

На правах рукописи

Кочин Дмитрий Юрьевич

**ПОСТРОЕНИЕ БАЗ ЭКСПЕРТНЫХ ЗНАНИЙ ДЛЯ
ИНТЕЛЛЕКТУАЛЬНЫХ ОБУЧАЮЩИХ СИСТЕМ**

Специальность 05.13.01
Системный анализ, управление и обработка информации

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Москва – 2006

Работа выполнена в Институте системного анализа Российской академии наук

Научный руководитель – доктор технических наук
профессор Петровский Алексей Борисович

Официальные оппоненты – доктор технических наук, профессор
Стефанюк Вадим Львович

кандидат физико-математических наук
Янишевский Игорь Михайлович

Ведущая организация – Институт программных систем
Российской академии наук

Защита состоится «20» ноября 2006 г. в 11⁰⁰ час.
на заседании специализированного совета Д.002.086.02 при Институте
системного анализа РАН по адресу:
Москва, 117312, Проспект 60 лет. Октября, д. 9.

С диссертацией можно ознакомиться в библиотеке ИСА РАН.

Автореферат разослан «19» октябре 2006 г.

Ученый секретарь
диссертационного совета
Д.002.086.02



д.т.н. А.И. Пропой

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность работы. Стремительный рост знаний, создаваемых человеком, делает все более актуальными задачи обучения и переобучения. Во многих важных областях человеческой деятельности, путь развития от новичка до опытного специалиста занимает значительное время. Поэтому актуальной является задача поиска путей более быстрого и более эффективного обучения, а также предоставления возможности неограниченному числу начинающих специалистов овладеть опытом лучших мастеров своей профессии – экспертов.

Доступность персональных компьютеров делает возможным их повсеместное использование в учебных заведениях и для самостоятельной подготовки, однако недостаток соответствующих обучающих программных систем и методов их разработки не позволяет решить эту задачу в полной мере. Одной из причин такого положения вещей являются значительные теоретические трудности в построении интеллектуальных обучающих систем (ИОС), основанных на знаниях опытных специалистов для так называемых слабоструктурируемых предметных областей, в частности, медицинской диагностики, где доминирующими факторами при принятии решений являются профессиональный опыт и интуиция.

Эффективность ИОС в первую очередь определяется объемом и качеством содержащихся в них формализованных экспертных знаний, как декларативных (теоретических), так и процедуральных (практических навыков). В то же время процесс получения этих знаний от экспертов остается одной из наиболее сложных и актуальных задач при построении как ИОС, так и обычных экспертных систем, являясь так называемым «узким местом» при создании систем, основанных на знаниях. На необходимость создания специальных методов выявления экспертных знаний и их теоретического обоснования указывается в работах многих отечественных и зарубежных авторов: Л. Заде, Д. Канемана, О.И. Ларичева, Э.В. Попова, Д.А. Поспелова, Б. Руа, Г. Саймона, В.Л. Стефанюка, А. Тверски, Э. Фейгенбаума, В.К. Финна, Р. Хоффмана и др.

Трудности выявления экспертных знаний во многом обусловлены подсознательным характером процедуральных знаний, когда эксперт не может вербализовать полную и непротиворечивую систему правил, которую он использует при принятии решений. В совокупности с повышенными требованиями к точности и полноте знаний, которые предъявляются к экспертным знаниям для построения на их основе ИОС, эта проблема делает актуальной разработку методик выявления не только полных и непротиворечивых систем экспертных решающих правил, но и

методик повышения достоверности извлеченных знаний, минимизирующих ошибки эксперта.

Цели и задачи исследования. Целью диссертации является развитие методов и программных средств построения экспертных баз знаний для их использования в компьютерных обучающих системах для слабоструктурируемых предметных областей, где определяющими являются профессиональный опыт и интуиция эксперта при отсутствии объективных моделей проблемной ситуации.

Для достижения поставленных целей были сформулированы и решены следующие задачи:

- разработка алгоритма построения полных и непротиворечивых баз экспертных знаний для слабоструктурируемых предметных областей;
- разработка методов формального представления экспертных решающих правил;
- разработка методики извлечения достоверных экспертных знаний, минимизирующей неточность информации и ошибки эксперта;
- разработка программных средств, реализующих предложенные методы и алгоритмы.

Методы исследования. В работе используются методы теории принятия решений, теории графов, теории информации, искусственного интеллекта, теории построения алгоритмов и систем.

Результаты, выносимые на защиту.

- Алгоритм решения задачи порядковой классификации многопризнаковых объектов, имеющий широкую область применимости и высокую эффективность опроса эксперта.
- Метод и алгоритмы формализованного представления экспертных решающих правил.
- Методика извлечения достоверных экспертных знаний.
- Программная реализация предложенных методов и алгоритмов.

Научная новизна работы. В диссертации получены следующие новые научные результаты:

- Разработан новый метод КЛАРА (КЛАссификация Реальных Альтернатив) построения полных и непротиворечивых баз экспертных знаний для слабоструктурируемых предметных областей, позволяющий повысить эффективность опроса эксперта и применимый к широкому кругу задач экспертной классификации.
- Формализовано понятие экспертного решающего правила и построен алгоритм их выявления.
- Введено понятие зоны неустойчивых экспертных знаний и предложена методика выявления этой зоны.
- Предложена общая методика построения достоверных баз экспертных знаний, минимизирующая случайные ошибочные суждения эксперта,

а также выявляющая суждения эксперта с низкой степенью уверенности.

- Построены базы знаний по диагностике острого инфаркта миокарда и расслаивающей аневризмы аорты. На их основе построены соответствующие интеллектуальные обучающие системы.

Практическая ценность работы. Метод КЛАРА реализован в виде программной системы, с помощью которой можно в короткий срок создавать полные и непротиворечивые базы экспертных знаний для построения ИОС и экспертных систем для задач порядковой классификации. Метод КЛАРА эффективно работает в случаях разреженных пространств объектов, подлежащих классификации, что особенно востребовано в задачах медицинской диагностики, геологии, технической диагностики, банковского дела и др.

На основе предложенной методики построения достоверных баз экспертных знаний для интеллектуальных обучающих систем с использованием реализованной системы КЛАРА построены система ОСДИМ (Обучающая Система по Диагностике острого Инфаркта Миокарда) для обучения диагностике острого инфаркта миокарда и система обучения диагностике расслаивающей аневризмы аорты, которые применяются в Учебно–научном центре по внедрению передовых медицинских технологий городской клинической больницы (ГКБ) им. С. П. Боткина, г. Москва, а также в ГКБ № 33 им. проф. А.А.Остроумова, г. Москва.

Обоснованность и корректность предложенных методов и алгоритмов подтверждены на примере практического построения систем обучения медицинской диагностике острого инфаркта миокарда и расслаивающей аневризмы аорты.

Реализация результатов. Результаты диссертации использованы при выполнении проекта 2.21 программы фундаментальных исследований президиума РАН «Математическое моделирование и интеллектуальные системы» (2001-2005 годы), проекта 1.2 программы фундаментальных исследований ОИТВС РАН «Фундаментальные основы информационных технологий и систем», проектов 01-01-00514, 03-01-06283, 04-01-00290 Российского фонда фундаментальных исследований, гранта Президента Российской Федерации для поддержки ведущих научных школ НШ1964.2003.1.

