УДК 004.358:681.518

З.Л. Костина, С.С. Танянский, В.А. Филатов

ОПЕРАЦИОННАЯ СПЕЦИФИКАЦИЯ РЕЛЯЦИОННОЙ МОДЕЛИ ДАННЫХ В ЗАДАЧАХ ПОДДЕРЖКИ НЕЧЕТКИХ СИСТЕМ

Введение

Компьютерные технологии с организацией интеллектуальных вычислений переживают свой расцвет. Это связано, главным образом, с потоком новых идей, исходящих из области компьютерных наук, которая образовалась на пересечении искусственного интеллекта, статистики и теории баз данных. Предметные области, в которых могут применяться нечеткие базы данных, разнообразны, они требуют соответствующих методов манипуляции нечеткими данными. На практике встречается множество объектов, которые можно оценить не количественными величинами, а только качественными. Такие типы данных часто встречаются в экспертных системах, системах поддержки принятия решений и интеллектуальных базах данных. Поэтому сегодня проблема проектирования нечеткой модели базы данных и методик обработки неточной и абстрактной информации средствами реляционных систем становится все более актуальной.

1. Основные свойства реляционной модели

Рассмотрим классический подход к построению реляционного отношения и выделим основные свойства отношений при расширении множества доменов.

Основной структурной компонентой данных в реляционной модели данных (РМД) является n-арное отношение, представляющее собой подмножество кортежей декартова произведения доменов, то есть множества значений элементов данных. Для заданных конечных множеств $D_1,\ldots, D_n$ (не обязательно различных по типу) декартовым произведением $D_1\times\ldots\times D_n$ называется множество произведений вида $d_1,\ldots, d_n$, где $d_1\in D_1,\ldots, d_n\in D_n$. Отношением R, определенным на множествах $D_1,\ldots, D_n$, называется подмножество произведения (декартово произведение) $D_1\times\ldots\times D_n$, то есть $R\subseteq D_1\times D_2\times\ldots\times D_n$.

Множество $D = \{D_1,\ldots, D_n\}$ называется доменами. Домены − это однотипные семантически однозначные (одинаковых по смыслу) значения элементов данных. Элементы декартова произведения $d_1,\ldots, d_n$ называются кортежами, число n определяет степень отношения, количество кортежей определяет мощность отношения.

Схемой отношения R будем называть выражение $S\ (A_1,\ldots A_n)$, в котором все атрибуты A_i различны. При этом экземпляр отношения $R(S)$ определяется как подмножество декартова произведения доменов $r_i \subseteq \rho(a_i)\times\ldots\times\ \rho(a_n)$.

Экземпляр отношения со схемой R_i будем обозначать как $R_i\,(r_i)$. Отметим, что перестановка атрибутов в схеме не порождает нового состояния БД. Таким образом, множество атрибутов $\{A_1,\ldots, A_n\}$ задает тип отношения и определяет его свойства. Схему БД будем обозначать как множество схем отношений $U = \{R_1,\ldots, R_n\}$, где $R_i\in R$ и все R_i различны. Соответственно, экземпляр БД будем обозначать как множество экземпляров отношений $U\,(r_1,\ldots,r_n)$.

Концептуально реляционная БД является информационной моделью предметной области (ПО), такой, что каждый экземпляр соответствует некоторому состоянию ПО в определенный момент времени. Каждое состояние моделируется упорядоченной совокупностью значений элементов данных, соответствующих значениям свойств объектов ПО. Объекту определенного типа соответствует кортеж отношения. Объекты обладают определенным набором свойств, которые задаются схемой отношения, а свойства имеют определенные наборы возможных значений, которые задаются отображением ρ.

2. Реляционная алгебра

Доступ к реляционным данным осуществляется при помощи реляционной алгебры или эквивалентного ей реляционного исчисления. Реляционная алгебра представляет собой набор операторов, использующих отношения в качестве аргументов, и возвращающих отношения в качестве результата. Каждое отношение обязано иметь уникальное имя в пределах базы данных. Традиционно, определяют восемь реляционных операторов, объединенных в две группы. Теоретико-множественные операторы: объединение, пересечение, вычитание, декартово произведение. Специальные реляционные операторы: выборка, проекция, соединение, деление. Не все они являются независимыми, т.е. некоторые из этих операторов могут быть выражены через другие реляционные операторы.

