



Изв.вузов «ПНД», т.4, № 1, 1996

УДК 681

**МОДЕЛИРУЮЩИЕ СИСТЕМЫ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ  
ДЛЯ ЗАДАЧ КЛАССИФИКАЦИИ:  
новая информационная технология для работы  
с плохо формализованными знаниями**

*M. Г. Крейнес*

Рассмотрены методологические основы, принципы реализации и практического применения новой информационной технологии работы с плохо формализованными знаниями при решении задач классификации. Основной особенностью предлагаемой технологии является предоставление пользователю эффективной помощи в поиске актуальных для решения конкретной задачи знаний и данных в результате моделирования объектов профессионального мира пользователя.

**Введение**

Компьютерные системы представления знаний, которым посвящена данная работа, моделируют профессиональные ситуации и объекты профессионального мира и создают на основе моделей своеобразное игровое пространство, которое с удовольствием используют в качестве среды для упражнений и поддержки принятия решений обучаемые, преподаватели и профессионалы. На сегодня системы, созданные в рамках рассматриваемой технологии, регулярно и эффективно используются при обучении и профессиональной переподготовке студентов-медиков и врачей (более подробно см. [1]).

Основное содержание статьи - обобщение результатов и конкретного опыта, накопленных при разработке информационной технологии и при создании конкретных моделирующих систем представления знаний для задач медицинской диагностики. Не вина автора, что временами это обобщение приобретает общефилософский характер: такова специфика предмета. Многолетние усилия математиков, программистов и инженеров сделать что-нибудь полезное с плохо формализованными знаниями в рамках имеющихся навыков и умений требуют достаточно широкого осмыслиения. Уже простое понимание причин нежелания или невозможности для врачей использовать статистические методы или методы искусственного интеллекта при решении практических задач далеко выходит за рамки чисто технической проблемы. Так получилось, что это осмыслиение шло практически параллельно с формированием отчетливых представлений о функциях, структуре, возможностях моделирующих систем представления знаний для задач классификации и с накоплением опыта реализации этой новой информационной технологии для задач медицинской диагностики. Поэтому значительное место в статье занимают рассуждения о принципиальных основах

и базовых структурах профессиональной интеллектуальной деятельности в плохо формализованных предметных областях.

Человек настолько ограничен (интеллектуально и во времени) и так устроен, что ему необходимо прибегать к чужому опыту и знаниям для решения сложных задач. При поверхностном взгляде на вопрос кажется, что компьютер, имитирующий рассудочную деятельность специалиста, - идеальное средство, облегчающее применение чужих или же отчужденных знаний к решению задач, возникающих перед данным пользователем.

Реальные проблемы появляются, если речь идет о сложной и плохо формализованной предметной области, когда имеется необходимость использования для эффективной деятельности определенной системы теоретических знаний в качестве руководства и отсутствуют надежные критерии отбора моделей и методов, способных привести к успеху в решении конкретной задачи. Ведь тогда пользователь, чтобы получить результат, под которым он мог бы подписаться сам, вынужден выбрать или индивидуализировать применительно к своей задаче компьютерные способы имитации рассудочной деятельности специалиста - те именно способы, которые он планирует использовать. Однако, решение этих проблем наталкивается на два препятствия. Во-первых, осмысленный выбор неявно предполагает, что пользователь не менее мудр, опытен и профессионален, чем источник применяемого знания. (А тогда почему бы ему не решить задачу самому?) Во-вторых, осмысленная индивидуализация возможна лишь в случаях, когда компьютерная система ориентирована на конструктивную гибкую модель представления знаний, допускающую учет ранее не предусмотренных в модели факторов и обстоятельств. Но существование таких моделей представления знаний противоречит гипотезе об отсутствии достоверных критериев отбора эффективных методов решения практических задач в плохо формализованных предметных областях. (Действительно, если предположить существование подобной модели, то в ее рамках естественным образом формулируются условия адекватного применения способов имитации рассудочной деятельности и нет оснований говорить о предметной области как о плохо формализованной.)

