INFORMATION SYSTEMS

ИНФОРМАЦИОННЫЕ СИСТЕМЫ

УДК 004.2

Г.П. Токмаков, А.Л. Савкин

ФУНКЦИОНАЛЬНАЯ СВЯЗЬ, РЕКВИЗИТЫ И ПОКАЗАТЕЛИ КАК СРЕДСТВА РАЗРАБОТКИ СЕМАНТИЧЕСКИ КОРРЕКТНЫХ СХЕМ БАЗ ДАННЫХ

Токмаков Геннадий Петрович, доктор технических наук, окончил радиотехнический факультет Ульяновского политехнического института. Главный научный сотрудник ФНПЦ АО «НПО «Марс». Профессор кафедры «Вычислительная техника» Ульяновского государственного технического университета. Имеет монографию, учебные пособия, статьи и изобретения в области моделей данных. [e-mail: mars@mv.ru].

Савкин Александр Леонидович, кандидат военных наук, доцент, окончил Ульяновское высшее военное командное училище связи, Военную академию связи им. С.М. Буденного, адъюнктуру при Военной академии связи. Начальник службы обеспечения научно-технической деятельности ФНПЦ АО «НПО «Марс». Имеет научные работы, учебные пособия, статьи в области моделирования систем управления и связи, разработки информационных баз данных. [e-mail: mars@mv.ru].

Аннотация

В статье рассматриваются вопросы моделирования данных, связанные с неоднозначностью представления реальных предметной области (ПрО) через схему отношений. Попытки устранения этого недостатка реляционной модели путем определения функциональных зависимостей (ФЗ) между атрибутами отношений решили эту проблему лишь частично. Объясняется это неформальным характером процесса установления ФЗ, реализуемого при нормализации схем баз данных (БД). В ходе процесса нормализации устраняются аномалии операций обновления, причиной которых является некорректность моделей ПрО, вызванная недостатком семантической информации в реляционных моделях. Для повышения уровня представления семантической информации, применяемой при разработке модели ПрО, предлагается использовать понятия «реквизиты» и «показатели», всесторонне изученные еще в эпоху бумажного документооборота и применяемые для формализации содержания документов. С этой целью в статье представлено описание базовых компонентов реляционной модели (ИнфМ) ПрО в терминах реквизитов и показателей. На основе этого описания для выполнения операций агрегации и обобщения, используемых при формировании реляционной модели, определены формальные критерии, приводящие к снижению субъективного фактора в процессе проектирования концептуальной схемы БД.

Ключевые слова: отношение, реквизит, показатель, агрегация, обобщение, функциональная связь, смысл, семантика, объект, сущность, составная сущность.

doi: 10.35752/1991-2927-2020-3-61-39-49

FUNCTIONAL CONNECTION (BINDING), ESSENTIAL ELEMENTS AND RATING AS TOOLS FOR DEVELOPING A SEMANTIC CORRECT SCHEME FOR DATABASES

Gennadii Petrovich Tokmakov, Doctor of Sciences in Engineering; graduated from Radioengineering Faculty at Ulyanovsk Polytechnic Institute; Chief Scientist of FRPC JSC 'RPA 'Mars', Professor of the Department of Computer Engineering of Ulyanovsk State Technical University; an author of articles, study guides, a monograph and inventions in the field of data models. e-mail: mars@mv.ru.

Aleksandr Leonidovich Savkin, Candidate of Military Science, Associate Professor; graduated from the Ulyanovsk Higher Military Command School of Communications and the Marshal Budjonny Military Academy of Signal Corps; completed his postgraduate studies in the Military Academy of Communications; Head of Science and Engineering Support Department of FRPC JSC 'RPA 'Mars'; an author of scientific works, manuals, and articles in the field of the development and modeling of computer-aided control systems of communications and statistical analysis of supplier appraisal. e-mail: mars@mv.ru.

Abstract

The paper considers an ambiguous representation of the subject area reality in data modelling in terms of the relation scheme. There were some attempts to rectify the relational model drawbacks by determining functional dependences among the relation attributes. Those attempts did not tackle the issue completely due to the specific nature of determining functional dependences during standardization of database scheme. The standardization allows eliminating an anomaly of operation updating caused by incorrect models of the subject area associated with the lack of semantic information in the relation model. To improve the representation process of semantic information, which is applied in designing of subject area, the authors suggest using such terms as essential elements and rating. These terms had been scrutinized ever since the hard-copy document workflow, when they were applied to formalize the document contents. To this purpose, the article describes the main components of the relational model in terms of essential elements and rating. For accomplishing aggregation and generalization operations used in developing the relational model and based on the description, the paper identifies formal criteria enabling to reduce the subjective factor while designing a conceptual database scheme.

Key words: relation, essential element, rating, aggregation, generalization, functional connection, meaning, semantics, object, nature, components essence.

ВВЕДЕНИЕ

Современные информационные системы (ИнФС), используемые для автоматизации деятельности социальных систем в различных предметных областях (ПрО), в качестве одного из основных компонентов содержат централизованную структуру данных, называемую базой данных (БД). Доминирующей на сегодняшний день моделью данных для реализации БД является реляционная, которая в качестве основной единицы представления данных использует понятие «отношения», задаваемого списком имен атрибутов.

На стадии проектирования БД аналитик изучает моделируемую ПрО, выявляет и группирует ее свойства, представляемые как объекты, и определяет связи между ними, руководствуясь положениями реляционной модели. При этом формальные способы группировки атрибутов, необходимые для проверки правильности выбранных схем, в рамках реляционной модели не предлагаются. Поэтому выбор объектов, их атрибутов и фиксация взаимосвязей между ними выполняется аналитиком субъективно в соответствии с его уровнем знаний о моделируемой ПрО и навыками проектирования схем БД.

В этих обстоятельствах для одной и той же ПрО можно создать множество различных схем БД. Это приводит к проблеме выбора корректной схемы реляционной БД, максимально подходящей для решения задач создаваемой ИнФС. В качестве критерия корректности схемы представляется наличие функциональных зависимостей (ФЗ) между атрибутами отношений схемы БД, что гарантирует отсутствие в ней связей, не отвечающих реальному положению дел, и приводящих к побочным эффектам и аномалиям, возникающим при ведении БД.

От таких связей избавляются путем выполнения процедуры последовательного приближения схемы отношений к корректной схеме БД, основанной на теории нормализации. Процедура нормализации предписывает пошаговое устранение нежелательных зависимостей определенных видов, например, частичной и транзитивной. При этом одни схемы БД заменяются на другие, обладающие лучшими свойствами по сравнению с предыдущей. Данная процедура обеспечивает приведение схемы отношений в нормализованное состояние и широко используется на практике. Конкретные алгоритмы нормализации схем отношений описаны в многочисленных работах, например в [1, 2].

Проблема заключается в том, что теорию нормализации очень сложно использовать для моделирования новой схемы ПрО со сложной семантикой. Если даже при этом все необходимые ФЗ установлены, нет никаких гарантий, что они не носят случайный характер и будут действительны на всем множестве значений ее атрибутов. Поэтому данная процедура имеет ограниченное применение и, как правило, используется для модификации существующих схем отношений, но при проектировании новых схем она малоэффективна [3].