3. Основные свойства расширения реляционной алгебры

Рассмотрим один из подходов к расширению операций реляционной алгебры (РРА) на отношениях фаззификации [1]. Через $\widetilde{\mathrm{R}}$ будем обозначать множество всех отношений фаззификации и через и R множество обычных реляционных (натуральных) отношений, чьи схемы выбираются из некоторой фиксированной схемы базы данных. Для каждого отношения $\tilde{r} \in \tilde{R}$ определим функцию *Natural*($\tilde{r}$) = $\{r \mid r$ - реляционное (натуральное) отношение$\}$. Эта функция будет использоваться для исследования различных расширений области действия реляционных операций на $\tilde{R}$. Если, например, необходимо доопределить операцию соединения на отношении с нечеткими данными, то для расширенной операции соединения $\tilde{\bowtie}$ должно выполняться равенство (1) для $\tilde{r}, \tilde{s} \in \tilde{R}$.

$$Natural(\tilde{r} \;\tilde{\bowtie}\; \tilde{s}) = Natural(\tilde{r}) \bowtie Natural(\tilde{s}) \quad (1)$$

Сформулируем ряд условий, необходимых для строгого выполнения расширенной операции.

Утверждение 1. Пусть задана операция ξ, определенная на R и операция $\tilde{\xi}$ на $\tilde{R}$. Операция $\tilde{\xi}$ является естественным расширением ξ при выполнении следующих условий:

1) если ξ и $\tilde{\xi}$ - унарные операции, то $\xi(r) = \tilde{\xi}(r)$ для каждого $r \in R$, для которого $\xi(r)$ определена;

2) если ξ и $\tilde{\xi}$ - бинарные операции, то $r\xi s = r\tilde{\xi}s$ для каждого $r,s \in R$, для которого $r\xi s$ определена.

Утверждение 2. Пусть заданы операция ξ на R и операция $\tilde{\xi}$ на $\tilde{R}$. Операция $\tilde{\xi}$ является точным расширением ξ относительно функции *Natural* при выполнении следующих условий:

1) если ξ и $\tilde{\xi}$ - унарные операции, то $Natural(\tilde{\xi}(\tilde{r})) = \xi(Natural(\tilde{r}))$ для каждой $\tilde{r} \in \tilde{R}$.

2) если ξ и $\tilde{\xi}$ - бинарные операции, то $Natural(\tilde{r}\;\tilde{\xi}\;\tilde{s}) = Natural(\tilde{r})\xi Natural(s)$ для всех $\tilde{r}, \tilde{s} \in \tilde{R}$.

Утверждение 3. Пусть заданы операция ξ на R и операция $\tilde{\xi}$ на $\tilde{R}$. Операция $\tilde{\xi}$ является адекватной для ξ относительно функции *Natural* при выполнении следующих условий:

1) если ξ и $\tilde{\xi}$ - унарные операции, то $Natural(\tilde{\xi}(\tilde{r})) \supseteq \xi(Natural(\tilde{r}))$ для каждой $\tilde{r} \in \tilde{R}$.

2) если ξ и $\tilde{\xi}$ - бинарные операции, то $Natural(\tilde{r}\;\tilde{\xi}\;\tilde{s}) \supseteq Natural(\tilde{r})\xi Natural(\tilde{s})$ для всех $\tilde{r}, \tilde{s} \in \tilde{R}$.

Утверждение 4. Пусть заданы операция ξ на R и операция $\tilde{\xi}$ на $\tilde{R}$. Операция $\tilde{\xi}$ является ограниченной для ξ относительно функции *Natural* при выполнении следующих условий:

1) если ξ и $\tilde{\xi}$ - унарные операции, то для каждого $\tilde{r} \in \tilde{R}$ не существует такого $\tilde{S} \in \tilde{R}$, что $Natural(\tilde{\xi}(\tilde{r})) \supset Natural(\tilde{s}) \supseteq \xi(Natural(\tilde{r}))$.

