широкое использование библиотек стандартных подпрограмм связано с уже упомянутым модульным принципом построения больших программ, при котором новые сложные программы строятся в основном из готовых кусков — модулей с добавлением к ним необходимых переходных и нестандартных кусков. В ЭВМ третьего поколения предусматриваются средства для обеспечения гибкости использования модулей программ. Эти модули строятся в виде так называемых макроопределений, обращение к которым осуществляется с помощью макрокоманд. В самих модулях предусматриваются участки, обеспечивающие их настройку на тот или иной конкретный вариант работы. Эта настройка в большинстве случаев сводится к выбору из самого модуля определенных (достаточно автономных) участков и

соединению их в нужной последовательности. Все эти подготовительные процессы, связанные с настройкой отдельных модулей, а также с выбором нужных модулей и соединения их между собой, носят название генерации модулей и выполняются под управлением операционной системы.

С помощью операционной системы происходит управление процессом трансляции программ с входных языков (АЛГОЛ, КОБОЛ, ФОРТРАН, ПЛ/1 и др.) на промежуточный машинный язык, процессом ввода (загрузки) программ, распределения памяти для размещения программ и данных, а также процессом выполнения программ. При этом операционная система обеспечивает контроль за защитой зон памяти, т. е. исключает возможность обращения одной программы к зоне памяти, отведенной для другой программы. Важным свойством операционных систем ЭВМ третьего поколения является обеспечение возможности объединения в одну программу различных кусков, написанных на различных входных языках, т. е. приведение их к единому машинному виду и согласование между собой. Эта функция выполняется частью операционной системы, называемой «редактором». Так как программы, осуществляющие трансляцию (перевод) программ с входных языков на машинные, работают также под управлением операционной системы, то часто и эти программы относят к операционной системе в широком смысле этого понятия.

Операционная система осуществляет взаимодействие между ЭВМ и оператором, управляющим машиной. Она воспринимает указания оператора, набираемые на пульте или на пишущей машинке, выдает ему сообщения о ходе вычислений, о возможных сбоях, неисправностях оборудования и т. п.

Операционная система существенно упрощает работу оператора ЭВМ, освобождая его от текущего контроля за порядком решения задачи, вводом и выводом данных и программ. Особенно важными являются функции операционной системы при работе ЭВМ в так называемом режиме разделения времени одновременно со многими удаленными абонентами (пользователями), имеющими в своем распоряжении дисплеи (терминалы). Таких терминалов может быть подключено к одной достаточно мощной ЭВМ несколько десятков и даже сотен. Все эти пользователи могут запрашивать ЭВМ о выдаче

той или иной информации, вводить свои программы и исходные данные, вводить задания на решение определенных задач, корректировать свои информационные массивы, хранящиеся в запоминающих устройствах ЭВМ и т. п., причем все эти действия совершаются в произвольные моменты времени.

Операционная система должна принимать поступающие сообщения, разбираться с ними, разделять их по видам работ и очередности выполнения, организовывать исполнение соответствующих программ и выдачу результатов, вести учет работы ЭВМ по заказам отдельных абонентов.

При достаточно высоком быстродействии ЭВМ каждый пользователь может получать свои ответы практически почти без задержки, при этом другие абоненты ему не мешают и у него создается впечатление, что он один работает с ЭВМ. Такой режим работы называют работой в реальном масштабе времени. Для работы в этом режиме ЭВМ должна обладать устройствами, обеспечивающими возможность прерываний работы машины для ввода и вывода данных и для изменения порядка (очередности) решения различных задач.

При работе ЭВМ в системах управления технологическими процессами, когда она воспринимает исходную информацию не от людей-пользователей, а от электрических датчиков и выдает результаты не людям, а исполнительным механизмам, к режиму работы в реальном масштабе времени предъявляются более жесткие требования, обусловленные величиной допустимой задержки во времени реакции управляющего органа на изменение состояния системы. Как уже упоминалось, операционные системы могут иметь различную мощность и сложность — системы ДОС и ОС; каждая из них настраивается (генерируется) на определенный состав оборудования ЭВМ (ее конфигурацию). Так, операционная система ДОС рассчитана на работу со средними ЭВМ третьего поколения из единой серии (ЕС-1022, ЕС-1033), имеющими в своем составе магнитные диски. Операционная система ОС, будучи более широкой по своим функциям и более сложной, предназначена, в основном для более мощных моделей ЭВМ единой серии (ЕС-1040, ЕС-1050).

Следующей существенной особенностью ЭВМ третьего поколения является более высокая степень унифика-

ции логической структуры ЭВМ. Как уже упоминалось, в этих машинах в подавляющем большинстве случаев принято байтовое представление информации. При этом способе минимальной единицей информации в машине является не слово (машинная ячейка), а 8-разрядный двоичный код — байт, который выражает один буквенный символ.

В ЭВМ третьего поколения принята почти во всем мире единая система кодирования символов. Из байтов могут формироваться слова по 4 байта, соответствующие ячейкам памяти по 32 двоичных разряда; 2 байта образуют полуслово. В противоположность ЭВМ второго поколения, в которых наблюдалась полная разнокалиберность в размерах ячеек памяти (24, 37, 48, 50 и др.), в ЭВМ третьего поколения машинное представление данных унифицировано.

Для ЭВМ третьего поколения характерна также унификация структуры информационных массивов, в частности упоминавшаяся выше иерархическая организация групп данных; наибольшей единицей является том данных. Он делится на массивы; массивы состоят из однотипных записей документов. Эти записи могут объединяться в пределах массива в блоки, размеры которых определяются исходя из условий переписи данных.

Массивы информации, расположенные на магнитных дисках или лентах и организованные одним из стандартных способов, являются одним из примеров файла, так как они имеют однотипную структуру записей, одно назначение и общее наименование.

Для работы с информационными массивами, размещаемыми на магнитных лентах или дисках, предусматривается единообразная методика (наличие заголовков, служебных записей и меток, стандартные процедуры обращения к массивам, их переписи, корректировки, контроля и т. д.). Некоторые из этих стандартных приемов будут описаны при рассмотрении способов организации массивов на магнитных дисках.

Особенности обработки информации с использованием ЗУ на магнитных дисках. Запоминающее устройство на магнитных дисках является характерным типом ЗУ для ЭВМ третьего поколения. Это ЗУ сочетает большую емкость и высокую скорость обращения. Оно называется ЗУ с прямым доступом, так как время обращения к различным позициям информации в этом ЗУ практиче-

ски не зависит от места их расположения на физическом носителе.

Типовым представителем является ЗУ на МД со следующими характеристиками. Алюминиевые диски, покрытые ферромагнитным материалом, объединяются в пакеты по 6 дисков. Расстояние между дисками в пакете 1 дюйм (2,54 см); диаметр диска 35,56 см; масса пакета 5,6 кг. Для хранения информации используются 10 поверхностей пакета. Верхняя поверхность верхнего диска и нижняя поверхность нижнего диска не используются. Поверхности перенумерованы сверху вниз от 0 до 9. Каждая поверхность имеет 203 концентрические дорожки, которые нумеруются от 0 (наружная дорожка) до 202 (внутренняя дорожка). Все дорожки независимо от их фактической длины одинаковой информационной емкости — по 3625 байтов. Таким образом, емкость всего пакета составляет 7,25 млн. байтов. На нижнем диске пакета имеется отверстие, представляющее собой отметку начала отсчета информации на дорожках и называемое маркером начала оборота.

Запись и чтение информации на МД осуществляется с помощью 10 магнитных головок, расположенных на 5 держателях. Эти держатели могут перемещаться в радиальном направлении, при этом головка устанавливается под любой из 203 дорожек. Дорожки с одинаковыми номерами, расположенные на разных дисках (10 дорожек), образуют как бы цилиндр. Каждый цилиндр просматривается при одном положении держателей головок, и это обстоятельство учитывается при выборе способов расположения информации на магнитных дисках.

Для перемещения головок от одного цилиндра к другому требуется около 85 мс. Головки не касаются дисков, а плавают над ними на воздушной подушке. На случай выхода некоторых дорожек из строя предусмотрены три резервные дорожки, расположенные внутри, т. е. ближе к оси вращения дисков. Устройство управления работой ЗУ на МД обеспечивает выполнение следующих основных операций:

1. «Установка» — для подвода головки к нужному цилиндру.
2. «Поиск» — для сравнения заданного поля данных в оперативной памяти с полями данных просматриваемой дорожки.
3. «Чтение» — для считывания данных.

4. «Запись» — для записи данных.
5. «Управление» — для стирания данных, установки в начальное положение головок, разрешения и запрета обращения и др.

К одному устройству управления может подключаться до 16 окончательных устройств на МД, но практически подключаются 4 таких устройства. Пакеты дисков являются сменными и могут храниться отдельно, образуя машинный архив на МД.

Расположение данных на МД. На всех дорожках МД данные располагаются однотипным способом, обусловленным наличием стандартных программ обращения к МД. В типовом случае информация на МД размещается в виде записей одинакового размера и структуры. Каждая запись относится к какому-то одному объекту или явлению. В составе записи выделяется обычно один какой-то признак (свойство) объекта, по которому производится его поиск. Этот признак принято называть ключом. Существуют три основных типа файлов на МД: стандартный последовательный, прямой и индексно-последовательный.

Стандартный последовательный файл строится, как правило, без ключей. Он может состоять из записей фиксированной или произвольной длины. Этот файл используется как быстрая магнитная лента, т. е. обращение к записям производится путем последовательного просмотра всего файла. Допускается объединение записей в блоки, т. е. в группы записей, переписываемых за одну операцию обмена.

Файл с прямым доступом строится как с ключами, так и без них. При помещении новой записи нужно указать только адрес дорожки, так как адрес свободного места на дорожке будет определен автоматически самой ЭВМ.

При чтении записи нужно указать как адрес дорожки (цилиндра и головки), так и номер записи (т. е. адрес на дорожке). Если используются ключи, то поиск записи в пределах дорожки может производиться по ключу.

При прямом способе разные записи могут иметь одинаковые ключи, если эти записи находятся на разных дорожках.

Для адресации записи принято следующее распределение байтов: номер цилиндра — 2 байта; номер головки — 2 байта; номер записи — 1 байт.

В случае применения индексно-последовательного файла использование ключей обязательно. Записи могут быть как сгруппированными в блоки, так и несгруппированными. Обращение к записям может идти по порядку изменения ключей (например, по возрастанию или по алфавиту и др.), хотя сами записи могут при этом располагаться на МД в произвольном порядке. Можно обратиться сразу к нужной записи, зная ее ключ.

Рассмотрим принцип построения индексно-последовательного файла на примере (табл. 2)

Пусть имеем 40 записей с длиной поля данных в 1000 байтов и длиной поля ключа в 2 байта. Ключом будет служить целое положительное двоичное число, имеющее 16 разрядов. Всего файл в таком составе содержит 40 500 байтов. Разместим записи в двух цилиндрах с номерами 5 и 6. На каждой дорожке будет умещаться по три записи; таким образом, в цилиндре может поместиться 30 записей, но мы будем помещать в 5-м цилиндре 24 записи и в 6-м цилиндре 16 записей.

На нулевой дорожке 5-го цилиндра вначале поместим записи так называемых индексов. Для надежности записи индексов дублируются. Запись индекса содержит сам индекс — значение ключа и адрес отсылки, т. е. адрес начала той последовательности записей, в которой этот ключ является максимальным. Так как файл занимает два цилиндра, то сначала будут идти четыре записи индексов цилиндров (два индекса по два раза). Эти индексы займут записи с номерами R_1 — R_4 . Запись R_5 содержит в поле ключа все единицы и означает конец записей индекса цилиндров. Затем пойдут (также попарно) записи индексов дорожек для 5-го цилиндра; всего их будет 16 (для дорожек с номерами $0 \div 7$).

Дорожки с номерами 8 и 9 резервируются на случай переполнения дорожек $0 \div 7$; 22-я запись — запись конца индексов дорожек, а 5-я запись — запись конца индексов цилиндров. С 23-й записи нулевой дорожки начнутся информационные записи файла.

В индексе цилиндров адреса отсылок указывают начало индекса дорожек соответствующего цилиндра, а в индексе дорожек адреса отсылок — начало соответствующих записей на соответствующей дорожке.

Номера записей определяют при считывании путем счета адресных маркеров. Записи индексов 12-байтовые: 2 байта — ключ, 5 байтов — адрес (номер цилиндра, но-

Принцип построения индексо-последовательного файла

	R ₆	Индекс цилиндров			Запись конца индекс-сов цилиндров			Индекс дорожек			Запись конца индекс-сов дорожек			Запись с данными		
		R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	R ₇	R ₈	R ₉	R ₁₀	R ₁₁	R ₁₂	R ₁₃	R ₁₄	R ₁₅
0		255	255	397	397	11...1	35	35	...	255	255	11...1	17	22	35	
5-й цилиндр Дорожка	Поле данных Дорожка	5, 0, 6	5, 0, 6	6, 0, 1	6, 0, 1		5, 0, 23	5, 0, 23	5, 7, 1	5, 7, 1	5, 7, 1					
			R ₁	(45)				R ₂	(50)				R ₃	(57)		
			R ₁	(61)				R ₂	(70)				R ₃	(75)		
			R ₁	(82)				R ₂	(87)				R ₃	(90)		
			R ₁	(93)				R ₂	(99)				R ₃	(105)		
			R ₁	(111)				R ₂	(130)				R ₃	(150)		
			R ₁	(161)				R ₂	(178)				R ₃	(190)		
	R ₁	(200)				R ₂	(240)				R ₃	(255)				

	R ₀	Индекс дорожек				Запись конца индекса дорожек		Запись с данными				
		R ₁	R ₂	R ₃	R ₄	R ₁₁	R ₁₂	R ₁₄	R ₁₄	R ₁₅	R ₁₅	
0		275	275	300	300	397	397	11...1	260	270	275
Поле ключа		6, 0, 14	6, 0, 14	6, 1, 1	6, 1, 1		6, 5, 1	6, 5, 1				
Поле данных												
Дорожка												
1		R ₁	(280)				R ₂	(290)			R ₃	(300)
2		R ₁	(305)				R ₂	(340)			R ₃	(342)
3		R ₁	(348)				R ₂	(353)			R ₃	(370)
4		R ₁	(380)				R ₂	(385)			R ₃	(390)
5		R ₁	(397)				R ₂	(EOF)			R ₃	(390)
								Запись конца файла				

6-й цилиндр
Дорожка

мер дорожки, номер записи), 5 байтов — информация о том, что эта запись является индексом. Ниже схематично показано размещение записей на дорожках 5-го и 6-го цилиндров. Более подробно показаны нулевые дорожки. Выражение вида $R_1(k)$ условно означает i -запись на дорожке, имеющую ключ k . Например, на 5-м цилиндре на 0 дорожке под символом записи R_1 написано: 255, 5, 0, 6. Это означает, что запись R_1 является записью индекса 5-го цилиндра, имеющего наибольшее значение ключа 255; записи индексов дорожек для этого цилиндра начинаются с записи R_6 , расположенной в 5-м цилиндре, на 0 дорожке; порядковый номер записи 6. Аналогично под символом записи R_5 указано максимальное значение ключа для 6-го цилиндра (397) и указан адрес начала записей индексов дорожек для этого цилиндра (6, 0, 1). Запись R_6 указывает, что наибольшее значение ключа для записей на 0 дорожке 5-го цилиндра равно 35, а эти записи начинаются с 23-й записи на этой дорожке.

При включении новых записей, если дорожка заполнена, происходит перенос записи на резервную дорожку (8) с указанием об этой в записи R_0 данной дорожки и происходит корректировка индексов дорожек (и цилиндров).

При поиске записи с заданным значением ключа сначала просматривается индекс цилиндров и находится цилиндр, в котором находится нужная запись. После этого происходит переход к данному цилиндру и на нем просматривается индекс дорожек, по которому определяется дорожка, содержащая искомую запись. Эта дорожка просматривается сплошь и на ней находится нужная запись.

В данном примере сами информационные записи упорядочены на дорожках по возрастанию ключа. Это упрощает построение индексов, но требует упорядоченного размещения записей и их передвижек при пополнении файла. Можно ввести еще один уровень индексов — индекс записей, в котором указать ключи записей и фактические адреса записей на дорожках, при этом сами записи могут располагаться в произвольном порядке и их не требуется перемещать при вводе и исключении ненужных записей.

Приведенный пример конкретного расположения записей на дорожках магнитных дисков служит только

для пояснения существа такого распространенного способа размещения массивов записей в памяти ЭВМ, каким является индексно-последовательный способ.

Практически пользователям ЭВМ и даже программистам, составляющим программы поиска данных индексно-последовательным способом, не приходится иметь дело с размещением и адресацией записей на МД. Они задают только исходные данные, определяющие наименование массива, значения ключей (признаков), размеры записей и их строение, число записей. Все необходимые действия по формированию файла на МД, а также по поиску и выдаче найденных записей осуществляет операционная система. Для задания указанных исходных данных, для формирования файла и для поиска записей в языках программирования (ПЛ—1, язык Ассемблера и др.) существуют специальные обобщенные операторы, применение которых не требует детального знания технической стороны реализации указанных процессов. Однако знакомство с принципами построения индексно-последовательного файла оказывается полезным даже лицам, не занимающимся непосредственно программированием, а осуществляющим лишь формулировку исходных условий для решения и постановку задач, связанных с поиском данных в больших массивах информации.

Такое знакомство позволяет оценить возможность использования данного способа в конкретных условиях, иногда «подогнать» задачу к этому способу.

Следует заметить, что основные черты этого типового способа (многоуровневая система индексов-ключей, использование ссылок — адресов для переходов между уровнями, расположение индексов в возрастающем порядке в пределах каждого уровня) являются характерными для различных модификаций этого способа, а также для многих специальных способов организации файлов на магнитных дисках.

Пакеты прикладных (пользовательских) программ. При использовании ЭВМ третьего поколения большое значение имеет применение пакетов прикладных программ, которые, как уже упоминалось, представляют собой полностью отработанное программное обеспечение определенных классов задач и рассчитанное на потребителей, не являющихся специалистами по системному программированию. Ниже приводятся краткие сведения

по четырем таким пакетам, разработанным научно-производственным объединением «Центрпрограммсистем»¹.

Банк данных универсальной структуры представляет собой пакет прикладных программ для ЕС ЭВМ, рассчитанный на работу с дисковой операционной системой ДОС ЕС при составе оборудования модели ЕС 1020 и выше. Этот пакет позволяет создавать и корректировать до пяти файлов информации на МД (добавлять и исключать записи) и производить в них поиск.

В качестве физической единицы данных в файлах служит страница — блок фиксированной длины, состоящий из записей. Максимальный размер записи не должен превышать длины дорожки. Предусматривается возможность иметь до 255 разных типов записей.

Описываемый банк данных строится по так называемому ассоциативно-адресному узловому способу организации поисковых массивов, сущность которого описывается в разделе 2 главы № IV). Каждая запись может входить в произвольное число организуемых по этому способу цепных списков, причем структура записей, состав поисковых признаков и цепных списков, состав перекрестных связей между массивами определяются и задаются пользователем в зависимости от конкретных условий применения данного пакета программ.

Обращения к записям могут осуществляться либо поиском по заданному ключу, либо прямым способом по адресу записи, либо поиском по цепному списку.

В состав указанного пакета прикладных программ входят следующие части: набор макроопределений; набор процедурных модулей; программа формирования таблиц констант, описывающих структуру базы данных; резидентная программа, обеспечивающая все связи между данной системой и программой пользователя; обслуживающие программы, обеспечивающие учет работы системы (время работы, число обращений и т. д.) и ведение системного журнала, в котором отражаются все изменения, вносимые в банк данных.

¹ Помимо описанных, известны также другие пакеты прикладных программ, как отечественные, так и зарубежные, предназначенные для реализации на ЭВМ информационных поисковых систем и банков данных. Среди отечественных следует назвать пакеты АСПИД, ПЕГАС, БАНК; из зарубежных — пакеты AIDOS (ГДР), УСС (ЧССР), DPS, STAIRS, IMS—2 (США). Много пакетов находится в разработке (прим. ред.).

На основе указанного журнала и копии файлов специальная восстанавливающая программа может восстановить данные ее искажении или сбоях аппаратуры.

К обслуживающим программам относится также программа подготовки МД, служащая для формирования (разметки) страниц на МД в областях, предназначенных для размещения базы данных. Имеется также специальная программа реорганизации банка данных, предназначенная для включения новых записей, организации цепных списков, исключения записи.

Набор макроопределений служит для описания структуры базы данных и используется при подготовке программы пользователя к работе с базой данных.

Процедурные модули могут настраиваться на конкретные структуры базы данных и обеспечивают возможность пользователю обращаться к банку данных на логическом уровне, т. е. без непосредственного программирования операций обработки цепных списков, вычисления адресов записей и т. д. Предусматривается возможность включения процедурных модулей в программу пользователя с использованием языка Ассемблера.

Процедурные модули служат для реализации следующих процедур: включения записей в файлы в соответствии с их типом и описанием; исключения записей из файлов с исключением их из всех цепных списков, в которые они входят, и удаления всех записей из цепных списков, отходящих от исключаемой записи; поиска и выдачи записей по их признакам; корректировки значений полей данных, образующих записи.

Для работы с данным пакетом прикладных программ требуется предварительное описание конкретной структуры базы данных пользователя, задание в рабочей области памяти буферов, полей и ячеек связей, необходимых для работы пакета, и включение в программу пользователя оператора банка данных.

Программа, составленная на исходном языке, проходит этапы трансляции и редактирования, причем возможны отдельная трансляция макрокоманд описания базы данных и использование полученного объектного модуля другими программами. Применение пакета программ универсального банка данных упрощает процесс программирования задач обработки больших массивов информации, приводит к унификации подобных программ у различных пользователей, повышает четкость

Ведения массивов. Этот пакет обеспечивает достаточно высокую скорость поиска и обновления файлов, экономию объема памяти на МД (за счет устранения дублирования данных). Важной является также защита от случайных искажений в базе данных, так как обращение к ней осуществляется через процедурные модули.

Банк данных АСУП (общим объемом 25 000 машинных команд)— это пакет прикладных программ для создания, корректировки, поиска и обработки массивов информации, представленных в виде четырех файлов (предметный файл, файл состава изделий, файл пооперационных трудовых нормативов, файл рабочих мест).

Между файлами имеются адресные ссылки, обеспечивающие возможность построения древовидных поисковых структур, объединяющих данные об изделиях, их структуре, технологии изготовления, рабочих местах. Указанные файлы строятся на магнитных дисках, причем главный предметный файл строится по индексно-последовательному способу с управлением, а остальные файлы (связывающие)— по прямому способу. Адреса их записей находятся либо в главном, либо в других связывающих файлах. Поиск информации в файлах осуществляется с помощью программ пользователя, написанных на языках Ассемблера, КОБОЛ или ПЛ/1. Структуру любого изделия записывают в банк данных один раз независимо от того, сколько раз данное изделие входит как составная часть в другие изделия. Состав и содержание записей определяются пользователями и задаются с помощью описаний при генерации системы. Указанный банк данных работает под управлением ДОС ЕС и рассчитан на 64К, т. е. 64 тыс. байтов оперативной памяти. Количество МД определяется в зависимости от объема файлов. Банк данных АСУП ориентирован на решение задач поиска справочно-нормативной информации, технической подготовки производства, оперативного управления и т. д.

Генератор программ совместной обработки массивов (ГЕСОМ) осуществляет выработку программ на языке Ассемблера, обеспечивающих ввод информации с МЛ и МД; вывод на МЛ, МД; организацию последовательной обработки массивов (причем один из массивов является ведущим); логический и синтаксический контроль исходной программы; подключение нестандартных блоков обработки.

Исходными данными для работы генератора ГЕСОМ являются описание входного файла, выходных файлов и подключаемых блоков программ пользователя. Сгенерированная рабочая программа должна быть оттранслирована и отредактирована. При работе рабочей программы выдаются на АЦПУ сообщения об ошибках.

Генератор программ ввода — вывода информации из документов сложной структуры, записанных на машинных носителях ЕС ЭВМ (перфоленга, перфокарты, магнитные ленты, магнитные диски). В качестве выхода строятся файлы на МЛ, ПК и МД, причем на МД строятся выходные файлы с прямой и последовательной организацией.

Осуществляются логический контроль вводимой информации, исключение или корректировка ошибочных входных данных.

Вопросы организации обработки медицинской информации. Важным вопросом прикладной медицинской кибернетики является внедрение в практику различных методов решения медицинских задач и обработки информации. При этом необходимо обеспечить переход от старых (неавтоматических) методов к новым методам, что требует проведения ряда организационно-технических мероприятий и психологической подготовки медицинских работников, связанных с этими задачами.

Внедрение представляет собой процесс постепенного перехода от существующей системы к автоматизированной системе решения задач и частично проводится параллельно с разработкой проектной документации.

Основными этапами внедрения отдельных задач (подсистем) являются подготовка управленческого аппарата и специалистов медицинских учреждений к внедрению новых задач и опытная эксплуатация подсистем, заканчивающаяся сдачей их в рабочую эксплуатацию.

Порядок внедрения отдельных подсистем в опытную эксплуатацию и сдачи их в эксплуатацию в общих чертах может быть следующим.

К началу опытной эксплуатации подсистем должны быть: сданы в эксплуатацию технические средства, установленные в медицинских учреждениях или в Главном вычислительном центре (ГВЦ) системы; переданы специалистам медицинских учреждений инструкции по заполнению и перфорации нормативно-справочной информации (НСИ); закончены работы по организации систе-

мы ведения НСИ и ГВЦ (запись информации на машинных носителях, проверка и внесение изменений); обучены специалисты медицинских учреждений выполнению функций, связанных с подготовкой информации и решением задач: подготовлен контрольный пример, который может характеризовать правильность функционирования сдаваемой подсистемы (задачи).

Сдаваемые программы должны соответствовать техническому заданию и включать следующие документы: описание алгоритмов и программ; инструкцию по перфорации входных документов; инструкцию по эксплуатации программ; контрольный пример и инструкции по эксплуатации программ этого примера. Описание алгоритмов и программ состоит из введения, общего описания процесса решения задачи, описания рабочих программ.

Во введении указываются тип ЭВМ и необходимый состав внешних устройств, особенности организации и ведения массивов информации, применяемые средства внутреннего математического обеспечения (языки, трансляторы, библиотеки, операционные системы и т. д.), исполнители и перечень выполненных ими работ.

В общем описании процесса решения задачи приводятся математическая постановка задачи, принципиальная блок-схема процесса решения (с указанием программ, наименований и шифров массивов и их структур) и особенности технологии решения задачи на ЭВМ; описание машинного алгоритма (в соответствии с блок-схемой) или метода решения задачи.

В описании рабочих программ содержится принципиальная блок-схема программы; алгоритм в описательной форме, описание входной информации и результатов работы программы (или ссылка на них); распределение оперативного накопителя и внешних накопителей; состав устройств ввода и вывода; перечень используемых стандартных процедур с ссылкой на источники; особенности программ.

В инструкции по перфорации входных документов приводятся общие требования к методам перфорации (принятая система перфорации, методы исправления ошибок перфорации, маркировка машинных носителей после перфорации), пример заполнения документа (для особых случаев) и при необходимости пример перфорации, ссылка на принятый метод перфорации, объем (в

словах, строках, документах) входной информации, включаемой в одну пачку документов, схема (макет) перфорации и необходимые к ней пояснения, формы входных документов (приложение).

Инструкция по эксплуатации программ включает инструкцию по подготовке программ к эксплуатации и технологии их обслуживания и собственно инструкцию по эксплуатации программ. В инструкции по подготовке программ к эксплуатации и технологии их обслуживания проводятся перечень и объем внешних накопителей для массивов и программ, правила организации массивов и программ на внешних накопителях, их дублирования, корректировки, восстановления и тестирования (или ссылка на них, если они предусматриваются внутренним математическим обеспечением). Инструкция по эксплуатации программ включает распределение используемых внешних устройств и варианты их применения (замены); действия оператора (в том числе вызов программ в оперативную память, правила их запуска); режимы работы и правила их задания; вид справочной информации, получаемой в процессе работы программы, и необходимые к ней пояснения.

Для решения контрольного примера исходная информация и специальные программы при необходимости предъявляются на машинном носителе; для задач с запросом в качестве примера приводится контрольный запрос.

Инструкция по эксплуатации программ контрольного примера (запроса) содержит описание специальных программ (при их наличии), инструкцию по организации исходной информации на внешних носителях, инструкцию по эксплуатации программ (или ссылку на нее).

Программы, которые должны быть отперфорированы на входных носителях информации (перфокартах, перфолентах), предъявляются в двух экземплярах.

Для сдачи подсистемы в опытную эксплуатацию обычно назначается комиссия с указанием сроков приемки конкретной подсистемы.

Цель опытной эксплуатации подсистемы — комплексная проверка всех ее элементов, проверка подготовленности информационной базы, отладка технологического процесса сбора и обработки информации, проверка готовности персонала ГВЦ (ВЦ) и работников медицинских учреждений к работе в условиях функционирования подсистемы.

ПОНЯТИЕ О МАТЕМАТИЧЕСКИХ МЕТОДАХ, ИСПОЛЬЗУЕМЫХ В МЕДИЦИНСКОЙ КИБЕРНЕТИКЕ¹

1. Системный анализ и исследование операций

Системный анализ — это комплексный научный подход к изучению сложных динамических систем в биологии, медицине, технике, экономике, в социальных науках.

Под динамической системой понимают совокупность взаимодействующих объектов, обладающих способностью воспринимать и обмениваться информацией и участвующих в общем целенаправленном процессе. Понятие динамической системы по определению является рекурсивным (многоступенчатым). Объекты, являющиеся элементами данной динамической системы, при более детальном рассмотрении сами могут оказаться целыми динамическими системами, состоящими из объектов более мелкого масштаба, и т. д.

Системный анализ охватывает широкий круг явлений (не только деятельность коллективов людей, использующих технику) и имеет целью выяснение закономерностей этих явлений как для лучшего использования существующих систем, так и для проектирования новых систем.

Исследование операций — научное направление, суть которого сводится к разработке методов оптимального руководства деятельностью коллективов людей, использующих технику, при выполнении сложных организационных, хозяйственных, военных и других мероприятий.

Ясно, что выводы, полученные в процессе исследования операций, могут быть использованы и для лучшей организации новых коллективов людей и разработки новых или совершенствования существующих технических средств.

Системный анализ и исследование операций — два взаимосвязанных научных направления. Различие между ними состоит в том, что в системном анализе упор дела-

¹ В данной главе не рассматриваются такие традиционные математические методы, используемые в медицине и здравоохранении, как методы теории вероятностей и математической статистики.

ется на методологию изучения сложных структур, на раскрытие внутренних связей и закономерностей явлений, в том числе явлений живой природы, а в исследовании операций — на анализ организационных форм деятельности коллективов людей и выработку оптимальных решений по руководству этой деятельностью. В медицине и здравоохранении все более широкое применение получают и системный анализ, и исследование операций. Типичными областями применения указанных научных методов являются задачи комплексного изучения состояния здоровья человека в норме и патологии, разработка систем автоматизации управления медицинскими учреждениями, оценка эффективности различных диагностических и лечебных методов и средств, изучение вопросов охраны внешней среды и разработка мероприятий по ее защите от загрязнения и т. д.

При решении всех этих задач необходим всесторонний учет как внутренних закономерностей изучаемого явления или процесса, так и внешних связей и зависимостей данного явления от окружающей среды. Характерными для указанных задач являются также необходимость экономических оценок и анализ возможностей реального осуществления тех или иных мероприятий. Например, при решении вопроса о приобретении дополнительного диагностического оборудования нужно оценить расходы на его покупку, ожидаемую степень его загрузки, необходимые размеры помещений и численность обслуживающего персонала, а также влияние, которое окажет наличие этой аппаратуры на поток и структуру больных; средние сроки их обследования и лечения, возможности комплексного использования нового оборудования совместно со старым.

Характерной особенностью организации работ по системному анализу и исследованию операций является включение в состав рабочих групп представителей разных специальностей, способных оценить различные стороны изучаемого явления, — врачей, математиков, инженеров, экономистов, юристов и др. Перечисленные представители для успешной работы в составе групп исследования операций должны обладать общим языком, знать методологию исследования операций и в то же время иметь высокую профессиональную подготовку. Важным моментом в организации работы подобных групп является также установление правильных взаимо-

отношений между группой исследования операций и ответственным руководителем учреждения, имеющим право принимать окончательные решения.

Группа исследования операций должна представлять руководителю объективные данные, обосновывающие оптимальное решение, но не должна навязывать это решение. Руководитель, ставя задачу перед группой исследования операций, должен четко и объективно формулировать исходные условия и цели исследования, при этом он не должен предрешать желаемого варианта решения и ориентировать группу на обоснование такого решения.

В работе групп по исследованию операций выделяется ряд типовых этапов:

1. Постановка задачи и сбор исходных данных.
2. Выяснение взаимосвязей между различными величинами, т. е. построение модели исследуемого явления.
3. Выбор метода решения и выполнение необходимых расчетов вариантов.
4. Анализ полученных вариантов и выбор оптимального варианта.
5. Представление решения руководителю, обеспечение мероприятий по доведению решения до исполнителей, контроль за правильностью понимания этого решения заинтересованными лицами.

На практике при реализации описанного порядка исследований часто встречаются отклонения, возвраты к предшествующим этапам для уточнения или изменения построенной модели, изменения исходных данных или метода решения, выбора нового критерия оценки решения и т. д.

Таким образом, в исследовании операций можно выделить две основные стороны: 1) уяснение задачи и сбор необходимых исходных данных; 2) определение оптимального решения и доведение его до заинтересованных лиц.

Следует заметить, что зачастую выполнение только первой части работы (уяснение постановки задачи, сбор и систематизация относящейся к задаче информации) уже позволяет получить полезные результаты, т. е. дает возможность руководителю принять целесообразное решение.

На основе систематизированной информации, представленной в компактной наглядной форме (например,

в виде графиков или диаграмм на экране дисплея), опытный руководитель или специалист может оценить тенденции процесса и интуитивно правильно выбрать подходящий вариант решения. Поэтому не преуменьшая роли количественных математических методов нахождения оптимальных решений, следует не допускать недооценки роли и возможностей опыта и интеллекта человека и стремиться к оптимальному сочетанию возможностей ЭВМ и человека при решении сложных задач исследования операций.

ЭВМ, снабженная дисплеем, может быстро и точно обработать огромный объем первичной информации и представить его в удобном для анализа виде.

Человек способен быстро оценить общую обстановку и наметить стратегию поиска оптимальных решений, указав ЭВМ методы и исходные данные для расчетов. Результаты расчетов, т. е. различные варианты решений, могут оцениваться и отбираться самой ЭВМ, если ей заданы критерии отбора, либо этот процесс возьмет на себя человек, когда для выбора вариантов еще нет полностью формализованных критериев. Выбор оптимального решения должен производиться из какого-то множества возможных решений, поэтому сначала нужно определить это множество либо простым перечислением всех решений (если это практически возможно), либо указанием способа получения этих решений. Оптимальное решение выбирается всегда для достижения какой-то цели, и поэтому необходимо четко сформулировать в количественной (математической) форме критерий оценки выбранного решения и показать, что оптимальное решение обеспечивает достижение максимального (или минимального) значения этого критерия. Такой количественный критерий при решении задач исследования операций принято называть целевой функцией. Таким образом, в исследовании операций большую роль играют математические методы нахождения оптимальных решений или, как принято говорить, методы решения экстремальных задач. Но исследование операций не сводится только к применению этих методов. Помимо них, в исследовании операций применяются методы моделирования, методы натуральных испытаний, методы эвристических решений, методы ситуационных игр. В задачах исследования операций зачастую бывает важно не только найти оптимальный вариант распределения ма-

териальных, денежных, людских и других ресурсов между участниками работы, но и определить необходимый (или оптимальный) порядок выполнения различных действий или работ.

Для решения первой задачи применяются методы поиска решений экстремальных задач (методы математического программирования). Для решения второй задачи используются методы теории расписаний и методы сетевого планирования и управления (СПУ).

Сущность перечисленных методов нами будет рассмотрена ниже. Системный анализ и исследование операций относятся к прикладной кибернетике и, естественно, используются как в медицине и здравоохранении, так и в других областях человеческой деятельности. Эти научные направления являются новыми и развивающимися. В них еще нет иногда четко установившихся терминов и определений. Допускаются различные толкования и самих названий этих научных направлений: прикладная кибернетика, системный анализ, исследование операций. Несмотря на то что эти три направления тесно связаны между собой, между ними имеются различия.

Прикладная кибернетика занимается применением общих идей и методов кибернетики к решению практических задач и, таким образом, представляет собой теоретическую базу для системного анализа и исследования операций.

2. МАТЕМАТИЧЕСКИЕ МЕТОДЫ нахождения оптимальных решений

Как уже говорилось раньше, оптимальным решением принято называть одно из возможных решений заданной проблемы, которое наилучшим образом позволяет достичь цель. Оценка того, в какой мере то или иное решение является оптимальным, может производиться человеком либо интуитивно, на основе своего опыта, либо с помощью расчетов конкретных показателей (суммы расходов, сроков выполнения работ, объема полученных результатов и т. д.). Если проблема допускает математическое описание исходных условий и критерия качества решения, то для нахождения оптимального решения могут быть использованы математические методы. Обычно применение этих методов связано с большим объемом вычисления и их практическая реализация требует

использования ЭВМ. В зависимости от особенностей математического описания (постановки) задачи применяются различные математические методы; каждый из методов служит для решения задач определенного класса. Обычно уже при формулировке задачи и построении ее математического описания ориентируются на использование определенного метода. При выборе метода учитываются также требования по желаемой точности нахождения оптимального решения и по ожидаемой точности получения исходных данных. Часто первое приближенное решение определяет интуитивно человек, а уточнение этого решения (его оптимизация) производится с помощью математических методов. Бывает и наоборот, что, применяя приближенный математический метод, определяют ориентировочный вариант «оптимального» решения, окончательный же выбор решения производит человек.

Известно, что в классической математике существуют методы решения экстремальных задач — методы нахождения максимумов или минимумов функций одной или нескольких переменных. Эти методы используются и при поиске оптимальных решений в автоматизированных системах управления в экономике, в здравоохранении и в других областях народного хозяйства. Однако эти методы имеют ряд принципиальных ограничений, препятствующих их широкому применению при поиске оптимальных решений в сложных системах управления. Во-первых, для применения указанных методов функции, выражающие связь между исходными условиями (аргументами) и результатами оптимизируемых процессов, должны обладать определенными свойствами — иметь первые и вторые производные, экстремальные значения функции не должны находиться на границе области изменения переменных, не допускается наличия неравенства среди ограничений, определяющих область изменения переменных. Во-вторых, практические возможности применения классических методов решения экстремальных задач резко убывают при увеличении числа переменных.

В связи с изложенным в течение последних четырех десятилетий создано много новых специальных методов нахождения оптимальных решений, которые в настоящее время принято объединять общим названием «математическое программирование». Характерными особенностями

ми методов математического программирования являются большое количество переменных, от которых зависит исследуемая функция, дискретный характер изменения переменных; наличие неравенств среди ограничений, определяющих область изменения переменных; обязательность неотрицательности переменных. Одним из первых методов математического программирования, предложенным в 1939 г. советским математиком Л. В. Канторовичем, был метод линейного программирования¹. Этот метод и в настоящее время является наиболее популярным методом нахождения оптимальных решений.

Следует подчеркнуть две стороны проблемы применения указанных математических методов в медицине и здравоохранении:

- 1) постановку задач для решения этими методами;
- 2) выбор метода решения и его реализацию.

Если вторая сторона дела является делом математиков и программистов, то с первой стороной, т. е. с методикой постановки задач, должны быть знакомы врачи, организаторы здравоохранения, биологи, экономисты и другие специалисты, которые собираются применять в своих конкретных областях ЭВМ и математические методы.

В связи с этим в данной книге мы не затрагиваем вопросов построения методов решения, а постараемся на нескольких примерах пояснить сущность постановок задач для решения указанными методами.

В качестве примера постановки задачи линейного программирования рассмотрим задачу о выборе рационального (оптимального) варианта оборудования некоторой лаборатории, обеспечивающего заданное количество различных процедур в заданное время при минимальной стоимости оборудования (приводим один из возможных вариантов постановки этой задачи).

¹ Термины «линейное программирование» и «математическое программирование», как и упоминаемые в дальнейшем термины «динамическое программирование» и «стохастическое программирование», не следует смешивать с программированием для ЭВМ. Термином «линейное программирование» (введенным еще до появления ЭВМ) обозначается метод решения некоторого класса задач, взятых первоначально из области экономического планирования. В настоящее время этот термин, как и упомянутые выше термины, получил большое распространение, так что замена его другим термином, например термином «линейное планирование», нецелесообразно.

Пусть нам задано n видов приборов Π_i , $i=1, 2, \dots, n$, которые можно комбинировать, составляя оборудование лаборатории. Каждый из этих видов приборов может выполнять определенный набор процедур из общего числа m различных процедур Q_j , $j=1, 2, \dots, m$, необходимых для функционирования лаборатории. Наборы процедур, выполняемых отдельными приборами, задаются матрицей a_{ij} , где a_{ij} означает производительность прибора, т. е. количество процедур вида Q_j , выполняемых в день прибором вида Π_i .

Задается дневная норма, т. е. необходимое количество процедур каждого вида, которые должны выполняться лабораторией в день N_j .

Задается стоимость каждого вида приборов:

$$c_i, \quad i=1, 2, \dots, n$$

Задаются наличные запасы каждого вида приборов F_i (максимально допустимая заявка на каждый вид приборов).

Задаются максимально допустимые количества приборов данного вида в лаборатории S_i (например, крупногабаритные приборы могут занимать слишком много места или приборы могут требовать трудоемкого обслуживания и т. д.).

Требуется найти такой набор приборов в лаборатории $F(x_i)$, который бы выполнял необходимые количества процедур каждого вида, имел бы минимальную стоимость:

$$F = \sum_{i=1}^n c_i x_i$$

и удовлетворял бы перечисленным ограничениям:

$$\sum_{i=1}^n a_{ij} x_i \geq N_j; \quad S_i \geq x_i \geq 0; \quad x_i \leq p_i$$

$$i=1, 2, \dots, n; \quad j=1, 2, \dots, m.$$

При этом накладывается дополнительное условие, чтобы величины x_i принимали целочисленные значения, так как приборы по физическому смыслу должны использоваться целыми единицами. Можно решать эту задачу и не накладывая дополнительного условия на целочисленность x_i . При получении в качестве ответа дробных значений x_i производится округление этих значений,

исходя из общих интуитивных соображений. Очевидно, что такой путь позволяет получить приближенное решение, что на практике зачастую оказывается вполне достаточным.

Таким образом, мы выразили математически поставленную задачу. После того как будут заданы конкретные значения всех величин, входящих в приведенные выражения, методами линейного программирования можно получить искомое решение задачи.

Для решения задач линейного программирования существует общий метод, кроме того, для многих частных случаев задач линейного программирования имеются специальные эффективные методы решения.

Таким образом, линейное программирование позволяет находить оптимальные варианты решений в тех случаях, когда целевая функция и все ограничения являются линейными.

Второй пример постановки задачи линейного программирования относится к области лучевой терапии. Рассмотрим один из возможных методов формирования оптимального плана облучения злокачественной опухоли.

Ограничимся для простоты расчетом в одной плоскости при постоянной диафрагме источника энергии (т. е. постоянным размером и формой отверстия излучателя). Будем считать, что источником энергии является гамма-терапевтический аппарат. Соответствующий расчет дозы при других источниках энергии (нейтронная или электронная энергия) принципиальных отличий не имеет, увеличиваются только технические трудности расчета.

Облучение опухоли может производиться с различных направлений и расстояний (из различных точек) и различное время (экспозиция t) из каждой точки. Необходимо определить оптимальный план облучения, т. е. время облучения из каждой точки.

При расчете оптимального плана облучения должны учитываться следующие физические ограничения.

Для здоровых тканей (обозначим их через A_T) доза облучения не должна превосходить предельно допустимой величины Q_T (предельно допустимой дозой называется доза, которая еще не влечет за собой необратимых вредных последствий для ткани). Для тканей злокачественной опухоли (обозначим их через F) доза облучения должна быть не меньше значения Q_F , необходимого для поражения опухоли. При этом общая интегральная доза

на весь организм должна быть минимальной. Это требование кладется в основу метода оптимизации.

При облучении варьируются следующие параметры: 1) расстояние источник — поверхность (РИП); 2) направления облучения; 3) время экспозиции.

Обозначим все допустимые направления облучения через M_1, M_2, \dots, M_n . Обозначим соответствующие РИП для направления M_i через $B_i^1, B_i^2, \dots, B_i^{s_i}$. Обозначим через t_{ik} время экспозиции при облучении с направления M_i и РИП B_i^k . Приведем мысленно плоский срез тела человека в районе злокачественной опухоли приблизительно перпендикулярно (в основном) к большинству возможных направлений облучения. Разобьем данный срез на элементарные области ω_{jq} радиальными линиями, проходящими через центр опухоли, и концентрическими линиями с центром в центре опухоли. Таким образом весь срез будет разбит на элементарные площадки ω_{jq} , расположенные в плоскости среза.

Пусть теперь E_{jq}^{ik} — доза, полученная за единицу времени площадкой ω_{jq} при облучении опухоли с направления M_i и расстояния B_i^k .

Очевидно, что произведение $t_{ik} \cdot E_{jq}^{ik}$ будет представлять собой соответствующую дозу, полученную этой площадкой при экспозиции t_{ik} ($t_{ik} \geq 0$, равенство нулю t_{ik} означает, что в плане отсутствует облучение площадки ω_{jq} из точки B_i^k).

Задача определения плана облучения заключается в расчете матрицы (t_{ik}) . Нам необходимо получить оптимальный план, т. е. план, обеспечивающий минимальную суммарную дозу облучения при допустимых дозах облучения здоровых тканей и заданных дозах облучения больных тканей.

Подсчитаем дозу, полученную площадкой ω_{jq} при некотором плане (t_{ik}) облучения. Она будет равна:

$$D_{jq} = \sum_i \sum_k t_{ik} \cdot E_{jq}^{ik}.$$

Если площадка ω_{jq} принадлежит зоне здоровых тканей A_r , то должно выполняться условие:

$$D_{jq} \leq Q_r. \quad (I)$$

Если же площадка ω_{jq} принадлежит зоне злокачественной опухоли, то должно выполняться условие:

$$D_{jq} \geq Q_r. \quad (II)$$

Требуется определить матрицу (t_{ik}) такую, чтобы при соблюдении условий (I) и (II) обеспечивалось минимальное значение це-

левой функции, т. е. общей интегральной дозы облучения, полученной организмом.

$$R = \min \sum_j \sum_q D_{jq}. \quad (III)$$

Определив искомую матрицу (t_{ik}), мы получили оптимальный план лучевой терапии. Существенным в данной обстановке задачи является допущение о том, что ограничения (I) и (II) и целевая функция (III) являются линейными функциями времени t_{ik} , т. е. зависимости представляются многочленами первой степени относительно t_{ik} .

Описанная задача решается общим методом решения задач линейного программирования — так называемым симплекс-методом и, несмотря на ряд упрощающих предположений, имеет практическое применение. При использовании подобной методики обеспечивается уменьшение общей дозы облучения организма на 15—20% при сохранении того же терапевтического эффекта, что и при облучении оптимального плана.

После ознакомления с конкретными примерами рассмотрим общую постановку задач математического программирования.

Пусть требуется найти вектор¹ $X(x_1, x_2, \dots, x_n)$, обеспечивающий экстремум функции $F(X) = F(x_1, x_2, \dots, x_n)$, называемой целевой функцией, при наличии следующих ограничений на значения независимых переменных x_1, x_2, \dots, x_n .

$$q_1(x_1, x_2, \dots, x_n) = 0, \quad i = 1, 2, \dots, m_1$$

$$q_i(x_1, x_2, \dots, x_n) \leq 0, \quad i = m_1 + 1, \dots, m$$

$$x_j \geq 0, \quad j = 1, 2, \dots, n.$$

Функция $F(X)$ и ограничения $g_1(X)$ могут иметь различный характер, и в зависимости от этого меняется методика решения задачи и определяется тот или иной раздел математического программирования. В случае, если

¹ В математике величина (длина) некоторого вектора X , выходящего из начала координат, является функцией координат x_1, x_2, \dots, x_n его конца (т. е. проекции конца вектора на оси координат). В нашем случае искомая нами величина x (дневной рацион, план облучения и т. д.) тоже определяется некоторым набором величин x_1, x_2, \dots, x_n , которые обозначают соответствующие задаче параметры (калорийность продуктов питания, направление облучения и т. п.). Наличие этой аналогии позволяет в общем случае называть искомую величину вектором, а параметры, от которых она зависит, — координатами.

$F(X)$ и $g_1(X)$ являются линейными функциями переменных x_1, x_2, \dots, x_n , т. е. имеет место пропорциональная зависимость этих функций от аргументов в первой степени, то задача относится к линейному программированию. В рассмотренных выше примерах как раз имели место такого рода зависимости.

Если целевая функция $F(X)$ или некоторые из ограничений $g_1(X)$ нелинейны, то задача относится к так называемому нелинейному программированию.

Задачи нелинейного программирования, вообще говоря, решать значительно труднее, чем задачи линейного программирования, и поэтому путем упрощающих допущений их всегда стремятся свести к задачам линейного программирования, если это не приводит к существенной потере точности.

Наиболее разработан в нелинейном программировании случай, когда целевая функция является полиномом второй степени, а ограничения линейны. Для задач такого рода принято название квадратичное программирование.

Если среди параметров, определяемых условием задачи, имеются случайные величины, то задача относится к стохастическому программированию, которое применяется для выработки оптимальных планов, учитывающих вероятностные характеристики процессов.

Примером постановки задачи стохастического программирования в области здравоохранения может служить задача долгосрочного планирования развития сети лечебно-профилактических учреждений. При этом, естественно, имеет место некоторая неопределенность в исходных данных, определяющих потребность в различных видах медицинского обслуживания, а также возможные ресурсы и затраты. При практической реализации любого перспективного плана обязательно будут иметь место неувязки, для устранения которых необходимо заранее предусмотреть резервы.