Для сложных и плохо формализованных областей деятельности распространение компьютерной техники остро ставит вопрос о создании безопасных информационных технологий: таких информационных технологий, использование которых не маскирует (в том числе, ссылками на авторитет, научность или традицию) безответственного применения имитаций чужого сознания. В докомпьютерную эру, пока знания передавались с помощью книги и живой традиции, ученик должен был перерasti учителя, чтобы стать профессионалом. С другой стороны, приобретенное человеком знание включалось в общую систему индивидуальных знаний и представлений о мире и, в принципе, не существовало вне фактов и обстоятельств, напрямую с конкретным знанием не связанных. В такой ситуации обращение к авторитетному учителю с конкретной задачей и просьбой решить ее или подсказать способ решения было естественным и эффективным. Прежде всего потому, что при этом происходило обращение не к определенной совокупности знаний, а к целостной личности, не только владеющей профессиональными знаниями и навыками их применения, но и со сформировавшимся взглядом на мир в целом - взглядом, являющимся основой для интерпретации и индивидуализации профессиональных знаний применительно к данной задаче.

### **Базовая идея создания моделирующих систем представления знаний**

Рассмотренная в данной работе технология работы со знаниями гарантирует сохранение духа европейской книжной традиции и живой преемственной связи с учителем. Принципиальная идея этой информационной технологии заключается

в следующем. Мы научились с помощью компьютера строить и поддерживать базы знаний, отражающие профессиональные взгляды конкретных специалистов (см.[1, 2]). Далее, мы научились решать на компьютере на основании таких баз знаний задачи моделирования объектов профессионального мира нашего специалиста и строить их описания (см. [1, 2]). Также мы научились описывать реальные объекты на профессиональном языке нашего специалиста и освоили методы измерения соответствующих параметров у объектов реального мира. При этом можно использовать результаты измерений лишь некоторых параметров для задания частичного описания реальных объектов (в противоположность полному описанию, когда для конкретного объекта фиксированы значения всех параметров, на основе которых формулируются профессиональные взгляды специалиста). Наконец, мы научились, используя решение задачи моделирования, достраивать частичные описания реальных объектов до полного описания смоделированных объектов гипотетических объектов профессионального мира нашего специалиста (см. [2]). В принципе, решение такой задачи домоделирования не единственно: им может быть любой модельный объект, не противоречащий знаниям из базы знаний, описывающим моделируемый класс объектов. Теперь если выбрать конкретный реальный объект и начать работать с его частичными описаниями, то выясняется удивительное свойство «популяции» смоделированных объектов (для каждого частичного описания мы будем много раз решать задачу домоделирования). Чем меньше мы знаем об объекте (или предполагаем, что знаем, ибо мы можем включить свои предположения в описание реального объекта как гипотетические результаты измерений), тем более разнообразными будут представители популяции смоделированных объектов. По-существу, чем больше мы знаем о решаемой задаче (реальном объекте, с которым мы столкнулись), тем более однозначной будет реакция нашей моделирующей системы. В частности, если мы знаем о реальном объекте все, что можно знать, - его полное описание, то мы сможем смоделировать только это же описание, если, разумеется, оно соответствует профессиональным знаниям, включенными в базу знаний. Если такого соответствия нет, то в данном случае решить задачу домоделирования окажется вообще невозможным из-за противоречивости имеющихся данных и знаний. (На практике, в междисциплинарном коллективе, руководимом автором, поэтапная процедура моделирования для различных по степени детализации описаний реальных объектов широко используется в качестве средства верификации медицинских баз знаний).

Если взглянуть на описанные принципы работы моделирующих систем представления знаний с точки зрения математика, то легко увидеть, что наши системы не прибегают к решению обратных задач, а решают только прямые задачи. Более того, наши системы структурно не приспособлены к решению обратных задач - над базами знаний в них отсутствуют машины логического вывода.