В качестве альтернативного подхода к проектированию реляционных БД для выявления корректных ФЗ схем отношений предлагается использовать методы структурирования информации документов, разработанные задолго до появления теории БД, СУБД и языка SQL. Следует отметить, что эти методы уже используются в современных технологиях информационного моделирования, в частности, при разработке унифицированной системы документов ИнФС [4, 5].

1 СОСТОЯНИЕ ПРОБЛЕМЫ

Рассмотрим базовые элементы реляционной модели, которая базируется на понятии элементарной единицы, каковой является значение, зафиксированное на каком-либо материальном носителе и называемое элементом данных [3]. Природа значений, представляемых в модели данных, различается, поэтому они группируются в различные множества, называемые доменами. Набор доменов $D = \{d_1, \dots, d_\delta, \dots, d_\Delta\}$, каждый из которых представляется в виде множества значений $d_\delta = \{z_{1(\cdot)}^\delta, \dots, z_{i(\cdot)}^\delta, \dots, z_{I(\cdot)}^\delta\}$, определяется для каждой моделируемой ПрО.

Примечание. С целью упрощения индексных обозначений в статье:

- используется символ $(\cdot)$ – «точка в скобках», в котором точка заменяет символы верхнего индекса;

- выражения вида $\delta(n)$ заменяются на δn . Количество вложенных символов может быть произвольным, например, выражение $\delta n\phi$ следует читать как $\delta(n(\phi))$.

Домены в моделях используются в качестве области значений атрибутов, представляющих собой информационные модели свойств реальных (предметов и событий) ПрО. Определим множество атрибутов как $A = \{a_{\delta(\cdot)}^1, \dots, a_{\delta(\cdot)}^\alpha, \dots, a_{\delta(\cdot)}^A\}$, где $a_{\delta(\cdot)}^\alpha$ – отдельный атрибут, $\alpha \in 1, \dots, A$ – его имя, а $\delta(\alpha)$ – его домен как множество значений. Данное множество является базовым элементом реляционной модели, на основе которого далее строятся более сложные конструкции, называемые объектами, а также отношения, объединяющие классы однородных объектов.

В реляционной модели объекты определяются как агрегаты [6], построенные на множе-

стве атрибутов $\{a_{\delta(\cdot)}^1, \dots, a_{\delta(\cdot)}^n, \dots, a_{\delta(\cdot)}^N\} \subset A$. Выполнив операции агрегации для всех допустимых наборов атрибутов ПрО, определим множество объектов $T = \{T_1, \dots, T_o, \dots, T_O\}$, соответствующих реализам ПрО, где $T_o \rightarrow \langle a_{\delta(\cdot)}^{1o}, \dots, a_{\delta(\cdot)}^{no}, \dots, a_{\delta(\cdot)}^{No} \rangle$, а $a_{\delta(\cdot)}^{no}$ – атрибут объекта T_o с именем n . Для обзора всех возможных объектов отношения допустимые наборы атрибутов ПрО опишем в виде декартова произведения $a_{\delta(\cdot)}^1 \times \dots \times a_{\delta(\cdot)}^n \times \dots \times a_{\delta(\cdot)}^N$, представляемого множеством кортежей значений атрибутов $\langle z_{i(\cdot)}^{\delta 1}, \dots, z_{i(\cdot)}^{\delta n}, \dots, z_{i(\cdot)}^{\delta N} \rangle$, для всех, $i = 1, \dots, I$.

Но не все сочетания значений атрибутов соответствуют объектам ПрО, а только те из них, которые удовлетворяют некоторым условиям. Путем отбора кортежей декартова произведения, соответствующих этим условиям, формируется набор однородных кортежей, представляющих экземпляры некоторой реализации ПрО, называемый в реляционной модели *отношением*. В современной практике информационного моделирования условия, которым должны соответствовать кортежи, представляющие реализации ПрО, формально не определены, а соблюдаются пользователем, осуществляющим ввод наборов данных кортежей, соответствующих этим условиям. При этом он исходит из своих субъективных представлений о природе моделируемой реализации. Описанное отношение определяет *сущность* E , объединяющую в себе однородные объекты, описываемые отобранными кортежами. При этом для каждой сущности формируется отдельное описание, включающее в себя только общие для всех ее объектов элементы. Процедура, в соответствии с которой множество однородных объектов представляется как единая сущность, называется *операцией обобщения*. В результате выполнения операции обобщения ИнФМ ПрО множество объектов T преобразуется в множество сущностей $E = \{E_1, \dots, E_\xi, \dots, E_\Xi\}$, каждая из которых описывает множество однородных объектов $E_\xi \rightarrow (T_1^\xi, \dots, T_\varepsilon^\xi, \dots, T_E^\xi)$. При этом формируется концептуальное представление ПрО, согласно которому каждый объект $T_o \in T$ описывается как экземпляр T_ε^ξ некоторой сущности E_ξ .

По приведенному краткому обзору реляционной модели можно установить источники проблем с моделированием ПрО, связанные с неоднозначностью ее представления в виде схем отношений. Эти проблемы вызваны, прежде всего, недостатком семантической информации, учитываемой при проектировании схем отношений, и заключаются в том, что:

- набор атрибутов для выполнения операции агрегации определяется проектировщиком, исходя из потребностей конкретных задач ПрО и его понимания

природы моделируемых реалий ПрО. Формальные критерии, на основе которых реализуется операция агрегации, отсутствуют [7];

- ограничения, накладываемые на значения атрибутов при выполнении операции обобщения, т. е. определении сущностей, не формализованы, и реализуются при вводе данных пользователем субъективно, исходя из его понимания природы ПрО.

2 ФЗ и ПОКАЗАТЕЛИ КАК СРЕДСТВА ФОРМАЛЬНОГО ПРЕДСТАВЛЕНИЯ СМЫСЛОВЫХ СВЯЗЕЙ МЕЖДУ АТТРИБУТАМИ ОТНОШЕНИЙ

Многочисленные теоретические исследования в области моделирования данных, выполненные с целью совершенствования реляционной модели, привели к пониманию того, что для улучшения ее качества следует обратить внимание на закономерности, не зависящие от специфики отдельных ПрО. В качестве такой закономерности в настоящее время рассматривается понятие ФЗ данных, накладывающей ограничения на текущие значения схемы отношений.

2.1 ФЗ как средство формализации схем отношений

ФЗ атрибута $a_{\delta(\cdot)}^2$ отношения E_ξ от атрибута $a_{\delta(\cdot)}^1$ того же отношения обозначается выражением $E_\xi \cdot a_{\delta(\cdot)}^1 \rightarrow E_\xi \cdot a_{\delta(\cdot)}^2$ и определяет, что в каждый момент времени любому элементу проекции $E_\xi \left[a_{\delta(\cdot)}^1 \right] = z_{i(\cdot)}^1$ соответствует только один элемент проекции $E_\xi \left[a_{\delta(\cdot)}^2 \right] = z_{i(\cdot)}^2$, входящий вместе с ним в какой-либо кортеж отношения E_ξ .

ФЗ отражает не только текущее состояние схемы отношений, а определяется всей совокупностью возможных состояний ее атрибутов. Поэтому если в другой момент времени значение $z_{i(\cdot)}^1$ однозначно определяет значение $z_{j(\cdot)}^2$, то это говорит о том, что значения $z_{i(\cdot)}^1$ и $z_{j(\cdot)}^2$ представляют атрибут $a_{\delta(\cdot)}^2$ в различные моменты времени [8].

