

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УДК 004.93(043.3)+004.855(043.3)

ШУТ
Ольга Викторовна

**МОДЕЛИ И АЛГОРИТМЫ ДЛЯ ЗАДАЧ РАСПОЗНАВАНИЯ С
КОМБИНИРОВАННОЙ ОБУЧАЮЩЕЙ ИНФОРМАЦИЕЙ**

Автореферат
диссертации на соискание ученой степени
кандидата технических наук

по специальности 05.13.17 – Теоретические основы информатики

Минск, 2016

Работа выполнена в Белорусском государственном университете.

Научный руководитель – Образцов Владимир Алексеевич,
кандидат физико-математических наук, доцент,
доцент кафедры информационных систем
управления
Белорусского государственного университета.

Официальные оппоненты: Головки Владимир Адамович,
доктор технических наук, профессор,
заведующий кафедрой интеллектуальных
информационных технологий
УО «Брестский государственный технический
университет»;

Новоселова Наталья Анатольевна,
кандидат технических наук, доцент,
старший научный сотрудник лаборатории
биоинформатики ГНУ «Объединенный институт
проблем информатики НАН Беларуси».

Оппонирующая организация – УО «Белорусский государственный
университет информатики и
радиоэлектроники».

Защита состоится **30 сентября** 2016 г. в **10.00** часов на заседании совета по защите диссертаций Д 02.01.02 при Белорусском государственном университете по адресу: 220030, г. Минск, ул. Ленинградская, 8 (корпус юридического факультета), ауд. 407.

Телефон ученого секретаря – (017) 209-57-09.

С диссертацией можно ознакомиться в Фундаментальной библиотеке Белорусского государственного университета.

Автореферат разослан «24» августа 2016 г.

Ученый секретарь
совета по защите диссертаций
кандидат физ.-мат. наук доцент

Е.С. Чеб

ВВЕДЕНИЕ

Теория и практика распознавания образов привлекают внимание многих исследователей из самых различных областей – от математиков до специалистов в конкретных приложениях. Как результат, возникают различные постановки и алгоритмы решения задачи распознавания.

Важным элементом постановки прикладных задач, влияющим на выбор алгоритма и эффективность организации процесса получения решения, является способ представления обучающей информации. В качестве основных для задач распознавания можно выделить логический способ (описание объектов с использованием принципа свертки) и прецедентный (непосредственное перечисление объектов и их классов).

Логический и прецедентный способы представления, как правило, используются отдельно и независимо друг от друга. Однако существуют задачи распознавания (например, медицинская диагностика и др.), в которых по разным причинам возникает необходимость использования указанных способов представления информации одновременно.

Вопросы решения таких задач рассматривались в работах отечественных (Абламейко С.В., Закревский А.Д., Краснопрошин В.В., Образцов В.А. и др.) и зарубежных (Гренандер У., Журавлёв Ю.И., Рязанов В.В., Фор А. и др.) ученых. На основе полученных ими теоретических результатов разработан ряд компьютерных систем, ориентированных на решение практических задач распознавания и поддержки принятия решений в различных предметных областях. Однако, несмотря на важность проблематики и многолетнюю ее историю, многие вопросы, к числу которых можно отнести связь между различными видами представления информации и теоретическую аргументированность результатов решения конкретных задач, во многом остаются открытыми. Поэтому комплексное рассмотрение подобных проблем является, несомненно, актуальным для дальнейшего развития теории и практики распознавания образов.

В диссертационной работе предложен один из возможных подходов к решению задач распознавания, в которых используются различные варианты описания информации. Подход основан на объединении логического и прецедентного способов представления информации, что позволяет при решении задач распознавания использовать различные типы алгоритмов и строить комбинированное решение, сочетающее результаты, полученные этими алгоритмами.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Связь работы с крупными научными программами (проектами) и темами

Результаты, изложенные в диссертации, получены в рамках исследований и разработок, проводимых на кафедре информационных систем управления факультета прикладной математики и информатики Белорусского государственного университета, в соответствии с:

1) ГНТП «Информационные технологии», утвержденная приказом ГКНТ Республики Беларусь от 2 июня 2006 г. № 137: задание 5-07 «Разработать технологию построения компьютерных систем для решения задач распознавания образов в области медицинской диагностики при условии неполноты информации и внедрить ее в спортивной травматологии и реабилитологии» (№ госрегистрации 20064847, срок выполнения 2006-2012 гг.), исполнитель;

2) государственной программой научных исследований на 2011-2015 гг. «Информатика и космос»: задание 1.4.05 «Разработать технологии и компьютерные системы для решения задач распознавания образов со сложной структурой информации» (№ госрегистрации 20114342, срок выполнения 2010-2013 гг.), исполнитель;

3) договором с Белорусским республиканским фондом фундаментальных исследований № Ф15РМ-037: задание «Разработка и исследование многоуровневых моделей решения многоклассовых задач распознавания» (договор с БРФФИ от 04 мая 2015 г., № госрегистрации 20150928, срок выполнения 04.05.2015 – 31.03.2017 гг.), ответственный исполнитель;

4) научно-исследовательской программой Белорусского государственного университета: задание «Разработать теоретические основы и технологии построения информационных систем управления с приложением в интеллектуальном анализе данных» программы НИР БГУ на 2011-2015 гг. (№ госрегистрации 20120332, срок выполнения 2011-2015 гг.), исполнитель.