2) если ξ и $\tilde{\xi}$ - бинарные операции, то для каждого $\tilde{r}, \tilde{q} \in \tilde{R}$ не существует такого $\tilde{s} \in \tilde{R}$, что $Natural(\tilde{r}\;\tilde{\xi}\;\tilde{q}) \supset Natural(\tilde{s}) \supseteq Natural(\tilde{r})\xi Natural(\tilde{q})$.

Очевидно, что если операция $\tilde{\xi}$ является точным расширением ξ, то операция $\tilde{\xi}$ адекватна и ограничена для ξ. В тех случаях, когда точное расширение невозможно, будем пользоваться адекватными и ограниченными расширениями. Также необходимо учитывать, что расширенные операции сохраняют свойства обычных реляционных операций, такие как коммутативность и ассоциативность.

4. Операционная спецификация РРА для отношений фаззификации

Рассмотрим расширения некоторых реляционных операций для отношений фаззификации.

Теоретико-множественные операции, рассматриваемые в статье, являются расширением операций над нечеткими множествами, которые применяются к исходным схемам отношений $\tilde{R}$, в результате чего получаются новые отношения, определенные на тех же данных. Таким образом, теоретико-множественные операции будем рассматривать применительно к схемам (2).

$$\tilde{R} = \{\mu_{\tilde{R}}(x) / x\} = \tilde{R}[A(a_i)], \quad (2)$$

где $\mu_{\tilde{R}}(x)$ - функция принадлежности и $A(a_i) = \{a_i \mid i = \overline{1,n}\}$, $x = dom(A(a_i))$.

Основные операции РРА $C_{\tilde{R}}(\tilde{R}[A(a_i)])$, как правило, взаимодействуют не с одной схемой отношений, а с двумя, то есть являются бинарными. Если $A(a_i)$ и $B(b_i)$ носители схем отношений $\tilde{R}$, то ее можно определить выражением (3).

$$C_{\tilde{R}}\,(\tilde{R}_A[A(a_i)]) = \tilde{R}_X[A(a_i \cup b_i)], \quad (3)$$

где $a_i \in A$, $b_i \in B$, $\tilde{R}_X$ – отношение, определенное на нечетком домене $dom(A(a_i) \cup B(b_i))$.

Функция принадлежности множества $\tilde{R}_X$ определяется как выражение (4).

$$\mu_{\tilde{R}_X}(u, v) = \mu_{\tilde{R}_A}(u) \mid \forall (u \in dom(A(a_i), (u, v) \in dom(a_i \cup b_i))) \quad (4)$$

Операция *проекции* отношения фаззификации $\tilde{R}_A[A(a_i) \cup B(b_i)]$, где $\tilde{R}_A$ – отношение, определенное на нечетких доменах $(u, v) \in dom(A(a_i)) \times dom(B(b_i))$, на нечеткое множество $dom(A(a_i))$, определяется выражением (5).

$$\pi_{B_i}\,(\tilde{R}_A[A(a_i \cup b_i)] = \tilde{R}_X[A(a_i)]), \quad (5)$$

где $\tilde{R}_X$ - отношение фаззификации с нечетким доменом $dom(A(a_i))$ и функцией принадлежности (6).

$$\mu_{\tilde{R}_X}(u) = max\{\mu_{\tilde{R}_A}(u, v) \mid u \in dom(A(a_i)), v \in dom(B(b_i)), (u, v) \in (dom(A(a_i)) \times dom(B(b_i)))\}. \quad (6)$$

Графическая интерпретация операции *проекции* в контексте нечетких множеств представлена на рис. 1.