Из этого следует, что ФЗ инвариантна к изменению состояний модели данных во времени [6]. Именно это свойство ФЗ позволяет использовать ее в качестве универсального знания о ПрО, используемого при нормализации схем отношений. Для проверки наличия ФЗ между атрибутами отношения необходимо:

- определить ее аргумент (т. е. ключевой атрибут отношения);
- установить однозначное соответствие между значением ключевого атрибута и значениями других атрибутов отношения.

Если проверка наличия ФЗ дает положительный результат со всеми неключевыми атрибутами отношения, то считается, что отношение определено корректно. В противном случае, атрибуты, не зависящие от ключевого

го атрибута отношения, выделяются в отдельную сущность. Данная процедура повторяется до тех пор, пока для всех атрибутов отношений ПрО не будет установлено ФЗ.

Но приведенное определение ФЗ позволяет проверить ее наличие только для отношений, которые уже существуют. Определить новую сущность на основе ФЗ довольно сложно, т. к. она определяется, как уже было отмечено, не только по текущему состоянию ПрО, а на всем множестве ее настоящих, прошлых и возможных будущих состояний. Но в реальной жизни это условие невыполнимо. Поэтому при отсутствии полного описания отношений сложно установить ФЗ, отвечающие природе данных моделируемой ПрО, и можно выявить ФЗ, носящие случайный характер и не отражающие реальные связи между свойствами ПрО.

Фактически ФЗ представляют собой смысловые связи между свойствами ПрО и по этой причине не могут быть получены формальными методами. Они устанавливаются аналитиками после глубокого изучения ПрО, но «... являясь высказываниями о сущностях ПрО, они не могут быть доказаны [6]». Именно это обстоятельство порождает неоднозначность представления ПрО отношениями реляционной БД. Эта неоднозначность в полной мере не устраняется даже после проведения процедуры нормализации.

Для решения проблем проектирования корректных схем БД могут оказаться полезными понятия, сформированные еще в докомпьютерную эпоху, когда ИнФМ создавались на естественном языке при составлении документов, используемых для управления деятельностью той или иной организации. В ходе довольно длительного периода существования этой технологии информационного моделирования были выработаны такие понятия, как реквизиты и показатели, описывающие свойства реалий внешнего мира и их взаимосвязей, отражаемых в документах.

2.2 Реквизит как элемент данных реляционной модели

Свойства предметов и событий внешнего мира, отражаемые в документах с помощью переменных величин, принято называть реквизитами. В документах реквизиты представляются в текстовом виде, например, «Вес – 100 кг», «Цена – 200 р». Как видно из примеров, каждый реквизит имеет имя и значение из заданного домена, поэтому обозначим реквизит символом $r_{\delta(\cdot)}^p$, где индекс:

- p определяет его название (например, «Вес»);

- δ определяет домен $d_\delta = \{z_{i(\cdot)}^\delta, \dots, z_{i(\cdot)}^\delta, \dots, z_{I(\cdot)}^\delta\}$,

i -е значение $z_{i(\cdot)}^\delta$ которого представляется как значение реквизита в некоторый момент времени (например, «100 кг»).

Полное обозначение реквизита, в котором отражается его наименование p , домен $\delta(p)$ и значение $z_{i(\cdot)}^{p\delta}$, будем описывать выражением $r_{\delta(\cdot)}^p \left(z_{i(\cdot)}^{p\delta} \right)$. При совмест-

ном представлении наименования и значения реквизита можно использовать сокращенные записи индексов и представить реквизит выражением $r_{\delta(\cdot)}^p(z_i)$, где z_i – значение из домена $\delta(p)$, определенного областью значений для реквизита с наименованием p .

Понятие реквизита является аналогом понятия атрибута реляционной модели, поэтому в дальнейшем изложении в зависимости от контекста будем использовать эти термины как синонимы. Значимость понятия реквизита с точки зрения цели данной статьи заключается в том, что в рамках его определения свойства делятся на реквизиты-признаки и реквизиты-основания:

- реквизиты-признаки определяют качественные свойства объекта и отражают неизменные, инвариантные характеристики объектов ПрО. Значения таких свойств используются для выполнения логической обработки их составных единиц при поиске, выборке, сортировке и т. д. Наименование реквизита-признака будем представлять символом $k_{\delta(\cdot)}^\pi$, а полное обозначение с указанием наименования и значения – выражением $k_{\delta(\cdot)}^\pi(z_{j(\cdot)}^{\delta\pi})$ или в сокращенном варианте $k_{\delta(\cdot)}^\pi(z_j)$ ($\pi = 1, \dots, \Pi, j = 1, \dots, J$), где символ k обозначает реквизит-признак, используемый в реляционной модели в качестве ключевого атрибута отношения;

- реквизиты-основания представляют свойства, определяющие состояние объекта или процесса. Значения таких свойств можно определить инструментально и описать количественно, либо выбрать из конечного перечня, не обязательно числового. Такие свойства принято называть параметрами [7]. Над ними могут выполняться арифметические или логические операции. Наименование реквизита-основания будем представлять символом $o_{\delta(\cdot)}^\omega$, а полное обозначение с указанием наименования и значения – выражением $o_{\delta(\cdot)}^\omega(z_{i(\cdot)}^{\delta\omega})$ или в сокращенном варианте $o_{\delta(\cdot)}^\omega(z_i)$ ($\omega = 1, \dots, \Omega, i = 1, \dots, I$), где символ o обозначает реквизит-основание, используемый в реляционной модели в качестве неключевого атрибута отношения.

Принципиальное различие описанных видов свойств заключается в том, что при допустимом изменении состояния, определяемого значениями $o_{\delta(\cdot)}^\omega(z_i)$, объект сохраняется как таковой, тогда как при изменении значения качественного свойства, т. е. $k_{\delta(\cdot)}^\pi(z_j)$, идентифицируется другой объект. Поэтому каждое значение реквизита-признака (идентификационного свойства) определяет отдельный объект и используется для его идентификации, а с помощью реквизитов-оснований (характеристических свойств) определяются состояния этих объектов или отдельных свойств в различные моменты времени.

2.3 Показатель как средство представления ФЗ

Взятое в отдельности свойство, т. е. реквизит, не имеет смысла и не может служить информацией о наблю-

даемых объектах и процессах ПрО. Для придания смысла значениям реквизитов необходима совокупность реквизитов, определяющая соотношенность свойств с объектами. Если некоторый реквизит-признак представить совместно с реквизитом-основанием, то его значение будет восприниматься как характеристика объекта, идентифицируемого данным реквизитом-признаком. Таким образом, совокупность реквизита-признака и реквизита-основания образует структурную единицу представления информации, называемую показателем и относящую свойство к его носителю, т. е. объекту [7]. В соответствии с приведенным определением показатель представляется выражением

$$q_{i\omega}^{\pi} \rightarrow k_{\delta(\cdot)}^{\pi}(z_j), o_{\delta(\cdot)}^{\omega}(z_i),$$

где $k_{\delta(\cdot)}^{\pi}$ – реквизит-признак показателя с наименованием π , а $z_{j(\cdot)}^{\delta\pi}$ представляет j -е значение домена $d_{\delta\pi}$, идентифицирующее некоторый объект ПрО;

$o_{\delta(\cdot)}^{\omega}$ – реквизит-основание показателя с наименованием ω , а $z_{i(\cdot)}^{\delta\omega}$ представляет i -е значение домена $d_{\delta\omega}$, определяющее состояние свойства ω объекта, идентифицируемого j -м значением реквизита-признака $k_{\delta(\cdot)}^{\pi}$;

$q_{i\omega}^{\pi}$ – обозначение показателя, реквизит-признак которого имеет имя π и j -е значение домена, а реквизит-основание – имя ω и i -е значение домена.