Другими словами, рассматриваемая в данной работе информационная технология - не способ имитировать сознание специалиста, а способ моделировать профессиональный мир, каким он представляется специалисту. Для человека это - удивительное средство осознания и сохранения собственного внутреннего профессионального мира. Для его учеников это - средство обучения и поддержки принятия решений, обладающее знаниями учителя и, одновременно, не пытающееся влиять на формирование целей и на оценки полезности способов поведения учеников при решении конкретной профессиональной задачи. Вопрос об ответственности учителя за поддержку принятия решений не возникает - система консультирует только по поводу моделей. Это средство, моделирующее объекты профессионального мира учителя, соответствующие задачам учеников, - своеобразная виртуальная реальность, порождаемая интеллектом учителя.

Дальнейшее рассмотрение будет проведено на примере моделирующих систем, ориентированных на задачи медицинской диагностики. По существу, для обобщения изложенного на произвольные задачи классификации требуются лишь терминологические изменения.

## **Особенности знаний и рассуждений в плохо формализованных предметных областях (на примере медицинской диагностики)**

С точки зрения моделирования целенаправленного поведения особенности медицинского мышления определяются непростыми соотношениями формализованного знания типа «знаю как» с персональным глубинным знанием типа «знаю что». По существу, речь идет, с одной стороны, о формализованных описательных моделях (о возможных вариантах проявления болезней) и о процедурных моделях (о ступеньках процесса понимания совокупности симптомов как проявлений конкретного заболевания) и, с другой стороны, о моделях прагматики поведения (об оценках полезности для данного врача тех или иных действий по отношению к конкретному больному в условиях имеющейся симптоматики, базирующихся на неформализованных образах, личных впечатлениях и опыте), которые могут служить основой не только для отбора и объяснения описательных или процедурных моделей, но и для их пересмотра. Для врача «знаю что» без «знаю как» бессмысленно и немыслимо «знаю как» без «знаю что».

В клинической медицине основные формализованные структуры медицинского знания: описательные модели (модели семиотики заболеваний) и процедурные модели, в сущности, модели способов понимания (модели семантики) - характеризуются значительным разнообразием, противоречивостью и иерархичностью взаимосвязей. Ведь процедуры понимания клинической ситуации в конечном итоге основаны на определенном упрощении и идеализации описательных моделей заболеваний [2 - 4], а формы проявления заболеваний весьма вариабельны. Разнообразие и противоречивость моделей семиотики и семантики делает для врача каждый раз необходимым выбор фиксированной модели из множества существующих моделей или модификацию выбранной модели для конкретного клинического случая. Применение нескольких моделей одновременно либо бессмысленно (если все модели дадут один результат), либо все равно потребует выбора (если какие-либо из использованных моделей дадут разные результаты). Выбор моделей или их модификация в рамках возможностей, предоставляемых формализованным медицинским знанием, реализуются врачом на основании персонального знания - личных моделей прагматики поведения. Модели прагматики поведения определяются ценностными установками врача и причинно-следственными связями. К последним естественно отнести и представления о принципиально возможном и невозможном. К сожалению, в медицине не удается строить причинно-следственные модели, соответствующие стандартам точных наук. (В медицине связь причины и следствия почти в той же степени вопрос веры и интуиции, что и в Истории.)

Таковы, в общих чертах, основные содержательные факты, которыми недопустимо пренебрегать при создании и внедрении в медицину, как и в практически любую плохо формализованную предметную область, новых методов и средств информатики. Дело в том, что только ориентация на особенности профессионального мышления и знания способна обеспечить реальный успех создаваемых компьютерных систем. Например, широко и разносторонне обсуждается факт непопулярности у практиков медицинских компьютерных систем, ориентированных на поддержку принятия решений и профессиональное обучение [3 - 7]. Высказываются разнообразные и, зачастую, противоречивые причины такой непопулярности - от недостаточных технических возможностей ЭВМ до консерватизма врачей [5 - 7]. Следует отметить, что большинство подобных аргументов выглядит откровенно неубедительно, хотя бы рядом с историей стремительного внедрения и распространения в медицинской практике компьютерной томографии.