В этом случае задача долгосрочного планирования решается методом стохастического программирования в два этапа, при этом рассчитываются как затраты на собственно выполнение плана, так и затраты (резервы), необходимые для корректировки этого плана в процессе реализации. При математической постановке задачи стохастического программирования необходимо ввести некоторую случайную величину ω , определяющую степень

отклонения фактического состояния внешней среды в период реализации плана от предполагаемого состояния в период его составления. Для учета влияния этой случайной величины на решение нужно задать зависимости коэффициентов уравнений от величины ω . Рассмотрим постановку задачи стохастического программирования на примере распределения ресурсов при долгосрочном планировании строительства объектов.

$$\sum_{j=1}^n a_{ij}(\omega)x_j + \sum_{l=1}^r b_{il}(\omega)y_l = C_i(\omega),$$

$$i=1, 2, \dots, m; \quad x_j \geq 0, \quad y_l \geq 0.$$

$$j=1, 2, \dots, n;$$

$$l=1, 2, \dots, r.$$

x_j — это суммарные плановые затраты по j -ой статье (j -й ресурс) на строительство всех m объектов. C_i — суммарные плановые и дополнительные затраты по всем статьям на строительство i -го объекта. y_l — суммарные дополнительные затраты по l -й статье на строительство всех m объектов.

План $X (x_1, x_2, \dots, x_n)$ рассчитывается при определенном допущении относительно значения ω . После того как станет известным фактическое значение ω , неувязки (расхождения) в уравнениях устраняются за счет вектора коррекции плана $Y (y_1, y_2, \dots, y_r)$ при заданных X и ω .

Суммарные затраты на реализацию и коррекцию плана строительства всех m объектов определяют целевую функцию:

$$F(X) = \sum_{j=1}^n d_j x_j + M \sum_{l=1}^r s_l(\omega) y_l(X, \omega),$$

где M, d_j, s_l, ω — некоторые коэффициенты.

В результате решения задачи необходимо найти минимум суммарных затрат $F(X)$ при условии $X \geq 0$, для чего требуется распределение значений компонент случайного вектора

$$y_l(X, \omega).$$

Практически задаются зависимостью решения x от случайной величины ω и некоторых параметров $Z_1 (z_1, z_2, \dots, z_k)$, т. е. полагают $X(\omega) = f(\omega, Z)$. Например, X может зависеть от случайного колебания цен ω на материалы и доли затрат на материалы (параметр Z) в общей стоимости строительства.

Иногда параметры по физическому смыслу могут быть только целыми числами. Подобные задачи относятся к целочисленному программированию, решаемому в основном комбинаторные задачи. Приведенный выше пример задачи о выборе оптимального набора приборов для некоторой лаборатории относится к целочисленному программированию.

Распространенным частным случаем такого рода задач являются задачи, в которых искомые переменные могут принимать одно из двух значений (0 или 1, «да» или «нет»).

В медицине к подобным задачам относятся задачи составления оптимальных расписаний обслуживания больных различными кабинетами, лабораториями и другими службами больницы, задачи выбора оптимальных наборов диагностических тестов, обеспечивающих минимальное время и достаточную надежность диагностики при условии минимального вреда больному и т. п.

Для решения задач целочисленного линейного программирования существует общий метод, разработанный американским ученым Р. Гомори. Если искомое решение является двоичным, т. е. представляет собой логическую матрицу, то применяются в основном два метода — метод упорядоченного перебора и метод ветвей и границ.

Характерной особенностью задач перспективного планирования развития отраслей народного хозяйства, а также многих задач планирования деятельности предприятий является то, что весь планируемый период может быть разбит на ряд последовательных этапов выполнения плана и план для каждого последующего этапа зависит от состояния системы в начале этого этапа и не зависит от того пути, по которому система пришла в это состояние. Задачи этого класса решаются с помощью общей методики, получившей название динамического программирования.

Для таких задач существенным является допущение о том, что значение целевой функции в конце полного периода планирования может быть представлено в виде суммы значений этой функции для отдельных этапов этого периода.

Обычно начальное и конечное состояние системы считаются известными и требуется определить оптимальный путь для перевода системы из начального состояния в конечное.

Управление движением системы осуществляется с помощью некоторых воздействий, выбираемых для каждого этапа из множества воздействий, допустимых на этом этапе.

Суть общего подхода к решению задач динамического программирования заключается в последовательном рассмотрении возможных переходов системы из одного состояния в другое, начиная с ее конечного состояния.

Таким образом, метод динамического программирования используется в тех случаях, когда процесс функционирования системы можно разбить на этапы так, что «будущее» системы полностью определяется ее «состоянием» в данный момент и не зависит от того, каким «путем» она достигла этого состояния.

Особый класс методов нахождения оптимальных решений разработан в теории игр. Методы теории игр позволяют находить оптимальные варианты решений в конфликтных ситуациях, в которых участвует несколько сторон, преследующих различные цели. Практическое применение теории игр началось в конце второй мировой войны при планировании военных операций, в процессах конкурентной борьбы различных капиталистических фирм. Особенности постановок подобных задач являются неполная информация о действиях противоположной стороны и необходимость учитывать интересы этой стороны при анализе возможных вариантов ее действий. В медицине и здравоохранении методы теории игр находят применение при планировании мероприятий по предотвращению эпидемий и борьбе с ними, при выборе оптимальных вариантов проведения диагностических и лечебных процедур, при машинной постановке диагнозов. В качестве «противоположной» стороны в данных случаях выступает природа. С помощью методов теории игр делаются фактически попытки как-то компенсировать недостаток наших знаний закономерностей рассматриваемых явлений и выбрать решение, рассчитанное на наиболее вероятный худший вариант течения процесса.

3. Методы моделирования и сетевого планирования и управления

В разделе 1-й главы II мы упоминали о наличии класса аналоговых машин, действие которых основано на принципе аналогий между явлениями различной физиче-

ской природы. Наличие этих аналогий позволяет широко применять для изучения различных явлений метод моделирования. Различают два основных вида моделирования: физическое и математическое.

При физическом моделировании вместо реального объекта исследуют другой физический объект — модель, либо той же физической природы, но другого масштаба, либо другой физической природы, но подчиняющейся аналогичным законам функционирования.

При математическом моделировании строят математическое описание — модель изучаемого явления, представляющее собой систему уравнений и различных соотношений, определяющих зависимости основных характеристик процесса от исходных данных. В этих уравнениях в качестве параметров фигурируют различные постоянные и переменные коэффициенты, значения которых характеризуют свойства изучаемого объекта или процесса.

Суть метода математического моделирования сводится к тому, чтобы путем многократного решения указанных уравнений определить течение процесса во времени при различных исходных данных и различных параметрах.

Следует четко различать роль параметров, определяющих свойства исследуемого процесса, и роль исходных данных, определяющих воздействие внешней среды на исследуемый объект.

При заданном наборе параметров, т. е. для заданного образца объекта, производят многократное варьирование исходных данных, определяющих условия протекания процесса, и путем моделирования определяют характеристики качества функционирования объекта (его устойчивость, быстродействие и т. д.). Затем варьируют параметры объекта и снова проводят моделирование при различных исходных данных и т. д. В результате определяют оптимальные значения параметров объекта для различных исходных данных.

Основу метода математического моделирования составляет математическая модель внешней среды, в которой должно происходить функционирование исследуемой (или проектируемой) управляющей системы. Подобные системы должны воспринимать информацию из внешней среды, перерабатывать эту информацию и выдавать определенные команды исполнительным органам. Эти системы рассчитываются на восприятие определен-

ных видов информации из внешней среды — так называемых заявок на обслуживание, и на выработку определенного набора команд, которые могут выполняться исполнительными органами в определенное время и с заданными ограничениями (например, на допустимые последовательности команд).

Математический аппарат, используемый для представления потока заявок на обслуживание и описания процессов функционирования управляющих систем и исполнительных органов, для многих типовых случаев содержится в теории массового обслуживания. В частности, имеются математические формулы для описания различных типов потоков независимых и зависимых заявок, для описания различных типов систем обслуживания (с очередями на обслуживание, с отказами и т. д.). Для представления процесса функционирования исследуемой управляющей системы и исполнительных органов (называемых системой обслуживания) используется метод имитации, т. е. представления в виде алгоритма на определенном алгоритмическом языке операций и процедур, выполняемых исследуемой системой при определенных внешних воздействиях. В этот моделирующий алгоритм включаются и дополнительные служебные части, учитывающие время обслуживания разных заявок, частоту их появления, степень занятости отдельных частей системы, а также осуществляющие выдачу результатов.

Важной областью применения методов моделирования является моделирование отдельных функций организма и, возможно, в будущем — моделирование основных функций организма в целом для оценки состояния организма в данный момент и прогнозирования его изменений. Моделирование осуществляется путем выяснения зависимостей между различными физиологическими, биохимическими и другими параметрами организма (температура, частота пульса, давление крови, состав газов крови и других жидкостей, наличие определенных патологических симптомов и т. д.) и представления этих зависимостей в виде дифференциальных и алгебраических уравнений, с помощью специальных математических функций, в виде табличных или графических функций, в виде логических уравнений и т. д.

Основу математической модели составляет система уравнений, позволяющая по совокупности значений раз-

личных параметров организма и их производных по времени определять интегральный критерий (или ряд критериев) оценки общего состояния организма и прогнозировать наступление опасного для организма состояния, что имеет актуальное значение, в частности, космической биологии и медицине. Кроме того, в модель включаются программы поиска оптимальных решений, осуществляющие автоматическое варьирование внешних параметров и определение таких значений этих параметров, при которых уменьшается риск возникновения опасной ситуации.

Вся эта сложная совокупность различных уравнений, представляемая в виде программ для ЭВМ с необходимыми табличными данными и коэффициентами, и образует программную модель функций организма. В состав исходных данных включаются и внешние параметры, характеризующие внешние воздействия на организм. Используя эту программную модель, можно по определенным (заданным) исходным данным и параметрам рассчитать остальные параметры организма, включенные в модель. Изменяя многократно исходные данные и параметры и рассчитывая каждый раз интересующие нас выходные параметры, можно выяснить общие закономерности и характеристики функций и состояния организма и прогнозировать состояние организма при тех или иных воздействиях. Подобное программное моделирование может применяться (и уже применяется) при контроле за состоянием организма при критических ситуациях, например: во время хирургической операции и в послеоперационном периоде, при определенных видах медикаментозного лечения, во время космического полета или при подготовке космонавта и т. д. Практически это осуществляется путем использования ЭВМ, в которую вложена в виде программ и табличных данных модель организма и к которой подключена система датчиков различных физиологических параметров, установленных на контролируемом организме. Кроме того, к ЭВМ подключено одно или несколько экранных устройств отображения информации (дисплеев), совмещенных с клавишными устройствами ввода данных в ЭВМ и устройствами печати данных, отображаемых на экране.

В процессе автоматического контроля за состоянием организма информация снимается датчиками с организма, преобразуется из аналоговой формы в цифровую и

вводится в ЭВМ, которая рассчитывает все остальные параметры организма, а также интегральные критерии оценки общего состояния организма.

При превышении интегральным критерием допустимого значения ЭВМ автоматически выдает сигнал тревоги (звуковой и световой — загорается красная лампочка). В процессе работы на экранах дисплеев непрерывно отображаются выдаваемые ЭВМ значения тех физиологических параметров, которые контролируются соответствующими специалистами. Например, во время хирургической операции на дисплей, предназначенный для анестезиолога, выдаются данные, характеризующие глубину наркоза и состояние оперируемого.

Каждый специалист контролирует относящиеся к нему параметры и в случае необходимости может вызвать на этот же экран из ЭВМ дополнительные данные, изменить масштаб графика, вызвать повторно ранее выданные, снять копию с отображаемых данных на бумагу.

Таким образом, моделирование в реальном масштабе времени сочетается с проведением натурального медико-биологического эксперимента либо диагностического или терапевтического мероприятия, и модель в этом случае работает параллельно и используется для анализа и контроля реально протекающих процессов в живом организме. Возможна и такая постановка задач моделирования, когда живой организм, снабженный соответствующими датчиками физиологических параметров и приборами для воздействия на состояние организма, включается в общую схему моделирования в качестве одного из элементов замкнутого контура обратной связи. При этом программная модель, реализованная с помощью ЭВМ, используется для анализа сигналов, поступающих с живого организма, и выработки необходимых воздействий, изменяющих состояние данного организма.

Важным классом математических моделей биологического характера являются модели процессов развития популяций живых организмов, в которых с помощью математических уравнений описываются зависимости, определяющие темпы размножения, гибели и взаимодействия между различными видами живых организмов. Используя подобные модели, исследуют условия, благоприятствующие развитию того или иного вида организма, а также характеристики внешней среды, оказывающие решающее влияние на процессы биологической эволю-

ции. Подобные исследования приобрели особенно большое значение в последние годы в связи с серьезной опасностью загрязнения внешней среды радиоактивными и другими промышленными отходами.

Математическое моделирование медико-биологических процессов представляет собой чрезвычайно важное средство исследования в космической биологии и медицине. Оно позволяет сократить дорогостоящие натурные эксперименты и повысить качество анализа и контроля за проведением особо ответственных исследований и медицинских процедур. С помощью методов математического моделирования исследуются процессы переработки информации и управления в биологических системах, в том числе процессы хранения и передачи генетической информации, процессы передачи информации в нервной системе и в живом организме в целом. Естественно, что эти процессы в силу чрезвычайной сложности могут быть представлены для моделирования в настоящее время в весьма упрощенном виде. Однако даже такое упрощенное представление позволяет получить с помощью моделирования ряд важных теоретических данных, подтверждаемых реальными экспериментальными данными. Совместное использование методов математического моделирования и реального эксперимента, представляющее сущность системно-кибернетического подхода к изучению явлений живой природы, следует считать весьма перспективным при решении сложных медико-биологических проблем.

Сетевое планирование и управление (СПУ) представляет собой эффективное средство планирования и контроля хода выполнения сложных проектов или больших организационных мероприятий, включающих значительное число исполнителей, различных этапов и конкретных работ.

Хотя основные преимущества этого метода проявляются в полной мере при проектах или мероприятиях при планировании и организации исполнения небольших (локальных) мероприятий.

Применение метода СПУ обеспечивает существенное повышение качества планирования и управления, выражающееся в сокращении сроков выполнения проектов, экономии материальных, денежных и людских ресурсов.

Издавна при планировании сложных и даже не очень сложных разработок, мероприятий, строительства различных объектов и т. д. применяются различные графики, представляемые в виде диаграмм, таблиц, ленточных схем и др.

Метод СПУ появился в конце 50-х годов как дальнейшее развитие графических способов представления планов работ. Основу этого метода составляют тщательный анализ всех элементарных компонентов (этапов работ, шагов процесса, отдельных процедур и т. д.), необходимых для выполнения всего проекта, и выявление причинно-временных зависимостей между ними. Каждая такая компонента носит название работы. Она характеризуется тремя сторонами: длительностью, совокупностью условий для начала, конечным результатом. Все работы и их взаимосвязи изображаются в виде графика, причем работы представляются стрелками. Совокупность условий для начала работы называется начальным событием работы, результат работы — конечным событием работы. Начальное и конечное события работ изображаются кружками, в которых указываются порядковые номера событий. Стрелка (работа) идет от начального к конечному событию; номер работы состоит из номеров ее начального и конечного событий. При объединении разных работ в единый график (так называемый сетевой график) конечные события одних работ являются начальными событиями других (последующих) работ. Для всего сетевого графика (всего проекта или мероприятия) существует одно начальное событие, называемое исходным. Весь проект, как правило, завершается одним конечным событием, так называемым завершающим. Заметим, что современные методы СПУ могут быть применены и к таким проектам или мероприятиям, которые имеют не один, а несколько конечных результатов. В подобных случаях на сетевом графике будет несколько завершающих событий.

Путь на сетевом графике — это последовательность сцепленных стрелок, представляющих работы. Каждая работа характеризуется определенным временем ее выполнения (длительностью), и сумма этих времен для последовательности сцепленных стрелок является общим или суммарным временем выполнения всех работ данного пути. Наибольший интерес при анализе сетевых графиков представляет так называемый критический путь.

Критический путь — это путь между исходным и завершающим событием, имеющий наибольшую суммарную продолжительность работ или, как принято говорить, наибольшую длину.

Длина критического пути является основной характеристикой сетевого графика, так как она определяет сроки окончания всего проекта или выполнения строительства и т. п. Исполнитель каждой работы, содержащейся в сетевом графике, должен дать на основе своего опыта и расчетов две или три оценки ожидаемого времени выполнения своих работ: минимальное время, за которое можно выполнить работу при благоприятном стечении обстоятельств (оптимистическая оценка), наиболее вероятное время выполнения работы (реалистическая оценка) и максимальное время, если события сложатся наиболее неблагоприятно (пессимистическая оценка). Разработаны математические способы учета этих оценок, позволяющие вычислить три аналогичные оценки для времени выполнения всего проекта или мероприятия.

Для всех путей, кроме критического, имеются резервы времени событий и работ. Анализ сетевого графика позволяет разделить все работы в зависимости от их напряженности (наличия или отсутствия резервов времени) на ряд категорий. Во-первых, как говорилось уже, выделяются все работы, принадлежащие критическому пути, и на них сосредоточивается внимание руководителей проекта. Во-вторых, определяются так называемые подкритические пути, которые по своей длительности приближаются к критическому пути и могут при определенных условиях стать критическими. При перераспределении ресурсов работы, принадлежащие подкритическим путям, обычно не затрагиваются. Стараются изыскать дополнительные ресурсы за счет работ, принадлежащих другим, так называемым ненапряженным путям. Очевидно, что при перераспределении ресурсов в выделении дополнительных сил и средств на выполнение работ критического пути продолжительность этих работ сокращается, что приводит к общему уменьшению длительности данного пути. При этом критический путь может перестать быть критическим и его место займет один из подкритических путей. Такое явление может иметь место и без перераспределения ресурсов, а в связи с конкретным ходом различных работ. Особенно это

относится к выполнению больших научно-исследовательских и опытно-конструкторских разработок, при которых предварительное определение продолжительности отдельных работ является ориентировочным и может уточняться в ходе работы.

Анализ сетевого графика позволяет определить не только критический путь, но и исполнителей соответствующей этому пути последовательности работ.

Ясно, что длительность критического пути определяет и длительность выполнения всего проекта в целом. Таким образом, если необходимо сократить сроки выполнения всей работы, то сетевой график дает возможность сосредоточить внимание только на тех работах, которые лежат на критическом пути, и отыскать способы ускорения их выполнения. Этого можно достичь, например, путем привлечения на эти участки работы больших сил и средств, в том числе и за счет тех сил и средств, которые предусмотрены для ненапряженных путей. Конечно, такая переброска должна выполняться в разумных размерах, чтобы от этого ненапряженные пути сами не стали критическими.

После того как в результате этих мероприятий продолжительность критического пути будет сокращена, возможно, что критическим станет уже другой путь, т. е. другая последовательность работ. Этот путь также можно подвергнуть анализу для изыскания возможностей его сокращения и т. д. Таким образом, метод СПУ приводит к повышению четкости в планировании работ, дает возможность оптимальным образом использовать имеющиеся ресурсы, осуществлять систематический конкретный контроль за ходом выполнения проекта. Для очень сложных проектов, включающих тысячи и десятки тысяч работ, составление, расчет и анализ сетевых графиков представляют достаточно сложное дело, и с этой целью эффективно используются ЭВМ. Имеются типовые методики, инструкции и программы для ЭВМ для обработки сетевых графиков. ЭВМ осуществляет составление и корректировку сетевого графика в соответствии с исходными данными, получаемыми от исполнителей работ, расчеты критического пути и резервов времени для событий и работ, лежащих на других путях, и печатает выходные документы, позволяющие руководителю проекта или мероприятия принимать решения по управлению этим процессом.

4. Элементы прикладной математической логики

При анализе и постановке различных задач обработки информации и принятия решений в медицине и здравоохранении приходится часто формулировать соотношения и зависимости между различными величинами (признаками, показателями, причинами, следствиями и т. д.). Например, при постановке задач машинной диагностики заболеваний необходимо записать предварительно в виде таблицы связи между симптомами и диагнозами: при выборе с помощью ЭВМ рациональных методов лечения необходимо иметь зависимости рекомендуемых методов лечения от диагнозов заболеваний, состояния больного, его физических и других характеристик, а также от условий лечения (в стационаре, поликлинике, на дому и т. п.). Многочисленные случаи описания различных соотношений между величинами возникают также при решении задач планирования и организации здравоохранения (структура заболеваемости, демографические данные обслуживаемого населения, специализация и размещение сети лечебно-профилактических учреждений и т. п.).

В математике давно уже разработан общий подход к составлению формальных выражений (формул), показывающих логические зависимости между различными переменными и постоянными величинами. Раздел математики, занимающийся этими вопросами, называется математической логикой.

Математическая логика — сложная математическая дисциплина, возникшая под влиянием внутренних потребностей самой математики в обосновании ее исходных положений и принципов. Однако разработанные в математической логике символика и методика составления и преобразования логических выражений удобны для практики постановки и описания различных информационно-логических задач. Основным принципом математической логики является использование логических величин, т. е. величин, которые могут иметь только одно из двух значений: «да» или «нет» (истинно или ложно). Эти значения представляются также единицей («да») и нулем («нет»). Для применения методов математической логики к задачам, в которых участвуют величины, имеющие различные численные количественные или качественные (признаковые) значения, необходимо каждую

величину представить в виде ряда логических величин. Например, если нужно представить возраст больного в виде логической величины, то весь диапазон возможных возрастов делится на интервалы (например, до 1 года, от 1 года до 2 лет, от 2 до 3 лет и т. д.) и возраст больного обозначается постановкой крестика (или подчеркиванием) соответствующего интервала. Таким же способом в виде набора логических величин можно представить температуру, давление, цвет и т. д. Очевидно, что чем точнее мы желаем представить ту или иную количественную величину в виде логической величины, тем меньшими должны быть интервалы, на которые делится весь диапазон изменения данной количественной величины.

Многие признаки и симптомы, рассматриваемые при диагностике различных заболеваний, по существу являются логическими величинами, т. е. допускают только один из двух возможных ответов («да» или «нет»). Например, сведения о перенесенных заболеваниях, наследственности, бытовых условиях, проведенных операциях, наличии осложнений, прививок и т. д. Часто при опросах больных возникает ситуация, когда, помимо ответов «да» или «нет», фигурирует третий вариант ответа «не знаю». В этих случаях каждый вопрос можно представить в виде двух последовательных вопросов, на которые должны даваться такие ответы: «да» или «не знаю», «нет» или «не знаю», или же в таком варианте: «знаю» или «не знаю» и при ответе «знаю» задается второй вопрос, на который должен быть ответ «да» или «нет». Таким образом любые величины (с предварительным преобразованием — разделением на интервалы или непосредственно) могут быть представлены в виде логических величин. Как и обычные арифметические или алгебраические величины, логические величины могут быть постоянными или переменными. Постоянные логические величины имеют либо значение 1 (истинно), либо 0 (ложно), переменные логические величины так же, как и в обычной алгебре, обозначаются буквами (А, В, С, D и т. д.). Каждая переменная величина может принимать определенное значение (0 или 1) в зависимости от условий задачи. Для обозначения независимых (исходных) логических величин (аргументов) обычно используются начальные буквы латинского алфавита, а для обозначения функций (зависимых величин) — последние буквы

латинского алфавита (x, y, z). Иногда применяются и другие буквы (греческие, русские и др.). Логическая величина — это всегда некоторое утверждение (высказывание), которое может быть истинным или ложным. Когда мы обозначаем эту величину буквой, то смысл этого высказывания мы подразумеваем и мысленно связываем с данной буквой. Иногда эту связь мы записываем для памяти явно на бумаге, например: «пусть A обозначает, что больной — это ребенок моложе 14 лет». Если данный больной действительно ребенок моложе 14 лет, то величина A будет иметь значение 1, если нет — то 0.

Подобные высказывания делаются обычно по отношению к некоторой группе или классу объектов, причем по отношению к некоторым объектам из данного класса (например, группа больных) высказывание может быть истинным, а по отношению к другим объектам из того же класса — ложным. Такого рода высказывания в математической логике принято называть предикатами. Представитель класса объектов, обозначаемый буквой x или некоторым общим наименованием (например, «больной»), называется предметной переменной. Эта предметная переменная может принимать конкретное значение (например, больной Петров А. И.), тогда высказывание (предикат) принимает тоже конкретное значение «истинно» или «ложно» (0 или 1). Предикат может быть одноместным, т. е. касаться одного объекта; тогда он выражает свойство этого объекта. Предикат может быть двухместным, т. е. касаться двух объектов, трехместным и вообще многоместным. В последнем случае предикат выражает отношение между объектами. Например, предикат: «больной A и больной B имеют одинаковый диагноз» будет истинным (т. е. будет иметь значение 1), если вместо общих обозначений больных A и B будут подставлены фамилии (или номера) больных с одинаковыми диагнозами, и ложное значение (0), если в качестве больных A и B будут фигурировать больные с разными диагнозами.

Применение букв придает высказываниям общность, т. е. позволяет формулировать законы логики для любых высказываний.

Из одного или нескольких высказываний, принимаемых за простые, можно составлять сложные высказывания, которые будут функциями простых высказываний.

Используются определенные логические операции (или логические связи), позволяющие объединять некоторые данные высказывания (постоянные или переменные). Ниже приводятся определения основных логических операций и дается в наглядной табличной форме их пояснение.

Таблица 3

Операция отрицания	
A	\bar{A}
1	0
0	1

Таблица 4

Операция конъюнкции		
A	B	$A \wedge B$
1	1	1
0	1	0
1	0	0
0	0	0

К основным логическим операциям относятся операции отрицания, конъюнкции, дизъюнкции, эквивалентности и импликации.

Отрицание высказывания A — это высказывание, которое истинно, когда A ложно, и ложно, когда \bar{A} истинно; обозначается через \bar{A} и читается «не A ». Операция отрицания задается табл. 3. В этой и последующих таблицах истинное значение представляется единицей, а ложное — нулем.

Конъюнкция двух высказываний — сложное высказывание, которое истинно в случае истинности обоих высказываний, его образующих, и ложно в остальных случаях; обозначается через $A \wedge B$ и читается « A и B »; знак логической операции \wedge имеет смысл союза «и» и называется знаком конъюнкции. Часто конъюнкция называется логическим умножением. Операция конъюнкции задается табл. 4.

Дизъюнкция двух высказываний — сложное высказывание, которое ложно в случае ложности обоих составляющих его высказываний и истинно в остальных случаях; обозначается $A \vee B$ и читается « A или B ». Дизъюнкцию часто называют логическим сложением. Знак логической связи \vee имеет смысл союза «или» и называется знаком дизъюнкции. Дизъюнкция задается табл. 5.

Эквивалентность двух высказываний — сложное высказывание, истинное тогда, когда значения истинности составляющих высказываний одинаковы, и ложное в про-

Таблица 5

Операция дизъюнкции		
A	B	$A \vee B$
1	1	1
0	1	1
1	0	1
0	0	0

Таблица 6

Операция эквивалентности		
A	B	$A \equiv B$
1	1	1
0	1	0
1	0	0
0	0	1

тивном случае; обозначается $A \equiv B$ и читается: «А эквивалентно В». Табл. 6 иллюстрирует существо операции эквивалентности.

Следует заметить, что логическая операция эквивалентности дает результат истинно, если она применена к двум исходным логическим величинам, значения истинности которых совпадают, т. е. одновременно истинны или одновременно ложны. Эта операция может быть последовательно распространена и на большее число исходных логических величин.

В качестве последней операции математической логики, представляющей интерес с точки зрения записи алгоритмов решения задач поиска и обработки информации на ЭВМ, мы рассмотрим операцию импликации.

Импликация двух высказываний — это сложное высказывание, связывающее два высказывания (предшествующее и последующее). Импликация ложна в том и только в том случае, если предшествующее высказывание истинно, а последующее ложно. Импликация обозначается стрелкой $A \rightarrow B$. Суть этой операции показана в табл. 7.

Импликация по смыслу может истолковываться как связь между условием А и следствием В (хотя в математической логике эта операция определяется формально и ее смысл как причинной зависимости не подчеркивается).

Введение буквенных обозначений логических величин и четкое формальное определение основных логических операций позволяют не только записывать сложные ло-

Таблица 7

Операция импликации		
A	B	$A \rightarrow B$
1	1	1
0	1	1
1	0	0
0	0	1

гические выражения (логические формулы), но и осуществлять их формальные преобразования — упрощения или приведения к типовому стандартному виду. При этом используются следующие основные правила преобразований.

1. $A = \bar{\bar{A}}$. Два и вообще четное число операций отрицания могут быть отброшены.

2. $A \wedge B = B \wedge A$; $A \vee B = B \vee A$. Перестановка местами членов конъюнкции или членов дизъюнкции не изменяет значения этого логического выражения.

3. $(A \wedge B) \wedge C = A \wedge (B \wedge C)$; $(A \vee B) \vee C = A \vee (B \vee C)$. Для нескольких операций конъюнкции и для нескольких операций дизъюнкции не имеет значения порядок выполнения этих операций.

4. $A \wedge B = A \vee \bar{B}$; $A \vee B = A \wedge \bar{\bar{B}}$, операция отрицания над конъюнкцией, дизъюнкцией) может быть заменена операцией отрицания над членами конъюнкции (дизъюнкции) при условии, что эти члены будут затем соединены знаком дизъюнкции (конъюнкции).

5. $A \wedge 1 = A$; $A \vee 0 = A$; $A \vee 1 = 1$; $A \wedge 0 = 0$. Эти формулы показывают, как определяется значение конъюнкции или дизъюнкции, если один из членов является постоянно истинным (1) или постоянно ложным (0).

Рассмотрим некоторые примеры применения математической логики при решении на ЭВМ задач обработки информации. В качестве первого примера можно привести задачу составления сводных отчетов по массиву первичных учетных документов. Такой случай имеет место, например, при решении на ЭВМ задач учета кадров. При этом в качестве исходных данных в ЭВМ вводят анкетные сведения о каждом работнике. Эти данные записывают на магнитную ленту или магнитный диск. На основе этих первичных данных с помощью ЭВМ получают сводные отчеты о составе кадров и их распределении по возрасту, специальности, стажу работы, прохождению переподготовки и т. д. Рациональным является способ, при котором массив записей первичных документов просматривают один раз. При этом каждый первичный учетный документ (каждую анкету в нашем примере) проверяют сразу же на соответствие наборам признаков всех выходных отчетных форм. В случае положительных ответов производят прибавление единиц в соответствующие графы промежуточных таблиц, соответствующих вы-

ходным формам. При этом логическая функция поиска может иметь, например, такой вид:

$$\Phi(p_1, p_2, \dots, p_s) = p_1 \wedge \underset{\text{I}}{\dots} \wedge p_j \wedge (p_k \vee \underset{\text{II}}{\dots} \vee p_l) \wedge \\ \wedge (p_m \vee \underset{\text{III}}{\dots} \vee p_n) \wedge (p_r \vee \underset{\text{IV}}{\dots} \vee p_s).$$

Здесь признаки I представляют собой общую группу признаков, которыми должны обладать все объекты, относимые к определенной категории (графе выходной формы). Кроме того, в скобках указаны несколько групп переменных признаков, соединенных знаками \vee — дизъюнкции. Каждая из этих групп признаков будет характеризовать определенную подграфу данной выходной формы (II, III или IV).

Вторым примером применения логических матриц является поиск объектов, отвечающих заданному набору признаков.

Пусть мы имеем M объектов и N признаков (логических величин). Для каждого признака строится логическая матрица с M элементами.

Логическая матрица — это прямоугольная таблица элементов, каждый из которых может принимать значения 0 или 1; логический вектор — строка элементов, принимающих значения 0 или 1. Над логическими матрицами (векторами) можно выполнять ряд операций, необходимых при решении задач поиска заданных объектов по признакам или при определении связей между объектами.

Во всех матрицах одинаковые по порядку номеров элементы соответствуют одним и тем же объектам. Если данный объект обладает данным признаком, то в соответствующей матрице на месте его элемента ставится единица, если не обладает, то ставится ноль.

Поиск объектов, имеющих заданный набор признаков $p_1 \wedge p_j \wedge \dots \wedge p_k$, осуществляют путем переписи с МЛ (или МД) в ОЗУ матриц, соответствующих этим признакам (т. е. i -й, j -й, ..., k -й матриц в данном случае), и их логического перемножения (операция конъюнкции). Затем определяют порядковые номера ненулевых элементов в матрице — произведении, которые будут соответствовать порядковым номерам объектов, обладающих заданным набором признаков.

Используя логические матрицы подобного типа, можно производить поиск объектов по наборам признаков, связанных различными логическими связями. Например, можно искать объекты, обладающие признаком p_i или p_j и не обладающие признаком p_k , т. е.

$$\Phi(p_i, p_j, p_k) = (p_i \vee p_j) \wedge \bar{p}_k,$$

при этом над j -й и i -й логическими матрицами должна быть выполнена операция дизъюнкции, затем должно быть получено дополнение (отрицание) k -й логической матрицы, и над матрицами логической суммы и дополнения должна быть выполнена операция конъюнкции. Номера найденных объектов определяют так же, как в предыдущем случае.

Примером использования логических высказываний может служить автоматизация постановки диагноза. В этом случае простыми высказываниями будут утверждения о симптомах больного, например: А — температура повышена, В — пульс учащенный, С — лейкоцитоз повышенный, D — язык обложен, E — РОЭ повышено, F — дыхание затруднено, G — имеется сыпь и т. д.

Из этих высказываний можно составить сложное высказывание, утверждающее наличие у пациента некоторого заболевания, положим аппендицита. Это сложное высказывание будет справедливым в том случае, если некоторые из перечисленных высказываний (подтверждающие гипотезу об этом заболевании) будут истинными, а некоторые (исключающие вероятность этого заболевания) — ложными. Естественно, что для каждого заболевания будет своя логическая формула, т. е. своя комбинация истинности или ложности тех же простых утверждений (симптомов).

В действительности формула, выражающая диагноз, имеет гораздо более сложный вид, так как необходимо, например, учитывать многие особенности проявления одного и того же симптома (это тоже можно сделать с помощью логических формул): насколько повышена температура и каков ее суточный ход, когда появилась сыпь, ее характер и место появления и т. д.

АВТОМАТИЗАЦИЯ ПОИСКА ИНФОРМАЦИИ

1. Назначение и состав информационно-поисковой системы (ИПС)

Поиск информации является одним из основных кибернетических процессов, характерных для управляющих систем любой природы (живых организмов или автоматических устройств).

Информация, которая отыскивается, обычно содержится в письменных материалах разных видов — в книгах, журнальных статьях, патентах, научно-технических отчетах, рекламных сообщениях и т. д. В области информационного поиска все подобные источники информации обозначаются общим термином — **документы**. В настоящее время объем понятия «документ» существенно расширился и под документом понимают любые материальные носители информации, т. е. любые объекты поиска. Таким образом, в зависимости от области применения, документом будут называться история болезни, медицинское предписание, анкета личного состава, чертеж и т. п.

ИПС такого рода называются документальными. В ответ на запрос они указывают документ или документы, в которых содержатся запрашиваемые сведения (книги, статьи, истории болезни).

Существуют также **фактографические ИПС**, которые в ответ на запрос выдают непосредственно сами запрашиваемые сведения (как правило, числовые характеристики или другие конкретные данные).

Огромные массивы исходной информации, среди которой нужно отыскивать сведения, привели к необходимости создания автоматизированных ИПС, получающих все большее и большее распространение. Они являются важной составной частью АСУ, а также любой системы машинной диагностики, ведения историй болезней, выбора методов лечения, контроля использования медикаментов и т. д.

Основными требованиями, предъявляемыми к ИПС, являются обеспечение скорости, точности и полноты поиска документов, удобство обновления фонда документов и словаря, надежность в работе, простота эксплуатации.

Типовая автоматизированная ИПС состоит из следующих основных частей:

1. ЭВМ с большим объемом памяти и широким набором внешних устройств для подготовки входной информации, отображения и размножения выходной информации.

2. Комплекс (пакет) алгоритмов и программ, обеспечивающих ввод, контроль и формирование массивов информации в ЭВМ, а также поиск, редактирование и выдачу информации из ЭВМ.

3. Комплекс информационных массивов, включающий массивы терминологических словарей, поисковые и справочные массивы, а также вспомогательные массивы наименований различного рода (учреждений, журналов и т. п.).

4. Система технологической документации, включающая инструкции по отбору и подготовке (индексации) исходной информации ко вводу в ЭВМ, бланки для записи этой информации, инструкции для работы на ЭВМ при разных режимах (ввод данных, корректировка, поиск и выдача данных и др.), методические указания по использованию ИПС и т. д.

5. Внемашинный архив документов, включающий первичные документы и выходные данные на микрофильмах, перфокартах, перфолентах и бумажных документах.

Основной задачей любой ИПС является так называемый ретроспективный поиск информации, т. е. поиск информации по тематическим запросам в определенном объеме исходных данных, введенных предварительно в память ЭВМ.

Обычно большие массивы данных, в которых должен производиться поиск, хранятся на магнитных лентах. Массивы средних размеров, а также словари различного рода (см. ниже) хранятся на магнитных дисках. В процессе работы ИПС эти массивы частями переписываются для просмотра в оперативное запоминающее устройство ЭВМ. Кроме ретроспективного поиска, ИПС реализуют поиск в других режимах или другого назначения,

например: распределение документов по заранее установленным рубрикам и подрубрикам (подготовка библиографических указателей), избирательное распределение информации (ИРИ), т. е. распределение текущих поступлений документов по заранее сформулированным тематическим профилям постоянных абонентов, и др.

В настоящее время основным способом представления научных, технических, экономических, медицинских и других текстов в ИПС, обеспечивающим эффективную автоматическую обработку и поиск информации, является перевод их основного смыслового содержания на некоторый формализованный искусственный язык, называемый информационно-поисковым языком (ИПЯ).

ИПЯ включает словарь терминов, называемых дескрипторами, и набор формальных грамматических правил для связи дескрипторов в предложения. Эти правила могут быть очень простыми (например, правило: запись дескрипторов последовательно через запятые) или достаточно сложными (запись дескрипторов в предложении с указанием грамматической роли каждого дескриптора и смысловых связей между ними). Дескрипторы имеют свои порядковые номера, называемые кодами дескрипторов (КД). Для обеспечения возможности машинного поиска и обработки документов необходимо основное смысловое содержание документа выразить с помощью ИПЯ или, как принято говорить, составить поисковый образ документа (ПОД). Этот процесс выполняется либо людьми, которые называются индексаторами, либо ЭВМ автоматически по специальной программе. Запросы от потребителей информации, поступающие на естественном языке, также должны представляться на ИПЯ, как правило, теми же индексаторами, которые осуществляли индексацию документов данного профиля.

Индексация документов заключается в основном в выборе из текста документа наиболее характерных (информативных) терминов или в составлении фраз определенной структуры.

Замена словесных представлений дескрипторов их кодами (номерами), а также представление грамматических связей между дескрипторами с помощью числовых кодов называются кодированием ПОД и поисковых

образов запросов (ПОЗ). Обратная замена кодов дескрипторов и кодов грамматических связей их словесными представлениями называется декодированием.

Поисковые образы документов в кодированном виде вводятся в ЭВМ при создании ИПС и в последующем — по мере появления новых документов. При поиске кодированный поисковый образ запроса также вводится в ЭВМ и сопоставляется с ПОД всех введенных документов.

Сопоставление ведется по определенным, чисто формальным правилам, которые в информационном поиске называются критерием соответствия, или критерием выдачи (КВ). Если эти правила для ПОД некоторого документа удовлетворяются, этот документ выдается, т. е. абоненту сообщается, что содержание этого документа соответствует поставленному им запросу. При проверке всех выданных документов может оказаться, что некоторые документы действительно отвечают запросу (такие документы называются релевантными), а некоторые документы не соответствуют — они нерелевантны поставленному запросу. Говорят, что человек при установлении факта релевантности пользуется критерием смыслового соответствия (КСС), т. е. некоторыми неформализуемыми правилами, основанными на его эрудиции, инженерном опыте и т. п. Если среди выданных документов окажется чрезмерное количество нерелевантных документов, то это значит, принятый в ИПС критерий выдачи малоэффективен, т. е. плохо согласован с критерием смыслового соответствия. В таких случаях КВ подвергаются корректировке, или улучшают правила индексирования документов, или делают и то и другое. Эти же операции производят тогда, когда среди выданных на запрос документов отсутствуют многие релевантные документы, содержащиеся в поисковом массиве. В таких случаях говорят, что велики потери поиска.

Выше описан один из способов представления в ЭВМ сведений об искомым документах — в свернутом виде, в виде ПОД и ПОЗ.

Существует также другой способ представления информации в ИПС — использование непосредственно естественного языка научных или деловых текстов. Однако этот способ пока не получил широкого распространения из-за своих ограниченных возможностей и трудности ре-

ализации достаточно точных алгоритмов анализа текстов на естественном языке.

Следует подчеркнуть, что вопрос о выборе форм и способов представления информации, подготавливаемой ко вводу в ИПС и хранимой в ИПС (внешний и внутренний языки ИПС), является одним из основных вопросов создания ИПС, определяющих эффективность взаимодействия людей и ИПС.

Для человека наиболее удобными формами являются текстовая и графическая (картинная) формы, в то время как для машинной обработки необходима цифровая (кодированная) форма представления информации. Можно назвать четыре основных пути решения этой проблемы:

1. Ручное индексирование и кодирование информации для ввода в ЭВМ и ручное декодирование ее при выводе из ЭВМ. Этот способ является несовершенным, трудоемким и сопряжен с большим количеством ошибок. Однако в некоторых ИПС, особенно узкоспециального назначения, указанный способ применяется и в настоящее время.

2. Ручное индексирование и автоматическое кодирование и декодирование информации. При этом человек-индексатор должен представить информацию перед вводом в ЭВМ в формализованном виде, но в удобной для него текстовой форме, а ЭВМ переведет ее в цифровую кодированную форму; при выводе данных ЭВМ осуществит обратное преобразование из кодированной формы в текстовую.

3. Автоматическое индексирование и кодирование текстов при вводе в ЭВМ и автоматическое декодирование данных при выводе их из ЭВМ. Этот способ является развитием второго способа в направлении уменьшения степени формализации исходных текстов (допустимость использования простых назывных предложений, возможность использования падежей и форм множественного числа, применение большего или меньшего набора предлогов, возможность варьирования в предложениях порядка слов и т. д.).

Автоматическое индексирование текстов позволяет привести тексты к виду, обеспечивающему возможность автоматического анализа, поиска и обобщения информации. При индексировании документов производится не только замена встречающихся терминов и грамматических связей их стандартными эквивалентами (норма-

лизация текста), но и сжатие текста, т. е. исключение из него второстепенных деталей. Таким путем получается поисковый образ документа, имеющий значительно меньший размер. Однако ПОД, в отличие от самого документа, пригоден в основном только для осуществления поиска документов.

4. Хранение, поиск и обработка информации в ЭВМ на естественном языке. Этот способ реализован в ряде систем; он позволяет пока что осуществлять поиск документов по запросам, содержащим наборы дескрипторов (с учетом или без учета грамматических связей между дескрипторами). В отличие от предыдущего способа, основанного на приведении различных фраз естественного языка, выражающих одинаковый смысл, к единообразной кодированной форме, в данном случае производится, наоборот, размножение фразы, представляющей запрос на поиск; получение из одной фразы запроса нескольких эквивалентных по смыслу вариантов (модификаций запроса). Эти варианты проверяются на сравнение с фразами естественного языка, выражающими содержание документов. При этом одновременно используются некоторые приемы нормализации фраз запросов и документов.

Перечисленные варианты представления информации используются, вообще говоря, как в документальных, так и в фактографических ИПС. Реальные ИПС не всегда можно четко отнести к тому или другому типу. В зависимости от назначения ИПС в ней могут быть в различной степени представлены документальные или фактографические свойства.

Основные принципы и особенности построения перечисленных выше основных вариантов ИПС удобнее всего уяснить на примере конкретной ИПС. В качестве такого примера возьмем документально-фактографическую ИПС, реализованную в Московском электротехническом институте связи на ЭВМ «Минск-32».

2. Документально-фактографическая ИПС, использующая формализованный естественный язык назывных предложений

В данной ИПС сочетаются в равной степени свойства и возможности документальных и фактографических ИПС, в основном благодаря использованию формализо-

ванного языка, пригодного для записи как поисковых образов документов, так и фактографических справок. Последние могут записываться и на обычном естественном языке.

В подобных системах обеспечивается гибкость применения: в зависимости от условий использования, требований заказчиков, наличия исходной информации и т. д. возможны накопление и выдача как ссылок на документы, так и фактических справок, представляющих собой конкретные ответы на поставленные вопросы. В процессе эксплуатации возможны пополнение документальных ссылок фактическими сведениями, корректировка ранее введенных справочных данных, установление перекрестных ссылок между различными документами и т. д.

При построении формализованного языка для индексации документов следует исходить из трех основных и в значительной мере противоречивых требований. С одной стороны, язык должен обладать достаточной «семантической силой», т. е. средствами для выражения содержания индексируемых документов, с другой — должен быть полностью формализованным и по возможности более простым, что диктуется стремлением к упрощению практических алгоритмов анализа и синтеза предложений. И, наконец, с третьей стороны, язык должен как можно ближе подходить к естественному языку медицинской литературы и документов, что должно обеспечить удобство использования его людьми.

Выбор структуры языка (состава смысловых связей и допускаемой структуры предложений) основывается в первую очередь на анализе структуры и состава типовых предложений медицинских текстов. В этих текстах основными являются назывные предложения, в которых описывается какое-то событие (процесс, явление), происходящее с каким-то (над каким-то) объектом, с определенным назначением (целью), при помощи каких-то средств или методов и при наличии определенных условий. Кроме того, указываются результаты или последствия этих действий. При названиях процессов, объектов, средств и других элементов событий обычно строят определения в виде прилагательных. Анализ показывает возможность отказаться при формализации медицинских текстов вообще от использования глаголов, имея в виду назначение языка для описания существа содержания

научных и практических медицинских документов и возможность значительно упростить за счет исключения глаголов структуру языка.

Статистический анализ фраз медицинских документов подтверждает целесообразность выбора перечисленной структуры типового назывного предложения и соответствующей системы смысловых связей.

Следует подчеркнуть особенность использования формализованного языка в ИПС: формализованные фразы поисковых образцов документов соответствуют требованиям автоматической обработки с помощью ЭВМ (сравнение, эквивалентные преобразования, подсчеты статистических характеристик и др.) и в то же время они вполне пригодны (понятны) для восприятия людьми при выдаче их на печать.

В описываемом формализованном языке допускаются только назывные фразы определенной структуры (без глаголов) и фиксированный набор смысловых связей. Каждая формализованная фраза должна содержать одно или несколько подлежащих — существительных в именительном падеже и может содержать определения в форме сокращенных прилагательных и дополнения в родительном падеже. Применение сокращенных прилагательных вместо полных (например, красн. вместо красный, красная, красное, красные и т. д.) резко упрощает грамматический анализ и кодирование прилагательных. То же можно сказать и об ограничении форм существительных только именительным и родительным падежами единственного числа; получающиеся при этом формализованные фразы обеспечивают достаточную полноту и точность представления ПОД и ПОЗ с точки зрения автоматического поиска и обработки информации.

Допустимые варианты смысловых связей данного языка реализуются с помощью следующих 10 видов синтаксических признаков или отношений:

- признак подлежащего (главного члена предложения);
- признак определения (сокращенного прилагательного или причастия);
- признак дополнения (существительного в родительном падеже);
- слово «без» (отношение отрицания);
- союз «или» (отношение равнозначности);
- запятая (отношение перечисления);

- слово «для» (отношение назначения);
- слово «после» (причинно-временное отношение);
- слова «при наличии» (отношение условия);
- слова «при помощи» (отношение метода, средства).

Помимо ограничений в отношении состава смысловых связей и допустимых словоформ в языке формализованных фраз, имеются ограничения на допустимые конструкции предложений.

В наглядной форме структура формализованного сообщения¹ показана на схеме 6.

Иногда формализованные языки подобного рода называют гибридными языками, имея в виду, что в них сочетаются две стороны: достаточно полная лексика естественного языка для определенной области знаний и набор формальных грамматических средств.

Для записи формализованных фраз в памяти машины используется набор цифровых кодов — указателей роли и указателей связи. Все слова в пределах предложения (за исключением служебных слов) нумеруются подряд слева направо, и порядковый номер слова используется в качестве указателя связи (УС) к этому слову от другого слова того же предложения. Смысловые отношения также перенумерованы, и эти номера используются в качестве указателей роли

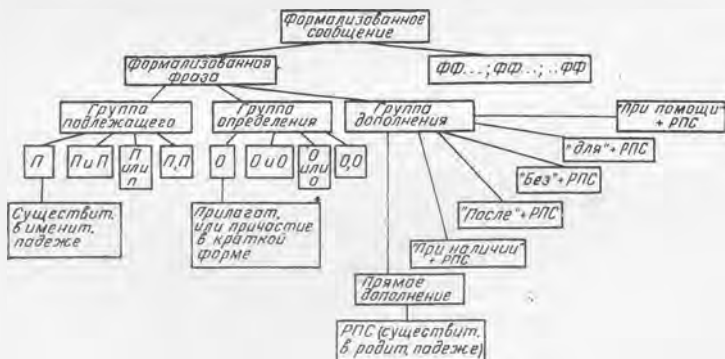


Схема 6. Структура языка формализованных фраз.

¹ В этой структуре не показано, что группы определения могут также относиться к подлежащим и дополнениям и что дополнения могут относиться также к дополнениям.

(УР). Таблица кодирования смысловых отношений приведена ниже (табл. 8). Исключениями являются обозначения подлежащего ($УР=0$, $УС=0$) и союзов «и», «или». Эти исключения вызваны стремлением ограничиться при кодировании одноразрядными числами (восьмеричными цифрами).

