

## ТЕОРИЯ ПОДСОЗНАТЕЛЬНЫХ РЕШАЮЩИХ ПРАВИЛ И ЕЕ ПРИМЕНЕНИЕ В ДИАГНОСТИЧЕСКИХ ЗАДАЧАХ\*

© 2003 г. О. И. Ларичев

Академик РАН, Институт системного анализа РАН

Анализируется феномен экспертного знания, излагаются экспериментальные факты, позволяющие найти объяснение процессу распознавания “целостных” ситуаций в рамках теории подсознательных решающих правил (ПРП). На основании этого диагностика представляется как задача классификации, решая которую эксперт относит объект к одному из нескольких классов решений. Имитируя экспертные решения, эта задача может быть представлена несколькими решающими правилами, почти однозначно восстанавливаемыми из граничных объектов. Количество решающих правил не превышает объема кратковременной памяти, что объясняет быстроту их использования. Обосновываются выводы о “ситуативности” экспертного мышления и о том, что механизм распознавания объектов экспертом может быть правдоподобно описан на основе символической теории. ПРП можно рассматривать как специальные символы, имеющие подсознательный характер.

*Ключевые слова:* мышление, символический подход, медицинская диагностика, сохранение знаний, клинические ситуации, экспертные знания.

### 1. ЧЕЛОВЕЧЕСКОЕ МЫШЛЕНИЕ КАК МАНИПУЛЯЦИЯ СИМВОЛАМИ

Построение систем искусственного интеллекта тесно связано с исследованиями человеческой системы переработки информации. Это вполне объяснимо, так как основная цель исследований в этой области состоит в создании искусственных систем, равных или даже превосходящих человека по своим интеллектуальным возможностям.

Наиболее известным в когнитивной психологии подходом к анализу человеческого мышления является теория символической переработки информации. Согласно этому подходу, мышление – это преобразование символов [20]. Всякую задачу можно представить как совокупность этапов переработки информации, на каждом из которых человек выделяет, сравнивает, преобразует символы, адекватные содержанию задачи. Так, в задаче о ханойской башне [7] каждый символ может представлять либо диск, либо перемещение диска или совокупности дисков. Согласно Г. Саймону с соавторами, система физических символов состоит из множества объектов, которые могут быть также компонентами другого объекта, называемого символической структурой [21]. При решении математических задач каждый из символов – это либо элемент описания объектов, либо

какое-то преобразование элементов, совершаемое с помощью тех или иных правил или теорем.

Теория символической переработки информации подтверждена экспериментами Саймона и его коллег по сравнению поведения человека с моделью при решении таких задач, как запоминание и узнавание слогов [20], нахождение математических зависимостей по ряду чисел и т.д. Модели ЕРАМ, ВАСОН и др. получили мировую известность.

Применительно к задачам обучения символический подход является основой АСТ-теории Дж. Андерсона [10]. Если мышление – преобразование символов, то опытный человек осуществляет эти преобразования быстрее новичка. Согласно подходу “компиляции знаний” Д. Андерсона, новичка следует научить использовать одновременно несколько символов. Например, при решении геометрических задач студент учится осуществлять несколько преобразований одновременно.

Символический подход явно или неявно подразумевает, что сложный объект или паттерн всегда разложим на отдельные составляющие. “Важно лишь, чтобы система могла оперировать паттернами всех видов – входными (считываемыми) и сохраняемыми в памяти, создавать сложные паттерны из более простых компонентов, выводить (записывать) паттерны, сопоставлять их с тем, чтобы установить степень сходства или различия” [9].

\*Работа частично поддержана РФФИ (гранты 01-01-00514 и 0015-96053) и РАН (проект 2-1 программы “Интеллектуальные компьютерные системы”).

По мнению авторов символического подхода, он достаточно универсален и может быть распространен на задачи любого типа. И в самом деле, этот подход нашел свое выражение в столь различных методах, как построение продукционных экспертных систем, машинное обучение, распознавание образов, нейронные сети [22]. Во всех этих подходах можно увидеть разложение целого на части и создание целого как объединение этих частей.

Известный феномен быстрого распознавания человеком различных образов (лиц, ситуаций и т.д.) символический подход объясняет двумя способами.

При нейросетевом подходе предполагается, что человеческая система распознавания работает параллельно, сравнивая образ со многими другими и выделяя наиболее близкое сходство. При этом сходство определяется как функция от составляющих, частей образа.

Однако Саймон полагает, что человеческая система переработки информации работает последовательно, а не параллельно [8], и эту точку зрения подтверждают многочисленные эксперименты.

Другая модель предполагает мгновенное распознавание объектов с помощью интуитивного суждения [9]. Саймон считает, что интуиция сводится к распознаванию и может быть объяснена с помощью модели, которая “реализуется на дискриминирующей сети, представляющей собой вариант сортирующей системы”. При этом сортировка осуществляется по отдельным составляющим образа. Мы опять видим здесь гипотезу о разложимости символа на его составляющие.

