

РАСШИРЕННАЯ МОДЕЛЬ „СУЩНОСТЬ-СВЯЗЬ”: ТИПЫ СУЩНОСТЕЙ СУПЕРКЛАСС И ПОДКЛАСС, ТИП СВЯЗИ СУПЕРКЛАСС/ПОДКЛАСС

Дмитрий Буй, Людмила Сильвейструк

Аннотация: Рассматриваются и формализуются такие основные понятия расширенной модели „сущность-связь”: тип сущности суперкласс, тип сущности подкласс, тип связи суперкласс/подкласс, тип связи isa. В терминах теории отношений уточняются ограничения, которые накладываются на тип сущности суперкласс и его типы сущностей подкласс.

Ключевые слова: тип сущности суперкласс, тип сущности подкласс, тип связи суперкласс/подкласс, тип связи isa.

ACM Classification Keywords: E.4 Coding and information theory – Formal models of communication.

Conference: The paper is selected from XIVth International Conference "Knowledge-Dialogue-Solution" KDS 2008, Varna, Bulgaria, June-July 2008

Введение

В работах [Буй, Сильвейструк, 2006; Буй, Сильвейструк, 2007] рассматривались и формализовались основные понятия модели „сущность-связь” (entity-relationship model): типы сущностей, типы связей, ограничения типов связей, роли, сильные и слабые типы сущностей, сильные и слабые типы связей.

В 80-х годах понятий модели „сущность-связь” стало недостаточно для того, чтобы создавать модели данных. Такая ситуация послужила стимулом к разработке дополнительных концепций семантического моделирования. Таким образом, модель „сущность-связь” была пополнена дополнительными концепциями и получила название *расширенной модели „сущность-связь”* (Enhanced Entity-Relationship model) или *EER-модели*.

Расширенная модель „сущность-связь” включает все концепции модели „сущность-связь”, а также собственные концепции. Для расширенной модели „сущность-связь”, как и для модели „сущность-связь”, не существует единого общепринятого стандарта, но имеется набор общих конструкций, которые лежат в основе большинства вариантов (интерпретаций) модели.

В данной работе рассматриваются и уточняются (на основе теории отношений) такие понятия этой модели: тип сущности суперкласс, тип сущности подкласс, тип связи суперкласс/подкласс, тип связи isa, согласно [Гарсиа-Молина, Ульман, Уидом, 2004, гл. 2; Дейт, 1998, часть III, гл. 14; Коннолли, Бегг, Страчан, 2003, часть III, гл. 12; Крэнке, 2003, часть II, гл. 3; Elmasri, Navathe, 2004, часть IV].

Следует отметить, что понятия расширенной модели „сущность-связь” не имеют общепринятой и четкой интерпретации; более того, существует существенный терминологический разнобой; поэтому в табл. 1 приведены варианты, встречающиеся в русскоязычной литературе [Гарсиа-Молина, Ульман, Уидом, 2004; Дейт, 1998; Коннолли, Бегг, Страчан, 2003; Крэнке, 2003].

Такой терминологический разнобой может быть обусловлен неточностями перевода соответствующих источников. Если придерживаться терминологии, которая была принята в работах [Буй, Сильвейструк, 2006; Буй, Сильвейструк, 2007], целесообразно суперкласс называть типом сущности суперкласс, подкласс – типом сущности подкласс, связь суперкласс/подкласс – типом связи суперкласс/подкласс, связь isa – типом связи isa.

Табл. 1 – Ключевые понятия расширенной модели „сущность-связь” и их названия

Понятие	1 вариант	2 вариант	3 вариант	4 вариант
Суперкласс	Супертип	Базовый класс	Суперкласс	Надтип
Подкласс	Подтип	Подкласс	Подкласс	Подтип
Связь суперкласс/подкласс	–	–	Связь суперкласс/подкласс	–
Связь isa	Связь isa	Связь isa	–	Связь типа „Есть”
Источник	[Дейт, 1998, часть III, гл. 14]	[Гарсиа-Молина, Ульман, Уидом, 2004, гл. 2]	[Коннолли, Бегг, Страчан, 2003, часть III, гл. 12]	[Крёнке, 2003, часть II, гл. 3]

Типы сущностей суперкласс и подкласс

Понятия типов сущностей суперкласс и подкласс были добавлены в расширенную модель „сущность-связь” после публикации работы Дж. Смита и Д. Смита [Smith, Smith, 1977]. Причиной стали случаи, когда тип сущности содержит определенные сущности, имеющие специальные свойства, которые не имеют другие сущности данного типа. Поэтому полезно разделять такой тип сущности на несколько специальных типов сущностей, каждый из которых называется *типом сущностей подкласс (subclass entity type)*, а исходный тип сущности называется *типом сущности суперкласс (superclass entity type)*.

Тип сущности суперкласс – тип сущности, который включает одну или несколько разных вспомогательных совокупностей его сущностей, которые должны быть представленный в модели данных.

Тип сущности подкласс – вспомогательная совокупность сущностей некоторого типа сущности (типа сущности суперкласс), которая должна быть представлена в модели данных.

Уточнение одного и того же типа сущности может проводиться на основе разных отличительных особенностей, то есть один тип сущности суперкласс может иметь несколько совокупностей типов сущностей подкласс, которые (совокупности) отображают разные способы группировки сущностей исходного типа сущности суперкласс.

