

## ОСОБЕННОСТИ ФОРМАЛИЗАЦИИ КЛАССИФИЦИРУЮЩИХ СТРУКТУР В СРЕДЕ ПРИОБРЕТЕНИЯ ЗНАНИЙ

к.т.н. Т.Н. Новожилова  
(представила д.т.н., проф. Р.Т. Волколупова)

Рассмотрены особенности языка представления знаний интеллектуальной базы данных (ЯПЗ ИБД) для задач бизнес - планирования.

Знания в интеллектуальных системах образуют упорядоченные структуры, для представления и обработки которых используются различные классифицирующие процедуры, лежащие в основе ЯПЗ. Существующие ЯПЗ не опираются на аналоги когнитивных структур для представления знаний, которыми пользуются люди, что связано с плохой изученностью форм представления знаний у человека. Соответствующий раздел психологии – когнитивная психология – развивается быстро, но достигнутые результаты, которые могли бы оказать влияние на создание новых моделей представления знаний, ещё слишком скромны. На наш взгляд, проблемы разрозненности простых структур знаний в ЯПЗ можно решить не только отказавшись от них (реляционно-фреймовые ЯПЗ [1]), но и организовав управляющие структуры, моделирующие эволюцию этапов осмысления фактов экспертами [2, 3]. В разрабатываемом ЯПЗ ИБД предпринята попытка снятия противопоставления ситуативного типа классификации родовидовому: для ЯПЗ ИБД, как модели обобщения по выборкам, характерно априорное разделение фактов по ситуациям (**DS**). Дальнейшая детализация классификации **DS** и сочетания вероятностных и логических признаков позволяет исследовать возможности экспертов по осмыслению **DS** при разработке бизнес-планов. Основные результаты исследования взаимосвязи интерпретаций классифицируемых структур ЯПЗ ИБД с основными понятиями теории вероятности сведены в табл. 1. В ЯПЗ ИБД различаются два способа классификации  $R_{na}(H_2)$ :

1) простое перечисление всех его элементов (*определение класса расширения*) [4]: чтобы существовало отображение

$$R_{na}(H_2)_i \xrightarrow{\Delta R_{na}} R_{na}(H'_0) \quad (1)$$

для наборов правил неопределённого класса  $R_{na}(H_2)_i$ , необходимо и достаточно, чтобы  $\exists h \in H_1$ , для которой

$$f(CF(H_2, h)) \geq f(CF(H_1)) > E.UNKN, \quad (2)$$

где **E.UNKN** – порог неопределённости;

2) определение класса в терминах взаимосвязи элементов (*определение*

Взаимосвязь интерпретаций теории вероятностей и ЯПЗ ИБД

Обозначения	Интерпретация теории вероятностей	Интерпретация ЯПЗ ИБД
$DS = \langle DS_1, \dots, DS_i, \dots, DS_{72} \rangle$ $DS_i = \langle \Delta KS, \Delta AN, \Delta IS \rangle$	Аксиоматическая установка несовместимости событий	Группировка фактов по <b>DS</b> [2-13]
<b>CF(O)</b>	Вероятностная оценка	Коэффициент уверенности
<b>X</b>	Комплекс условий	Условия, при которых имеет смысл <b>CF(O)</b>
<b>{X-O}</b>	Опыт	Опрос экспертов
<b>RXO<sub>1</sub></b> <b>R O<sub>2</sub></b> и др.	Способы выделения равновозможных случаев	Планы экспертных опросов [9,12,14]
<b>f(CF)</b>	Мера несовместимости	Степень уверенности, определяемая <b>f(CF)</b>
<b>R<sub>na</sub>(H<sub>2</sub>)={X-O}<sub>1</sub></b>	Совокупность соотносящихся событий	Набор правил неопределённого класса
<b>h</b>	Гипотеза	Гипотеза о существовании связи <b>{X-O}</b>
$\left\{ R_{na} \begin{bmatrix} h & H_2 \\ X & O \end{bmatrix} \right\}$	Параметры преобразований равновозможных случаев	Логические координаты [4]
$\cup, \cap, \otimes, \oplus, /$	Операции преобразования событий	Операции преобразования <b>R<sub>na</sub>(H<sub>2</sub>)</b> [5]
<b>R<sub>na</sub>(H<sub>1</sub>)</b>	Закон распределения событий	Механизм целеполагания [2,3,6]
Для <b>R<sub>na</sub>(H<sub>2</sub>)<sub>1</sub>, ..., R<sub>na</sub>(H<sub>2</sub>)<sub>n</sub></b> $\exists \lim_{h \rightarrow r} CF(R_{na}(H_2)) \geq E.UNKN$	Закон больших чисел	Сходимость к устойчивому множеству рассуждений [4]