Апробация работы. Результаты диссертации и материалы исследований докладывались и обсуждались на восьмой национальной конференции по искусственному интеллекту с международным участием (КИИ'2002), Коломна, Россия, 7-12 октября 2002 г.; Международных научных конференциях «Интеллектуализация обработки информации» (ИОИ-2006) Крым, Алушта, 4-11 июня 2006 г. и (ИОИ-2004) Крым,

Алушта, 14-19 июня 2004 г.; 20-ой европейской конференции по исследованию операций «Исследование операций и управление электронными службами», Родос - Греция, 4-7 июля, 2004 г.; первой международной конференции «Системный анализ и информационные технологии» (САИТ-2005), Переславль-Залесский, 12-16 сентября 2005 г.; 10-й конференции по искусственному интеллекту в медицине (АИМЕ 05) 23 - 27 июля 2005 г., Абердин, Шотландия; научных семинарах ИСА РАН (2001-2006 годы)

Публикации. По материалам диссертации опубликовано 10 работ (в том числе 2 публикации в ведущих рецензируемых научных изданиях, рекомендованных ВАК, 5 публикаций в трудах научных конференций).

Личный вклад соискателя. Результаты, выносимые на защиту, получены автором самостоятельно. Личный вклад соискателя в совместно опубликованных работах составляет 1.9 п.л.

Структура и объем работы. Диссертация состоит из введения, четырех глав, заключения, списка литературы (69 наименований), и приложения. Общий объем работы составляет 115 страниц, включая 8 таблиц и 15 рисунков.

СОДЕРЖАНИЕ РАБОТЫ

Во введении обосновывается актуальность работы, формулируется ее цель, научная новизна, полученные результаты и структура диссертации.

В первой главе рассматриваются проблемы извлечения и последующей передачи экспертных знаний с помощью интеллектуальных обучающих систем с точки зрения информационного подхода – одного из основных направлений в современной когнитивной психологии. Отмечается принципиальное различие в подходах к построению ИОС для хорошо структурируемых областей знаний, к которым относятся, например, задачи математики, физики, программирования, и слабоструктурируемых областей, таких как медицинская диагностика.

В отличие от хорошо структурируемых областей знаний, где основные усилия исследователей при построении ИОС сфокусированы на построении моделей обучаемого и процесса обучения, основной проблемой построения ИОС для слабоструктурируемых областей остается проблема извлечения экспертных знаний, подход к решению которой определяет все остальные аспекты построения каждой конкретной обучающей системы.

Рассмотрены особенности мышления экспертов, порождающие основную проблему построения ИОС для слабоструктурируемых областей – невербализуемость знаний, когда эксперты не могут сформулировать правила, которыми они пользуются при принятии решений. Те немногие

правила, которые удается получить в явном виде, охватывают лишь наиболее простые задачи из тех, которые эксперт способен правильно решить. Тем самым можно считать, что, по крайней мере, часть правил распознавания, которые формируются у экспертов в результате многолетней практики, находятся на подсознательном уровне.

Различаются два типа подсознательных навыков (процедуральных знаний): подсознательные навыки первого типа, первоначально имевшие явное, декларативное представление, применение которых стало автоматическим в результате продолжительной практики, и подсознательные навыки второго типа, которые даже изначально не были представлены в явном, декларативном виде.

На основе анализа наиболее известных обучающих систем, основанных на экспертных знаниях (GUIDON, QMR, ILIAD, TRAINER, ACT-R, QUEST) делается вывод, что эффективность обучающей системы в слабоструктурируемой области зависит, главным образом, от полноты и достоверности базы экспертных знаний. Все недостатки базы знаний, в том числе неполнота, наличие противоречий и возможные несовпадения с решениями эксперта, непосредственно отражаются на качестве обучающей системы и, следовательно, на качестве обучения. При этом главную сложность представляет именно извлечение процедуральных знаний, практических навыков, потому что эти знания носят подсознательный характер и не могут быть вербализованы и переданы экспертом явно.

Для построения баз знаний в слабоструктурируемых областях хорошо применим подход порядковой экспертной классификации. Формальная постановка задачи порядковой классификации многопризнаковых объектов состоит в следующем:

Дано:

1. G – свойство, отвечающее целевому критерию задачи (наличие и степень тяжести заболевания, критичность неисправности в технической системе, ценность кредитного проекта и т.д.).
2. $K = \{K_1, K_2, \dots, K_N\}$ – множество критериев (признаков), по которым оценивается каждый объект исследования.
3. $S_q = \{k_1^q, k_2^q, \dots, k_{\omega_q}^q\}$, $q = 1, \dots, N$ – шкала оценок по критерию K_q ; ω_q – число градаций на шкале критерия K_q . Оценки $k_1^q, k_2^q, \dots, k_{\omega_q}^q$ упорядочены по убыванию характерности для свойства G , то есть, на каждом множестве S_q определено рефлексивное антисимметричное транзитивное отношение Q_q (необязательно связное) такое, что $(k_i^q, k_j^q) \in Q_q \Rightarrow i \leq j$.
4. $Y = S_1 \times S_2 \times \dots \times S_N$ – декартово произведение шкал критериев, которое определяет пространство гипотетически возможных

состояний объектов, подлежащих классификации. Каждый объект описывается набором оценок по критериям K_1, \dots, K_N и представляется в виде векторной оценки $y \in Y$, где $y = (y_1, y_2, \dots, y_N)$, y_q – одна из оценок из множества S_q .

5. $L = |Y| = \prod_{q=1}^N \omega_q$ – мощность множества Y .
6. $Y^* \subseteq Y$ – множество объектов, которые требуется классифицировать. Это множество назовем *допустимым*, а составляющие его объекты *допустимыми*.
7. $C = \{C_1, C_2, \dots, C_M\}$ – множество классов решений, упорядоченных по убыванию выраженности свойства G , то есть, на множестве C определено рефлексивное антисимметричное транзитивное отношение Q_C такое, что $(C_i, C_j) \in Q_C \Leftrightarrow i \leq j$.

Вводится бинарное отношение строгого доминирования векторных оценок:

$$P = \left\{ (x, y) \in Y \times Y \mid \forall q = 1 \dots N (x_q, y_q) \in Q_q \text{ и } \exists q_0 : x_{q_0} \neq y_{q_0} \right\} \quad (1)$$

которое для краткости может записываться как $(x, y) \in P \Leftrightarrow x P y \Leftrightarrow x \succ y$.

Требуется: на основе знаний эксперта построить разбиение множества допустимых объектов Y^* на M непересекающихся классов решений C_i ($C_i \cap C_j = \emptyset \forall i \neq j$, $\cup C_i \supseteq Y^*$) так, чтобы выполнялось свойство непротиворечивости классификации:

$$\forall x, y \in Y \text{ таких, что } x \in C_i, y \in C_j, (x, y) \in P \Rightarrow i \geq j. \quad (2)$$

Сделан обзор подходов к решению поставленной задачи, которые можно разбить на четыре группы. К первой группе относятся подходы к построению экспертных баз знаний, когда эксперт явно формулирует логические рассуждения и заключения, которые ложатся в основу базы знаний. Недостаток этого подхода в том, что значительная часть экспертных навыков, как отмечалось ранее, находится на подсознательном уровне и не может быть корректно извлечена в явном виде.

В рамках подходов второй группы, к которым примыкает известный метод ELECTRE TRI, эксперт определяет либо субъективные вероятности влияния отдельных признаков на конечное решение, либо задает свои предпочтения в виде балльных оценок. Недостаток этих подходов в том, что согласованное назначение вероятностных, балльных оценок является слишком сложной задачей для системы переработки информации человека, что обычно ведет к ошибкам и противоречиям.

К третьей группе примыкают подходы, где эксперт формирует наборы решенных примеров, на основе которых различные алгоритмы машинного обучения, например, нейронные сети, строят базу экспертных знаний.

Точность имитации эксперта сильно зависит от полноты обучающей выборки, что не всегда удаётся обеспечить на практике.