Анализируя различные объяснения феномена распознавания, можно сделать вывод, что “разложимость” целого (образа, паттерна, символа) на части является неотъемлемым элементом символического подхода.

Однако, как известно, существуют целостные объекты, разложение которых на части невозможно без существенных искажений. Примером распознавания такого объекта является принятие медицинских диагностических решений. Относительная “сила” диагностических признаков и их значений (симптомов) зависит от конкретного сочетания симптомов и не может быть определена вне зависимости от общей ситуации. Следовательно, может быть другое представление об экспертном знании: как о чем-то, возникающем только при целостном рассмотрении объектов.

Цель данной статьи состоит в анализе известных объяснений феномена экспертного знания, а также в изложении экспериментальных фактов, позволяющих найти новое теоретическое объяс-

нение феномену распознавания “целостных” ситуаций.

## 2. ПРАКТИЧЕСКИЙ ПРИМЕР – МЕДИЦИНСКАЯ ДИАГНОСТИКА

Различие двух представлений о человеческом мышлении можно показать на примере построения экспертных диагностических медицинских систем.

Как известно, одной из первых таких систем была система МАЙСИН [12], разработанная В. Кланси, а первой обучающей системой, основанной на экспертных знаниях, была НЕОМАЙСИН [25]. В системе МАЙСИН экспертное знание представлено в виде продукции типа:

ЕСЛИ (условие),... ТО (вывод).

В качестве символов рассматривались значения диагностических признаков. Важным элементом системы был блок логического вывода. При предъявлении нового объекта этот блок осуществлял поиск решения, комбинируя различные части известных продукций – символы, рассматриваемые как независимые. Следовательно, используется предположение, что можно выделить так называемые примитивы, и потом составить из них любые объекты. Отметим, что в системе НЕОМАЙСИН объяснения для студентов давались в виде продукций, т.е. предполагалось, что экспертное знание может быть “составлено” из частей продукций.

Таким образом, система была основана на символическом подходе. Приведенный пример показателен тем, что автор системы В. Кланси в 1990-е годы отказался от своей системы, признав как единственно правильный подход ситуативного мышления [13].

На наш взгляд, задачи медицинской диагностики являются по своей природе “ситуативными” в том смысле, что целостный образ ситуации, возникающий у эксперта, определяет диагноз. Информативность значений различных диагностических признаков зависит от их конкретных сочетаний. Заметим, что слово “ситуативный” используется в когнитивной психологии в том смысле, что знание порождается конкретной ситуацией и не может быть абстрагировано от нее [16].

При развитии экспертных систем в принципе мог бы быть использован иной подход. Рассмотрение в качестве символов целостных образов ситуаций никак не противоречит символическому подходу. Однако легко понять, почему развитие экспертных систем не пошло по этому пути. Сбор и анализ целостных образов ситуаций представляет собой трудную и долговременную задачу. В то же время подход, позволяющий судить о целом по его частям, гораздо легче реализовать. Вспомним, например, систему INTERNIST, охватывающую 500 заболеваний и 6000 диагностических признаков [24].

Итак, имеются два возможных представления механизма мышления врача-эксперта. Эти объяснения противоречивы. Необходимы дополнительные эксперименты, позволяющие выбрать наиболее правдоподобное из них. Но предварительно необходимо провести классификацию задач.

### 3. РАЗНЫЕ ТИПЫ ЗАДАЧ

Неоднократно отмечалось, что есть классы задач, для которых полностью справедлив символичный подход. Даже его противники [14] согласны в том, что задачи математического типа: головоломки, доказательства теорем, задачи в алгебре и геометрии – имеют вполне адекватное представление в виде независимых символов. Так, упомянутая выше задача о ханойской башне допускает вполне адекватное описание в терминах пространства состояний и совокупности преобразований символов.

Для такого класса задач Саймон и коллеги [8, 9, 20] предложили совокупность эвристических и мощных по возможностям программ, таких, как “общий решатель задач”, “логик-теоретик” и другие.

Если целостный, ситуационный характер задач противоречит “декомпозиционному” символьному подходу, то символичный подход может рассматриваться как средство нахождения аппроксимации неизвестного поведения эксперта при целостном рассмотрении им ситуации. Иначе говоря, использование гипотезы о независимости и аддитивности символов в задачах, где роль символов зависит от ситуации, дает возможность найти какое-то приближенное решение проблемы.

Насколько точна эта аппроксимация? Можно ли выделить классы задач, где она вполне удовлетворительна и где она не работает? Рассмотрим эти вопросы подробнее.

### 4. ЧЕЛОВЕЧЕСКОЕ И МАШИННОЕ МЫШЛЕНИЕ

Есть одна область искусственного интеллекта, где определенное сопоставление символического и целостного (ситуативного) подхода осуществлялось на протяжении ряда лет, – это шахматы. Саймон называл игру в шахматы “дрозофилой” искусственного интеллекта [28].