Цель и задачи исследования

Цель диссертационной работы – разработка моделей и алгоритмов решения задач распознавания с комбинированной обучающей информацией, объединяющих результаты и преимущества различных видов алгоритмов.

Для достижения цели были поставлены и решены следующие задачи:

- провести анализ подходов к решению задач распознавания и установить алгоритмическую связь между логическим и прецедентным способами представления информации;
- модифицировать алгоритм резолюций для объектов с конечным числом значений признаков;
- разработать гибридный алгоритм, объединяющий результаты решения алгоритмов с логическим и прецедентным способами представления информации;
- провести теоретическую и экспериментальную оценку качества предлагаемых алгоритмов и на базе проведенных исследований разработать программную технологию для решения комбинированных задач.

Объект исследования – задачи распознавания, в которых одновременно используются различные типы обучающей информации.

Предмет исследования – методы и алгоритмы решения задач распознавания с комбинированной обучающей информацией.

Научная новизна

Научная новизна исследования заключается в разработке оригинального подхода к решению задач распознавания с комбинированной обучающей информацией.

Разработана алгебра объектов, с помощью которой установлена связь между логической и прецедентной моделями представления информации. Показано, что алгебра объектов изоморфна алгебре функций k -значной логики. Предложена оригинальная кодировка объектов, учитывающая неопределенность значений признаков.

Разработаны модификация метода резолюций на случай задач распознавания с прецедентным представлением информации и гибридный алгоритм, ориентированный на различные способы представления обучающей информации.

Предложены эффективные реализации разработанных алгоритмов. Экспериментально показано, что уменьшение времени работы алгоритмов в среднем экспоненциально зависит от структуры информации об объектах.

Положения, выносимые на защиту

1. Модель алгебры объектов, устанавливающая связь между логическим и прецедентным способами представления информации, и алгоритмы перехода (в конечномерных пространствах) от одного представления информации к

другому, что в совокупности позволяет решать задачи распознавания независимо от исходного способа представления информации.

2. Алгоритм объектных резолюций, позволяющий использовать метод резолюции в случаях, когда информация представлена прецедентным способом.

3. Гибридный алгоритм, построенный на базе алгоритмов объектной резолюции и индуктивного вывода, что позволяет одновременно использовать логический и прецедентный способы представления обучающей информации.

4. Эффективные (с вычислительной точки зрения) реализации предложенных алгоритмов, на базе которых программно реализовано алгоритмическое ядро для построения прикладных систем, ориентированных на решение специальных задач распознавания с не полностью заданной и сложной по структуре информацией.

Личный вклад соискателя

Основные результаты и положения, выносимые на защиту, получены лично соискателем. В публикациях с соавторами вклад соискателя определяется рамками излагаемых в диссертации результатов. Постановка целей и задач исследовательской работы, обсуждение и интерпретация полученных результатов, определение структуры диссертации выполнялись совместно с научным руководителем, кандидатом физико-математических наук Образцовым В.А.

Апробация результатов диссертации

Основные научные и практические результаты работы докладывались и обсуждались на следующих международных и отечественных конференциях: Computer Data Analysis and Modeling conference (CDAM'2010, Minsk); IV Машеровские чтения: международная научно-практическая конференция студентов, аспирантов и молодых ученых (2010, Витебск); Международный форум учащейся и студенческой молодёжи «Первый шаг в науку – 2011» (2011, Минск); Pattern Recognition and Information Processing (PRIP'2011, Minsk); Республиканская научная конференция студентов и аспирантов Республики Беларусь «НИРС-2011» (2011, Минск); International Congress of Computer Science: Information systems and Technologies (CSIST'2011, Minsk); Modeling and Simulation (MS'2012, Minsk); International Congress of Computer Science: Information systems and Technologies (CSIST'2013, Minsk); Pattern Recognition and Information Processing (PRIP'2014, Minsk).

По результатам диссертации получено 3 акта внедрения. Результаты диссертации используются в Республиканском научно-практическом центре спорта, а также в учебном процессе и при выполнении различных НИР на кафедре информационных систем управления ФПМИ БГУ.

Опубликованность результатов диссертации

Результаты диссертации опубликованы в 15 научных работах, включая: 5 статей в рецензируемых научных журналах в соответствии с п.18 Положения о присуждении ученых степеней и присвоении ученых званий в Республике Беларусь (общим объемом 2,7 авторского листа), 9 статей в сборниках материалов научных конференций и 1 тезисы доклада.