Наконец, в четвертой группе представлены подходы, в которых эксперт последовательно делает заключения относительно предъявляемых ему описаний объектов из рассматриваемой предметной области. Стратегия выбора объектов для предъявления эксперту гарантирует построение согласованной и полной базы знаний. К этой группе относится и подход вербального анализа решений (ВАР). Выявление знаний таким способом максимально близко к реальной профессиональной деятельности эксперта, поэтому полученное знание содержит, в том числе, и подсознательный пласт экспертного знания, который может быть далее проанализирован и формализован.

Приводится обзор методов ВАР, как наиболее подходящих для построения баз экспертных знаний для интеллектуальных систем обучения, описаны методы ОРКЛАСС, КЛАНШ, STEPCLASS и ЦИКЛ, указаны их достоинства и недостатки. Среди недостатков отмечаются высокая вычислительная сложность, узкая область применения или неэффективная работа с разреженными пространствами объектов, когда допустимые объекты Y^* составляют малую часть объектов полного пространства Y . Последний недостаток существенен для задач медицинской диагностики, когда множество комбинаций симптомов исключаются из рассмотрения ввиду невозможности такого сочетания в реальности.

Далее рассматривается проблема достоверности извлеченных знаний. Указывается, что искажение знаний может происходить из-за ошибок эксперта, связанных с рядом причин, в том числе с усталостью, редкой встречаемостью случая на практике, недостаточностью информации в формализованной задаче. Различаются два вида экспертных ошибок – грубые ошибки, приводящие к логическим противоречиям, то есть нарушению условия (2) и ошибки, не приводящие к логическим противоречиям, но искажающие границы классов решений. Ошибки второго вида не могут быть выявлены автоматически, и рассмотренные методы не предоставляют средств для их обнаружения.

Описана гипотеза об использовании экспертами при решении профессиональных задач подсознательных решающих правил, имеющих счетно-аддитивную структуру. Скорость, с которой эксперт принимает решения, позволяет предположить, что такие правила размещаются в кратковременной памяти человека, имеющей ограниченный объем. То есть, число экспертных решающих правил должно быть невелико.

На основании проведенного анализа делаются выводы и формулируются цели исследования.

Во второй главе описывается новый алгоритм КЛАРА решения задачи экспертной классификации, имеющий широкую область применения, низкую вычислительную сложность и эффективно применимый для случаев разреженных пространств классифицируемых объектов. Различаются два вида задач экспертной классификации. В задаче номинальной экспертной классификации необходимо определить для каждого элемента множества Y наличие или отсутствие некоторых свойств, т. е. решить задачу покрытия. В задаче порядковой экспертной классификации необходимо определить степень выраженности единственного свойства G , т. е. решить задачу разбиения. В работе используется такая декомпозиция задачи экспертной классификации, которая позволяет решать оба вида задач с помощью единого алгоритма.

Вводятся следующие определения.

Определение 1. Альтернативы $x, y \in Y$ называются *сравнимыми*: $x \sim y$, если $x \succeq y$ или $y \succeq x$, иначе они называются *несравнимыми*: $x \not\sim y$.

Любые две альтернативы, принадлежащие одному классу, либо находятся в отношении доминирования, либо несравнимы, следовательно, в каждом классе можно выделить подмножества недоминируемых и недоминирующих альтернатив.

Определение 2. Подмножество альтернатив $B^U(C_n)$ класса C_n называется *верхней границей* этого класса, если $\forall x \in C_n \exists y \in B^U(C_n)$ такой, что $y \succeq x$ и $\forall x, y \in B^U(C_n), x \neq y \Rightarrow x \not\sim y$.

Определение 3. Подмножество альтернатив $B^L(C_n)$ класса C_n называется *нижней границей* этого класса, если $\forall x \in C_n \exists y \in B^L(C_n)$ такой, что $x \succeq y$ и $\forall x, y \in B^L(C_n), x \neq y \Rightarrow x \not\sim y$.

Определение 4. Альтернатива $x \in Y$ *непосредственно доминирует* альтернативу $y \in Y$, если $x \succ y$ и $\nexists z \in Y: x \succ z \succ y$.

Определение 5. Альтернатива $x \in Y$ *непосредственно доминируется* альтернативой $y \in Y$, если $x \prec y$ и $\nexists z \in Y: x \prec z \prec y$.

Множество альтернатив, непосредственно доминирующих альтернативу x , обозначается как $Z^U(x)$, а множество непосредственно доминируемых – как $Z^L(x)$.

Определение 6. *Орграфом* $G(Y, E)$ *непосредственного доминирования альтернатив* называется ориентированный ациклический граф, в котором множество вершин есть множество альтернатив Y , а в множество дуг $E \subseteq Y \times Y$ состоит из элементов (x, y) , в которых альтернатива $x \in Y$ *непосредственно доминирует* альтернативу $y \in Y$.

Определение 7. Последовательность альтернатив $w = \langle y_1, y_2, \dots, y_l \rangle$, в которой $y_{i+1} \in Z^L(y_i)$, $1 \leq i \leq (l-1)$, называется *цепью*. Число $L(w) = l$

альтернатив цепи w называется *длиной* цепи. Отдельная альтернатива является цепью длины 1.

Алгоритм КЛАРА основан на идее дихотомии цепей альтернатив, начиная с цепи максимальной длины, впервые использованной в алгоритме ДИФКЛАСС, а затем в КЛАНШ, и адаптирует эту идею на случай разреженных пространств Y . Кроме того, в алгоритме КЛАРА используется новая идея адаптивной дихотомии, позволяющая быстрее находить границы классов решений и ускоряющая классификацию.

Задача порядковой классификации может быть решена путем предъявления эксперту всех допустимых альтернатив, число которых равно $\text{card } Y^*$, для получения искомого разбиения на классы. Использование отношения доминирования (1) и условия непротиворечивости (2) позволяет существенно сократить число непосредственно предъявляемых альтернатив и тем самым ускорить процедуру построения классификации. Действительно, согласно этим условиям, альтернатива x , имеющая более характерные для свойства G оценки, чем y , то есть $x \succ y$, не может быть отнесена к классу с менее выраженным свойством G . Таким образом, благодаря использованию информации об уже классифицированных альтернативах для классификации оставшихся, становится возможным сократить число предъявляемых альтернатив.

Приводится общая схема алгоритма КЛАРА.

1. В начале классификации коэффициенты дихотомии d_i полагаются равным $\frac{1}{2}$. Коэффициенты d_i представляют собой отношение, в котором цепи, проходящие через множество классифицируемых альтернатив Y , делятся границами классов C_i и C_{i+1} .
2. Орграф $G(Y, E)$ доминирования альтернатив может иметь несколько компонент связности, поэтому последовательно исследуются все допустимые, но еще не классифицированные альтернативы из множества Y^* . Очередная выбранная альтернатива x_s называется *исходной*.
3. В текущей компоненте связности (к которой принадлежит альтернатива x_s) орграфа $G(Y, E)$ строится цепь w_{max} альтернатив максимальной длины, проходящая через исходную альтернативу x_s и содержащая максимальное число ещё неклассифицированных альтернатив из допустимого множества Y^* .
4. Поскольку классы $\{C_n\}$ упорядочены по качеству, границы между классами на цепи строятся последовательно, отделяя класс большего качества C_n от класса меньшего качества C_{n+1} .