если 
$$\exists M(h) \in \{R_{na}(H_2)\} \quad (3)$$

и 
$$\lim_{h \rightarrow r} \bigcup_h M(h) = R_{na}(H_1), \quad (4)$$

то 
$$\{\exists h | R_{na}(H_1) \oplus R_{na}(h) = R_{na}(H_2)\}. \quad (5)$$

Предполагается, что определение расширением всегда можно свести к определению содержанием (обратный процесс возможен не всегда). Поэтому достаточно задать только определение содержанием [4]. При этом определение расширением всегда имеет фиксированную интерпретацию

$$\left[ R_{na}(h) \rightarrow R_{na}(H_1).RSS \rightarrow R_{na} \left[ \begin{array}{cc} h & H'_0 \\ X & O \end{array} \right] \rightarrow R_{na}(H_2) \right]. \quad (6)$$

С точки зрения структуры признаков, отражаемых классификационной моделью ЯПЗ ИБД, в среде приобретения знаний различаются интенциональный и экстенциональный аспекты классификации.

**Определение 1.** С экстенциональной точки зрения классификация экспертных знаний представляет собой описание классификационной структуры связями **EXT(X-O)** между наборами правил [2, 3, 9, 11] и теоретико-множественных отношений между ними [4].

Для учета интенционального аспекта знаний эксперта в ЯПЗ параллельно поддерживаются две классификации: родо-видовая классификация наборов правил (перевод категорий математического анализа на ЯПЗ ИБД [3, 4]) и классификация часть-целое модели эксперта [6].

**Определение 2.** С интенциональной точки зрения **INT(X-O)** классификация в ЯПЗ ИБД связана с признаками, на основании которых учитывается параметричность знаний в логических формах [2, 9, 11, 13].

Явное выражение интенционального аспекта классификации в ЯПЗ ИБД состоит в учёте в её структуре степени осмысленности экспертом свойств объектов классифицируемого соотношения, придающего описанию классификации необходимую параметрическую форму [14, 15]:

$$\bigcup_i^j R_{na}(H_1) : DS_i \rightarrow DS_j. \quad (7)$$

**Определение 3.** Требование единства интенционального и экстенционального аспектов метазнаний отражается эпитеоремой **f(R<sub>ni</sub>)** [2].

Если эксперт обладает опытом, который позволяет ему мысленно надстроить над единичным фактом большую генеральную совокупность [7-14], то его субъективные оценки используются для построения модели эксперта.

**Определение 4.** Под моделью эксперта будем понимать набор объектов **H'\_2(DS, R<sub>na</sub>(H<sub>2</sub>), f(CF))**, позволяющий для **DS<sub>i</sub> → DS<sub>j</sub>** задавать вероятностными рассуждениями **R<sub>na</sub>(H<sub>1</sub>)** [15].

ЯПЗ ИБД даёт аксиоматическую установку приобретаемым моделям эксперта и позволяет в абстрактной форме отражать закономерности (фиксируемые экспертами), присущие случайным событиям массового характера [15-18]. Способность эксперта формализовывать равновозможные случаи отражает интегральный признак сходимости приобретаемых наборов правил.

**Утверждение** (*интегральный признак сходимости модели эксперта*).

Если  $\sum_{i=1}^n R_{na}(H_{2i})$  сходится к устойчивому множеству рассуждений, то

для  $H'_2(DS, R_{na}(H_2), f(CF))$ ,  $\exists R_{na}(H_1): DS_i \rightarrow DS_j$  такой, что

$$\lim_{h \rightarrow r} R_{na}(H_2) \otimes R_{na}(H_1) = R_{na}(H'_2), \quad (8)$$

где  $R_{na}(H'_2)$  - подпространство гипотез [11].