Структура и объем диссертации

Диссертация состоит из перечня сокращений и условных обозначений, введения, общей характеристики работы, 4 глав, заключения, библиографического списка и 3 приложений. Полный объем диссертации составляет 114 страниц, в том числе 7 рисунков на 7 страницах, 15 таблиц на 10 страницах, 3 приложения на 11 страницах. Библиографический список содержит 68 наименований (включая собственные публикации соискателя).

ОСНОВНАЯ ЧАСТЬ

В первой главе исследуется задача распознавания образов с обучением. Выделены два основных способа представления обучающей информации: логический и прецедентный. Сформулирована задача распознавания с комбинированной обучающей информацией и определены основные проблемы, связанные с ее решением.

В рамках основных направлений теории распознавания рассматриваются дедуктивный и индуктивный подходы, использующие логический и прецедентный способы представления информации соответственно. Проведен анализ существующих алгоритмов решения задачи и показано, что каждый из них обладает определенными недостатками. При использовании алгоритмов дедуктивного типа невозможно получать решение за пределами ситуаций, ограниченных областью действия логического вывода. А при использовании алгоритмов индуктивного типа нельзя обосновать правильность получаемого решения. Вместе с тем существуют задачи распознавания (например, медицинская диагностика и др.), в которых по разным причинам необходимо использовать оба способа представления информации одновременно. Поэтому актуальной является проблема построения метода решения комбинированных

задач распознавания, позволяющего заимствовать все лучшее в существующих алгоритмах.

Рассмотрены методы решения задач распознавания, в которых обучающая выборка содержит оба типа обучающей информации. На рисунке 1 приведены схемы процесса получения решения, которые на сегодняшний день можно считать общепринятыми. В первом случае происходит комбинирование результатов решения задачи, полученных алгоритмом резолюции¹ A_0 и алгоритмом индуктивного вывода² A_2 , а во втором – выборка, представленная логическим способом, переводится к прецедентному представлению, и задача решается алгоритмом A_2 .

Для решения комбинированных задач предлагается подход, основанный на объединении различных типов алгоритмов (рисунок 2).



Рисунок 1. – Схемы решения задач для различных типов информации

Общая концепция подхода заключается в следующем. Вводится специальная алгебра объектов, в рамках которой устанавливается соотношение между логическим и прецедентным представлениями обучающей информации и строятся алгоритмы перехода от одного вида представления к другому. Результатом является адаптация классического метода резолюций к прецедентному представлению (алгоритм A_1), на основе которого строится гибридный алгоритм распознавания A_C .

Далее в главе формулируются основные проблемы, решение которых позволяет реализовать изложенный подход и тем самым достичь поставленной в диссертации цели.

¹ Barwise, J. Handbook of Mathematical Logic / J. Barwise. – Elsevier Science B.V, 1977. – 1165 p.

² Абламейко, С. В. Модели и технологии распознавания образов с приложением в интеллектуальном анализе данных / С. В. Абламейко, В. В. Краснопрошин, В. А. Образцов // Вестник БГУ. Серия 1, Физика, Математика, Информатика. – 2011. – № 3. – С. 62–72.



Рисунок 2. – Предлагаемая схема решения комбинированных задач

Во второй главе исследуется связь между логическим и прецедентным способами представления информации, и предлагаются алгоритмы перехода между ними.

Пусть Z – произвольная задача распознавания. Для того, чтобы установить указанное соответствие, вводится понятие алгебры объектов.

Произвольный объект предлагается рассматривать в виде отображения:

$$p(s_1, s_2, \dots, s_n) = (D_1^p, D_2^p, \dots, D_n^p),$$

где $S = \{s_1, s_2, \dots, s_n\}$ ($n < \infty$) – наименования признаков, D_j^p – множество их значений ($D_j^p \neq \emptyset$).

Вводится специальный объект-признак:

$$p^j(s_1, s_2, \dots, s_n) = (D, \dots, D, \{d_j^{p^j}\}, D, \dots, D),$$

который получается фиксацией значения одного признака ($|D_j^p| = 1$).

В процессе построения объектов возникают ситуации, когда значение некоторых признаков не определено либо число их значений $|D_j^p| > 2$ (т.н. шкалированные или непрерывные признаки). Для применимости алгоритмов в описанных выше случаях вводится несколько видов кодировки: регулярным образом расширяется признаковое пространство, но при этом сами значения признаков выбираются из множества $B_2 = \{0,1\}$. Кодировка, в результате которой устраняется неопределенность значений признаков, называется нормализацией.

Далее определяется базовая операция произведения двух произвольных объектов p и q (обозначается pq):

$$D_j^{pq} = D_j^p \cap D_j^q,$$

с использованием которой вводятся операции над наборами объектов $V \subset X$ и $W \subset X$:

$$\begin{aligned} \bar{V} &= X^{norm} \setminus V \text{ – отрицание,} \\ V \wedge W &= \bigcup_{p \in V, q \in W} \{pq\} \text{ – умножение,} \\ V \vee W &= V \cup W \text{ – сложение,} \end{aligned}$$

где X^{norm} – множество нормализованных объектов.