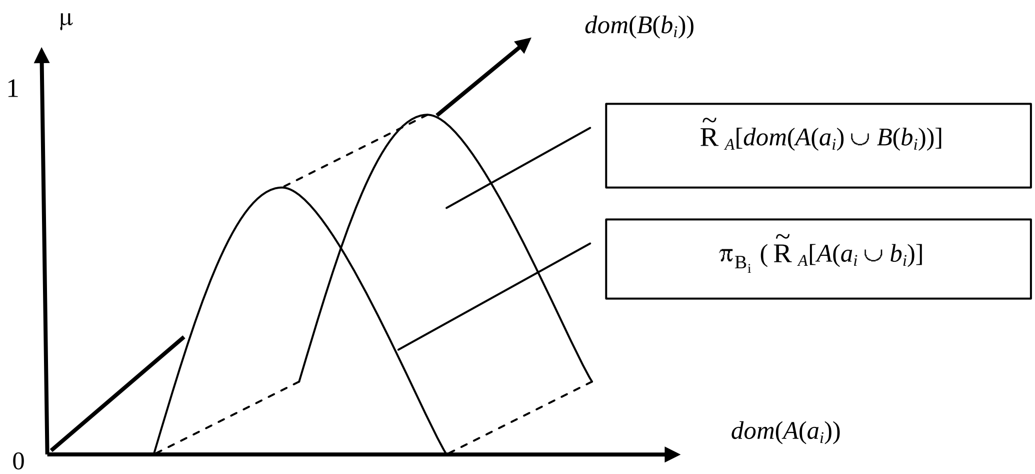

Рис. 1 - Графическая интерпретация операции проекции на отношениях фаззификации.

Операция *произведения* отношений фаззификации $\tilde{R}_A[A(a_i)]$ и $\tilde{R}_B[B(b_i)]$, где $A(a_i) \cap B(b_i) = \varnothing$, представлена выражением (7).

$$\tilde{R}_A[A(a_i)] \times \tilde{R}_B[B(b_i)] = C_{b_i}\,(\tilde{R}_A[A(a_i)]) \cap C_{a_i}\,(\tilde{R}_B[B(b_i)])). \quad (7)$$

Операция *суммы* тех же отношений фаззификации определяется объединением вида (8).

$$\tilde{R}_A[A(a_i)] + \tilde{R}_B[B(b_i)] = C_{b_i}\,(\tilde{R}_A[A(a_i)]) \cup C_{a_i}\,(\tilde{R}_B[B(b_i)])). \quad (8)$$

Операция *селекции* позволяет сформировать новое отношение фаззификации $\tilde{R}_X[B(b_i)]$ по отношению $\tilde{R}_A[A(a_i)]$ на основе проверки условия θ, где θ - формула, построенная с помощью операций логического сравнения (<, >, =, ≤, ≥, ≠) и логических связок ¬ (не), ∧ (и), ∨ (или).

В результате выполнения операции селекции из множества нечетких доменов $dom(B(b_i))$ выбираются домены, удовлетворяющие ограничению (9), для которых условие θ истинно.

$$\{ d_1^1,\ d_2^1, \ldots,\ d_n^1,\ d_1^2,\ d_2^2, \ldots,\ d_m^2, \ldots,\ d_1^p,\ d_2^p, \ldots,\ d_k^p \} \subset dom(A(a_i)), \tag{9}$$

При этом функции принадлежности этих доменов не меняются. Общее выражение операции селекции можно представить в виде формулы (10).

$$\sigma_\theta(\tilde{R}_A[A(a_i)]) = \tilde{R}_X[B(b_i)]. \tag{10}$$

Операция *сравнение* двух отношений фаззификации $\tilde{R}_A[A(a_i)]$ и $\tilde{R}_B[B(b_i)]$ по операции θ истинна, если выполняется условие $A(a_i)\ \theta\ B(b_i)$ и справедливо выражение (11).

$$\forall u \in dom(A(a_i)) \mid \mu_{\tilde{R}_A}(u) = \mu_{\tilde{R}_B}(u). \tag{11}$$

В общем виде операцию сравнения представим выражением (12).

$$\tilde{R}_A[A(a_i)]\ \theta\ \tilde{R}_B[B(b_i)], \tag{12}$$

где θ = {<, >, =, ≤, ≥, ≠}.