**Определение 5.** Под алгеброй коэффициентов уверенности  $f(CF)$  в ЯПЗ понимается способ анализа эффективности логических форм представления знаний в  $H'_2(DS, R_{na}(H_2), f(CF))$ .

**Определение 6.** Логическая форма ЯПЗ ИБД отражает степень осмысления комплекса условий посредством параметров обобщения функциональных зависимостей в  $H'_2(DS, R_{na}(H_2), f(CF))$  [2, 13].

При определении класса  $R_{na}(H_2)$  учитывается определяющее свойство логической формы, специфицируемое множеством всех элементов класса, и поэтому используемое для проверки принадлежности объектов классу (необходимое и достаточное условие) [4].

**Определение 7.** Под логическими координатами ЯПЗ задач классификации будем понимать параметричность  $H'_2(DS, R_{na}(H_2), f(CF))$  или условия существования логических форм ЯПЗ ИБД, где логические связи и их сочетания выделяются как параметры  $R_{na}(H_2)$ , позволяющие получать классификацию эвристических решений и обоснование их правдоподобности.

Таким образом, под логическими координатами понимаются логические условия (признаки) и их сочетания, а также фрагменты структур, которые способствуют решению задач классификации. Понятно, что условия, помогающие или ускоряющие решение одних задач, могут быть неприемлемыми для других. Поэтому нельзя говорить об абсолютных логических координатах классификации, независимо от решаемых задач [9-14]. Анализируя знания эксперта, когнитолог не ограничивается изучением знаний лишь со стороны их структуры [10, 16-19]: логическими координатами ЯПЗ ИБД нацелена идея перехода  $DS_i \rightarrow DS_j$ . В этом выражается формализованная в  $R_{na}(H_1)$  готовность логических форм к развитию, обобщению, дифференцированию в различных ситуациях с помощью тех или иных параметров обобщения. Эта особенность ЯПЗ ИБД позволяет оценивать одно и тоже эвристическое решение с разных точек зрения, используя методы:

- обработки логических форм с иерархической структурой [5, 6];

- моделирования сложных рассуждений [2, 3, 7, 11-13];
- обнаружения содержательных признаков (параметров) в классифицируемых объектах и саму систему этих признаков [8, 9, 12, 14].

Таким образом, интенциональный подход позволяет расширить классификационное поле ЯПЗ ИБД до классификационного универсума (эпитеоремы [2, 3]):

$$\lim_{h \rightarrow r} R_{ni}(\text{INT}(X-O), \text{EXT}(X-O)) = f(R_{ni}(X-O)). \quad (9)$$

Эта особенность логических форм ЯПЗ ИБД к развитию в процессе анализа приобретаемых знаний у эксперта доказывает, что подход к приобретению знаний через сформированные исходные структуры (фреймы, семантические сети, продукции), по-видимому, не единственный. Существует другой путь – *переход от одной логической формы к другой, в порядке предметности, развития одной или нескольких её логических координат*. Если формальная логика изучает мышление лишь со стороны его логических форм и рассматривает формы мысли как готовые, сформировавшиеся, отвлечаясь от их изменения, развития, то формализмы ЯПЗ ИБД взаимосвязаны с возможностями их преобразования. В [1] представлены результаты многолетнего опыта проектирования проблемно-ориентированных баз знаний научной школы профессора Кедруса В.А., которому принадлежала идея о необходимости разработки специализированного ЯПЗ для приобретения и обработки экономической информации. В настоящее время согласно договору о научно-техническом сотрудничестве в сфере прогрессивных информационных технологий с целью внедрения разработанной методики приобретения знаний [2-19] анализируются бизнес-планы кредитной организации [19, 20] для различных сфер бизнеса.