Показано, что система операций $\{\neg, \wedge, \vee\}$ над наборами объектов обладает свойством полноты в том смысле, что любой набор может быть построен из объектов-признаков с помощью только этих операций.

Построены алгебры объектов $G = \langle \rho(X), \{\neg, \wedge, \vee\} \rangle$ и их нормализованных вариантов $G^{norm} = \langle \rho(X^{norm}), \{\neg, \wedge, \vee\} \rangle$, где $\rho(X)$, $\rho(X^{norm})$ – множество всех возможных наборов и нормализованных вариантов соответственно. Доказано, что алгебра G^{norm} является подалгеброй G .

С использованием введенных кодировок построены также алгебры объектов C и C^{norm} вида G и G^{norm} соответственно. Показано, что алгебры G и C , а также G^{norm} и C^{norm} изоморфны. Кроме того, доказано, что эти алгебры изоморфны классическим алгебрам k -значной логики.

В результате проведенных исследований построены алгоритмы перехода от логической модели представления информации к прецедентной (T_{12}) и обратно (T_{21}). Вкратце их можно описать следующим образом.

Пусть φ – логическая формула, описывающая класс X_i , где $X_i \neq \emptyset$. Рассмотрим случай $k=2$. Поскольку φ – булева функция, то ее можно представить в виде совершенной ДНФ:

$$\varphi(x_1, \dots, x_n) = \bigvee_{(d_1, \dots, d_n)} x_1^{d_1} \dots x_n^{d_n}, \quad x_i^{d_i} = \begin{cases} x_i, & d_i = 1 \\ \bar{x}_i, & d_i = 0 \end{cases}$$

$$\varphi(d_1, \dots, d_n) = 1$$

Пусть t – число элементарных конъюнкций в данном представлении формулы φ . Обозначим j -ю элементарную конъюнкцию (ЭК) через $\varphi_j(x_1, \dots, x_n) = x_1^{d_1} \dots x_n^{d_n}$. Тогда

$$\varphi(x_1, \dots, x_n) = \bigvee_{j=1}^t \varphi_j(x_1, \dots, x_n).$$

Алгоритм T_{12} :

Шаг 1. Каждой ЭК $\varphi_j(x_1, \dots, x_n) = x_1^{d_1} \dots x_n^{d_n}$ ставим в соответствие объект $p_j(s_1, \dots, s_n) = (\{d_1\}, \dots, \{d_n\})$.

Шаг 2. Строим набор $V = \{p_1\} \vee \{p_2\} \vee \dots \vee \{p_t\}$.

Шаг 3. Алгоритм завершает работу.

Алгоритм T_{21} :

Шаг 1. Нормализуем набор V . Пусть $V^{norm} = Norm(V)$.

Шаг 2. Каждому объекту $p_j(s_1, \dots, s_n) = (\{d_1^{p_j}\}, \dots, \{d_n^{p_j}\})$, $p_j \in V^{norm}$, ставим в соответствие ЭК $\varphi_j(x_1, \dots, x_n) = x_1^{d_1^{p_j}} \dots x_n^{d_n^{p_j}}$.

Шаг 3. Набору V ставим в соответствие ДНФ $\varphi(x_1, \dots, x_n) = \bigvee_{j=1}^t \varphi_j(x_1, \dots, x_n)$, где t – количество объектов в наборе V^{norm} .

Шаг 4. Алгоритм завершает работу.

Для случая $k \geq 3$ вводится операция:

$$I_i(x) = \begin{cases} k-1, & x = i \\ 0, & x \neq i \end{cases},$$

которая позволяет строить совершенную ДНФ в виде:

$$\varphi(x_1, \dots, x_n) = \bigvee_{\substack{(d_1, \dots, d_n) \\ \varphi(d_1, \dots, d_n) = k-1}} I_{d_1}(x_1) \dots I_{d_n}(x_n)$$

Показано, что для построенных алгоритмов T_{12} и T_{21} справедливо следующее утверждение.

Теорема. Для произвольных формул φ , наборов V и объектов p справедливы соотношения:

$$\varphi(p) = 1 \Leftrightarrow p \in T_{12}(\varphi)$$

$$T_{21}(V)(p) = 1 \Leftrightarrow p \in V$$

В третьей главе предложена модификация метода резолюций (метод объектных резолюций), исходными данными которой являются объекты из множества X . На его основе строится гибридный алгоритм.

Пусть r – произвольный объект, значения признаков которого удовлетворяют условию:

$$D_j^r = \begin{cases} D_j^p \cup D_j^q, & j = h \\ D_j^p \cap D_j^q, & j \neq h \end{cases}$$

где h – номер признака $s_h \in S$, а p и q – исходные объекты.

Процедура $Or_h(p, q)$, в результате которой может быть получен такой объект, названа операцией построения объектной резольвенты.

На основе $Or_h(p, q)$ построен алгоритм решения задачи Z .

Пусть $X_i^0 = X^0 \cap X_i$ ($i \in \{1, \dots, l\}$).