Выделение в структуре медицинского диагностического знания трех типов моделей: семиотики, семантики и прагматики, - позволяет предложить простое и логичное объяснение указанной группы явлений. Врач не хочет, чтобы компьютер определял прагматику поведения при принятии решений (в том числе, диагностических решений). А именно прагматику поведения, по существу,

навязывают байесовские классификаторы и системы искусственного интеллекта, «помогающие» в формировании стратегий диагностического поиска или критикующие поведение врача [8]. Врачу нужна информация о моделях семиотики и семантики, связанная с конкретным клиническим случаем и выдвигаемыми самим врачом диагностическими гипотезами. Эти знания врач использует для формирования стратегии сбора данных о больном и для интерпретации собранных данных. В принципе, необходимую информацию врач получает из литературы, но нужно суметь найти в книгах и в журналах конкретные знания, связанные с конкретной задачей. В рамках технологии моделирующих систем представления знаний такая постановка задачи позволяет ограничить функции вычислительной системы решением только прямых задач [2,4,9,10] (см. подробнее следующий раздел).

Моделирующие системы представления знаний позволяют эффективно работать с моделями семиотики и семантики, соответствующими конкретной практической ситуации, и при этом не вмешиваются в прагматику поведения врача. Речь идет о системах представления знаний, фактически, реализующих новый вариант активной, но не переходящей границ навязчивости, компьютерной книги. Существенно, что в рамках рассматриваемой технологии компьютер не просто перелистывает страницы, а реализует совершенно новые, ранее недоступные функции. Например, наша «обобщенная» компьютерная книга способна сама рисовать иллюстрации к собственному содержанию, позволяет читателю стать действующим лицом и самому сочинять сюжет. Эта книга готова проконсультировать читателя по поводу разумных способов поведения при различных перипетиях сюжета и проверить рациональность поведения читателя и соответствие поведения излагаемым в книге моделям (в нашем случае - моделям семиотики и семантики заболеваний).

Использование нашей нестандартной компьютерной книги обеспечено ее содержанием - описанием на некотором формальном языке предметных знаний - моделей семиотики и семантики отдельных заболеваний и их групп, связей между значениями параметров, представлений о возможном в природе и о распространенности заболеваний и различных форм их проявления. Такое описание содержательных знаний - база знаний системы для конкретной предметной области - выполняется в рамках модели представления медицинских диагностических знаний [2,9,11].

### **Основные задачи, решаемые моделирующей системой представления знаний**

В рассматриваемой нами технологии основной механизм поиска и применения экспертных знаний, актуальных при решении конкретных задач пользователя, реализуется решением задачи моделирования профессиональной ситуации на основании базы знаний. Алгоритмы такого моделирования подробно рассмотрены в [2, 9, 12] и ориентированы на конкретную структуру баз знаний [2, 9], основанную на модели представления диагностических знаний [2, 11]. Последняя отражает не только модели семиотики и семантики для конкретных заболеваний и их групп (и, по существу, включает знания об иерархии проявления заболеваний), но и знания о иерархической организации и структуре самого пространства параметров, в рамках которого описываются представления эксперта - знания о том, какие параметры являются определяющими для других параметров и каковы функциональные или статистические зависимости между параметрами. Для описания структуры пространства параметров важно выделить группы параметров, для значений которых существенно наличие совместных ограничений: вектор параметров может принимать значения лишь из подмножества декартова произведения множеств значений компонентов. Следует подчеркнуть, что задание зависимостей само по себе не определяет иерархическую организацию пространства параметров, которая характеризует причинно-следственные связи между параметрами. Например, значения двух параметров

могут быть пропорциональны друг другу, но один из них причинно определяет значение второго, и его место в иерархии будет выше. Подобные иерархические отношения возможны и для групп параметров.

Решение базовой (для нашей технологии) задачи моделирования ситуации положено в основу достижения пяти главных для пользователя целей создания моделирующих систем представления знаний, формализованных как следующие задачи:

**Задача 1.** Моделирование клинической ситуации, соответствующей выбранному заболеванию (нескольким заболеваниям, группе заболеваний).