Т а б л и ц а 8

Кодирование смысловых отношений

Предикат	Содержание предиката	УР	УС	Номер дескриптора, у которого стоят УР, УС
$P_1(x_i)$	x_i —подлежащее	0	0	i
$P_2(x_i)$	x_i —однородное подлежащее	5	0	i
$P_3(x_i, y_j)$	x_i —определение для y_j	0	j	i
$P_4(x_i, y_j)$	y_j —прямое дополнение x_i	2	i	j
$P_5(x_i, y_j)$	y_j —однородный член, связанный с x_i союзом „и“	5	i	j
$P_6(x_i, y_j)$	y_j —однородный член, связанный с x_i союзом „или“	i	j	j
$P_7(x_i, y_j)$	x_i —предназначен для y_j	4	i	j
$P_8(x_i, y_j)$	x_i —без y_j	3	i	j
$P_9(x_i, y_j)$	x_i —при помощи y_j	1	i	j
$P_{10}(3x_i, y_j)$	x_i —после y_j	6	i	j
$P_{11}(x_i, y_j)$	x_i —при наличии y_j	7	i	j

Примечание: i, j—порядковые номера дескрипторов во фразе.

Союз «и» указывается постановкой $УР=5$ и $УС$, равного номеру (i) данного слова, союз «или» обозначается с помощью $УР=i$ (номеру того слова, к которому направлена связь) и $УС=j$ (номеру данного слова). Заметим, что $УР$ и $УС$ записываются перед кодом того дескриптора, к которому они относятся.

Формализованные фразы при вводе в ИПС подвергаются автоматическому анализу и кодированию, в результате чего получают машинные кодированные записи этих фраз (в основном машинные ПОД). В них дескрипторы (термины) заменены кодами дескрипторов (КД), а смысловые отношения представлены кодами указателей ролей ($УР$) и указателей связи ($УС$). При выдаче из ИПС эти машинные записи декодируются и

из них синтезируются формализованные фразы. Ниже будут рассмотрены алгоритмы синтаксического анализа и синтеза формализованных фраз. Так как в этих алгоритмах используется машинный грамматический словарь, то предварительно рассмотрим кратко один из вариантов такого словаря.

Организация словаря ИПС на ЭВМ «Минск-32».

В рассматриваемой ИПС используются три вида машинного словаря: кодовый, алфавитный и иерархический.

В начале массива словаря размещается так называемый каталог, представляющий собой адресный отсылочный список для указания сегментов алфавитных указателей, начинающихся с определенного символа (буквы).

Кодовая часть машинного словаря представляется областью кодовых указателей, в которой для каждого дескриптора отведена одна ячейка. Адрес этой ячейки численно равен коду дескриптора (КД); содержимым ячейки является так называемый адрес связи, т. е. адрес, указывающий начало соответствующей лексемы, т. е. буквенного представления основы дескриптора, и трех окончаний (именительного падежа, родительного падежа и формы сокращенного прилагательного). Эта ячейка называется ко д о в ы м у к а з а т е л е м.

Затем следует область алфавитного словаря, содержащая так называемые алфавитные указатели. Каждый указатель соответствует одному дескриптору и размещается в двух последовательно расположенных ячейках. В первой ячейке находятся: признак вида дескриптора (простой или сложный), число букв в основе, количество букв в окончании родительного падежа, количество букв в окончании сокращенной формы прилагательного. В остающихся разрядах этой ячейки размещается адрес связи, т. е. отсылка к соответствующей лексеме. Во второй ячейке алфавитного указателя размещается свертка дескриптора, т. е. условный код, полученный суммированием кодов букв дескриптора. Для любого слова можно произвести заданные арифметические действия с кодами букв (складывать их, перемножать и др.) и получить в результате какое-то одно число. Такое число, полученное по определенным правилам, называется сверткой.

Алфавитные указатели размещаются по сегментам в соответствии с первыми буквами их дескрипторов — каждой букве алфавита, которая может стоять в начале сло-

ва (кроме «ы», «ъ», «ь»), отводится один сегмент. Эти сегменты имеют, естественно, разные размеры, так как в словаре количество слов, начинающихся на разные буквы, различно. Внутри каждого сегмента алфавитные указатели могут идти в любом порядке; важно, чтобы все они соответствовали дескрипторам, начинающимся с одной буквы.

Код свертки основы дескриптора занимает 36 двоичных разрядов и получается путем суммирования со сдвигом влево семиразрядных кодов символов основы (без учета символов окончаний). При поиске заданного термина в алфавитном словаре сначала производится обращение к каталогу, в котором по коду символа находится ячейка с отсылкой к сегменту алфавитных указателей, содержащему все дескрипторы, начинающиеся на данный символ. Этот сегмент переписывается в оперативную память с МЛ, и путем сплошного просмотра и сравнения кодов сверток заданного дескриптора и дескрипторов, имеющих в сегменте, производится поиск заданного дескриптора. Просмотр сегмента начинается с первой пары ячеек. У опознаваемого понятия отделяется заданное число букв и по ним получается свертка, которая сравнивается со сверткой, находящейся во второй ячейке пары. Если имеет место совпадение, то по отсылке, находящейся в первой ячейке пары, происходит переход к лексическому представлению этого дескриптора.

Если свертки не совпали, то происходит переход к следующей паре ячеек сегмента и т. д.

Лексическая часть содержит лексемы, т. е. текстовые (словесные) представления дескрипторов в трех их формах, и является общей как для кодового, так и для алфавитного словарей. К лексеме можно перейти либо от кодового, либо от алфавитного указателей. Лексема дескриптора состоит из четырех частей, разделенных разделителями — знаками тождества. Первая часть — основа дескриптора, вторая часть — окончание именительного падежа, третья часть — окончание родительного падежа и четвертая часть — окончание сокращенной формы прилагательного.

Примеры исходной записи лексем при подготовке их для перфорации и ввода в словарь:

<диагности == ка == ки == чesk. ==.
<ЭКГ == == == ==
<внутриклеточн. == == == ==

В последнем случае основа используется только для образования сокращенного прилагательного и поэтому записывается сразу в нужной форме.

В таком виде словарь готовится для перфорации и ввода в ЭВМ.

Заметим, что при построении описываемого машинного словаря такие термины, как основа и окончание, используются условно, только с точки зрения образования трех указанных форм дескриптора. В частности, если используется сложный дескриптор, состоящий из нескольких слов, то основой его может считаться основа первого слова, а «окончаниями» — продолжение этой фразы с изменением в ней одной или нескольких букв. В принципе допустим и такой случай, когда основа имеет нулевое количество букв, а «окончания» представляют собой различные фразы, имеющие смысл дескриптора в именительном падеже, в родительном падеже и форме сокращенного прилагательного.

В подобных словарях могут указываться также для каждого дескриптора отсылки к более общим (родовым) понятиям и к более частным (видовым) понятиям. Эти отсылки образуют иерархическую структуру словаря, которая используется в основном при модификации запросов на поиск.

Синтаксический анализ и кодирование формализованных фраз. Существуют два основных метода синтаксического анализа предложений:

1) метод последовательного анализа, когда для каждого слова последовательно определяются его управляющее слово и тип связи (управления);

2) метод фильтров, когда строится ряд гипотез о расстановке связей между словами фразы и производится проверка пар слов на их сочетаемость, при этом используются синтаксические таблицы. Примеры фильтров будут приведены при рассмотрении методов автоматического индексирования. Для естественных языков удастся лишь выделить группы слов в предложениях, но отношения внутри групп и между группами могут быть неоднозначными. При использовании формализованного языка могут быть также неоднозначности в подчиненности групп «при помощи», «без», «для», «при наличии». Возможность появления синтаксической неопределенности учитывается в алгоритме эквивалентных преобразований запросов в процессе поиска.

В описываемой ИПС на ЭВМ «Минск-32» применен последовательный анализ формализованных фраз по следующим причинам:

— ограниченный набор смысловых связей позволяет использовать в процессе анализа сравнительно небольшое количество правил;

— необходимость минимизации времени обработки документов и запросов в пакетном режиме требует применения наиболее простого метода анализа;

— возможные неточности определения связей в предложениях могут быть компенсированы введением эквивалентных преобразований фраз в запросах.

С помощью описываемого алгоритма синтаксического анализа за один просмотр фразы определяется ее синтаксическая структура.

Алгоритм анализа использует словарь основ и окончаний слов, а также словарь служебных слов — связок («при помощи», «без» и т. д.).

Алгоритм анализа включает следующие этапы: 1) опознавание главного члена предложения и однородных подлежащих, 2) опознавание прямых дополнений к подлежащим и предыдущим дополнениям, 3) опознавание групп «при помощи», «для», «без», «после», 4) опознавание определений; опознавание однородных членов предложений (кроме подлежащих).

В процессе опознавания используются правила следующего вида. Первым существительным в назывном русском предложении может быть только подлежащее. Если перед ним имеются слова, то они являются определениями к нему. Всем определениям присваивается код указателя роли, равный нулю, код направления связи, равный номеру того слова во фразе, к которому эти определения относятся.

Все встретившиеся во фразе определения помещаются временно в буфер, с тем чтобы уменьшить общее число слов во фразе (значение УС, для которого отведен один разряд. Нумерация слов в предложении идет, исключая определения, служебные слова и предлоги).

В формируемой синтаксической цепочке предложения вместо самих дескрипторов ставятся их текущие адреса в тексте предложения. Любое существительное, встретившееся непосредственно за подлежащим (если не встретились запятая и (или) какой-либо из предлогов или служебное слово), может быть только прямым дополне-

нием к подлежащему. Любое существительное, встретившееся непосредственно за дополнением, может быть также только дополнением к дополнению. Таким дополнением присписывается код УР «2» и код УС, равный номеру слова, к которому это дополнение относится.

Если перед рассматриваемым словом стоит служебное слово (например, «после») или какой-либо из предлогов, то УР получает код соответствующего вида связи, а УС — значение предыдущего слова (т. е. слова, которому подчинена эта группа). При наличии в тексте союзов «и», «или», запятой производится анализ (с помощью словаря основ и окончаний), какие классы слов (существительные или прилагательные) — подлежащие, дополнения или определения) соединяются этими союзами. Все они получают одинаковое значение УС и УР.

После обработки предложения все определения (их адреса) из буфера записываются в конец синтаксической цепочки и соединяются с помощью УР и УС с соответствующими словами. Например, в результате синтаксического анализа следующей формализованной фразы:

Методика определения лучевой дозы для диагностики и лечения злокачественной опухоли; максимальная погрешность расчета дозы.

получится следующая кодированная запись (табл. 9).

Таблица 9

Пример кодирования фразы

УР	УС	Адрес лексем	Лексема	УР	УС	Адрес лексем	Лексема
0	0	КД1	Методика	0	6	КД8	Злокачественной
2	1	КД2	Определения	0	0	КД9	Погрешность
2	2	КД3	Доза				Расчета
4	3	КД4	Диагностики	2	1	КД10	Дозы
5	4	КД5	Лечение	2	2	КД3	Максимальная
2	5	КД6	Опухоли	0	1	КД11	ная
0	3	КД7	Лучевой				

Структурная схема алгоритма синтаксического анализа и кодирования формализованных фраз приведена на схеме 7.

Синтез формализованных фраз. Алгоритм синтеза включает три основных этапа: 1) построение цепного

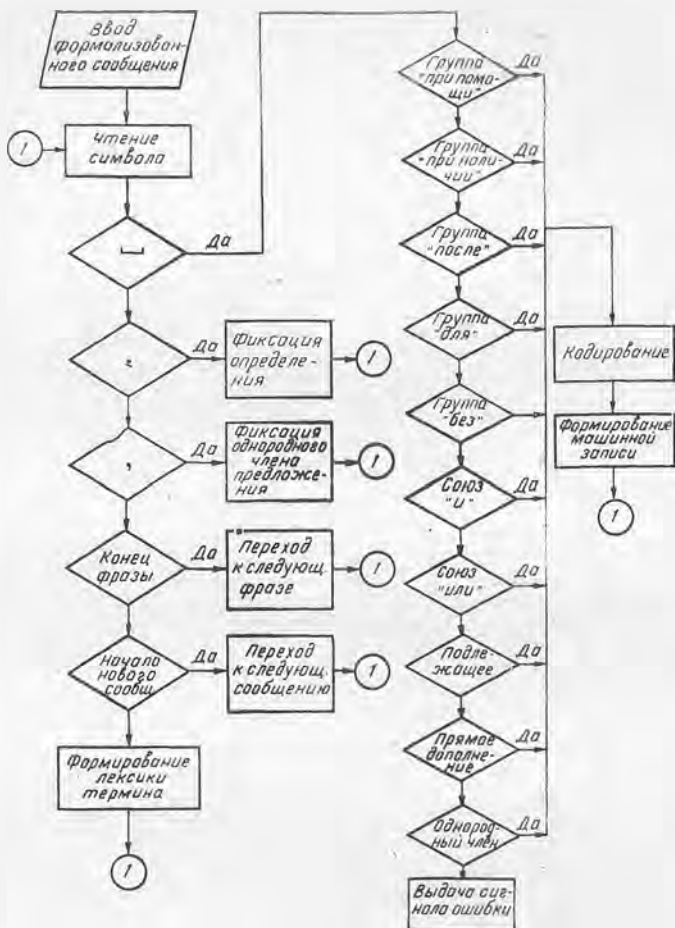


Схема 7. Синтаксический анализ и кодирование формализованных фраз.

списка следования дескрипторов в предложении; 2) поиск по кодам дескрипторов и считывание лексем в ОЗУ из машинного словаря; 3) замена кодов смысловых связей и формирование предложения.

Адреса связи в цепном списке предложения определяются на основе значений указателей связи УС с учетом указателей ролей УР. Значения УР определяют также, какое служебное слово или предлог должны быть

включены в соответствующие места цепного списка и в какой словоформе (из трех возможных) следует поставить лексику данного дескриптора в формулируемое предложение. Последовательность работы данного алгоритма синтеза формализованных фраз проиллюстрируем на примере декодирования того же предложения, которое было использовано ранее в примере синтаксического анализа.

Методика определения лучевой дозы для диагностики и лечения злокачественной опухоли, максимальная погрешность расчета дозы.

Исходный для декодирования машинный ПОД приведен в табл. 10. В нем опущены несущественные для данного процесса части (единый инвентарный номер документа, коды рубрик и подрубрик, последняя ячейка с позиционной строкой, адреса связи дескрипторных цепных списков).

Т а б л и ц а 10

Исходные данные для декодирования фразы

Адрес ячейки	Адрес связи	УР	УС	Коды дескрипторов	Дескрипторы
К+4	+1	0	0	КД1	Методика
К+5	+1	2	1	КД2	Определение
К+6	+1	2	2	КД3	Доза
К+7	+1	4	3	КД4	Диагностика
К+10	+1	5	4	КД5	Лечение
К+11	+1	2	5	КД6	Опухоль
К+12	+1	0	3	КД7	Лучевая
К+13	+1	0	6	КД8	Злокачественная
К+14	+1	0	0	КД9	Погрешность
К+15	+1	2	1	КД10	Расчет
К+16	+1	2	2	КД3	Доза
К+17	+1	0	1	КД11	Максимум

В предпоследней графе условно показаны коды дескрипторов (КД_і), а в последней — соответствующие им дескрипторы. Процесс декодирования начинается с того, что производится обращение к словарю по кодам дескрипторов (КД_і) и считывание из области кодовых указателей этого словаря адресов лексем (кодовых указателей лексем). Затем производится сортировка этих адресов по возрастанию номеров зон на МЛ, при этом получается рабочий массив (табл. 11).

Результат сортировки лексем по номерам зон

Адрес ячейки	Номер зоны	Адрес в зоне	Пояснения
Л+1	НЗ ₁	A _j	Методика
Л+2	НЗ ₁₊₁	A _{j+1}	Определение
Л+3	НЗ ₁₊₂	A _{j+2}	Диагностика
Л+4	НЗ ₁₊₃	A _{j+3}	Лечение
Л+5	НЗ ₁₊₄	A _{j+4}	Лучевая
Л+6	НЗ ₁₊₅	A _{j+5}	Доза
Л+7	НЗ ₁₊₆	A _{j+6}	Расчет
Л+10	НЗ ₁₊₇	A _{j+7}	Опухоль
Л+11	НЗ ₁₊₈	A _{j+8}	Максимум
Л+12	НЗ ₁₊₉	A _{j+9}	Погрешность
Л+13	НЗ ₁₊₁₀	A _{j+10}	Злокачественная

Сортировка по номерам зон необходима для того, чтобы уменьшить время обращения к магнитной ленте.

По адресам лексем (НЗ₁ — номер зоны *i*-й и A_j — адрес в зоне *j*-й) считываются из области лексем машинного словаря, расположенного на МЛ, лексемы в оперативную память. После этого в декодируемый машинный ПОД, находящийся в оперативной памяти, записываются на месте кодов дескрипторов (КД₁) адреса ячеек оперативной памяти, занимаемых соответствующими лексемами (АЛ₁). Эти адреса (АЛ₁) являются как бы временными заместителями самих лексем. После этого на основе указателей связи (с учетом указателей роли) строят цепной список следования слов во фразе, причем в этот список включаются и служебные слова, адреса лексики которых указаны в ячейках с адресами А+1, ..., А+5 (табл. 12).

Адреса связи цепного списка предложения записываются в машинный ПОД на места адресов связи дескрипторных цепных списков, которые используются только при поиске и в данном случае не нужны.

На основе преобразованного таким образом машинного ПОД производится окончательное формирование и выдача на печать (если требуется) формализованного предложения.

Двигаясь по цепному списку предложения (начиная с фиксированной ячейки А+0), выбираем адреса лексем, переписанных уже в оперативную память (АЛ₁), по

Цепной список следования слов во фразе

Адрес ячейки	Адрес связи	УР	УС	Адрес лексемы	Пояснения
К+4	К+5	0	0	АЛ ₁	Методика
К+5	К+12	2	1	АЛ ₂	Определение
К+6	А+1	2	2	АЛ ₃	Доза
К+7	А+2	4	3	АЛ ₄	Диагностика
К+10	К+13	5	4	АЛ ₅	Лечение
К+11	А+3	2	5	АЛ ₆	Опухоль
К+12	К+6	0	3	АЛ ₇	Лучевая
К+13	К+11	0	6	АЛ ₈	Злокачественная
К+14	К+15	0	0	АЛ ₉	Погрешность
К+15	К+16	2	1	АЛ ₁₀	Расчет
К+16	А+4	2	2	АЛ ₃	Доза
К+17	К+14	0	1	АЛ ₁₁	Максимум
А+0	К+4			Начало цепного списка	
А+1	К+7			Адрес „для“	
А+2	К+10			" (и)	
А+3	К+17			" (;)	
А+4	А+5			" (.)	
А+5	КС			Конец цепного списка	

коду указателя роли (УР) определяем нужную словоформу и вставляем ее в данное место предложения, которое записывается в ОЗУ в выходной буфер. Укрупненный алгоритм декодирования и синтеза фраз показан на схеме 8.

Таким образом, синтаксический анализ позволяет получить для заданной формализованной фразы ее машинную запись с помощью КД, УР и УС. Такая запись представляет собой по существу дерево, в вершинах которого стоят КД, направления ветвей и виды связей определяются УС и УР. Описанный выше способ синтаксического анализа является последовательным в том смысле, что слова фразы рассматриваются поочередно и к каждому слову применяется ряд правил анализа. Эти правила являются своего рода «фильтрами», и описанный способ можно считать комбинацией последовательного способа и способа фильтров. Подобные способы применяются при анализе и трансляции программ с алгоритмических языков на машинные. Такие трансляторы принято называть блочными, так как они состоят из блоков, реализующих отдельные виды проверок и трансляции.

Существует другой, так называемый синтаксически направляемый способ, при котором анализ фраз входного языка осуществляется на основе описания синтаксиса этого языка. При этом алгоритм и программа синтаксического анализа не зависят от конкретного входного языка:

структура этого языка задается таблицей металингвистических правил (определений). Одна и та же программа анализа может быть применена для обработки различных входных языков при условии, что для каждого из них задана таблица металингвистических правил. Такие правила строятся так же, как описано выше. Для ускорения процесса просмотра таблицы эти правила располагаются в порядке последовательного укрупнения определяемых ими понятий, причем первые правила таблицы определяют различные классы основных элементов (разде-



Схема 8. Декодирование и синтез формализованных фраз.

лители, служебные слова, дескрипторы и др.). Следующие правила определяют различные классы простейших конструкций, которые могут строиться из основных элементов. Затем определяются сложные конструкции и так далее до получения определения законченной конструкции языка (сообщение, программа). Процесс анализа состоит в последовательной выборке элементов входной (анализируемой) строки и формировании из этих элементов синтаксических конструкций.

В результате анализа получается синтаксическое дерево (граф), представляющее собой структуру фразы или сообщения.

Организация поиска информации. Для ускорения поиска информации целесообразно машинные информационные массивы ИПС разделить на два вида: поисковые образы документов (ПОД), образующие поисковый массив (ПМ), и справки (формуляры) документов, образующие справочный массив (СМ). Это разделение носит технологический характер: в поисковом массиве информация представлена в формализованном виде на ИПЯ и предназначена для машинного анализа и отбора; в справочном массиве информация представлена на естественном языке и предназначена для выдачи человеку.

Поисковый образ документа имеет единый инвентарный номер документа (ЕИН) и содержит набор формализованных фраз, определяющих суть вопроса, рассматриваемого в документе. Справка или формуляр имеет ЕИН (тот же, что и соответствующий ПОД) и содержит название документа, фамилии авторов, краткую аннотацию и библиографические данные (год издания, название источника, страницы и др.). Справка документа выдается после того, как в результате поиска в поисковом массиве определен ЕИН документа. Справка может содержать подробные фактические сведения по вопросу, определенному ПОД.

Так как поисковый массив используется значительно чаще, то его целесообразно хранить в запоминающих устройствах с прямым доступом — на магнитных дисках. Учитывая меньшую частоту обращения к справочному массиву, а также его больший объем, целесообразно его хранить на магнитных лентах. Создание отдельного массива ПОД приводит к более плотной записи ПОД. При этом переписываемые в процессе поиска с МД в ОЗУ отдельные зоны или блоки поискового массива содержат

большее число записей (ПОД) по сравнению со случаем совместного хранения ПОД и справок в одном массиве. Уплотнение поискового массива сокращает число обращений к внешним накопителям, что существенно увеличивает скорость поиска.

Наряду с этим наличие отдельного массива справок упрощает процесс пополнения и корректировки справок, что часто бывает необходимо особенно в фактографических ИПС.

Поиск документов в общем случае строится в виде нескольких последовательных этапов. Такой подход целесообразен при достаточно большом объеме словаря дескрипторов и большом объеме фонда документов.

Для осуществления первого этапа поиска весь фонд документов делится на ряд тематических подмассивов, соответствующих различным разделам документации. Первый этап поиска заключается в выборе в соответствии с запросом таких подмассивов, в которых могут содержаться искомые документы.

Второй этап поиска состоит в более детальном просмотре отобранных подмассивов и поиске в них нужных документов с использованием достаточно простого критерия соответствия документа запросу. Например, удовлетворяющими условиями поиска считаются документы, в поисковые образы которых входят все дескрипторы, заданные ПОЗ.

Третий этап поиска сводится к логическому анализу поисковых образов, найденных на втором этапе, при этом анализируются смысловые связи между дескрипторами поискового образа документа, которые сравниваются со смысловыми связями между дескрипторами запроса. Третий этап осуществляется только в том случае, если на втором этапе поиска было найдено достаточно большое число документов. Если исходный массив документов не велик (меньше 10 000), что может быть на начальной стадии комплектования ИПС документами, то можно не проводить и первый этап поиска и не делать деления массива на подмассивы. В этом случае выполняется сразу второй этап поиска, который является, таким образом, основным.

Можно представить себе четвертый, еще более детальный этап отбора документов, когда для анализа привлекаются различного рода эквивалентные преобразования поисковых образов документов и запросов, учи-

тываются статистические характеристики дескрипторов, запросов и документов и т. д.

Первый этап поиска — предварительный отбор тематических подмассивов, в которых могут находиться требуемые документы, — достаточно просто реализуется с помощью логических матриц, фиксирующих использование дескрипторов в различных тематических поисковых подмассивах. Для каждого дескриптора отводится строка матрицы, в которой фиксируется использование этого дескриптора в разных тематических подмассивах. Каждому из разрядов этой строки поставлен в соответствии один тематический поисковый подмассив; если в этом подмассиве имеется хотя бы один документ, в ПОД которого фигурирует данный дескриптор, то в этом разряде ставится единица, в противном случае — ноль. На первом этапе поиска для каждого поискового образа запроса выбираются строки логической матрицы, соответствующие дескрипторам запроса, и путем их логического умножения определяются тематические подмассивы, в которых могут быть документы, отвечающие данному запросу. Так как поиск обычно идет одновременно по большому числу запросов (20—30), то такая операция проводится сразу для всех ПОЗ. Затем осуществляется сортировка ПОЗ по тематическим подмассивам. После этого поочередно проводится поиск внутри каждого тематического подмассива, отобранного на первом этапе, при этом поиск идет только по тем ПОЗ, которые относятся к данному тематическому поисковому подмассиву.

Для машинной реализации второго этапа поиска весьма эффективным оказывается ассоциативный узловой метод, сочетающий преимущества прямого и инверсного способов организации поисковых массивов. Поисковый образ каждого документа при этом способе представляется ассоциативным узлом, в котором каждому дескриптору ПОД соответствует одно списковое слово, состоящее из кода дескриптора, кода смысловой связи и адреса связи. Адреса связи объединяют в единые цепные списки, документы, обладающие одинаковыми дескрипторами. Каждому дескриптору соответствует один цепной список. Поиск заданных документов ведется путем прослеживания одного цепного списка, соответствующего тому дескриптору из числа имеющихся в ПОЗ, который содержит наименьшее число документов в своей цепочке.

Благодаря выбору такого дескриптора сокращается время поиска.

Механизм поиска заключается в том, что при обращении к очередному ассоциативному узлу выбранного цепного списка проверяется наличие в этом узле всех остальных дескрипторов, содержащихся в запросе. Если в узле содержатся не все дескрипторы запроса, следовательно, документ не отвечает принятому критерию соответствия. В этом случае выполняется переход к следующему ассоциативному узлу этого списка, где эта процедура повторяется. Если узел содержит все дескрипторы запроса, до документ отвечает запросу, его ЕИН запоминается и выполняется переход к очередному узлу.

В случае, если на этом этапе поиска будет обнаружено слишком большое количество документов и есть необходимость сократить их число, производится третий этап поиска — сплошной просмотр ПОД, найденных на втором этапе, и отбор среди них документов, наиболее полно отвечающих содержанию запроса. Для более точного отбора используются большее количество дескрипторов ПОЗ и смысловые связи (указатели отношений между дескрипторами), а также более тонкие критерии смыслового соответствия.

После того как в поисковом массиве найдены нужные единые номера документов, в справочном массиве отыскиваются соответствующие им справки документов, которые выдаются (частично или полностью) заказчику.

Для поиска документов запросы задаются на том же формализованном языке, на котором записаны ПОД, причем ЭВМ по специальной программе диалога может помочь неопытному абоненту правильно сформулировать запрос. Запрос должен представлять собой набор формализованных фраз, характеризующих тематику искомых документов.

Для повышения полноты и точности поиска документов может быть использован модуль программы эквивалентных преобразований фраз запросов. Он позволяет выразить представленный запрос в нескольких модификациях. По каждой из модификаций могут производиться отдельный поиск и выдача документов. В число эквивалентных преобразований входят, например, замена прилагательного перед данным существительным существительным в родительном падеже, стоящим после данного существительного.

Заказчик-абонент, задавая запрос, может указать в нем форму выдачи, например: выдавать весь документ или его определенную часть, выдавать все найденные документы или только их количество, выдавать не больше, чем заданное количество документов, и т. д.

Помимо поиска и выдачи информации, документально-фактографическая ИПС обеспечивает возможность логической и арифметической обработки отобранных данных (упорядочение, суммирование, определение средних, максимальных и минимальных значений и др.). Если ИПС должна производить логическую обработку отдельных пунктов справок, то информация в них должна быть представлена в виде формализованных фраз.

Существенное повышение точности и полноты поиска обеспечивается за счет использования в процессе поиска и в процессе индексации документов родо-видовых и ассоциативных отношений между дескрипторами. Для реализации такой возможности строится так называемый иерархический словарь, позволяющий для данного дескриптора находить более общие (родовые) и более частные (видовые) понятия, а также понятия, связанные с ним тематической близостью (ассоциативной связью). Примером ассоциативной связи может служить связь понятий «ЭВМ» и «программирование». В иерархическом словаре понятий используются только коды дескрипторов (КД), которые связываются между собой с помощью адресных отсылок и образуют набор деревьев понятий. Вершиной каждого дерева является наиболее общий дескриптор данной предметной области; от него отходит подписок более частных дескрипторов, от каждого из которых отходят подписки еще более частных дескрипторов и т. д. В этих списках и подписках у каждого члена должен быть обратный адрес связи, отсылающий к более общему понятию, а также адрес связи (или несколько адресов связи) к понятиям (дескрипторам), связанным с данным дескриптором ассоциативными отношениями¹.

В зависимости от состава запроса и поискового массива на запрос абонента ИПС может выдать слишком

¹ В области информационного поиска дескрипторный словарь вместе со словарем, указывающим различного рода взаимосвязи между дескрипторами, называется тезаурусом. В переводе с греческого тезаурус — «сокровище».

мало или слишком много (по мнению абонента) документов; могут быть выданы документы, не вполне отвечающие его требованиям. В этом случае ЭВМ по специальной программе (и указанию абонента или оператора) может обратиться к иерархическому словарю и выбрать из него дескрипторы более общие (если нужно увеличить выдачу) или более частные, а также ассоциативные (если нужно уточнить и сократить выдачу), переформулировать запрос и снова произвести поиск. Переформулированный запрос может быть либо показан абоненту для подтверждения перед повторным поиском, либо ЭВМ произведет поиск и выдаст абоненту готовые результаты поиска. При индексации документов часто бывают случаи, когда индексы-люди используют в ПОД термины, отсутствующие в словаре. В этом случае ЭВМ после поиска в словаре печатает отсутствующие термины, с тем чтобы индексатор (человек) заменил их другими эквивалентными терминами или более общими или частными терминами.

Общая структура ИПС и управление ее работой. Важными принципами построения ИПС являются модульность построения программ и использование специализированной операционной системы для управления их работой. Программные модули — это автономные участки программы, выполняющие законченные функции по формированию поискового, справочного и словарного массивов, осуществлению поиска, преобразованию поисковых образов документов и запросов, формированию и выдаче ответов и т. д. Вопрос о рациональном разделении сложной программы на отдельные модули является одним из важных вопросов теории программирования. Существует ряд методик решения подобных задач, основанных на теории массового обслуживания и статистическом моделировании процесса функционирования сложной программы. Особенно важна оптимизация структуры программ, работающих в режиме реального времени, т. е. осуществляющих ввод новых данных и выдачу ответов на запросы в моменты их поступления.

Практически часто выбирают варианты выделения модулей программ, исходя из критерия их функциональной законченности и автономности.

Управление работой отдельных модулей программ осуществляется специальной программой, которая является специализированной операционной системой ИПС.

Эта программа анализирует вид входной информации (запрос на поиск, информация для пополнения словарей, ввод новых документов и др.) и автоматически формирует нужную последовательность программных модулей, и задает им требуемые режимы работы, т. е. выполняет функции диспетчера ИПС. Например, если идет ввод новых документов, то сначала работает программный модуль инвентаризации, который присваивает очередному документу единый инвентарный номер и осуществляет контроль и перекодирование поискового образа документа, затем работает модуль грамматического анализа и кодирования дескрипторов, после чего работают модули формирования поискового массива и справочного массива.

Стратегия управления работой ИПС, реализованная в диспетчере ИПС, должна сочетать принцип пакетной обработки запросов с принципом приоритетного обслуживания абонентов. Это означает, что при отсутствии срочных запросов (первого приоритета) ИПС начинает работу по поиску информации только после того, как накопится определенное количество запросов; до этого ИПС будет выполнять другие работы (ввод новых документов, пополнение словаря и др.).

На схеме 9 показана в качестве примера функциональная схема документально-фактографической информационно-поисковой системы (ДФИПС), разработанной в Московском электротехническом институте связи. В этой ИПС реализованы такие основные принципы построения, как ассоциативно-адресный способ организации поискового массива, использование отдельного справочного массива с фиксированными значениями пунктов справок-формуляторов, использование формализованных фраз при выдаче на печать и при проведении поиска, автоматическое кодирование формализованных фраз при вводе информации в ИПС и др.

Блоки, показанные в верхней части схемы, являются типовыми расчетными задачами для АСУ административного назначения (составление отчетов, учет кадров, планирование и т. д.). Исходные данные для решения этих задач хранятся в ИПС, которая находит их и выдает в требуемые моменты для использования. Результаты решения задач вводятся в качестве новой информации в ИПС, которая хранит их для последующего использования и выдачи заказчикам.

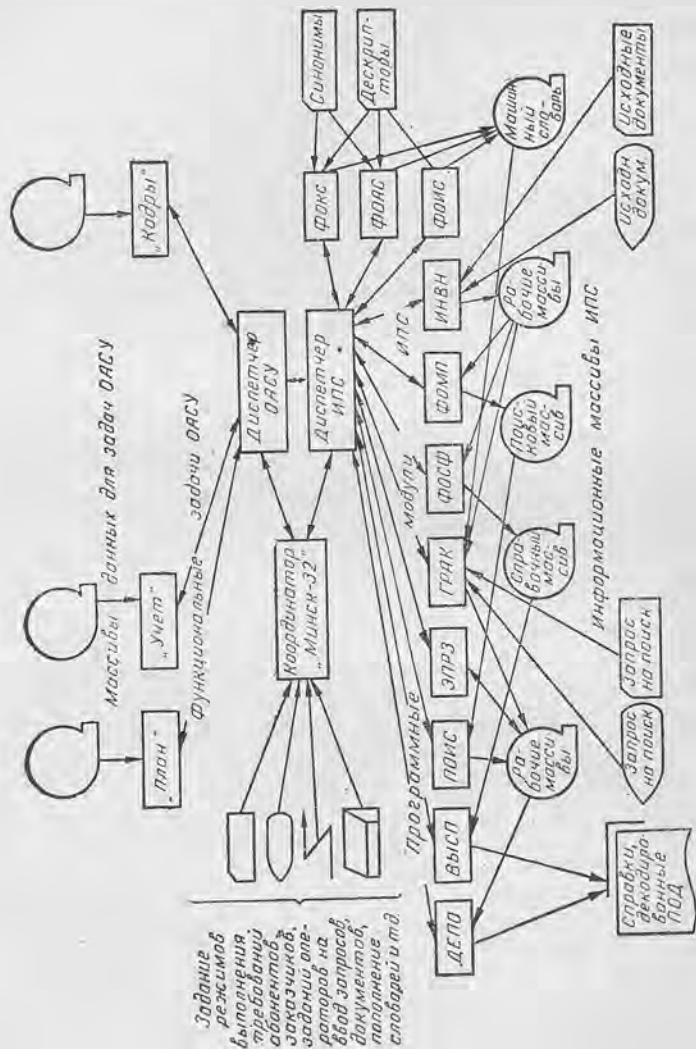


Схема 9. Функциональная схема ДФИСС.

Управление работой отдельных программных модулей ИПС, а также порядком решения расчетных задач АСУ осуществляется специализированной операционной системой АСУ и ИПС, которая показана в средней части схемы. Эта операционная система сопрягается со стандартной операционной системой ЭВМ «Минск-32» и работает во взаимодействии с ней.

В нижней части схемы представлены блоки (прямоугольники), соответствующие отдельным программным модулям ИПС:

- ГРАК — грамматический анализ фраз и их кодирование
- ВЫСП — выдача справок (целиком и по пунктам)
- ДЕПО — декодирование ПОД и формирование фраз
- ФОМП — формирование поискового массива
- ФОСФ — формирование справочного фонда
- ПОИС — программа поиска
- ЭПРЗ — программа эквивалентных преобразований поисковых образов
- ФОАС — формирование алфавитного словаря
- ФОКС — формирование кодового словаря
- ФОИС — формирование иерархического словаря
- ИНВИ — инвентаризация документов. Программа производит автоматическое присвоение ЕИН и контроль исходных документов.

Нижний ряд символических изображений магнитных лент представляет состав основных информационных массивов ИПС. На этой же схеме с помощью условных фигур показаны различные способы ввода и выдачи информации (выдача на печать, на экраны дисплеев, ввод с перфокарт, с клавиатур дисплеев, непосредственный дистанционный ввод вопросов с каналов связи).

3. Запись ПОД с использованием семантико-синтаксических категорий и линейных кодов

Информационно-поисковый язык для ИПС, применяемых в области медицины, имеет некоторые особенности.

Рассмотрим две особенности ИПС, используемой для накопления и поиска информации о химических веществах и медикаментах: применение семантико-синтаксических категорий и запись структурных химических формул.

Применение семантико-синтаксических кодов, сущность которых будет ясна из дальнейшего, вызвано мно-

гимий причинами, одну из которых мы приведем в качестве примера. Химические вещества, описываемые в документах, могут выступать в разных ролях: в качестве исходного продукта, катализатора, конечного или побочного продукта и т. д. Если отыскивать документы только по названию вещества, то среди отысканных будет большое количество не соответствующих необходимому аспекту запроса. Отсеять ненужные документы можно, если в поисковом образе документа и запросе указывать необходимую синтаксическую и семантическую категорию (ССК), т. е. роль, выполняемую данным дескриптором. Для выявления необходимых категорий на основе статистического анализа совместной встречаемости дескрипторов в ПОД и ПОЗ строится классификация дескрипторов. В этой классификации дескрипторы распределяют по категориям в зависимости от той семантической и синтаксической функции, которую они выполняют в ПОД и ПОЗ.

Статистический анализ проводился на основе результатов поисков по ПОЗ, составленным без учета синтаксических отношений между дескрипторами. При этом ставилась задача выяснить следующие вопросы: а) типовые случаи ложных сочетаний дескрипторов, позволяющие сформулировать категории дескрипторов, допускающие такие сочетания; б) необходимые виды логических связей, которые дают возможность устранить такие сочетания; в) типовые синтагмы (словосочетания).

Например, в ИПС по химии и химической технологии, разработанной в Научно-исследовательском институте технико-экономических исследований химической промышленности Министерства химической промышленности (НИИТЭХИМ МХП) совместно с Азербайджанским институтом нефти и химии имени Азизбекова выделены следующие 13 категорий:

1. Вещества, материалы, продукты.
2. Функциональные определители роли элементов I категории в технологических процессах (форма, агрегатное состояние и т. п.).
3. Вещественные определители роли.
4. Свойства и другие характеристики элементов 1, 2, 3 категорий, явления.
5. Процессы, связанные с веществами.
6. Параметры процессов 5 категории.
7. Оборудование и приборы.

8. Методы анализа (химические, физические, физико-химические и др.).

9. Характеристики и параметры оборудования, приборов и методов анализа.

10. Процессы, связанные с оборудованием.

11. Отрасли промышленности и производства.

12. Процессы и действия, связанные с п. 11.

13. Библиографические данные.

В этой же ИПС применяются следующие 12 видов логико-семантических отношений:

1. Иметь функциональный определитель роли.

2. Быть функциональным определителем роли.

3. Иметь вещественный определитель роли.

4. Быть вещественным определителем роли.

5. Иметь характеристику (свойство или параметр).

6. Быть характеристикой (свойством или параметром).

7. Подвергаться процессу, действию.

8. Быть процессом.

9. Иметь место.

10. Быть местом.

11. При помощи.

12. Предназначаться для.

Поисковый образ документов в этой ИПС представляется в виде матрицы, у которой столбцы соответствуют синтаксическим категориям, а строки — отдельным формальным фразам — линейным записям дескрипторов, связанным между собой логико-семантическими отношениями. Эти отношения определяют возможность сочетания дескрипторов фразы в заданной последовательности.

Использование системы семантических категорий позволяет в словаре дескрипторов хранить только корневые (обобщенные) понятия и иметь только для них отдельные коды дескрипторов. Помещая это обобщенное понятие в ту или иную категорию, можно получить из него различные конкретные модификации. Например, для понятий «абсорбенты», «абсорбция», «абсорбер» будет существовать один код дескриптора. Этот код, будучи отнесенным к 1 категории, даст понятие «абсорбенты», к 5 категории — «абсорбция», к 7 категории — «абсорбер».

Подобная система синтаксических категорий может быть предложена для медицины, социологии, экономики и других областей знаний.

Язык линейных кодов для записи химических соединений. В ИПС, применяемых в химии, медицине, биологии и т. п., часто возникает необходимость записи в ПОД или фактографических справках названий или структурных формул химических соединений. Известно огромное количество (около 4 млн.) различных химических соединений, и число их увеличивается ежегодно приблизительно на 100 тыс. В связи с этим не представляется возможным включать в словарь дескрипторов названия химических соединений даже для сравнительно узких классов веществ.

С этой целью разработан ряд способов записи структурных формул химических соединений в виде так называемых линейных кодов — линейных последовательностей символов, допускающих ввод их и хранение в памяти ЭВМ, а также поиск и обработку этих линейных кодов.

Система записи структурных формул в виде линейных кодов основана на следующих принципах:

— разделение неорганических соединений и органических соединений на классы и подклассы; символы классов и подклассов указываются в начале линейного кода данного соединения.

— разделение всей формулы соединения на фрагменты и составление линейных кодов фрагментов и кодов связей между фрагментами.

— выделение специальных символов для указания состава и расположения атомов в главных цепях, циклах, заместителях и т. п.:

При поиске соединений по формулам (линейным кодам) применяют два способа:

1) разделение массива на подмассивы, соответствующие классам и подклассам соединений, и поиск внутри подмассивов;

2) применение «фильтров», представляющих собой записи наиболее важных структурных особенностей соединений, и поиск соединений по наличию у них заданных особенностей.

В частности, в качестве первого (грубого) фильтра может выступать символ класса или подкласса соединений.

Описанный ниже язык линейных кодов разработан и практически используется в НИИТЭХИМ МХП СССР.

В указанной ИПС возможен поиск как по дескрип-

торной, так и по структурной части ПОД (по линейным кодам), а также по комбинации этих частей. Поисковые образы документов состоят из двух частей: дескрипторной и кодовой (химической). Эти части хранятся совместно и образуют единый поисковый массив. В данной ИПС практически реализован режим избирательного распределения информации (ежемесячно между 400 абонентами).

При использовании семантико-синтаксических категорий (ССК) упрощается не только строение словаря и уменьшается его объем, но и упрощается поиск ключевых слов, так как при этом достаточно получить только совпадение заданного количества (например, 6) первых букв слов. Неточности, связанные с расхождениями в суффиксах и окончаниях, будут устраняться за счет влияния ССК.

При составлении ПОД во многих случаях целесообразно встречающиеся в документе сложные слова и словосочетания разделять на составные элементы. Особенно это относится к префиксам («вибро», «гидро» и т. д.). В словаре подобные элементы нужно представлять в качестве отдельных дескрипторов, объединяя их со словами с такой же основой. Это дает возможность комбинировать подобные элементы с другими терминами и получать большое количество сложных понятий без увеличения объема словаря. В процессе поиска это позволяет осуществлять иерархические переходы к более общим понятиям (уровень «выше»).

4. Информационно-поисковая система, работающая на естественном языке

Принципы построения. В описываемой ИПС естественный язык научных текстов используется для представления в памяти ЭВМ как документов, так и запросов на поиск этих документов. В этой ИПС отсутствует разделение информационных массивов на два вида (поисковый и справочный), а используется единый массив машинных записей документов. Каждый машинный документ (машинная запись) состоит из набора пунктов фиксированного значения (смысла)¹. Например, пункт № 1

¹ С целью унификации записи сведений о документах в нашей стране и за рубежом разрабатываются так называемые предмашинные и коммуникативные форматы, которыми регламентируется порядок и содержание записи всех вводимых в ЭВМ сведений о

означает заглавие документа, пункт № 2 — фамилии авторов, пункт № 3 — название журнала (для статей), пункт № 4 — краткую аннотацию статьи, пункт № 5 — название издательства (для монографий), пункт № 6 — год издания и т. д. Помимо таких общих библиографических пунктов, в записях могут быть и конкретные тематические пункты. Не обязательно все пункты должны присутствовать в каждой записи. Общий набор различных пунктов может составлять несколько сотен. Из этого набора пунктов в каждой конкретной машинной записи используется в среднем 10—15 пунктов. Для обеспечения эффективного взаимодействия ИПС с пользователями предусматривается иерархическая система организации пунктов: отдельные пункты являются детализацией (подпунктами) других пунктов, что отражается в их десятичной нумерации. Например, пункты с номерами 41, 42, ..., 49 являются детализацией пункта № 4 и т. д. При использовании ИПС в конкретной области знаний указанная иерархическая система тематических пунктов образует ССК пунктов, образует ССК вопросов, содержащихся в документах, вводимых в ИПС.

Поиск документов может производиться в принципе по любому из пунктов или подпунктов. Для поиска задаются номер пункта и набор признаков — терминов, по которым должен производиться поиск. Критерием выдачи является вхождение терминов запроса в состав терминов соответствующего пункта документа. Если, например, необходимо произвести поиск по тематическим признакам, то для поиска можно выбрать либо пункт № 1 (заглавие), либо пункт № 4 (аннотация). В первом случае поиск будет идти быстрее, но неточно; во втором случае поиск будет идти более точно, но значительно медленнее. Можно комбинировать оба эти способа: сначала произвести поиск по заглавиям, а затем прибегнуть

документе (библиографические данные, аннотация, реферат или поисковый образ документа). Терминология, применяемая авторами ниже, отличается от той, которая принята при описании форматов. В частности, термин «пункт» в качестве названия элемента библиографического описания в форматах не используется.

Поскольку вопрос о форматах является специальным, подробное изложение его в данной книге нецелесообразно. Учитывая это, а также то, что в книге описывается конкретная ИПС, разработанная под руководством одного из авторов (проф. А. И. Китова), мы считаем возможным сохранить терминологию авторов (*Прим. ред.*).

к более точному отбору — по аннотациям. Если нужно произвести поиск по авторам, то задается пункт № 2 и фамилия авторов. Можно аналогичным образом производить поиск по названию издательства, названию журнала и т. д. В частности, пункт с нулевым номером резервируется для обеспечения возможности включения в состав машинного документа специально составленной фразы (фраз), выполняющей роль поискового образа документа. Необходимость в таком пункте может возникнуть в тех случаях, когда поиск по заголовкам или аннотациям документов не обеспечивает нужной полноты и точности. В этом же пункте (нулевом) указывается и единый инвентарный номер документа.

Термины в запросе могут быть связаны тремя видами логических связей: конъюнкцией (пробел или запятая), дизъюнкцией (\vee), отрицанием ($-$). Все признаки запроса, связанные конъюнктивной связью, должны обязательно присутствовать в соответствующем пункте документов. Из терминов запроса, связанных дизъюнктивной связью, достаточно присутствие хотя бы одного. Термины запроса, помеченные отрицанием, должны отсутствовать в соответствующем пункте документа.

Принципиальной особенностью описываемой системы является отсутствие в ней каких-либо машинных словарей-тезаурусов и в связи с этим отсутствие процессов кодирования документов и запросов. Документы, представленные наборами пунктов указанных типов, вводятся в ЭВМ и хранятся в ее памяти в виде фраз естественного языка, перед которыми стоят номера пунктов. Запросы вводятся в виде номера пункта и последовательности слов (признаков), связанных операциями конъюнкции, дизъюнкции и отрицания¹.

Обычно содержание пунктов и подпунктов машинных документов представляется простыми назывными фразами естественного языка, без использования глаголов. В принципе для поиска могут использоваться и

¹ Следует сразу же отметить, что отсутствие тезауруса и возможность непосредственного использования текста документов и запросов без их индексирования хотя и кажутся весьма привлекательными, однако связаны с серьезными недостатками: ИПС, построенная на таком принципе, имеет весьма низкую полноту поиска. В такой ИПС каждый раз при поиске не будут найдены документы, в которых проблема, выраженная в запросе, изложена другими терминами (*прим. ред.*).

обычные фразы с глаголами. Но при поиске глаголы будут пропускаться, так как не будут давать совпадений с дескрипторами запросов. Использование в ИПС фраз естественного языка как для поиска, так и для выдачи справочной информации имеет следующие преимущества: а) исключается процесс индексирования; б) при выдаче на печать сохраняется удобная для человека форма фразы. Однако при этом увеличивается объем общего информационного массива и в связи с этим растут затраты времени на поиск.

Принципиальным вопросом создания любых ИПС, работающих с естественным языком, является выбор алгоритма сопоставления запросов с машинными записями документов и в первую очередь выбор алгоритма отождествления отдельных слов запроса и записи. В описываемой ИПС реализован простой алгоритм сопоставления, в котором не учитываются синтаксические (и тем более семантические) связи между словами, а осуществляется только проверка вхождения слов запроса в записи. При этом учитываются три основные логические операции: «и», «или», «не». Для отождествления слов запроса и записей применяется комбинация двух известных способов морфологического анализа слов («справа налево» — отделение окончаний и «слева направо» — отделение основ). При этом производится отделение окончаний у слов запроса (путем побуквенного анализа справа налево) и побуквенное (слева направо) сравнение оставшихся основ (корней) слов со словами соответствующего пункта машинной записи документа.

Условие совпадения двух дескрипторов (ПОД и ПОЗ) считается выполненным, если основа слова из запроса совпала с левой начальной частью некоторого слова в документе. При этом не рассматривается вопрос о том, что собой представляет оставшаяся (правая) часть данного слова документа. Таким образом, в качестве «основ» и «окончаний» слов могут использоваться довольно произвольные их левые и правые части. В ИПС могут храниться два списка: список окончаний и список суффиксов, и в зависимости от заданного режима будут отделяться более длинные или более короткие части. Алгоритм сравнения допускает также такую возможность, когда отделение (окончаний) в дескрипторах запроса может производить сам автор запроса при его формулировке. При этом слова с отделенными оконча-

ниями записываются с точкой (вместо отделенной части). В этом случае программа не производит отделения окончаний, а использует часть слова, стоящую перед точкой в качестве основы при сравнении. Используя эту возможность, автор запроса может в сомнительных случаях оставлять в качестве основы только несколько начальных букв слова (если, например, слово допускает чередование букв в основе). Алгоритм сравнения позволяет также осуществлять исключение из фраз запроса лишних слов (так называемых паразитных слов — вопросы, рассмотрение и т. п.), а также предлогов, союзов и т. д. Для этого достаточно такие слова в полном виде включить в список «окончаний», и они будут полностью отбрасываться программой отделения окончаний.