Алгоритм A_1 :

Шаг 1. Введем множество $Y = X_i^0$.

Шаг 2. Если $x \in Y$ – переходим к шагу 6, иначе – к шагу 3.

Шаг 3. Выберем из Y тройку (p, q, h) , где p и q – объекты, h – номер признака. Если все тройки уже рассмотрены – переходим к шагу 6.

Шаг 4. Вычислим $r = Or_h(p, q)$. Если $\exists j D_j^r = \emptyset$ – возвращаемся к шагу 3.

Шаг 5. Если $r \notin Y$ – добавляем к Y объект r . Возвращаемся на шаг 2.

Шаг 6. Алгоритм завершает работу.

Показано, что алгоритм A_1 может применяться для прямого и обратного выводов, а также работать не только с нормализованными объектами.

Пусть Z – произвольная задача распознавания с комбинированной информацией, в которой выборка X^{01} получена из логической части, а X^{02} – из прецедентной. Предлагается следующий алгоритм решения задач Z .

Гибридный алгоритм A_C :

Шаг 1. Выбираем произвольный объект $x \in c(X)$.

Шаг 2. Фиксируем номер класса $i = \overline{1, l}$.

Шаг 3. Применяем алгоритм A_1 к объекту x и набору X_i^0 из X^{01} .

Шаг 4. Если $P_i^{A_1}(x) = 1$ – переходим к шагу 7. Иначе – к шагу 5.

Шаг 5. Если классы исчерпаны – переходим к шагу 6, иначе – к шагу 2.

Шаг 6. Применяем алгоритм A_2 к объекту x и X_i^0 из X^{02} .

Шаг 7. Если классы исчерпаны – переходим к шагу 8, иначе – к шагу 1.

Шаг 8. Алгоритм завершает работу.

Доказано, что результаты работы алгоритма A_C , не хуже результатов решения Z алгоритмами A_1 и A_2 :

$$\Phi_{A_C}(X^q) \geq \max\{\Phi_{A_1}(X^q), \Phi_{A_2}(X^q)\},$$

где Φ_A – функционал качества работы алгоритмов, а $X^q \subset X$.

В четвертой главе исследуется вычислительная сложность алгоритмов, проводится их экспериментальная оценка и рассматриваются проблемы применения полученных в диссертации результатов на практике.

Построены эффективные (по числу операций) модификации алгоритмов A_1^B , A_2^B и A_C^B , основанные на идеологии метода ветвей и границ. При модификации алгоритма A_1 признаки рассматриваются последовательно и удаляются те объекты, которые заведомо не могут быть использованы при выводе x из множества X^0 . Для модификации алгоритма A_2 при построении оценки близости используется сокращенное пространство признаков. Что касается модификации алгоритма A_C , то она основана на использовании алгоритмов A_1 или A_2 , в зависимости от области компетентности. Показано, что сложность алгоритмов A_1^B , A_2^B и A_C^B экспоненциально зависит от структуры информации.

Проведены экспериментальные исследования свойств алгоритмов T_{12} и T_{21} , а также оценка качества и времени работы алгоритмов A_1 , A_2 , A_C , A_1^B , A_2^B , A_C^B . Тестирование проводилось с использованием данных из известных репозиторий UCI и KEEL. Результаты исследований подтвердили корректность теоретических результатов диссертации.

Во второй части главы описана программная технология построения прикладных систем для решения специальных задач распознавания с неполной (не полностью заданной) и сложной по структуре информацией. Технология основана на использовании унифицированного алгоритмического ядра и инструментальных средств для ввода формализованных знаний и настройки на предметную область, что в совокупности дает возможность серийного построения и модификации прикладных систем. Алгоритмическое ядро построено на базе программных реализаций алгоритмов, разработанных в диссертации.

Разработана методология практического использования алгоритмического ядра системы, которое базируется на следующих процедурах.

Представление алгоритма принятия решений:

Этап 1. Анализ информации.

Шаг 1.1. Приведение обучающей выборки, представленной с использованием логической модели, к прецедентному виду.

Шаг 1.2. Приведение описания классов (диагнозов) к прецедентному представлению.

Этап 2. Обследование.

Шаг 2.1. Формирование входной информации, включающей описание объектов и классов (см. этап 1).

Шаг 2.2. Вычисление и интерпретация результатов диагностики.

Этап 3. Завершение работы.

Схема практического использования системы:

Шаг 1. Ввод информации о пациенте (паспортные данные и результаты предварительного обследования).

Шаг 2. Определение локализации заболевания.

Шаг 3. Осмотр пораженной области и выявление дополнительных симптомов.

Шаг 4. Постановка диагноза.

Шаг 5. Представление результатов обследования в удобном для пользователя виде.

Шаг 6. Формирование схемы лечения.

Шаг 7. Завершение работы.

Приводятся результаты практического применения интеллектуальной системы EXTRA, построенной с использованием указанной выше технологии. Общая структура системы приведена на рисунке 3.

