

## МЕТОДИКА ПРИОБРЕТЕНИЯ ЗНАНИЙ ДЛЯ ЗАДАЧ ЭКСПЕРТНОЙ КЛАССИФИКАЦИИ

к.т.н. Т.Н. Новожилова  
(представил д.т.н., проф. Г.Г. Асеев)

В работе рассматриваются вопросы автоматизации процесса приобретения знаний в интеллектуальной базе данных (ИБД).

Методология моделирования документоведения на основе ИБД находится на этапе становления. В настоящее время лишь начинают складываться единые концепции и взгляды на данную проблематику. Теоретические разработки по проектированию ИБД носят фрагментарный характер и в ряде случаев противоречивы. Это объясняется сложностью учета совокупности параметров, характеризующих схемы баз данных и знаний, а также связью данной проблематики с анализом и критикой мышления [1]. Но *основные проблемы* приобретения знаний по анализу экономической информации связаны с непривычным для экспертов способом извлечения информации. Даже опытные эксперты затрудняются определить априорные и апостериорные вероятности событий, испытывают затруднения при построении функций принадлежности аппарата нечетких множеств [2]. В связи с этим *актуальными* являются вопросы анализа эффективности методов экспертной классификации в ИБД, что, безусловно, сказывается на функциональной схеме среды приобретения знаний (СПЗ) [3]. В данной работе СПЗ рассматривается *как инструмент когнитолога*, позволяющий автоматизировать процесс выявления, моделирования и логического вывода отношений между фактами, сгруппированными по динамическим ситуациям ( $DS_i$ ). ИБД состоит из генератора ИНТЕР-ЭКСПЕРТ; редактора знаний; планировщика; интегрированного вывода; системы словаря справочника базы метазависимостей (СССБМЗ) [3].

Поскольку суждения экспертов, успешно справляющихся с проблемами классификации при анализе  $DS_i$ , являются в основном результатом логических рассуждений, а не использования численных процедур и построения классифицирующих правил, базирующихся на формальных моделях, ИБД позволяет в процессе приобретения знаний осуществлять: анализ логики эксперта содержательными методами; получение от эксперта логической схемы вывода  $DS_i \rightarrow DS_j$ ; обоснование точки зрения эксперта; учет требований эксперта к содержанию и форме аналитических документов; поиск наиболее информативных  $DS_i$  ( $\max CF(DS_i)$ ); учет возможностей эксперта при обработке информации; поиск и исключение ошибок в ответах эксперта; учет статистических закономерностей экспертной классификации.

В ИБД предусмотрены средства структурирования  $DS_i$  по совокупности признаков, используемых экспертом, и стандартные логические схемы решений. На рис.1 представлены соотношения шкал ИБД для классификации экспертных оценок основных соотношений:

$$AN+AF=KS+KZ ; \quad (1)$$

$$IS=KS - AN, \text{ при } IS >0, IS <0, IS =0, \quad (2)$$

где  $A$  – экономические активы:  $AN$  – нефинансовые ( $AND$  - долгосрочные,  $ANO$  - оборотные);  $AF$  – финансовые ( $AFD$  - долгосрочные,  $AFO$  - оборотные);  $K$  – капитал:  $KS$  - собственный;  $KZ$  - заёмный;  $IS$  - индикатор финансово-экономической устойчивости;  $f, nf$  – индексы, соответственно, денежная, не денежная формы.