Преимуществом указанного метода поиска является отсутствие машинных словарей и процедур кодирования и декодирования поисковых образов документов и запросов.

Примененный способ сопоставления слов запроса и документа (отделение окончаний в запросе и сопоставление слева направо основ слов из запроса со словами в документе) может обеспечить в принципе более высокую точность сопоставления по сравнению со способом независимого кодирования слов запроса и слов документа с помощью машинного тезауруса. В последнем случае процесс отделения окончаний должен выполняться дважды — для документа и для запроса, а в первом случае он выполняется один раз и, естественно, вероятность ошибок при этом резко уменьшается. Следует иметь в виду, что ошибка в определении основы слова в документе и неправильное присвоение ей кода дескриптора (класса эквивалентности) будут приводить к многократным ошибкам сопоставлений запросов и документов. Этот метод не предусматривает кодирования слов документов, и поэтому такие ошибки исключаются.

Недостатками рассмотренного метода является, во-первых, отсутствие машинного устранения синонимии (что достигается в ИПС, использующих тезаурусы) и, во-вторых, увеличенный расход памяти для хранения словесных (а не кодированных) машинных записей документов.

Для учета синонимии необходимо в запросе с помощью дизъюнкции задавать все синонимы каждого тер-

мина. Это можно делать вручную с помощью ручного словаря либо автоматически с помощью специального машинного словаря синонимов. Ясно, что при пользовании ручным словарем для учета синонимии слов при поиске время составления запроса увеличивается, так как автор должен выбрать из словаря все синонимы терминов запроса и включить их в запрос, соединив знаками дизъюнкции. Ручной словарь синонимов должен вестись по существу вручную. Периодически по мере пополнения ИПС документами, содержащими новые слова, эти новые слова выдаются на печать, и специалисты-индексаторы должны анализировать их с точки зрения отнесения их либо к новым словам в словаре, либо к синонимам слов, уже введенных в словарь. Для этой цели должны вестись вручную следующие картотеки слов словаря:

— алфавитная картотека, в которой все слова расположены по алфавиту, каждое слово находится на отдельной карточке,

— предметная систематическая картотека, в которой карточки со словами располагаются группами синонимов. Сами группы располагаются в виде пермутационного указателя по алфавиту, по каждому слову группы путем их последовательной перестановки — пермутации.

Подобные словари (алфавитный и пермутационный) могут периодически выдаваться из ЭВМ по специальной программе. Ясно, что при выпуске пермутационного словаря с помощью ЭВМ само решение вопроса об отнесении того или иного слова к синонимам или к новым словам останется за человеком, так как это решение связано с пониманием смысла слова и его значения в определенной области знаний. Однако ручная работа со словарем синонимов и в связи с этим увеличение времени на составление запроса могут рассматриваться не только как отрицательное, но и как положительное явление, так как в процессе этой работы индексатор более подробно знакомится с общим составом терминов документов, имеющих в ИПС по данному вопросу, более тщательно продумывает запрос и обеспечивает более четкую его постановку.

Уточнение запроса может быть произведено в дальнейшем по результатам поиска, при этом использование естественного языка расширяет возможности диалогово-

го режима работы пользователя с ИПС и приводит к улучшению результатов поиска.

Достоинствами описанной ИПС, работающей на естественном языке, являются также возможность непрерывного совершенствования критериев и стратегии поиска, отсутствие формальных преобразований и кодирования документов при вводе в ИПС и хранение их в естественном виде, готовом для выдачи и восприятия их человеком.

Заметим, что сочетание описанной ранее иерархической структуры (тематических пунктов и подпунктов) машинных записей с представлением их содержания на естественном языке позволяет осуществить также семантико-синтаксическую классификацию терминов, т. е. отнесение терминов к определенным семантико-синтаксическим категориям в зависимости от их роли в типовых текстах данной отрасли знаний. Так как состав терминов и способы их использования в сильной степени зависят от конкретной области знаний, то указанная категоризация терминов должна производиться на основе статистического анализа текстов. При наличии ССК в ИПС, использующих словари, существенно сокращается объем словаря дескрипторов за счет объединения в одну группу эквивалентности всех терминов, имеющих общую основу и выражающих одно общее корневое понятие (например, испытание, испытатель, испытуемый, средство для испытаний, испытательский и т. д.). Все эти термины будут иметь единый код дескриптора. Присоединяя к этому коду код той или иной ССК, можно получить все конкретные значения имеющихся терминов. В ИПС, работающей непосредственно на естественном языке, введение ССК позволит получить более высокую точность опознавания терминов запросов и машинных записей за счет приписывания этим терминам кодов ССК. При этом для повышения степени совпадений терминов-запросов и машинных записей у терминов-запросов могут отделяться не только окончания, но и словообразующие суффиксы. Потеря части смысла этих слов будет компенсироваться наличием ССК у них, а также тем, что при поиске будут использоваться пункты записей, представляющие те же ССК.

Рассмотренные до сих пор вопросы по существу относятся к процессу анализа и отождествления отдельного документа (либо по его заголовку, либо по анно-

тации, по авторам и т. д.). Сам же процесс поиска — отбор документов для такого анализа остался пока вне поля зрения. При небольших массивах можно производить сплошной просмотр и анализ документов описанным методом без предварительного отбора. При больших массивах сплошной просмотр занимает слишком много времени и делается нерациональным. В этом случае необходим какой-то механизм поиска и предварительного отбора документов. Можно использовать ассоциативно-адресные способы поиска, способы логических матриц и др. В частности, достаточно простым и эффективным способом является формирование всего массива документов не сплошным потоком, а по зонам на магнитной ленте, соответствующим тематическим подмассивам. Запись документов в те или иные зоны должна производиться в соответствии с рубриками и подрубриками, присваиваемыми документам авторами или индексаторами. Так как один и тот же документ может относиться к нескольким рубрикам и подрубрикам и количество документов в них заранее трудно предвидеть, то закрепление зон магнитной ленты за рубриками и подрубриками должно носить не жесткий, а вероятностный характер. При этом в одной зоне могут оказаться документы, относящиеся к разным рубрикам и подрубрикам, но основная часть документов должна обладать тематической близостью.

По мере наполнения системы документами должны проводиться периодические ревизии массива, переформирования зон и перестроения рубрикатора с целью обеспечения большего соответствия между составом документов в зонах и рубрикатором. При этом возможна автоматическая корректировка рубрикатора в соответствии с фактическим поступлением документов.

Соответствие между зонами и рубрикатором постоянно фиксируется в специальном машинном каталоге, представляющем собой упрощенный вариант инверсного поискового массива, в котором вместо номеров документов фигурируют номера зон магнитной ленты. При прямом входе в каталог можно по номеру рубрики и подрубрики найти номера зон, в которых записаны документы, относящиеся к заданным рубрике и подрубрике. При обратном входе в каталог можно по заданному номеру зоны определить номера рубрик и подрубрик документов, имеющих в этой зоне. Прямой вход использу-

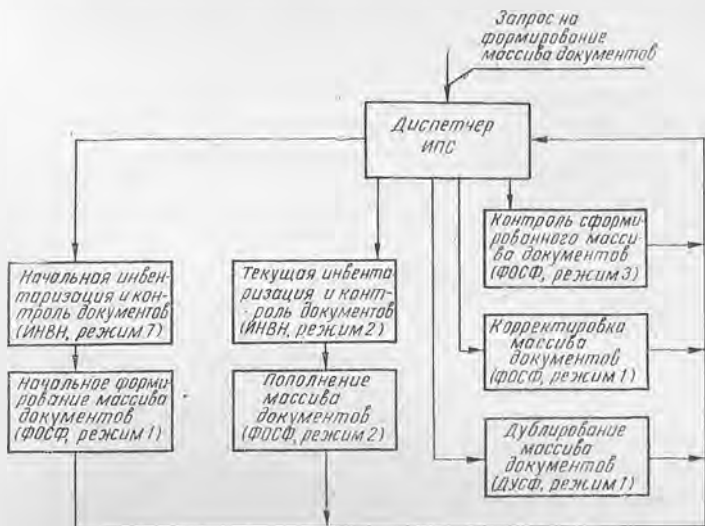


Схема 10. Формирование массива документов ИПС.

ется при поиске документов, когда по заданным рубрикам и подрубрикам отбираются зоны, в которых могут находиться искомые документы. Поиск документов в этих зонах осуществляется сплошным просмотром. Обратный вход в каталог используется при реорганизации массива документов и корректировке рубрикатора. Отметим, что периодическая реорганизация и перезапись массива документов с корректировкой рубрикатора и каталога имеют важное значение как для сохранности массива, так и для изучения его состава и структуры специалистами, ведущими этот массив, что является специфической особенностью человеко-машинных систем.

Программная реализация. Формирование массива документов рассматриваемой ИПС осуществляется четырьмя программами, работающими под управлением диспетчерской программы.

Формирование этого массива (схема 10) включает начальное формирование, пополнение, корректировку, дублирование и контроль сформированного массива документов. Основными являются первые два режима, которые реализуются программой инвентаризации и контроля входных документов (программа ИНВН) и про-

граммой формирования справочного фонда (программа ФОСФ), загружаемыми и включаемыми в работу диспетчером ИПС. Программа инвентаризации осуществляет формальный контроль поступающих документов и обеспечивает автоматическое присвоение вводимым документам единых инвентарных номеров. Проверенные документы с помощью программы ФОСФ вносятся в информационный фонд, организованный на магнитных лентах.

Программа, осуществляющая поиск документов в информационном фонде (схема 11), позонно считывает с магнитной ленты в оперативную память документы массива и, последовательно просматривая их, выдает пользователю на алфавитно-цифровое печатающее устройство или на дисплей документы, отвечающие критерию соответствия поисковому запросу.

Для проведения поиска пользователь с терминального устройства вводит запрос, общий вид которого имеет вид:

А ХХХ. ПОЗ >

где А — признак вида запроса (один символ), ХХХ — номер пункта, по содержанию которого будет производиться поиск (три десятичные цифры), ПОЗ — поисковый образ запроса, представляющий собой фразу (фразы) на естественном языке; отдельные элементы фразы могут быть связаны с помощью конъюнкций, дизъюнкций и отрицаний.

Признак вида запроса определяет режим автоматического преобразования исходного запроса к виду, необходимому для осуществления операций сравнения. Разработанный в описываемой ИПС алгоритм реализует следующие четыре вида запросов:

1. Отсутствие преобразования (символ «Ж»),
2. Исключение лишних слов (символ «П»),
3. Отделение окончаний (символ «О»),
4. Отделение суффиксов (символ «С»).

Первый вид реализуется в том случае, если пользователь сразу подает запрос в готовом виде и не желает подвергать его каким-либо преобразованиям. В этом случае пользователь ставит в запросе в качестве признака А символ «Ж», что означает «жесткий запрос».

Режим исключения лишних слов ограничивается исключением слов, заданных списком. Режим отделения

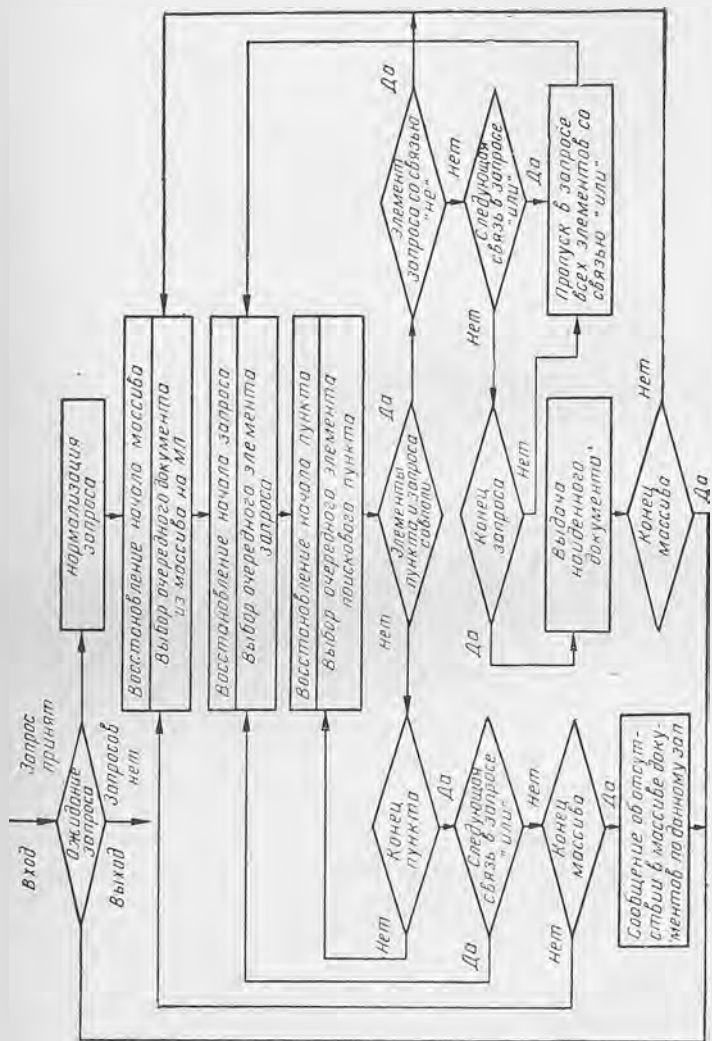


Схема 11. Алгоритм поиска документов.

окончаний, кроме того, выполняет отделение окончаний. В режиме отделения суффиксов, сверх того, отделяются суффиксы. Отделение суффиксов и окончаний осуществляется в соответствии с их списками, заданными заранее в памяти ЭВМ.

После преобразования запрос представляет собой последовательность отдельных элементов (основ или корней слов), связанных логическими операциями «и», «или» и «не». Для представления операции «и» используется пробел между словами, для представления операции «или» — символ «V» и для представления отрицания перед соответствующим термином ставится символ «—» (минус). Заметим, что в исходном запросе для обозначения конъюнкции («и») используется также пробел, для обозначения дизъюнкции — союз «или», а отрицания — слово «кроме», которое ставится перед отрицаемым термином.

Для проведения поиска из справочного массива выделяется определенная зона. Из нее выбирается очередной документ, фиксируется начало пункта с заданным в запросе номером. Элементы этого пункта (слова) по-символьно сравниваются с элементами нормализованного запроса.

Элемент запроса, имеющий перед собой символ, обозначающий отрицание, не должен давать совпадения ни с одним элементом заданного пункта документа. Если он совпадает хотя бы с одним элементом данного пункта, то программа предусматривает переход к проверке следующего документа. Если проверяемый элемент (слово) запроса не имеет перед собой отрицания и проверка его с очередным элементом документа не дает совпадения, то выполняется проверка вида связи данного элемента запроса с последующим. Наличие в этом месте связки «V» требует продолжения сравнения запроса с документом, а именно следующий элемент запроса сравнивается с элементами поискового пункта. Если элемент запроса, не совпавший ни с одним элементом пункта, соединен с последующим элементом запроса логической связью «и», то документ считается нерелевантным запросу и осуществляется переход к выбору и просмотру следующего документа. При совпадении неотрицаемого элемента запроса с каким-либо элементом документа процесс сравнения запроса и документа продолжается, причем все последующие элементы запроса, связанные

с данным элементом логической связью «или», пропускаются.

Если поисковый пункт отвечает критерию выдачи, он выдается пользователю на алфавитно-цифровое печатающее устройство или на дисплей и начинается анализ следующего документа. Если при просмотре всего справочного массива не было обнаружено совпадения поискового пункта с нормализованным запросом, пользователю выдается сообщение о том, что в справочном массиве по данному запросу документов нет. Рассмотренный алгоритм сравнения запросов и документов на естественном языке реализован на ЭВМ «Минск-32».

5. Методы и возможности автоматического индексирования документов и запросов

Нами были рассмотрены два варианта ИПС, в которых использовались два различных способа представления поисковой информации: применение формализованного языка и непосредственное использование естественного языка.

Основным недостатком первого способа является необходимость выполнения людьми-индексаторами работы по записи основного смыслового содержания документов и запросов на формализованном языке. Главным недостатком второго способа следует считать низкую точность поиска документов, обусловленную неприспособностью естественного языка к автоматической обработке.

Большое значение в настоящее время приобретает третий способ — автоматический перевод основного смыслового содержания документов и запросов с естественного языка на формализованный¹.

Непосредственное сравнение запросов и документов на естественном языке не позволяет получить удовлетворительных результатов поиска, так как приводит к высокому уровню информационного шума. Это связано с такими свойствами естественного языка, как:

1) неравномерное распределение смыслового содержания текста по его отдельным словам, что требует разделения слов на значащие и незначащие;

¹ Проблема автоматического индексирования весьма сложна и многогранна. В данном разделе авторы излагают один из подходов к ее решению (прим. ред.).

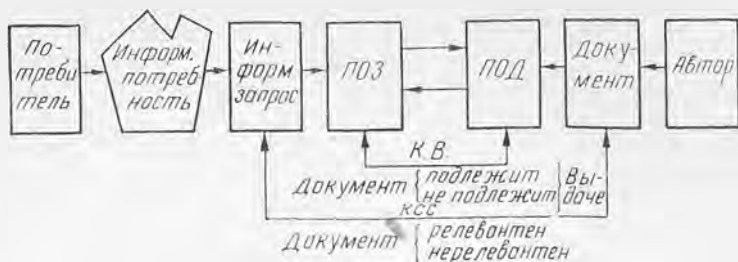


Схема 12. Взаимодействие пользователей с ИПС.

2) синонимия и многозначность как лексических, так и грамматических единиц языка;

3) наличие скрытых внутренних связей между понятиями как в пределах каждого конкретного текста, так и в общей системе языка.

4) наличие различий в системах грамматических категорий в разных языках.

Принципиальная схема взаимосвязи с ИПС представлена на схеме 12.

Автоматическим индексированием называется автоматическое определение основного смыслового содержания документов и запросов и выражение его в виде ПОД или ПОЗ на некотором ИПЯ.

Программы автоматического индексирования, реализуемые на ЭВМ третьего поколения, строятся в соответствии со следующими принципами, характерными для систем программного обеспечения современных ЭВМ вообще.

1. Программа должна быть рассчитана на различные области применения и различные аппаратурные средства (конфигурации вычислительной системы). Это требует создания гибкой модульной программы. Благодаря параметрической универсальности программ пользователь получает возможность задавать конкретные параметры работы, касающиеся, например, конфигурации системы, требуемого качества индексирования и т. д. Функциональная избыточность обеспечивает возможность реализации противоречивых требований с достаточной эффективностью, а функциональная избирательность позволяет выбирать соответствующие программные модули для конкретной области применения.

2. Программа должна быть открыта для последующих расширений и усовершенствований. Новые достижения, например в области языкознания, не должны приводить к коренной переделке программы. Это предъявляет особые требования к модулям синтаксического и семантического анализа текстов.

3. В программах используется возможность тесной связи человека с машиной, представляемой ЭВМ третьего поколения. Это означает включение в порядок работы ИПС режимов полуавтоматической работы, позволяющих применять методы синтаксического анализа, при которых проблемы многозначности еще не решаются полностью автоматически. Совершенствование синтаксических методов позволит впоследствии увеличить степень автоматизации. Это соответствует методу диалогового поиска информации, при котором участие человека способствует повышению качества информационного поиска, осуществленного ЭВМ.

С целью облегчения обмена научной медицинской информацией в международном масштабе целесообразно иметь программы, пригодные для анализа и обработки ряда языков, на которых печатается большинство публикаций и структура которых позволяет применять единые средства анализа. В настоящем разделе описывается система автоматического индексирования (САИ), разработанная применительно к трем языкам — русскому, английскому и немецкому. Для этой группы языков построены унифицированные алгоритмы автоматического индексирования текстов и представления их на ИПЯ. Унификация облегчается некоторыми общими свойствами этих языков и тем обстоятельством, что язык научных текстов отличается от литературного языка большей формализованностью и однозначностью средств выражения.

С точки зрения разработки алгоритмов автоматического индексирования можно выделить следующие две основные стороны этого процесса:

1. Определение понятий, характеризующих смысловое содержание документов или запросов, заданных в виде текстов на естественном языке.

Известно, что в документальных ИПС в поисковых образах документов не содержится непосредственно та информация, в которой нуждается пользователь. ПОД содержит только понятия, затрагиваемые автором в ра-

боте и называющие (именующие) проблемы, которые рассматриваются в тексте. Процесс определения понятий, характеризующих смысловое содержание текста, и устранение синонимии, полисемии и омонимии в тексте будем называть нормализацией текста;

2) выделение основного смыслового содержания текстов.

В связи с тем что для целей информационного поиска не требуется, да и не представляется возможным отразить в ПОД все аспекты данного документа, необходимо выделить из совокупности всех понятий, определяющих смысловое содержание текста, наиболее значащие и существенные понятия. Этот процесс назовем сжатием текста. Определение двух названных процессов позволяет классифицировать существующие методы автоматического индексирования и разработать структуру необходимой программной системы.

Конкретные процедуры, применяемые для нормализации текста, зависят, с одной стороны, от данного естественного языка, с другой — от применяемого ИПЯ.

Словарный состав ИПЯ образуется из дескрипторов, которые выступают как представители некоторого класса условно эквивалентных ключевых слов (условно эквивалентных по отношению к определенному понятию). Таким образом, дескриптор может быть определен как лексическая единица информационно-поискового языка (ИПЯ), выражающая определенное понятие. Применение дескрипторов исключает такие выше-названные свойства естественного языка, как синонимия, омонимия и полисемия. В ИПС с грамматикой используются, кроме самих дескрипторов, также средства для выражения смысловых отношений между ними. Анализ многих ИПЯ с грамматикой позволяет выделить некоторые основные смысловые отношения, присутствующие в этих языках:

- 1) отношение между агентом и процессом (agt),
- 2) отношение между процессом и объектом, на который он направлен (gl),
- 3) отношение между процессом и адресатом процесса (dat),
- 4) отношения между дескрипторами, в которых один дескриптор определяет свойства другого дескриптора (att),

5) сочинительные отношения между однородными членами приложений (со).

Для разработки алгоритмов перевода текстов с естественного языка на ИПЯ используется определенная модель естественного языка. В модели, разработанной советскими учеными за последнее десятилетие, различаются следующие уровни естественного языка:

- 1) графический (ГУ);
- 2) морфологический (МУ);
- 3) поверхностно-синтаксический (ПСУ);
- 4) глубинно-синтаксический (ГСУ);
- 5) семантический (СУ).

Указанная модель естественного языка предусматривает ряд последовательных этапов анализа и преобразования текстов в процессе его нормализации. В частности, описываемая система автоматического индексирования, разработанная П. Науманом, включает ряд этапов, реализуемых соответствующими программными модулями.

Указанная система автоматического индексирования может применяться как к полным текстам, так и к рефератам документов (статей). В первом случае целесообразно сначала произвести сжатие документов. После сжатия можно осуществить автоматическую нормализацию текста. Естественно, что автоматическая нормализация может применяться и к полным текстам, что потребует больших затрат машинного времени, а также к рефератам или аннотациям документов, полученным вручную.

В процессе самой нормализации текстов также предусматривается их некоторое сжатие за счет исключения так называемых общих слов (см. ниже).

Указанная система автоматического индексирования осуществляет индексирование именных (назывных) предложений, причем в общем тексте части, не являющиеся назывными предложениями, будут пропускаться.

Таким образом, САИ может использоваться для индексирования специально подготовленных рефератов, содержащих именные простые предложения. САИ предусматривает предредактирование текста, которое сводится к выделению заголовка (символами в начале и в конце), выделению предложений или их частей, которые не должны подвергаться автоматическому индекси-

рованию (например, фамилии авторов, названия журналов, номера томов, стр. и т. п.), эти сегменты выделяются символами \diamond в начале и $\&$ в конце.

Исходный текст после указанного предредактирования перфорируется и вводится в ЭВМ, после чего начинают работать модули пакета программ автоматического индексирования. Рассмотрим основные принципы построения алгоритмов упомянутых этапов автоматического индексирования.

1. **Этап графического анализа** обеспечивает представление текста в виде слов (последовательностей букв), пробелов, знаков препинания. Программа графического анализа выполняет ввод текста и разделение его на слова; при этом также определяются границы глав, параграфов, абзацев, предложений. Графическое представление получается непосредственно из текста анализируемых документов.

Структура текста выражается на ГУ с помощью отношения линейного упорядочения слов внутри предложений, предложений внутри абзаца и т. д. Лексические единицы этого представления — слова естественного языка. Использование слов текста в качестве лексических единиц ИПЯ, а также использование отношения линейного упорядочения в качестве отношения между элементами ИПЯ исключается из-за названных выше свойств естественного языка. Так, например, синонимия устраняется на этом уровне только для элементов текста, последовательности букв которых совпадают.

Программный модуль графического анализа (ГРАФ), переводящий текст в графическое представление, выполняет ввод текста с некоторого носителя информации (перфокарт, перфолент, магнитных дисков или лент), устанавливает начало и длину любого слова текста, тип разделительного знака между словами, обеспечивает другим модулям доступ к любым словам текста, а также производит некоторое редактирование введенного текста.

Модуль ГРАФ формирует так называемую таблицу обращения к тексту (ТОТ), в которой каждому слову или знаку препинания соответствует одна строка — вектор, и заносит в вектор следующие данные: адрес начала слова (знака) в ОЗУ, длину слова в байтах, номер абзаца, номер предложения (фрагмента), номер слова в предложении, признак положения слова (из за-

головка — Т, из тела текста — В). Фрагмент — любой участок текста (предложения), заключенный между разделителями (знаками препинания). Под сегментом понимается последовательность символов, заключенных между пробелами или другими разделителями. Обычно это бывает слово. Следующей работает программа отбора общих слов, которая сравнивает каждое слово обрабатываемого текста со словами заданного заранее словаря словоформ общих слов и исключает найденные общие слова из дальнейшей обработки. К общим словам относятся незначащие слова (такие, как рассмотрение, очевидно, описывается и т. п.), высокочастотные значащие слова, различительная роль которых в данной тематике документов невелика, ряд служебных слов. Список общих слов может корректироваться и пополняться в процессе эксплуатации САИ.

2. Этап морфологического анализа обеспечивает представление каждого слова предложения в виде последовательности морфем наименьших единиц языка, выполняющих смысловую или структурную роль. Программы морфологического анализа, разделяющие слова на основы и окончания, могут работать по двум принципам:

— анализ справа налево, при этом находится наибольшее окончание из словаря окончаний, совпадающее с правой частью слова;

— анализ слева направо, при этом находится наибольшая основа из словаря основ, совпадающая с левой частью слова.

На этапе морфологического анализа различные словоформы одного слова, а также словоформы синонимов сводятся к одной и той же смысловой единице — лексеме. Эта лексема снабжается набором так называемых морфологических переменных, определяющих класс слов (существительное, прилагательное, причастие, местоимение, числительное, союз, предлог), род. Сама лексема (ее код и побуквенное представление) берется из словаря основ, а значения морфологических переменных — как из словаря основ, так и из словаря окончаний. Иногда в этих словарях указываются и семантические признаки, например, для окончаний «ирование» и «ировка» может стоять признак «процесс», а для окончаний «атор» и «ист» — признак «субъект» или «устройство». В связи с тем что описываемая система автоматического индексирования рассчитана на анализ имен-

ных (назывных) фраз, в ней не учитываются глаголы.

С целью уменьшения ошибок при отделении окончаний (излишнее отделение или недостаточное отделение), связанных с чередованием букв в основе, применяются дополнительные приемы (проверка на совместимость и отсутствие перекрытия основ и окончаний, поиск в списке основ, допускающих чередование и др.).

Для английского языка можно использовать простой алгоритм отделения наиболее длинного окончания. Для русского языка целесообразно модифицировать этот алгоритм следующим образом: сначала отделить наиболее длинное флексивное окончание, а затем — наиболее длинный словообразовательный суффикс. При таком двухступенчатом способе резко сокращаются списки окончаний и суффиксов. Для немецкого языка удобно сначала попробовать выделить суффикс, потом окончание (короткое), а затем суффикс и т. д.

Этап морфологического анализа принято делить на два подэтапа: поверхностно-морфологический анализ, заканчивающийся разделением слов на морфемы (отделением окончаний и суффиксов), и глубинно-морфологический анализ, обеспечивающий получение значений морфологических переменных.

Исходными данными для глубинно-морфологического анализа являются полученные при поверхностно-морфологическом анализе номер основы в словаре основ (код дескриптора или код лексемы), номер окончания в списке окончаний, некоторые грамматические признаки словообразовательных суффиксов, указанные в списке суффиксов, а также флективный класс основы, который также указывается в словаре основ.

По номеру флективного класса определяются класс слова, род, одушевленность, а по сочетанию номера флективного класса и номера окончания — число и падеж. В САИ используется флективная классификация изменяемых слов, данная в работе Г. Г. Белоногова (1973). Возможен также вариант приближенного глубинно-морфологического анализа в случаях отсутствия указаний флективных классов основ. При этом используются признаки словообразовательных суффиксов, показывающие, к каким флективным классам может принадлежать данная основа.

Возникающая при этом морфологическая многозначность может быть либо устранена на этапе синтаксиче-

ского анализа, либо она приведет к синтаксической многозначности, которая должна быть снята при участии человека (в диалоговом режиме).

Таким образом, при представлении текста на морфологическом уровне указываются лексемы предложения текста; с каждой лексемой упорядоченный кортеж морфологических переменных с фиксированными значениями, порядок лексем в предложении. При этом к одной и той же лексеме сводится множество слов текста, различающихся значениями своих морфологических переменных.

Использование элементов морфологического представления в качестве лексических элементов ИПЯ приведет к ИПЯ с достаточно низкой семантической силой. В этом языке обнаруживалась бы только синонимия слов, отличающихся флексийными окончаниями и, если это допускается принятым критерием смыслового соответствия, то и словообразовательными суффиксами. На морфологическом уровне можно использовать только отношение линейного упорядочения элементов текста, что не позволяет применить грамматические средства в ИПЯ, получающемся при переводе с этого уровня. Очевидно, что значение линейного расположения слов в тексте еще не определяет все связи между этими словами. Семантическую силу такого ИПЯ можно повысить, снабжая основы и окончания определенной грамматической информацией, характеризующей роль отдельных слов. Однако применение указателей роли слов без указания связи между словами (т. е. без синтаксического анализа) проблематично и может сопровождаться ошибками.

Следует заметить, что найденные на этапе морфологического анализа основы по сути дела являются псевдоосновами, так как не всегда совпадают с основами, которые получил бы лингвист при морфологическом анализе. Но для целей информационного поиска это допустимо, если при этом соответствующие словоформы всегда сводятся данным алгоритмом отделения окончаний к одной и той же псевдооснове. Метод анализа справа налево целесообразно применять при хранении основ слов в виде сверток.

При анализе слева направо проверяется наличие данного слова в списке побуквенным сравнением слова текста с основой в списке основ, начиная с первой бук-

вы до получения максимального совпадения. Остающиеся несовпадающие части слова текста сравнивают со списками окончаний.

3. Этап поверхностно-синтаксического анализа обеспечивает определение управляющих слов в предложении и видов связи между ними. Этот этап анализа тесно связан с последующим этапом глубинно-синтаксического анализа, что проверяется, в частности, в выборе систем отношений (связей) для этих двух этапов. Сначала, исходя из семантики языка, интуитивно выбираются основные типы синтаксических связей между членами предложения для глубинно-синтаксического уровня. В качестве таких связей выбраны следующие четыре типа связей: 1) субъектная — связь между действием и органом (лицом, устройством и т. д.), выполняющим это действие; 2) объектная — связь между действием и объектом, на который направлено это действие; 3) определительная — связь определяемого и определяющего членов предложения; 4) сочинительная — связь между однородными членами предложения.

Так как эти основные глубинно-синтаксические связи могут иметь место между словами различных классов, то возникает большое количество поверхностно-синтаксических связей. Ввиду того что в данной САИ обрабатываются только назывные фразы, то в ней учитываются возможные синтаксические связи между словами следующих четырех классов: существительные, прилагательные, причастия, предлоги.

Алгоритм поверхностно-синтаксического анализа предусматривает последовательный просмотр слов фразы, для каждого из них находят управляющее слово и вид связи. Возможны различные подходы к построению такого алгоритма. В данной САИ реализован алгоритм, согласно которому на основе достаточно простой системы правил (представленных в виде так называемой таблицы конфигураций) определяются все потенциально возможные варианты связей между словами анализируемого предложения, а затем с помощью набора фильтров, представленных в виде подпрограмм, устраняются лишние связи. Потенциально возможные связи (отношения подчинения) фиксируются в виде матрицы, в которой номера строк соответствуют номерам зависимых слов, номера столбцов — номерам управляющих слов, а на их пересечении указывается код связи.

Таблица конфигураций представляет собой матрицу, в которой указываются классы зависимого и управляющего слов, условие управления (например, согласование в роде, числе, падеже), расположение управляемого слова по отношению к управляющему в пределах фразы (вправо, влево или в обе стороны), тип связи. Подпрограммы фильтров устраняют связи, противоречащие определенным ограничениям. Так, например, проверяется условие проективности синтаксического дерева подчинения, заключающееся в том, что различные связи не должны пересекаться. Если какое-либо слово имеет одну потенциальную связь (подчиняется другому слову), то эта связь считается правильной и устраняются другие связи, противоречащие данной. Если, например, прилагательное имеет две потенциальные связи и в одной из них выступает как прилагательное, в другой — как существительное (что допустимо в русском языке), а за ним следует существительное, то вторая связь у прилагательного устраняется.

Далее, любой предлог в русском языке должен управлять существительным, и один из фильтров проверяет это условие насыщенности предлогов.

Устанавливаемые при поверхностно-синтаксическом анализе синтаксические отношения в общем случае не являются достаточными для использования их в качестве грамматических средств ИПЯ, так как они служат в основном для отражения сходства внешнего грамматического оформления, а не содержательных связей между элементами предложения.

4. Этап глубинно-синтаксического анализа обеспечивает представление содержательных связей между членами предложения, независимо от их грамматического оформления, т. е. представление в обобщенном виде множества эквивалентных поверхностно-синтаксических структур. Так, например, при поверхностно-синтаксическом анализе будут определены различные отношения для следующих фраз:

Расчет конструкций
Рассчитанная конструкция
Рассчитываемая конструкция.

При глубинно-синтаксическом анализе эти фразы будут сведены к одному отношению «процесс — объект».

Глубинно-синтаксическая структура — это дерево, вершины которого представляются символами обоб-

щенных лексем. Ветви этого дерева помечены символами типов глубинно-синтаксических связей, характерных для естественного языка. В описываемой САИ используются следующие семь видов глубинно-синтаксических связей: субъектная связь, четыре вида объектной связи, определительная связь и сочинительная связь. При переходе к глубинно-синтаксическому уровню представления фразы все субъектные поверхностно-синтаксические связи сводятся к одной субъектной глубинно-синтаксической связи, столбцы объектных связей — к четырем глубинным объектным связям, все определительные поверхностные связи — к одной глубинной определительной связи, все сочинительные поверхностные связи — к одной глубинной сочинительной связи.

В общей теории глубинно-синтаксического анализа предусматривается использование лексических функций, позволяющих свести разные лексемы к одной обобщенной лексеме (например, «сильный», «бурный», «категорический» и т. п.) к «Magn», означающей «высокая степень». В описываемой САИ этот метод не применяется, за исключением того, что синонимичные слова с помощью словаря сводятся в классы эквивалентности.

При обработке запросов на поиск возможно также включение более общих, частных или ассоциированных дескрипторов, если это требуется выбранной стратегией поиска.

При практической реализации описываемой системы автоматической индексации коды понятия присваиваются предметным терминам программным модулем, условно названным синоним, сразу после представления анализируемого предложения на морфологическом уровне. Дополнительный модуль семан используется в качестве фильтра, пропускающего только релевантные смысловые связи.

В результате анализа процесса нормализации текста П. Науманном были составлены шесть программных модулей. Эти модули обеспечивают, с одной стороны, сведение слов текста к нормированным понятиям, являющимся лексическими элементами ИПЯ. Для этого используются списки основ предметных терминов, словообразовательных суффиксов и флексийных окончаний вместе с грамматической и семантической информацией, а также тезаурус. С другой стороны, полученные модули устанавливают смысловые отношения между

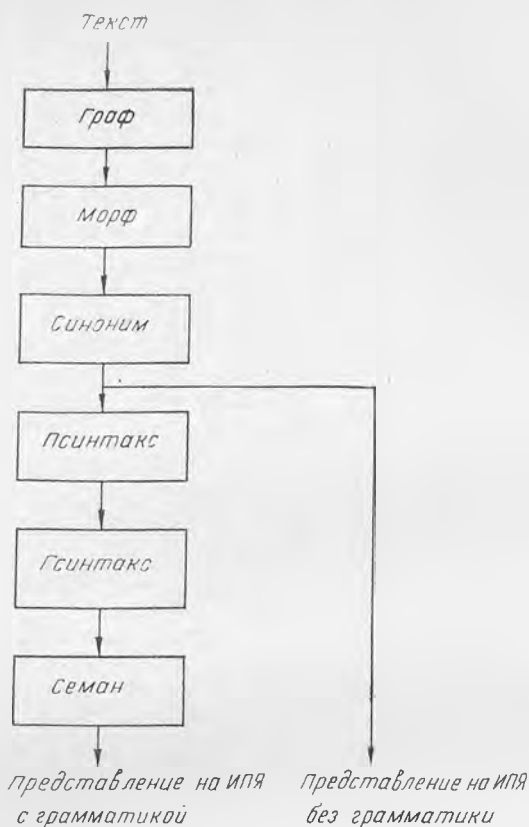


Схема 13. Структура программы автоматической индексации без сжатия текста.

дескрипторами на основе грамматики для поверхностно-синтаксического анализа, правил преобразований для перехода к представлению на глубинно-синтаксическом уровне (ГСУ) и системы анализа смысловых связей, выраженных с помощью предлогов. Таким образом была получена структура программы нормализации текста, представленная на схеме 13.

Рассмотренные программные модули могут настраиваться на разные варианты работы путем задания определенных параметров. Так, к модулям предъявляются разные требования в зависимости от того, нужен ли

анализ смысловых отношений между элементами индексируемых предложений или нет. Модуль (Морф) может выполнять отделение только флексийных окончаний или вместе с ними и словообразовательных суффиксов в зависимости от требуемой точности представления содержания текста.

Предусмотренный в системе диалоговый режим работы позволяет практически реализовать такие нерешенные задачи, как углубленный синтаксический или семантический анализ, при которых вмешательство человека ограничивает затраты в рамках разумного.

Шаги нормализации текста должны сочетаться с мерами его сжатия. Можно выделить два способа сжатия текста.

1. Сжатие текста путем выделения отрезков (предложений, абзацев) текста, в которых содержатся самые важные и существенные сведения о содержании текста (квазиреферирование).

На базе списков индикативных слов, критериев связности предложений между собой или с заголовками, сведений о расположении предложений в тексте, статистических соображений и др. выделяются предложения текста, которые затем подвергаются нормализации. В основе этого метода лежит утверждение, подкрепленное результатами экспериментов, что результаты поиска не ухудшаются, если индексируются только рефераты вместо полных текстов. Такой подход позволяет начать практическое использование системы автоматического индексирования без модулей, осуществляющих сжатие текста, используя в качестве исходного материала рефераты документов, составленные обычным способом. Затем возможно постепенное введение средств сжатия текста в уже работающую и опробованную систему нормализации текста.

2. Сжатие текста путем определения значимости предметных терминов текста, проводимое на основе вероятностно-статистического анализа всего текста. Этот метод основан на подтверждаемом практикой допущении того, что авторы используют в своих материалах чаще всего понятия, которые ассоциированы с темой статьи.

Возможная структура программы автоматического индексирования с использованием первого способа сжатия текста представлена на схеме 14. Программный мо-

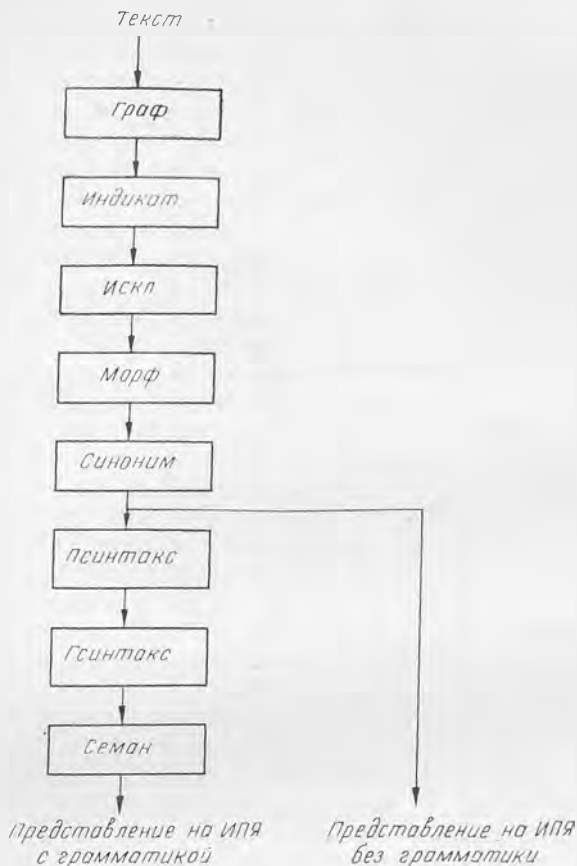


Схема 14. Структура программы автоматической индексации со сжатием текста.

дуль **Индикат** выделяет или исключает предложения текста на основе названных выше критериев индикативности, а модуль **Искл** исключает все предметные термины текста, дискриминационная сила которых в связи с большой частотой употребления в текстах по соответствующей тематике очень низка (высокочастотные слова), и маркирует все функциональные слова (предлоги, союзы), которые учитываются при анализе смысловых отношений.

6. Медицинская лингвистика

В связи с увеличением сложности и объема медицинских данных возникает необходимость в разработке точных формализованных методов записи и использования этих данных, обеспечивающих возможность их быстрого анализа и использования для принятия решений. Это необходимо также для улучшения обмена информацией между специалистами в области медицины, объединения усилий в исследованиях, обеспечения взаимосвязи медицинских учреждений с экономическими и социальными-административными учреждениями.

Использование общих лингвистических принципов и подходов в области медицинской информации должно способствовать созданию медицинского информационного языка. При этом следует иметь в виду, что в лингвистике, к сожалению, еще далеко не решены проблемы представления семантики в естественных языках. Медицина объединяет большое количество различных специальностей, каждая из которых имеет специальную терминологию, в большей или меньшей степени пересекающуюся с терминологией других медицинских специальностей.

Основным принципом развития систем медицинской информации должно быть представление ее в форме текстов естественного языка, удобных для восприятия людьми — медицинскими работниками и пригодных для автоматической обработки на ЭВМ с использованием методов морфологического и синтаксического анализа с учетом смыслового значения слов.

Медицинские предписания, записи физических и физиологических обследований и врачебных заключений в большинстве случаев представляют собой назывные фразы или группы назывных фраз естественного языка.

В медицине традиционно сложился и продолжает развиваться подход к построению семантических систематизированных систем для формулировки диагноза — построение классификаторов и номенклаторов болезней, операций, специальностей, медикаментов и т. д. Наиболее известными из подобных систем являются Международная классификация болезней, травм и причин смерти (восьмого пересмотра), Американская систематизированная номенклатура патологий (SNOP — Systematized Nomenclature of pathology), предназначена для организации хранения и поиска клинической информации.

В SNOP все термины разделены на четыре семантические категории: морфология (M), характеризующая изменения в тканях, органах (опухоль, разрыв, воспаление и т. д.); топология (T), дающая представление о локализации патологического процесса (органы, ткань, система, организм); этиология (E), раскрывающая причину заболевания (микроорганизм, химическое вещество или медикамент и др.); функция (F), характеризующая физиологическое проявление заболевания, включая симптоматику. Каждый термин в SNOP может принадлежать только к одной семантической категории. При использовании SNOP для записи медицинских заключений используемые в них термины снабжаются указателями категорий. В Национальном институте здоровья США А. Праттом ведется работа по созданию автоматизированной системы индексации медицинских диагнозов. В качестве основы для построения формализованного медицинского языка принята указанная систематизированная номенклатура патологий (SNOP). Каждый термин в SNOP имеет свой номер (код) в пределах определенного информационного класса.

С помощью данной семантической модели медицинский диагноз может быть представлен как семантическое произведение четырех указанных выше семантических категорий (T, M, E, F) и набора предикатов отношений (R_1, R_2, \dots, R_n). Предикатами отношений являются различные типы лингвистических единиц (предлоги, глаголы и т. д.), основной функцией которых является выражение семантических отношений между категоризованными семантическими единицами.

Например, утверждение «Пневмония из-за стафилококков» будет представлено в формализованном виде как $MR_{np}E$, где R_{np} представляет собой причинный предикат отношения «из-за»; пневмония является семантической единицей M и стафилококки — семантической единицей E.

Введение семантических категорий существенно для обработки медицинских данных на естественном языке. После того как семантические категории указаны, дальнейший анализ и обработка приводят к ряду отношений между словами, фразами и семантическими категориями. Семантическая интерпретация медицинских заключений основана на морфо-синтаксической и семантической информации, присвоенной каждому слову в фразе или пред-

ложении. Предполагается, что введение дополнительных дедуктивных различительных правил в существующую логическую модель семантики медицинского формализованного языка может повысить его точность и общность.

Основные проблемы в области автоматической обработки медицинских текстов связаны в настоящее время с решением следующих вопросов: 1) устранением различий в написаниях и сокращениях медицинских текстов, т. е. унификацией медицинской терминологии; 2) исключением синонимии, полисемии и антонимии; 3) унификацией правил использования знаков пунктуации, особенно запятых; 4) разработкой методов автоматического опознавания анафорических, эллиптических и идиоматических конструкций; 5) разработкой методов автоматического анализа не прямых отсылок, встречающихся в текстах; 6) опознаванием семантики эквивалентных, но синтаксически различных конструкций; 7) исключением многозначности полисемических слов, фраз и предложений; 8) анализом параграфов и больших единиц текстов; 9) разработкой единой семантической модели естественного медицинского языка. Все существующие теории мало подходят для построения единой семантической теории естественного языка.

По-видимому, общий подход должен включать в себя построение формальных логических структур, которые бы отражали смысл содержания текстов, заданных на естественном медицинском языке. Переменные и операторы в этом формальном языке должны отражать модель мыслительного процесса врачей, пользующихся данным естественным медицинским языком. Допускается, что внутренняя глубинная структура мыслительного процесса людей, говорящих на одном языке, имеет сходное строение и поэтому может быть представлена на одном и том же формальном языке. Следовательно, разработка формального промежуточного языка может служить базой для машинного перевода исходного языка в промежуточный и из промежуточного языка в ряд конкретных языков.

Перечисленные вопросы относятся к перспективам полной автоматизации обработки текстовой информации, выраженной на естественном медицинском языке. Однако уже в настоящее время имеется практическая возможность автоматической обработки медицинской информации, представленной на естественном языке с определен-

ными ограничениями. Основным ограничением является, как уже говорилось, возможность использования только простых назывных предложений.

К важным практическим задачам, при решении которых уже используется автоматическая обработка текстов, относятся автоматическое индексирование и кодирование медицинских диагнозов, врачебных заключений и назначений. Эти задачи являются составными частями более общей задачи применения естественного языка для ведения машинных историй болезней. Очевидно, что естественный язык медицинских записей при ведении историй болезней на ЭВМ наиболее удобен для врачей, однако в настоящее время он еще не может без ограничений быть использован при машинном ведении историй болезней. Машинная история болезни в общем случае должна включать формализованный раздел для языка анкетных и общих (количественных) сведений и раздел, заполняемый на естественном языке и содержащий информацию о состоянии больного и его лечении, вводимую в хронологическом порядке. Используемый при этом машинный словарь должен содержать общепринятые медицинские слова и выражения, смысл которых должен быть строго фиксирован и не должен зависеть от контекста. Предложения должны строиться с помощью простой системы синтаксических правил, легко усваиваемых медицинским персоналом. Вводимые данные должны представляться порциями — последовательностями простых предложений и цифровых показателей (результатов измерений), подобных обычным фразам телеграфного стиля.

Порции данных (отдельные параграфы или абзацы) должны удовлетворять следующим требованиям логического и формального контроля.

1. Порция должна начинаться с даты. Если она не указана, то берется дата предыдущей порции.

2. Основой каждого предложения должно являться главное существительное — первое встретившееся в нем простое или сложное существительное. Например, правильная запись фразы имеет вид: «Печень, увеличенная в объеме» (главное слово — печень). Неправильная запись этой же фразы: «Увеличение объема печени». Фразы могут содержать дополнительные существительные, но все вопросы при поиске информации о больных будут ориентированы в основном на проверку главных существительных фраз. Только после совпадения глав-

ных существительных во фразах запроса и во фразах соответствующей порции истории болезни будет производиться уточнение соответствия данной порции запросу с помощью дополнительных слов, имеющихсЯ во фразах.

Естественно, что с развитием методов автоматического индексирования текстов, в частности по мере реализации тех моделей глубинно-синтаксического и семантико-синтаксического анализа текстов, которые были описаны раньше (в разделе 5-м главы IV), упомянутые ограничения на использование естественного медицинского языка для машинного ведения историй болезней будут постепенно сниматься.

Наряду с этим очевидно, что будет проходить и обратный процесс уточнения и систематизации терминологии естественного медицинского языка, что будет способствовать повышению эффективности методов автоматической обработки медицинских текстов.

ОСНОВНЫЕ НАПРАВЛЕНИЯ ПРИМЕНЕНИЯ ЭВМ В МЕДИЦИНЕ

В настоящее время наблюдается усиленное внедрение ЭВМ в медицину и здравоохранение. Во многих странах эти машины широко используются не только при выполнении исследовательских работ и решении задач медицинской статистики, но и непосредственно в лечебном и диагностическом процессах, при управлении больницами и другими медицинскими учреждениями и т. д. Под термином «медицинская информатика» принято понимать широкий круг вопросов автоматизации процессов сбора, обработки и использования информации в медицине и здравоохранении. Основным средством автоматизации информационных процессов являются ЭВМ.