**Задача 2.** Проверка соответствия конкретных данных о больном принятому врачом диагностическому решению (успешное решение задачи домоделирования частично описанной ситуации для выбранного врачом диагноза).

**Задача 3.** Проверка достаточности конкретных данных о больном для формального подтверждения предложенного врачом диагностического решения (успешное решение задачи домоделирования частично описанной ситуации только для выбранного врачом диагноза и невозможность ее решения для других диагнозов).

**Задача 4.** Выявление (локализация на клиническом алгоритме) ошибок пользователя (отступлений от предусмотренных в системе моделей семиотики и семантики), допущенных при анализе конкретной клинической ситуации.

**Задача 5.** Определение диагностических критериев для гипотез, предусмотренных в системе, и проверка соответствия указанным критериям данных о клиническом случае.

Решение задачи моделирования клинической ситуации (**задача 1**) основано на отыскании произвольной точки в пространстве параметров, удовлетворяющей условиям, описанным в базе знаний, и выполняется простым методом случайного поиска. Эти условия налагаются моделями семиотики (или семантики) выбранного заболевания и включающих его группы заболеваний, а также связями между значениями параметров.

Трудоемкость вычислений оказывается приемлемой благодаря иерархической организации пространства параметров. Действительно, задача поиска допустимой точки каждый раз решается для подмножества пространства параметров, определяемого совместными ограничениями на значения параметров, которое обычно имеет небольшую размерность. Вначале рассматривается набор взаимосогласованных параметров с максимальным уровнем иерархии, затем - со следующим уровнем иерархии и т.д. При этом найденные допустимые значения параметров из подпространств более высокой иерархии считаются зафиксированными. Если на каком-либо уровне иерархии система не может отыскать допустимую точку за априори заданное число случайных шагов, то происходит возврат к набору параметров с максимальным уровнем иерархии. Такая, по существу, эмпирическая процедура хорошо работает на практике [2, 9, 12].

Проверка соответствия конкретных данных о больном принятому врачом диагностическому решению (**задача 2**) реализуется с помощью решения задачи 1 для заболевания, выбранного пользователем. В процессе домоделирования база знаний формально дополняется данными, описывающими конкретный клинический случай. Если такая задача домоделирования решается успешно, это свидетельствует о соответствии данных о конкретном больном модели выбранного врачом заболевания. Неудача в решении может означать, что данные противоречат условиям, предусмотренным базой знаний.

Проверка достаточности конкретных данных о больном для формального подтверждения предложенного врачом диагностического решения (**задача 3**) реализуется многократным безуспешным решением задачи 2 для диагностических гипотез, исключающих ту, которая была выдвинута врачом. Если же решение задачи моделирования для некоторой альтернативной гипотезы оказалось успешным, то это - свидетельство ее необоснованного исключения пользователем

из дальнейшего рассмотрения, то есть недостаточности имеющихся данных для доказательства гипотезы врача.

Выявление и интерпретация ошибок (отступлений от предусмотренных в системе моделей семиотики и семантики или связей между параметрами), допущенных пользователем при анализе конкретной клинической ситуации (**задача 4**), также основаны на решении задачи моделирования. Только теперь, в отличие от предыдущего случая, отыскивается ближайшая вышестоящая по отраженной в базе знаний иерархии диагностическая гипотеза, для которой решение задачи моделирования еще является успешным, и тем самым определяется локализация ошибки.

Необходимо отметить, что использование для решения задачи моделирования алгоритма случайного поиска делает безусловно значимым только удачное решение задачи моделирования. Ведь то, что методом случайного поиска за определенное число итераций не удалось отыскать допустимую точку, строго говоря, можно трактовать лишь в вероятностном смысле. Однако, несколько повторных неудач для конкретной гипотезы должны заставить пользователя задуматься. При регистрации числа шагов, необходимых для отыскания допустимой точки в диагностическом тренажере по анемиям (обучающей системе, использующей моделирование клинических ситуаций [3,4,9]), неожиданно выявилось достаточно хорошее соответствие относительных объемов многомерных параллелепипедов, представляющих модели семиотики различных групп заболеваний, описанных в базе знаний системы, оценкам относительной распространенности этих групп заболеваний. Может быть, мы хвораем от того, что природа раскидывает нас по допустимым областям в пространстве параметров?