Показатель как средство представления данных имеет следующий состав:

- значение $z_{i(\cdot)}^{\delta\omega}$ реквизита-основания;
- наименование значения (или его представление)

$k_{\delta(\cdot)}^{\pi}(z_j)[o_{\delta(\cdot)}^{\omega}]$, составленное из наименования и значения реквизита-признака, а также наименования реквизита-основания в терминах естественного языка.

Пары значений реквизита-признака и реквизита-основания, составляющие показатель, определяются человеком на этапе разработки документов, исходя из его знаний внешнего мира. Но если множество таких пар значений занести в память ИнФС, то в дальнейшем ФЗ между ними можно будет определять программными средствами. Например, если значение «12345» свойства «табельный номер» идентифицирует некоторого сотрудника, а в некотором документе оно связывается со значением «инженер» свойства «должность» сотрудника, то можно сказать, что между этими значениями имеет место ФЗ. Если эту связь зафиксировать в какой-либо форме в памяти ИнФС, то ее можно будет использовать для программной интерпретации значения «инженер», т. е. по значению «12345» свойства «табельный номер» программа однозначно определит значение «инженер» свойства «должность».

Следовательно, все, что говорилось о ФЗ справедливо и для показателя, но понятие показателя по сравнению с понятием ФЗ обладает следующими преимуществами:

- реквизит-признак показателя моделирует свойство-идентификатор объекта ПрО и по определению используется в качестве аргумента ФЗ;

- субъективный фактор, имеющий место при определении смысла содержания документов, представленных совокупностями показателей, минимизируется. Это объясняется тем, что формы типовых документов, представленные совокупностями показателей, проходят процедуры унификации и стандартизации, т. е. объективированы. Это снижает степень субъективизма при определении ФЗ ПрО.

Важно иметь в виду, что при информационном моделировании должны отражаться не только видимые проявления свойств объектов, но и взаимосвязи между ними. Эти взаимосвязи нематериальны и существуют только в сознании человека, которые он устанавливает, используя свою систему знаний о внешнем мире [7]. Сказанное относится и к связи «объект – свойство», которую человек устанавливает по факту одномоментного проявления группы свойств, представляемой как объект.

Когда специалист, составляя по текущему состоянию ПрО некоторый документ, определяет ФЗ между конкретными значениями реквизитов-признаков и реквизитов-оснований, он устанавливает смысловые связи между реквизитами-признаками и реквизитами-основаниями безотносительно их конкретных значений. Этим отражается то обстоятельство, что некоторые пары свойств-идентификаторов и свойств-параметров представляют один объект, идентифицируемый значением реквизита-признака.

Таким образом, при составлении документов с помощью показателей материализуется и представляется в виде знаков типовое смысловое значение «объект и его свойство» [7]. Эти смысловые значения выявляются аналитиком исходя из содержания анализируемых документов и фиксируются в унифицированных формах документов в виде множества показателей или единиц семантической информации. Фиксация в ИнфМ элементарных семантических конструкций позволяет на их основе строить более сложные семантические конструкции, что будет продемонстрировано ниже.

3 ОПИСАНИЕ ОБЪЕКТОВ РЕЛЯЦИОННОЙ МОДЕЛИ В ТЕРМИНАХ РЕКВИЗИТОВ И ПОКАЗАТЕЛЕЙ

В приведенном выше определении показателя не содержится явного указания на объект, чьи идентификаторы и свойства-параметры представляются реквизитами-признаками и реквизитами-основаниями, хотя известно, что «данные наделяются смыслом только при их интерпретации через понятие объекта [7]». Это противоречие разъясняется, если учесть, что классическое определение показателя было выработано в эпоху бумажного документооборота, когда информация, представленная в документах, воспринималась человеком, который недостающую информацию мог «домысли-

вать», исходя из своих представлений о внешнем мире. Например, по реквизиту-признаку *Фамилия (Петров)* человек делает вывод, что *Петров* – это значение свойства *Фамилия* сущности *Человек*.

При информационном моделировании ПрО такие «умолчания» не допускаются, и вся информация, необходимая для интерпретации элементов данных должна быть представлена явно. Поэтому с целью использования понятия «показатель» при информационном моделировании ПрО, когда любое значение, представляющее состояние свойства некоторого объекта, должно быть однозначно идентифицировано, сформулируем трактовку определения показателя, в котором содержится явное указание на «объект», имея в виду «экземпляр сущности». Для этого в качестве основы возьмем классическое определение показателя, описываемое выражением, приведенным подразделе 2.3, а затем на его основе выведем определения понятий «объект», «сущность» и «составная сущность». В ходе вывода этих определений проследим трансформацию описания показателя, в котором сначала определяется ссылка на «объект», а затем – на «экземпляр сущности».

3.1 Определение исходных элементов описания ИнфМ

В качестве первого шага на пути к поставленной цели определим исходные элементы создаваемой ИнфМ, представленные множествами:

- реквизитов ПрО, разделенных на категории;
- показателей ПрО, определяющих ФЗ реквизитов-оснований от реквизитов-признаков.

Исходные данные для формирования ИнфМ определяются в ходе анализа системы документов ПрО путем выявления реквизитов, определяющих содержание ее документов. При этом предварительно проводится унификация наименований реквизитов, отражающих в разных документах одни и те же свойства ПрО. В результате формируется множество реквизитов

$R = \{r_{\delta(.)}^1, \dots, r_{\delta(.)}^p, \dots, r_{\delta(.)}^P\}$, которое делится на подмножества:

- реквизитов-признаков

$$K = \{k_{\delta(.)}^1(z_j), \dots, k_{\delta(.)}^{\pi}(z_j), \dots, k_{\delta(.)}^{\Pi}(z_j)\};$$

- реквизитов-оснований

$$O = \{o_{\delta(.)}^1, \dots, o_{\delta(.)}^{\omega}, \dots, o_{\delta(.)}^{\Omega}\}, \text{ таких, что } K \cup O = R.$$

Индексам $\pi = 1, \dots, \Pi$ и $\omega = 1, \dots, \Omega$, отражающим порядковые номера, соответствуют наименования реквизитов, поэтому в дальнейшем изложении будем представлять их именами реквизитов. Множество доменов D при делении множества R на подмножества O и K остается неизменным и используется для определения как реквизитов-оснований, так и реквизитов-признаков.

Множество показателей Q определяется в ходе содержательного анализа системы документов ПрО путем выявления ФЗ между элементами множеств O и K и описывается выражением

$$Q = \{q_{i\omega_1(\cdot)}^{1\pi}, \dots, q_{i\omega_n(\cdot)}^{j\pi}, \dots, q_{i\omega_N(\cdot)}^{J\pi}\},$$

где $q_{i\omega(\cdot)}^{j\pi}$ – показатель, наименование $k_{\delta(\cdot)}^{\pi}(z_j)[o_{\delta(\cdot)}^{\omega}]$ которого используется для представления значения $z_{i(\cdot)}^{\delta\omega}$. Индекс i при символе показателя указывает на значение его реквизита-основания, которое при одном и том же значении реквизита-признака j может меняться с течением времени. Поэтому в некоторых случаях этот индекс в обозначении показателя опускается.

