

## ЛИТЕРАТУРА

1. Касперович А. Н., Шалагинов Ю. В. Некоторые вопросы проектирования АЦП с использованием амплитудной свертки сигнала. — Автометрия, 1978, № 4.
2. Полубабкин Ю. В., Прозоров Ю. П., Ломтев Е. А., Шляндин В. М. Быстродействующий АЦП. (Автор. свид-во № 750722). — БИ, 1980, № 27.
3. Смолов В. Б. Микроэлектронные цифроаналоговые и аналого-цифровые преобразователи. М.: Энергия, 1976.

Поступила в редакцию 26 февраля 1980 г.

УДК 62-50 : 519.14

К. Р. ВИКНА, З. П. МАРКОВИЧ

(Рига)

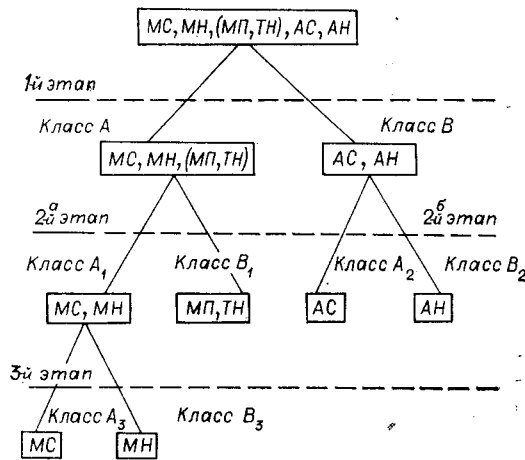
### СИСТЕМЫ РЕАЛЬНОГО ВРЕМЕНИ ДЛЯ РЕШЕНИЯ КЛАССИФИКАЦИОННЫХ ЗАДАЧ В КЛИНИКЕ

К настоящему времени как в технической, так и в медицинской сфере имеется значительное количество алгоритмов диагностирования, позволяющих ставить правильные диагнозы с вероятностью 0,92—0,95. Однако количество лечебно-профилактических учреждений, где вычислительная техника обслуживает область медицинской диагностики, весьма ограничено. Трудности практического использования ЭВМ в этом направлении вызваны, на наш взгляд, следующим. С одной стороны, для решения диагностических задач необходим диалоговый режим «врач — ЭВМ» без привлечения операторов и алгоритмических языков, наличие быстродействующих ЭВМ с достаточно большой памятью, возможность доступа к ЭВМ в произвольный момент времени. С другой стороны, связь с вычислительными центрами через терминальные устройства теоретически хотя и возможна, но практически реализована лишь в отдельных случаях. Кроме того, учреждения здравоохранения, как правило, не имеют мощных вычислительных средств, а обладают мини-ЭВМ типа «Искра-125», «Электроника-100», «Wang» и др.

Преодоление имеющихся трудностей возможно при комбинированном использовании больших и малых ЭВМ: подготовка и частичное решение задачи осуществляются на больших ЭВМ в ВЦ; введение конкретных входных данных и получение результатов распознавания — на мини-ЭВМ, работающей в режиме реального времени (на базе медицинского учреждения), с последующим возвращением опыта решения конкретных задач в универсальную ЭВМ. Контакт с мини-ЭВМ выполняется непосредственно медицинским персоналом без использования алгоритмических языков. При этом возможно как автоматическое, так и диалоговое решение задачи.

Разработка двух систем названного типа (система дифференциальной диагностики пяти вариантов приобретенных пороков сердца «Диагностика ППС» и система скрининга основных сердечно-сосудистых болезней ССБ-АВТО) практически реализована в клинике Латвийского НИИ кардиологии. В обеих системах в качестве мини-ЭВМ применен отечественный вычислительный комплекс «Искра-125».

**Диагностика ППС.** Система предназначена для распознавания следующих заболеваний: митрального стеноза (МС), митральной недостаточности (МН), митрального порока в сочетании с трикуспидальной недостаточностью (МП, ТН), аортального стеноза (АС), аортальной недостаточности (АН). Многоклассная задача решается поэтапно дихотомической процедурой распознавания образов с формированием искусственных



(объединенных) классов ППС на каждом этапе (кроме последнего) на основе дерева решений (см. рисунок). Диагностирование производится путем движения по дереву решений от начала ствола до конца какой-либо ветви, т. е. одинарного класса.