Задача выбора мастером очередного шага в шахматной партии является, безусловно, ситуативной. Роль признаков, используемых хорошим шахматистом, таких, как вероятность атаки на короля, открытая линия, сила фигур, зависит от ситуации. Шахматист не может просматривать позиции на много шагов вперед, он использует целостные образы ситуации.

Совсем не так играет компьютер. Кроме огромного запаса прошлых партий и некоторых правил игры (в окончаниях), у него имеются лишь

формальные способы оценки ситуаций, использующие заранее заданные “веса” фигур. Зато компьютер может осуществлять далекий (на много полуходов) просчет за себя и за противника.

Мы видим здесь два абсолютно разных пути мышления – человеческий (ситуационный) и машинный (символьный). Результаты сравнения известны: машинные программы играют на уровне гроссмейстера, и одна из них (Deer Blue) даже победила Г. Каспарова.

Но посмотрим теперь на поле сравнения. Шахматы – достаточно “узкая” игра с четко заданными правилами. Отмечалось, что применение подходов просмотра и формальной оценки в игре “Го” занимает огромное машинное время (несколько лет на один ход при любом компьютере) [19]. К тому же любые игры существенно отличаются от реальных задач, где условия определены нечетко, а объект исследования достаточно сложен. К таким задачам относится и медицинская диагностика.

### 5. ЧЕРТЫ ПОВЕДЕНИЯ ЭКСПЕРТА И ПОИСК ИХ ОБЪЯСНЕНИЯ

К настоящему времени накоплено много сведений о поведении эксперта при решении практических задач. Обратимся к медицинской диагностике. Задачу диагностики можно адекватно представить как задачу классификации, решая которую эксперт относит объект к одному из нескольких классов решений.

Известно, что опытный врач использует при диагностике малое число обобщенных диагностических признаков, существенно меньшее, чем начинающий врач [11]. Известно также, что первый быстро и почти безошибочно ставит диагноз, используя “прямой”, а не обратный вывод [23]. Одна из общепризнанных черт экспертного мышления – это подсознательный характер экспертных знаний: эксперт не может вербализировать используемые им правила решения задач [17]. Решая диагностические задачи, врачи содержательно применяют клинические описания ситуаций, истории болезни и т.п.

В свете этих накопленных фактов мы можем взглянуть на оба представленных выше взгляда на экспертное мышление.

Можно предположить, что символами являются не значения диагностических признаков, а что-то иное. Но как найти такие символы? Принятый в символическом подходе метод вербальных протоколов не приведет к успеху из-за подсознательного характера экспертных знаний. Действительно, в наших экспериментах [3] мы просили врачей “рассуждать вслух” при предъявлении им клинической ситуации (описание больного). Эксперт просто называл значения ряда признаков и затем говорил, что этого достаточно для постановки диагноза. В других клинических ситуациях

с иными значениями признаков диагноз мог бы быть таким же. Из этих ответов можно было установить, что та или иная клиническая ситуация характерна для данного диагноза. Иногда удавалось найти так называемый “синдром” – сочетание ряда значений диагностических признаков. Но обычно этот синдром был характерен только для части клинических ситуаций с тем же диагнозом.

Для поиска правдоподобного объяснения поведения эксперта нам пришлось построить новые системы извлечения знаний и проанализировать найденные результаты [4].

## 6. ПОСТРОЕНИЕ ПОЛНЫХ И НЕПРОТИВОРЕЧИВЫХ БАЗ ЭКСПЕРТНЫХ ЗНАНИЙ: ПОДХОД СОХРАНЕНИЯ ЗНАНИЙ (СЗ)

Нами был разработан и реализован в виде совокупности компьютерных программ новый подход к выявлению экспертных знаний [18]. Мы не будем вдаваться в детали подхода, основанного наряду с когнитивной психологией на результатах, полученных в областях прикладной математики и компьютерных наук [5]. Отметим лишь основные идеи, необходимые для дальнейшего изложения. Как и ранее, все примеры даны для задач медицинской диагностики.

### 6.1. Описание множества проблем, для решения которых необходимо экспертное знание

Вполне обосновано предположение, что для сравнительно узкой профессиональной области эксперт может указать набор диагностических признаков, необходимых ему для постановки диагноза. Поясним понятие сравнительно узкой области профессиональной деятельности: диагностика одного заболевания, дифференциальная диагностика между двумя заболеваниями.

Для каждой такой области эксперт называет необходимые диагностические признаки и их возможные значения. Заметим при этом, что множество значений диагностического признака всегда дискретно. Даже в случае, казалось бы, непрерывной шкалы признака (например, артериальное давление), эксперт выделяет несколько информативных для него значений или диапазонов значений (например, повышенное, нормальное или пониженное артериальное давление). Чаще всего врач-эксперт использует при диагностике 2–4 значения на шкале каждого диагностического признака.