Семантика в определении показателя заключается в том, что для пар свойств-идентификаторов и свойств-параметров ПрО в ИнфМ для соответствующих пар атрибутов фиксируется смысловая связь «идентификатор-параметр», указывающая на отнесенность любого свойства-параметра к некоторому объекту.

3.2 Агрегация показателей – определение объектов

При рассмотрении базовых элементов реляционной модели было отмечено, что определяя отношение как подмножество декартова произведения множеств значений атрибутов, авторы реляционной модели не дают никаких указаний относительно критериев выбора атрибутов этого произведения. Эту задачу решает разработчик концептуальной схемы ПрО, полагаясь на свой практический опыт и интуицию, одним словом, субъективно.

Для установления формального критерия для выборки атрибутов операции агрегации обратимся к множеству показателей Q и выделим в нем показатели, реквизиты-признаки которых имеют одинаковые наименование π и значение z_j . Выделенная таким образом группа показателей представляется выражениями вида:

$$q_{\omega_1(\cdot)}^{j\pi} \rightarrow k_{\delta(\cdot)}^{\pi}(z_j), o_{\delta(\cdot)}^{\omega_1}(z_i),$$

...

$$q_{\omega_n(\cdot)}^{j\pi} \rightarrow k_{\delta(\cdot)}^{\pi}(z_j), o_{\delta(\cdot)}^{\omega_n}(z_i),$$

...

$$q_{\omega_N(\cdot)}^{j\pi} \rightarrow k_{\delta(\cdot)}^{\pi}(z_j), o_{\delta(\cdot)}^{\omega_N}(z_i),$$

где $\{q_{\omega_1(\cdot)}^{j\pi}, \dots, q_{\omega_n(\cdot)}^{j\pi}, \dots, q_{\omega_N(\cdot)}^{j\pi}\} \subset Q$, а верхний индекс $j(\pi)$ следует интерпретировать как «реквизит-признак $k_{\delta(\cdot)}^{\pi}$ со значением z_j^{π} ».

Выполнение описанной группировки показателей на всем множестве Q приводит к его разделению на подмножества

$$Q = \begin{bmatrix} Q_{1(\cdot)}^1 & \dots & Q_{j(\cdot)}^1 & \dots & Q_{J(\cdot)}^1 \\ \vdots & & \vdots & & \vdots \\ Q_{1(\cdot)}^{\pi} & \dots & Q_{j(\cdot)}^{\pi} & \dots & Q_{J(\cdot)}^{\pi} \\ \vdots & & \vdots & & \vdots \\ Q_{1(\cdot)}^{\Pi} & \dots & Q_{j(\cdot)}^{\Pi} & \dots & Q_{J(\cdot)}^{\Pi} \end{bmatrix},$$

где $Q_{j(\cdot)}^{\pi}$ представляет группу показателей с общим реквизитом-признаком $k_{\delta(\cdot)}^{\pi}(z_j)$, где $\pi = 1, \dots, \Pi$ и $j = 1, \dots, J$. Значение $J(\pi)$ определяет количество значений для каждого реквизита-признака π , которые суммарно, т. е. для всех $\pi = 1, \dots, \Pi$, определяют количество групп показателей на множестве Q .

Анализируя выражение, описывающее группу показателей, примем во внимание то, что в соответствии с определением показателя (см. подраздел 2.2) пара реквизита-признака и реквизита-основания представляет значение свойства объекта, идентифицируемого значением реквизита-признака. Учитывая это обстоятельство, можно утверждать, что реквизиты-основания данной группы представляют атрибуты-параметры (т. е. неключевые атрибуты) объекта, идентифицируемого значением общего реквизита-признака. Для отражения этого факта в обозначениях объектов ПрО $T_{j(\cdot)}^{\pi}$ будем применять индексы при соответствующих группах показателей.

Для вывода выражения, описывающего объект $T_{j(\cdot)}^{\pi}$, представим показатели группы в виде последовательности и выведем за скобки общий реквизит-признак, оставив в скобках зависящие от него реквизиты-основания. В результате получим реквизитный состав объекта, описываемый выражением:

$$T_{j(\cdot)}^{\pi} \rightarrow k_{\delta(\cdot)}^{\pi}(z_j) [o_{\delta(\cdot)}^{\omega_1}(z_i), \dots, o_{\delta(\cdot)}^{\omega_n}(z_i), \dots, o_{\delta(\cdot)}^{\omega_N}(z_i)],$$

где $T_{j(\cdot)}^{\pi}$ – наименование объекта ПрО, идентифицируемого значением z_j реквизита-признака $k_{\delta(\cdot)}^{\pi}$;

$o_{\delta(\cdot)}^{\omega_1}, \dots, o_{\delta(\cdot)}^{\omega_n}, \dots, o_{\delta(\cdot)}^{\omega_N}$ – реквизиты-основания группы показателей $Q_{j(\cdot)}^{\pi}$, представляющие атрибуты-параметры объекта $T_{j(\cdot)}^{\pi}$;

$z_{i(\cdot)}^{\delta\omega_1}, \dots, z_{i(\cdot)}^{\delta\omega_n}, \dots, z_{i(\cdot)}^{\delta\omega_N}$ – значения реквизитов-оснований группы показателей, определяющие i -е состояние объекта $T_{j(\cdot)}^{\pi}$ на текущий момент времени. Это говорит о том, что каждый объект $T_{j(\cdot)}^{\pi}$ может пребывать в различных состояниях, оставаясь самим собой, т. е. при сохранении значения z_j реквизита-признака $k_{\delta(\cdot)}^{\pi}$. В реляционной модели реквизит-признак объекта называется первичным ключом.

Выполнив описанную операцию агрегации для всех выявленных групп показателей, определим множество объектов

$$\tilde{T} = \begin{bmatrix} T_{1(.)}^1 & \dots & T_{j(.)}^1 & \dots & T_{J(.)}^1 \\ \vdots & & \vdots & & \vdots \\ T_{1(.)}^\pi & \dots & T_{j(.)}^\pi & \dots & T_{J(.)}^\pi \\ \vdots & & \vdots & & \vdots \\ T_{1(.)}^\Pi & \dots & T_{j(.)}^\Pi & \dots & T_{J(.)}^\Pi \end{bmatrix},$$

представляющих всю ПрО. Обратите внимание на то, что составы множества объектов T , определенного в рамках описания реляционной модели (см. раздел 1), и множества объектов $\tilde{T}$, описанного в данном подразделе, различаются. Причина этого явления – неоднозначность представления ПрО через понятие отношения.

Нетрудно заметить, что описанная процедура формирования реквизитного состава объекта представляет собой операцию агрегации, в ходе выполнения которой совокупность компонентов (в данном случае реквизитов) выражается одним именем, т. е. именем объекта $T_{j(.)}^\pi$. Использование понятия показателей при описании операции агрегации дает нам то преимущество, что в этом случае состав атрибутов $k_{\delta(.)}^\pi, o_{\delta(.)}^{\omega_1}, \dots, o_{\delta(.)}^{\omega_n}, \dots, o_{\delta(.)}^{\omega_N}$ объекта $T_{j(.)}^\pi$ выводится по формальному критерию. Таким критерием является значе-

ние z_j общего реквизита-признака $k_{\delta(.)}^\pi$ группы показателей $q_{\omega_1}^{j\pi}, \dots, q_{\omega_n}^{j\pi}, \dots, q_{\omega_N}^{j\pi} \subset Q$.