Первоначальное описание классов представлено набором 42 бинарных признаков (наличие или отсутствие симптомов заболеваний). На каждом этапе распознавания используется различное сочетание не более 25 признаков, конкретный со-

став которых определен на основе расчета количества информации по Шеннону.

В данном случае в качестве алгоритма распознавания заболевания выбран алгоритм логического анализа дискретной информации\*, не требующий столь большого объема выборки, как статистический алгоритм.

Алгоритм основывается на вычислении так называемых различающих весов признаков  $R(x_i)$  на основе обучающей выборки:

$$R(x_i) = k_i/k, \quad (1)$$

где  $k$  — число тупиковых тесторов таблицы  $T$  обучающей выборки (столбцы  $T$  — симптомы, строки — реализации);  $k_i$  — число тупиковых тесторов, в которые входит признак  $x_i$ ,  $i = 1, \dots, m$ .

Тестором является набор столбцов таблицы  $T$  и соответствующий ему набор признаков  $(x_1, \dots, x_r)$ , если после удаления из  $T$  всех столбцов, не вошедших в  $(x_1, \dots, x_r)$ , все строки подтаблицы  $T_A$  класса  $A$  будут различны со всеми строками подтаблицы  $T_B$  класса  $B$ . Тестор для  $T_A$  и  $T_B$  окажется тупиковым в том случае, если после удаления из него любого столбца он перестанет быть тестором для  $T_A$  и  $T_B$ .

Решение принимается по трем правилам. Первое правило:

$$\rho_A = \frac{1}{t} \sum_{i=1}^m \sum_{j=1}^t R(x_i) (\alpha_{ij} \oplus \gamma_i),$$

$$\rho_B = \frac{1}{l} \sum_{i=1}^m \sum_{j=1}^l R(x_i) (\beta_{ij} \oplus \gamma_i). \quad (2)$$

Здесь  $\alpha_{ij}$  —  $i$ -й элемент  $j$ -й реализации в обучающей выборке класса  $A$ ,  $j = 1, \dots, t$ ; знак  $\oplus$  — суммирование по модулю 2;  $\beta_{ij}$  — то же, в выборке класса  $B$ ,  $j = 1, \dots, l$ ;  $S_x = \{\gamma_1, \dots, \gamma_m\}$  — бинарный вектор распознаваемой реализации. Если  $\rho_A < \rho_B$ , то  $S_x$  относится к классу  $A$ ; если  $\rho_A > \rho_B$  — к классу  $B$ ; если  $\rho_A = \rho_B$ , то распознавание не производится.

Второе правило:

$$(\rho_{\min})_A = \min_j \sum_{i=1}^m R(x_i) (\alpha_{ij} \oplus \gamma_i),$$

$$(\rho_{\min})_B = \min_j \sum_{i=1}^m R(x_i) (\beta_{ij} \oplus \gamma_i). \quad (3)$$

\* Дмитриев А. Н., Журавлев Ю. И., Кренделев Ф. П. Об одном принципе классификации и прогноза геологических объектов и явлений. — Геология и геофизика, 1968, № 5.

Таблица 1

Оператор (врач)	ЭВМ	На дисплее
<p>Вызывает программу</p> <p>Вводит данные больного с клавиатуры</p> <p>Просматривает результаты и а) соглашается, б) не соглашается и сообщает свое решение</p>	<p>Сигнализирует о готовности к работе</p> <p>Распознает ситуацию по правилам (1)—(3) на первом этапе. Принимает решение</p> <p>В случае «а» выбирает по дереву решений необходимое направление и распознает ситуацию 2-го этапа. В случае «б» выбирает по дереву решений указанное врачом направление и распознает ситуацию 2-го этапа</p>	<p>Прошу ввести</p> <p>По (1)—класс А, по (2)—класс В, по (3)—класс А</p> <p>И т о г о: класс А Ваше решение?</p> <p>По (1)—класс А<sub>1</sub>, по (2)—класс А<sub>1</sub>,</p> <p>И т о г о: класс А<sub>1</sub> Ваше решение?</p>

Если  $(\rho_{\min})_A < (\rho_{\min})_B$ , то  $S_x$  относится к классу А; если  $(\rho_{\min})_A > (\rho_{\min})_B$  — к классу В; если  $(\rho_{\min})_A = (\rho_{\min})_B$ , то распознавание не производится.