## ЛИТЕРАТУРА

1. Головешко С.А., Кедрус В.А., Новожилова Т.Н., Соха Ю.А. Экспертные системы – перспективное направление развития информационных технологий на транспорте // Железнодорожный транспорт. – 1998. – №1. – С. 5 - 9.
2. Новожилова Т.Н. Особенности функциональной схемы обобщения процесса капиталобразования в интеллектуальной базе данных // Системи обробки інформації. – Харків: НАНУ, ПАНМ, ХВУ. – 2000. – Вип.1(7). – С. 8 - 15.
3. Новожилова Т.Н. Анализ техники мышления в интеллектуальной базе данных // Системи обробки інформації. – ХФВ “Транспорт України”. – 2000. – Вип.4(10). – С. 107 - 114.
4. Новожилова Т.Н. Классификация объектов в интеллектуальной базе данных // Системи обробки інформації. – ХФВ “Транспорт України”. – 2001. – Вип.3(13). – С. 141 - 145.
5. Новожилова Т.Н. Алгебраические свойства операций интеллектуальной базы данных // Системи обробки інформації. – Харків: НАНУ, ПАНМ, ХВУ. – 2001. – Вип. 4(14). – С. 111 - 115.

6. Новожилова Т.Н. Анализ эффективности логических форм в интеллектуальной базе данных // Системи обробки інформації. – ХФВ “Транспорт України”. – 2001. – Вип.1(11). – С. 22 - 28.

7. Новожилова Т.Н. Интегрированный логический вывод интеллектуальной базы данных // Системи обробки інформації. – Харків: НАНУ, ПАНМ, ХВУ. – 1999. – Вип. 2(6). – С. 9-22.

8. Новожилова Т.Н. Интеллектуальная база данных как средство приобретения знаний // Системи обробки інформації. – Харків: НАНУ, ПАНМ, ХВУ. – 2000. – Вип.1(7). – С. 41 - 49.

9. Новожилова Т.Н. Помехоустойчивость кодирования эвристик в интеллектуальной базе данных // Системи обробки інформації. – Харків: НАНУ, ПАНМ, ХВУ. – 2000. – Вип.1(8). – С. 12 - 18.

10. Новожилова Т.Н. Эвристическая сила структурности экспертных знаний // Вестник ХГПУ. – 2000. – Вып. 78. – С. 8 - 11.

11. Новожилова Т.Н. Структурирование подпространства гипотез // Новые решения в современных технологиях. – Вестник ХГПУ. – 2000. – Вып. 79. – С. 10.

12. Новожилова Т.Н. Многоуровневая схема анализа многоальтернативных представлений // Новые решения в современных технологиях. – Вестник ХГПУ. – 2000. – Вып. 80. – С. 16 - 19.

13. Новожилова Т.Н. Обобщённое индуктивное определение функциональных зависимостей в семантических отношениях // Новые решения в современных технологиях. – Вестник ХГПУ. – 2000. – Вып. 81. – С. 26 - 30.

14. Новожилова Т.Н. Планирование экспертных опросов в интеллектуальной базе данных // Системный анализ, управление и информационные технологии. – Вестник ХГПУ. – 2000. – Вып. 99. – С. 125 - 129.

15. Новожилова Т.Н. Методика приобретения знаний для задач экспертной классификации // Системи обробки інформації. – ХФВ “Транспорт України”. – 2001. – Вип. 2(12). – С. 15 - 20.

16. Новожилова Т.Н. Анализ структуры знаний // Всеукраїнська науково - методична конференція “Гуманізація і гуманітаризація вищої технічної освіти” / ХТУРЕ. – Харків. – 2000. – С. 250 - 252.

17. Новожилова Т.Н. Эвристическая сила структурности // 4-й Международный молодёжный форум «Радиоэлектроника и молодёжь в XXI веке» / ХТУРЭ. – Харьков. – 2000. – С. 313 - 314.

18. Новожилова Т.Н. Язык представления знаний для задач экспертной классификации // Матеріали 2-ї міжнародної конференції «Сучасні проблеми науки та освіти». Ч. 1. / Харків: УАЖНО, ХНУ, ХІУ. – 2001. – С. 204 - 205.

19. Новожилова Т.Н. Экспертная оценка бизнес - планов в интеллектуальной базе данных // 5-й Международный молодёжный форум «Радиоэлектроника и молодёжь в XXI веке» / Харьков: ХТУРЭ, 2001. – С. 193 - 194.

20. Новожилова Т.Н., Галушко И.Ю. Структурный анализ бизнес-планов в среде приобретения знаний. // 7-я Международная конференция «Теория и

техника передачи, приёма и обработки информации» – Харьков: ХТУРЭ, 2001. – С. 374 - 375.

*Поступила в редколлегию 13.08.2001*

---