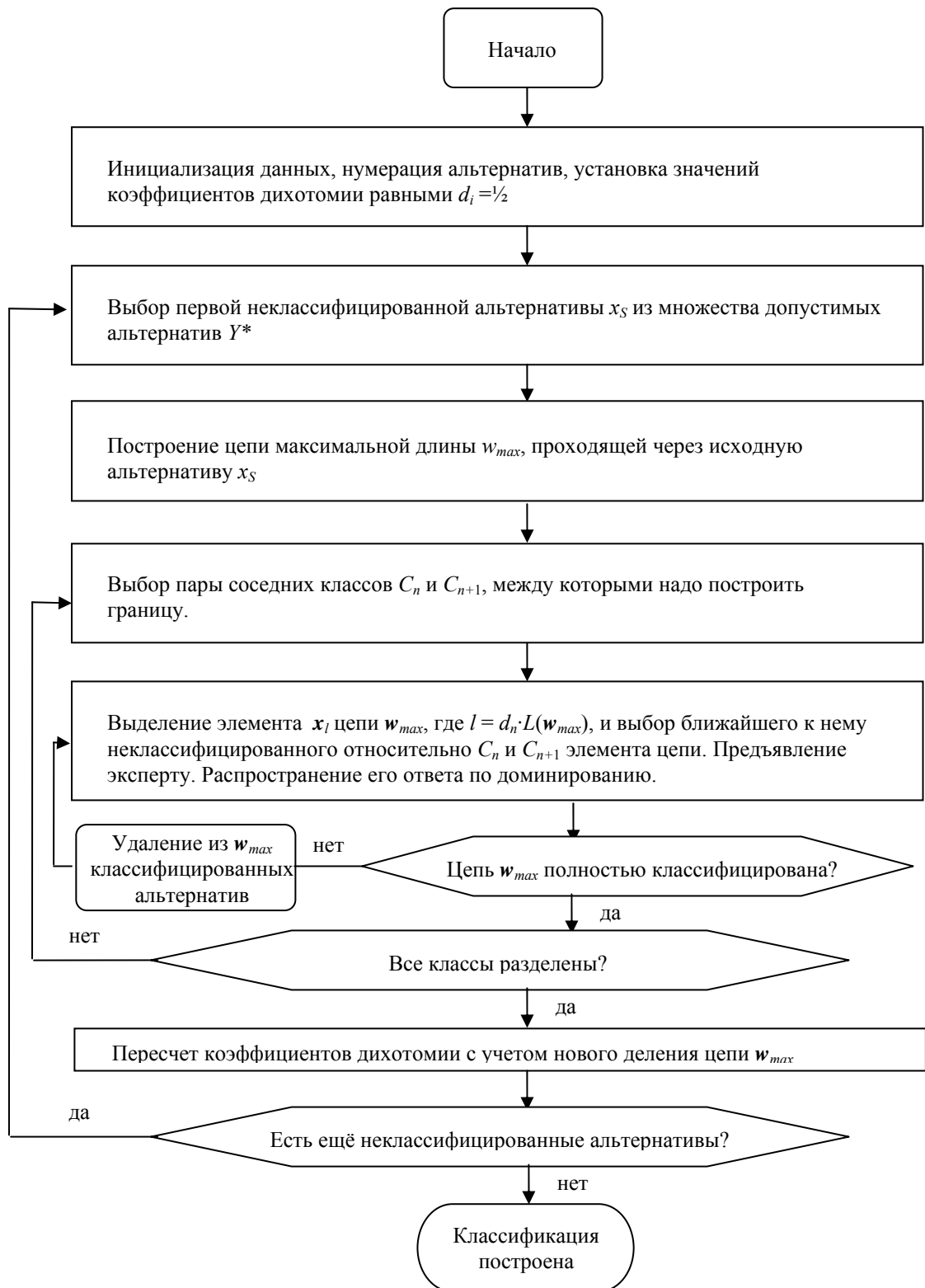


Рис. 1. Блок-схема алгоритма КЛАРА

5. Для предъявления эксперту выделяется элемент x_l цепи w_{max} , где $l = d_n \cdot L(w_{max})$, причем если альтернатива x_l оказалась недопустимой или уже классифицированной, то в качестве нового x_l берется неклассифицированный допустимый элемент цепи с ближайшим индексом.
6. Эксперту предъявляется допустимая альтернатива x_l цепи w_{max} и проводится распространение его решения на максимально возможное число элементов, принадлежность которых к классам C_n и C_{n+1} остается неопределенной.
7. Если w_{max} еще содержит допустимые неклассифицированные элементы, то продолжается дихотомия цепи w_{max} , которая завершается, когда все входящие в нее допустимые альтернативы оказываются прямо или косвенно классифицированы относительно классов C_n и C_{n+1} . В противном случае, на цепи ищется следующую границу между классами (производится возврат на шаг 4). Если же цепь классифицирована относительно всех классов, то для каждого класса находится индекс k в цепи w_{max} , где происходит смена класса с C_n на C_{n+1} . Полагается $d_{nw} = k/L(w_{max})$. На каждом последующем шаге d_n есть среднее арифметическое всех посчитанных до этого d_{nw} .
8. Цикл выполняется до тех пор, пока все допустимые альтернативы из допустимого множества Y^* не окажутся классифицированными относительно этой пары классов.

Описываются все входящие в общий алгоритм вспомогательные процедуры, в том числе распространение по доминированию, процесс адаптивной дихотомии максимальной цепи, поиск максимальной цепи. Идея адаптивной дихотомии цепи состоит в том, что для поиска границы классов на ней применяется не деление пополам, а деление в определенной пропорции, которая подсчитывается исходя из истории деления цепи на классы.

Процедура поиска максимальной цепи основана на процедуре, аналогичной методу КЛАНШ, и адаптирована для поиска цепи, содержащей максимальное число допустимых ещё не классифицированных альтернатив. Доказана математическая корректность этой процедуры.

Показано, что если каждый критерий характеризуется только двумя оценками, упорядоченными по характерности, а число классов решений фиксировано, то вычислительная сложность процедуры поиска очередной альтернативы для предъявления эксперту, которая представляет собой итерацию главного цикла алгоритма КЛАРА, не превышает $O(|Y| \cdot \log_2 |Y|)$. Эта оценка совпадает с аналогичным показателем алгоритмов КЛАНШ и ДИФКЛАСС.

Рассмотрен механизм проверки согласованности ответов эксперта с условием непротиворечивости классификации (2), выявления и оперативного устранения возникающих противоречий по мере построения полной классификации.

Проведено сравнение алгоритма КЛАНШ с другими алгоритмами порядковой классификации по следующей методике сравнения.

Исследования поведения человека при решении задач классификации показывают, что его решения могут хорошо аппроксимироваться решающими правилами, имеющие структуру двухуровневого дерева, корень которого соответствует конъюнкции значений некоторых наиболее важных (основных) критериев, а концевые вершины суть сочетания характерных для рассматриваемого класса оценок по менее важным (дополнительным) критериям. Например, альтернатива y принадлежит классу C_1 , если она имеет 1-оценку по первому и второму критерию и не менее двух 1-оценок по остальным критериям.

Пусть даны два класса решения: C_1 и C_2 . Зафиксируем N – число двоичных критериев. Обозначим через r – число основных критериев, а через t – число 1-оценок по дополнительным критериям, причем $1 \leq t < (N-r)$. Если альтернатива y имеет 1-оценки по всем r основным критериям и не менее t 1-оценок по дополнительным, то она принадлежит классу C_1 , в противном случае – классу C_2 . Сформулированное правило используется при статистическом моделировании работы алгоритма в качестве решающего правила для подсчета числа обращений к эксперту, необходимого для построения полной классификации. Будем называть средним числом обращений $Q(R, r, t)$ алгоритма R среднее число обращений данного алгоритма среди всевозможных вариантов выбора r основных критериев, число которых равно C_M^r .