Третье правило:

$$N_A = \rho_A(\rho_{\min})_A, \quad (4)$$

$$N_B = \rho_B(\rho_{\min})_B.$$

Если  $N_A < N_B$ , то  $S_x$  относится к классу А; если  $N_A > N_B$  — к классу В; если  $N_A = N_B$ , распознавание не производится.

В применяемом алгоритме наиболее трудоемкими оказываются нахождение тупиковых тесторов и расчеты различающих весов  $R(x_i)$  признаков по выражению (1), требующие длительной (до 12 ч) работы универсальной ЭВМ. Следовательно, классификация распознаваемых реализаций в диалоговом режиме с непосредственным вычислением  $R(x_i)$  невозможна даже на очень быстродействующих ЭВМ. Данный алгоритм обладает тем преимуществом, что расчеты  $R(x_i)$  могут быть проведены заранее, а распознавание реализаций с использованием правил (2)—(4) на заключительном этапе осуществимо в диалоговом режиме на мини-ЭВМ. Перерасчет различающих весов выполняется универсальной ЭВМ на основе накопленной статистики и опыта.

Система «Диагностика ППС» работает в автоматическом и диалоговом режимах по выбору. При работе в автоматическом режиме после ввода основной программы и симптоматики по 42 признакам (с клавиатуры) мини-ЭВМ осуществляет распознавание неизвестной реализации на каждом этапе, двигаясь по дереву решений автоматически. Результаты каждого этапа решения, рассчитанные величины  $\rho$ ,  $\rho_{\min}$ ,  $N$ , а также окончательный результат выводятся на дисплей и печатающее устройство.

Решение принимается голосованием правил. В случае невозможности распознавания ввиду тождественности значений  $\rho$ ,  $\rho_{\min}$  или  $N$  осуществляются переход в диалоговый режим и высвечивание на дисплее вопроса об указании направления дальнейшего продвижения (табл. 1).

Автоматический режим распознавания дает так называемый средневзвешенный результат на базе обучающей выборки.

**ССБ-АВТО.** Система скрининга используется при проведении профилактических обследований населения для раннего выявления пяти синдромов: ишемической болезни сердца (ИБС), артериальной гипертензии (АГ), поражения периферических сосудов (ППС), дисциркуляторной энцефалопатии (ДЭ), нейроциркулярной дистонии (НЦД).

Таблица 2

Номер вопроса	Высвечивается на экране
H01	Бывают ли у Вас по ночам приступы одышки (нехватки воздуха)?
H02	Бывает ли у Вас одышка при ходьбе, при умеренной работе?
...	...
...	...
H42	Болеют (болели) ли многие Ваши близкие родственники сердечно-сосудистыми заболеваниями?
<p>Спасибо. Теперь обратитесь к присутствующему медицинскому работнику. Он пояснит Вам, что рекомендует эта вычислительная машина.</p>	

В качестве метода сбора данных в системе использован принцип автоинтервьюирования (диалог «обследуемый — ЭВМ») на основе субъективных данных в результате ответов «да — нет» на 42 наиболее информативных двоичных признака, сформулированных в виде вопросов (табл. 2).

На подготовительном этапе определены диагностические веса признаков и пороговые значения суммы весов относительно каждого заболевания.

Система имеет два режима: автоматический и диалоговый по принципу «пациент — ЭВМ» на основе адаптивного алгоритма, имитирующего врачебную логику. При работе в автоматическом режиме обследуемый вводит ответы на все 42 вопроса. Превышение максимального порогового значения ведет к установлению достоверного диагноза, превышение минимального — вероятного (заболевание не исключено). Количество и последовательность задаваемых вопросов при работе во втором режиме определяются содержанием ответов на предыдущие.

**Результаты.** Апробация систем в клинике и базовой поликлинике ЛНИИГ подтверждает возможность и эффективность их применения в медицинской практике.