Определение диагностических критериев или параметров, включенных в диагностические критерии, для диагностических гипотез, предусмотренных в системе (**задача 5**) и проверка соответствия указанным критериям конкретного набора клинических данных выполняется непосредственно по информации из базы знаний системы.

Возможно также самостоятельное формирование пользователем с помощью решения задач 1-5 достаточно сложных сценариев анализа клинических ситуаций. Вот пример такого сценария. Выявлена недостаточность данных о больном для формального подтверждения принятого врачом решения и выявлен перечень альтернативных решений, не противоречащих имеющимся данным. Используя решение задачи 1, можно просмотреть типичные клинические ситуации для каждого из найденных диагнозов. Задача 5 подскажет, какие данные представляют наибольший интерес. При формировании стратегии сбора дополнительной информации полезно сначала умозрительно предположить, что определенные обследования уже выполнены и дали тот или иной результат, и для разных вариантов решить задачи 2, 3, 4. Таким образом, пользователю моделирующей системы предоставляется возможность игры (познавательной, обучающей и увлекательной) с модельными или гипотетическими клиническими ситуациями, а также с «полумодельными» ситуациями, полностью соответствующими реальному больному в значениях зарегистрированных у него параметров (оставшиеся неизвестными для реального больного показатели фиксируются в результате решения задачи домоделирования).

Применяя технологию моделирующих систем представления знаний, мы создаем новый тип компьютерной книги, предназначенный одновременно для человека читающего и для человека играющего. Способность системы моделировать клинические ситуации позволяет обеспечить квазинтеллектуальное поведение системы.

Использование различий моделей семиотики и семантики наряду с допустимостью игнорировать собранную о больном информацию или определенные компоненты моделей дают пользователю возможность продумывать модели прагматики собственного поведения и даже экспериментировать с ними в «игровом» режиме работы моделирующей системы. Целесообразно еще раз подчеркнуть важную особенность функционирования моделирующих систем

представления знаний в качестве систем поддержки принятия решений. Наши системы не навязывают пользователю «собственное мнение», а в различных формах предоставляют знания, связанные с конкретной ситуацией принятия решений.

## Заключение

Сегодня мало внимания уделяется серьезной опасности, связанной с развитием и проникновением во все сферы жизни средств и методов информатики - от простейших калькуляторов до современных вычислительных комплексов, позволяющих моделировать сложнейшие процессы и явления. Следствием такого проникновения информатики в жизнь является создание комфортных условий для применения чужих знаний к конкретным задачам пользователя, или, что тоже самое, создание привлекательной модельной среды, которая готова подменить собой реальный мир. По существу, речь идет о важной общефилософской и моральной проблеме: об ответственности за адекватность и за результаты применения отчужденного знания. Корень проблемы в том, что информация и знания, применение которых само по себе требует определенных профессиональных навыков, становятся доступными для использования широкой публикой, не имеющей необходимых навыков. Заметим, это было не так, пока основным источником информации и знаний были книги и живая традиция.

Много десятилетий назад великий лингвист Сепир [13] говорил о естественном стремлении человека ограничить сферу распространения порождаемых им текстов. Сепир видел за этим стремление человека-автора сохранить свою индивидуальность. Быть может, сегодня опасность для человечества - потеря индивидуальности людьми, с удовольствием и безответственно пользующимися отчужденными информацией и знаниями и готовыми охотно их распространять.

Речь должна идти о создании новой научной дисциплины - экологии информатики, то есть экологии информационных технологий. Ее основная задача - исследование условий безопасного, в широком смысле слова, применения отчужденных знаний и информации. Здесь следует изучать с точки зрения функционирования в рамках человеческих сообществ: профессиональные знания и социально значимую информацию, рекламу и призывы к насилию.