Можно указать три основных раздела медицинской информатики:

1) общие вопросы и методология внедрения ЭВМ в здравоохранение (в основном касающиеся работников управлений и служб здравоохранения и отдельных больниц). Сюда же относятся вопросы подготовки кадров и оценка эффективности применения ЭВМ в медицине и здравоохранении;

2) обработка медицинской информации и организация работы вычислительных центров. В этом направлении заняты главным образом специалисты в области программирования, ЭВМ и алгоритмирования, а также медики, непосредственно связанные с работой ЭВМ. В этом разделе важными являются вопросы применения ЭВМ в диагностике, ведение машинных историй болезней, планирование обслуживания пациентов, а также вопросы построения информационной базы АСУ медицинского назначения, требования к техническим средствам и математическому обеспечению указанных АСУ;

3) обработка электрофизиологических сигналов с помощью ЭВМ, применение ЭВМ в клинических лабораториях, рентгеновских кабинетах и лабораториях, отделе-

ниях интенсивного ухода, радиологических отделениях, аптеках и т. д. Здесь имеется ряд конкретных задач, решаемых с помощью ЭВМ с высокой степенью эффективности.

Практика показывает, что внедрение ЭВМ должно осуществляться в первую очередь для сбора, хранения, обработки больших объемов медицинской информации в больницах, амбулаториях и других медицинских учреждениях.

Важной областью применения ЭВМ являются также экономические и плановые расчеты в здравоохранении.

1. Автоматизация сбора и обработки данных о состоянии больных и машинное ведение истории болезни

Современная организация больничной работы характеризуется широким использованием сложной диагностической, терапевтической и другой медицинской техники. Высокая техническая оснащенность и возрастающая специализация отделений, служб и медицинского персонала повышают эффективность лечения и обслуживания больных, но вместе с этим усложняют работу медицинского персонала и повышают стоимость лечения. Современная больница должна рассматриваться как большое производственное предприятие со сложной внутренней кооперацией. В связи с этим важными являются автоматизация сбора и обработки данных в больнице и системный подход к планированию обслуживания больных, контролю работы всех служб и анализу результатов. Организация больницы должна быть приспособлена к возрастающим и изменяющимся задачам, и здесь должны быть использованы методы системного анализа.

Указанные трудности привели к необходимости создания больничных автоматизированных информационных систем (БАИС), предназначенных для решения лечебно-диагностических задач и обработки планово-экономической информации.

Техническая база БАИС включает в себя в основном ЭВМ, работающую в режиме разделения времени и в реальном масштабе времени, связанную сетью терминалов-дисплеев с различными подразделениями больницы.

В составе БАИС выделяют обычно три основные подсистемы:

- 1) подсистема сбора и обработки информации о больных (лечебная подсистема),
- 2) административная подсистема, обеспечивающая учет и планирование работы больницы,
- 3) подсистема обработки лабораторных данных и результатов электрофизиологических и рентгенологических обследований.

С помощью терминалов в память ЭВМ могут вводиться как числовые, так и текстовые данные, получаемые в различных пунктах больницы или поликлиники.

Информация из ЭВМ может запрашиваться также с помощью терминалов. Информацию о больном могут запрашивать врачи перед осмотром больного и во время его, во время конференций, во время операции; необходимые сведения могут запрашивать сестры при выполнении назначений врачей и т. д.

Первые данные о больном вводятся в БАИС еще до его госпитализации, во время первого амбулаторного приема в связи с данным заболеванием. БАИС по указанию врача планирует сроки необходимых лабораторных, рентгенологических и других обследований больного, собирает и хранит все результаты обследований и выдает их врачу, осматривающему больного. При госпитализации больного эти данные пополняются направлением на госпитализацию и врачебными комментариями. При первом амбулаторном приеме больного ему присваивается так называемый идентификационный номер, включающий год, месяц, день рождения и его инициалы. С помощью такого условного номера данные о любом больном могут быть быстро найдены в памяти ЭВМ без риска их перепутать с данными другого больного. Таким образом формируется и постоянно пополняется машинная история болезни для каждого госпитализированного или амбулаторного больного. Особенностью этой истории болезни является высокая степень оперативности информации, а также ее четкость, однозначность и унификация форм представления.

Особое внимание при построении больничных информационных систем уделяется вопросу обеспечения надежности функционирования. Для этого используются резервирование аппаратуры, технические и программные способы контроля работы.

При поступлении новых данных ЭВМ осуществляет их логическую проверку, исключает нереальные значения

параметров и запрашивает подтверждение сомнительных данных.

Такие информационные системы, помимо непосредственной лечебной работы, применяются при подготовке кадров, в частности при разборах случаев болезней, методов лечений, видов осложнений и т. д.

Для пользования подобной информационной системой не требуется длительной подготовки персонала. Эти системы удобны для работы врачей и медицинских сестер, в них широко используется обычная текстовая форма представления информации.

При построении подобных систем обеспечивается конфиденциальность информации, т. е. разрешается доступ к данным о больных только тем лицам, которым это положено. Вопрос обеспечения конфиденциальности информации в автоматизированных медицинских системах имеет важное практическое значение.

Для получения доступа к определенным данным, хранящимся в памяти ЭВМ, соответствующее лицо набирает свой персональный код или вставляет в специальное отверстие терминала свою персональную кодовую карточку. Все попытки запросов неположенной информации регистрируются ЭВМ с выдачей на печать сообщений о таких случаях.

Лечению больного обычно предшествуют диагностические обследования, позволяющие определить его заболевания. Определенные стадии обследования, такие, как сбор анамнеза и выполнение типовых анализов, могут быть сделаны перед врачебным осмотром. Более того, ряд диагностических исследований при плановом поступлении больного в стационар может и должен быть проведен в стационаре. Это может быть выполнено при условии полной преемственности лабораторных, рентгенологических и других исследований между данной поликлиникой и больницей.

Для сбора анамнеза и предварительных сведений о состоянии больного в БАИС широко используется метод типовых вопросников различного профиля (ориентированных на различные классы заболеваний). Получив заблаговременно такой вопросник, больной дома отвечает на перечисленные в нем вопросы.

Вопросники построены с учетом двух возможных видов ответов — «да», «нет». Кроме того, больной может добавлять свои замечания в виде произвольного текста.

После заполнения регистрационный бланк и вопросник возвращаются в больницу, из которой больной получает через несколько дней письмо, содержащее напечатанное ЭВМ приглашение в больницу с указанием времени и места приема. Подробнее существо задачи планирования (диспетчеризация) приема и обслуживания больных будет рассмотрено ниже.

После выполнения обследований больной поступает в больницу для лечения. В приемном отделении дежурный врач составляет предварительный план его дополнительного обследования и лечения, учитывая новую и ранее полученную информацию о больном, взятую из БАИС. Предварительный план лечения и обследования сразу же вводится в БАИС, определяющую врача и сестру, которые будут лечить больного, номер его отделения, палаты и койки. Для контрольных целей дежурный врач устанавливает ориентировочно конечную дату излечения больного. Медицинская сестра приемного отделения дает больному соответствующий вопросник и сообщает ему фамилию доктора и медицинской сестры, назначенных для его лечения. После заполнения ответов больным и ввода этих сведений в БАИС формируется первая часть машинной истории болезни стационарного больного. В дальнейшем к этим сведениям добавляются результаты обследований, сведения о состоянии больного, назначения врачей, данные о выполнении назначений и т. д. Следует заметить, что уже перед первым осмотром больного лечащим врачом этот врач должен ознакомиться с помощью дисплея со всей информацией о данном больном, имеющейся в БАИС. Это позволит врачу глубже уяснить состояние больного и наметить правильный диагноз и план лечения.

Каждый раз при поступлении больного медицинская сестра приемного отделения запрашивает БАИС, имеется ли история болезни этого больного в памяти ЭВМ. Если этот больной поступает в больницу впервые, то медицинская сестра с помощью клавиатуры дисплея вводит в ЭВМ его регистрационные данные. После ввода и проверки данных ЭВМ выдает сообщение о том, что регистрация больного выполнена. Если пациент был в этой больнице ранее, то машина сообщает прежние сведения о больном: в каком отделении он был и по какому поводу. Эта информация учитывается при госпитализации больного.

Машинное ведение историй болезней обычно сочетается с полуавтоматическим ведением бумажных историй болезней. Медицинская сестра приемного отделения сразу же заготавливает папку для сбора информации о больном в течение всего периода его пребывания в больнице. Эта папка будет играть роль обычной (бумажной) истории болезни. В папку вкладываются все распечатки, выдаваемые ЭВМ, листы регистрации, вопросники с ответами, бланки обследований, данные лабораторных анализов, заключения врачей и т. д. Сюда же записываются фамилии лечащих врача и медицинской сестры, предполагаемая дата окончания лечения. После окончания лечения все эти бумажные документы уничтожаются, но информация остается в так называемом банке данных (в машинном архиве на магнитных лентах), откуда ее всегда можно получить по запросу.

На всех документах, относящихся к больному, а также во всех машинных записях о нем присутствует его идентификационный номер, позволяющий быстро и точно находить в БАИС необходимую информацию о больном.

Параллельное ведение машинных и бумажных историй болезней обеспечивает сохранность информации в случае неисправностей ЭВМ, а также юридическую сторону ведения медицинской документации (подписи врачей). Однако это не приводит к дублированию записей и работы, так как в бумажную историю болезни вкладываются документы, напечатанные ЭВМ; врач только должен их просмотреть и подписать.

Данные о больных хранятся в БАИС в виде банка данных, который состоит из основной части, расположенной в памяти ЭВМ на магнитных дисках, и архива на магнитных лентах. Архив используется для выдачи справок и для периодической ретроспективной обработки. Основная запись историй болезни, хранимая на МД, содержит информацию, которая должна быть доступна в течение нескольких секунд. Архивные записи содержат информацию, обращения к которой совершаются 2—3 раза в день. В архиве накапливается медицинская и анкетная информация о всех предыдущих пребываниях больного в больнице. Часть этих данных может квалифицироваться как конфиденциальная. Связь между основной и архивной записями о больном обеспечивается благодаря общему идентификационному номеру. В записях историй

болезней допускается использование различных видов буквенно-цифровой информации, причем данные могут храниться как в кодированном, так и в текстовом виде. Стандартные первичные регистрационные данные и результаты обследований обычно хранятся в кодированном цифровом виде, предусмотренном стандартными бланками обследований, вопросниками для сбора анамнеза и т. д.

При заполнении подобных бланков и вопросников допускается дополнение кодированных данных произвольной информацией.

Банк данных, содержащий истории болезней, непрерывно пополняется новыми сведениями. Из этого банка можно получать по запросам как целые истории болезни, так и их определенные части: можно легко находить информацию по любому набору признаков и обрабатывать ее для получения обобщенных выводов; можно выбирать данные по комбинации симптомов физического состояния, по значениям лабораторных данных и по видам диагнозов. Для каждого набора признаков можно получать статистические оценки — абсолютные и относительные значения, средние, максимальные и минимальные значения и характеристики отклонений. Ниже будет подробно рассмотрена задача статистической обработки данных о выбывших из больничного стационара, решаемая практически в течение ряда лет в одном из американских госпиталей.

Дополнительными функциями БАИС, связанными с машинным ведением историй болезней, являются следующие.

Регулярно ЭВМ выдает контрольные листы, т. е. списки больных, которые не выписаны до назначенной контрольной даты. Эти листы позволяют привлечь внимание заведующего отделением к данным больным.

С помощью ЭВМ осуществляется оформление назначений лекарств и контроль за их выполнением. Для регистрации назначений лекарств используются предварительно отперфорированные и отпечатанные карты. Когда больному предписывается необходимое лекарство, медицинская сестра вводит определенную перфокарту вместе с перфокартой, содержащей номер больного, в ЭВМ. На основе назначений лекарств, сделанных врачами, ЭВМ составляет дневные заявки на медикаменты для всех отделений в больничную аптеку, ведет учет назначений

лекарств для всех больных, проверяет допустимость доз и сочетаний лекарств, ведет учет расходов и наличных запасов лекарств в аптеке.

Работа по применению ЭВМ для ведения историй болезней в СССР. Существующие методика и форма ведения бумажных историй болезней приспособлены к ручным способам заполнения и использования медицинских данных. Бумажная история болезни выполняет в основном две роли: роль памятных записей врачей, которые дополняют личные впечатления и сведения, хранящиеся в памяти врача, и роль юридического документа, фиксирующего состояние больного и проводимые медицинские мероприятия. Традиционная история болезни содержит в основном несистематизированную описательную информацию, которая не может быть непосредственно использована для машинной обработки, поиска ответов на запросы и выработки оптимальных решений, касающихся диагностики и лечения больного. В связи с этим в последние 10 лет во многих организациях в СССР и за рубежом усиленно ведутся работы по созданию машинно-ориентированных форм историй болезней, удобных для заполнения и использования врачами в составе автоматизированных информационных систем медицинского назначения. В СССР наиболее успешно эта работа проводится в Институте кибернетики АН УССР, во Всесоюзном научно-исследовательском институте медицинского приборостроения (ВНИИМП), а также в Институте экспериментальной и клинической хирургии Министерства здравоохранения СССР. Основным принципом построения машинной истории болезни является то, что она должна представлять собой типовую стандартную систему записей, входящих основной частью в унифицированный банк медицинской информации, предназначенный для решения разнообразных клинических, статистических и административно-хозяйственных задач в крупной многопрофильной больнице.

Выделяются три основных вида историй болезней: для стационара, для амбулаторно-поликлинического учреждения, для родильно-акушерского учреждения. История болезни стационарного больного должна иметь общую универсальную часть, включающую общие данные о больном, анамнез и некоторые другие данные, и ряд специализированных частей, соответствующих профилям отделений больниц.

Составной частью системы машинного ведения историй болезней является диалоговый режим опроса больных с помощью автоматизированных вопросников, реализуемых либо на ЭВМ в реальном масштабе времени, либо путем предварительного заполнения больным стандартизированных опросных листов. В указанных листах, как правило, допускаются ответы трех типов: «да», «нет», «не знаю». Благодаря такому характеру ответов можно построить программы обработки ответов на ЭВМ с целью выдачи сжатых сводок ответов больного для анализа врачом и для фиксации их в машинной истории болезни. Машинная история болезни имеет стандартизованную структуру, включающую основные части традиционной истории болезни, представленные в формализованном виде, позволяющем врачу заполнять их с минимальной затратой времени (путем постановки ответов «да», «нет» или цифровых данных). При этом сразу же получается закодированная информация, пригодная для перфорации и ввода в ЭВМ. Эти же данные могут вводиться с помощью клавиатур терминалов экранного типа (дисплеев).

Кроме формализованных частей, в машинной истории болезни обязательно должна присутствовать текстовая описательная информация, дающая возможность врачу выражать индивидуальные особенности больного и специфику применяемых методов лечения и диагностики. В этих записях отражаются логика мышления врача и его индивидуальный опыт и подход к данному больному. С целью автоматизации поиска и обработки подобного рода медицинских записей, в первую очередь диагнозов и предписаний врачей, должна использоваться система автоматизации индексирования, описанная достаточно подробно в главе IV. Использование текстовых медицинских записей особенно важно в сложных случаях, когда у больного имеется несколько заболеваний или осложнений, что характерно для больных пожилого возраста, удельный вес которых в общем составе больных резко увеличился за последние годы.

Машинная история болезни должна строиться с учетом общей технологической схемы прохождения больного через все стадии обследования и лечения, и в этом отношении должна быть обеспечена преемственность между историей болезни амбулаторного и стационарного больного.

При первом посещении врача в поликлинике (профилактическом центре) больной получает постоянный идентификационный шифр (номер) своей истории болезни, который является общим как для поликлиники, так и для стационара (если больной в будущем будет госпитализирован).

Первыми данными, вносимыми в историю болезни, являются данные анамнеза, лабораторных анализов, рентгеноскопии. Последующие результаты осмотров, диагностических и лечебных процедур, диагноза, предписания и т. д. в хронологическом порядке вводятся в историю болезни с помощью терминалов или перфокарт. После выписки больного сжата сводка основных данных вводится в специальный статистический файл, а полные данные переносятся из текущего накопителя на магнитных дисках в накопитель на магнитных лентах (в машинный архив) для длительного хранения.

Во время нахождения истории болезни в текущем накопителе (что соответствует периодам амбулаторного или стационарного обследования и лечения больного) она используется для выдачи оперативной информации врачам по следующим основным разделам:

а) общие данные больного, группа крови, лекарственная переносимость, идентификационный номер, данные анамнеза;

б) диагнозы, врачебные заключения, включая записи из диагностических кабинетов и лабораторий;

в) врачебные назначения и сведения о их выполнении и результатах, в первую очередь сведения о лекарствах;

г) проведенные операции, осложнения, результаты лечения;

д) все патологические проявления.

Весь текущий файл историй болезней больных данного учреждения должен строиться в виде ИПС с использованием ассоциативно-адресных способов организации поисковых массивов, обеспечивающих быстроту поиска данных и их корректировки и пополнения. Этот файл используется не только для лечебных целей, но и для планирования работы больницы и для административно-хозяйственных расчетов. Основными требованиями к машинной истории болезни являются удобство использования ее непосредственно лечащими врачами и надежность функционирования. Последнее обстоятельство требует периодической выдачи на печать основных данных

машинной истории болезни и хранения их в виде бумажных документов на случай выхода из строя ЭВМ и других технических средств, а также применения при ведении машинной истории болезни разветвленной системы логического и формального контроля входных и выходных показателей.

Важным требованием к машинной истории болезни является сочетание полноты информации с селективностью и компактностью выдаваемых ответов. Это достигается за счет иерархической структуризации историй болезни. Программная реализация массива машинных историй болезней может быть осуществлена в виде ИПС, описанной в главе IV, с использованием методов обработки формализованных фраз и автоматической индексации документов.

Существенным требованием к машинному ведению историй болезней является обеспечение конфиденциальности информации о больных и разделение ее на виды в соответствии с различием в категориях персонала, имеющего доступ к этой информации. Для этих целей применяются специальные жетоны — ключи, вставляемые в терминалы и идентифицирующие личность запрашивающего лица, а также специальные условные шифры или слова, вводимые в ЭВМ с терминалов в виде паролей.

Рассмотрим основные принципы создания машинных историй болезней, разработанные в Институте кибернетики АН УССР.

Первым этапом построения системы сбора и обработки медицинской информации является домашняя обработка формализованных историй болезней (ФИБ), заключающаяся в построении иерархической классификации историй болезней по клиническим профилям и в стандартизации признаков, симптомов, синдромов, диагнозов, методов лечения, методов исследования. Сюда же относятся разработка клинических классификаторов болезней, согласованных с Международным классификатором ВОЗ; разработка номенклатур лекарственных средств; анализ потоков информации в типовых больницах и определение структуры медицинской информационной системы, в которой будет использована данная ФИБ; обучение кадров и подготовка инструктивно-методических материалов для работы с ФИБ.

Вторым этапом является непосредственное использование ФИБ в практике работы данного лечебного учреж-

дения, связанное с последовательным формированием банка медицинской информации и решением на его основе лечебных, административных и научно-статистических задач.

Для амбулаторных больных должны применяться в качестве истории болезни формализованные амбулаторные карты (ФАК), обеспечивающие удобство работы врачей амбулаторно-поликлинических учреждений и пригодные для машинной обработки.

В ФАК должны фиксироваться данные всех посещений, все диагнозы и врачебные назначения в хронологическом порядке, особенно назначения лекарств и данные о их приеме.

Данные основных лабораторных анализов (мочи, крови, желудочного сока, желчи) должны фиксироваться последовательно, отражая характер их изменения во времени.

Записи о временной и стойкой нетрудоспособности, записи ВКК, ВТЭК, направления на госпитализацию должны быть формализованы и иметь постоянное поле (раздел) ФАК. В ФАК должны переноситься основные сведения из формализованной амбулаторной карты развития ребенка (ФАКРР). При направлении в стационар из ФАК должны передаваться необходимые сведения в ФИБ. Кроме того, лечащий врач стационара должен иметь возможность запросить любые более подробные сведения из ФАК, если они ему потребуются. Особенности ФАКРР являются большая динамика изменения основных антропометрических показателей, подробный учет прививок и реакций на них, учет особенностей формирования личности ребенка. Формализованная амбулаторная карта диспансерного наблюдения за беременной (ФАК ДНБ), помимо идентификатора и общих данных, должна содержать диагнозы перенесенных заболеваний (взятые из ФАК и ФИБ), данные социологических обследований, записи консультаций специалистов и подробные сведения о ходе беременности и родах.

На основе перечисленных машинно-ориентированных документов должна формироваться в будущем унифицированная история здоровья гражданина СССР, которая будет содержать всю информацию о здоровье человека начиная с рождения и кончая смертью.

Важными требованиями к этой истории здоровья являются соблюдение преемственности при обращениях в

различные лечебные учреждения и при перемене места жительства, а также высокая надежность хранения информации.

2. Автоматизированная система накопления и обработки данных о выбывших больных в госпитале

В качестве примера практически действующей системы обработки информации о выбывших больных рассмотрим систему, реализованную в одном из американских госпиталей Вест Джеферсон. Этот госпиталь обслуживает пригороды Нового Орлеана, имеет 400 коек, в 1973 г. принял 15 тыс. больных, кроме того, 30 тыс. было принято в отделении неотложной помощи.

Бюджет госпиталя составляет 12,5 млн. долл. США в год. Области применения ЭВМ в данном госпитале являются обработка выписок из историй болезней, экономические расчеты, зарплата, учет кадров. Опишем основную задачу, решаемую с помощью ЭВМ в данном госпитале, — сбор и обработку выписок из историй болезней. В ноябре 1972 г. в этом госпитале была введена в действие ЭВМ типа ИБМ 370 модель 135, на которой с 1973 г. решается указанная задача.

Система накопления и обработки выписок из историй болезней была разработана фирмой ИБМ как пакет пользовательских программ 57-96-AFA.

Исходной информацией для решения указанной задачи являются сообщения о всех больных, выписанных из госпиталя или обслуженных неотложной помощью. Такие сообщения подготавливаются к вводу в ЭВМ специальной группой работников медицинского архива. Эта группа отвечает также за составление периодических сводок о движении больных и работе врачей, за анализ и контроль данных, включенных в истории болезней. Перечисленная информация используется как администрацией, так и врачами госпиталя для лечебных, организационных и научных целей.

Раньше обработка выписок велась с помощью ручных карт. Составлялись списки различных больных, списки умерших, сводки выполненных процедур и работы врачей. Эта трудоемкая ручная работа была сопряжена с большим количеством ошибок и не позволяла использовать сложные критерии для выборки и анализа данных. Например, если нужно было получить распределе-

ние по возрасту больных с определенным заболеванием, которые лечились в определенное время и у определенного врача, то требовалось сделать большое число ручных операций поиска и сортировки, что практически было невозможно.

При внедрении ЭВМ в данном госпитале ставились следующие цели:

1) все сводки и списки, ранее составлявшиеся вручную, должны готовиться с помощью ЭВМ,

2) сведения о выписанных больных и о работе неотложной помощи должны обрабатываться ЭВМ для получения оценок работы различных отделений и служб госпиталя,

3) на основе применения ЭВМ с сетью терминалов в госпитале должна быть сокращена ручная работа по медицинской статистике.

Благодаря однократному вводу и соответствующей проверке повышается точность вводимых данных. Конфиденциальность информации обеспечивается тем, что доступ к определенным данным имеет только соответствующий медицинский персонал. ЭВМ позволяет выбирать данные и составлять списки, определять суммарные и сравнительные данные, выяснять корреляционные связи.

Для всех больных в больнице и в службе неотложной помощи с помощью ЭВМ ведется общий журнал учета, выписки и перевода, который ежедневно обновляется и используется в основном бухгалтерией госпиталя. На основе этих ежедневных общих учетных данных ЭВМ составляет рабочий лист карт выбывших больных. На этой стадии карта выбывшего больного, представленная на рабочем листе, заполнена только частично. Она содержит лишь данные, полученные из общего бухгалтерского журнала учета больных. Такой рабочий лист посылается в медицинский архив, где дополняется из соответствующей истории болезни, заполняемой в основном медицинскими сестрами. Ответственная за ведение историй болезней медицинская сестра переписывает данные, содержащиеся на первой странице истории болезни, в рабочий лист. На этой стадии рабочий лист заполняется почти полностью. В нем отсутствуют коды болезней и коды операций. Эта информация будет внесена в рабочий лист после того, как лечащий врач полностью заполнит историю болезни.

Рабочий лист передают оператору терминала, который вводит данные с него в массив карт выбывших больных. Оператор вызывает на экран формат карты и вводит информацию в соответствующие поля карты, указывая номер больного и признак того, что карта содержит полные данные (в том числе код болезни и код операции) или что она содержит неполные данные. После визуальной проверки (на экране) оператор вводит карту в машинный массив и переходит к заполнению следующей карты, набирая соответствующий код карты и номер больного.

Рабочий лист, с которого производился ввод данных в ЭВМ, вкладывается в историю болезни, которая передается в архив на хранение. Ниже приводятся формы (состав вопросов) карты выбывшего больного и первой страницы истории болезни, принятые в этом госпитале.

Форма рабочего листа карты выбывшего больного

- Фамилия, имя, отчество
- Адрес
- Номер больного
- Служба приема больницы
- Дата рождения
- Количество дней в отделении интенсивной терапии
- Раса
- Религия
- Пол
- Семейное положение
- Финансовая категория
- Дата приема
- Дата выписки
- Количество дней отсутствия
- Служба выписки больницы
- Лечащий врач
- Второй врач
- Врачи консультировавшие
- Тип выписки
- Послеоперационная смерть
- Коронарная болезнь
- Коды болезней (до 5, по 6 знаков с точкой)
- Коды операций (до 4)
- Повторный прием
- Тип приема
- Переливание крови ($\frac{\text{да}}{\text{нет}}$)
- Тип опухоли
- Тип анестезии
- Лабораторные анализы ($\frac{\text{да}}{\text{нет}}$)
- Даты операций (до 6)

Форма первого листа истории болезни

Общая часть

Номер больного, фамилия, имя, отчество

Номер палаты, фамилия врача, номер врача, возраст больного, адрес, телефон, семейное положение (Ж — женат, О — одинок, В — вдов, Р — разведен, Л — живет отдельно).

Пол, раса (белый, черный, цветной), девичья фамилия, предыдущая

госпитализация $\left(\begin{array}{c} \text{да} \\ \text{нет} \end{array} \right)$ и ее дата и причина, место рождения, дата рождения.

Ближайший родственник, степень родства, телефон.

Религия, занятие, место работы, с какого времени, адрес и телефон места работы.

Дата приема, время приема, дата выписки.

Диагноз при приеме	код
Окончательный диагноз	код
Вторичный диагноз или осложнения	код
Операции	коды
Причина смерти	
Консультанты	

Выписан: живой	умерший	вскрытие: $\frac{\text{да}}{\text{нет}}$
----------------	---------	--

Проверка комиссией по историям болезней	Я проверил и подтверждаю эту заполненную историю болезни на на 19
---	---

подпись:

Лечащий врач

База данных карт выбывших больных состоит из трех массивов. 1. Собственно массив карт выбывших из стационара. Эти карты вводятся в массив в момент выписки больного из больницы. Массив содержит необходимую информацию для составления разных сводок и выдачи справок на текущие запросы. К массиву могут обращаться с помощью терминалов допущенные лица. Карты хранятся в массиве на магнитных дисках 18 мес и затем переписываются на магнитную ленту для архивного хранения. Состав данных карты соответствует рабочему листу карты выписанного больного, приведенному выше.

2. Массив карт амбулаторных больных. Содержит карты для каждого обращения больного в службу неотложной помощи данного госпиталя. Состав информации в этой карте почти такой же, что и в карте выбывшего из стационара.

3. Общий справочный массив всех больных. Содержит записи для всех стационарных и всех амбулаторных больных, расположенные в алфавитном порядке фамилий, и используется в основном для выдачи справок о больных по запросам, вводимым с терминалов лечащими врачами, врачами-консультантами и администрацией больницы.

Выходные документы (отчеты). Имеются два вида выходных документов, выдаваемых системой:

— Регулярные месячные отчеты, используемые администрацией больницы, и указатели (списки), используемые вышестоящим органом здравоохранения.

— Целевые справки-доклады, выдаваемые по запросам отдельных врачей или комиссий для специального изучения болезней.

Месячные отчеты выдаются ЭВМ на 2-й день после окончания месяца, причем в них учитываются все выписанные за месяц больные (даже если на некоторых из них еще не введены в ЭВМ коды болезней, операций и осложнений).

Ниже приводится состав информации в пяти различных формах месячных отчетов, выдаваемых системой.

1. Статистика выписанных больных (за месяц и с начала года). Общее число выписанных больных, в том числе взрослых и детей, новорожденных; из них белых, черных, других. Общее число койко-дней, в том числе взрослых и детей, новорожденных.

Средняя длительность пребывания: взрослые и дети, новорожденные.

Общая смертность.

Общее число вскрытий, уровень кросс-вскрытий.

2. Частотное распределение выписанных больных по районам, из которых поступили больные за месяц и с начала года.

3. Регистр смертности, содержащий перечисление умерших за месяц. Указываются дата смерти, номер и фамилия больного, код причины смерти, до или после 48 ч, пол, раса, семейное положение, дата рождения, служба (отделение) больницы, лечащий врач (номер).

4. Анализ работы врачей, показывающий для каждого врача число выписанных больных и общее число койко-дней за месяц с начала года.

5. Врачебный перечень, представляющий собой подробный список больных каждого врача с указанием но-

мера и фамилии больного, пола, расы, возраста, даты выписки, длительности пребывания, финансового положения, выписавшего отделения (службы), второго врача (его номера), консультировавших врачей (их номеров).

Целевые справки-доклады. Одна из наиболее важных возможностей системы состоит в выдаче по запросам целевых справок для анализа работы госпиталя и структуры заболеваемости. Запрашивающий может выбрать один или несколько из следующих семи типов справок.

1. Структура (профиль) больных: распределение больных по полу, виду госпитализации (обычная, неотложная, срочная), общее число койко-дней, продолжительность пребывания (средняя, максимальная, минимальная), возраст больных (минимальный, максимальный до 30 лет, 30—39 лет, 40—49 лет, 50—59 лет, 60—69 лет, старше 69 лет), прием и выписка по дням недели с указанием средней продолжительности пребывания.

2. Средняя длительность пребывания по операциям. Указываются код операции, код хирурга, число выписанных, число койко-дней, число умерших, средняя длительность операции.

3. Средняя длительность пребывания по болезням. Данные суммируются для каждого врача в пределах кода болезни. Если поиск ограничен, например, одним кодом операции, то в справке фигурируют только те болезни, для которых эта операция является частью лечебного процесса. Указываются код болезни, код врача, число выписанных, число дней, средняя длительность пребывания, число умерших.

4. Частотное распределение по болезням. Показывается, сколько раз данный код болезни появляется как первичный и как вторичный диагноз в заданной группе больных. Указываются код болезни, ее название, число первичных диагнозов, число вторичных диагнозов, общее число диагнозов.

5. Частотное распределение по операциям. Код и название операции и число.

6. Операционный перечень (указатель), содержащий подробные сведения о каждом больном, подвергшемся операции,— код и название операции, номер и фамилия больного, продолжительность пребывания, дата выписки, возраст, пол, тип выписки, номер хирурга, раса, коды операций (основной и дополнительных). Подсчитываются число операций и средняя длительность пребывания.

7. Перечень (указатель) больных по видам болезней. Указываются код болезни с признаком (основная — первичная или сопутствующая — вторичная), номер и имя больного, длительность пребывания, дата выписки, возраст, пол, тип выписки, номер лечащего врача, раса больного, коды болезней (всех у данного больного), коды операций.

Для каждой болезни подводится итог: всего больных с основным диагнозом и средняя длительность их пребывания, всего с сопутствующим диагнозом и средняя длительность пребывания, общее число и средняя общая длительность. Один и тот же больной может появиться в этом указателе несколько раз, если он имеет несколько болезней, и каждый раз будет указано, основная это болезнь или сопутствующая.

Ввод запросов с терминалов на выдачу целевых справок. Запрос должен содержать указание типа справки, которую требуется получить, и набор критериев (признаков) для отбора группы больных, по которым будет составляться требуемая справка. В качестве таких критериев можно указывать период времени, фамилию врача (в том числе принимавшего, консультировавшего, хирурга), основной диагноз, сопутствующий диагноз выполненную операцию (операции), продолжительность пребывания в днях, возраст больных.

Можно указывать сразу несколько из этих признаков, и они при поиске будут объединяться в запросе с помощью логической операции конъюнкции («и»).

Для ввода запроса на экране терминала изображается стандартная форма запроса, содержащая наименования перечисленных признаков и пустые поля для их заполнения автором запроса. Универсальная программа — генератор рапортов — выполняет анализ запроса, производит поиск и отбор нужных больных, согласно заданным признакам, суммирует и выдает на экран требуемую целевую справку. В верхней части этой справки указываются заданные критерии поиска. Гибкость такой программы позволяет пользователям — врачам применять разные виды статистического анализа.

Для вызова на экран карты выбывшего больного оператор должен ввести сначала свой код права доступа, потом код вида вызываемого сообщения и затем идентификационный номер больного. В данной системе идентификационные номера присваиваются при каждом слу-

чае госпитализации или обращения, а также оказания неотложной помощи, поэтому по заданному идентификационному номеру больного выдается одна карта, соответствующая этому случаю обращения. Состав отображаемой карты соответствует составу данных рабочего листа карты больного.

Иногда идентификационный номер больного неизвестен, или требуется выдать все записи обращений определенного больного в больницу, или нужно просмотреть данные, содержащиеся в общем справочном массиве всех больных. В этих случаях оператор набирает идентификатор запроса на поиск в справочном массиве больных. ЭВМ в ответ выдает на экран перечень наименований признаков для поиска больного (больных), которые должен заполнить запрашивающий: фамилия, год рождения (в пределах 5 лет), пол, раса больного. Обязательной для поиска в общем справочном массиве больных является фамилия больного; остальные признаки могут задаваться (если они известны) или не задаваться. Но даже если фамилия больного задана неточно, ЭВМ производит охватывающий поиск, выдавая данные всех больных, фамилии которых начинаются на те буквы, которые заданы для поиска в качестве фамилии больного. В результате поиска выдаются на экран данные ряда больных, причем каждая строка соответствует одному больному и содержит его фамилию, год рождения, пол, расу больного. Строки на экране перенумерованы подряд, и если оператор хочет вызвать карту некоторого больного из перечисленных на экране, он должен набрать номер этой строки. Если число больных не помещается на экране, то ЭВМ указывает оператору, что он может вызвать следующую страницу экрана и т. д. Практическая эксплуатация описанной автоматизированной системы обработки карт выживших больных (стационарных и амбулаторных) показала ее эффективность, а также позволила определить пути ее дальнейшего развития.

Перспектива развития системы заключается в дополнении карт выживших клиническими данными и во включении в ее состав процессов приема и перемещения больных.

Недостатком данной системы является сравнительно большой объем ручной работы, связанной с заполнением рабочих листов, ручным ведением историй болезней и подготовкой карт выживших больных по вводу в ЭВМ.

3. Диспетчеризация и учет медицинского обслуживания больных с помощью ЭВМ

Применение современных методов диагностики и лечения невозможно без квалифицированного персонала и дорогостоящего оборудования, количество которого в больших больницах исчисляется десятками и сотнями единиц. При большом числе больных неизбежно проведение в нескольких местах одинаковых анализов и процедур. Без автоматической системы управления и контроля очень трудно обеспечить эффективное использование персонала и оборудования, а также равномерное распределение нагрузки между ними. Для решения этой проблемы необходимо использовать ЭВМ с сетью терминалов, обеспечивающей эффективную оперативную связь между различными пунктами лечебного учреждения.

Обычно планированием медицинского обслуживания больных занимается квалифицированный медицинский персонал, время которого следует использовать с возможно большей эффективностью.

Обеспечение эффективной работы медицинского учреждения невозможно без правильного использования коек. Плохое планирование может привести к тому, что большое число коек окажется занятым больными, ожидающими обследования.

Качество обслуживания больного повышается при сведении к минимуму времени на ожидание процедур или на переходы из одного подразделения в другое.

Описываемая ниже система машинного планирования медицинского обслуживания больных была разработана Шведским агентством административного планирования совместно с Каролинским госпиталем и фирмой ИБМ в 1971 г.

Составление расписаний работы лабораторий, кабинетов, специалистов осуществляется ЭВМ по специальной программе на основе двух массивов информации: постоянного, содержащего сведения о составе и пропускной способности этих служб больницы, и переменного, содержащего все заказы на текущее медицинское обслуживание больных. Эти заказы вводятся ежедневно в ЭВМ с помощью терминалов, находящихся в разных отделениях больницы.

Постоянные массивы, сохраняемые на магнитных дисках, содержат такую информацию, как длительность

каждого обследования, перечень установок, с помощью которых может быть выполнено каждое обследование, периоды времени, предшествующие обследованию и следующие за ним, а также периоды времени, необходимые для проведения обследований на каждой установке. Эти массивы составляются заранее и могут по мере необходимости корректироваться. Переменные параметры, которые используются программой для составления расписаний, зависят от специфики больного и назначений врача: состояния здоровья, срочности назначения, других процедур, предписанных больному, и т. д.

Варианты планов-графиков обслуживания больных, составленные ЭВМ, выдаются на экраны дисплеев, с тем чтобы врач или медицинская сестра могли выбрать подходящий вариант. Отобранные варианты печатаются ЭВМ и рассылаются в отделения и в соответствующие кабинеты и лаборатории для выполнения. В случае необходимости имеется возможность отменить выполнение той или иной процедуры для данного больного или перенести срок ее выполнения. С помощью дисплея можно запросить ЭВМ в любой момент о степени загрузки той или иной диагностической или лечебной установки, о загрузке врачей-специалистов, о ходе выполнения составленных планов-графиков.

В результате решения этой задачи ЭВМ выдает ежедневно планы-графики работы каждой диагностической или лечебной установки и листы расписаний дневного обслуживания отдельного больного.

Для решения этой задачи требуется создание следующих файлов на магнитных дисках.

Файл имеющегося времени содержит данные о пропускной способности установки и о времени, которое может быть использовано для выполнения назначений на каждой установке. Он охватывает период в 3 мес.

Файл процедур содержит данные о длительности типовых диагностических и лечебных процедур таких, как рентгенография и рентгеноскопия, электрокардиография, лабораторные клинические и биохимические анализы, лучевая терапия и т. д. В этом же файле содержатся данные о необходимом предпроцедурном и послепроцедурном времени, а также о регулярности (повторяемости) некоторых процедур. Здесь же указываются установки (их коды), на которых выполняются эти процедуры.

Файл характеристик установок содержит специфическую информацию по каждой установке: название, время для экстренных назначений, координаты размещения установки по отношению к другим установкам. Информация файла располагается в соответствии с кодами установок; те же коды установок используются в записях файлов процедур.

Файл кодов для выбора сочетаний процедур (установок) содержит информацию по комбинационным кодам, представляющим группы установок.

Файл текущих назначений содержит сводку всех заказов на обслуживание, т. е. перечень планируемых назначений по конкретным больным.

Ввод данных текущих назначений в ЭВМ для каждого больного осуществляет оператор (медицинская сестра) с помощью дисплея. На экране дисплея высвечивается стандартная форма заявки на обследование, содержащая 7 строк. Первая строка служит для заполнения идентификационного номера больного. В следующих двух строках указываются две возможные группы процедур обследования и необходимые перерывы после них. Четвертая строка предназначена для указания третьей возможной группы обследований (без указания последующего перерыва). В пятой строке приводится комбинационный код для выбора группы установок. Седьмая строка служит для указаний, касающихся дат проведения процедур или их срочности.

Идентификационный номер присваивается больному при его поступлении в больницу и вводе его данных в ЭВМ. По этому номеру ЭВМ может выдать фамилию, адрес и другие сведения о больном, а также указать отделение, палату (и койку), в которых больной находится. Следует заметить, что в Швеции все граждане имеют постоянные регистрационные номера, которые обычно используются в качестве идентификационных номеров больных. Это 10-разрядный код, состоящий из данных рождения (год, месяц, день), 3-разрядного порядкового номера и контрольного разряда. Если такой номер отсутствует, то в качестве идентификационного номера берется временный 10-разрядный номер, содержащий две последние цифры года, когда больной лечился в больнице в первый раз, 2-разрядный временный номер (между 90 и 99), 5-разрядный порядковый номер и контрольный разряд. Каждая процедура представляется 4-разрядным

кодом. Группа обследований включает обследования, которые должны быть запланированы на один и тот же день. В каждую группу могут включаться 1—3 обследования. Код каждого обследования отделяется запятой. На одну форму заявки может быть записано до 9 обследований одного больного. Если необходимы добавочные назначения, то для их ввода может быть использован любой добавочный номер.

Код перерыва, включаемый после одного или обоих первых двух групп обследований, ограничивает порядок выполнения назначений. Обычно каждая «группа» назначений планируется на различные дни и на самый ранний возможный день для каждой группы. Путем ввода цифровой величины от 1 до 96 в качестве кода перерыва две группы процедур разделяются соответствующим числом дней. Код перерыва «0» или «А» разрешает планировать обе группы процедур на один и тот же день. Разница между этими двумя кодами заключается в том, что код «0» требует проведения всех обследований первой группы до проведения обследований второй группы, а код «А» указывает на отсутствие необходимости такого порядка. Обычно система выбирает одну установку из группы всех тех, на которых может быть проведен данный вид обследования. Если это нежелательно, выбор может быть ограничен определенным типом установок. Такой выбор проводится путем ввода кода, обозначающего определенную установку или группу установок.

При составлении расписаний алгоритм предусматривает использование первого же свободного дня. День, в который проводится решение задачи, не считается доступным днем. Для задания специальных условий на дни проведения процедур имеются три возможности. Если вводят ГГММДД¹, то система рассматривает только дни после указанной даты. Если задано ГГММДД, то рассматриваются дни до этой даты. Задание двух ГГММДД—ГГММДД определяет период, в течение которого должна быть выполнена данная процедура.

Для каждой установки может быть зарезервировано время с расчетом на чрезвычайный случай. Путем ввода специального указания срочности (символ F) после спецификации дат обеспечивается первоочередное обслуживание данного больного.

¹ ГГ — две последние цифры года, ММ — номер месяца, ДД — номер дня.

После того как заполненная форма посылается с терминала, программа формирует таблицу, в которой содержатся основные данные по каждому назначению. Каждая таблица содержит возможное время назначения, которое остается свободным после устранения периодов, ограниченных параметрами терминала. В случае взаимосвязанных назначений выбираются три самых ранних по времени дня из всех возможных. В случае взаимосвязанных назначений, которые должны выполняться в один и тот же день, выбирается оптимальное решение, т. е. самый ранний день, на который могут быть запланированы все обследования и время выполнения, обеспечивающие самые маленькие интервалы между процедурами.

Для каждой группы процедур посылаются на дисплей три альтернативных варианта. Оператор терминала выбирает соответствующий вариант для каждой группы, и этот вариант регистрируется в файле назначений. Оставшиеся альтернативы хранятся в файле свободного времени. Процесс решения завершается выдачей на печать выбранного варианта выполнения процедур. Рассмотрим следующий пример.

Предположим, что в форме заявки зафиксированы следующие обследования:

1101 рентгенологическое обследование плеча
1123 рентгенологическое обследование сердца и легких
0901 химический лабораторный анализ
0525 ЭКГ
0010 консультация у врача (д-р Андерсон)

Форма заявки на экране терминала будет иметь вид:

Δ 3305067369
1101, 1123, 0901, 0Δ
0525, 3Δ
0010Δ
*, *, *, *, 1050Δ
700112 — Δ

Здесь в первой строке указан идентификационный номер больного, во второй — первая группа процедур, в третьей — вторая группа процедур и в четвертой строке — третья группа процедур. Пятая строка содержит код выбора конкретной установки, шестая — указание по ограничению дат проведения процедур.

Первые четыре обследования должны быть сделаны в один и тот же день, но анализ ЭКГ должен быть последним видом обследования в этот день. Это указывается кодом интервала «0». Интервал не менее 3 дней между первыми четырьмя обследованиями и визитом к

врачу указан кодом интервала «3». Код отбора установки «1050» ограничил выбор среди всех возможных врачей доктором Андерсоном. Спецификация даты показывает, что назначения не должны начинаться до 12 января 1970 г.

После ввода заполненной формы заявки в ЭВМ она рассчитывает три возможных варианта обслуживания и выдает их на экран терминала. Расчет вариантов производится из условия, чтобы интервалы между процедурами были по возможности меньшими. При этом учитывается информация, имеющаяся в файле обследования: время, которое должно предшествовать процедуре; время, которое необходимо выдержать после процедуры; время, необходимое на проведение процедуры. Если не подходит ни один из трех вариантов, то можно дать команду ЭВМ рассчитать три новых варианта. После того как выбран подходящий вариант и указания об этом введено с терминала в ЭВМ, система включает этот вариант в файл назначений.

В данном примере ЭВМ напечатает следующее извещение о выбранном варианте выполнения процедур, которое будет направлено в отделение больному или по почте амбулаторному больному на дом:

Уведомление о назначениях для 3305067369, г. Стокгольм

*Кинг стрит, 48
Джонс Мэри*

Для Вас установлены следующие назначения. В случае если Вы не сможете явиться, сообщите об этом, пожалуйста, в больницу как можно быстрее.

Химический анализ 12 января 8 ч 30 мин

Медицинская химическая лаборатория Клиники ОРД, 1-й этаж

Рентгенологическое обследование плеча
12 января 9 ч 00 мин

Центральная рентгенологическая лаборатория, основное здание, 3-й этаж

Рентгенологическое обследование сердца и легких 12 января 9 ч 15 мин

Центральная рентгенологическая лаборатория, основное здание, 2-й этаж

ЭКГ 12 января 10 ч 00 мин

Центральная физическая лабора-

Первый визит к врачу 16 января 10 ч
30 мин

тория, основное
здание, 1-й этаж
Отделение для ам-
булаторных боль-
ных, д-р Андер-
сон, 2-й этаж

В данной системе предусматривается возможность отмены каких-либо назначений. Для этого на экран терминала вызывается список всех назначений для данного больного и путем ввода с терминала специального кода (буква N) в строку отменяемого назначения ЭВМ сообщается информация об отмене этого назначения. Время, выделенное для данного назначения, переводится в резерв свободного времени данной установки или врача-специалиста.

На основе решения данной задачи ЭВМ выдает также дневные графики загрузки диагностических и лечебных установок (кабинетов, лабораторий) и списки приема больных каждым врачом-специалистом (консультантом). Эти сведения могут высвечиваться на экране терминала и выдаваться на печатающее устройство. Например, список приема больных врачом-консультантом может иметь вид:

Назначения 70.01.16

Место : д-р Андерсон, ОРД

Консультация, первый визит	1204290710	Хемр Рудольф	9.00
Специальная консультация	2605071006	Андерс Пиа	9.30
Консультация, первый визит	3305067369	Джонс Мэри	10.30
Консультация, повторный визит	021203225	Прус Ив	14.00
»	»	» 2509258998	Хансон Джо 14.15
»	»	» 2602277184	Торас Лаз 14.30

Это все назначения.

Можно запросить ЭВМ выдать график свободного времени для каждой установки по дням месяца и по часам работы (начиная с 7.00 до 20.00).

Система допускает также ввод с терминала заранее определенных времени и места (установки, кабинета) конкретного назначения. Такое назначение включается в файл назначений, и соответствующее время исключается из резерва свободного времени. Эта возможность чаще всего используется для ввода назначений больным, находящимся в списках ожидания, т. е. таким больным, которым сделаны предварительные назначения вручную, выходящие за пределы трехмесячного периода планирования назначений с помощью ЭВМ. Когда предваритель-

но назначенное время входит в этот период, то с терминала данное назначение включается в файл реальных назначений.

Ряд не связанных с системой отдельных программ поддерживает работу программ, помогая содержать файлы на соответствующем уровне.

Как указывалось раньше, файл имеющегося времени охватывает период 3 мес. Содержание этого файла пересматривается в начале каждого месяца, когда файл сдвигают, отбрасывая «первый» месяц и прибавляя новый «последний» месяц. Отдельные изменения в соответствии с периодами времени, зарегистрированными в файле, могут быть внесены посредством внутрисистемной программы корректировки файлов.

Файл назначений должен время от времени уплотняться. Из него должны исключаться все неактивные назначения (выполненные или отмененные), которые помещаются в статистический файл для анализа загрузки установок и врачей, сроков и стоимости обслуживания и т. п.

Опыт двухлетнего использования данной системы планирования медицинского обслуживания показал ее действенность. Ввиду трудностей, связанных с учетом деятельности персонала получить точные статистические данные о ее эффективности пока не удалось.

Внедрение системы привело к упрощению связи внутри больницы и повышению ее надежности, к сокращению затрат времени на планирование обслуживания больных, более равномерной загрузке установок и врачей, уменьшению среднего срока пребывания больных в больнице, улучшению контроля за использованием оборудования и более обоснованному приобретению нового оборудования.

4. Автоматизация сбора и обработки информации о состоянии больного во время операции и в период интенсивной терапии

Особенно важно быстрое и точное получение информации о больном во время операции и в период интенсивного ухода. В течение операции может фиксироваться 1000 и больше значений различных параметров, характеризующих состояние больного. Быстро сопоставлять и выделять нужные данные довольно сложно, осо-

бенно в критических ситуациях, когда основное внимание сосредоточено на самом больном. Трудно при этом учесть и предшествующие записи в историях болезней, данные о наличии медикаментов и т. д.

При использовании ЭВМ с дисплеями данные о текущем состоянии больного (артериальное давление, пульс, температура, газы крови, химические данные и др.) могут наблюдаться на экранах дисплеев в операционной и в других пунктах отделения и больницы в виде графиков и таблиц. При необходимости можно вызвать другие относящиеся к больному данные, изменить масштаб кривых, отпечатать копии данных, отображаемых на экране.

Если какие-либо данные выходят за критические пределы, то ЭВМ дает звуковые и световые предупреждающие сигналы дежурным с выдачей на экран информации, поясняющей причину тревоги.

Так как одни и те же данные могут выдаваться ЭВМ одновременно на разные дисплеи, то представляется возможность срочных обсуждений сложных вопросов по телефону, без сбора заинтересованных лиц в одной комнате.