Комбинируя значения диагностических признаков, можно создавать *клинические ситуации* (КС). Каждая КС содержит описание (на языке, данном экспертом) одного из возможных пациентов. Легко подсчитать, что общее число КС равно произведению числа значений на шкалах всех диагностических признаков. Число КС – это коли-

чество различных пациентов (размерность пространства), для анализа и диагностики которых нам необходимо выявить экспертное знание.

Множество КС в данной профессиональной области создают *пространство проблем*. Приведем численные примеры. При пяти диагностических признаках с тремя значениями на шкале каждого из них пространство содержит  $3^5 = 283$  проблемы. А при 10 признаках (часто встречаются в задачах медицинской диагностики) и двух значениях на шкалах имеется 1024 проблемы, при трех значениях – 80089 проблем.

Приведем конкретный пример. В [1] рассматривается проблема построения экспертной системы для диагностики острых медикоментозных отравлений. Для построения одной из подсистем необходимо создать базу знаний, имитирующую решение эксперта при диагностике отравлений барбитуратами. Типичным лекарством этой группы является амитриптилин. Предположим, что пациент находится в коматозном состоянии.

Ниже приведен список диагностических признаков и их значений, необходимых для эксперта при диагностике отравлений барбитуратами.

### *Признаки и их значения, используемые экспертом при диагностике отравлений барбитуратами*

1. Поведение пациента до комы
  - 1.1. Моторное возбуждение, агрессивность
  - 1.2. Галлюцинации, тактильные и слуховые
2. Цвет кожи
  - 2.1. Нормальный
  - 2.2. Гиперемия
3. Судороги
  - 3.1. Есть судороги
  - 3.2. Нет судорог
4. Зрачки
  - 4.1. Расширенные зрачки
  - 4.2. Суженные зрачки
5. Пульс
  - 5.1. Тахикардия (пульс больше 110)
  - 5.2. Умеренная тахикардия (пульс 90–100)
6. Характеристика пульса
  - 6.2. Аритмия
  - 6.3. Нормальный, нет аритмии

### *6.2. Упорядочение значений признаков по характерности*

Другим видом экспертного знания, которое можно получить при первичном анализе профессиональной области, является упорядочение значений признаков по характерности для различных заболеваний. Так, при дифференциальной диагностике выдвигается гипотеза, что для каждого из диагностических признаков эксперт может упорядочить значения по характерности.

Упорядочение по характерности вводит определенный порядок в *пространстве проблем*: выделяются более и менее типичные КС по отношению к разным заболеваниям.

Отметим, что в списке первое значение каждого признака более характерно для отравления барбитуратами, чем второе.

### 6.3. Получение экспертных знаний

Центральным элементом подхода СЗ является выявление знаний путем предъявления эксперту КС – целостного образа объекта, описанного на языке значений диагностических признаков. Отметим, что именно с такой задачей сталкивается эксперт в своей практической деятельности. Специальный алгоритм формирует КС из значений диагностических признаков и предъявляет эксперту на экране компьютера.

Используя приведенный выше пример, представим возможный вопрос эксперту. Состояние пациента описывается следующим образом:

*“Пациент в коме; до комы было моторное возбуждение, агрессия; Цвет кожи – гиперемия; Судороги; Расширенные зрачки; Пульс – тахикардия, аритмии нет. Определите один из двух диагнозов:*

*S<sub>1</sub> – есть подозрение на отравление барбитуратами;*

*S<sub>2</sub> – нет подозрения на отравление барбитуратами”.*

В отличие от широко распространенного способа построения экспертных систем мы не пытаемся выявлять и представлять экспертное знание в виде продукций, исходя из того, что это знание “ситуативно” и каждый образ КС определяет для эксперта принадлежность КС к одному из классов решений.

Конечно, компьютерная система не предъявляет все КС эксперту, так как, в общем случае, их оценка потребовала бы исключительно больших затрат времени. *Используется гипотеза о том, что классификация менее характерных (для данного класса) КС определяет классификацию более характерных (с точки зрения значений диагностических признаков) КС. Эта гипотеза проверяется при построении полной классификации КС. Мы называем ее гипотезой доминирования по характерности.*

Обратимся к приведенному выше примеру. Пусть эксперт определил для предъявленной КС первый класс решений: *S<sub>1</sub> – есть подозрение на отравление барбитуратами*. В этой КС значения двух признаков: “Цвет кожи” и “Характеристика пульса” не характерны для первого класса. Используя решение эксперта и гипотезу доминирования по характерности, можно отнести к первому классу еще три КС, отличающихся от приве-

денной выше более характерными значениями одного или двух этих признаков.

Подход СЗ реализован в виде совокупности компьютерных программ, которые продолжают опрос эксперта до построения полной и непротиворечивой классификации всех КС. Эффективность программ такова, что они позволяют классифицировать 500–700 КС за час работы эксперта [2]. При этом программы позволяют получить точную копию экспертных знаний (см. далее).

Отметим еще раз существенное отличие систем, построенных на основе подхода СЗ, от обычных экспертных систем. Прежде всего КС не рассматриваются как продукция. Поэтому нет нужды в блоке логического вывода: для всех КС установлен класс решений. Гипотеза доминирования по характерности предполагает распространение классификации одних КС на другие, причем данная гипотеза проверяется.

