

Полтавцев А.А. Декомпозиция связей высокого порядка в информационном моделировании. // Проблемы информатики в образовании, управлении, экономике и технике: Сб. статей XVI Междунар. научно-техн. конф. – Пенза: ПДЗ, 2016. – С. 32-37.

УДК 004.657

ДЕКОМПОЗИЦИЯ СВЯЗЕЙ ВЫСОКОГО ПОРЯДКА В ИНФОРМАЦИОННОМ МОДЕЛИРОВАНИИ

А.А. Полтавцев

HIGH ORDER RELATIONSHIPS DECOMPOSITION IN INFORMATION MODELLING

A.A. Poltavtsev

Аннотация. Статья посвящена проблеме отображения в логическую реляционно ориентированную модель семантической модели данных, содержащей взаимосвязи высокого порядка (>2) между объектами. Используется семантическая модель данных класса ER или EER. Формулируется понятие адекватности преобразования и исследуются условия возможности его выполнения.

Ключевые слова: информационное моделирование, ER диаграммы, реляционная схема, база данных, отношение.

Abstract. Paper is devoted to a problem of conversion, into logic relational focused model of semantic data model which contains high order (> 2) relationships between objects. The data semantic model of classes ER or EER is used. The concept of transformation adequacy is formulated and opportunity of its performance conditions are investigated.

Keywords: information modeling, ER diagrams, the relational scheme, database, relation.

Концептуальное моделирование данных – основа большинства проектов создания информационных систем. Цель концептуальных моделей двояка: дать семантически правильное концептуальное представление данных, а также обеспечить платформу для разработки логической схемы БД. Следовательно, методология должна иметь семантически оптимальную форму и предоставлять эвристический набор правил, который позволяет преобразовать результат моделирования к требуемой инструментом форме. Одна из наиболее широко используемых методик – ER модель или расширенная EER модель. Однако есть многочисленные конкурирующие модели, каждая из которых имеет свои сильные и слабые стороны. Все они продолжают теоретически и практически развиваться. Способности каждой из структур и нотаций постоянно сравниваются и обсуждаются.

Определенные проблемы в них остаются вокруг использования N-ary отношений, которые более адекватно отражают семантику, но порождают проблемы адекватной реализации их в реляционной базе данных [1, 2].

Эти проблемы происходят как из-за трудности в идентификации N-ary отношений на практике, так и из-за недостатка понимания конструкций, отвечающих требованиям нормализации, на которой основана реляционная модель. В N-ary отношении мы получаем не одну, как в случае бинарной связи, а несколько пар сущностей, и эти пары взаимно пересекаются и наследуются. Тем не менее в соответствии с требованиями реляционной модели требуется в обязательном порядке представить N-ary связь в виде набора бинарных связей и сущностей, при переходе от семантической модели к логической модели базы данных.

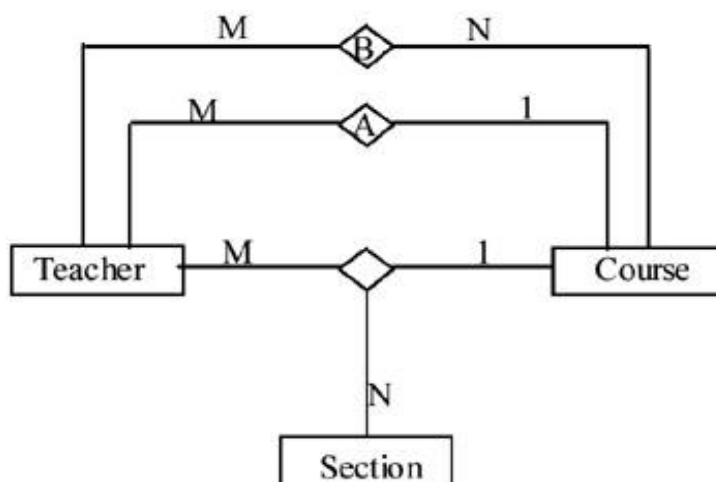
К сожалению, относительно небольшое число исследований было выполнено по обоснованию и использованию N-ary связей [2]. Большинство из них выполнено в контексте исследований процедур нормализации до 4NF и 5NF без привязки к концептуальному моделированию.