Врач может заблаговременно ввести в память ЭВМ указания, касающиеся необходимых мер в критических ситуациях, и эти указания будут в нужный момент автоматически выданы на экраны дисплеев в тех пунктах, где это нужно. Особенно важными являются данные, полученные во время операции и в процессе интенсивного лечения, так как они отражают динамику лечения тяжелобольных. Поэтому эти данные должны быть доступны заинтересованным лицам как можно быстрее — практически в момент их появления. Эти сведения позволяют принимать быстрые решения. Кроме того, такая система используется для подготовки персонала и исследовательских работ.

Система осуществляет в основном сбор, накопление и выдачу данных; обработка их играет меньшую роль. Информация может вводиться в любое время дня и ночи из всех пунктов, где установлены соответствующие терминалы. Данные выдаются главным образом на терминалы экранного типа. Кроме того, могут получаться в необходимых случаях печатные копии отображаемой на экранах информации, которые вклеиваются в истории болезней. Печатные данные содержат обычно дату ввода.

Регистрация больных с помощью дисплея ведется путем заполнения стандартных форм, представляемых на экране. ЭВМ проверяет аббревиатуру ответов и в случае ошибок печатает «Не определено». При правильной аббревиатуре специальный световой указатель дисплея, называемый меткой, или курсором, передвигается и указывает проверенные позиции. ЭВМ проверяет также вводимые значения нахождение их в допустимых пределах, и в случае выхода их за эти пределы выдает их на экран с примечанием «Ненормально». Оператор может подтвердить эти ненормальные значения путем повторного ввода.

Кроме патологических значений, ЭВМ может обнаружить и неприемлемые значения. Эти значения выдаются ЭВМ на экран с соответствующим примечанием. Они не могут быть введены в память ЭВМ даже путем повторного подтверждения. Пределы значений, которые должны считаться ненормальными или невозможными, заранее задаются в ЭВМ врачами и при необходимости могут корректироваться с использованием дисплея. Врач может вводить свободный текст с помощью клавиатуры дисплея. Вводимые данные о состоянии больных хранятся на магнитных дисках в специализированном банке данных, состоящем из двух частей: 1) массива цифровых данных, снимаемых с определенной частотой и характеризующих параметры организма (температура, давление и т. д.), 2) массива текстовых данных (заключения врачей о состоянии больного, предписания врачей, замечания по ходу операций или ухода и т. д.).

Выдача данных осуществляется по командам с дисплея. Выдаваемые данные изображаются набором форм, организованных в древовидную иерархическую структуру. С помощью специальных команд — указаний, подаваемых с дисплея, можно двигаться по этому дереву вверх или вниз, вызывая более общие или более частные формы. На каждой форме отдельные участки имеют свои сокращенные наименования, указывая которые можно вызвать более детальное изображение этих участков. Отдельные части форм включают в себя кривые, таблицы и их комбинации, а также словесные пояснения, помещаемые внизу формы. Такое расположение позволяет отобразить связь табличных или графических данных со словесными пояснениями. Например, форма для представления больного предусматривает выдачу в ви-

де кривых, зависящих от времени: систолического артериального давления, центрального венозного давления, частоты пульса, температуры. Имеется возможность в любой момент прочитать в цифровой форме ранее сделанные измерения, т. е. цифровые значения в любой точке кривой. Для этого с помощью специальной клавиши на пульте и светового пера оператор переводит ось «У» (вертикальную ось графика) в нужную точку оси времени «Х». Справа от оси У в виде колонки цифр будут выданы соответствующие показатели. При желании можно увеличить или уменьшить масштаб по оси времени. Два последних значения каждой переменной всегда изображаются в цифровом виде в двух колонках справа. Момент времени, которому соответствуют эти значения, показывается на кривой при помощи стрелки. В нижней части формы в виде таблицы даются значения переменных, характеризующих состояние дыхания и крови.

Общие регистрационные данные об оперируемых больных вводятся в ЭВМ заблаговременно с помощью дисплеев, находящихся в приемном отделении. Регистрация больного выполняется путем высвечивания и заполнения специальной регистрационной формы, в которой перечислены все вопросы и оставлены места для ответов. Эти ответы набирает на клавиатуре дисплея оператор (сестра), затем они высвечиваются на экране и после проверки вводятся в память ЭВМ.

Ввод данных о текущем состоянии больного начинается немедленно по прибытии больного в операционную. Анестезиолог (или сестра) свободным текстом набирает на клавиатуре характеристику состояния больного, а также отвечает на вопросы, используя форму карты для описания общего состояния больного. Затем с помощью специальных форм карт вводятся сведения о времени, виде и дозах анестезии, вводимых медикаментах. Непрерывно по мере получения информации фиксируются с помощью дисплея частота пульса, артериальное давление крови, температура, состояние дыхания, результаты анализа газов и т. д. При необходимости вводятся замечания о примененных медикаментах, времени взятия лабораторных анализов, результатах наблюдения, указания врача по дальнейшему лечению и т. д.

Кровь, забираемую для проведения анализа газов и других химических анализов, посылают в лаборатории

в разные интервалы. Результаты анализов вводят в ЭВМ через лабораторный терминал по мере их получения. Эти результаты могут сразу же наблюдаться на экранах дисплеев в операционной и других отделениях больницы.

С момента ввода первых сведений о больном и в течение всего времени операции анестезиолог имеет доступ к любым фактическим данным о состоянии больного. Например, из общего набора форм в процессе операции на сердце используются в основном три формы: общего состояния, описания анестезии, характеристик работы аппарата искусственного кровообращения. Эти формы содержат большую часть информации, которой пользуется анестезиолог. Кроме того, для получения более полных сведений о состоянии больного используются две формы, являющиеся детализацией формы общего состояния: состояния дыхания и форма, характеризующая циркуляцию крови. Все эти формы в процессе операции выдаются также на дисплей в палате интенсивного ухода, в которую будет доставлен больной после операции, что позволяет персоналу этой палаты наблюдать за ходом операции и своевременно подготовиться к приему больного.

В палате интенсивного ухода за больным, прибывшим из операционной, ведется непрерывное наблюдение, и все данные вводятся в ЭВМ по мере их получения. Частота ввода определяется состоянием больного и инструкциями врача.

В течение всего периода пребывания в этой палате измеряется частота пульса, артериальное давление, температура, определяются характеристики дыхания. Эти данные вводит медицинская сестра палаты интенсивного ухода, заполняя в соответствии с инструкциями нужные формы на экране дисплея. Периодически вводятся в ЭВМ результаты анализов химической лаборатории, в том числе анализов газов крови. Дежурный врач или медицинская сестра в палате интенсивного ухода следит на экране дисплея за изменениями значений параметров состояния больного и периодически вызывает на печать и на экран сводные рапорты о состоянии больного. Кроме того, ЭВМ постоянно проверяет нахождение этих параметров в допустимых пределах и при выходе их за критические значения автоматически подает сигнал тревоги. После перевода больного из отделения интенсив-

ного ухода все его данные выдаются из ЭВМ на перфоленту и на магнитную ленту по команде от дисплея. В случае возвращения больного его данные могут быть вновь введены в ЭВМ.

Каждые 24 ч ЭВМ выдает рапорты, клинические заключения, сведения о примененных и назначенных лекарствах.

В случае сбоев в системе эти распечатки могут использоваться как база для ведения обычных ручных историй болезни.

5. Применение ЭВМ в диагностике заболеваний

Значительное внимание в настоящее время уделяется проблеме машинной диагностики заболеваний — теории и практике этого дела. Первым шагом в этом направлении является построение методики формального (алгоритмического) описания состояния здоровья пациента — представление его в виде упорядоченного множества элементов (компонентов вектора) и отнесение к определенному классу болезней. Для этого нужно иметь формальное математическое описание классов болезней, что является чрезвычайно сложной задачей ввиду ее многомерности.

Следующим шагом является формализация процесса выбора для заданной болезни соответствующего вида лечения или дополнительного обследования, что требует также формализации описаний видов лечений. При этом выборе должна быть обеспечена максимизация значения функции полезности с учетом стоимости (и опасности для больного) различных видов лечения или обследования. Существуют конкретные примеры программ и методов машинной диагностики, а также разрабатываются общие математические и медицинские аспекты этой проблемы.

Можно сделать следующий вывод о состоянии и перспективах машинной диагностики. В настоящее время ЭВМ уже практически применяются при диагностике ряда заболеваний и являются средством, помогающим врачу быстрее и точнее установить диагноз и тем самым сократить или исключить излишние и сложные исследования и даже операции. Одновременно с решением частных задач во многих исследовательских центрах широко ведутся исследования по общей теории и методологии машинной диагностики.

Для более конкретного представления направлений работ и достигнутых успехов в области машинной диагностики заболеваний приведем несколько конкретных примеров, взятых из докладов на международной конференции, посвященной методам принятия решений в медицине и роли кибернетики. Эта конференция проходила в мае 1976 г. в Дижоне (Франция) и была посвящена в основном проблеме машинной диагностики заболеваний.

Важнейшую роль в машинной диагностике играет достоверность исходных данных, получаемых либо путем опроса больного и врачебными осмотрами, либо с помощью лабораторных анализов и функциональных измерений. Данные врачебных осмотров и опросов больных занимают в настоящее время значительное место в диагностике многих болезней. По своей природе эти данные являются субъективными и характеризуются значительными расхождениями в определении и интерпретации.

Еще Н. Винер, основатель кибернетики, отмечал, что расхождения в медицинских наблюдениях являются характерной чертой клинической медицины, которую он отнес к «полуточным» наукам.

Интересные сведения по вопросу объективности медицинских данных были приведены в докладе проф. F. T. Dombal из университета в Лидсе (Англия). Он рассматривал следующие три вопроса: масштаб расхождений в медицинских наблюдениях, меры, которые могут быть приняты для уменьшения этих расхождений, и эффект, который оказывают расхождения в наблюдениях на внимание и действия врачей.

В течение нескольких лет начиная с 1973 г. проводились опыты по анализу расхождений в наблюдениях, касающихся «острого живота» (болей в области живота). Результаты опроса больных одним врачом записывались независимо тремя медицинскими сестрами на специальные бланки. Анализ записей показал наличие больших расхождений: 1/5 всех вопросов записи расходились в том смысле, что был ли задан этот вопрос вообще, а в 1/6 вопросов — был ли ответ пациента положительным или отрицательным.

Уровень подобных ошибок уменьшается при использовании четкой структуры и формализации системы опроса.

Второй пример относится к врачебным осмотрам на рак молочной железы. Были проанализированы результаты двойных осмотров 11 168 женщин, причем первый осмотр делал один врач, а второй осмотр — другой. Расхождения по вопросу о наличии увеличенных лимфатических узлов и размерах первичной опухоли составляли 30—40%. Выводом из этого опыта явилось замечание о необходимости более четкой классификации и разделения на стадии процесса появления и развития рака молочной железы и применении многовариантного анализа наблюдаемых клинических признаков. Отмечаются также большие расхождения в оценке состояния пациентов после лечения, особенно после хирургических операций, что связано в основном с различными подходами к оценке этого состояния в различных медицинских учреждениях. Например, при варикозном расширении вен применяется ряд различных методов лечения (включая хирургические операции) и в связи с большими расхождениями в оценке состояния пациентов после лечения отсутствуют и определенные объективные оценки различных методов лечения. Считается, что все они дают приблизительно одинаковый результат. В действительности «полезный сигнал» (различия в результатах применения разных методов) тонет в «шуме» (в расхождениях в оценке состояний больных). Общий вывод сводится к тому, что перед применением точных математических методов в клинической медицине необходимо проделать большую работу по уточнению определений клинических параметров и признаков, используемых для диагностики, а также по определению результатов лечения.

Первыми шагами в применении машинных методов диагностики в любой области медицины должны быть тщательный анализ достоверности первичных данных, минимизация расхождений и обеспечение воспроизводимости описаний и определений. Должно быть отменено использование для оценки результатов лечения таких терминов, как «успешно», «улучшено», «удовлетворительно».

Различные сопоставления и сравнения серий диагностических или терапевтических процедур должны производиться на основе заранее определенной системы достаточно точных и воспроизводимых параметров. Таким образом, нужно стремиться не только к тому, чтобы иметь больше информации о состоянии пациента, но и к тому,

чтобы иметь больше уверенности в той информации, которая имеется в распоряжении врача.

В другом докладе F. T. Dombal приведены конкретные результаты машинной диагностики «острого живота» (боли в верхней части брюшной полости) на основе пятилетнего опыта, в 10 000 случаев и показывается практическая целесообразность и эффективность машинных методов.

На первой стадии исследований (1971—1973) изучались больные, страдавшие острыми болями в животе меньше недели (серия включала 552 больных). Диагноз этим больным ставили обычным образом (дежурным врачом, хирургом и заведующим отделением) и параллельно с помощью ЭВМ, в которую вводились данные, составленные по специальной форме дежурным регистратором. Диагностический алгоритм, заложенный в ЭВМ, был основан на использовании теоремы Байеса и условных вероятностей, рассчитанных предварительно путем обработки 700 случаев. Диагноз ЭВМ был готов через 5 мин и выдавался на экран дисплея, но он не показывался врачам до установления ими своего диагноза.

Окончательный диагноз, как правило, определялся после хирургического вмешательства.

Сравнение результатов врачебного и машинного диагнозов для указанной серии больных (552) показало, что точность диагностики различных врачей составляла 42—82%, а точность машинной диагностики — 90%.

Несмотря на более высокую точность диагностики, разработанная на первом этапе система в Лидсе не была признана пригодной для практического использования, так как базировалась на устарелой ЭВМ и могла использоваться только лицами, которые ее создали, а не рядовыми врачами. Кроме того, в случае «острого живота» наиболее важным является принятие правильного решения по необходимым мерам помощи (чем установление точного диагноза), а это в разработанной системе предусмотрено не было. В связи с этим в 1974—1975 гг. была разработана и внедрена новая диагностическая система, которая была рассчитана также на ряд других заболеваний (диспепсия, заболевания нижней части желудочно-кишечного тракта, гинекологические и урологические заболевания). При этом точность диагностики характеризовалась следующими показателями. По диспепсии для 400 случаев точность диагноза врачами состав-

ляла 55%, а ЭВМ—81%, по заболеваниям нижней части желудочно-кишечного тракта для 300 случаев точность диагноза была соответственно 64% и 87%, по гинекологическим и урологическим заболеваниям для 393 случаев — соответственно 68% и 81%.

Был сделан также вывод, что обработка на ЭВМ подробной структурированной истории болезней для больных с диспепсией имеет значение для скрининга ранней стадии рака желудка (gastric intestinal cancer).

Новая система была рассчитана на использование ее рядовыми врачами, для обучения которых достаточно было от 6 до 8 нед, причем обучение заключалось в основном в усвоении терминологии.

Важными для применения любой машинной диагностической системы являются точность и удобство терминологии.

Опыт применения описываемой диагностической машинной системы показал, что четкость и дисциплина в сборе и фиксации структурированной информации, которая требуется в этом случае от врачей, и наличие «обратной связи» от системы к врачам привели к значительному повышению уровня правильно принимаемых врачами решений, что особенно важно в случае болезней «острого живота». Так, уровень предоперационных перфораций аппендицита снизился с 36 до 6%, а негативных лапаротомий — с 20 до 7%.

Интересно, что при временном прекращении сопоставления врачебных диагностических и терапевтических решений с машинными данными (при отсутствии обратной связи) уровень врачебных ошибок поднимался до прежнего значения.

Эта система продемонстрировала возможности применения многовариантного анализа для прогнозирования индивидуального течения болезни, в частности для прогнозирования состояния больных раком молочной железы после адrenaлэктомии.

Опыт использования описанной системы в других местах, в частности в Копенгагене, показал, что при изменении географического района несколько снижаются показатели точности работы системы. Это объясняется изменением состава пациентов, некоторым изменением статистических характеристик распределения признаков и заболеваний. На основе большого практического опыта авторы системы делают заключение, что подобные ма-

шинные диагностические системы весьма полезны в клинической практике.

В докладе А. Alperovitch, М. Le Minor, J. Lellouch (Институт INSERM, Франция) на упомянутой выше конференции были приведены три подхода к машинной диагностике на примере трех типов заболеваний: опухолей легких, заболеваний печени и заболеваний щитовидной железы. Для получения практически полезной диагностической системы необходимы глубокое изучение и формализация медицинского существа задач.

Диагностика опухолей легких производилась на основе анализа рентгеновских изображений грудной клетки. Пациенты, у которых есть затемнения на рентгеновских изображениях грудной клетки, могут иметь заболевания 30 видов, из которых наиболее частым (30% случаев) и опасным является рак легких. При этом недопустима как ошибка, так и неопределенность в диагнозе. Однако клинические признаки и свойства рентгеновских изображений не дают полной возможности различать возможные диагнозы. Для этого всегда требуются дополнительные обследования различной сложности, стоимости и тяжести (для больного). Наиболее надежен во многих случаях метод торакотомии, который сочетает диагностические и терапевтические возможности, но является наиболее тяжелым. Возникает задача определения оптимальной последовательности дополнительных тестов, приводящих наиболее быстро к окончательному диагнозу и исключающих по возможности дорогие, неприятные или опасные для больного процедуры. При этом необходимо учитывать специфику каждого конкретного больного.

Практически реализованный алгоритм, основанный на байесовском подходе выбора кратчайшего пути в графе, обеспечил заметное сокращение числа тестов.

Данные об общем числе дополнительных тестов, назначенных врачами и ЭВМ, приведены в табл. 13.

Для второй группы заболеваний — заболеваний печени и билиарной системы — рассматривались почти все возможные заболевания. В связи с большим числом этих заболеваний и большим числом возможных обследований важна задача оптимизации последовательности тестов.

В отличие от предыдущего случая здесь дополнительные тесты иногда оказываются ненужными, так как окончательный диагноз может быть установлен сразу на ос-

Сравнение числа тестов

Тест	Назначено	
	врачами	ЭВМ
Анализ патогенов	20	11
» бацилл Коха	61	20
Серологические реакции аспергиллеза	3	6
Серологические реакции при гидатидных заболеваниях	1	7
Ангиография	6	4
Сцинтиграфия ^{131I} щитовидной железы	4	5
Биопсия периферийного лимфатического узла	4	4
Рентгеноскопия пищевода	16	5
Бронхоскопия	62	44
Транспариетальная пункция	7	4
Торакотомия	20	23

нове клинических признаков и результатов систематических биологических тестов. Кроме того, здесь допустима и некоторая неопределенность в различии возможных видов заболеваний. В связи с этим в данном случае сразу вычислялись вероятности предполагаемых диагнозов на основе теоремы Байеса и решался вопрос о том, проводить ли дополнительные обследования.

Особенностью данной группы заболеваний является также то, что иногда невозможно вообще установить окончательный диагноз однозначно, так как отсутствуют соответствующие тесты. В связи с тем что диагностическая неточность в определении отдельных видов заболеваний не фатальна, диагностический алгоритм для данной группы заболеваний был рассчитан в основном на оптимизацию выбираемых тестов с учетом стоимости обследований и «цены» диагностических ошибок.

Для третьей группы заболеваний — заболеваний щитовидной железы — разработанный алгоритм обеспечивал оптимизацию скрининга. Задача состоит в вычислении вероятностей гиперфункции щитовидной железы на основе клинических признаков и биологических тестов. Обычно обследование больного с заболеванием щитовидной железы выполняется в два этапа. Сначала проводят анализы гормонов щитовидной железы и с введением радиоактивного йода, а затем — дополнительные тесты

(если они необходимы). С помощью обычных тестов хорошо различаются два основных диагноза: повышенная функция щитовидной железы и недостаточная функция щитовидной железы. Неопределенность возникает при слабо выраженном повышении функции этой железы.

Совершенствование обычного диагностического процесса путем применения машинного алгоритма достигается уменьшением избыточности систематических тестов и сокращением назначений дополнительных тестов. При построении алгоритма учитывалось, что 80% больных в данном госпитале с предполагаемым диагнозом увеличения функции щитовидной железы оказывались здоровыми. В связи с этим первой задачей алгоритма был отсеивание больных с такими ошибками диагноза при сохранении возможно более низкого уровня ошибок для случаев с недостаточной функцией щитовидной железы.

Из систематических тестов проверялся только индекс свободного тироксина. Вычислялись две вероятности: на основе клинических наблюдений и на основе индекса свободного тироксина. Программа диагностики, построенная на основе байесовского подхода с учетом взаимозависимости клинических признаков, обеспечила существенное повышение точности диагностики больных с повышенной и с пониженной функцией щитовидной железы и сокращение числа дополнительных тестов.

Рассмотренные примеры показывают практическую целесообразность и большие возможности применения методов машинной диагностики. В СССР во многих институтах и больницах успешно применяются методы машинной диагностики (в Научно-исследовательском институте хирургии имени Вишневского, Научно-исследовательском институте сердечно-сосудистой хирургии имени Н. А. Бакулева, в Институте экспериментальной и клинической онкологии и др.).

Применение ЭВМ при массовых осмотрах населения. Важной областью применения ЭВМ в медицине является автоматизация массовых обследований населения.

Автоматизированные системы для массовых медицинских обследований (АСММО) населения — так называемого скрининга применяются как самостоятельно, так и в виде подсистем, входящих в состав БАИС.

Примером АСММО является система, действующая практически в течение ряда лет (с 1969 г.) в г. Гавле (Швеция). В этом городе проживает 75 000 человек. Об-

следованиям подвергают лиц, родившихся в 1944 г. и раньше. Всего 24 000 человек.

При создании этой системы ставились следующие цели: обеспечить контроль здоровья жителей; разработать и проверить методику и технику обследований; оценить возможность системы по выявлению лиц, нуждающихся в медицинском обслуживании и в первую очередь оценить эффективность многофазовых лабораторных тестов.

Методика обследования предполагает два этапа: первый визит в процессе скрининга; второй визит в необходимых случаях для осмотра врачом.

Во время второго визита подозреваемых на рак легких проводятся также рентгеновская маммография, цветная термография и биопсия тонкой иглой.

Скрининг включает:

1. Заполнение вопросников на специальных, автоматически читаемых бланках (общий вопросник, вопросник для выявления коронарных сердечных заболеваний, вопросник для выявления рака легких).

2. Анализ мочи на глюкозу, белки и др.

3. Антропометрию (рост, масса и др.).

4. Рентгеноскопию легких и сердца.

5. Проведение автоматизированных лабораторных тестов крови с помощью трех видов аутоанализаторов: общего анализа крови, гемоглобина, определения гематокрита, подсчета числа эритроцитов и лейкоцитов, биохимического анализа крови; белков сыворотки, анализа содержания альбумина, билирубина, мочевины крови, кальция, глюкозы и ряда других веществ крови; анализа содержания железа сыворотки, кальция сыворотки.

Данные от анализаторов передаются автоматически в устройство сбора данных ИБМ 1080. После преобразования данные поступают на обработку в специализированную ЭВМ ИБМ 1130.

6. Электрокардиография (ЭКГ). Используются четыре вида отведений: отведения Франка с обработкой данных на ЭВМ; 12 скалярных отведений с обработкой на ЭВМ; 12 скалярных отведений с записью на магнитную ленту и последующей обработкой на ЭВМ; данные по одному отведению.

Важно, что для одного человека снимаются ЭКГ по всем четырем видам отведений почти в одно и то же время. Эти данные сравниваются между собой и данными, характеризующими сердечно-сосудистую систему.

7. Измерение давления крови по методу Рива-Роччи.

8. Термография для скрининга с целью выявления рака легких. Термограммы посылают для анализа специалистам, а обследуемого направляют (если нужно) на цветную термографию и биопсию.

Во время второго визита врач внимательно осматривает обследуемого, предварительно ознакомившись с резюме общего вопросника, данными анализа крови и мочи. О других результатах скрининга врач ничего не знает. Свои наблюдения и диагноз врач записывает на специальный бланк, приспособленный для оптического чтения и ввода в ЭВМ, причем диагноз записывается в виде кода международной классификации.

Врач указывает, нуждается ли обследуемый в дальнейшем обслуживании. После этого обследуемого осматривает другой врач, имеющий полные результаты скрининга и записи первого врача. Имеется программа для сравнения результатов скрининга в двух вариантах: 1) скрининг с участием врача, но без специальных лабораторных тестов, 2) скрининг с участием врача, располагающего данными лабораторных анализов.

Первоначально в ЭВМ вводят идентификационные коды всех жителей округа, что используется для автоматического контроля вводимых данных скрининга. ЭВМ выдает письма (приглашения), списки приглашаемых и очередей с датами, данные о результатах скрининга и т. д.

Входные данные для ЭВМ вводятся тремя способами в зависимости от характера информации:

а) вопросники с ответами — с помощью автоматически читаемых бланков,

б) результаты лабораторных анализов — с помощью перфокарт или клавиатур дисплеев с записью на магнитные диски,

в) ЭКГ и результаты их анализа — с помощью магнитных лент.

Еженедельно ЭВМ осуществляет: ввод новых данных скрининга; выбор лиц, которые должны быть приглашены на второй визит; выдачу списков этих лиц; печать писем-приглашений; хранение и обновление данных о больных; печать справок врачам для второго визита; выработку текущих итоговых статистических сводок.

Данная система ориентирована на выявление рака легких. Подобные системы существуют по гинекологии (г. Упсала, Швеция), по туберкулезу и т. д.

В нашей стране ведутся работы по созданию и применению автоматизированных комплексов в медицинской технике для раннего выявления заболеваний. В директивах ЦК КПСС к плану развития народного хозяйства на 1976—1980 гг. отмечена необходимость создания и освоения в производстве автоматизированных унифицированных электронных приборов и аппаратов для массовых медицинских обследований населения.

Признано перспективным для практического здравоохранения и промышленности медицинской техники направление работ, связанное с созданием автоматизированных унифицированных электронных приборов и комплексов (включающих блоки обработки информации), которые можно эксплуатировать в условиях поликлиник и диспансеров, без участия специального обслуживающего технического персонала.

Следует разрабатывать автоматизированные медицинские приборы, обеспечивающие экономию врачебного труда при массовых медицинских обследованиях определенных возрастных и профессиональных контингентов населения путем автоматизированного разделения обследованных по группам риска различных заранее определенных заболеваний. В первую очередь необходимо направить усилия на автоматизацию предварительных и периодических обследований рабочих.

Итак, относительно возможностей применения ЭВМ в медицине можно сделать следующие выводы:

1. В настоящее время во всех передовых странах широкое практическое применение в медицине и здравоохранении получили ЭВМ, а также автоматизированные системы сбора и обработки данных.

2. Основным направлением применения ЭВМ в медицине является обработка лабораторных данных, сведений о больных, находящихся в критических условиях, машинная диагностика, а также обработка данных, касающихся обслуживания пациентов и внутрибольничного управления.

3. Основными проблемами, связанными с применением ЭВМ в медицине и здравоохранении, являются построение банка медицинских данных, автоматизации получения первичных сведений, обеспечение надежности функционирования систем, вопросы сохранности и конфиденциальности информации.

АВТОМАТИЗАЦИЯ ПРОЦЕССОВ УПРАВЛЕНИЯ В ЗДРАВООХРАНЕНИИ

1. Назначение и структура АСУ «Здравоохранение»

Здравоохранение является одной из важных отраслей народного хозяйства, поэтому автоматизация процессов управления в этой отрасли имеет большое государственное значение.

В Отчетном докладе ЦК КПСС XXV съезду КПСС сказано: «...очень важное звено в улучшении руководства экономикой — это **совершенствование организационной структуры и методов управления**»¹.

В «Основных направлениях развития народного хозяйства СССР на 1976—1980 гг.», принятых XXV съездом КПСС, отмечается необходимость обеспечить дальнейшее развитие и повышение эффективности автоматизированных систем управления и вычислительных центров, последовательно объединяя их в единую общегосударственную систему сбора и обработки информации для учета, планирования и управления.

Здравоохранение является не только важной, но и сложной отраслью народного хозяйства, в которой расходуются большие ресурсы. Хотя здравоохранение не выпускает материальной продукции, от качества и уровня его функционирования зависят производительность труда и уровень работы всех других отраслей народного хозяйства. В системе здравоохранения действует большое количество медицинских и вспомогательных учреждений, трудятся большое число врачей, средних медицинских работников, обслуживающий персонал и специалисты инженерных и других специальностей. В связи с важностью и сложностью этой системы весьма актуальной становится проблема автоматизации процессов управления

¹ Брежнев Л. И. Отчет Центрального Комитета КПСС и очередные задачи партии в области внутренней и внешней политики. М., Политиздат, 1976, с. 60.

в здравоохранении. Над этой проблемой работает ряд научно-исследовательских институтов союзного и республиканского подчинения. Головным институтом по данной проблеме является Всесоюзный научно-исследовательский институт социальной гигиены и организации здравоохранения имени Н. А. Семашко Министерства здравоохранения СССР.

Основу оптимального планирования и управления здравоохранением, как и другими отраслями народного хозяйства, должна составлять комплексная модель функционирования и развития здравоохранения, в которой должны быть отражены не только внутренние взаимосвязи в системе здравоохранения между различными ее функциональными частями, но и внешние связи здравоохранения с народным хозяйством и внешней средой.

Вопросы системного моделирования здравоохранения и построение комплексной универсальной модели здравоохранения наиболее успешно развиваются в нашей стране группой специалистов под руководством доктора мед. наук Д. Д. Венедиктова. Эта модель может быть использована как общая методологическая основа для построения локальных моделей здравоохранения в союзных и автономных республиках, в областях, городах, а также для общей ориентировки при разработке моделей здравоохранения в разных странах, несмотря на различия в политическом и экономическом строе. Заметим, что указанная работа выполняется в рамках биомедицинского проекта в Международном институте прикладного системного анализа в Вене.

Основные черты разрабатываемой модели здравоохранения сводятся к следующему.

Вводится формализация и типизация функциональной структуры здравоохранения путем определения трех основных функциональных блоков модели (профилактика, лечение, медицинская наука), двух обеспечивающих блоков (кадры, ресурсы) и одного блока управления здравоохранением. Для каждого блока определен набор подфункций, или задач. Например, блок профилактики предназначен для решения следующих проблем: иммунизация населения, массовые медицинские осмотры, медицинское просвещение и др. Блок управления здравоохранением включает сбор и анализ информации о состоянии управляемой совокупности объектов, выработку решений, доведение решений до исполнителей и контроль

за выполнением решений, анализ результатов и корректировка решений и т. д.

В модели в формализованном виде представляются взаимосвязи между системой здравоохранения и экономическими, социальными факторами и характеристиками условий внешней среды.

Взаимосвязи между блоками модели численно выражаются с помощью четко определенной системы показателей, общее число которых составляет 340. Из них 19 показателей характеризуют уровень здоровья, 40 — профилактическую деятельность, 19 — лечебную деятельность, 24 показателя — медицинскую науку, 21 показатель — управление здравоохранением, 16 показателей — ресурсы здравоохранения. Для характеристики социально-экономических факторов намечены 120 показателей и для характеристики условий внешней среды (природных условий) — 34 показателя. Естественно, что состав и количество перечисленных показателей будут изменяться и корректироваться в процессе разработки и отладки модели, однако приведенные числа дают некоторое представление об информационной базе модели и степени детализации описания общей системы здравоохранения с помощью указанной модели.

В рассматриваемой модели здравоохранения предусматривается формализация и стандартизация организационной структуры здравоохранения путем определения основных типов медицинских учреждений и видов административного подчинения этих учреждений руководящим органам. Общее число различных типов медицинских учреждений составляет 113 (больничные стационары, поликлиники, санитарно-эпидемиологические станции, родильные дома, диспансеры и т. д.). В модели представляются два вида подчинения медицинских учреждений: местным административным органам (территориальное подчинение) и центральным административным органам (отраслевое подчинение).

Такие виды деятельности системы здравоохранения, как развитие медицинской науки и внедрение ее достижений в практику, а также планирование подготовки кадров, являются прерогативой центрального органа.

Важной составной частью модели являются матрицы информационных взаимосвязей между:

а) различными функциями и подфункциями (задачами). В этой матрице показываются виды входной и вы-

ходной информации для каждой функции и задачи, фиксируются источники и потребители информации,

б) различными типами медицинских учреждений (включая органы управления), различными функциями и задачами.

В этих матрицах фиксируются два типа информационных связей: сильные и обычные. К сильным связям относятся регламентированные потоки информации, обусловленные иерархическим подчинением медицинских учреждений, к обычным связям — остальные виды обмена информацией.

В модели реализован общий алгоритм планирования здравоохранения методом последовательных приближений в диалоговом режиме взаимодействия человека и ЭВМ. Разработка и реализация модели осуществляется по этапам. Первый этап предусматривает выработку решений по следующим аспектам: а) планирование подготовки и оптимального распределения кадров медицинских специалистов, б) оптимизация процессов массовых медицинских осмотров населения (при этом ставится задача обеспечить оптимальное соотношение между стоимостью и эффективностью указанных мероприятий), в) определение нормативов и стандартов в системе здравоохранения, обеспечивающих оптимальное отношение между стоимостью и эффективностью как всей системы, так и ее отдельных секторов или отдельных медицинских учреждений.

Описываемая модель отражает не только статистические, но и динамические свойства системы здравоохранения. Определено, что наиболее подходящей единицей приращения модельного времени является 1 мес. С точки зрения программной реализации комплексная модель строится в виде набора взаимодействующих программных модулей, представляющих отдельные блоки модели («сеть медицинских учреждений», «население», «ресурсы», «кадры», «управление здравоохранением» и др.). Например, предусматриваются три разных модуля для представления заболеваемости: хронические и длительные болезни, острые и инфекционные болезни; отравления и несчастные случаи.

Все население делится на три группы: здоровые, невыявленные больные и зарегистрированные больные. Каждая из групп подразделяется на подгруппы в зависимости от пола и возраста.

В результате профилактических мероприятий, массовых медицинских осмотров населения, лечения в стационарах и поликлиниках, иммунизации и других мер происходят переходы населения из одних групп и подгрупп в другие группы и подгруппы.

На основе описанной модели можно изучать различные варианты отношений между объемом профилактической и лечебной деятельности, влияние уровня технической оснащенности на стоимость и эффективность массовых осмотров, анализировать нормы загрузки медицинского персонала и т. д. Для функционирования модели необходимо задаться и ввести в нее сведения, характеризующие предполагаемые тенденции роста и распределения населения, данные о перспективах развития диагностических и терапевтических средств и их эффективности.

В результате расчетов согласно описанной модели из ЭВМ будет выдаваться набор альтернативных решений по планированию здравоохранения. Окончательный выбор и принятие решений должны осуществляться лицами, ответственными за систему здравоохранения или ее соответствующие участки.

Подобные комплексные модели могут использоваться также для исследования теоретических проблем: количественной оценки уровня здоровья, определения видов и характеристик взаимосвязей системы здравоохранения и других отраслей народного хозяйства, определения эффективных критериев оценки деятельности системы здравоохранения.

Полное построение комплексной модели здравоохранения и реализация на ее основе задач оптимального планирования и управления здравоохранением представляют собой весьма сложную и трудоемкую проблему, которая потребует нескольких лет работы. Реализация этой проблемы может быть отнесена ко второй очереди АСУ «Здравоохранение».

Первая очередь АСУ «Здравоохранение» включает автоматизацию процессов сбора и обработки статистической информации по основным направлениям деятельности и решение отдельных задач оптимизации управленческих процессов.

Ниже будут рассмотрены более подробно основные задачи первой очереди АСУ «Здравоохранение».

2. Задача сбора и обработки данных о лечебной деятельности больничных стационаров

В условиях централизованного управления сетью лечебно-профилактических учреждений (в масштабе страны, республики, области или в рамках ведомственных систем здравоохранения — железнодорожный транспорт, аэрофлот и др.) большое значение имеет анализ деятельности больничных стационаров. Это обусловлено важной ролью стационаров в сохранении здоровья обслуживаемого контингента населения и значительными расходами на их строительство, содержание и развитие. Предполагается, что в функции центрального органа управления входят контроль и оценка деятельности стационаров, а также планирование их специализации и развития, распределение денежных, материальных и других ресурсов.

Как известно, лечебная деятельность стационаров определяется прежде всего данными о численности, составе больных, сроках и результатах их лечения.

Анализ показывает, что необходимый для центрального органа управления объем информации о лечебной деятельности стационаров получается с достаточной степенью точности и оперативности путем сбора и централизованной обработки на ЭВМ карт о каждом больном, выбывшем из стационара. Для центрального органа управления нет необходимости получать более подробные данные о всех промежуточных этапах и стадиях пребывания каждого больного в стационарах. Однако для центрального органа оказываются недостаточными годовые, квартальные и даже месячные отчеты стационаров, содержащие суммарные данные о больных, выписавшихся из стационаров.

Оптимальным является вариант ведения централизованного массива карт о каждом выписанном из стационара. Такой массив позволяет оперативно получать в центре не только заранее установленные сводные отчеты, но и любые абсолютные или сопоставляемые данные, касающиеся разных стационаров, заболеваний, методов лечения, категорий больных, периодов деятельности и т. д.

Принципы построения подобных систем легче всего уяснить из рассмотрения конкретного примера одной из реализованных задач.

В описываемой автоматизированной системе в качестве первичного документа принята карта выбывшего из стационара, со-

держающая следующие 14 признаков: регистрационный номер, специальность (профиль коек), возраст, житель города или села, порядок поступления (экстренно или нет), время доставки, исход заболевания, дата выписки или смерти, число проведенных дней, основной клинический диагноз, патологоанатомический диагноз, название операции, наличие послеоперационного осложнения, метод лечения злокачественных новообразований.

Все эти сведения передаются в ВЦ с помощью аппаратуры передачи данных, от близлежащих стационаров данные передаются в виде документов и перфорируются в самом ВЦ.

В результате обработки выдаются сводные и сопоставительные таблицы, характеризующие использование коечного фонда и движение больных в стационарах, большую заболеваемость и летальность, длительность пребывания в стационарах больных туберкулезом органов дыхания, использование различных методов лечения злокачественных новообразований, хирургическую деятельность стационаров, состав, сроки и исходы лечения больных, госпитализированных по экстренным хирургическим показаниям.

Эти сведения выдаются по каждому стационару и по всем стационарам каждые полгода. Кроме того, ВЦ оперативно выдает по запросам центрального органа или главных врачей отдельных стационаров различные справочные данные, необходимые для текущего управления стационарами.

Программное обеспечение описываемой централизованной системы обработки данных о лечебной деятельности стационаров состоит из трех унифицированных программ:

- 1) контроль входной информации и формирование машинного массива документов,
- 2) формирование и печать таблиц абсолютных и относительных показателей,
- 3) поиск отдельных документов или их групп в машинном массиве по заданным наборам признаков.

Важная роль в обеспечении надежного функционирования данной системы принадлежит автоматизации контроля входной информации. Это связано прежде всего с тем, что в систему непрерывно поступает большое число первичных документов, причем большинство из них поступает по каналам связи от удаленных абонентов.

Программа контроля входной информации и формирования машинного массива документов состоит из пяти

основных программ, выполняющих следующие функции: формальный контроль документов, логический контроль документов, печать ошибочных документов с указанием выявленных ошибок, формирование машинного массива из проверенных документов, печать всех включенных в машинный массив документов в виде «квитанций», высылаемых для подтверждения в соответствующие стационары (абонентам).

Формальный контроль документа заключается в проверке соответствия кодов стационара и лица, пославшего сообщение, проверке количества реквизитов в документе, их разрядности, нахождения значений реквизитов в заданных пределах, например, дата 30 февраля невозможна, шифр диагноза не может быть больше заданного числа и т. д.).

Кроме формального контроля, существует логический контроль, смысл которого заключается в том, например, что при определенных заболеваниях не могут иметь место операции и послеоперационные осложнения. В общем случае логический контроль состоит в проверке соблюдения смысловых зависимостей, подобных упомянутой выше.

Так, наличие числовых значений одних реквизитов может требовать обязательного наличия числовых значений других реквизитов, если один из реквизитов имеет определенное значение, то некоторые другие могут иметь только определенные значения, или, наоборот, не могут принимать определенных значений, или же должны быть обязательно больше (или меньше) некоторых значений и т. п.

Возможно применение и других видов контроля без переделки программы. Программа является унифицированной, т. е. может быть использована для контроля различных документов по различным видам проверок (отношениям между реквизитами). Программа настраивается на контроль конкретных документов с помощью задания, содержащего в табличной форме описание контролируемых соотношений документа.

Ошибки, выявляемые с помощью программы логического контроля, делятся на условные и безусловные. К условным ошибкам относятся случаи появления редких, но, вообще говоря, допустимых отношений между реквизитами. Документы с такими ошибками пропускаются в массив с пометкой «условно» и выдаются на печать для

анализа специалистами. Безусловными ошибками являются явно недопустимые соотношения. Документы с такими ошибками в главный машинный массив не включаются, а возвращаются абонентам для исправления. После исправления такие документы вводятся в массив как новые документы. Если же исправляется условная ошибка, то такие документы вводятся в массив с признаками «повторно».

Работой отдельных подпрограмм программы контроля и формирования массива управляет специальный диспетчерский блок, который осуществляет ввод группы документов — сообщения в ЭВМ из периферийного устройства, контроль служебных признаков в сообщении, формирование из сообщения блоков машинных документов, вызов подпрограмм формального и логического контроля и других подпрограмм.

В программе контроля около 5000 общих команд. Эта программа ведет статистический учет ошибок, распределяя их по видам ошибок, источникам получения, периодам работы и т. д. Кроме того, в системе ведется общий учет ошибок, допускаемых на различных стадиях технологического процесса. Анализ ошибок, проведенный в процессе практической эксплуатации описываемой системы, показал, что удельный вес ошибочных документов составляет 1,4%, причем около 1% ошибок приходится на шифровку и перфорацию первичных документов и 0,4% ошибок — на передачу данных по каналам связи и их ввод в ЭВМ. С помощью описанной системы контроля обнаруживается приблизительно 90% ошибок. Остальные ошибки выявляются на заключительной стадии обработки документов, а иногда даже в процессе использования выданных из ЭВМ результатов.

Машинный массив исходных документов в сводные таблицы абсолютных и относительных показателей как по отдельным стационарам, так и по группе стационаров перерабатывается с помощью унифицированной программы формирования и печати таблиц (УФП). Эта программа пригодна для выдачи разных таблиц при различных массивах исходных документов. Настройка программы на выдачу конкретной таблицы по конкретному машинному массиву исходных документов осуществляется с помощью задания, представляемого в табличной форме. Программа выдает таблицы, снабженные текстовыми заголовками столбиков и колонок (и общим заголовком)

для каждого значения или заданных наборов значений ключевых признаков.

Задание определяет все заголовки, форму таблицы, алгоритм ее составления.

Обычно при обработке массива первичных документов требуется составлять сразу большое число выходных таблиц (отчетных форм). Описываемая унифицированная программа составляет все необходимые таблицы за один просмотр магнитной ленты с исходным массивом. Каждый просматриваемый документ разносится по всем матрицам, соответствующим различным выходным документам. Наличие унифицированной программы формирования таблиц исключает необходимость в программировании расчетов каждой новой таблицы. Для этого нужно составить новое задание, что гораздо проще и легче, чем программирование.

Для контроля за правильностью составления заданий существует специальный блок контроля, входящий в состав УФП. Благодаря применению указанной программы стало возможным участие врачей, экономистов и других лиц, не являющихся программистами, в составлении алгоритмов формирования нужных им таблиц.

Третья составная часть математического обеспечения рассматриваемой системы — это программа поиска документов в главном массиве по заданным наборам признаков. Необходимость в подобных поисках в условиях центрального органа управления возникает тогда, когда нужно более подробно разобраться с некоторыми исключительными случаями (острые осложнения, недопустимо высокая летальность и т. д.), а также при анализе узких вопросов лечебной деятельности отдельных стационаров.

В связи с тем что главный массив формируется по отдельным стационарам в порядке возрастания регистрационных номеров карт выбывших из стационаров, для предварительного отбора документов при поиске в качестве ключевых признаков используются номер стационара и дата выписки. Между датами выписки и регистрационными номерами имеется соответствие, которое фиксируется в специальной таблице. Кроме того, для ускорения поиска документов по диагнозам заболеваний применяются цепные списки. Окончательный отбор документов осуществляется путем прямого перебора предварительно найденных документов по трем перечисленным признакам.

Опыт практической эксплуатации описанной системы показал ее эффективность и выявил целесообразность некоторого расширения сведений о выбывших из стационаров и проведения дополнительной обработки собираемых данных.

Для повышения качества управления стационарами предполагается наряду с усредненными статистическими данными выдавать по отдельным стационарам и отделениям персональные сведения о больных, показатели которых значительно отклоняются от средних значений, а также выдавать на печать вместо полных отчетных таблиц, составляемых сейчас по всем стационарам и нозологическим формам, таблицы отчетных данных только по тем стационарам и нозологическим формам, которые существенно отличаются от нормы и требуют дополнительного анализа и принятия управляющих решений со стороны центрального органа.

3. Комплекс задач подсистемы управления эпидемическим обслуживанием населения

Цель создания подсистемы заключается в повышении эффективности управления путем проведения профилактических и противоэпидемических мероприятий для обеспечения дальнейшего снижения инфекционной заболеваемости, повышения экономической эффективности деятельности эпидемиологической службы, оптимального использования материально-технических ресурсов и улучшения тем самым эпидемиологического обслуживания населения.

Первая очередь подсистемы включает следующие четыре задачи:

1. Оперативное слежение за инфекционной заболеваемостью с целью оценки интенсивности течения эпидемического процесса на контролируемой территории, выделения вспышек и оказания оперативной помощи учреждениям в расследовании причин их возникновения и их ликвидации.

2. Анализ инфекционной заболеваемости с целью определения особенностей течения эпидемического процесса на контролируемой территории, необходимых для рационального планирования и обеспечения профилактических и противоэпидемических мероприятий.

3. Оценка состояния иммунизации обслуживаемого населения с целью контроля полноты и своевременности проведения иммунизации.

4. Контроль неотложных противоэпидемических мероприятий при эпидемическом неблагополучии на контролируемой территории с целью анализа эффективности принятых мер и их обоснованности.

Методики, используемые в процессе решения этих задач, можно также разделить на четыре группы: слежение за уровнем контролируемого параметра, комплексная оценка состояния, распознавание образов (характера эпидемического процесса), прогнозирование развития процесса. Эти методики разработаны Г. В. Романовским.

Общим для всех методик является использование системы контрольных уровней и оценочных нормативов, полученных в результате обработки фактических данных методами математической статистики и теории вероятностей. Такой подход объясняется тем, что элементам физических процессов, как и процессов управления, свойственна изменчивость двух видов: органически присущая данному элементу (физическое состояние заболевшего, качество вакцины, совершенство инструментария и т. п.) и изменчивость, которая может быть доведена до определенного минимума путем совершенствования системы управления (сроки расследования вспышек, норма расхода бактериологических препаратов, графики работы персонала больницы).

Методика количественной оценки интенсивности течения эпидемического процесса и выделения вспышек при оперативном слежении за инфекционной заболеваемостью на уровне центрального органа управления основана на сравнении данных о зарегистрированных заболеваниях за истекший день и с начала месяца с системой однодневных, трехдневных и месячных контрольных уровней, в результате чего различаются и устанавливаются три интенсивности и состояния эпидемического процесса: «возможна вспышка», «вспышка» и «заболеваемость превысила месячный контрольный уровень».

Контрольные уровни в описываемой системе определялись для 29 нозологических форм в результате статистической обработки фактических данных об инфекционной заболеваемости на контролируемых территориях за последние 5 лет методами математической статистики и

теории вероятностей в предположении применимости распределения Пуассона вида:

$$P_k = \frac{M^k e^{-M}}{k!}$$

для описания спорадической составляющей и обычного уровня эпидемического процесса, где P_k — вероятность регистрации случайного числа заболеваний K , M — параметр распределения Пуассона, e — основание натуральных логарифмов.

Приемлемость этих моделей для описания обычной заболеваемости определялась путем обработки статистических данных на основе критерия, известного в математической статистике как критерий χ^2 (хи-квадрат).

В качестве контрольных уровней использовались числа заболеваний за 1 день, 3 дня и 1 мес, вероятность превышения которых при обычной для данной местности и времени года интенсивности течения эпидемического процесса за счет причин чисто случайного характера, не связанных со вспышкой, была бы достаточно малой, т. е. не превышала 5%:

$$p(k \geq k_y) = 1 - \sum_{k=0}^{k_y} \frac{M^k e^{-M}}{k!} \leq 0,05,$$

где k_y — заданный контрольный уровень.

Методика текущей оценки интенсивности течения эпидемического процесса и выделения эпидемических подъемов при оперативном слежении за инфекционной заболеваемостью на уровне учреждения основана на использовании планов последовательного анализа. Для построения такого рода планов использовались количественные оценки среднего и контрольного уровней заболеваемости, характеризующих обычное течение эпидемического процесса на контролируемых территориях по месяцам года. С помощью этих оценок были решены линейные уравнения вида:

$$k^0 = \frac{\ln \frac{\beta}{1-\alpha}}{\ln \frac{k_y}{M}} + \frac{k_y - M}{\ln \frac{k_y}{M}} \cdot T,$$

$$k^1 = \frac{\ln \frac{1-\beta}{\alpha}}{\ln \frac{k_y}{M}} + \frac{k_y - M}{\ln \frac{k_y}{M}} \cdot T,$$

где M — среднее «многолетнее число» заболеваний за рассматриваемый период, k_y — контрольный уровень обычной заболеваемости, α — вероятность ошибки первого рода, заключающейся в том, что благоприятное в эпидемическом отношении течение эпидемического процесса на основании анализа текущих данных будет оценено как неблагоприятное, β — вероятность ошибки второго рода, сводящейся к тому, что неблагоприятное в эпидемическом отношении течение эпидемического процесса на основе анализа текущих данных будет оценено как благоприятное, T — время в месяцах.

В практике медицинских исследований обычно принимают

$$\alpha = \beta = 0,1.$$

Графики этих уравнений делят плоскость плана на три зоны: зону обычного течения эпидемического процесса, зону неустойчивого благополучия и зону подъема эпидемических заболеваний.

Характер течения эпидемического процесса определяется положением точек графика с координатами $(\Sigma k, T)$ текущей заболеваемости на плане, т. е. нахождением его в одной из трех упомянутых выше зон.