Число обращений к эксперту никогда не может быть меньше суммы числа элементов нижней границы первого класса решений $B^L(C_1)$ и верхней границы второго класса $B^U(C_2)$. Поэтому можно определить *эффективность* алгоритма классификации следующим образом:

$$E(R, r, t) = \frac{|B^U(C_1)| + |B^L(C_2)|}{Q(R, r, t)};$$

Приводятся данные и графики, отражающие среднее число обращений к эксперту, а также эффективность сравниваемых алгоритмов при решении задач классификации с различными типами границ между двумя классами решений. График эффективности алгоритма КЛАРА при 11 двоичных признаках показан на Рис. 2:

Алгоритм КЛАРА

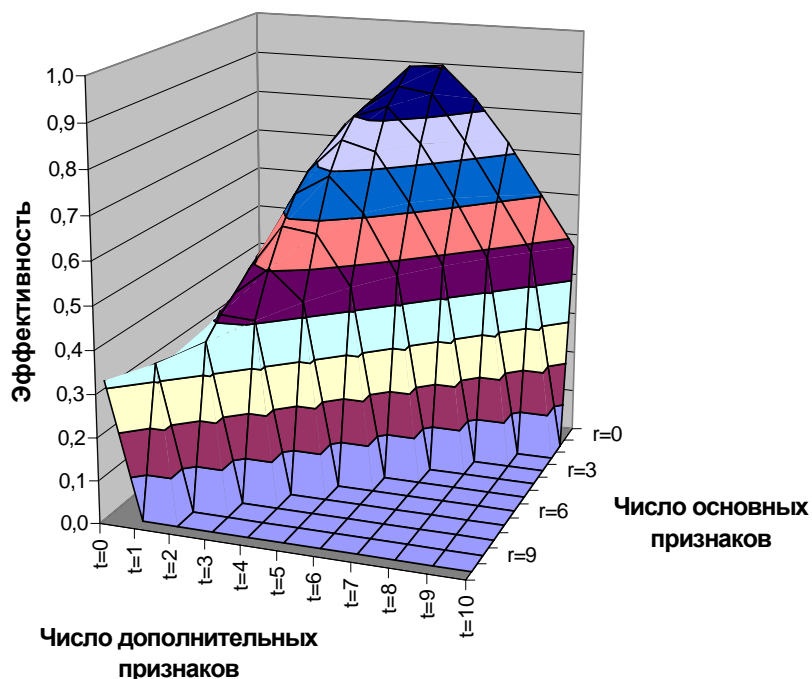


Рис. 2: Эффективность алгоритма КЛАРА при 11 двоичных признаках.

Для однозначного сравнения алгоритмов определим *среднюю эффективность* $E_{cp.}(R, N)$ алгоритма классификации R на N двоичных критериях как его эффективность $E(R, r, t)$, усредненную по всевозможным сочетаниям r и t , где $r = 0, \dots, (N-1)$ и $t = 0, \dots, (N-1)$ за исключением пары $(r = 0, t = 0)$.

Численное моделирование работы алгоритма КЛАРА по общей методике сравнения алгоритмов классификации показало, что по средней эффективности КЛАРА превосходит алгоритм КЛАНШ, имеющий такую же область применимости, и уступает алгоритмам ДИФКЛАСС и ЦИКЛ, имеющим более узкую область применимости.

В третьей главе рассматривается проблема извлечения достоверных экспертных знаний. Результатом работы всякого метода порядковой классификации являются границы классов, по которым любую альтернативу из классифицированного множества можно отнести к одному из классов. Путем анализа получившихся границ классов возможно выявить правила, которые породили данную классификацию. Например, простое решающее правило «к классу C_1 относятся альтернативы, у которых не менее трех высших оценок» порождает следующую классификацию со следующей нижней границей первого класса: $V^L(C_1) = \{22111, 21211, 12211, 21121, 12121, 11221, 21112, 12112, 11212, 11122\}$. Легко заметить, что все элементы этой границы суть все возможные перестановки из трех единиц и двух двоек, и её можно

записать одним выражением $P_5^{3(1),2(2)}$. Такая запись весьма близка к исходному смыслу решающего правила, но представлена в более формализованном виде.

Приводится формальная постановка задачи описания границы класса в виде набора решающих правил. Вводится формальная запись

$$ab^{***} + P_n^{k_1[x_1] \dots k_m[x_m]}, \text{ кроме } \{abcde, \dots, abpqr\}, \quad (3)$$

которая называется *шаблоном* правила. Шаблон правила состоит из трех частей.

1. Первую часть шаблона правила, то есть ab^{***} , назовём *фиксированной частью* правила. Фиксированная часть задаёт общие оценки по заданным критериям для всего множества альтернатив, описываемых правилом. По критериям, где в фиксированной части явно указаны оценки, эти же оценки стоят во всех альтернативах правила. По критериям, на месте которых указаны звездочки, у альтернатив описываемого правилом множества стоят перестановки оценок, задаваемые второй частью правила.
2. Вторую часть шаблона $P_n^{k_1[x_1] \dots k_m[x_m]}$ назовём *перестановочной частью* правила. Она задаёт параметры перестановок, которые будут осуществляться на местах, помеченных *, и соответствует сочетаниям значений дополнительных признаков. Здесь n равно числу звёздочек, k_i – числу оценок x_i , участвующих в перестановке.
3. Третья часть шаблона представляет собой перечисление исключений – альтернатив, которых не хватает до полной перестановки.

Ставится задача описания некоторой заданной совокупности альтернатив B минимальным числом правил вида (3) так, чтобы каждая альтернатива попадала ровно в одно правило. Требование минимальности разложения основывается на гипотезе, что решающие правила содержатся в кратковременной памяти эксперта, имеющей ограниченный объём.

Вводятся определения.

Определение 8. Составом оценок альтернативы называется вектор $\mathbf{s} = (k_1, \dots, k_n)$, где $n = \max_q(\omega_q)$, k_i – количество i -ых оценок в альтернативе.

Определение 9. Перестановочную часть правила, порождённую составом оценок $\mathbf{s} = (k_1, \dots, k_n)$, назовём перестановочную часть $P_s = P_m^{k_1[1] \dots k_n[n]}$, где m – число ненулевых компонент вектора \mathbf{s} .

Доказаны следующие утверждения.

Утверждение 1. Правило вида (3) описывает множество альтернатив с одинаковыми составами оценок.

Утверждение 2. Множество альтернатив с одинаковым составом оценок не содержит альтернатив, находящихся в отношении доминирования (1).

Утверждение 3. Если мощность множества альтернатив с одинаковым составом $\mathbf{s} = (k_1, \dots, k_n)$ равна $\frac{m!}{k_1! \dots k_n!}$, где m – число ненулевых компонент вектора \mathbf{s} , то этому множеству соответствует правило, состоящее только из перестановочной части $P_{\mathbf{s}}$.

Утверждение 4. Если мощность множества альтернатив с одинаковым составом $\mathbf{s} = (k_1, \dots, k_n)$ меньше $\frac{m!}{k_1! \dots k_n!}$, где m – число ненулевых компонент вектора \mathbf{s} , то этому множеству соответствует правило, состоящее из перестановочной части $P_{\mathbf{s}}$, кроме {некоторое множество альтернатив}.

На основании этих утверждений решающее правило следует искать только для множества альтернатив с одинаковым составом оценок. Любое множество альтернатив с одинаковым составом оценок может быть описано правилом вида (3) с перечислением альтернатив исключений. Однако правила с большим числом исключений интереса не представляют, так как желательно найти только краткие правила. Поэтому предлагается ввести ограничение на мощность множества исключений, а именно, оно должно содержать не более k_c элементов, и k_c составляет от общего числа элементов, описываемых перестановочной частью правила, не более k_p процентов. В реализованной системе КЛАРА было принято $k_c = 3$ и $k_p = 25\%$.

Приводится ряд алгоритмов для решения задачи описания произвольного множества альтернатив B с одинаковым составом оценок в виде набора решающих правил.

Алгоритм 1 – поиск единственного правила.