## 7. ГРАНИЧНЫЕ ЭЛЕМЕНТЫ КЛАССИФИКАЦИИ

После построения полной классификации КС в пространстве проблем можно выделить все КС, относящиеся к каждому из классов решений. Рассмотрим КС, относящиеся к одному классу, используя отношение характерности. Можно выделить наиболее характерные для данного класса (содержащие весьма часто встречающиеся для данного класса значения диагностических признаков) и наименее характерные. Формально говоря, можно выделить верхнее и нижнее множество Парето. Так, при дифференциальной диагностике двух заболеваний в каждом классе имеется одна КС, наиболее типичная для соответствующего заболевания: сочетание всех наиболее типичных для него значений диагностических признаков. Нижнее множество Парето представляют КС, содержащие лишь несколько характерных для каждого заболевания значений признаков. Целостный образ КС определяет, по мнению эксперта, ее принадлежность к тому или иному диагнозу. Назовем такие КС *граничными* [3].

Формально граничные КС определяются следующими условиями:

1) Граничная КС не доминирует по характерности над другой КС, принадлежащей к этому же классу (иначе говоря, она наименее характерна для данного класса).

2) При изменении значения хотя бы одного из диагностических признаков возникает новая КС, принадлежащая к другому классу (иначе говоря, две КС, отличающиеся лишь значением одного диагностического признака, находятся как бы “по соседству”).

Отметим, что граничные объекты всех классов полностью представляют собой результат проведенной экспертом классификации. Действительно, принадлежность произвольного объек-

та к классу решений можно установить путем сравнения с граничными объектами. Объект принадлежит определенному классу, если он менее характерен для него, чем какой-либо из верхних граничных объектов, и/или более характерен, чем какой-либо из нижних граничных объектов.

Нижние граничные КС представляют собой наиболее трудные (для данного эксперта) для диагностики случаи, т.е. наименее типичные для данного класса решений. Отметим, что принадлежность граничных КС к классу решений всегда выявляется при их непосредственной классификации экспертом.

## 8. ПРОВЕРКА ТОЧНОСТИ ИМИТАЦИИ ЭКСПЕРТНЫХ СУЖДЕНИЙ

Так как граничные объекты представляют наиболее сложные КС, то естественно проверить точность имитации экспертных суждений путем повторного предъявления их эксперту. Таким способом неоднократно проверялись построенные классификации. В задачах с небольшим количеством проблем (до 200) качество имитации проверялось путем сравнения классификации при предъявлении каждого из объектов с классификацией, построенной через 2–3 недели изложенным выше способом. Совпадение было практически полным: различия в 2–3 объектах. При большом количестве проблем качество имитации проверялось путем повторного предъявления эксперту наиболее сложных из них через какое-то время. Совпадение было практически полным.

Можно сделать вывод, что при прочных экспертных знаниях подход СЗ позволяет получить их точную имитацию.

## 9. ПОДСОЗНАТЕЛЬНЫЕ РЕШАЮЩИЕ ПРАВИЛА

Анализ выявленных граничных объектов показал, что они не представляют собой случайный набор КС. Напротив, группы КС имеют одинаковую структуру, одинаковое число наиболее типичных и нетипичных значений диагностических признаков. Объединение КС в группы позволило выдвинуть гипотезу о скрытых, подсознательных решающих правилах, существующих в памяти эксперта и используемых им при классификации [3]. Эта гипотеза многократно проверялась при построении различных классификаций [18].

Оказалось, что в общем случае решающее правило эксперта может быть представлено в виде дерева значений признаков:

$$f(x_1 \wedge x_2 \wedge \dots \wedge x_n) + C_{N-n}^k \quad (1)$$

где  $N$  – количество диагностических признаков;  $f(x_1 \wedge x_2 \wedge \dots \wedge x_n)$  – конъюнкция характерных и наиболее важных для данного класса значений  $n$  диагностических признаков;

$C_{N-n}^k$  – сочетания по  $k$  из  $N-n$ , характерных для данного класса значений из остальных диагностических признаков.

В частных случаях первая или вторая часть в этом правиле может отсутствовать.

Так, для проблемы острого отравления барбитуратами (см. выше) одно из экспертных решающих правил звучит так:

*“Имеется подозрение на острое отравление барбитуратами в случаях, когда до комы было моторное возбуждение, агрессия и не менее чем для трех диагностических признаков из оставшихся пяти имеются характерные значения такого отравления”.*

Чаще всего граничные объекты одного класса описываются 1–3 подобными решающими правилами. В экспериментах, приведенных в [4], было выявлено, что небольшая часть граничных объектов, соответствующих правилу, может отсутствовать. Однако при их внимательном анализе и повторной классификации экспертом решающее правило подтверждалось. На наш взгляд, даже самый опытный эксперт не может быть уверен полностью при классификации ранее не встречавшихся и трудных для него КС.