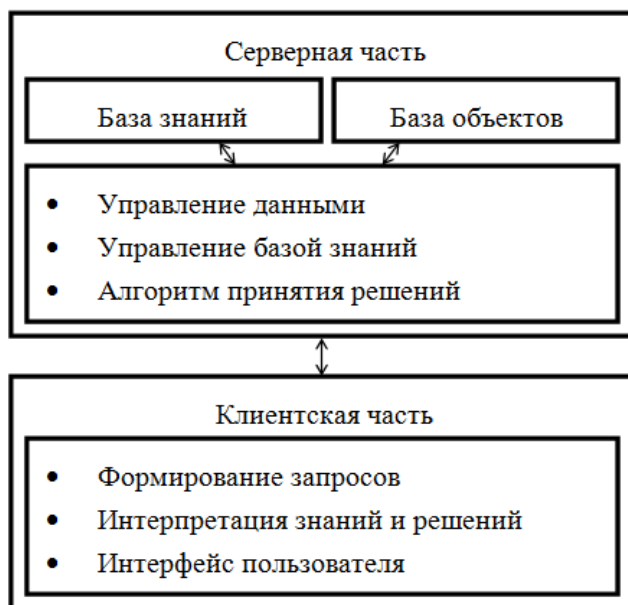


Рисунок 3. – Структура системы EXTRA

Система предназначена для поддержки принятия решений в области ортопедии и спортивной травматологии. Она обеспечивает лечебно-диагностический процесс от обследования пациента до выбора тактики лечения различных патологий. Система EXTRA внедрена в центре спортивной медицины РБ.

Таким образом, решены все поставленные задачи и достигнута цель диссертационной работы: разработан подход к решению задач распознавания с комбинированной обучающей информацией, позволяющий объединить результаты и преимущества различных типов алгоритмов.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

1. Построена модель алгебры объектов, устанавливающая связь между логическим и прецедентным способами представления информации, и алгоритмы перехода (в конечномерных пространствах) от одного представления информации к другому, что в совокупности позволяет решать задачи распознавания независимо от исходного способа представления информации. Показано, что алгебра объектов изоморфна алгебре функций k -значной логики [1, 5-8, 15].

2. Разработан алгоритм объектных резолюций, позволяющий использовать метод резолюций в случаях, когда информация представлена прецедентным способом [2, 3, 5, 9-12].

3. Разработан гибридный алгоритм, построенный на базе алгоритмов объектной резолюции и индуктивного вывода, что позволяет одновременно использовать логический и прецедентный способы представления обучающей информации. Показано, что качество данного алгоритма не хуже алгоритмов, образующих комбинацию [2, 3, 9-12].

4. Разработаны эффективные (с вычислительной точки зрения) реализации предложенных алгоритмов и проведены их экспериментальные исследования. На основе теоретических результатов построено алгоритмическое ядро для создания прикладных систем, которые ориентированы на решение специальных задач распознавания с не полностью заданной и сложной по структуре информацией [4, 13, 14].

Рекомендации по практическому использованию результатов

Разработанные в диссертации алгоритмы могут применяться для решения задач медицинской диагностики. На базе программных реализаций таких алгоритмов построено алгоритмическое ядро интеллектуальной системы

EXTRA, предназначенной для диагностики заболеваний и выбора тактики лечения в области ортопедии и спортивной травматологии. Система внедрена в центре спортивной медицины РБ.

Полученные в диссертационной работе результаты использовались также в учебном процессе и при выполнении различных НИР на кафедре информационных систем управления ФПМИ БГУ.

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ

Статьи в научных журналах

1. Шут, О.В. Метод решения задач распознавания на основе логической и прецедентной моделей / О.В. Шут // Информатика – 2012. – №3. – С. 35–50.
2. Shut, O.V. A model of Solving of Pattern Recognition Problems Based on Resolution Method / O.V. Shut // Lectures on Modeling and Simulation – 2012. – Vol. 14, Issue 1. – P. 7–19.
3. Шут, О.В. Синтез алгоритмов для решения задач распознавания образов в конечномерных дискретных пространствах / О.В. Шут // Вестник БГУ. Серия 1, Физика, Математика, Информатика. – 2014. – № 1. – С. 56–62.
4. Шут, О.В. Сложность решения задач распознавания с обучением в классе алгоритмов индуктивной резолюции / О.В. Шут // Вестник БГУ. Серия 1, Физика, Математика, Информатика. – 2014. – № 2. – С. 107–113.
5. Shut, O.V. Code Descriptions of Classes Method for Pattern Recognition with Multiple Classes / A. Dokukin, V. Ryazanov, O. Shut // International Journal “Information Content and Processing”. – 2015. – Vol. 2. – P. 128–137.