Операция *принадлежности* заключается в проверке на включение (содержание) множества $\tilde{R}_A[A(a_i)]$ на множестве $\tilde{R}_B[B(b_i)]$. Операция определяет истинный результат, если $A(a_i) = B(b_i)$ и справедливо выражение (12).

$$\forall u \in dom(A(a_i)) \mid \mu_{\tilde{R}_A}(u) \le \mu_{\tilde{R}_B}(u). \tag{12}$$

В общем виде операцию принадлежности можно представить выражением (13).

$$\tilde{R}_A[A(a_i)] \subseteq \tilde{R}_B[B(b_i)]. \tag{12}$$

Кроме операций, советующих реляционной алгебре, к отношениям фаззификации можно применять традиционные теоретико-множественные операции: дополнение, объединение и пересечение [2, 3].

Операция *дополнения* на отношении фаззификации $\tilde{R}$ для нечеткого домена позволяет получить новое отношение при выполнении ограничений вида (13).

$$\neg \tilde{R} = \bigcup_{u \in \tilde{R}} (1 - \mu_{\tilde{R}}(u) / u). \tag{13}$$

Операция *объединения* двух отношений фаззификации $\tilde{R}_A$ и $\tilde{R}_B$ по нечетким доменам позволяет сформировать отношение, полученное в соответствии с выражением (14).

$$\tilde{R}_A \cup \tilde{R}_B = \bigcup_{u \in \tilde{R}} max[(\mu_{\tilde{R}_A}(u),\ \mu_{\tilde{R}_B}(u)] / u. \tag{14}$$

Операция *пересечения* двух отношений фаззификации $\tilde{R}_A$ и $\tilde{R}_B$ по нечетким доменам отображается в отношение, полученное в соответствии с выражением (15).

$$\tilde{R}_A \cap \tilde{R}_B = \bigcup_{u \in \tilde{R}} min[(\mu_{\tilde{R}_A}(u),\ \mu_{\tilde{R}_B}(u)] / u. \tag{15}$$

Пример

Для иллюстрации предлагаемого подхода сформируем два реляционных отношения фаззификации со схемами: Ra(***SK1,P1***,Z1) и Rb(***SK2,P2***,Z2). Курсивом с подчеркиванием выделены атрибуты, входящие в двойной составной ключ. Представленные отношения позволяют решить задачу хранения значений двух лингвистических переменных A и B, заданных на общей шкале данных {2,4,6,8,10}. Структуры отношений фаззификации с данными приведены на рис.2.

Ra		
SK1	**P1**	Z1
2,00	A	1,00
4,00	A	0,70
6,00	A	0,50
8,00	A	0,30
10,00	A	0,10

Rb		
SK2	**P2**	Z2
2,00	B	1,00
4,00	B	1,00
6,00	B	1,00
8,00	B	0,30
10,00	B	0,20

Рис.2 - Отношения фаззификации Ra и Rb

Для рассматриваемого примера операции объединения (14) и пересечения (15) можно будет эффективно реализовать при помощи запроса:

SELECT Ra.SK1, Ra.P1, Rb.P2, IIf([P1]>[P2],[P1],[P2]) **AS Ob**, IIf([P1]<[P2],[P1],[P2]) **AS Pr**
FROM Ra **INNER JOIN** Rb ON Ra .SK1 = Rb.SK2.

Результат выполнения запроса представлен отношением RESULT, логическая структура которого приведена на рис.3.

RESULT				
SK1	P1	P2	Ob	Pr
2,00	1,00	1,00	1,00	1,00
4,00	0,70	1,00	1,00	0,70
6,00	0,50	1,00	1,00	0,50
8,00	0,30	0,30	0,30	0,30
10,00	0,10	0,20	0,20	0,10

Рис. 3 - Отношение RESULT

Выводы

Таким образом, в статье получены условия расширения основных операций реляционной алгебры для класса систем, спроектированных на основе нечеткой логики. Используя полученные результаты, можно эффективно применять преимущества систем управления баз данных, поддерживающих классическую реляционную модель, для интеллектуальных систем.