Как можно показать, что эксперт действительно использует выявленные решающие правила при классификации? При проведении экспериментов фиксировалось время, затраченное экспертом при классификации каждого объекта. Оказалось, что время, затрачиваемое при классификации граничных объектов, существенно больше, чем при классификации остальных.

Табл. 1 характеризует решения эксперта при задаче классификации с 6-ю диагностическими признаками: среднее время, затрачиваемое экспертом при классификации КС.

Наиболее правдоподобным объяснением такого поведения эксперта является предположение, что при классификации граничного объекта он сравнивает его с несколькими решающими правилами (двух соседних классов), прежде чем принять окончательные решения.

Данные правила, описывающие группы граничных объектов и выявленные с помощью подхода СЗ, предъявлялись экспертам – их “авторам”. После размышлений и анализа решающие правила обычно получали одобрение такого экс-

**Таблица 1.** Среднее время (в с), затрачиваемое экспертом на классификацию одной КС при двух классах решений и шести бинарных диагностических признаках

КС	Класс	
	I	II
Внутри класса	24.6	24
На границе класса	48.7	43.75

Таблица 2. Количество экспертных правил для проблем различной размерности

L	64	64	128	256	256	1708	4096	5184	20736	32300
F	4	5	5	4	5	3	5	6	7	5

L – размерность проблемы (произведение количества значений на шкалах диагностических признаков); F – число решающих правил эксперта.

перта. В одном случае при 6-ти диагностических признаках [4] все объекты одного класса были выписаны на отдельных карточках и переданы эксперту с просьбой указать общие правила, “объединяющие” эти объекты. Эксперт указал только часть правил, относящихся к некоторым объектам. Все это показывает, что выявленные с помощью подхода СЗ решающие правила являются подсознательными. Их нельзя обнаружить путем “мышления вслух”, использования вербальных протоколов (основное средство выявления экспертных знаний при символьном подходе). Они могут быть установлены только в результате анализа совокупности действий эксперта, выполняемых в конкретных ситуациях.

Интересно, что общее количество решающих правил, выявленных в различных задачах медицинской диагностики, невелико и не превышает 7–9 (объем кратковременной памяти).

Табл. 2 представляет опыт применения подхода СЗ в задачах медицинской диагностики: количество экспертных правил для задач с различным числом проблем [18].

Из табл. 2 следует, что количество решающих правил не определяется размерностью проблемы, оно характеризует человеческую систему переработки информации, возможности ее приспособления к задачам различной размерности. На наш взгляд, такая возможность является присущей человеку, но для ее реализации требуются определенные личностные характеристики и не менее десяти лет интенсивной практики [15].

## 10. ТЕОРИЯ ПОДСОЗНАТЕЛЬНЫХ РЕШАЮЩИХ ПРАВИЛ (ПРП-ТЕОРИЯ)

Итак, классификация, точно имитирующая экспертные решения, может быть представлена несколькими решающими правилами, почти однозначно восстанавливаемыми из граничных объектов. Количество решающих правил не превышает объема кратковременной памяти, что объясняет быстроту их использования.

Можно предположить, что такие правила подсознательно формируются в памяти эксперта в течение многих (не менее 10 [15]) лет интенсивной практики. Именно с их помощью происходит “распознавание” нового объекта.

На основании полученных данных можно выдвинуть теорию подсознательных решающих правил (ПРП), которая следующим образом описывает поведение эксперта в задачах классифи-

кации. На основе конкретного вербального описания объекта у эксперта возникает его целостный, содержательный образ. Далее эксперт подсознательно сравнивает его с имеющимися в оперативной памяти и выработанными за годы практики решающими правилами, которые находятся в рабочей памяти и могут быть быстро использованы. Таким образом, осуществляется распознавание объекта и отнесение его к одному из классов решений.

На основании ПРП-теории структура памяти эксперта может быть представлена следующим образом. На нижнем уровне находится слой из нескольких сотен или тысяч реальных объектов, представляющих практические случаи, встретившиеся эксперту за время его деятельности. Описание этих объектов легко могут быть вербализованы. Вспоминание трудных случаев, извлечение их из памяти происходит, когда эксперт в ходе диагностики встречает особенно сложный объект. Следующий слой – подсознательные решающие правила, выработанные годами на основе реальных объектов. Некоторые из них могут быть не до конца сформированы. Но быстрота принятия решений, типичная для эксперта, позволяет предположить, что недостающие объекты скорее будут “упакованы” в имеющиеся структуры, чем внесены в “лист исключений”. Как мы уже отмечали, подсознательные решающие правила не могут быть вербализованы.

Верхний слой представляет собой набор информативных признаков и их значений, также выработанный экспертом в результате многолетней практики. Он может быть вербализован.