Материалы конференций

6. Shut, O.V. Logical and Precedent-Related Representation of Information in Pattern Recognition / O.V. Shut // Proc. of the Computer Data Analysis and Modeling conf., Minsk, 7–11 September 2010. – Minsk, 2010. – P. 127–130.
7. Шут, О.В. Логическое и прецедентное представление информации в задачах распознавания образов / О.В. Шут // IV Машеровские чтения: материалы междунар. науч.-практ. конф. студентов, аспирантов и молодых ученых, Витебск, 28–29 октября 2010 г., в 2 т. / Витебский гос. ун-т; редкол.: А.П. Солодков (гл. ред.) [и др.]. – Витебск: УО «ВГУ имени П.М. Машерова», 2010. – Т. I. – С. 71–72.
8. Шут, О.В. Проблема обоснования алгоритмов распознавания: прецедентный и логический подход / О.В. Шут // Первый шаг в науку – 2011: сб. материалов Междунар. форума учащейся и студенческой молодёжи «Первый шаг в науку – 2011», Минск, 25–29 апр. 2011 г. / Нац. акад. наук

Беларуси; редкол.: М.Ю. Дегтярёва [и др.]. – Минск: Беларус. навука, 2011. – С. 570–573.

9. Shut, O.V. Logical and Precedent-related Approaches to Representation of Information in Pattern Recognition Problems: Comparative Analysis and Applications / O.V. Shut // Pattern Recognition and Information Processing (PRIP'2011): proc. of the 11th International Conference, Minsk, 18–20 May 2011. – Minsk: BSUIR, 2011. – P. 339–341.

10. Шут, О.В. Метод объектной резолюции в задачах распознавания образов / О.В. Шут // Международный конгресс по информатике: информационные системы и технологии: материалы междунар. науч. конгресса, Республика Беларусь, Минск, 31 окт. – 3 нояб. 2011 г.: в 2 ч. / редкол.: С.В. Абламейко (отв. ред.) [и др.]. – Минск: БГУ, 2011. – Ч. 2. – С. 228–232.

11. Shut, O.V. A Model of Solving of Pattern Recognition Problems Based on Resolution Method / O.V. Shut // Modeling and Simulation : MS'2012 : Proc. of the Intern. Conf., Belarus, Minsk, 2–4 May 2012. – Minsk : Publ. Center of BSU, 2012. – P. 76–81.

12. Шут, О.В. Алгоритмы решения задачи распознавания образов на основе метода резолюций / О.В. Шут // Международный конгресс по информатике: информационные системы и технологии: материалы междунар. науч. конгресса, Республика Беларусь, Минск, 4–7 нояб. 2013 г. / редкол.: С.В. Абламейко (отв. ред.), В.В. Казаченок (отв. ред.) [и др.]. – Минск: БГУ, 2013. – С. 482–487.

13. Shut, O.V. The complexity of inductive resolution algorithms in pattern recognition problems / O.V. Shut // Pattern Recognition and Information Processing (PRIP'2014): proc. of the 12th International Conference, Belarus, Minsk, 28-30 May 2014. – Minsk: UPP NASB, 2014. – P. 260–264.

14. Шут, О.В. Алгоритмы распознавания на основе метода резолюций и их применение в системе EXTRA / О.В. Шут // Актуальные проблемы развития современной науки и образования: Сб. научных трудов по материалам Междунар. науч.-практ. конф. Москва, 30 апреля 2015 г.: в 5 ч. – М.: «АР-Консалт», 2015 г. – Ч. III. – С. 139–140.

Тезисы докладов

15. Шут, О.В. Прецедентный и логический подходы представления информации в задачах распознавания образов с обучением / О.В. Шут // Сб. тезисов докладов Респ. науч. конф. студентов и аспирантов Республики Беларусь «НИРС-2011», Минск, 18 окт. 2011 г. / редкол.: С.В. Абламейко [и др.]. – Минск: Изд. центр БГУ, 2011. – С. 78.

РЭЗІЮМЭ

Шут Вольга Віктараўна

Мадэлі і алгарытмы для задач распазнання з камбінаванай навучальнай інфармацыяй

Ключавыя словы: распазнанне вобразаў, навучальная інфармацыя, булева алгебра, метады рэзалюцый, функцыянал якасці.

Мэта работы: распрацоўка падыхода да рашэння задач распазнання з камбінаванай навучальнай інфармацыяй, дазваляючага аб'яднаць вынікі і перавагі розных тыпаў алгарытмаў.

Метады даследавання: тэорыя распазнання вобразаў, метады рэзалюцый, метады агульнай алгебры, аб'ектна-арыентаванае праграмаванне.

Атрыманыя вынікі і іх навізна. Прапанаваны новы падыход да рашэння задач распазнання з навучэннем, які дазваляе сумесна выкарыстоўваць розныя спосабы прадстаўлення навучальнай інфармацыі.

Распрацавана алгебра аб'ектаў, якая устанаўлівае сувязь паміж лагічным і прэцэдэнтным спосабамі прадстаўлення інфармацыі, і алгарытмы пераходу (у канечнамерных прасторах) ад аднаго прадстаўлення інфармацыі да іншага, што дазваляе рашаць задачы распазнання незалежна ад зыходнага спосаба прадстаўлення інфармацыі.