Моделирующие системы представления знаний - первый шаг в данном направлении. Для широкого и практически интересного класса задач (задач классификации в плохо формализованных предметных областях) на базе методологии анализа целенаправленного поведения не только проведено теоретическое исследование проблемы, но и сформулированы критерии выделения «экологически грязных» информационных технологий, предложены методология и методы создания и практического применения экологически чистых информационных технологий. Эффективность разработанных информационных технологий в обучении около 2000 студентов-медиков продемонстрирована практикой [1,2,14,12].

Особенностью предлагаемых в данной работе информационных технологий является сохранение традиций активной сознательной деятельности человека, как профессионала и как личности, при широком использовании любых вспомогательных средств, в том числе и мощного арсенала средств и методов современной информатики.

*Данная работа выполнена благодаря поддержке Министерства науки и технической политики РФ (в рамках проекта «Разработка моделирующей системы представления знаний для поддержки диагностических решений», проект № 1110 федеральной научно-технической программы «Перспективные информационные технологии») и Российского фонда фундаментальных исследований (в рамках проекта «Разработка и исследование модели рассуждений при решении задач в плохо формализованных предметных областях», грант 95-01-01583).*

## **Библиографический список**

1. Крейнес М.Г., Сумароков А.В., Удальцов Ю.А., Погромов А.П., Ериков В.И. Активные компьютерные системы для медицинского образования: подход и реализация // Вестник РАМН. 1994. № 9.
2. Крейнес М.Г. Модели семиотики, семантики и прагматики в компьютерных системах для медиков // Проблемы информатизации. 1993. Вып. 3-4. С. 57.
3. Крейнес М.Г. Использование средств и методов информатики в обучении медиков: аналитический обзор // Медицинский реферативный журнал. Раздел XXII. 1987. № 12. С. 30.
4. Gremy F. Person and computers in medicine and health // Meth.Inform. 1988. Vol.27, № 1. P. 3.
5. Eddy D.M. Clinical decision making: from theory to practice // JAMA. 1990. Vol. 263, № 2. P. 287.
6. Greens R.A., Shortliffe E.H. Medical Informatics - an emerging academic discipline and institutional priority // JAMA. 1990. Vol.263, № 8. P. 1114.
7. Seelos H.-J. A new paradigm of medical informatics // Meth. Iform. Med. 1992. Vol.31, № 2. P. 79.
8. Selected topics in medical artifical intelligence / Ed. P.L. Miller. N-Y.: Springer-Verlag. 1988. 231 р.
9. Крейнес М.Г. Компьютерные моделирующие системы представления знаний для врача и больного // Вестник Всесоюзного общества информатики и вычислительной техники. 1991. № 1. С. 52.
10. Miller R.A., Masarie F.E. Use of the Quick Medical Reference(QMR) program as a tool for medical education // Meth. Inform. Med. 1989. Vol. 28, № 4. P. 340.
11. Крейнес М.Г. Математическая модель диагностического процесса как средство представления содержательных медицинских знаний // Проблемы проектирования экспертных систем. М.,1988. С. 144.
12. Kreines M., Purtov V., Udalzov Y. Diagnostic trainer-approach, model and software // East-West Conference on Emerging Computer Technologies in Education,6-9 April 1992,Moscow,Russia,Conference Proceedings. M., 1992. С. 181.
13. Сепир Э. Избранные труды по языкоизнанию и культурологии. М.: Издательская группа «Прогресс». 1993. С. 215.
14. Крейнес М.Г., Сучков А.В., Ериков В.И. К задаче создания машинной обучающей системы для студентов медиков на основании математической модели диагностического процесса // Вопросы применения экспертиных систем. Минск, 1988. С. 206.