1. Критерии, по которым у всех альтернатив множества одинаковые оценки, исключаются из рассмотрения.
2. Множество альтернатив с уменьшенным числом критериев подставляются в условие утверждения 4 с учетом ограничений на k_c и k_p .
3. Если условие не удовлетворяется, то множество альтернатив B одним правилом не записывается.
4. В противном случае, объединяя результат применения утверждения 4 с исключенными на шаге 1 критериями с фиксированными оценками, получим искомое правило.

Общий Алгоритм 2 – заключается в переборе всех возможных подмножеств множества B и применения к ним алгоритма 1 поиска единственного правила. Выбираются все варианты разбиения множества B

на минимальное число подмножеств, описываемых единственным правилом, то есть описание множества B минимальным числом правил.

Указывается, что алгоритм полного перебора подмножеств имеет очень высокую вычислительную сложность. Поэтому предлагается **алгоритм 3**, который заключается в рекурсивном разбиении множества альтернатив B на подмножества по каждому из критериев так, что в каждом подмножестве оценки по данному критерию будут общими для всех альтернатив подмножества. Алгоритм 3 применяется к подмножествам алгоритма 1 до тех пор, пока алгоритм 1 не выдаст единственное правило. Результатом работы алгоритма 3 также являются минимальные разбиения множества B из тех, которые были получены в процессе его работы. Показывается, что вычислительная сложность алгоритма 3 равна $\Theta = O(|B| \cdot N^{\log_2 |B|})$. Алгоритм 3 перебирает не все возможные разбиения. Это не является существенным недостатком, поскольку исследования показывают, что решающие правила в большинстве своём имеют иерархическую структуру, основанную на различных значениях основного признака. То есть, различные правила, входящие в иерархию, имеют различные значения основного признака.

Попытаться уменьшить неточность алгоритма 3 призван **алгоритм 4**, который является синтезом алгоритмов 2 и 3. В процессе рекурсивного разбиения множества B на подмножества для подмножеств с небольшим числом элементов этот алгоритм использует алгоритм 2, иначе – алгоритм 3.

Полученные в результате анализа границ классов решающие правила требуется проверить. Вообще говоря, минимальное разложение множества альтернатив на описывающие его правила не единственно. Чаще всего это обусловлено первоначальным выбором ключевых признаков (критериев), которые составляют фиксированную часть правила в записи (3). Алгоритм поиска решающих правил производит поиск всех минимальных разложений. Поэтому, если разложение не единственно, выбор ключевых признаков осуществляет эксперт, отдавая предпочтение одному из наборов правил.

Перепроверка состоит в повторном предъявлении наиболее «подозрительных» альтернатив эксперту. Наиболее подозрительны альтернативы, которые описываются тривиальными правилами (содержащими только фиксированную часть), и исключения. Если эксперт относит предъявленную альтернативу к другому классу, то может возникнуть три ситуации:

1. *Противоречий нет и всё классифицируемое множество осталось классифицированным.* В этой ситуации заново производится поиск правил и предъявление не вписывающихся в правила альтернатив, за исключением уже предъявленных.

2. *Противоречий не возникло, но часть альтернатив оказалась не классифицированной.* Такая ситуация возможна, когда предъявленная альтернатива позволила косвенно классифицировать некоторое множество альтернатив, находящихся с ней в отношении доминирования. После изменения её класса косвенная классификация распространяется на них иначе. В этой ситуации запускается механизм классификации, чтобы довести классификацию до конца. После окончания осуществляется переход в ситуацию 1.
3. *Возникло логическое противоречие.* В этом случае запускается обычный механизм устранения логических противоречий. После этого возможен переход в ситуацию 1 или 2.

Частой причиной ошибок эксперта является его усталость из-за большой длительности работы по классификации. Приводится схема декомпозиции больших задач классификации на несколько упрощенных. При этом результаты решения каждой упрощенной задачи описываются в виде решающих правил, которые могут быть использованы позднее при решении исходной задачи, сокращая время её решения.

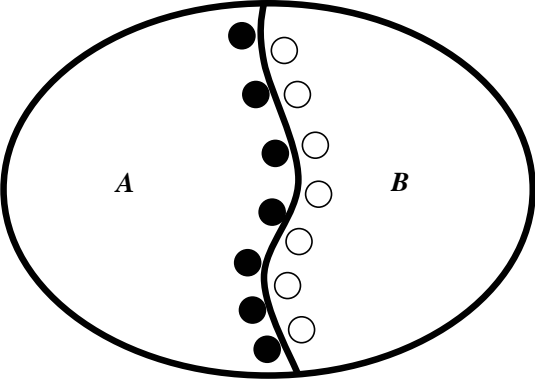
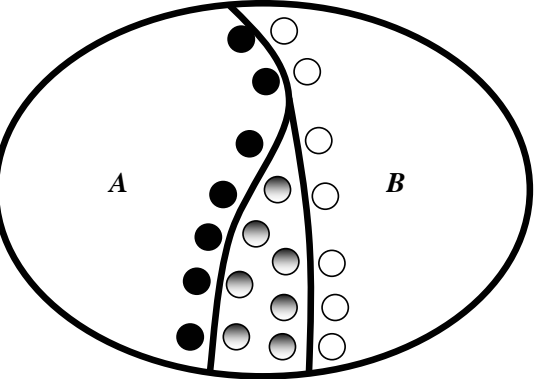
Пусть имеется N критериев с ω_q оценками по q -му критерию ($q=1, \dots, N$) и M классов решений C_1, \dots, C_M . Классы решений объединяются в два класса, обозначенные \underline{C} и \overline{C} , следующим образом: $\underline{C} = \bigcup_{i=1}^k C_i$, $\overline{C} = \bigcup_{i=k+1}^M C_i$ для некоторого k . Назовем оценки, более характерные для класса \underline{C} , «верхними», а оценки, более характерные для класса \overline{C} , «нижними». Вместо исходной задачи классификации в полном пространстве предлагается решить упрощенные задачи, в которых выбираются два класса решений \underline{C} и \overline{C} , а оценки выбираются одним из следующих способов:

1. По каждому критерию берутся две нижние оценки.
2. По каждому критерию берутся две верхние оценки.
3. По каждому критерию берется самая верхняя и самая нижняя оценки.

Приводятся результаты, показывающие, что для исходной задачи остаются верны, в первом случае – правила нижней границы класса \underline{C} , во втором случае – правила, полученные для верхней границы класса \overline{C} , и в третьем случае – правила, полученные для нижней границы \underline{C} и верхней границы класса \overline{C} .

Помимо упрощения задачи классификации и надежного сохранения промежуточных результатов в виде набора решающих правил, метод упрощенных задач в некоторых случаях может существенно сократить общее число обращений к эксперту. В работе даётся соответствующая теоретическая оценка.

Эксперт может ошибаться и тогда, когда данный конкретный случай редко встречается на практике, и у эксперта недостаточно опыта для его классификации. Такие знания называются «неустойчивыми» и не подходят для использования в ИОС. В работе выдвинута гипотеза, что эксперт неявно держит в памяти не правила разделения классов, а именно правила отнесения ситуации к конкретному классу. Предполагается, что не только нижняя граница класса A , но и верхняя граница класса B описывается небольшим числом экспертных правил. Между ними располагаются «неустойчивые знания» (см. Рис. 3 и Рис. 4).

Рис. 3. Простая задача классификации	Рис. 4. Сложная задача классификации. Присутствуют «неустойчивые знания» на границах классов
	
<ul style="list-style-type: none"> ● – Граничные элементы класса A ○ – Граничные элементы класса B 	<ul style="list-style-type: none"> ● – Граничные элементы класса A ○ – Граничные элементы класса B ◐ – Нестабильные ситуации

Была проведена серия экспериментов, подтвердивших гипотезу. На основании гипотезы становится возможным выявлять зону неустойчивых знаний эксперта.