Методика комплексной оценки инфекционной заболеваемости и эпидемической обстановки на контролируемой территории основана на использовании для оценки общего уровня инфекционной заболеваемости обобщенного показателя, общий вид которого следующий:

$$W = f(g_i, w_i) \\ i = 1, 2, \dots, n,$$

где W — обобщенный показатель инфекционной заболеваемости, w_i — частные показатели инфекционной заболеваемости i -й болезнью за рассматриваемый период, g_i — коэффициент эпидемической весомости i -го частного показателя, n — число показателей, учитываемых при комплексной оценке инфекционной заболеваемости.

Этот обобщенный показатель W сравнивается с заранее заданным оценочным числом $W_{\text{от}}$.

Для комплексной оценки эпидемической обстановки применяют обобщенный показатель

$$Z = f(z_i); \quad i = 1, 2, \dots, m.$$

где z_i — частные показатели, характеризующие эпидемическую обстановку на контролируемой территории, m — число показателей, учитываемых при комплексной оценке эпидемической обстановки, и соответствующей оценочный норматив $Z_{\text{оц}}$, установленный с учетом результатов статистической обработки соответствующих данных за последние 5 лет.

Методика количественной оценки состояния иммунизации обслуживаемого населения. При использовании количественных методов в управлении прививочной работой приходится иметь дело с системой из нескольких показателей, характеризующих различные стороны этого процесса, что создает определенные трудности. Эти трудности устраняются путем замены системы частных показателей одним обобщенным. Методика предусматривает использование для комплексной оценки состояния иммунизации и выполнения схемы по всем возрастным группам и видам прививок средних взвешенных геометрических показателей вида:

$$V = \sqrt[n]{\prod_{k=1}^n v_k^{a_k}},$$

где V — значение обобщенного показателя состояния иммунизации, v_k — частные показатели состояния иммунизации, характеризующие полноту и своевременность иммунизации населения различных возрастных групп, n — число возрастных групп, a_k — предельное число прививок на одного человека k -й возрастной группы (коэффициенты весомости).

Коэффициенты весомости определяются принятой схемой иммунизации, а контрольные нормативы устанавливаются в результате статистической обработки фактических данных за последние 2 года.

Методика расследования вспышек дизентерии с помощью ЭВМ. Постановка эпидемиологического диагноза наиболее затруднительна при анализе инфекционной заболеваемости и установлении причин возникновения вспышек.

Разработку методов, облегчающих анализ вспышек и расследование причин их возникновения, целесообразно производить в следующих двух основных направлениях:

1. Совершенствование системы расследования вспышек путем разработки и внедрения в практику санитарно-эпидемиологических станций и медико-санитарных частей типовых схем описания и методов расследования вспышек с использованием унифицированных форм учетно-отчетных документов и статистических методов обработки исходных данных.

2. Разработка методов машинной диагностики типов эпидемических процессов, а также определения наиболее вероятных путей и факторов передачи инфекций.

Из всего многообразия методов распознавания образов для решения этой проблемы представляется целесообразным апробировать методы, основанные на использовании детерминистской логики, логики максимального информационного правдоподобия и логики, основанной на принципе фазового интервала.

Связь признаков с образами может задаваться посредством условных вероятностей или частот встречаемости, определенных в результате обработки данных типовых описаний вспышек с помощью экспертной оценки. По результатам апробации можно сделать выбор метода, дающего наилучшие результаты по решению проблемы машинной диагностики типа эпидемического процесса, и принять решение о целесообразности его практического использования.

Методика построения медико-математических моделей. Характерная особенность анализа функционирования санитарно-эпидемиологической станции состоит в том, что в основе изучения процессов управления должны лежать не только логические посылки медицинского плана, но и определенные количественные зависимости. Результатом анализа должно явиться построение медико-математических моделей, описывающих некоторую реальную взаимосвязь зависимых переменных и определяющих факторов.

В рассматриваемых случаях медико-математические модели представляют собой уравнения множественной регрессии с соответствующими статистическими оценками параметров и дисперсии.

В качестве аппроксимирующих используется линейная функция вида:

$$y = a_0 + a_1x_1 + \dots + a_nx_n$$

или степенная функция вида:

$$y = a_0x_1^{a^1}, \dots, x_n^{a^n},$$

которая путем логарифмического преобразования приводится к линейной.

При этом характерным является подход, сущность которого состоит в том, что модели, описывающие более крупные процессы, не комбинируются из моделей входящих элементов, а строятся на основе первичных данных, но в более укрупненном виде.

Построение медико-математических моделей осуществляется с помощью ЭВМ по программам линейного многофакторного анализа.

4. Оптимизация решения управленческих задач

Рассмотрим пример управленческой задачи, связанной с планированием подготовки и распределением кадров специалистов. Целью задачи является получение оптимального плана комплектования лечебных учреждений врачами определенной квалификации и специализации. В основу модели, описывающей динамику кадрового состава, положена функциональная схема основных потоков, комплекующих врачебный контингент лечебного учреждения. Анализ факторов, влияющих на формирование потоков схемы, показывает определенные возможности управления (воздействия) на эти потоки. В соответствии с принятой практикой планирования в СССР в качестве основного цикла планирования кадров выбран годовой цикл планирования. Долгосрочное планирование (на 5-летие и более) рассматривается как задача последовательного многофазного планирования (с привлечением аппарата марковских процессов) с коррекцией (учетом реального состояния системы) на каждом цикле.

В основе расчета требуемого числа врачей определенных специальностей лежит изучение и прогнозирование характеристик обслуживаемого контингента (демографические данные, структура заболеваемости, рост производственных мощностей обслуживаемых предприятий и т. д.).

По этим данным с учетом нормативов, принятых в отрасли, возможностей финансирования, состояния и планов снабжения по основным фондам, а также с учетом «истории» некоторых потоков, комплекующих кадровый состав, определяются численность и структура (специализация, квалификация и прочее) врачебного персонала, требующегося на определенном цикле планирования.

Автоматизированная подсистема учета, планирования и управления кадрами выдает информацию о реальном состоянии кадрового состава, используемую как при годовом цикле планирования, так и при корректировке долгосрочного прогноза.

Разность между прогнозированной потребностью в кадрах и реальным их наличием представляет собой потребность данного лечебного учреждения в кадрах на планируемый период. Эта потребность удовлетворяется за счет врачей, прошедших переподготовку и усовершенствование внутри системы, и за счет молодых специалистов.

Эти кадры могут быть перераспределены между подведомственными учреждениями соответственно целевой функции, выбор которой должен осуществляться центральным органом управления, исходя из состояния кадровой системы в целом.

На схеме 15 показана модель прогнозирования и планирования кадрового состава. Содержание отдельных блоков этой схемы в соответствии с их номерами следующее.

1 — обслуживаемый контингент предшествующего цикла расчета прогноза, 2 — обслуживаемый контингент текущего цикла прогноза, 3, 4 — обслуживаемый контингент следующего цикла прогноза, 5, 6 — демографические изменения, 7, 8 — изменение контингента в зависимости от планов развития обслуживаемой отрасли, 9 — структура заболеваемости, 10 — прогноз структуры заболеваемости, 11 — учет нормативов, 12 — прогноз кадров, 13 — экспертные оценки и ограничения, 14 — скорректированный экспертами прогноз, 15 — ограничения: контрольные нормативы на год прогноза (по видам помощи, по основным фондам, по финансированию, по переподготовке, по приему молодых специалистов), 16 — прогноз ограниченный на следующий цикл прогноза, 17 — планируемое (требуемое) кадровое состояние организаций, 18 — прогноз кадров следующих циклов, 19, 22 — реальное кадро-

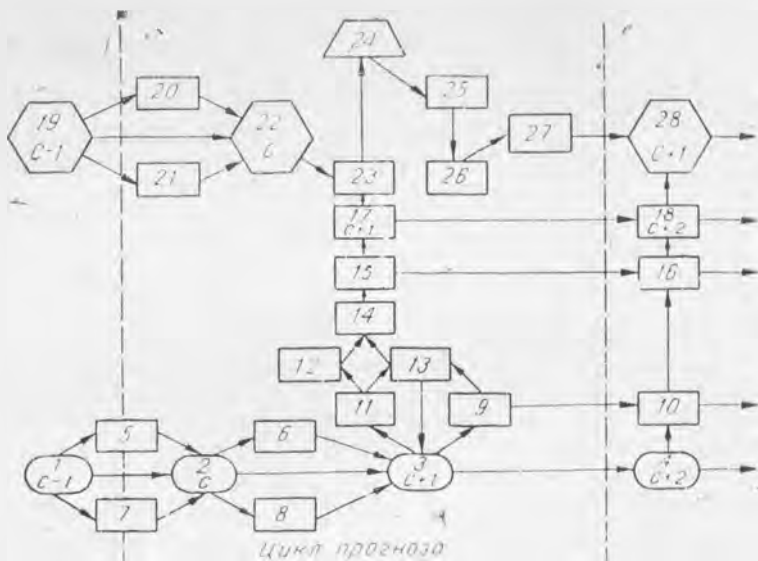


Схема 15. Модель прогнозирования и планирования кадрового состава.

вое состояние в соответствующие годы, 20 — учет переподготовки врачей, 21 — учет динамики состава, 23 — планируемая потребность в кадрах, 24 — решения директивных органов, 25 — состав и численность молодых специалистов, 26 — удовлетворение заявок, 27 — оптимальное распределение, 28 — реальное состояние кадров на следующий цикл планирования. Буквой C на схеме обозначен исходный цикл расчета прогноза. $C-1$ означает предшествующий, с $C+1$ последующий цикл расчета прогноза.

Помимо указанной общей задачи прогнозирования и планирования кадров, с помощью данной подсистемы решается ряд частных задач.

Расчет плана подготовки специалистов через ординатуру. В качестве целевой функции при этом расчете принимается условие обеспечить в каждом учреждении по заданному центральным органом набору видов медицинской помощи (узким специальностям) определенное число специалистов, окончивших ординатуру по соответствующей специальности. Порядок расчета приведен на схеме 16.

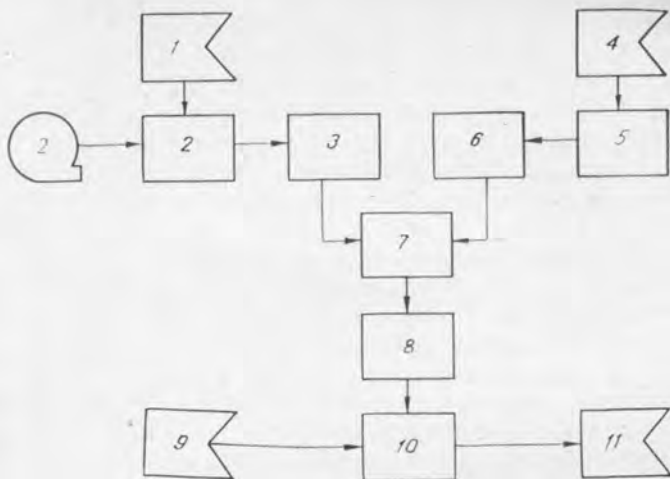


Схема 16. Расчет плана подготовки специалистов в ординатуре.

Задается набор видов помощи каждому учреждению, для которых необходимо подготовить специалистов с ординаторской подготовкой (1). Затем определяют наличное число врачей заданных специальностей в учреждении независимо от того, имеют ли они ординаторскую подготовку. При этом используют массив данных о кадрах учреждений (2). Отдельно определяют число врачей заданных специальностей в учреждениях, окончивших ординатуру (3).

Задается число коек по видам помощи для данного учреждения (4). Для каждого учреждения рассчитывают коэффициенты k_i — число «средних» отделений в данном учреждении по определенному (i -ому) виду помощи (5).

Число типов помощи, умноженное на число «средних» отделений, дает необходимое для учреждения число врачей с ординаторской подготовкой (6). Затем сравнивают по всем видам помощи и по всем учреждениям потребное и наличное число специалистов и определяют реальную потребность в них (7). Суммирование потребностей отдельных учреждений дает план подготовки ординаторов по специальностям (число мест в ординатуре) по всей отрасли (8). Проводят экспертную оценку плана с учетом дополнительных ограничений (экономических, организационных и т. д.) (9). Производят окончательный пересчет плана с учетом введенных ограничений (10).

При этом пересчете используют один из двух критериев: достижение некоторого среднего (общего) уровня обеспеченности кадрами с ординаторской подготовкой всеми отделениями всех учреждений; удовлетворение потребностей учреждений в кадрах в порядке приоритетности, заданной центральным органом управления. Конечным этапом является утверждение плана и его реализация (11).

Следующей частной задачей является расчет плана переподготовки и усовершенствования врачей.

Исходной информацией служит таблица чисел врачей по специальностям, прошедших переподготовку более 5 лет назад по каждому медицинскому учреждению отрасли. Производятся суммирование по отрасли и сортировка по специальностям в порядке убывания числа врачей, которые должны пройти переподготовку.

В качестве целевой функции для отбора на переподготовку врачей из каждого учреждения может использоваться один из трех критериев: отбор врачей, дольше всех не проходивших переподготовки, отбор из каждого подразделения равного числа врачей, отбор врачей по специальностям, определенным экспертной оценкой.

В результате получается таблица численности врачей, отобранных для переподготовки по специальностям и учреждениям. Проводится экспертный анализ этой таблицы и учет дополнительных экономических и организационных ограничений.

Примером третьей частной задачи является оценка текучести кадров в учреждениях и выработка исходных данных для формирования плана приема молодых специалистов. При этом производят общий расчет текучести кадров (поквартальный учет уволенных в каждом учреждении) и отдельно расчет уволившихся по всей отрасли с определением вероятностных коэффициентов увольнений и расчет естественной убыли (пенсия, перевод, прочее). Затем рассчитывают средние коэффициенты текучести по всем учреждениям, после чего определяются суммарные данные по отрасли с разнесением их по группам специальностей и по территориальным районам, а также осуществляют вероятностный расчет распределений по специальностям и районам. Полученные данные позволяют дать оценку численности и квалификации уволившихся из отрасли в целом.

КИБЕРНЕТИКА И МЕДИЦИНСКАЯ НАУКА

В настоящее время в научных учреждениях нашей страны работает около 4 млн. человек, в том числе более 1 млн. научных сотрудников. Ежегодно на развитие науки отпускается до 20 млрд. руб. Ученые ведут громадное число фундаментальных, поисковых, прикладных исследований. Вклад науки в повышение производительности труда, увеличение национального дохода чрезвычайно велик. В этих условиях особенно важное значение приобретает управление развитием науки, с тем чтобы эффективность труда ученых и, следовательно, эффективность вкладываемых в научные исследования средств была максимальной. Именно поэтому в «Основных направлениях развития народного хозяйства СССР на 1976—1980 гг.», принятых XXV съездом КПСС, в качестве важнейшей задачи советской науки в целом и медицинской науки в частности предусмотрено: «Совершенствовать организацию и повышать эффективность труда научных работников. Повысить ответственность научных коллективов и их руководителей за уровень и качество исследований, обоснованность предлагаемых рекомендаций. Усилить заинтересованность коллективов и работников научных учреждений в повышении результативности их деятельности»¹.

Не секрет, что в работе научных учреждений еще имеются серьезные недостатки и трудности. Зачастую научные исследования необоснованно затягиваются на многие годы, длительное время занимает внедрение в практику полученных достижений. Слабо еще поставлены организация комплексных работ и взаимодействие не только между отдельными научными учреждениями, но и внутри институтов — между лабораториями, отделами.

Как известно, темпы роста аппаратного оснащения современной науки должны примерно в 2 1/2—3 раза превышать темпы роста числа работающих в этой обла-

¹ Материалы XXV съезда КПСС. М., Политиздат, 1976, с. 213.

сти. Во многих научных учреждениях этот показатель значительно ниже единицы, что приводит также к фактическому снижению эффективности работы ученых. В этой связи в «Основных направлениях развития народного хозяйства СССР на 1976—1980 гг.» указывается на необходимость «укреплять материальную базу научных, проектных и конструкторских организаций путем улучшения их оснащенности приборами, оборудованием, материалами, средствами вычислительной и организационной техники, а также развития научно-экспериментальных подразделений»¹.

Недостаточная разработанность теоретических и практических вопросов управления развитием науки начинает сказываться на эффективности деятельности научных учреждений. По данным академика В. М. Глушкова (1972), большинство обследованных отраслевых научно-исследовательских учреждений связывает причины своих трудностей именно с недостатками в управлении научно-исследовательской деятельностью. При этом выяснилось, что в 26% случаев отсутствовали ясно сформулированные цели исследований и разработок, в 35% случаев было признано неудовлетворительным составление и контролирование программ работ, в 26% случаев отсутствовало обоснованное распределение ресурсов и был неудовлетворительным экономический анализ. Остальные 13% случаев неудач объяснялись специфическими индивидуальными причинами, не подпадающими под указанные критерии.

Для преодоления подобных недостатков и трудностей в развитии науки необходимо развернуть широкий круг исследований в области управления наукой.

Основы управления наукой стали развиваться лишь в последние десятилетия, и в настоящее время выделилась целая отрасль науки — науковедение. Науковедение обобщает опыт функционирования научных систем с целью повышения эффективности управления научным процессом с помощью средств организационного, экономического, информационного и социального воздействий.

К сожалению, отдельные ученые не придают еще достаточно серьезного значения проблемам управления, считая, что в науке нельзя применять индустриальные методы управления. Здесь, по-видимому, необходимо

¹ Материалы XXV съезда КПСС. М., Политиздат, 1976, с. 213.

прежде всего определить основные задачи науки управления. Среди первоочередных задач следует поставить совершенствование организации научного процесса, улучшение подготовки научных кадров по управлению, изучение резервов повышения эффективности исследований с целью оптимизации сроков выполнения научных работ и экономии расходования средств, изучение и внедрение комплексных системных методов в научных исследованиях.

В условиях научно-технической революции наука становится непосредственной производительной силой. Для повышения ее эффективности основной проблемой становится автоматизация и механизация труда в сфере исследований и управления наукой. С этой целью широким фронтом ведутся работы по использованию ЭВМ в научных исследованиях и при решении задач управления. Внедрение автоматизированных систем в управление наукой включает также механизацию переработки информации с возможностью использования ЭВМ для сбора, накопления, хранения, переработки, поиска и выдачи документальной информации.

Одной из наиболее насущных задач теории управления наукой является поиск путей оптимизации научно-исследовательской деятельности, перспективных методов совершенствования самого процесса управления наукой на всех уровнях.

Эффективность научной работы во многом зависит также от такого фактора, как правильность выбора направления научных исследований. Поэтому необходимо коренным образом улучшить работу по научному прогнозированию наиболее перспективных проблем и направлений научно-исследовательской работы с учетом всех основных факторов развития народного хозяйства, нашего коммунистического общества. При этом следует помнить, что современный этап развития науки как никогда требует системного подхода к решению любой научной задачи. Так, в медицинской науке нужна интеграция усилий физиологов, биохимиков, морфологов, математиков, а часто и физиков, чтобы обеспечить все стороны исследований воздействия на организм того или иного изучаемого патологического фактора. Самое главное и самое трудное заключается в том, что намечается переход от изучения изменений отдельных органов организма к изучению физиологических систем и как верши-

не всего научного творчества к изучению всех основных взаимосвязей в целом организме.

Наконец, высокая эффективность деятельности научных учреждений в значительной степени зависит от хорошо поставленной службы научно-технической информации. Именно полная достоверная и своевременная научно-техническая информация дает ученому исходные материалы для творческой работы, а руководителю данные для предотвращения дублирования при планировании новых работ.

1. Вопросы эффективности научной работы

До недавнего времени наука развивалась в основном за счет экстенсивного ее роста — увеличения числа научно-исследовательских институтов, проектно-конструкторских организаций, научных работников. Дальнейшее экстенсивное развитие науки, а также отдельных отраслей народного хозяйства почти полностью себя исчерпало. В то же время возможность интенсивного развития науки (повышения производительности труда ученого, коэффициента использования дорогостоящего оборудования, приборов, внедрения методов научного управления в институтах и т. д.) используется еще недостаточно. Отсюда возникает острая необходимость в анализе и разработке методов количественных оценок эффективности труда ученых и научных институтов в целом. Необходимо изучение технологии труда ученого, чтобы вскрыть «узкие» места в его деятельности и наметить основные пути повышения производительности труда.

Академик С. Г. Струмплин еще в 1932 г. писал, что наука, начинающаяся там, где мы имеем дело с мерой, весом и счетом, не разрешила еще задачу измерения объективной ценности даже своих собственных достижений. Действительно, до настоящего времени не отработана методология количественной оценки результатов научного труда. Существующие методы оценки являются качественными и к тому же они весьма далеки от совершенства. Более того, нередко они не стимулируют повышения эффективности труда ученых и научных коллективов.

Отыскание правильных критериев следует начинать с анализа недостатков существующих методов оценки. Известно, что труд исследователей должен завершиться определенным экономическим эффектом. Величина этого

эффекта, по-видимому, может служить одним из показателей производительности труда ученого. Отсюда ряд авторов высказывается в пользу учета экономического эффекта как единственного количественного критерия продуктивности научного труда. В действительности эффективность деятельности научно-исследовательских институтов и конструкторских бюро нельзя оценить по числу научных и опытно-конструкторских работ, объему затрат на их проведение, фонду зарплаты или по затратам на одного работающего. Практика показала, что при формальном благополучии этих показателей действительная эффективность работы научно-исследовательских институтов и конструкторских бюро может оказаться невысокой. В частности, действовавшая до последнего времени система оценки их деятельности приводила к тому, что, например, выполнение тематики с меньшими затратами, чем предусмотрено сметой, означало невыполнение плана.

При оценке деятельности творческих организаций нельзя также пользоваться показателем прибыли, полученной научно-исследовательским институтом и конструкторским бюро, как главным критерием их работы. До настоящего времени методы определения сметной стоимости научных работ далеки от совершенства, поэтому такой критерий часто приводит к завышению планируемой стоимости проведения работ. В результате этого фактическая стоимость работы получается значительно ниже плановой, а разница между ними, оцениваемая как прибыль, оказывается не соответствующей истинной прибыли. Таким образом, этот критерий не стимулирует повышения эффективности деятельности научно-исследовательских институтов и конструкторских бюро. Объективные показатели эффективности работы этих учреждений должны морально и материально стимулировать творческую активность ученых, их связи, повышать отдачу практике в более короткие сроки при минимальных затратах. К сожалению, премии устанавливаются в определенном проценте к фонду зарплаты и при этом не учитывается эффективность выполненных работ. Применяемые показатели оценки результатов деятельности научно-исследовательских институтов и конструкторских бюро (число работ и их стоимость) стимулируют рост численности и фонда зарплаты, но несколько не влияют на качество выполненных работ и отдачу каждого сотрудника.

Оценка этих учреждений только по экономическому эффекту неправильна и по следующим соображениям. Необходимо различать теоретические исследования, прикладные исследования и разработки. Только к последним можно целиком применить оценку результатов по экономическому эффекту. Результаты теоретических и прикладных исследований принципиально не могут быть оценены только экономическим эффектом.

Как правило, научные исследования ведутся не на пустом месте (не от нуля). Их исполнители используют научные работы, опубликованные или еще не опубликованные в печати, при этом такие исследования могут охватывать широкий круг вопросов, часто не связанных друг с другом и даже из разных областей знаний. Оценка же экономического эффекта нельзя сделать по разработке в целом без учета экономического вклада использованных научных материалов, т. к. это не будет отражать действительного положения вещей. Полученные результаты научных исследований, опубликованные в печати, могут быть использованы не только любым лицом, читающим литературу, но и в зависимости от их ценности — многими исследователями в течение ряда лет, пока они не потеряли актуальность. Это также не позволяет оценить экономический эффект научного исследования.

В связи с проблемой долговечности творческих идей, как отмечал акад. С. Г. Струмилин, возникают неразрешимые задачи, касающиеся индивидуального учета их эффективности. Полезное действие научной продукции выходит на неопределенные сроки за пределы года. Экономический эффект дает не непосредственно научная продукция, а результат ее использования. Оценка по экономическому эффекту противоречит принципу оценки непосредственно по количеству произведенной продукции.

Вследствие различия отраслей и масштабов применения работ величина экономического эффекта зависит от факторов, не подвластных исследователю.

Чрезмерное увлечение работами, дающими непосредственный экономический эффект, наносит непоправимый вред фундаментальным исследованиям, приносящим новые идеи и открытия. Голый практицизм ведет к вырождению истинных научных исследований. Необходимо развивать и фундаментальные исследования, и приклад-

ные исследования, и разработки, не отдавая предпочтения тем или другим. Только при этом условии можно добиться быстреего развития научно-технического прогресса.

Наука является творчеством. Результаты научных исследований в основном зависят от самих ученых — их таланта, и, конечно, здесь требуется громадный труд. Отсутствие талантливых ученых в институте ничем восполнить нельзя, поэтому главными задачами организации научной работы являются подбор и подготовка ученых, способных решать на современном уровне поставленные задачи. Высокая эффективность труда научных работников должна стимулироваться морально и материально. О работе ученого-теоретика можно судить только по его трудам, публикациям. Они также необходимы и для дальнейшего развития самой науки. В связи с этим для оценки результатов труда теоретиков необходимо проанализировать их научные публикации. Капитальное монографическое исследование, как правило, является важным вкладом в науку и практику. Именно поэтому в последнее время ряд авторов выдвигают в качестве количественной и непосредственной меры научной продукции публикации в печати. По-видимому, с определенными оговорками научные публикации могут быть приняты как один из критериев научной продукции, но не больше. При прохождении конкурсов в научных учреждениях широко используется критерий численности опубликованных научных работ, однако к этому критерию следует относиться с большой осторожностью. Нередко в опубликованных работах нет никаких новых научных фактов, а содержится лишь пересказ уже известного. Такие работы ничего общего с наукой не имеют.

Некоторые научные сотрудники охотно публикуют свои материалы и выступают с докладами по малейшему поводу и по маловажным темам. Наряду с этим имеются научные сотрудники, которые, наоборот, очень редко публикуют свои работы, и как раз среди таких людей часто встречаются выдающиеся ученые.

Руководители институтов, лабораторий, отделов, естественно, как ученые участвуют в ряде работ и руководят ими и также, естественно, публикуются в составе коллективов авторов. Но отмечаются случаи (к счастью, не такие уж частые), когда руководитель не участвует в данном исследовании, хотя в силу своего служебного по-

ложения оказывается (без особого на то согласия коллектива) в числе авторов. В итоге такие «ученые» оказываются авторами 300—400 научных статей и монографий. И это должно вызывать явное недоверие к «ученому». Ведь физически невозможно создать самостоятельно такое громадное число работ! Иногда в перечне публикаций фигурируют газетные статьи, очерки, популярные брошюры, ничего общего не имеющие с наукой. Одно дело популяризировать науку и совсем другое дело ее создавать, разрабатывать новые теории, выдвигать оригинальные идеи, подтверждая их экспериментальными и теоретическими расчетами.

К сожалению, при оценке значимости научных работ отдельными учеными проявляется субъективизм. Объективная оценка опубликованной научной работы возможна только на основе мнения коллективов авторитетных ученых при строгом соблюдении свободы научной критики, при обмене и борьбе мнений. Истинная наука ничего не принимает на веру без убедительных фактов, материалов, поэтому при оценке результатов труда ученых надо повышать значение научных обсуждений, коллективного мнения.

Методика оценки научных публикаций с учетом не только объема публикации, но и ее качества не разработана. Рекомендованная акад. С. Г. Струмилиным в 1932 г. оценка публикаций в листо-баллах не применяется прежде всего потому, что чрезвычайно трудно определить количественные критерии балльной оценки публикации.

В настоящее время сформулированы некоторые общие требования к количественному критерию научной деятельности:

1) универсальность, применимость к любому виду исследования,

2) объективность, однозначность (одинаковые численные значения для любого наблюдателя),

3) непосредственность — критерий должен быть прямой мерой научной продукции, а не мерой ее использования или специального оформления для публикации, должен быть связан с затраченным трудом, должен допускать по возможности индивидуализацию научной продукции,

4) независимость от сроков, полноты и масштабов использования; должен позволять одинаково учиты-

вать как используемую, так и неиспользуемую продукцию,

5) исключение моментов, не зависящих от исследователей (например, возможности опубликования).

К сожалению, ни один из известных опубликованных количественных критериев научной продукции этим требованиям не удовлетворяет. По этой причине для оценки эффективности деятельности научно-исследовательских институтов и конструкторских бюро целесообразнее использовать системы из нескольких взаимосвязанных показателей. Эти показатели должны отражать основные задачи институтов и конструкторских бюро, быть объективными, иметь количественное выражение, базироваться на статистической отчетности, возможно более полно и точно характеризовать многогранную деятельность научно-исследовательских институтов и конструкторских бюро.

Следует иметь в виду, что сам выбор системы показателей не решает проблемы. Применяя систему показателей эффективности работы научно-исследовательских институтов и конструкторских бюро, надо уметь сделать правильный вывод на основе анализа полученных значений отдельных показателей.

Проблема разработки количественных критериев результатов научно-исследовательских работ крайне трудна. Однако ее решение позволит поставить оценку научной деятельности на подлинно научную количественную сторону. С этой целью, а также для разработки путей повышения эффективности научно-исследовательских организаций в нашей стране и за рубежом создаются научные организации, целью которых является изучение в институтах вопросов финансирования, планирования, использования дорогостоящего оборудования, форм и методов информации, повышения производительности труда ученых.

Организация труда ученых. Когда говорят о повышении производительности труда ученого, то следует отчетливо видеть отличия труда исследователя и труда производственника. Эти отличия коренные, принципиальные. В производственном процессе в технологическом ритме участвуют и рабочие, и инженеры, и служащие, труд же ученого строго индивидуален и не может быть втиснут в рамки строгого нормирования (в части творческого процесса). При совершенствовании производства, меха-

низации и автоматизации различных производственных процессов живой труд вытесняется автоматикой и роль человека в процессе производства все больше приобретает управляющее значение, творческий же труд ученого не поддается никакой автоматизации. За счет применения средств вычислительной техники можно повысить продуктивность труда ученого, но творческие процессы всегда будут выполняться человеком.

Повышению производительности труда научных работников способствуют наличие в коллективе крупных ученых, благоприятного психологического климата, правильный подбор кадров, соблюдение правил гигиены умственного труда, обеспеченность современным научным оборудованием, высокий уровень механизации и автоматизации подготовки, проведения и обработки результатов экспериментальных исследований, а также широкое использование научно-технической информации.

Для выполнения значительной научной работы, решения проблемных вопросов, безусловно, необходимо участие большого числа научных работников и вспомогательного персонала, но целенаправленное исследование должно проводиться под руководством крупного ученого, пользующегося заслуженным авторитетом в данной области, создавшего научную школу. При этих условиях можно ожидать выхода серьезных работ, оригинальных идей, имеющих большое значение для науки и практики. Когда же лабораторию или отдел возглавляют бесперспективные «остепененные» люди, для которых единственной целью в жизни было добиться ученой степени и стать руководителем научного подразделения, то от таких коллективов нельзя ожидать успехов в научной деятельности. Это свидетельствует о том, что при подборе научных кадров, выдвижении ученого на роль руководителя должно исключительно тщательно изучаться его научное и общественное лицо, на это не следует жалеть времени и сил. И здесь должна по-настоящему проявляться партийная принципиальность. Ошибок в выдвижении научных работников не должно быть! Проблема подготовки и выдвижения научных кадров всегда будет проблемой номер один.

Труд ученого далеко не легкий, поэтому ученых надо всемерно освобождать от забот по лаборатории, отделу, в частности хозяйственных. Совершенно недопустимо, когда заведующий лабораторией, крупный ученый, тра-

тит большую часть своего драгоценного времени на то, чтобы добиться получения необходимых животных, реактивов, оборудования для развертывания исследования. В то же время хозяйственный персонал института, обязанный заниматься этими вопросами, остается невозможным. Нет животных — ничего, подождите, отсутствуют реактивы — потерпите и т. д. Это безответственное отношение к хозяйственному обслуживанию в институтах должно быть преодолено.

Необходимость в квалифицированных хозяйственных руководителях чрезвычайно велика, и проблема их подготовки должна быть решена в ближайшее время. Настоящий хозяйственный руководитель — это профессия, и не простая.

Много времени ученых уходит на участие в заседаниях, совещаниях, конференциях, семинарах и т. д. Здесь средством экономии времени являются тщательная подготовка различных совещаний, их четкая регламентация, ограниченный состав участников, необходимых для квалифицированного обсуждения.

Совещания и другие мероприятия не должны дублировать друг друга. Нужно тщательно обдумать необходимость совещания, а в сомнительных случаях лучше вообще его не проводить. Очень полезно также ретроспективно за какой-то отрезок времени (полгода, год) тщательно проанализировать эффективность проведенных совещаний, оценить, что они дали для работы, какие предложения, выдвинутые на совещаниях, приняты к реализации или уже реализованы. Такой анализ поможет на будущее принимать действенные меры к повышению эффективности совещаний. Времени занимают они много, и отдача от них должна быть весомая.

Не менее важное значение имеет создание нормальных условий труда и прежде всего положительного психологического климата в работе. Здесь, помимо руководителя подразделения, значительная роль принадлежит авторитетным сотрудникам, которые, не занимая административных управленческих должностей, оказывают серьезное влияние на всю жизнедеятельность учреждения. Руководитель должен не только считаться с этими неформальными группами научных и практических работников, а работать с ними в тесном контакте, прислушиваться к их замечаниям и предложениям. Идеально, когда руководитель подразделения является одновремен-

но лидером в определенной области науки, но это бывает далеко не всегда.

Руководитель подразделения занимает особое место в структуре научного учреждения, так как в настоящее время комплексное решение большинства научных задач и проблем под силу только многочисленным высококвалифицированным коллективам. Первоочередной задачей руководителя являются кооперирование и координация исследований, проводимых в различных лабораториях и научно-исследовательских институтах. Одним из итогов работы коллектива является информация о новых фактах, идеях. Полученные результаты экспериментального исследования обсуждаются на научном совете лаборатории, отдела, института. Это обсуждение должно проводиться не только после завершения всей работы, а постоянно, в процессе исследования, при получении промежуточных данных, при этом обсуждению должен быть свойствен дух объективного научного анализа материала, доброжелательного отношения к его автору. Цель научной критики — помочь товарищу правильно интерпретировать представленные материалы, наметить дальнейшие пути выполнения работы. И здесь главная роль принадлежит руководителю лаборатории, отдела.

Положительный психологический климат, бережное, внимательное отношение друг к другу создают оптимальные условия для проявления максимума творчества в работе, желания работать еще лучше и добиваться большей отдачи в работе. И, наоборот, отрицательный психологический климат (окрики, административирование в работе и обсуждении хода выполнения научного исследования) вызывает угнетенное настроение, нежелание работать и нарушает нормальную работу коллектива.

Роль научно-технической информации. Одним из важнейших факторов, повышающих эффективность деятельности научного учреждения, является научно-техническая информация. Только при исчерпывающем использовании научной информации по достижениям отечественной и зарубежной науки можно обеспечить актуальность и новизну планируемых работ, избежать их ненужного дублирования.

В вопросах планирования новых работ и в процессе их проведения служба научно-технической информации выявляет и своевременно информирует научных работников о наиболее перспективных научно-технических на-

правлениях отечественных и зарубежных исследований, при этом научные сотрудники не тратят своего времени на поиск и доставку научной информации.

Быстрый и все более ускоряющийся рост числа научных публикаций, увеличивающий необозримость печатной продукции, делает роль службы научно-технической информации все более важной.

Количество опубликованных в мире книг по естественным и техническим наукам в настоящее время превысило 30 млн. названий, издается более 100 тыс. журналов, публикующих ежегодно около 9 млн. статей, а также большое количество специальной литературы, трудов различных ассоциаций, институтов, конференций и симпозиумов. Ежегодно выпускается свыше 350 тыс. описаний к патентам и авторским свидетельствам. Объем мирового патентного фонда уже превысил 12 млн. описаний.

Мировая статистика показывает, что число публикаций ежегодно увеличивается на 8—10%. Это означает, что количество публикаций удваивается каждые 10—15 лет. Ясно, что ни один ученый или специалист не в состоянии самостоятельно выбрать из этого потока и прочесть все нужные ему публикации.

Важность роли службы научно-технической информации подчеркивается еще и тем обстоятельством, что, снабжая специалистов необходимой им информацией, эта служба ограждает их от огромного потока материалов, не представляющих для них интереса.

Основными требованиями к службе научной информации являются полнота и оперативность информационного обслуживания. Это достигается с помощью механизации и автоматизации информационных процессов, путем создания автоматизированных информационно-поисковых систем.

Научная информация, необходимая для планирования и выполнения научных исследований, включает следующие виды:

1. Библиографическая аннотированная информация. Обслуживание ведется путем оповещения о всех публикациях, появившихся за последнее время (например, за месяц), задержка не более 2 мес со дня появления статьи в журнале или книге; максимальная полнота охвата абонентов; краткость изложения сущности публикации, достаточная для того, чтобы специалист мог определить, нужна или нет ему эта публикация; удобная рубрика-

ция представляемых материалов, гарантирующая полноту поиска и просмотра библиографической информации.

2. Первичные источники информации. Такое обслуживание осуществляется путем высылки оригиналов или копий нужных статей, книг, отчетов, патентов и т. д. и их переводов по указанным библиографическим данным или единым инвентарным номерам.

3. Аналитическая обзорная информация. Обслуживание сводится к составлению специализированных аналитических обзоров высококвалифицированным специалистом, способным оценить данные, появившиеся за определенный период. В этих обзорах должны подводиться итоги, а также даваться прогнозы и рекомендации на будущее. Подготовка таких обзоров должна идти полуавтоматическим способом: с помощью ЭВМ могут получаться библиографические списки литературы, таблицы фактов и сжатые сводки прежних обзоров, которые затем должны анализироваться соответствующими специалистами.

4. Справочная фактографическая информация. Для наиболее полного обеспечения фактографическими данными необходимо использование ЭВМ, в которой должны храниться фактические данные о состоянии и результатах исследований, научных учреждениях и ученых, оборудовании, средствах, методах, химических веществах (медикаментах).

Автоматизированные информационно-поисковые системы могут использоваться при решении задач прогнозирования развития науки по основным направлениям и проблемам для накопления и систематизации данных о медикаментах, медицинских приборах и оборудовании, научных приборах, о тематике научных исследований, составе, профиле и уровне подготовки ученых.

2. О путях совершенствования структуры научно-исследовательских учреждений

Эффективность деятельности научно-исследовательских организаций и учреждений во многом зависит от их размера и от численности коллектива. В настоящее время еще много существует и создается карликовых научных учреждений, безуспешно пытающихся решать многочисленные научные задачи. Между тем крупные проб-

лемы под силу решать лишь крупным научным учреждениям и коллективам. В связи с этим необходимо добиваться укрупнения учреждений и создания новых, оснащенных современной техникой и оборудованием.

Определение оптимальных размеров научно-исследовательских и проектно-конструкторских организаций — задача сложная, и она должна решаться применительно к поставленным задачам.

В больших институтах (1000 человек и более) должна быть тщательно продумана и официально закреплена научная, административная и хозяйственная самостоятельность крупных подразделений, в противном случае такой организацией трудно управлять одному руководителю — директору института. Должны быть отработаны четкие (без мелочной опеки) функциональные обязанности сотрудников в подразделениях, где должны быть показаны не только обязанности, но и соответствующие обязанностям права.

Одним из важнейших вопросов повышения эффективности науки является также создание оптимальной структуры научно-исследовательского учреждения, отвечающей поставленным научным проблемам.

Как известно, структура научного учреждения может строиться по функциональному или проблемному принципу, а может быть и смешанной.

При функциональном принципе работы института в лабораториях объединены сотрудники однородных специальностей, при проблемном принципе — работники различных специальностей для комплексной работы над одной проблемой.

Функциональный принцип построения структуры, как правило, приводит к лучшему использованию сложной и дорогостоящей аппаратуры. Поэтому, например, в медико-биологических научных учреждениях (да и не только в них) целесообразно создание широкого круга централизованных лабораторий, обеспечивающих потребности всего учреждения. Так, в отдельных учреждениях уже создаются лаборатории электронной микроскопии, парамагнитного резонанса, биохимические, морфологические лаборатории и др., в которых концентрируется уникальная дорогостоящая аппаратура. Лаборатории укомплектовываются квалифицированными специалистами, владеющими сложными методами исследования. Создание подобных централизованных лабораторий дает

возможность ускорения научного исследования и определенный экономический эффект благодаря полной загрузке сложной аппаратуры и ее квалифицированному использованию. Вместе с тем функциональный принцип значительно усложняет организацию комплексных исследований. При функциональной структуре института комплексность, как правило, обеспечивается проблемными комиссиями, координирующими межлабораторные исследования по курируемой проблеме. К сожалению, обычно курирование носит в основном консультативный характер, так как проблемные комиссии административными правами не наделены.

Таким образом, во многих случаях наиболее эффективной оказалась бы, по-видимому, смешанная структура института (по функциональному и проблемному принципу). Смешанная структура института дает возможность комплексного решения проблем с использованием преимуществ, даваемых специализацией лабораторий. К сожалению, эта форма структуры еще недостаточно изучена.

Очень важным и сложным является вопрос о численности научных работников в отдельных подразделениях института, соотношении научных работников и обслуживающего персонала. Наш опыт показывает, что оптимально управляемая медико-биологическая лаборатория должна иметь 15—20 сотрудников, из них научных 5—6 человек.

Отечественные и зарубежные авторы в своих исследованиях указывают соотношение научных работников и вспомогательного персонала от 1 : 2 до 1 : 6. Это соотношение зависит от многих факторов, в том числе и от характера исследований. При внедрении новой техники, автоматизации исследований должна увеличиваться численность вспомогательного персонала, включая инженерно-технических работников. При фундаментальных исследованиях, естественно, больше привлекается научных работников, а при прикладных — вспомогательного персонала. На практике часто встречаются лаборатории, включающие фундаментальные и прикладные исследования.

На основании изучения отечественного и зарубежного опыта В. Соминский и М. Юделевич (1966) рекомендуют следующую структуру первичной научно-исследовательской группы (табл. 14).

Примерный состав научно-исследовательской группы
в зависимости от характера проводимых исследований

Вид исследования	Старшие научные сотрудники	Ведущие инженеры	Старшие инженеры	Инженеры	Младшие научные сотрудники	Лаборанты	Техники, чертежники, рабочие	Служащие
Фундаментальные	1	—	—	—	2—3	3—5	—	1
Прикладные	1	—	—	—	3—5	6—10	—	1
Опытно-конструкторские разработки	1	1	2—3	6—10	—	—	18—30	2—3

Для окончательного выбора состава научного подразделения должны быть тщательно изучены цель исследования, уровень научной и технической квалификации сотрудников, обеспеченность оборудованием, научной информацией и многое другое.

При организации института должны быть четко определены и утверждены положения о правах и обязанностях каждого подразделения, сотрудника.

Сотрудники обязаны знать Положение и строго руководствоваться им в своей организационной и научной деятельности. Отсутствие Положения приводит к излишнему дублированию работ, безответственности сотрудников. Ни в коем случае не должно быть у сотрудников двойственного руководства, т. е. никто не должен находиться в подчинении у нескольких руководителей.

В литературе длительное время дебатировался вопрос о том, какие качества руководителя научного учреждения более важны с точки зрения эффективности работы руководимого им учреждения — качество хорошего организатора или крупного ученого? В истории известны отдельные случаи, когда в одном лице сочетались крупный организатор науки и ведущий ученый. Такими качествами, например, обладали выдающиеся ученые и организаторы науки академики И. В. Курчатов и С. П. Королев. Но эти случаи очень редки. Обычно имеет место либо работа руководителя научной организации как организатора, либо как ученого. При существующем положении целесообразнее руководство научно-исследова-

тельским институтом возложить на ученого-администратора, прошедшего специальное обучение по современным методам управления. Такой ученый-администратор должен возглавить планирование исследований в институте, наладить связь с другими научными организациями, контроль за выполнением плана научно-исследовательских работ, финансирование, подбор и расстановку научных кадров, внедрение в практику результатов научных достижений в институте, выявление всех резервов в институте с целью их оптимального использования.

3. Кибернетические методы научного прогнозирования в медицине и здравоохранении

Одним из главных вопросов, определяющих всю деятельность научно-исследовательского института, является долгосрочное прогнозирование курируемой области науки. Прогнозирование развития научных исследований — сложнейший вопрос, требующий направления сил всего творческого коллектива научного учреждения. Научно обоснованный прогноз развития отдельных направлений науки служит основой перспективных планов развития всего народного хозяйства. Обоснованность прогнозов может быть обеспечена при условии, что они базируются на марксистской теории развития общества, глубоком познании и всестороннем учете действия экономических законов социализма. Прогнозы всегда имеют вероятностный характер, потому что в них невозможно учесть весь комплекс причин, влияющих на пути развития научно-технического прогресса.

Целью науки является познание законов развития природы и общества и использование их в практической деятельности. Научное предвидение основано на знании этих законов. Люди могут успешно осуществить лишь такую деятельность, которая протекает в соответствии с законами природы и общества. Предвидение и прогнозирование новых явлений основаны на раскрытии тенденций, т. е. направления, по которому идет развитие какого-либо явления.

При разработке прогноза первым этапом является тщательный анализ достигнутого уровня исследований на сегодняшний день и наметившейся тенденции его развития. Анализ изучаемой научной проблемы должен включать не только состояние отечественной науки, но и срав-

нительный анализ зарубежных данных, материалы по участию наших ученых в международных съездах, конференциях, симпозиумах, зарубежных научных командировках. Особое значение имеет прогнозирование при планировании работ в новых отраслях промышленности и науки, развитие которых идет бурными темпами.

Для составления прогноза должны также использоваться ежегодные итоги по научным исследованиям. Институты зачастую составляют ежегодный научный отчет по отдельным темам, чтобы показать выполнение плана. Нам представляется целесообразным итоги выполнения плана (для финансовых органов) по отдельным темам составлять очень кратко: тема, выполнение, финансирование, затраты. Основное же внимание, по-видимому, следует обратить на ежегодное составление проблемных научных отчетов, в которые должны входить темы, относящиеся к конкретной проблеме. Такой научный проблемный отчет должен также содержать сравнительный анализ состояния проблемы в институте, нашей стране и за рубежом, а также оценку тенденции развития изучаемой проблемы на будущее. Такие проблемные отчеты разрабатывать не просто, но они дадут возможность серьезно обсуждать ежегодно их в институте и намечать научно обоснованные тенденции их развития.

Отчет должен сопровождаться отдельным планом-графиком внедрения в практику полученных достижений института, при этом в одном случае это будут научные публикации, в другом — методика, предложенный новый препарат, прибор и т. д.

Достоверность прогнозирования в значительной степени зависит от правильности выбора метода прогнозирования. Каждый метод прогнозирования имеет определенную область применения. Известно более 130 способов и методов научно обоснованного прогнозирования. Несмотря на такое большое количество методов, вопросы научного прогнозирования не решены до настоящего времени. Не определены четкие показания к выбору того или иного способа прогнозирования применительно к конкретной ситуации.

Для составления прогноза используются самые различные материалы: стандарты, технические условия, паспорта, отчеты о законченных научно-исследовательских и опытно-конструкторских работах (последние имеют особое значение), патентная документация, которая опе-

режает все другие печатные виды информации в среднем на 3—4 года, материалы научных съездов, конференций, симпозиумов и т. п.

Для целей прогнозирования может также использоваться анализ динамики научных публикаций в той или иной области за определенный период, предшествующий времени составления прогноза.

При оценке точности прогноза используют как традиционные формулы теории вероятностей, математической статистики, так и более сложные математические методы. Следует подчеркнуть, что любой прогноз носит в определенной степени вероятностный характер, даже если при его составлении использована надежная первичная информация. Прогноз будущего развития, безусловно, включает ряд социальных и других факторов, которые требуется уточнять по мере накопления новых фактов.

В настоящее время в методике научного прогнозирования часто используется метод экстраполяции. Основой метода являются изучение существующей динамики в данной области и экстраполяция этих данных на будущее. При этом необходимо учитывать возможность появления новых факторов, которые окажут влияние на развитие изучаемого процесса. Экстраполяция, как правило, применима к прогнозированию изменений только отдельных параметров.

Метод экспертных оценок особое значение имеет при прогнозировании научно-технического прогресса и других процессов. По одному из методов этой группы — методу Делфи — подбирают группу экспертов. Всем экспертам задают определенный вопрос, на который они индивидуально отвечают. Данные ответов анализируют и выводят средний ответ по данному вопросу. Точность прогноза с помощью метода экспертных оценок невысока; она может быть значительно повышена при введении постоянно действующей системы прогноза с использованием комплекса методов.

Событийную систему используют для оценки времени наступления тех или иных событий, при этом для каждого события в отличие от метода экспертной оценки используют отдельную группу экспертов. Каждая группа оценивает время наступления изучаемого события в зависимости от ряда других событий. В итоге получается сеть взаимосвязанных экспертных оценок, которые составляют основу математической модели прогноза. Эти

данные вводятся в ЭВМ, где по определенным программам вычисляются оценки всех событий, подвергавшиеся экспертизе. После обработки на ЭВМ вновь привлекают экспертов для окончательной оценки полученных данных. Эксперты оценивают не только время наступления того или иного события, но и потребные для этого материальные ресурсы. Таким образом прогноз постепенно превращается в план.

Процессионная система используется для прогнозирования изменений разных параметров (например, годовой прирост населения) в изучаемом объекте. Применяя эту систему, вначале (экспертным путем) выбирают наиболее существенные параметры, характеризующие объект в целом, а затем устанавливают взаимосвязь этих параметров. Установленные взаимосвязи и их влияние друг на друга, значения параметров закладывают в ЭВМ. По определенной программе ЭВМ обрабатывает введенные данные и выдает возможные варианты развития объекта. В процессионной системе также происходит постоянное взаимодействие с экспертами.

В настоящее время оптимальными сроками прогноза экономического и социального развития общества приняты 10—15 лет, государственные пятилетние планы и государственные годовые планы.