Основой полноценного исследования N-ary отношений может служить анализ тернарных отношений. Этот уровень N-ary отношений достаточен, чтобы продемонстрировать проблемы пересечений подмножеств атрибутов и потенциальные противоречия между кардинальностью и декомпозицией отношения, в набор отношений более низкого порядка.

Кардинальность тернарного отношения может быть 1:1:1, 1:1:M, 1:M:N или M:N:P. Ограничение кардинальности сущности в тернарном отношении определяется парой сущностей, связанных с этой (единственной) сущностью. В тернарном отношении $R(X, Y, Z)$ с кардинальностью M:N:1, для каждой пары экземпляров (X, Y) есть только один экземпляр Z; но для каждой пары экземпляров (X, Z) существуют N экземпляров Y; для каждой пары (Y, Z) есть M X.

Согласно [2] будем использовать понятия "семантически связанная бинарная связь (SCB)" и "семантически не связанная бинарная связь (SUB)". Семантически связанная бинарная связь – это связь между двумя сущностями, порождаемая наличием тернарной связи (между тремя сущностями). Соответственно, SUB – бинарная связь между сущностями, существующая независимо от каких-либо тернарных связей, в которых участвуют эти сущности.

Рассмотрите тернарное отношение между Преподавателями, Курсами и Разделами (блоками курсов). Пусть отношение имеет кардинальность M:1:N соответственно и моделирует разделы, связанные с преподавателями курсов, которые в настоящее время их читают. Предположим, что необходимо наложить ограничение, что Преподаватель может преподавать только один курс (что не определено тернарным отношением). Это ограничивает потенциальные комбинации экземпляров (троек) тернарного отношения. Моделирование данной ситуации представлено на рисунке. Бинарная связь (A) между Преподавателем и Курсом имеет тип SCB. В модели существует также независимая, т.е. SUB связь (B) между Преподавателем и Курсом, которая отражает, например, бизнес-правило «курсы, которые преподаватель способен читать».



Такая схема порождает определенные трудности при отображении ее в реляционную модель, основанную на теории нормализации. Как известно, первые три нормальные формы имеют дело с функциональными зависимостями (FD).

Однако, рассматривая функциональные зависимости в пределах тернарных отношений, необходимо понимать, что функциональная зависимость (а она двойная) просто идентифицирует ограничение существования в пределах структуры троичной природы. Минимальный детерминант для тернарных отношений должен быть, по меньшей мере, конкатенацией идентификаторов двух сущностей.

Рассмотрим возможность проведения декомпозиции тернарных отношений в бинарные, которые выполняются без потерь, с сохранением функциональных зависимостей и отсутствием аномалий обновления. В работе Jones и Song (1993) доказано, что, если в пределах тернарной структуры существует по меньшей мере одно бинарное ограничение, тернарное отношение может быть без потерь (losslessly) приведено к структуре из бинарных отношений.

Однако простая декомпозиция без потерь недостаточна, чтобы обеспечить полную эквивалентность на концептуальном и последующих логическом и физическом уровнях.

Декомпозиция любой структуры моделирования предполагает, что получающаяся структура обладает всеми внутренними свойствами оригинальной (первоначальной) структуры. Таким образом, альтернатива, по меньшей мере равная первоначальной структуре, должна быть более приемлема, с точки зрения обработки и извлечения данных.

Рассматривая N-ary отношения с точки зрения сохранения эквивалентности при декомпозиции, можно выделить три области исследования:

- является ли декомпозиция декомпозицией без потерь;
- сохраняет ли декомпозиция функциональные зависимости;
- может ли декомпозиция эквивалентно иметь дело с требованиями выполнения ограничений в физической базе данных, получаемой из модели (то есть требования отсутствия аномалий вставки, удаления, обновления), что в дальнейшем будем объединять одним термином "ограничениями сохранения обновления".

Будем рассматривать только N-ary связи с обязательным участием на всех сторонах.

Рассмотрим сначала случай тернарного отношения 1:1:1.