Таксама распрацаваныя алгарытмы аб'ектных рэзалюцый, які рэалізуюць метады рэзалюцый для рашэння задач распазнання з прэцэдэнтным спосабам задання інфармацыі, і гібрыдны алгарытмы, які аб'ядноўвае рашэнні, атрымліваемыя пры выкарыстанні алгарытмаў аб'ектнай і індуктыўнай рэзалюцый. Паказана, што гэты алгарытмы рашае задачу распазнання з камбінаванай навучальнай інфармацыяй не горш за алгарытмы, якія ўтвараюць камбінацыю. Прапанаваны эфектыўныя (з вылічальнага пункту гледжання) рэалізацыі распрацаваных алгарытмаў, на базе якіх пабудавана алгарытмічнае ядро для стварэння прыкладных сістэм, арыентаваных на рашэнне спецыяльных задач распазнання з няпоўнай (не поўнасцю зададзенай) і складанай па структуры інфармацыяй.

Ступень выкарыстання: вынікі дысертацыйных даследаванняў ўкаранёны ў навучальны працэс на кафедры інфармацыйных сістэм кіравання факультэта прыкладной матэматыкі і інфарматыкі Беларускага дзяржаўнага ўніверсітэта. Прапанаваныя алгарытмы ў рамках сістэмы EXTRA, распрацаванай для дыягностыкі захворванняў у вобласці артапедыі, ўкаранёны ў рэспубліканскім Цэнтры спартыўнай медыцыны.

РЕЗЮМЕ

Шут Ольга Викторовна

Модели и алгоритмы для задач распознавания с комбинированной обучающей информацией

Ключевые слова: распознавание образов, обучающая информация, булева алгебра, метод резолюций, функционал качества.

Цель работы: разработка подхода к решению задач распознавания с комбинированной обучающей информацией, позволяющего объединить результаты и преимущества различных типов алгоритмов.

Методы исследования: теория распознавания образов, метод резолюций, методы общей алгебры, объектно-ориентированное программирование.

Полученные результаты и их новизна. Предложен новый подход к решению задач распознавания с обучением, который позволяет одновременно использовать различные способы представления обучающей информации.

Разработана алгебра объектов, устанавливающая связь между логическим и прецедентным способами представления информации, и алгоритмы перехода (в конечномерных пространствах) от одного представления информации к другому, что позволяет решать задачи распознавания независимо от исходного способа представления информации.

Также разработаны алгоритм объектных резолюций, реализующий метод резолюций для решения задачи распознавания с прецедентным способом задания информации, и гибридный алгоритм, объединяющий решения, получаемые при использовании алгоритмов объектной и индуктивной резолюций. Показано, что этот алгоритм решает задачу распознавания с комбинированной обучающей информацией не хуже алгоритмов, образующих комбинацию. Предложены эффективные (с вычислительной точки зрения) реализации разработанных алгоритмов, на базе которых построено алгоритмическое ядро для создания прикладных систем, ориентированных на решение специальных задач распознавания с неполной (не полностью заданной) и сложной по структуре информацией.

Степень использования: результаты диссертационных исследований внедрены в учебный процесс на кафедре информационных систем управления факультета прикладной математики и информатики БГУ. Предложенные алгоритмы в рамках системы EXTRA, разработанной для диагностики заболеваний в области ортопедии, внедрены в республиканском Центре спортивной медицины.

SUMMARY

Shut Olga Victorovna

Models and Algorithms for Pattern Recognition Problems with Combined Prior Information

Keywords: pattern recognition, prior information, Boolean algebra, resolution method, quality functional.

The aim of the research: development of an approach to solving of pattern recognition problems with combined prior information, which allows unification of results and advantages of different algorithm types.

Research methods: theory of pattern recognition, resolution method, methods of general algebra, object-oriented programming.

Obtained results and their novelty. A new approach to solving of pattern recognition problems with prior information, that allows to use simultaneously different ways of representation of prior information, is proposed.

An algebra of objects, that establishes the connection between logical and precedent-related ways of representation of information, and algorithms for conversion ways of representation of information (in finite-dimensional spaces) are developed. Thus, pattern recognition problems can be solved independently of the way of representation of information.

An object resolution algorithm, that implements resolution method for solving of pattern recognition problems, which have the precedent-related way of representation of information, and a hybrid algorithm, which unites solutions that are obtained via object and inductive resolutions algorithms, are developed. This algorithm is shown to solve pattern recognition problems with combined prior information not worse than any of two algorithms, which are making the combination. Effective (from the computational point of view) implementations of the developed algorithms are proposed. An algorithmic kernel, which is based on these implementations and is meant for creating applied systems, that are oriented to solving of special pattern recognition problems with incomplete (not fully specified) and complicatedly structured information, is built.

The level of application: the investigation results are applied in educational process at Information Control Systems department of Faculty of Applied Mathematics and Informatics of Belarusian State University. The proposed algorithms are also applied within the scope of the system EXTRA, which has been developed for diagnostics of orthopedics diseases, in Republican Centre of Sports Medicine.