## 11. СЛЕДСТВИЯ ДЛЯ ТЕОРИИ И ПРАКТИКИ

Изложенный выше подход к выявлению экспертных знаний не является, конечно, единственным. Возможности новых технологий искусственного интеллекта, таких, как, например, нейронные сети, генетические алгоритмы, многоагентные системы, еще полностью не исследованы. Но полученные результаты позволяют поставить ряд новых теоретических и практических задач.

Прежде всего желательно найти способ классификации реальных проблем окружающего нас мира на три класса.

Первый класс: проблемы допускают декомпозицию на отдельные компоненты их описания без ущерба для их содержания.

*Второй класс:* проблемы в целом или частично имеют целостный, "ситуативный" характер, но возможна достаточно точная аппроксимация решающих правил определенной функцией от отдельных компонент описания.

*Третий класс:* отдельные компоненты описания проблемы столь сильно взаимосвязаны, что такая аппроксимация искажает описание проблемы.

Представляется, что многие инженерные, геологические проблемы относятся к первым двум классам, а медицинские – к третьему классу.

Так как подсознательный характер экспертных знаний очевиден, то скорее всего экспертные знания в самых различных проблемах могут быть представлены в виде достаточно простых по своей структуре решающих правил. Такое предположение соответствует современному уровню знаний о человеческой системе переработки информации.

С чисто практической точки зрения важно выявление экспертных знаний на полном множестве проблем в рассматриваемой предметной области. Такой подход может помочь найти целостный облик экспертного знания.

## 12. ЗАКЛЮЧЕНИЕ

На основе опыта многолетней работы в области медицинской диагностики мы можем сделать вывод о "ситуативности" экспертного мышления. Следует в то же время подчеркнуть, что механизм распознавания объектов экспертом может быть правдоподобно описан на основе символической теории. Подсознательные решающие правила можно рассматривать как специальные символы, имеющие подсознательный характер. В рамках ПРП-теории оба подхода справедливы.

Наряду с построением экспертных систем путем имитации поведения эксперта на основе подхода СЗ ПРП-теория позволила создать компьютерную систему ОСТЭЛА [6]. Эта система предназначена для обучения искусству диагностики тромбозмболии легкой артерии на основе решающих правил эксперта.

В системе ОСТЭЛА используется принцип подсознательного обучения, когда обучаемый (молодой врач) должен сам воссоздать правила диагностики, а система предъявляет множество КС (начиная от легких и переходя к трудным) и сообщает ему лишь о правильности его решения, причем за эталон берутся выявленные ранее подсознательные решающие правила эксперта.

Помимо медицины подход СЗ и ПРП-теория могут быть распространены на другие области человеческой деятельности.

## СПИСОК ЛИТЕРАТУРЫ

1. Асанов А.А., Ларичев О.И., Нарыжный Е.В., Страхов С.И. ЭСТЕР – экспертная система для диагностики острых отравлений лекарствами // Тр.

- ВII национальной конференции по искусственному интеллекту. М.: Физматлит, 2000. С. 708–716.
2. Ларичев О.И., Асанов А.А. Метод ЦИКЛ порядковой классификации многокритериальных альтернатив // ДАН. 2000. № 5. С. 592–596.
3. Ларичев О.И. Структура экспертных знаний в задачах классификации // ДАН. 1994. № 6. С. 750–752.
4. Ларичев О.И. Структуры экспертных знаний // Психологический журнал. 1995. № 3. С. 82–89.
5. Ларичев О.И., Мечитов А.И., Мошкович Е.М., Фуремс Е.М. Выявление экспертных знаний. М.: Наука, 1989.
6. Ларичев О.И., Нарыжный Е.В. Компьютерное обучение процедуральным знаниям // Психол. журнал. 1999. № 6. С. 53–61.
7. Нильсон Н. Искусственный интеллект. М.: Мир, 1973.
8. Саймон Г. Науки об искусственном. М.: Мир, 1972.
9. Саймон Г., Чинг Ц. Распознавание, мышление и обучение как информационные процессы // Психол. журн. 1988. № 2. С. 33–46.
10. Anderson J.R. The Architecture of a Cognition. MA: Harvard University Press, 1983.
11. Boshuizen H.P.A., Schmidt H.G. On the role of biomedical knowledge in clinical reasoning by Experts, Intermediates, and Novices // Cognitive Science. 1992. № 16. P. 153–184.
12. Clancey W.J. The epistemology of a rule-based expert systems. Framework for explanation // Artificial Intelligence. 1983. № 20 (3). P. 215–225.
13. Clancey W.J. Notes on "Epistemology of a rule-based expert system" // Artificial Intelligence. 1993. № 59. P. 197–204.
14. Dreyfus H.L., Dreyfus S.E. Making a mind versus modeling the brain: artificial intelligence back at a branch-point // The philosophy of artificial intelligence / Ed. M. Boden. Oxford: Oxford University Press. 1985. P. 307–333.
15. Ericsson K.A. The Acquisition of Expert Performance: An Introduction to Some of the Issues // The Road to Excellence: The Acquisition of Expert Performance in the Arts and Sciences, Sports and Games / Ed. K.A. Ericsson. Hillsdale, NJ.: Lawrence Erlbaum Associates. 1996. P. 1–51.
16. Greeno J.G. The Situativity of Knowing, Learning and Research. American Psychologist. 1998, January. P. 5–26.
17. Kihlstrom J.F. The Cognitive Unconscious // Science. 1987. V. 237. P. 1445–1452.
18. Larichev O.I. Close Imitation of Expert Knowledge: The Problem and Methods // International J. of Information Technology & Decision Making. 2002. № 1. P. 27–42.
19. Mechner D.A. All Systems Go // The Sciences. 1998, Jan. P. 30–38.
20. Newell A., Simon H.A. Human Problem Solving. Englewood Cliffs, NJ: Prentice-Hall Inc., 1972.
21. Newell A., Simon H. Computer science as empirical enquiry: symbols and search // The philosophy of artificial intelligence / Ed. M. Boden. Oxford: Oxford University Press, 1990.
22. Nilsson N. Artificial intelligence. A new Synthesis. San Francisco: Morgan Kaufmann Publishers, 1998.
23. Patel V., Groen G. The general and specific nature of medical expertise: a critical look // Toward a general the-