Семантика в определении объекта заключается в том, что для некоторой группы свойств ПрО в ИнфМ для соответствующей группы атрибутов фиксируется смысловая связь «объект-свойство».

3.3 Обобщение объектов – определение сущностей

При определении реляционной модели используется еще одна операция, которая делает акцент на сходстве объектов, абстрагируясь от их различий. Эта операция позволяет представлять множество или класс однородных объектов одной сущностью и называется обобщением. С помощью операции обобщения бесконечное разнообразие объектов сводится к относительно небольшому набору сущностей, а каждый объект представляется как экземпляр некоторой сущности. При этом для каждой сущности определяется набор атрибутов, характеризующий класс однородных объектов.

Для выполнения операции обобщения на множестве $\tilde{T}$ выделим группы описаний объектов, идентификаторы которых представлены одним и тем же реквизитом-признаком, имеющим различные значения. При этом обнаружится, что объекты сформированных таким образом групп, имеют одинаковый реквизитный состав, т. к. каждый реквизит-признак вне зависимости от его значений определяет функциональные связи с одними и теми же реквизитами-основаниями. Объекты, идентифицируемые множеством значений одного реквизита-признака, можно описать выражениями:

$$\begin{aligned} T_1^\pi &\rightarrow k_{\delta(.)}^\pi(z_1) \left[o_{\delta(.)}^{\omega_1}(z_{i(.)}) \dots o_{\delta(.)}^{\omega_n}(z_{i(.)}) \dots o_{\delta(.)}^{\omega_N}(z_{i(.)}) \right], \\ &\vdots \\ T_{j(.)}^\pi &\rightarrow k_{\delta(.)}^\pi(z_j) \left[o_{\delta(.)}^{\omega_1}(z_{i(.)}) \dots o_{\delta(.)}^{\omega_n}(z_{i(.)}) \dots o_{\delta(.)}^{\omega_N}(z_{i(.)}) \right], \\ &\vdots \\ T_{J(.)}^\pi &\rightarrow k_{\delta(.)}^\pi(z_J) \left[o_{\delta(.)}^{\omega_1}(z_{i(.)}) \dots o_{\delta(.)}^{\omega_n}(z_{i(.)}) \dots o_{\delta(.)}^{\omega_N}(z_{i(.)}) \right], \end{aligned}$$

где $T_{j(.)}^\pi$ – объект-экземпляр сущности, идентифицируемый значением z_j атрибута $k_{\delta(.)}^\pi$;

$z_{i(.)}^{1j}, \dots, z_{i(.)}^{nj}, \dots, z_{i(.)}^{Nj}$ – набор значений атрибутов-параметров $o_{\delta(.)}^{\omega_1}, \dots, o_{\delta(.)}^{\omega_n}, \dots, o_{\delta(.)}^{\omega_N}$, описывающий текущее состояние экземпляра сущности. Индекс $i(nj)$ следует интерпретировать как «текущее значение реквизита-основания $o_{\delta(.)}^{\omega_n}$ для экземпляра $T_{j(.)}^\pi$ ».

Совокупность однородных объектов, имеющих одинаковый реквизитный состав и идентифицируемых зна-

чениями одного и того же реквизита-признака, в реляционной модели представляется как единый объект и называется сущностью. Обозначим сущность символом E_π , где индекс π – наименование реквизита-признака $k_{\delta(.)}^\pi$, значения z_j которого идентифицируют объекты $T_{j(.)}^\pi$ ($j = 1, \dots, J$), рассматриваемые в качестве экземпляров сущности E_π . Сущность E_π экземпляры которого идентифицируются множеством значений одного реквизита-признака, можно представить системой выражений:

$$E_{\pi} = \left\{ T_{j(\cdot)}^{\pi} \left[k_{\delta(\cdot)}^{\pi}(z_j) \left[o_{\delta(\cdot)}^{\omega_1}(z_i), \dots, o_{\delta(\cdot)}^{\omega_n}(z_i), \dots, o_{\delta(\cdot)}^{\omega_N}(z_i) \right] \right] \right\}, j = 1, \dots, J.$$

С одной стороны, полученная совокупность, представленная объектами с одинаковым реквизитным составом, рассматривается как единая сущность или класс однородных объектов. С другой стороны, экземпляры сущности – это отдельные объекты, описываемые значениями реквизитов. Поэтому выражение для описания сущности целесообразно разбить на:

- инвариантную часть (общую для всех экземпляров), выведенную из приведенной системы выражений путем «выведения за скобки» наименований реквизита-признака $k_{\delta(\cdot)}^{\pi}$ и реквизитов-оснований $o_{\delta(\cdot)}^{\omega_1}, \dots, o_{\delta(\cdot)}^{\omega_n}, \dots, o_{\delta(\cdot)}^{\omega_N}$ одинаковых для всех экземпляров и размещаемых в условном заголовке;

- параметрическую часть (переменную и индивидуальную для каждого экземпляра), представленную «оставшимися в скобках» значениями реквизита-признака z_j и реквизитов-оснований $z_{i(\cdot)}^{j1}, \dots, z_{i(\cdot)}^{jn}, \dots, z_{i(\cdot)}^{jN}$, описывающих текущие состояния экземпляров сущности, идентифицируемых значениями z_j реквизита-признака $k_{\delta(\cdot)}^{\pi}$. Обратите внимание на то, что в параметрической части значения представлены отдельно от соответствующих реквизитов, поэтому в выражении использованы их полные обозначения.

В развернутом виде инвариантно-параметрическое описание сущности можно представить следующим выражением:

$$\begin{array}{cccccc} E_{\pi} & \rightarrow & k_{\delta(\cdot)}^{\pi} & \left(o_{\delta(\cdot)}^{\omega_1} & \dots & o_{\delta(\cdot)}^{\omega_n} & \dots & o_{\delta(\cdot)}^{\omega_N} \right) \\ \downarrow & & \downarrow & \downarrow & & \downarrow & & \downarrow \\ T_{1(\cdot)}^{\pi} & \rightarrow & z_{1(\cdot)}^{\pi} & \left(z_{i(\cdot)}^{11} & \dots & z_{i(\cdot)}^{1n} & \dots & z_{i(\cdot)}^{1N} \right) \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ T_{j(\cdot)}^{\pi} & \rightarrow & z_{j(\cdot)}^{\pi} & \left(z_{i(\cdot)}^{j1} & \dots & z_{i(\cdot)}^{jn} & \dots & z_{i(\cdot)}^{jN} \right) \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ T_{J(\cdot)}^{\pi} & \rightarrow & z_{J(\cdot)}^{\pi} & \left(z_{i(\cdot)}^{J1} & \dots & z_{i(\cdot)}^{Jn} & \dots & z_{i(\cdot)}^{JN} \right), \end{array}$$

где E_{π} – обозначение сущности как класса однородных объектов, описываемой последовательностью атрибутов $k_{\delta(\cdot)}^{\pi}, o_{\delta(\cdot)}^{\omega_1}, \dots, o_{\delta(\cdot)}^{\omega_n}, \dots, o_{\delta(\cdot)}^{\omega_N}$;

$T_{1(\cdot)}^{\pi}, \dots, T_{j(\cdot)}^{\pi}, \dots, T_{J(\cdot)}^{\pi}$ – объекты-экземпляры сущности, идентифицируемые значениями $z_1, \dots, z_j, \dots, z_J$ атрибута $k_{\delta(\cdot)}^{\pi}$ и представленные матрицей значений $\left| z_{j(\cdot)}^{\pi}, z_{i(\cdot)}^{j1}, \dots, z_{i(\cdot)}^{jn}, \dots, z_{i(\cdot)}^{jN} \right|$ атрибутов $k_{\delta(\cdot)}^{\pi}, o_{\delta(\cdot)}^{\omega_1}, \dots, o_{\delta(\cdot)}^{\omega_n}, \dots, o_{\delta(\cdot)}^{\omega_N}$.