ЛИТЕРАТУРА:

1. Касаткина, Н.В. Методы хранения и обработки нечетких данных в среде реляционных систем / Н.В. Касаткина, С.С. Танянский, В.А. Филатов // Автоматика. Автоматизація. Електротехнічні комплекси та системи. – 2009. – Випуск 2(24). – С. 80 – 86.
2. Каргин, А.А. Введение в интеллектуальные машины. Книга 1. Интеллектуальные регуляторы / А.А. Каргин. – Донецк: Норд-Пресс, ДонНУ, 2010. – 526 с.
3. Девятов, В.В. Системы искусственного интеллекта: Учебн. пособие для вузов / В.В. Девятов. – М. : Изд-во МГТУ им. Н.Э. Баумана, 2001. – 352 с.

4. Касаткина, Н. В. Методы хранения и обработки нечетких данных в среде реляционных систем / Н. В. Касаткина, С. С. Танянский, В. А. Филатов // Автоматика. Автоматизація. Електротехнічні комплекси та системи. – Херсон : ХНТУ, 2009. – вип. 2 (24) . – С. 84–90.

ФИЛАТОВ Валентин Александрович
Доктор технических наук, профессор кафедры Искусственного интеллекта Харьковского национального университета радиоэлектроники
Научные интересы: базы данных и знаний, агентные технологии, мультиагентные системы, извлечение знаний из данных.

ТАНЯНСКИЙ Сергей Станиславович
Доктор технических наук, доцент кафедры электронно-вычислительных машин Харьковского национального университета радиоэлектроники.
Научные интересы: модели данных, базы данных и знаний, распределенные информационные системы.

КОСТИНА Зоя Леонидовна
аспирант кафедры Искусственного интеллекта Харьковского национального университета радиоэлектроники
Научные интересы: модели данных, базы данных и знаний, распределенные информационные системы.

ОПЕРАЦІЙНА СПЕЦИФІКАЦІЯ РЕЛЯЦІЙНОЇ МОДЕЛІ ДАНИХ У ЗАДАЧАХ ПІДТРИМКИ НЕЧІТКИХ СИСТЕМ

З.Л. Костіна, С.С. Танянський, В.О. Філатов

Стаття присвячена дослідженню підтримки нечітких систем засобами реляційних баз даних. Розглянуто основні проблеми підтримки операційної специфікації реляційної моделі даних. На основі аналізу основних операцій реляційної алгебри: об'єднання, перетину, різниці і декартова добутку пропонується розширити їх на особливий їхній клас відносин - відносини фазифікації. Наведено приклад, що підтверджує ефективність запропонованого підходу.

ОПЕРАЦИОННАЯ СПЕЦИФИКАЦИЯ РЕЛЯЦИОННОЙ МОДЕЛИ ДАННЫХ В ЗАДАЧАХ ПОДДЕРЖКИ НЕЧЕТКИХ СИСТЕМ

З.Л. Костина, С.С. Танянский, В.А. Филатов

Статья посвящена вопросам поддержки технологии нечетких систем средствами реляционных баз данных. Рассмотрены основные проблемы поддержки операционной спецификации реляционной модели данных. На основе анализа основных операций реляционной алгебры: объединения, пересечения, разности и декартова произведения предлагается расширить их на особый их класс отношений – отношения фаззификации. Приведен пример, подтверждающий эффективность предлагаемого подхода.

OPERATION SPECIFICATIONS THE RELATIONAL DATA MODEL IN PROBLEMS SUPPORTING OF FUZZY SYSTEMS

Z. Kostina, S. Tanyansky, V. Filatov

The article is devoted to supporting technology of fuzzy systems by means of relational databases. The main issues to support the operating specifications of the relational data model. Based on the analysis of the basic relational algebra operations: union, intersection, difference and Cartesian product are proposed to extend them to a special class of their relations - a relation fazzification. An example is given, confirming the effectiveness of the proposed approach.