С учетом изложенного строится общая методика извлечения достоверных экспертных знаний. Извлечение знаний происходит в диалоге с экспертом. **На первом этапе** производится структуризация предметной области: формулировка свойства G , задание экспертом множества критериев K , построение и упорядочение шкал критериев, определение экспертом множества упорядоченных классов решений C . Выделяются исключения, то есть из множества Y удаляются такие ситуации, которые, например, не встречаются в жизни по причине несочетаемости некоторых признаков. Невозможные сочетания могут исключаться как на первом, так и на втором этапе.

Второй этап методики выявления знаний – экспертная классификация, состоит в предъявлении эксперту последовательности векторных оценок. Этот этап является довольно длительной процедурой, поскольку число всевозможных сочетаний признаков обычно весьма велико. Методика предусматривает выделение из исходной задачи классификации упрощенных задач, которые получаются из исходной переходом только к двум значениям по каждому признаку. Решение упрощенных задач производится гораздо быстрее, что обеспечивает возможность получить более надежные результаты. Каждая упрощенная задача рассматривается как исходная задача с точки зрения применения данной методики, то есть, к упрощенным задачам также применяются описанные ниже третий и четвертый этапы. Полученные правила вводятся в полную задачу, и эксперт решает уже частично решенную задачу в полном пространстве.

На третьем этапе методики выполняется проверка границ классов.

На четвертом этапе полученные границы классов преобразуются в экспертные решающие правила специального вида (3) и согласуются с экспертом. Может оказаться, что правила окажутся слишком сложными. В этом случае, возможно, понадобится провести процедуру выявления зоны неустойчивых знаний, потому что сложные, «некрасивые» правила зачастую свидетельствуют о неустойчивости выявленных знаний. Для этого осуществляется возврат ко второму этапу методики.

В результате применения описанной методики получается полная и непротиворечивая база экспертных знаний в заданной предметной области. Лучше всего данная методика применима к слабоструктурированным областям, например, таким как медицина, что подтверждено многолетней практикой решения подобных задач.

В четвертой главе описываются система КЛАРА извлечения экспертных знаний, реализующая предложенный метод построения полных и непротиворечивых баз экспертных знаний; система ОСДИМ для обучения навыкам диагностики острого инфаркта миокарда (ОИМ), и система обучения диагностике расслаивающейся аневризмы аорты (РАА), построенные на основе предложенных методов.

С помощью системы КЛАРА можно задавать и изменять структуру задачи классификации, которая включает в себя:

1. Список классов решений, список критериев, списки оценок по шкале каждого критерия;
2. Орграфы доминирования оценок по критериям, задаваемые списками дуг;
3. Упорядочение классов решений по характерности для свойства G .

После задания структуры задачи классификации система КЛАРА позволяет:

1. Проводить опрос эксперта, последовательно выбирая альтернативы с последующим их предъявлением эксперту для классификации. Описание альтернативы представлено на Рис. 5.
2. Осуществлять контроль на непротиворечивость ответов эксперта.
3. Задавать явные правила классификации либо в диалоговом режиме по ходу построения классификации, либо загрузив файл с описанием правил.
4. Отображать и сохранять для дальнейшего анализа границы классов решений.
5. Представлять границы классов решений в виде правил.
6. Производить проверку границ классов решений и решающих правил путем предъявления их эксперту.
7. Строить разбиения классов решений на слои альтернатив различной степени сложности и сохранять их для дальнейшего использования в обучающей системе.

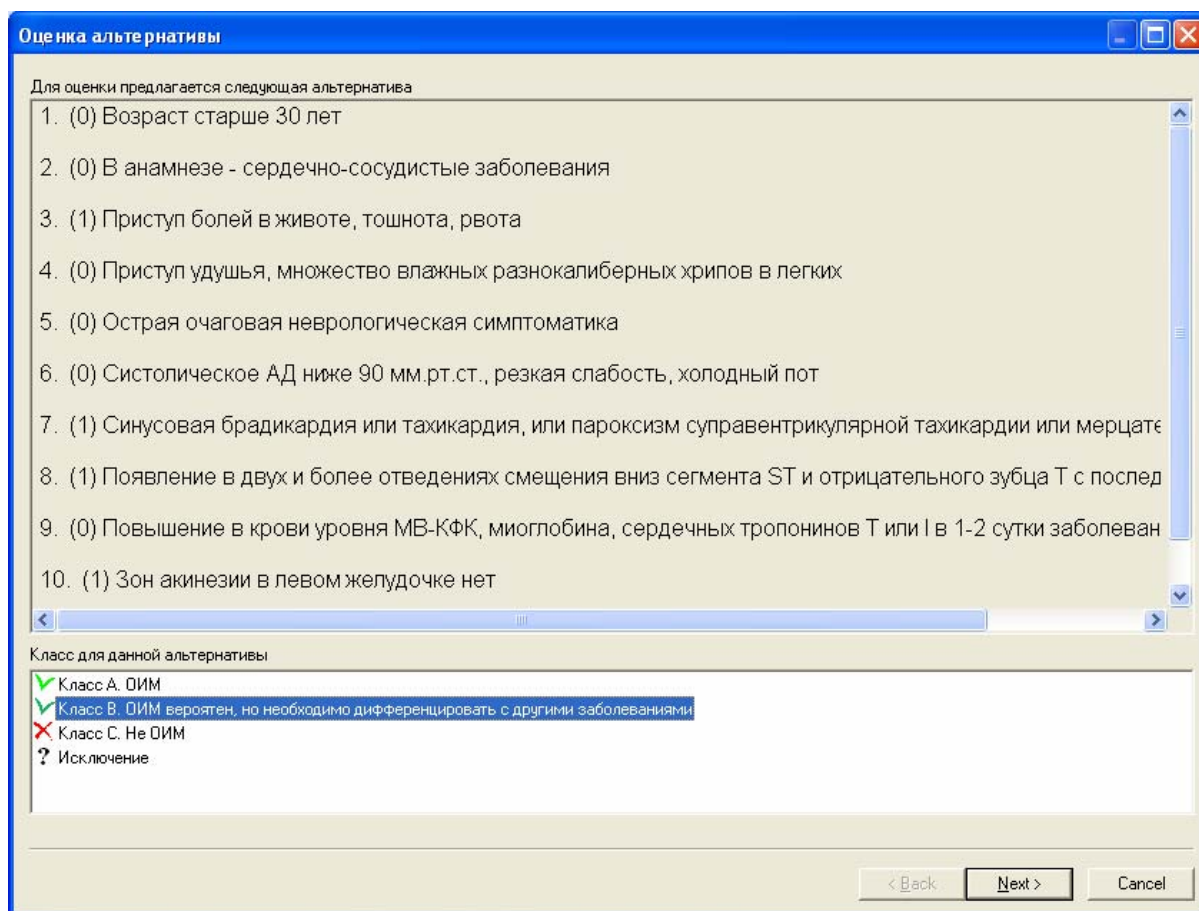


Рис. 5. Альтернатива, предъявленная эксперту для классификации в системе КЛАРА.

Задача диагностики заболевания рассматривается как задача порядковой классификации в многомерном пространстве, представляющем собой декартово произведение шкал диагностических признаков. Диагностика заболевания строится на клинических и

инструментальных признаках, каждый из которых имеет от 2 до 5 различных значений. Каждый элемент пространства может быть отнесен к одному классу решений. Для задачи диагностики расслаивающей аневризмы аорты: «РАА», «не РАА»; для задачи диагностики острого инфаркта миокарда: «ОИМ», «Для ОИМ требуется дообследование» и «не ОИМ». С помощью системы КЛАРА эксперт построил полные классификации за 4 сеанса работы по 3 часа с небольшими перерывами, ответив менее чем на 600 вопросов системы при общей суммарной размерности пространства свыше 10000 альтернатив. В первые два сеанса решались упрощенные задачи.