Пусть каждая сущность имеет по одному экземпляру. Тогда существует единственный $X_1:Y_1:Z_1$. Никакие другие комбинации не возможны, а эта удовлетворяет требованию кардинальности.

Пусть теперь каждая сущность имеет по два экземпляра. В этом случае допустимы только две комбинации: $\{X_1:Y_1:Z_1, X_2:Y_2:Z_2\}$. Дело в том, что т.к. связь 1:1:1, то для пары (X_1, Y_1) есть единственный экземпляр Z_1 ; для пары (X_1, Z_1) есть единственный Y_1 ; для (Y_1, Z_1) единственный экземпляр X_1 , и т.д.

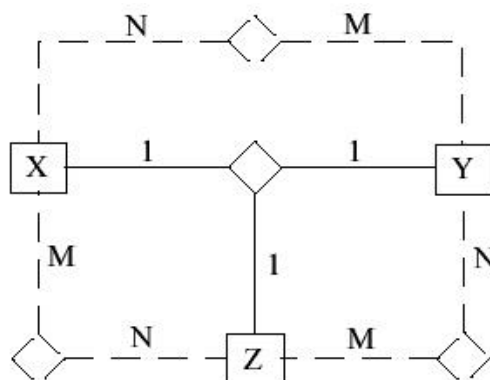
Тернарная связь естественным образом порождает бинарные связи между ее участниками. В данном случае все бинарные связи имеют тип 1:1. На данном множестве допустимых экземпляров $\{X_1:Y_1:Z_1, X_2:Y_2:Z_2\}$ любые другие типы приводят к нарушению кардинальности. Например, создав связь 1:M между парой (X, Y) и Z , придется добавить кортеж $X_1:Y_1:Z_2$, что приводит к нарушению кардинальности 1:1:1, которая становится 1:1:M.

Однако, когда число экземпляров каждой сущности увеличивается хотя бы до трех, ситуация кардинально меняется. Следующий набор экземпляров отвечает требованию 1:1:1: $\{X_1:Y_1:Z_1, X_2:Y_2:Z_2, X_1:Y_2:Z_3\}$.

В данном случае бинарное отношение между X и Y имеет кардинальность M:N. Отношения X:Z и Y:Z – оба 1:M. В то же время тернарное отношение имеет кардинальность 1:1:1. Таким образом, с ростом количества экземпляров все бинарные отношения принимают тип M:N. Например, тернарное 1:1:1 отношение с экземплярами

$$\{X1:Y1:Z1, \\ X2:Y2:Z2, \\ X1:Y2:Z3 \\ X4:Y4:Z3\}$$

определяет все бинарные связи как связи M:N.



Отсюда можно сделать следующий вывод. Так как при проектировании базы данных число возможных экземпляров каждой сущности не учитывается, то все бинарные связи, порождаемые тернарной (N-ary), должны рассматриваться как связи M:N.

Учитывая, что любая тернарная связь, отличная от 1:1:1, является менее ограничивающей, чем 1:1:1, то легко показать, что при отображении тернарной (и любой N-ary) связи в логическую модель все соответствующие бинарные связи должны рассматриваться как M:N.

Таким образом, ограничение кардинальности тернарной (N-ary) связи определяет только комбинации N-к и ничего не говорит о кардинальности соответствующих бинарных связей.

Поэтому в информационном ER моделировании «при отображении N-ary связи в логическую модель каждая бинарная связь, из комплекса бинарных связей, заменяющих N-ary связь, должна определяться как M:N».

Из этого следует, к сожалению, что N-ary связь без дополнительных ограничений не имеет эквивалентной ей декомпозиции в бинарные структуры.

Библиографический список

1. Jones T. and Song I. Y. Analysis of binary/ternary cardinality combinations in entity-relationship modeling // Data and Knowledge Engineering. 1996. №19. P. 39–64.
2. Полтавцева М.А. Хранение сложных структур данных в реляционной базе данных. Тверь: Изд-во ТвГТУ, 2013. 184 с.

Полтавцев Анатолий Алексеевич
Тверской государственный
технический университет,
г. Тверь, Россия
E-mail: aapolt@gmail.ru

Poltavtsev A.A.
Tver State Technical University,
Tver, Russia