*Московский специализированный центр  
новых информационных технологий  
на базе Московской медицинской  
академии*

*Поступила в редакцию 14.07.95  
после переработки 15.01.96*

## **MODELLING KNOWLEDGE BASED SYSTEMS FOR CLASSIFICATION PROBLEMS: new information technology for insufficiently formalized knowledge**

*M. G. Kreines*

We discuss methodological base, ways of realization and problems of practical usage of the new information technology for insufficiently formalized knowledge. This technology gives user the effective help in gathering essential for the problem knowledge and information and does not disturb the pragmatic of the user.



Крейнес Михаил Григорьевич - родился в 1953 году в Москве. Окончил Московский авиационный институт (1975). После окончания института работает в Московской медицинской академии над проблемами использования математического моделирования в биологии и медицине. Защитил диссертацию на соискание ученой степени кандидата физико-математических наук, посвященную моделированию систем поддержки решений, в Вычислительном центре АН (1989). Создал и возглавил Московский специализированный (по медицинским специальностям) центр новых информационных технологий на базе Московской медицинской академии (1993). В настоящее время основные области научных интересов: изучение и моделирование профессиональных рассуждений человека в плохо формализованных предметных областях, создание на базе этих теоретических исследований новых типов компьютерных обучающих систем и систем поддержки принятия решений для медиков.



## KNOWLEDGE BASED COMPUTER SYSTEMS FOR DECISION SUPPORT AND LEARNING DIAGNOSTIC TRAINER

Diagnostic Trainer is a new word in the educational computer systems in medical diagnostics. It is based on the original mathematical model of medical knowledge, human diseases, diagnostic methods and effective strategies of diagnostic problem solving.

A user can run Diagnostic Trainer in form of a game, during which the system simulates clinical cases, while user analyzes them. Diagnostic Trainer uses it's knowledge base for simulations and doesn't need any clinical cases to be prepared in advance. Diagnostic Trainer is a «patient» and «tutor» at the same time. As a patient it simulates the description of patient's clinical case, while, as a tutor, it consults user on the diagnostic process and simultaneously checks the validity of users actions and efficiency of decision making strategy. The combination of patient and tutor provides a really personal educational process, effective feedback, helps user to acquire practical skills and effective theoretical approaches to the diagnostic decision making - how to collect and to interpret clinical information.

The first member of the Diagnostic Trainer family is the Diagnostic Trainer for anaemia cases - Anaemia Diagnostic Trainer. It is in regular use in Internal Disease Department of the Moscow Medical Academy. Anaemia Diagnostic Trainer simulates infinite number of different clinical cases of some 90 diseases, leading to anaemia manifestations of different types. The Anaemia Diagnostic Trainer's knowledge base expresses the points of view of the leading Russian medical experts and educators.

Diagnostic Trainer can create realistically appearing simulated clinical cases using a large number of interrelated clinical parameters (medical history, laboratory data, special investigation data, clinical signs etc.). Diagnostic Trainer's algorithms guarantee the limitless variety of clinical cases for all diseases involved and intellectual behaviour of the system itself. Diagnostic Trainer always «knows» the right answer, i.e. the diagnose corresponding to the clinical case offered to user, thus it can effectively respond to the user's actions. The system's responses achieve thought-provoking learning experience. User interface of the Diagnostic Trainer is based on the hierachic menu based on the hypertext technology. If the user needs any help or advice, he can ask for system's opinion on the effective diagnostic strategy and it's interpretation of collected information.

Anaemia Diagnostic Trainer runs under MS DOS on all IBM compatibles with 640 Kb RAM or more. Anaemia Diagnostic Trainer installation needs only 360 Kb memory on Hard or Floppy disk.

Diagnostic Trainer never forces You, but always fixes your attention on the essential points and possible contradictions between the patient's data and your problem solving strategy. We have documentation in Russian, in Spanish and in English.

Mikhail Kreines,

Ph.D., the Head of Moscow Center for New Information Technology in Medical Education

Russia, 119435, Moscow, B.Pirogovskaya str., 2-6,  
Moscow Medical Academy, Moscow Center  
for New Information Technology in Medical Education  
E-mail:Nnovik@sms.ccas.msk.su