Выполнение процедуры обобщения на всем множестве $\tilde{T}$ приводит к его разбиению на множество сущностей

$$E = \{E_1, \dots, E_{\pi}, \dots, E_{\Pi}\},$$

экземпляры каждой из которых идентифицируются значениями $z_{j(\cdot)}$ соответствующего реквизита-признака $k_{\delta(\cdot)}^{\pi}$ ($\pi = 1, \dots, \Pi$).

Из приведенных выражений следует, что каждый первичный ключ или реквизит-признак $k_{\delta(\cdot)}^{\pi}$ из множества K определяет сущность E_{π} , а его значения z_j идентифицируют экземпляры $T_{j(\cdot)}^{\pi}$ этой сущности. По этой причине мощность множества сущностей E соответствует мощности множества реквизитов-признаков K .

Процедура, в соответствии с которой объекты представляются как экземпляры некоторой сущности, называется обобщением, в ходе выполнения которой совокупность однородных объектов получает общее имя, т. е. имя сущности. Использование понятия показателей при выполнении операции обобщения дает нам то преимущество, что в данном случае формирование описания сущности или определение состава объектов $T_{j(\cdot)}^{\pi}$ сущности E_{π} выполняется по формальному критерию. Таким критерием является общий реквизит-признак $k_{\delta(\cdot)}^{\pi}$, значения z_j ($j = 1, \dots, J$) которого идентифицируют объекты $T_{1(\cdot)}^{\pi}, \dots, T_{j(\cdot)}^{\pi}, \dots, T_{J(\cdot)}^{\pi}$.

Семантика в определении сущности заключается в том, что для некоторой группы однородных реальных ПрО в ИнФМ для соответствующей группы объектов фиксируется смысловая связь «сущность-объект».

В данном варианте выполнения операции обобщения проблема описания ограничений, на основе которых формируется класс однородных объектов, по-прежнему не решена. Но при этом снижается степень субъективизма при отборе экземпляров сущностей, определяемых значениями их реквизитов-признаков, и нерешенной остается только проблема ввода корректных текущих значений реквизитов-оснований.

3.4 Агрегация показателей сущностей – определение составных сущностей

В ПрО представлены не только простые, но и сложные сущности, содержащие так или иначе организованное множество сущностей. Определение сложной сущности предполагает наличие связи между простыми сущностями, составляющими ее части. Например, сущность *Отдел* связана с сущностью *Сотрудник*, содержащей данные о сотрудниках, работающих в соответствующих отделах. В документах такие связи явно не представляются и определяются аналитиком путем анализа его содержания. Допустим, что в документе

Сотрудники отделов сущности *Отдел* и *Сотрудник* представлены в списке, в котором под наименованиями отделов следуют перечни данных об их сотрудниках. Связи между экземплярами этих сущностей легко определяются человеком по относительному расположению их описаний на площади документа.

С наступлением эпохи автоматизированных систем, когда данные документов «перекочевали» в общую машинную память, потребовалось явное представление связей между сущностями, описываемыми в документах, т. к. фактор размещения их описаний в одном документе и места расположения компонентов описания на площади документа перестал существовать. В реляционной модели связь между сущностями, одна из которых называется родительской, а другая – дочерней, определяется путем ввода в дочернюю сущность атрибута вторичного

ключа, определяемого на том же домене, что и атрибут первичного ключа родительской сущности. В приведенном примере сущность *Отдел* является родительской, а сущность *Сотрудник* – дочерней, представляющей собой составной элемент родительской сущности *Отдел*, а в качестве вторичного ключа введем дополнительное свойство или реквизит-основание № *отдела*. Определенный таким образом реквизит-основание имеет тот же смысл, что и атрибут первичного ключа родительской сущности, а в его определении содержатся указатели на родительскую сущность связи и ее реквизит-признак.

Наша задача заключается в определении составной сущности в терминах показателей и реквизитов. Допустим, что родительская и дочерняя сущности, для экземпляров которых определена связь, выявлены и описываются выражениями:

$$E_{\pi(.)}^r = \left\{ T_{j(.)}^{\pi r} \left[k_{\delta(.)}^{\pi r} (z_j) \left[o_{\delta(.)}^{\omega_1 r} (z_{i(.)}^{j1}), \dots, o_{\delta(.)}^{\omega_n r} (z_{i(.)}^{jn}), \dots, o_{\delta(.)}^{\omega_N r} (z_{i(.)}^{jN}) \right] \right] \right\}, j = 1, \dots, J,$$

$$E_{\pi(.)}^d = \left\{ T_{l(.)}^{\pi d} \left[k_{\delta(.)}^{\pi d} (z_l) \left[o_{\delta(.)}^{\omega_1 d} (z_{i(.)}^{l1}), \dots, o_{\delta(.)}^{\omega_n d} (z_{i(.)}^{ln}), \dots, o_{\delta(.)}^{\omega_N d} (z_{i(.)}^{lN}), v_{\delta r \pi}^{\pi d} (z_{i(.)}^{ld}) \right] \right] \right\}, l = 1, \dots, L,$$

где $E_{\pi(.)}^r$ – родительская сущность, идентифицируемая реквизитом-признаком $k_{\delta(.)}^{\pi r}$ а $E_{\pi(.)}^d$ – дочерняя сущность, идентифицируемая реквизитом-признаком $k_{\delta(.)}^{\pi d}$;

$T_{j(.)}^{\pi r}$ – j -й экземпляр родительской сущности, идентифицируемый значением z_j реквизита-признака. Всего родительская сущность содержит J экземпляров;

$v_{\delta r \pi}^{\pi d}$ – реквизит-основание, добавленный в дочернюю сущность в роли атрибута вторичного ключа с наименованием πd и доменом $\delta(\pi(r))$, определенным для реквизита-признака $k_{\delta(.)}^{\pi r}$ сущности $E_{\pi(.)}^r$.

Особенностью определения связи между двумя сущностями является то обстоятельство, что реквизиты показателя связи $q_{\pi d}^{\pi r} \rightarrow E_{\pi(.)}^r (T_{j(.)}^{\pi r}) [k_{\delta(.)}^{\pi r} (z_{j(.)})]$, $E_{\pi(.)}^d (T_{j(.)}^{\pi d}) [v_{\delta r \pi}^{\pi d} (z_{i(.)})]$ размещены в разных сущностях: реквизит-признак представляет собой первичный ключ родительской сущности, а реквизит-основание – вторичный ключ дочерней сущности.