- ory of expertise / Ed. K.A. Ericsson and J. Smith. Cambridge: Cambridge Univ. Press, 1991. P. 93–125.
24. Pople H. The formation of composite hypothesis in diagnostic problem solving—An exercise in sythetic reasoning // IJCAI. 5. P. 1030–1037.
25. Rodoliz N.S., Clancey W.J. GUIDON – MANAGE: teaching the process of medical diagnosis // Medical Cognitive Science / Ed. D. Evans, V. Patel. Cambridge: Bradford Books, 1989. P. 313–348.
26. Simon H. Reason in Human Affairs. Stanford: Stanford University Press, 1983.

## THEORY OF SUBCONSCIOUS DECIDING RULES AND ITS APPLICATION IN DIAGNOSTICS

O. I. Larichev

*Academician of RAS, Institute of system analysis, RAS*

The phenomenon of expert knowledge is analyzed; empirical data that can explain the process of recognition of “holistic” situations within the framework of the theory of subconscious decision rules are presented. On this basis, diagnostics is considered as classification task when the expert rates an object at one of the types of decisions. Imitating expert judgments, this task can be presented with some decision rules that are reconstructed almost unambiguously from the objects. The number of decision rules doesn’t exceed the short-term memory span that’s why they are used so quickly. The conclusions are made about the situational character of expert thinking and that the mechanism of objects recognition can be described using the symbolic theory. The decision rules are considered as the special symbols of subconscious character.

*Key words:* thinking, symbolic approach, medical diagnostics, retention of knowledge, clinical situations, expert knowledge.

## ЭХО ВЗРЫВОВ: СРАВНИТЕЛЬНЫЙ АНАЛИЗ ВОСПОМИНАНИЙ МОСКВИЧЕЙ О ТЕРРОРИСТИЧЕСКИХ АКТАХ 1999 г. (Москва) и 2001 г. (Нью-Йорк)<sup>1</sup>

© 2003 г. В. В. Нуркова\*, Д. М. Бернштейн\*\*, Э. Ф. Лофтус\*\*\*

\*Кандидат психологических наук, доцент МГУ им. М.В. Ломоносова, Россия

\*\*Доктор философии, научный сотрудник Университета штата Вашингтон, США

\*\*\*Почетный профессор Университета Калифорнии, Ирвайн, штат Калифорния, США

Исследованы воспоминания москвичей о взрывах жилых домов в Москве 9 и 13 сентября 1999 г. и террористической атаке на Всемирный торговый центр в Нью-Йорке 11 сентября 2001 г. Показано, что в формировании субъективной отчетливости и стабильности воспоминаний об исторических событиях значительную роль играет фактор визуальной кристаллизации и означивания с помощью СМИ. Подтверждена гипотеза о том, что личностная включенность в событие ведет к снижению “качества” образа памяти, вследствие чего испытуемые в большей степени склонны доверять “социально сформированным” образам памяти, чем непосредственно пережитым.

*Ключевые слова:* автобиографическая память, историческая память, формирование и искажение воспоминаний.

Американский ученый У. Найссер заметил: “Личные воспоминания – это способ соотнесения индивидуальной истории и Истории” [13]. Несомненно, индивидуальное бытие всегда разворачивается в конкретном историческом контексте. М. Линтон в статье “Феникс и Химера: Изменчивое лицо памяти” [9] связывает фиксацию обще-

ственных событий в индивидуальной памяти, во-первых, с их значимостью для частной жизни, во-вторых, с их ролью “маркеров” временной оси жизни и, в-третьих, с повторяющейся процедурой “восстановления” воспоминания при помощи социальных воздействий. По нашему мнению, правомерно расширить перечень функций отражения информации об исторических событиях в автобиографической памяти и полагать, что исторические

<sup>1</sup> Работа выполнена при грантовой поддержке фонда Фулбрайт (США).