На основе проведенного анализа границ классов решений с помощью предложенного алгоритма был сформулирован набор решающих правил, с высокой точностью аппроксимирующих решения эксперта, и на их базе были построены решающие правила для диагностики рассматриваемых заболеваний. Эти правила были положены в основу обучающих систем, предназначенных для обучения начинающих врачей навыкам диагностики острого инфаркта миокарда и расслаивающей аневризмы аорты.

Функции обучающей системы ОСДИМ включают:

1. Обучение декларативным знаниям о данной предметной области.
2. Обучение процедуральным знаниям (практическим навыкам диагностики).
3. Ведение протокола обучения процедуральным знаниям.

Обучение процедуральным знаниям предполагает предварительное изучение декларативных знаний, то есть постановки рассматриваемой задачи диагностики, предлагаемых шагов ее решения, используемых диагностических признаков, их характерных значений для рассматриваемых заболеваний. Система обучения декларативным знаниям представляет собой электронный вариант краткого учебного пособия, после прочтения которого обучаемому предлагается решить ряд задач для проверки усвоенного материала.

Переход к обучению процедуральным знаниям, т.е. искусству диагностики, возможен только после овладения необходимым объемом декларативных знаний. До перехода ко второму этапу обучения пользователь предварительно экзаменуется на задачах максимальной сложности для определения уровня его начальной подготовки.

Обучение процедуральным знаниям начинается с наиболее простых задач. При правильных ответах обучаемого сложность задач постепенно увеличивается до максимальной. Каждая задача представляет собой развернутое описание гипотетического пациента, которому соответствует некоторая альтернатива в пространстве состояний задачи классификации. Вместе с текстовыми фрагментами расшифровки ЭКГ изображаются и их графики. При неправильных ответах обучаемому немедленно

предоставляются объяснения и комментарии, поясняющие выводы эксперта при решении аналогичных задач. Предоставляемые объяснения не могут непосредственно использоваться как решающие правила при решении предъявленной задачи, но способствуют развитию клинического мышления у пользователей обучающей системы.

После прохождения курса обучения пользователь еще раз сдает экзамен на той же выборке задач максимальной сложности, в результате чего делается вывод об эффективности обучения.

Эксперименты по обучению диагностике проходили на базе Учебно-научного центра по внедрению передовых медицинских технологий при ГКБ им. С.П. Боткина. В них принимали участие клинические ординаторы Российской государственной медицинской академии постдипломного образования и молодые врачи ГКБ им. С.П. Боткина, всего более 15 человек. Курс обучения состоял из двух сеансов по 4 часа. За это время каждый из испытуемых решал в среднем около 500 задач. Если на предварительном тесте процент правильных ответов в среднем совпадал с показателями случайного выбора, то после окончания курса обучения испытуемые демонстрировали на контрольном тесте 90-100% совпадений с ответами эксперта. При этом они не могли словесно сформулировать полную систему решающих правил, которые бы описывали все их решения, т.е. сформированный навык диагностики был подсознательным. Некоторые испытуемые проходили повторный тест через неделю и показатели снижались не более чем на 5%, что говорит о закреплении навыка.

В заключении изложены основные результаты и выводы по диссертационной работе.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

1. Проведен анализ современных ИОС, основанных на экспертных знаниях, показавший, что наиболее значимыми характеристиками ИОС как средства передачи экспертных знаний являются качество базы знаний и ее соответствие психологическим аспектам принятия решений. Имеются существенные трудности в построении баз знаний для ИОС в слабоструктурируемых областях знаний, удовлетворяющих указанным требованиям.
2. В рамках подхода вербального анализа решений разработан новый метод классификации КЛАРА для построения полных и непротиворечивых баз экспертных знаний для слабоструктурируемых областей знаний, который имеет широкую область применимости.

3. Показано, что алгоритм КЛАРА превосходит по эффективности и не уступает по вычислительной сложности алгоритмам порядковой классификации КЛАНШ, ОРКЛАСС, немного отстаёт по эффективности от алгоритма ЦИКЛ, но способен решать значительно более широкий круг задач, включая эффективную применимость к случаям разреженных пространств альтернатив для классификации, а также работу с частично упорядоченными шкалами критериев.
4. Введено понятие зоны неустойчивых экспертных знаний в задачах порядковой экспертной классификации и проведены эксперименты по исследованию возможности выявления этой зоны.
5. Разработан алгоритм поиска формальных решающих правил, описывающих порядковую классификацию многокритериальных объектов.
6. Предложена общая методика построения экспертных баз знаний с повышенными требованиями к достоверности знаний, пригодных для построения на их основе интеллектуальных обучающих систем.
7. С помощью системы КЛАРА построены полные и непротиворечивые базы знаний для задач диагностики острого инфаркта миокарда и расслаивающей аневризмы аорты.
8. На основе построенных баз знаний созданы интеллектуальные обучающие системы ОСДИМ и РАА, которые используются в Учебно-научном центре по внедрению передовых медицинских технологий ГКБ им. С. П. Боткина, г. Москва, и ГКБ №33 им. А.А. Остроумова, г. Москва, для повышения квалификации молодых врачей.

В приложении приведены вспомогательный алгоритм нумерации вершин графа доминирования альтернатив с оценкой его вычислительной трудоемкости, таблицы эффективности метода решения упрощенных задач, данные численного моделирования алгоритма КЛАРА, структура рассматриваемых задач.

СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Кочин Д.Ю. Метод классификации заданного множества многокритериальных альтернатив. // Методы поддержки принятия решений. Сборник трудов Института системного анализа РАН. — М.: Эдиториал УРСС, 2001. — С. 4-18.

2. Kochin D.Yu., Larichev O.I., Kortnev A.V. Decision support system for classification of a finite set of multicriteria alternatives. // *Journal of Decision Support Systems*, 2002, 33, — С. 13-21.
3. Кочин Д.Ю., Асанов А.А. Метод выявления решающих правил в задачах экспертной классификации. // *Искусственный интеллект*, 2002, №2. — С. 20-31.
4. Кочин Д.Ю., Асанов А.А. Выявление подсознательных экспертных решающих правил в задачах многокритериальной классификации. // *Восьмая национальная конференция по искусственному интеллекту с международным участием (КИИ'2002). Труды конференции. Том 1.* — М.: Физматлит, 2002. — С. 534-544.
5. Кочин Д.Ю., Подлипский О.К. Системы неявного обучения экспертным знаниям. // *Искусственный интеллект*, 2004, №2. — С. 305-309.
6. Кочин Д.Ю., Подлипский О.К. Построение баз экспертных знаний. // *Моделирование процессов управления, Сборник научных трудов.* — М.: Московский физико-технический институт, 2004.— С. 116-123.
7. Кочин Д.Ю., Подлипский О.К. О границах классов решений в задачах экспертной классификации. // *Некоторые проблемы фундаментальной и прикладной математики.* — М.: Московский физико-технический институт, 2004.— С.103-118.
8. Kochin D., Ustinovichius L., Sliesoraitiene V., *Implicit Learning System for Teaching the Art of Acute Cardiac Infarction Diagnosis.* // *10th Conference on Artificial Intelligence in Medicine (AIME 05).* S. Miksch et al., (Eds.): LNAI 3581, 2005.— С. 395-399.
9. Кочин Д.Ю. Построение систем неявного обучения экспертным знаниям. // *Труды I-ой международной конференции «Системный анализ и информационные технологии» (SAIT-2005).* В 2 т. Т. 1.— М.: КомКнига, 2005.— С. 263-267.
10. Кочин Д.Ю. Система неявного обучения диагностике острого инфаркта миокарда. // *Методы поддержки принятия решений. Сборник трудов Института системного анализа РАН. Т.12* — М.: Едиториал УРСС, 2005. — С. 26-42.