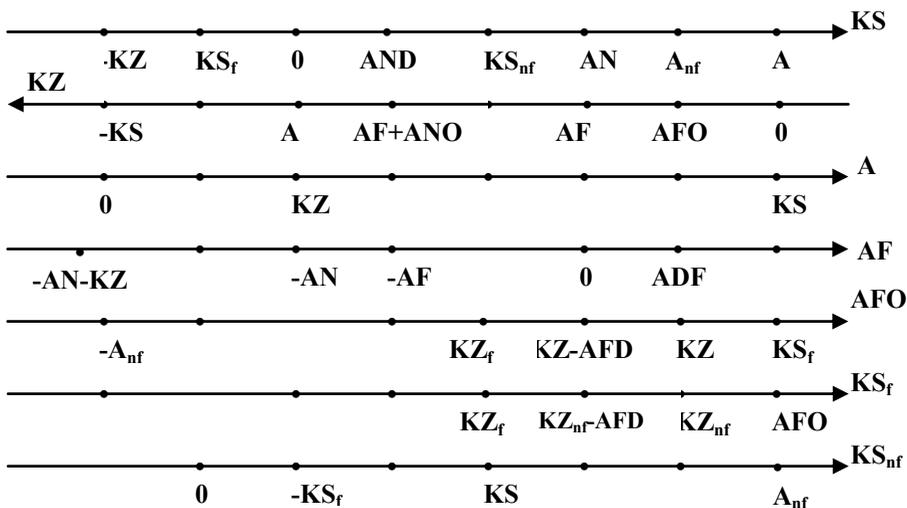


Рис.1. Соотношения шкал ИБД

Из уравнения (1) следует, что

$$(KS - AN)+(KZ-AF)=0. \quad (3)$$

Из (2) и (3) вытекает, что  $IS=AF-KZ$ . Тогда:

$$\text{если } IS>0, \text{ то } (KZ - AF)<0; \quad (4)$$

$$\text{если } IS<0, \text{ то } (KZ - AF)>0; \quad (5)$$

$$\text{если } IS=0, \text{ то } (KZ - AF)=0. \quad (6)$$

Опираясь на бухгалтерский баланс и промежуточные аналитические таблицы ИБД, эксперт прогнозирует финансовое благополучие предприятия на основе экспертной классификации соотношения

$$\frac{\Delta KS}{\Delta A} = \frac{\Delta KS}{\Delta AFO} \times \frac{\Delta AFO}{\Delta AF} \times \frac{\Delta AF}{\Delta KZ} \times \frac{\Delta KZ}{\Delta A}. \quad (7)$$

Механизм изменения **IS** в ИБД интерпретируется как изменение соотношений **AN** и **KS**, **AF** и **KZ**, **KS** и **KZ**, что требует единства формы и содержания экономической информации при анализе темпов роста показателей

$$T_{KS_f} > T_{KS} > T_{AFO} > T_{AF} > T_A > T_{AN} > T_{Anf} > T_{KZ} . \quad (8)$$

В понятия «платежеспособность» и «финансовая устойчивость» в ИБД можно вкладывать разные содержания (экономические), которым соответствуют различные методы анализа и прогнозирования финансовых отношений и, соответственно, различные формы промежуточных аналитических документов [4]. Дальнейшая дифференциация признаков классификации **DS<sub>i</sub>** и сочетание логических и вероятностных признаков при соотношении шкал экспертных оценок финансово-экономического **DS<sub>i</sub>** обеспечивают анализ:

- соотношения заёмных средств и их покрытия;
- определения достаточности или недостаточности **KS** для обеспечения производительных активов;
- раскрытия источников прироста **KS**;
- учёта прохождения имущества через стадии образования, распределения и использования;
- имущества, зафиксированного в балансе предприятия, дохода и собственности (анализ их распределения).

Схеме статистического анализа финансово-экономического состояния предприятия в ИБД соответствует анализ в приростной форме с распределением прироста капитала ( $\Delta K$ ) на прирост  $\Delta KS$  и прирост  $\Delta KZ$ , а активов – на  $\Delta AN$  и  $\Delta AF$ . При этом **IS** трансформируется в приростное чистое кредитование ( $\Delta IS > 0$ ) или в приростное чистое заимствование ( $\Delta IS < 0$ ):

$$\begin{aligned} \Delta IS_{t/0} &= IS_t - IS_0 = (KS_t - AN_t) - (KS_0 - AN_0) = \\ &= (KS_t - AN_0) - (AN_t - AN_0) = \Delta KS_{t/0} - \Delta AN_{t/0}. \end{aligned} \quad (9)$$

От экспертов требуется мотивация присвоения рангов различным типам процессов, с учётом статики и динамики. Ранжирование процессов требует другой последовательности, чем типы переходов: при приобретении эвристик возникает *задача соединения* двух систем показателей – 13 типов процессов и 72 динамических ситуаций. Для этого процессы в ИБД группируются по значениям  $\Delta IS$ . Таким образом, исчисление **DS<sub>i</sub>** позволяет анализировать причины изменений, зафиксированных в балансе, рассматривать добавленную стоимость, механизм её формирования и влияния на прирост активов и их структуру, т.е. проследить:

- как будут изменяться «несущие» экономические показатели в **DS<sub>i</sub>**;
- что представляет собой область финансового равновесия;
- как происходит образование, распределение и использование дохода и собственности;
- как платежеспособность увязывается с ликвидностью;
- как в данной ситуации можно управлять финансовыми ресурсами;
- почему получился именно такой итог бухгалтерского баланса;
- откуда появились «узкие» места и недостаток собственных средств.