Проблема планирования научных исследований (в основном на пятилетие) исключительно актуальна. Развитие науки на основе плана — естественное следствие социалистической плановой экономики, проявление ее преимущества перед капиталистической. В Советском Союзе создана общегосударственная система планирования развития науки и техники. Как отмечается в книге «Проблемы управления наукой в условиях научно-технической революции» (М., 1972), трудности планирования в науке обусловлены следующими факторами. Во-первых, исследование всегда носит вероятностный характер и связано с определенным риском, степень которого трудно определить. Во-вторых, фундаментальные и частично прикладные (поисковые, теоретические) исследования нельзя оценить непосредственно с помощью экономических критериев — прибыли, рентабельности и т. п. В-третьих, залогом успеха исследования является соответствие его целей призванию и интересам исследователя, которые не всегда можно учесть при планировании общегосударственных научных задач. И, наконец, расходы на

исследование осуществляются быстро (обычно в течение года), внедрение — медленно (на протяжении 2—3 лет), а «отдача» — и того больше (в течение 5—8 лет).

Основу прогнозирования составляет информационная модель системы, описывающая структуру и функции системы и зависимости между характеристиками системы и ее внутренними и внешними параметрами.

Характеристики системы — ее данные, важные с точки зрения ее использования или выполнения ею тех функций, для которых она предназначена. По сути дела характеристики — это критерии качества работы системы. Параметры системы — ее данные, определяющие процесс функционирования. Внутренние параметры определяют связи между определенными частями системы. Внешние параметры определяют связи системы с внешней средой.

Различают изменяемые и неизменяемые параметры. Изменяемые параметры — это параметры, значения которых могут меняться в процессе функционирования системы (по определенному закону — статистическому или детерминированному). Неизменяемые параметры не меняются в процессе работы системы, но могут иметь разные значения для разных вариантов систем. Основным методом прогнозирования сложных систем является метод моделирования с помощью ЭВМ.

Прогнозирование в таких широких областях человеческой деятельности, как медицина и здравоохранение, может осуществляться на основе метода последовательных приближений, с использованием ЭВМ. В соответствии с этим методом вся область исследований делится на узкие и более или менее автономные подобласти или направления, для которых могут быть построены модели и осуществлено их автономное прогнозирование. Затем на втором этапе учитываются взаимные связи между выделенными направлениями и осуществляется моделирование и прогнозирование развития всей области. Этот метод последовательного расчленения и поэтапного моделирования может быть применен и к прогнозированию выделенных на первом этапе подобластей путем расчленения их на направления третьего уровня и т. д. В области медицины и здравоохранения в качестве основных направлений (первого уровня) могут быть выделены: развитие медицинской науки; прогресс медицинской техники и промышленности; организация системы практического здравоохранения; подготовка медицинских кадров.

Каждое из этих крупных направлений, естественно, может быть подразделено на более конкретные направления, прогнозирование которых, однако, само по себе представляет собой сложную проблему. Так, для медицинской науки характерно деление ее на комплексные проблемы, направления, разделы, которые соответствуют тематическому содержанию работ и определены перспективными планами Министерства здравоохранения СССР и АМН СССР. Кроме того, в медицинской науке целесообразно выделить ряд аспектов накопления и систематизации информации следующего вида:

1. Развитие фундаментальных исследований по основным медико-биологическим проблемам и ожидаемые основные научные открытия и достижения.

2. Развитие и создание новых медицинских научных центров и учреждений, их оценка в количественном и качественном отношении, включая объемы финансирования, численность сотрудников, специализацию, оборудование, размещение, взаимные связи и т. д.

3. Развитие научных школ и направлений, формируемых ведущими учеными.

4. Разработка и применение новых методов и средств диагностики и лечения.

5. Развитие системы научной медицинской информации.

6. Разработка и внедрение новых медикаментов и прогнозирование их основных и побочных действий.

Исходной базой научного прогнозирования является достаточно полная, достоверная и систематизированная информация о состоянии медицинской науки. Основными принципами накопления, систематизации и анализа научной медицинской информации являются четкая классификация направлений медицинской науки и выделение узловых пунктов или объектов, характерных для каждого направления. Намеченные узловые пункты или объекты становятся «центрами» группирования информации, для которых могут быть составлены однотипные формализованные машинные описания (анкеты). При этом общий поток научной медицинской информации, содержащейся в статьях, отчетах, книгах, патентах и других источниках, разделяется по определенным аспектам и накапливается в стандартном формализованном виде, удобном для машинного анализа и обработки. При этом существенно наличие различных связей (ассоциативных, иерархических

и т. п.) между объектами внутри направлений и между ними.

Методы прогнозирования в медицинской науке аналогичны методам, используемым в других естественных и технических науках. Особенностью здесь является глубокая и разносторонняя связь и взаимозависимость как между разными направлениями и дисциплинами самой медицинской науки, так и с другими разделами науки и техники. На состояние медицинской науки оказывают непосредственное влияние состояние других наук, уровень техники, производства, экономики. Глубокая связь различных медицинских наук обусловлена прежде всего единством объекта изучения — человеческого организма, его центральным положением в окружающей нас природе. В связи с этим специфической особенностью задач научного прогнозирования в медицинской науке является необходимость накопления и обработки огромного количества фактических данных, что может быть сделано только с помощью ЭВМ.

В медицине и здравоохранении могут быть использованы следующие основные методы и направления научного прогнозирования:

1) метод экспертных оценок с последовательным этапно-временным разворачиванием проблем (метод Делфи);

2) метод логического анализа прямых причинно-временных зависимостей между стадиями развития и внедрения научных открытий и изобретений (метод «обратной связи»);

3) метод статистического анализа динамики изменения параметров изделий. Применим для прогнозирования медицинского оборудования, научных приборов и технических средств диагностики и лечения;

4) метод структурных родо-видовых моделей. В первую очередь используется для прогнозирования в области фармакологии с целью систематизации данных о медикаментах, их свойствах и строении и определения свойств новых медикаментов, относящихся к определенным видам. При этом возможно, например, прогнозирование побочных действий новых лекарств;

5) метод накопления и статистической обработки данных о финансировании и составе ученых основных зарубежных научных центров и определения таким путем направлений исследований, на которых сосредоточиваются основные силы;

б) метод статистического анализа потоков научной информации и определения изменений в объемах публикаций по годам по различным направлениям исследований, выявления внутренних междисциплинарных связей и ссылок.

Для прогнозирования науки необходимо прежде всего иметь достаточно полные сведения о текущем состоянии науки в форме, допускающей быстрый поиск данных по признакам, обновление, систематизацию и классификацию данных. Это может быть достигнуто на основе использования информационно-логических машин, хранящих в своих запоминающихся устройствах необходимые данные вместе с отсылками к другим данным, показывающим внутренние смысловые связи и структуру классификации информации.

Такие системы пока не созданы, но потребность в них ощущается все более остро. Они должны предназначаться для накопления фактографических данных о веществах, материалах, приборах, исследованиях, болезнях и т. д. и для выдачи по запросам как адресов первичных источников информации (документов, статей, отчетов и др.), так и конкретных фактических сведений, которые могут использоваться потребителями без дополнительных обращений к первичным источникам. Такие системы мы будем называть сокращенно АФИЛС (автоматизированная фактографическая информационно-логическая система).

В связи с большим объемом информации, подлежащей хранению, и узкоспециальным характером задачи прогнозирования развития науки целесообразно создавать АФИЛС для узких разделов знаний с тем, чтобы эти системы содержали максимально исчерпывающие, точные и достоверные данные по данному разделу. Естественно, что каждая отдельная АФИЛС может содержать необходимые отсылки к другим специализированным АФИЛС. Таким образом будет получаться система АФИЛС, охватывающая все более широкие области науки и техники. Подобным путем может быть осуществлен постепенный переход к машинному накоплению, отбору и систематизации всех научных знаний, приобретенных человечеством и содержащихся сейчас в виде различных письменных материалов. С учетом непрерывного развития науки АФИЛС должны строиться по гибкому принципу, допускающему непрерывное изменение систем

классификации, систем связей (отсылок) и состава терминологических словарей.

Основу программной реализации АФИЛС составляют ассоциативные поисковые деревья, ассоциативные узловы структуры и фактографические формулы объектов, обеспечивающие быстроту поиска, удобство ввода, накопления и вывода данных, возможность построения гибких перекрестных связей и многоаспектную классификацию и поиск данных.

В отличие от документальных поисковых систем, в которых поисковый массив образуется из поисковых образов документов, содержащих минимально необходимый для поиска набор дескрипторов и не содержащих, как правило, фактической информации, имеющейся в документах, в АФИЛС должна содержаться основная фактическая информация о данном объекте, собранная из разных источников. При этом вместо индексирования рефератов, применяемого при пополнении документальных систем, наполнение АФИЛС должно идти по пути тщательного изучения первичных документов специалистами и отбора из них ценной информации. Для ввода в АФИЛС эта информация должна соответствующим образом подготавливаться.

Таким образом, наполнение информацией АФИЛС в отличие от наполнения информацией документальных поисковых систем будет представлять собой трудоемкий процесс, требующий творческого участия квалифицированных специалистов. При вводе каждого нового сообщения в АФИЛС эти специалисты должны запрашивать, как правило, АФИЛС о наличии в ней аналогичных данных, сопоставлять новые сведения с уже имеющимися в АФИЛС, производить необходимые корректировки в старых сведениях, а новые данные вводить с учетом того, что уже имеется в системе.

В будущем, по-видимому, значительную часть такой работы по сопоставлению и корректировке данных сможет выполнять сама АФИЛС.

Таким образом, АФИЛС должна представлять собой комплексную человеко-машинную информационную систему, в которой функции творческого анализа, отбора и оценки информации выполняются человеком (во всяком случае в ближайшем будущем), а функции накопления, хранения, поиска, систематизации и обработки информации выполняются машиной.

4. Системный подход к накоплению и обработке медицинской информации

Рассмотрим возможности и особенности системного кибернетического подхода к накоплению и обработке медицинской информации на примере онкологии.

Проблема рака является одной из наиболее острых проблем современной медицины, и для ее решения необходима действенная координация усилий ученых разных стран.

Первым шагом в этом направлении может быть создание международной автоматизированной информационной системы по онкологии, обеспечивающей систематизацию, накопление и обработку экспериментальных и клинических данных по диагностике и лечению злокачественных новообразований и быстрое доведение сведений до научных работников и практических врачей разных стран. Рассмотренная система должна аккумулировать опыт и знания сотен тысяч специалистов в области онкологии и обеспечивать анализ и обобщение этого опыта и эффективный обмен знаниями в рассматриваемой области.

Система должна строиться в виде сети современных информационных центров, связанных между собой и с абонентами линиями связи с терминалами. Эта система не должна сводиться к одному из новых вариантов документальной системы (пусть даже исчерпывающей), хотя документальный поиск будет являться весьма важной ее функцией. Главной задачей подобной системы должно быть накопление систематизированных и проанализированных фактических данных по онкологии, полученных из разных стран и источников, при разных условиях наблюдения, постановки экспериментов, проведения лечебных и диагностических процедур. Для обеспечения совместимости исходных данных и сопоставимости результатов все эти данные должны извлекаться из литературных источников или представляться непосредственно специалистами разных стран в унифицированной форме, а ввод этих данных в систему должен производиться с использованием методов и средств автоматической нормализации научных текстов.

Это должно в конечном счете привести к созданию единого медицинского информационного языка по онкологии.

Новая информационная система должна обеспечить симбиоз творческих способностей ученых и технических возможностей современных ЭВМ по накоплению и переборке огромных количеств информации. Творчество людей должно проявляться прежде всего при анализе исходных данных и приведении их к единому сопоставимому виду, необходимому для последующей машинной обработки, при постановке вопросов к ЭВМ и анализе выдаваемой ЭВМ информации. Результаты анализа должны снова вводиться в ЭВМ для формирования информации следующего, более высокого уровня обобщения.

В связи с международным характером новой системы в ней должен быть реализован многоязычный ввод и вывод информации. Наиболее эффективно это может быть сделано в системах дескрипторного типа с разветвленной грамматикой. Информация в памяти ЭВМ должна храниться не на каком-либо конкретном естественном языке, а на унифицированном информационном языке, включающем коды понятий (дескрипторов) и коды синтаксических и семантических связей. Для построения новой системы не будут подходить ИПС, в которых данные представляются на каком-то одном естественном языке. Такие ИПС должны быть рассчитаны на компактную, стандартизированную, четкую запись научных, экспериментальных и клинических фактов в виде, пригодном для машинного анализа и обобщения.

Эта система должна включать документальную подсистему и ряд специализированных фактографических подсистем, соответствующих основным потокам входной информации и классам потребителей данных. Одним из основных входных потоков будет поток статистических данных о результатах лечения, разделенный по локализации и способам лечения, в частности сведения об испытаниях новых препаратов и способах лечения. Вторым потоком явятся статистические данные, касающиеся диагностики и условий возникновения злокачественных новообразований.

В решении проблемы рака, как и во многих других областях науки и техники, возможны и имеют место два параллельных и взаимно дополняющих пути развития. Первый путь заключается в сборе, накоплении, систематизации и анализе большого количества экспериментальных, клинических фактов и результатов теоретических исследований отдельных частных вопросов и сторон ме-

ханизмов рака, его диагностики и лечения. При этом огромное количество мелких результатов и фактов при постоянном их накоплении и анализе может и должно привести к глубоким качественным изменениям представлений о раке, к фундаментальным открытиям. Здесь действует принцип ценной реакции, так как полученные результаты стимулируют постановки новых исследований и клинических испытаний и новые результаты. Для ускорения этого процесса необходимо резкое ускорение процесса накопления, систематизации и доведения информации до широких кругов исследователей и практических врачей. Эту функцию и призвана решить всемирная автоматизированная информационная система по раку. Второй путь — это традиционный путь получения научных открытий, заключающийся в том, что в определенное время и в определенном месте появляется ученый, который при благоприятных обстоятельствах делает блестящие гипотезы. Очевидно, что резкое ускорение темпов движения науки по первому пути должно резко повысить вероятность открытий и на втором пути, так как гениальные открытия не рождаются на пустом месте и наличие систематизированной фактографической, легкодоступной информации по любому вопросу в области онкологии должно способствовать не только систематическому продвижению в этой области, но и появлению новых открытий.

Основными целями создания автоматизированной информационной системы по онкологии являются:

1. Обеспечение доступности всех имеющихся и получаемых научных и клинических данных по онкологии для исследователей и практических врачей, работающих в области онкологии. Это осуществляется благодаря резкому сокращению задержки во времени между выработкой новых данных и их доведением до всех заинтересованных лиц. В настоящее время такой формой доведения является публикация статей и такая задержка в лучшем случае составляет полгода. При создании ИПС эти данные должны быть доступны для использования на практике в момент их выработки. Время поиска старых данных, когда-либо ранее введенных в систему, также сокращается почти до нуля. Ввод новых данных и обращение к накопленным данным становятся одинаково быстрыми и доступными для всех научных и лечебных учреждений, занятых в системе, независимо от расстояний, на

которых расположены эти учреждения друг от друга и от информационных центров. При вводе новых данных обеспечиваются автоматическая или полуавтоматическая их систематизация, сопоставление с ранее накопленными данными и запись на хранение в запоминающие устройства ЭВМ в виде, удобном для последующей обобщенной выдачи и анализа. При этом все данные, относящиеся к одному вопросу, независимо от времени и источников получения группируются, дополняются критическими или обобщающими комментариями и выводами, а также ссылками на сведения по смежным вопросам. Периодически (раз в год или по мере необходимости) проводится ревизия накопленных данных и они заменяются обобщенными сведениями с пересылкой ранее хранимых данных в архив для длительного хранения.

Кроме того, обеспечивается ликвидация языкового барьера в обмене научными и клиническими результатами между учеными и врачами разных стран. Так как данные должны вводиться и выдаваться в виде формализованных фраз со строго фиксированным набором смысловых связей и синтаксических структур, то задача автоматического перевода и автоматического индексирования и классификации этих данных вполне реальна. Соответствующие алгоритмы уже существуют. Создание многоязычных терминологических машинных словарей (тезаурусов) по онкологии может производиться параллельно с вводом документов. Благодаря созданию таких словарей появится также возможность унификации терминологии в онкологии, что само по себе очень важно.

2. Обеспечение объективного и полного учета вклада каждого сотрудника или учреждения в целом в решение научных или практических вопросов онкологии. Все вводимые новые данные обязательно должны сопровождаться информацией о дате ввода и авторе (лице или учреждении). Вместе с этой авторской информацией данные постоянно хранятся и выдаются по запросам. Полученные на основе анализа выданных сведений обобщающие данные вводятся в систему как новые данные вместе с авторской информацией о лице, выработавшем это обобщение и критическое замечание. По желанию автора система может после ввода данных печатать ему автоматически «квитанцию» — копию введенных им данных с указанием даты ввода и единого инвентарного номера системы, присвоенного введенной порции данных. После

получения каких-либо замечаний (подтверждений, опровержений или сопоставлений) по введенным данным система высылает автору эти замечания. Кроме того, система периодически (например, раз в год) печатает сводки-бюллетени новых данных с указанием их авторства и ценности (правильности, новизны и т. д.).

3. Возможность автоматического анализа и обобщения накапливаемых данных. Так как все данные будут храниться в памяти ЭВМ в систематизированном и формализованном виде (формализованными кодированными фразами, таблицами, схемами и др.), то появляется возможность их автоматического анализа и обобщения с использованием статистических и логических методов, способов распознавания образов, автоматической классификации и др.

4. Непрерывное развитие, пополнение и уточнение системы информации по онкологии, т. е. поддержание ее в активном состоянии независимо от смены поколений ученых, врачей и других людей, создающих и использующих эту информацию.

Основным условием автоматизации процессов накопления, анализа и обобщения фактографической информации по онкологии (и вообще по медицине) является представление ее в памяти ЭВМ в нормализованном виде — в виде фраз с фиксированной структурой и лексикой. Процесс записи вводимой в систему научной информации на формализованном языке будет выполняться в будущем, как правило, людьми — специалистами и данной области знаний, так как с созданием подобных систем будет создаваться и развиваться формализованный язык данной отрасли науки. Этот язык будет включать единую терминологию, стандартную систему смысловых связей и структуру фраз и будет изучаться в соответствующих учебных заведениях. Таким образом, будущие специалисты будут в состоянии не только готовить тексты для ввода в ЭВМ на этом языке, они будут вообще использовать его для фиксации и для публикации научной информации и даже думать будут, используя этот язык.

В настоящее же время для ввода данных в ЭВМ специалисты будут представлять свои данные, как правило, на естественном научном языке, а для перевода этой информации в нормализованную форму будет использоваться система автоматической индексации с многоязычными входами и выходами.

В связи с изложенным следует подчеркнуть, что обязательным условием обеспечения эффективной работы указанной ИПС является наличие простого и надежного механизма разнесения вводимой информации по соответствующим позициям ИПС, обеспечивающего надежное отождествление и выдачу этих данных при поиске. При больших потоках входной информации самого различного характера это условие является решающим. Для обеспечения этого условия необходимо установить несколько основных признаков, которыми в обязательном порядке должна снабжаться любая вводимая в ИПС информация. Такими признаками могут быть: название органа, системы, организма, животного; название раздела медицины (диагностика, терапия, профилактика, этиология, механизм действия и т. д.); название вещества (медикаменты, препараты); название метода, типа прибора или установки. Эти четыре признака должны быть четкими, но не узкими терминами. Любые узкие термины, определения и специфические описания допустимы в последующих детализирующих пунктах вводимого сообщения.

Рассмотренная выше международная информационная система по онкологии будет представлять собой, по видимому, верхний уровень специализированной медицинской системы, в которой должны храниться в основном обобщенные научные данные, описания методов, приборов, медикаментов, результатов их применения и т. д. С помощью такой специализированной системы будет осуществляться обмен информацией с медицинскими информационными системами (МИС) научных, практических, общего и специальных профилей медицинских учреждений.

Автоматизированная информационная система медицинского учреждения должна выполнять следующие функции:

а) практическая информационная служба в больнице (отделении) по вопросам состояния, лечения и обслуживания больных;

б) обеспечение врачей фактической информацией, касающейся методик диагностики и лечения, инструктивно-методических указаний, медицинских приборов и инструментов, медикаментов, медико-биологических вопросов этиологии и патогенеза, а также статистическими данными об эффективности различных методов диагностики, лечения и т. д.;

в) обеспечение научных работников фактическими сведениями, необходимыми для выполнения исследований, координации, анализа, сопоставления и оценки уровня и направлений научно-исследовательских работ.

Данная система должна позволять находить детальные сведения по конкретным вопросам, которые можно использовать, как правило не прибегая к статьям, монографиям, отчетам, патентам и другим материалам. В частности, эта система должна выдавать сведения о методиках диагностики или обследования больных с помощью ЭВМ в виде, пригодном для использования практическими врачами, причем должны выдаваться не только сведения, почерпнутые из разработок и отчетов, но и комментарии и ссылки на другие аналогичные работы, имеющиеся в ИПС. Для планирования и оценок результатов НИР данная ИПС должна выдавать не общие названия проблем или тем, а конкретные формулировки вопросов, составляющих истинное содержание этих НИР. Это продиктовано той особенностью НИР, что в любой сколь угодно крупной проблеме в каждый данный момент движение вперед сдерживается отсутствием решения какого-то конкретного вопроса (например, не удается получить вещество с требуемыми характеристиками, нет решения какого-нибудь уравнения и т. д.).

Под общими формулировками тем НИР может скрываться самое различное содержание. По-видимому, использование таких общих формулировок при планировании и контроле НИР является одной из причин низкой эффективности управления этой сферой человеческой деятельности. Для обеспечения этой возможности в описываемой ИПС должен применяться метод иерархических развертывающихся пунктов справок, при котором в ответе на данный пункт справок должен указываться и перечень вопросов других пунктов, детализирующих данный пункт, и т. д.

Одной из основных особенностей указанной ИПС является автоматическое индексирование (нормализация) текстовой информации (документов и запросов), вводимой в медицинскую информационную систему, что должно обеспечить возможность врачам формулировать свои наблюдения, факты, выводы, назначения на естественном медицинском языке, не занимаясь дополнительной работой по формализации вводимой в ЭВМ информации — кодированием терминов, смысловых отношений, построе-

нием фраз определений структуры и т. д. В то же время в памяти ЭВМ эта информация (после автоматического индексирования) будет находиться в формализованном виде, допускающем ее автоматическую обработку. Такая обработка может включать:

а) поиск и сортировку данных по определенным признакам;

б) статистическую обработку текстовой информации с целью выявления корреляционных зависимостей между различными характеристиками состояния больного, методами диагностики и лечения, последствиями и осложнениями и т. д.;

в) проверку допустимости различных лекарств, их доз; контроль лучевых методов лечения, проверку суммарных и разовых доз при определенных характеристиках состояния больных. Подобный машинный анализ может делаться на основе нормализации текстовой информации о состоянии больного;

г) автоматический анализ состояний и тенденций развития болезни путем сопоставления текстовых описаний состояния больного, анализа нахождения комбинаций характеристик состояния и параметров больного в определенных пределах.

ИПС с развитой справочной частью и автоматическим индексированием (нормализацией) текстовой информации будет наиболее эффективной при ведении машинных историй болезней с использованием машинных методов диагностики, анализа и контроля состояния больного, а также контроля лечебного и диагностического процессов.

В машинных историях болезней некоторые части могут быть структуризованы, а другие части (записи врачей) могут вестись естественным текстом с последующей автоматической нормализацией.

В качестве примеров формулировок проблемных вопросов, по которым должны накапливаться сведения в ИПС верхнего уровня, приведем следующие:

1. Причины возникновения злокачественных новообразований.

2. Природа иммунитета к раку и меры по стимуляции иммунитета.

3. Эффективность лечения рака в разных случаях и разными методами.

4. Условия злокачественного роста клеток и природа рака.

5. Сопутствующие и способствующие заболевания и состояния организма, его отдельных органов и тканей.

6. Диагностирующие признаки для ранней диагностики.

7. Темпы роста и проявления опухолей.

8. Радиобиологические основы лучевой терапии.

Примерами конкретных постановок вопросов для ИПС нижнего уровня могут служить следующие вопросы: выбор дозы и способа облучения в зависимости от локализации опухоли и состояния больного; поиск сведений о конкретных методах лечения с указанием числа случаев, особенностей применения данных методов лечения, условий, сроков, результатов и побочных явлений; поиск обобщенных прецедентов и выдача рекомендаций по диагностике и лечению для конкретных больных; поиск фактических систематизированных, обобщенных данных по обработке историй болезней и регистров с целью оценки методов диагностики и лечения; поиск конкретных случаев болезни и методов лечения и диагностики, представляющих особый интерес; поиск сведений о тематике и результатах научно-исследовательских работ, о научных центрах и ученых, методах, приборах, медикаментах.

Наличие в одной ИПС практических описаний разных диагностических методов, относящихся к одному и тому же виду заболевания (локализации), позволит оперативно сравнивать результаты их применения к одним и тем же больным и будет способствовать как повышению качества диагностики конкретных больных, так и совершенствованию самих методов. Наряду с этим наличие в одной ИПС диагностических подсистем, относящихся к разным локализациям, будет способствовать комплексному использованию этих подсистем.

Применение в ИПС унифицированных методов и алгоритмов автоматического индексирования научных текстов, представленных на русском, немецком, английском (а в дальнейшем и на других) языках, позволит автоматически переводить основное научное содержание огромного количества кратких описаний конкретных случаев диагностики и лечения рака, публикуемых во многих журналах (а в дальнейшем специально собираемых в единый информационный центр), на единый машинный (т. е. формальный) ИПЯ. Представление этих данных (а также обобщенных данных, взятых из статей и отчетов) на ИПЯ обеспечит: а) быстрый поиск, сор-

тировку и статистическую обработку на ЭВМ указанных данных с выдачей выходных результатов на разных естественных языках; б) автоматическую логическую (качественную, т. е. причинно-следственную) обработку текстовой информации по онкологии с выявлением скрытых закономерностей исследуемых процессов и формированием диагностических и терапевтических таблиц, правил и т. д.; в) человеко-машинную творческую обработку накопленной информации с постановкой гипотез и вопросов человеком (ученым) и выдачей ответов (графиков, таблиц, обобщенных данных) ЭВМ с целью выявления закономерностей, касающихся самой природы рака и биологических процессов, лежащих в основе его профилактики и лечения.

Для выполнения перечисленных функций в ИПС должна вводиться следующая информация:

— описания (статьи, диссертации) случаев заболеваний и применявшихся методов диагностики и лечения с оценкой результатов;

— статьи со статистическими данными и анализом случаев заболеваний и методов диагностики и лечения;

— методические указания по диагностике и лечению;

— описания методов диагностики и лечения (сбора анамнеза, анализа признаков, подсчетов численных показателей и сравнения их с заданными критериями).

Информация должна представляться в виде трех уровней.

1. Библиографический уровень, включающий поисковый образ документа и библиографическую справку. Такой уровень обеспечивает только отсылку к литературе.

2. Фактографический уровень, содержащий поисковый образ темы (вопроса) и фактографическую справку. Этот уровень обеспечивает возможность практического использования информации, заключенной в справке.

3. Документальный уровень, включающий микрофильмы с исходными документами. Один и тот же документ может быть представлен на всех трех уровнях, а может быть — на двух или даже на одном уровне.

Для ИПС должен быть сделан рубрикатор по онкологии, в котором деление информации должно осуществляться по трем разрезам: показаниям заболеваний; методам диагностики и лечения (химический, хирургический, радиологический, радиоизотопный и т. д.); видам

информации (описания случаев, методов, методические указания, статистические данные и др.).

Рассмотренные выше задачи и принципы построения автоматизированной информационной системы по онкологии намечают, на наш взгляд, направления и возможности эффективного использования средств и методов кибернетики в решении одной из наиболее острых проблем современной медицины.

Безусловно, что разработка и практическая реализация подобных систем, ориентированных по существу на информационный симбиоз человека и кибернетической машины, потребуют длительного времени и будут осуществляться постепенно.

По-видимому, подобные автоматизированные фактографические информационно-логические системы будут разрабатываться для различных разделов медицины.

Не вызывает сомнения, что создание таких систем приведет к коренному изменению методов и форм использования научной информации, а также к резкому повышению эффективности научной работы и сокращению сроков внедрения научных достижений в практику медицины и здравоохранения.

ГЛОССАРИЙ

Автокод — машинно-ориентированный язык, в котором наиболее часто встречающиеся последовательности машинных команд для облегчения программирования объединены в специальные команды (так называемые макрокоманды).

Алгоритм — четкое предписание, определяющее процесс решения некоторого класса задач.

Алгоритмический язык — искусственный язык, специально приспособленный для записи алгоритмов. Часто под алгоритмическим языком понимают язык программирования.

Алфавитно-цифровое печатающее устройство — внешнее устройство ЭВМ, печатающее выводимую информацию на бумагу в виде букв, цифр и некоторых других символов.

Арифметическое устройство — устройство ЭВМ, в котором выполняется преобразование информации.

Ассемблер — один из машинно-ориентированных языков.

Ассоциативный узел — группа ячеек памяти ЭВМ, содержащая списковые слова всех дескрипторов данного ПОД.

База данных — совокупность машинных массивов информации, а также инструкций, определяющих порядок формирования, корректировки и выдачи этой информации.

Байт — единица хранения информации в ЭВМ, состоящая из 8 двоичных разрядов.

Библиотека программ (библиотека стандартных программ, библиотека подпрограмм, библиотека стандартных подпрограмм) — совокупность программ, решающих типовые, наиболее часто встречающиеся задачи.

Бит — единица количества информации, равная количеству ее в сообщении о наступлении одного из двух равновероятных событий. Битом называется также один

двоичный разряд в машинном слове или единичный импульс в группе импульсов.

Блок записей — группа записей, переписываемых в ОЗУ из внешних накопителей (или обратно) за одну операцию обмена данными.

Величина — элементарная единица информации, описывающая хранимые в информационной базе данные, имеющая определенный смысл и не допускающая расчленения на более малые части.

Внешнее запоминающее устройство (ВЗУ) — запоминающее устройство, позволяющее записывать, хранить и считывать информацию для процессора. Называется также внешним накопителем.

Входной язык — то же, что исходный язык.

Генерация модулей — выбор необходимых модулей, соединение их и настройка на решение нужной задачи (выполняется под управлением операционной системы).

Генерация системы — процесс настройки операционной системы на состав оборудования конкретной ЭВМ.

Граф — сеть линий, соединяющих заданные точки.

Дерево — граф, линии которого нигде не образуют замкнутой цепи.

Дескриптор — слово или словосочетание, выражающее определенное понятие в некотором ИПЯ.

Дисковая операционная система — операционная система; по сравнению с общей операционной системой обладает меньшими возможностями, но зато занимает в ЭВМ меньший объем памяти.

Диспетчер — программа, организующая работу ЭВМ по решению прикладных программ.

Дисплей — устройство для визуального представления на экране телевизионного типа вводимой в ЭВМ информации и выводимой из нее.

Документ — материальный объект (книга, статья, рукопись, история болезни и т. п.), содержащий информацию и предназначенный для передачи и использования этой информации.

Документальная ИПС — система поиска и выдачи адресов документов, отвечающих на поставленный запрос, библиографических описаний или самих документов.

Запись — систематизированная совокупность величин, характеризующих какой-либо объект (или акт).

Идентификатор записей — то же, что и ключ.

Иерархический словарь — словарь дескрипторов, в котором указаны родо-видовые и ассоциативные связи между дескрипторами.

Избирательное распределение информации — форма информационного обслуживания, при которой информационная служба направляет своим абонентам только ту информацию, которая интересует каждого из них (на основе сформулированных ими запросов).

Индекс — специальная запись, служащая для ускорения поиска содержащихся на МД записей.

Индексатор — лицо, выполняющее индексирование.

Индексирование — процесс составления поискового образа документа (или запроса).

Информационная база — то же, что и база данных.

Информационно-поисковый язык — информационный язык, предназначенный для описания и поиска документов.

Информационный шум — ошибки в поиске, которые выражаются выдачей на запрос документов, не соответствующих поставленному запросу.

Информационный язык — искусственный язык, используемый для представления и передачи информации в пределах некоторой кибернетической системы.

Искусственный язык — язык, созданный искусственным путем для достижения определенных практических целей.

Исходная программа — программа, составленная программистом на каком-либо языке программирования. В ЭВМ эта программа преобразуется с помощью транслятора в машинную программу.

Канал — устройство вычислительной машины, служащее исключительно для ввода и вывода информации, минуя устройство управления. Каналы бывают двух видов — мультиплексные и селекторные.

Ключ — группа знаков, служащих для опознавания и отыскания обозначаемых ими записей. Использование ключей ускоряет поиск в большом массиве записей.

Ключевое слово — слово или словосочетание, используемое при индексировании текста (документа или запроса).

Кодовый указатель — ячейка, адрес которой соответствует коду дескриптора. В ячейке содержится адрес связи, отсылающий к буквенному представлению основы дескриптора.

Команда — см. машинная команда.

Кортеж — последовательность некоторых объектов, связанных определенным положением, которое они занимают в данной совокупности объектов.

Критерий выдачи — признак (или совокупность признаков), согласно которому в информационно-поисковой системе устанавливается необходимость выдачи документа на поставленный запрос.

Критерий смыслового соответствия — признак или признаки, которыми пользуется человек при оценке смыслового соответствия между содержанием документа и содержанием поставленного запроса.

Лексема — слово, рассматриваемое в контексте (в совокупности всех его форм и значений).

Магнитная лента — накопительный элемент в виде ленты, покрытый ферромагнитным слоем.

Магнитный барабан — наполнительный элемент в виде вращающегося барабана, внешняя поверхность которого покрыта ферромагнитным слоем.

Магнитный диск — накопительный элемент в виде нескольких вращающихся дисков, плоские поверхности которых покрыты ферромагнитным слоем.

Макрокод — то же, что макрокоманда.

Макрокоманда — система кодирования, которая объединяет несколько машинных команд в одно кодовое слово.

Макроопределение — то же, что и программный модуль.

Маркер — специальная запись, служащая для ускорения поиска данных, записанных на МД. Она содержит значение индекса и указание на то, что данная запись является индексом.

Массив — совокупность однотипных записей, характеризующих однородные объекты.

Математическое обеспечение — совокупность программ, облегчающих работу человека с ЭВМ.

Сюда относятся программы операционной системы, языки программирования с трансляторами и библиотеки программ.

Матрицы — прямоугольная таблица элементов, обычно чисел или букв.

Машинная команда — условный двоичный код, заставляющий вычислительную машину выполнить определенную операцию.

Машинная программа — последовательность команд, которая может быть непосредственно прочитана и выполнена ЭВМ.

Машинно-ориентированный язык — язык программирования, приспособленный для записи программ, рассчитанных на машины определенного типа. Примером машинно-ориентированного языка является автокод.

Мнемокод — простейшая форма машинно-ориентированного языка, использующая вместо цифровых кодов машинных программ буквенные (мнемонические) коды, что несколько облегчает составление программ.

Модуль — то же, что и программный модуль.

Модульный принцип — метод экономного построения сложных объектов (в частности, программ) из отдельных единиц (модулей), допускающих их совместную работу в различных сочетаниях.

Морфема — минимальная значимая часть слова как естественного, так и искусственного языка.

Мультиплексный канал — канал связи процессора с вводными и выводными устройствами. Состоит из большего числа (128—256) подканалов, подключаемых попеременно.

Накопитель — запоминающее устройство ЭВМ.

Нормализация текста — устранение синонимии, полисемии и омонимии текстов на естественном языке. Иногда к нормализации относят и другие операции, связанные с переводом текста с естественного на формализованный язык, например приведение всех существительных к форме единственного числа, именительного падежа, мужского рода, удаление служебных слов и глаголов и т. д.

Омонимия — совпадение звучания слов, имеющих разное значение.

Оперативное запоминающее устройство (ОЗУ) — быстродействующее запоминающее устройство, входящее в состав процессора. ОЗУ, выполненное на магнитных сердечниках, называется магнитным запоминающим устройством (МОЗУ).

Оператор: 1) символ, используемый для обозначения некоторого соответствия или некоторой последовательности действий, которые должны быть выполнены с символами, стоящими за знаком оператора; 2) лицо, обслуживающее вычислительную машину и работающее за ее пультом.

Операционная система — программа, которая управляет работой ЭВМ — координирует ввод данных, решение задач, вывод результатов, управляет очередностью решения разных задач и т. п.

Основание — количественная характеристика величины.

Пакет прикладных программ (пакет пользовательских программ) — совокупность прикладных программ, решающих какую-либо сложную задачу.

Пакетная обработка — режим одновременного решения на ЭВМ нескольких задач, при котором исходные данные по всем задачам находятся во внешних накопителях.

Периферийное оборудование — совокупность периферийных устройств, т. е. устройств, не связанных с переработкой информации.

Периферийное устройство — устройство ЭВМ, не относящееся к процессору, т. е. не связанное непосредственно с процессом обработки информации, — перфораторы, УВВ, ВЗУ, и т. п.

Поисковый образ документа — описание документа, выраженное на информационно-поисковом языке и характеризующее основное содержание документа или какие-либо другие признаки документа, необходимые для последующего поиска этого документа.

Показатель — то же, что и величина.

Полисемия — свойство слов употребляться в разных значениях.

Полнота поиска — одна из характеристик качества поиска, выражающая полноту выявления всех документов, отвечающих поставленному запросу.

Пользовательская программа — то же, что и прикладная программа.

Потери поиска — ошибки при поиске, которые выражаются ненахождением документов, отвечающих на поставленный запрос. Потери поиска — особенно нежелательный вид ошибок.

Предикат (в математической логике) — содержание высказывания о некотором предмете или явлении.

Признак — качественная характеристика величины.

Прикладная программа — программа решения какой-либо научной, производственной или иной задачи.

Проблемно-ориентированный язык — язык программирования, рассчитанный на определенный класс решаемых задач.

емых задач и не зависящий от типа ЭВМ, на которой будет решаться записанная на нем задача.

Программа — см. машинная программа.

Программная совместимость — способность вычислительных машин разных типов выполнять программы, написанные для другой машины.

Программное обеспечение — то же, что и математическое обеспечение.

Программный модуль — программа, решающая некоторую конкретную задачу, допускающая ее использование в различных комбинациях с другими программами (модулями).

Процедурный модуль — программа, выполняющая вспомогательные задачи по организации вычислительного процесса, например включение записей в соответствующие файлы или исключение из них, поиск и выдачу записей по их признакам, и т. п.

Процессор — основная часть ЭВМ, выполняющая вычисления, т. е. ОЗУ, АУ, УУ. Процессор — это ЭВМ без периферийного оборудования.

Пульт управления (ПУ) — устройство ЭВМ, служащее для пуска, остановки машины и контроля ее работы.

Редактор — часть операционной системы, выполняющая объединение в одну программу кусков программ, написанных на разных исходных языках.

Режим разделения времени — режим одновременного решения нескольких задач, что достигается путем последовательного кратковременного решения каждой из задач.

Резидентная программа — программа, постоянно находящаяся в оперативной памяти машины.

Реквизит — качественная или количественная характеристика величины.

Релевантность — соответствие между содержанием документа и содержанием поставленного запроса.

Ретроспективный поиск — поиск документов заданного содержания, давность опубликования которых не превосходит заданной величины.

Свертка дескриптора — условный код, полученный суммированием по определенному закону кодов букв дескриптора.

Световое перо — устройство ввода информации в ЭВМ с помощью дисплея. Представляет собой указатель, позволяющий на экран терминала наносить (т. е. писать)

буквы, цифры или произвольные знаки, а также удалять их.

Селекторный канал — канал, выполняющий связь процессора с быстродействующими вводными и выводными устройствами.

Семантика — наука, изучающая смысл, содержащийся в текстах, написанных на естественных или искусственных языках.

Семантико-синтаксическая категория — синтаксическая и семантическая роль (функция), выполняемая данным дескриптором документа или запроса.

Символический язык — то же, что и язык программирования.

Синонимия — совпадение или близость значений слов с разными звучаниями.

Синтакс (для искусственных языков) — совокупность правил построения сигналов из морфем и из уже построенных сигналов.

Система команд — перечень всех команд вычислительной машины, их кодов и действий, выполняемых при их реализации.

Списковое слово — элемент цепного списка, состоящий из кода дескриптора (или другого объекта, для которого составлен цепной список), кода смысловой связи и адреса связи, т. е. адреса ячейки, где хранится следующее списковое слово.

Страница — физическая величина данных в файлах. Представляет собой блок фиксированной длины, состоящий из записей.

Тезаурус — словарь дескрипторов, дополненный указаниями на разного рода отношения между дескрипторами (родо-видовые, ассоциативные и т. д.).

Теория графов — математическая дисциплина, изучающая свойства геометрических схем, представляющих собой системы линий, соединяющих некоторые заданные точки. Изучение свойств таких сетей позволяет успешно решать задачи из различных областей знания.

Терминал — выносное оконечное устройство ЭВМ, снабженное дисплеем и клавиатурой для ввода в ЭВМ исходных данных для решения задачи и для получения результатов решения. Кроме клавиатуры, терминал может иметь световое перо.

Том данных — наибольшая единица записей, состоящая из нескольких массивов. Представляет собой стан-

дартный блок внешней записи (катушка магнитной ленты, пакет дисков и т. п.).

Точность поиска — одна из характеристик качества поиска, выражающая степень присутствия нерелевантных документов среди всех документов, выданных на запрос.

Транслятор — программа, которая преобразует программу, записанную на алгоритмическом языке, в машинную программу.

Трансляция — преобразование программы, записанной на алгоритмическом языке, в машинную программу.

Указатель роли — признак, обозначающий логическую роль данного дескриптора в поисковом образе документа.

Указатель связи — признак, указывающий на наличие некоторого рода связи между дескрипторами поискового образа документа, помеченными этим признаком.

Устройство ввода и вывода данных (УВВ) — периферийное устройство, позволяющее вводить в ЭВМ информацию, записанную на перфокартах, перфолентах или магнитных лентах, или выводить из ЭВМ информацию на печатающие устройства, перфораторы или на экран дисплея.

Устройство управления (УУ) — устройство ЭВМ, управляющее решением задач.

Фактографическая ИПС — система поиска и выдачи каких-либо сведений, данных, фактов.

Файл — набор данных одинаковой структуры, объединенных общим назначением. Файлы имеют названия, соответствующие содержанию данных, из которых состоит файл.

Формализованный язык — искусственный язык, построенный по формальным правилам.

Цепной список — способ записи в памяти ЭВМ сведений о группе (последовательности) некоторых объектов (например, о документах, в ПОД которых содержится определенный дескриптор), при котором соседние элементы цепного списка (списковые слова) могут располагаться в произвольном месте памяти.

Экстракод — то же, что и макрокоманда.

Язык программирования — искусственный язык, на котором программист составляет исходную программу для ЭВМ.

СПИСОК СОКРАЩЕНИЙ

- АЛ — адрес лексемы
 АПД — аппаратура передачи данных
 АСВТ — агрегатированные средства вычислительной техники
 АСММО — автоматизированная система для массового медицинского обследования
 АСУ — автоматизированная система управления
 АСУП — автоматизированная система управления предприятием
 АУ — арифметическое устройство
 АЦПУ — алфавитно - цифровое печатающее устройство
 БАИС — больничная автоматизированная информационная система
 ВЗУ — внешнее запоминающее устройство
 ВЦ — вычислительный центр
 ГВЦ — главный вычислительный центр
 ГСУ — глубинно-синтаксический уровень
 ГУ — графический уровень
 ДНБ — диспансерное наблюдение за беременной
 ДОС — дисковая операционная система
 ЕИН — единый инвентарный номер
 ЕС ЭВМ — единая серия ЭВМ
 ЗУ — запоминающее устройство
 ИБМ — крупнейшая американская фирма, производящая ЭВМ
 ИПС — информационно-поисковая система
 ИПЯ — информационно-поисковый язык
 ИРИ — избирательное распределение информации
 КВ — критерий выдачи
 КД — код дескриптора
 КСС — критерий смыслового соответствия
 МБ — магнитный барабан
 МД — магнитный диск
 МЛ — магнитная лента
 МИС — медицинская информационная система
 НЗ — номер зоны
 НМБ — накопитель на магнитном барабане
 НМД — накопитель на магнитном иске
 НМЛ — накопитель на магнитной ленте
 НСИ — нормативно-справочная информация
 ОАСУ — отраслевая автоматизированная система управления
 ОЗУ — оперативное запоминающее устройство
 ОС — операционная система
 ПМ — поисковый массив
 ПОД — поисковый образ документа
 ПОЗ — поисковый образ запроса
 ПУ — пульт управления
 САИ — система автоматического индексирования
 СМ — справочный массив
 СПМ — счетно-перворационная машина
 СПУ — сетевое планирование и управление
 ССК — семантико-синтаксическая категория
 СТА-2М — стартовый телеграфный аппарат
 Т-63 — телеграфный аппарат
 УВВ — устройства ввода—вывода
 УР — указатель роли
 УС — указатель связи
 УУ — устройство управления
 УПФ — унифицированная программа формирования и печати таблиц
 ФАК — формализованные амбулаторные карты
 ФАКРР — формализованная амбулаторная карта развития ребенка
 ФИБ — формализованная история болезни

ЛИТЕРАТУРА

- Брежнев Л. И.* Отчет Центрального Комитета КПСС и очередные задачи партии в области внутренней и внешней политики. Доклад XXV съезду КПСС 24 февраля 1976 г. М., Политиздат, 1976.
- Основные* направления развития народного хозяйства СССР на 1976—1980 гг. М., Политиздат, 1976.
- Автоматизированная* система управления (теория и методология). Том 1 и 2. М., «Мысль», 1972.
- Башин М. Л.* Планирование научно-исследовательских и опытно-конструкторских работ. М., «Экономика», 1969.
- Белоногов Г. Г., Богатырев В. И.* Автоматизированные информационные системы. М., «Советское радио», 1973.
- Бусленко Н. П.* Лекции по теории сложных систем. М., «Советское радио», 1973.
- Валентинов В. В.* Автоматизация поиска информации в АСУ. М., «Статистика», 1976.
- Венедиктов Д. Д.* Основные направления и этапы создания отраслевой автоматизированной системы управления здравоохранением (ОАСУ). — «Сов. здравоохран.», 1973, № 2, с. 3—8.
- Вентцель Е. С.* Введение в исследование операций. М., «Советское радио», 1964.
- Вишневский А. А., Артоболевский И. И., Быховский И. Л.* Машинная диагностика и информационный поиск в медицине. М., «Наука», 1969.
- Воробьев Е. И., Китов А. И.* Автоматизация обработки информации и управления в здравоохранении. М., «Советское радио», 1976.
- Глушков В. М.* Введение в АСУ. Киев, «Наукова Думка», 1972.
- Глушков В. М., Добров Г. М., Терещенко В. И.* Беседы об управлении. М., «Наука», 1974.
- Глушков В. М.* Кибернетика. — В кн.: Энциклопедия кибернетики. Т. 1. Киев, 1975, с. 440—445.
- Гордон Б.* Информационные системы в медицине. Сборник переводов. Под ред. Шнепс-Шнеппе. М., 1974.
- Китов А. И.* Программирование информационно-логических задач. М., «Советское радио», 1967.
- Китов А. И.* Программирование экономических и управленческих задач. М., «Советское радио», 1971.
- Консон А. С., Соминский В. С.* Экономика научных разработок. М., «Экономика», 1968.
- Медицинская* информационная система. Изд. 2-е. Киев, «Наукова думка», 1975.

- Петровский Б. В. Здоровье народа — важнейшее достижение социалистического общества. М., «Медицина», 1971.
- Попов Г. Х. Проблемы теории управления. М., «Экономика», 1970.
- Прогресс биологической и медицинской кибернетики. М., «Медицина», 1974.
- Эффективность научно-технического творчества. М., «Наука», 1968.
- Ямпольский С. М., Лисичкин В. А. Прогнозирование научно-технического прогресса. Методологические аспекты. М., «Экономика», 1974.
- (Clinical electrocardiography and computers). Вычислительные системы и автоматическая диагностика заболеваний сердца. Пер. с англ. М., «Мир», 1974.
- (Salton G.). Сэлтон Г. Автоматическая обработка, хранение и поиск информации. Пер. с англ. М., «Советское радио», 1973.
- (Wiener N.) Винер Н. Кибернетика, или управление, и связь в животном и машине. Пер. с англ. М., «Советское радио», 1958.

ОГЛАВЛЕНИЕ

	Стр.
Предисловие	3
Глава I. Основные сведения из кибернетики	6
1. Возникновение кибернетики	6
2. Информация и сигнал	9
3. Преобразование информации и алгоритмы	21
4. Кибернетические системы	29
Глава II. Принципы устройства ЭВМ и элементы программирования	37
1. Устройство ЭВМ	37
2. Элементы программирования	52
3. Особенности программирования для ЭВМ третьего поколения	60
Глава III. Понятие о математических методах, используемых в медицинской кибернетике	78
1. Системный анализ и исследование операций	78
2. Математические методы нахождения оптимальных решений	82
3. Методы моделирования и сетевого планирования и управления	92
4. Элементы прикладной математической логики	101
Глава IV. Автоматизация поиска информации	109
1. Назначение и состав информационно-поисковой системы (ИПС)	109
2. Документально-фактографическая ИПС, использующая формализованный естественный язык назывных предложений	114
3. Запись ПОД с использованием семантико-синтаксических категорий и линейных кодов	137
4. Информационно-поисковая система, работающая на естественном языке	141
5. Методы и возможности автоматического индексирования документов и запросов	153
6. Медицинская лингвистика	168
Глава V. Основные направления применения ЭВМ в медицине	173
1. Автоматизация сбора и обработки данных о состоянии больных и машинное ведение истории болезни	174
2. Автоматизированная система накопления и обработки данных о выбывших больных в госпитале	185
3. Диспетчеризация и учет медицинского обслуживания больных с помощью ЭВМ	193
4. Автоматизация сбора и обработки информации о состоянии больного во время операции и в период интенсивной терапии	200
5. Применение ЭВМ в диагностике заболеваний	205

Глава VI. Автоматизация процессов управления в здравоохранении	216
1. Назначение и структура АСУ «Здравоохранение»	216
2. Задача сбора и обработки данных о лечебной деятельности больничных стационаров	221
3. Комплекс задач подсистемы управления эпидемическим обслуживанием населения	226
4. Оптимизация решения управленческих задач	232
Глава VII. Кибернетика и медицинская наука	237
1. Вопросы эффективности научной работы	240
2. О путях совершенствования структуры научно-исследовательских учреждений	250
3. Кибернетические методы научного прогнозирования в медицине и здравоохранении	254
4. Системный подход к накоплению и обработке медицинской информации	263
Глоссарий	274
Список сокращений	283
Литература	284