В обеих системах время распознавания одной реализации определяется лишь скоростью ввода симптоматики с клавиатуры ЭВМ. Перезапись программ с магнитной ленты (МЛ) в ОЗУ осуществляется лишь в начале работы и не превышает 6 мин.

На экзамене системы «Диагностика ППС» по 250 верифицированным реализациям на 100 диагностируемых случаях обнаружилось в среднем 8 ошибок.

Качество работы системы ССБ-АВТО проверено на материале 250 верифицированных случаев сердечно-сосудистых заболеваний. В группе наиболее серьезных заболеваний — ИБС и АГ — не обнаружено ни одного случая пропуска цели (гиподиагностики). В остальных случаях выявлено 5% больных, диагноз заболевания которых не определен ЭВМ ввиду отсутствия субъективных признаков начавшейся болезни. В группе больных НЦД наблюдалась гипердиагностика более серьезного заболевания в 10% случаев, что клинически полностью оправдано, так как у части больных НЦД при повторных наблюдениях в течение ряда лет развивается органическая патология сердечно-сосудистой системы. В других группах гипердиагностика не превышала 3%.

#### ВЫВОДЫ

В настоящее время осуществление режима реального времени при решении задач медицинской диагностики наиболее рационально на базе

проблемного комплекса, состоящего как из мини-ЭВМ, так и из больших ЭВМ. Наличие соответствующего функционального обеспечения позволяет создавать небольшие вычислительные системы диагностики с доступными для лечебно-профилактических учреждений режимами эксплуатации.

*Поступила в редакцию 7 августа 1979 г.;  
окончательный вариант — 20 ноября 1980 г.*

УДК 578.088.5

**А. Л. КОЗИОНОВ, С. Ю. НОВОЖИЛОВ, В. Е. СОЛОБОЕВ,  
М. И. ШТОКМАН**  
(Новосибирск)

### **СВЕТОИНДУЦИРОВАННАЯ ДИФФУЗИЯ ДНК: ТЕОРИЯ И АВТОМАТИЗИРОВАННЫЙ ЭКСПЕРИМЕНТ**

**1. Введение.** Изучение взаимодействия мощного лазерного света с ДНК весьма важно как с фундаментальной, так и с прикладной точки зрения. Особый интерес представляет метод селективного лазерного воздействия на заданный тип макромолекул или на заданные участки макромолекулы, предложенный в [1]. Этот метод основывается на возможности селективной окраски макромолекул (в частности, ДНК) с помощью специфических красителей с последующим облучением мощным лазерным светом видимого или ближнего ультрафиолетового (УФ) диапазона. Такой свет ДНК непосредственно не поглощается, а красителем, фиксированным на макрополе, поглощается квазирезонансно. Поглощение молекулой красителя двух квантов света приводит к ее возбуждению в высшее синглетное состояние с энергией 6—7 эВ. Кванты такой энергии интенсивно поглощаются ДНК. Поэтому со значительной вероятностью происходит безрадиационный перенос возбуждения с молекулы красителя на пространственно близкие к ней участки ДНК. Оценки [1] показывают, что первичный акт возбуждения ДНК происходит в окрестности нескольких ангстремов вокруг молекулы красителя. Переданной ДНК энергии достаточно для эффективного разрыва любой химической связи.

Осуществление описанного процесса в эксперименте [2, 3] привело к обнаружению нового явления — нелинейного лазерного фоторазрезания (НЛР) ДНК. При облучении ДНК, окрашенной специфическим (для ДНК) красителем 8-метоксипсораленом (8-МОП), с высоким квантовым выходом наблюдаются разрывы ДНК на фрагменты. Этот процесс оказался существенно нелинейным по интенсивности возбуждающего света в соответствии с теоретическими предсказаниями [1, 3].

Обнаружение НРЛ сразу же выявило ряд проблем экспериментального плана. Это явление было идентифицировано в [2, 3] с помощью гель-электрофореза, который как инструмент исследования НЛР обладает существенными недостатками. Во-первых, гель-электрофорез — медленный метод: время, необходимое для измерения, — от нескольких часов до суток. Поэтому он непригоден для детального исследования НРЛ, когда необходимо провести большое число экспериментов. Во-вторых, гель-электрофорез является возмущающим методом. По этой причине эксперименты [2, 3] не могли дать ответа на фундаментальный вопрос, который