Статистическую обработку мнений экспертов предлагается использовать для выяснения *структурных закономерностей* в экспертном знании в области анализа финансовой деятельности предприятия. Поскольку равенство разностей не означает, что равны и сами приросты, внутри каждого блока  $DS_i$  различаются не только положительным или отрицательным приростом  $KZ$ , но также соотношением абсолютных значений приростов  $KS$  и  $KZ$  при равнонаправленном их изменении. При этом учитывается, что даже за краткосрочный период может произойти большое число разнонаправленных изменений общей величины активов и капитала. Чтобы сделать этот процесс управляемым, в параметры, формирующие  $DS_i$  в стандартные схемы синтеза, вводятся соотношения предопределяющие равнодействующую, которая выражается в положительном изменении  $IS$  и позволяет анализировать:

- абсолютную платежеспособность

$$IS > 0, KS - (ANF + AN) > 0; \quad (10)$$

- гарантированную платежеспособность

$$IS > 0, KS - (ANF + AN) < 0; \quad (11)$$

- потенциальную платежеспособность

$$IS < 0, KS - AND > 0; \quad (12)$$

- абсолютную неликвидность

$$IS < 0, KS - AND < 0. \quad (13)$$

Имеются три глобальных цели системы целеполагания: *нарастить* положительный  $IS$ ; *перевести* отрицательное значение  $IS$  в положительное; *ослабить* негативное значение  $IS$ . Таким образом, исчисление ситуаций в ИБД учитывает не только характер динамики индикатора устойчивости, но и типы переходов от одного состояния к другому по 13 рангам [4]:

$$CF(DS_i \rightarrow DS_j) = CF(<H_i, \Delta DS_i, \Delta IS, \Delta R_{na}>), \quad (14)$$

где  $H_i$  - механизм целеполагания,  $CF$  - коэффициент уверенности,  $\Delta R_{na}$  - общий вид изменения операций над  $DS_i$ .

При объяснении процесса аргументации генератор предоставляет когнитологу фрагмент дерева целей, в котором общие цели предшествуют частным. Это является *традиционным* в технологии экспертных систем и *не* является *эффективным* при автоматизации процесса приобретения знаний, поскольку среди общих сведений не выделяются признаки, характеризующие  $DS_i$ . При таком способе объяснений делается попытка упорядочения знаний эксперта в соответствии с *деревом целей*. Такой способ хранения знаний не характерен для эксперта: опытные эксперты хранят в памяти десятки реальных ситуаций и находят их по характерным признакам. Поэтому в ИБД логика структурированных процедур экспертного опроса в СССБМЗ взаимосвязана с логикой организации промежуточных аналитических таблиц, что делает целесообразным организацию процесса приобретения знаний в виде классификации соотношения

$$\bigcup_i^j R_{na}(H_{li}) : DS_i \xrightarrow{H_{li}} DS_j. \quad (15)$$

Поскольку при экспертной классификации используются навыки и приемы, которые эксперту трудно сформулировать, предлагаемый в ИБД способ приобретения знаний соответствует привычной для эксперта задаче анализа конкретных  $DS_i$  и применения различных упрощающих эвристических правил, направленных или на агрегацию информации, или на её последовательное рассмотрение. Но использование подобных правил может *искажать экспертные оценки*, уменьшать их точность, приводить к рассогласованию. Для исключения этих противоречий в ИБД вводятся двухуровневые метазнания. Метатеоремы ИБД позволяют анализировать получаемую от эксперта информацию с точки зрения её согласованности на основе алгебр  $f(CF)$ . Суть процедур экспертного опроса, основанных на построении промежуточных аналитических таблиц, заключается в использовании эпитеорем о структуре решаемой задачи и взаимосвязи между её отдельными элементами. При этом на основе анализа  $f(CF)$  можно делать определенные выводы относительно возможных пределов изменений ряда других экспертных оценок и контролировать неппротиворечивость приобретаемых наборов правил.