Процедура формирования составной сущности представляет собой операцию агрегации, выполняемой на множестве сущностей. Критерием выполнения операции агрегации на двух сущностях является наличие показателя, реквизиты которого входят в состав этих сущностей, как было описано выше.

Следует отметить, что показатели данного типа не определяются из содержания документов ПрО, а формируются и фиксируются в ИнфМ при составлении

концептуальной схемы ПрО в виде семантической информации. Объясняется это тем, что смысловая связь между сущностями и их составными частями в документах не указывается явно, а определяется человеком исходя из их размещения на соответствующих площадях документа.

ЗАКЛЮЧЕНИЕ

В реляционной модели базовое понятие отношения определяется как агрегаты, построенные на множестве атрибутов в виде подмножества декартова произведения, представляемого множеством кортежей значений атрибутов. При этом не указывается, как определяется множество атрибутов агрегата и каким образом отбирается подмножество декартова произведения. В этих условиях определение объектов, сущностей и составных сущностей при информационном моделировании ПрО в терминах реляционной модели во многом реализуется на основе интуиции и практического опыта разработчиков, что является источником недостатков, связанных со слабой формализацией понятий объект и сущность, составляющих основу реляционной модели. С целью повышения «удельного веса» формализованной составляющей процесса разработки ИнфМ в статье предложено использовать понятие ФЗ, сформулированное на основе определения показателя.

Семантические данные, зафиксированные на этапе анализа ПрО в виде множества показателей, отражают типовые смысловые связи «экземпляр сущности и значение его свойства». Именно из этих данных выводятся

критерии для выполнения операций агрегации и обобщения, используемые при определении объектов, сущностей и составных сущностей.

В качестве критерия, на основе которого формируется объект ПрО, определяется реквизит-признак, представленный одним значением. По этому критерию выполняется операция агрегации, выделяющая соответствующий набор ФЗ из множества показателей. По выделенному набору ФЗ определяется реквизитный состав объекта. *Вывод 1:* показатели или ФЗ с общим реквизитом-признаком и некоторым его значением определяют объект ПрО.

Выполнив операцию агрегации на всем множестве показателей, получаем множество объектов ПрО, на котором определим операцию обобщения. В качестве критерия, на основе которого группа объектов представляется как единая сущность, определяется реквизит-признак со всеми своими значениями. *Вывод 2:* показатели или ФЗ с общим реквизитом-признаком определяют класс однородных объектов, называемый сущностью ПрО.

Выполнив операцию обобщения на всем множестве объектов, получаем множество сущностей ПрО, на котором определим операцию агрегации, компонентами которой являются не реквизиты, как при определении объектов, а сущности. Операция агрегации в данном случае заключается в связывании двух или более сущностей. В качестве критерия для выполнения агрегации сущностей определяется показатель, реквизит-признак которого размещен в родительской сущности, а реквизит-основание – в дочерней сущности связи. Экземпляры составной сущности определяются по условию: значение реквизита-основания, представляющего вторичный ключ экземпляра дочерней сущности, равно значению реквизита-признака экземпляра родительской сущности. *Вывод 3:* показатель, в котором реквизит-основание дочерней сущности определяет ее ФЗ от реквизита-признака родительской сущности, формирует составную сущность ПрО. ФЗ экземпляра дочерней сущности от заданного экземпляра родительской сущности определяется равенством значений реквизита вторичного ключа с заданным значением первичного ключа родительской сущности.

Можно обратить внимание на то, что при конструировании объектов на основе семантической информации, содержащейся во множестве показателей, прослеживается следующая закономерность:

- реквизит-признак и некоторое его значение, определяет реквизитный состав (реквизиты-основания) объекта на множестве показателей;
- реквизит-признак со всеми своими значениями определяет класс однородных объектов, т. е. сущность, на множестве объектов;
- реквизит-признак, представляющий собой первичный ключ одной сущности, определяет на множестве сущностей составную сущность, если в другой сущности содержится реквизит-основание, функционально зависящий от данного реквизита-признака.

СПИСОК ЛИТЕРАТУРЫ

1. Пушников А.Ю. Введение в системы управления базами данных. URL: <https://stud-files.net/preview/7883250> (дата обращения: 24.08.2020).
2. Швецов В. Базы данных. URL: <https://www.intuit.ru/studies/courses/508/364/info> (дата обращения: 24.08.2020).
3. Цикритзис Д., Лоховски Ф. Модели данных : пер. с англ. М. : Финансы и статистика, 1985. 344 с.
4. Пономарева К.В., Кузьмин Л.Г. Информационное обеспечение АСУ : учеб. пособие. М. : Высш. шк., 1991. 222 с.
5. Смирнова Г.Н., Сорокин А.А., Тельнов Ю.Ф. Проектирование экономических информационных систем учебник / под ред. Ю.Ф. Тельнова. М. : Финансы и статистика, 2002. 512 с.
6. Туманов В. Основы проектирования реляционных баз данных. URL: <http://www.in-tuit.ru/studies/courses/1095/191/info> (дата обращения: 24.08.2020).
7. Вихнин А.Г., Сакипов Н.З. Штурм четвертого мегапроекта: кто будет новым Биллом Гейтсом? Системный анализ и выбор стратегии. М. : Издательство «Диалог-МИФИ», 2008. 288 с.
8. Карпова Т. Базы данных: модели, разработка, реализация. СПб. : Питер, 2001. 304 с.

REFERENCES

1. Pushnikov A.Iu. *Vvedenie v sistemy upravleniia bazami dannykh* [Introduction to the Database Control System]. Available at: <https://studfiles.net/preview/7883250> (accessed 24.08.2020).
2. Shvetsov V. *Bazy dannykh* [Databases]. Available at: <https://www.intuit.ru/studies/courses/508/364/info> (accessed 24.08.2020).
3. Tsikritzis D., Lokhovskiy F. *Modeli dannykh*. Per. s angl. [Data Models. Translated from Engl.]. Moscow, Finansy i statistika Publ., 1985. 344 p.
4. Ponomareva K.V., Kuzmin L.G. *Informatsionnoe obespechenie ASU*. Ucheb. posobie [Information Support of Computer-Aided Control Systems. Study Guide]. Moscow, Vyssh. shk. Publ., 1991. 222 p.
5. Smirnova G.N., Sorokin A.A., Telnov Iu.F. *Proektirovanie ekonomicheskikh informatsionnykh sistem*. Uchebnik pod. red. Iu.F. Telnova [Designing Economic Information Systems. Study Guide, eds. Iu.F. Telnov]. Moscow, Finansy i statistika Publ., 2002. 512 p.
6. Tumanov V. *Osnovy proektirovaniia relatsionnykh baz dannykh* [Fundamentals of Relation Database Designing]. Available at: <http://www.intuit.ru/studies/courses/1095/191/info> (accessed 24.08.2020).
7. Vikhnin A.G., Sakipov N.Z. *Shturm chetvertogo megaproekta: Kto budet novym Billom Geitsom? Sistemnyi analiz i vybor strategii* [The Storm of the Fourth Megaproject; Who will be a new Bill Gates? System Analysis and Strategy Selection]. Moscow, Dialog-MIFI Publ., 2008. 288 p.
8. Karpova T. *Bazy dannykh: modeli, razrabotka, realizatsiia* [Databases: Models, Development, Implementation]. St. Petersburg, Piter Publ., 2001. 304 p.