Поскольку возможно сравнивать различные  $DS_i$  на основе их описания по степени выраженности  $IS$ , то возможно на основе классификации экспертом  $DS_i$  делать выводы о возможных классах принадлежности других, сравнимых с ней ситуаций. Однако, диагностика  $DS_i$  на основе её описания возможна лишь в том случае, когда различные значения признаков имеют различные  $CF$  для каждого свойства:

$$DS_i R_x(CF_i, \Delta IS_i), DS_j R_x(CF_j, IS_j) \Rightarrow \begin{cases} \mu_0(DS_i) r_i \mu_0(DS_j), CF_i < CF_j; \\ \mu_0(DS_j) r_i \mu_0(DS_i), CF_i > CF_j; \\ DS_i X DS_j, IS_i + \Delta IS_i < IS_j; \\ DS_i X DS_i, IS_i + \Delta IS_i > IS_j; \\ DS_i X DS_j, IS_i + \Delta IS_i = IS_j, \end{cases} \quad (16)$$

где  $X = \sum_{i=1}^n X_i$  - набор правил;  $R_x$  - интенциональное отношение;  $r_i$  - обобщенное отношение,  $\mu <P, V, N, K>$  - учет шкал, на которых расположены

информационные единицы, при выводе новых фактов и оценке самой процедуры вывода зависимостей заданных отображений ( $P$  - по принципу классификации признаков,  $V$  - по отражению взаимосвязи признаков,  $N$  - по наличию начала отсчета,  $K$  - по принципу измерения) [2].

Итак, в процессе приобретения знаний в ИБД эксперту предъявляется бухгалтерский баланс, а исход консультации состоит в классификации  $DS_i$  и предложении типа перехода, согласно  $H_1$ . Для каждой  $DS_i$  определяются возможные переходы и выбираются переходы с  $\max CF(DS_i)$ . Поскольку ИБД учитывает не только характеристики  $DS_i$ , но и понимание роли отдельных

значений признаков в том или ином переходе, это позволяет анализировать возможности эксперта при автоматизации процесса приобретения знаний по анализу экономической информации методами экспертной классификации:

- использовать информацию о характерности  $\Delta DS_i$  для ряда динамических ситуаций без предъявления их эксперту;
- проверять согласованность оценок эксперта через определение классов принадлежности  $DS_i$ ;
- выделять группы близких критериев, определять понятия для этих групп и строить иерархию классификации  $DS_i$ .

Таким образом, классификация  $KS$  по зонам устойчивости дает в общем виде представление о финансовом положении предприятия и позволяет *автоматизировать* в ИБД процесс приобретения знаний по анализу экономической информации. Предлагаемая методика приобретения знаний охватывает широкий класс задач экспертной классификации и позволяет учитывать факторы и направления обобщений при поиске источника собственных средств в необходимой для погашения форме в используемой экспертом модели процесса капиталообразования [5]. Используя априорное разделение фактов по  $DS_i$ , ИБД осуществляет предварительную классификацию предметной области (как в распознавании образов) и позволяет когнитологу:

- соотносить процессы дифференцирования и интегрирования приобретаемых знаний [1,3];
- анализировать причинно-следственные связи [2];
- анализировать технику мышления эксперта при выборе модели капиталообразования [1];
- переводить анализируемые понятия с языка бухгалтерского учета на другие языки обработки экономической информации [2,3,4].

## ЛИТЕРАТУРА

1. Новожилова Т.Н. Анализ техники мышления в интеллектуальной базе данных // Системи обробки інформації. – Харків: НАНУ, ПАНМ, ХВУ. – 2000. – Вип. 4(10). – С. 107 - 114.
2. Новожилова Т.Н. Помехоустойчивость кодирования экспертных знаний // Системи обробки інформації. – Харків: НАНУ, ПАНМ, ХВУ. – 2000. – Вип. 2(8). – С. 12 - 18.
3. Новожилова Т.Н. Особенности функциональной схемы обобщения процесса капиталообразования в интеллектуальной базе данных // Системи обробки інформації. – Харків: НАНУ, ПАНМ, ХВУ. – 2000. – Вип. 3(9). – С. 8 - 15.
4. Новожилова Т.Н. Структурирование подпространства гипотез // Вестник ХГПУ. – Харків: ХГПУ. – 2000. – Вип. 79. – С. 10 - 14.
5. Новожилова Т.Н. Интеллектуальная база данных как средство приобретения знаний // Системи обробки інформації. – Харків: НАНУ, ПАНМ, ХВУ. – 2000. – Вип. 1(7). – С. 41 - 49.

*Поступила в редколлегию 25.12.2000*