Каждая сущность типа сущности подкласс является сущностью соответствующего типа сущности суперкласс, но сущность типа сущности суперкласс не обязательно является сущностью некоторого типа сущности подкласс. Заметим, что тип сущности подкласс должен содержать, как минимум, одну сущность, иначе нет смысла создавать данный тип сущности подкласс. Один и тот же тип сущности не может быть одновременно типом сущности суперкласс и типом сущности подкласс данного типа сущности суперкласс.

Типы сущностей суперклассы и подклассы будем интерпретировать как множества, а сущности указанных типов – как элементы данных множеств. Пусть множество \mathbf{E} соответствует типу сущности суперкласс, а непустые множества $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_n$ – соответствующим типам сущностей подкласс данного типа

сущности суперкласс, тогда $\bigcup_{i=1}^n \mathbf{E}_i \subseteq \mathbf{E}$.

Далее рассматривается тип сущности суперкласс и соответствующие типы сущностей подкласс, отвечающие некоторому фиксированному способу группировки сущностей суперкласса. При создании типа сущности суперкласс и соответствующих ему типов сущностей подкласс на эти типы сущностей накладываются ограничения. Рассмотрим два основных вида ограничений:

- *ограничение участия (participation constraint)*;
- *ограничение непересечения (disjoint constraint)*.

Ограничение участия определяет, каждая ли сущность типа сущности суперкласс относится к некоторому типу сущности подкласс.

Ограничение участия может быть *обязательным* (*mandatory*) или *необязательным* (*optional*). При обязательном участии каждая сущность типа сущности суперкласс должна быть сущностью некоторого типа сущности подкласс. При необязательном участии некоторая сущность типа сущности суперкласс может не быть сущностью ни одного типа сущности подкласс.

Ограничение непересечения описывает связь между сущностями типов сущностей подкласс (которые связаны с одним типом сущностей суперкласс) и указывает, может ли сущность типа сущности суперкласс принадлежать только одному или нескольким типам сущностей подкласс.

Ограничение непересечения может быть *непересекающим* (*disjoint*) или *пересекающим* (*nondisjoint*). В первом случае каждая сущность типа сущности суперкласс может быть сущностью не более одного типа сущности подкласс. Во втором – сущность типа сущности суперкласс может быть сущностью нескольких типов сущностей подкласс.

Понятно, что данное ограничение можно применять, если тип сущности суперкласс имеет более одного типа сущности подкласс.

Так как указанные два ограничения являются логически независимыми характеристиками образования типов сущностей суперкласс и подкласс, то при их совместимом использовании выделяют четыре следующих разных вида ограничений:

- обязательное (ограничение участия) и непересекающее (ограничение непересечения);
- необязательное и непересекающее;
- обязательное и пересекающее;
- необязательное и пересекающее.

При обязательном и непересекающем ограничении любая сущность типа сущности суперкласс должна принадлежать одному и только одному типу сущности подкласс данного типа сущности суперкласс.

При необязательном и непересекающем ограничении каждая сущность типа сущности суперкласс либо принадлежит единственному типу сущности подкласс данного типа сущности суперкласс, либо не принадлежит ни одному типу сущности подкласс данного типа сущности суперкласс.

При обязательном и пересекающем ограничении любая сущность типа сущности суперкласс должна принадлежать некоторому типу сущности подкласс (необязательно одному) данного типа сущности суперкласс.

При необязательном и пересекающем ограничении каждая сущность типа сущности суперкласс либо принадлежит некоторому типу сущности подкласс (необязательно одному) данного типа сущности суперкласс, либо не принадлежит ни одному типу сущности подкласс данного типа сущности суперкласс.

Рассматривая обязательное ограничение участия, приходим к теоретико-множественным понятиям покрытия и разбиения множества. Очевидно, что обязательное и пересекающее ограничение можно

уточнить с помощью покрытия подмножествами E_1, E_2, \dots, E_n множества E , то есть $E = \bigcup_{i=1}^n E_i$ и

$E_i \neq \emptyset$, где $i = 1, 2, \dots, n$.

Подмножества E_1, E_2, \dots, E_n образуют разбиение множества E , если рассматривать обязательное и

непересекающее ограничение, то есть $E = \bigcup_{i=1}^n E_i$, $E_i \neq \emptyset$ и $E_i \cap E_j = \emptyset$, где $i \neq j$ и

$i, j = 1, 2, \dots, n$.

Если рассматривать необязательное ограничение участия, то необязательное и пересекающее ограничение можно уточнить посредством покрытия подмножествами $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_n$ множества \mathbf{E} или

его собственного подмножества, то есть $\bigcup_{i=1}^n \mathbf{E}_i \subseteq \mathbf{E}$ и $\mathbf{E}_i \neq \emptyset$, где $i = 1, 2, \dots, n$.

Наконец, необязательное и непересекающее ограничение можно уточнить посредством разбиения подмножествами $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_n$ множества \mathbf{E} или его собственного подмножества, то есть $\bigcup_{i=1}^n \mathbf{E}_i \subseteq \mathbf{E}$,

$\mathbf{E}_i \neq \emptyset$ и $\mathbf{E}_i \cap \mathbf{E}_j = \emptyset$, где $i \neq j$ и $i, j = 1, 2, \dots, n$.

В работе [Буй, Сильвейструк, 2006] ограничения кардинальности, которые накладывались на бинарные типы связей, уточнялись посредством свойств соответствующих отношений (функциональность отношения и включение проекции по первой компоненте отношения в множество, на котором отношение задавалось). Поэтому по аналогии уточняются ограничения непересечения и участия.

Рассмотрим множество $\mathbf{E}' = \mathbf{E} \cup \{\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_n\}$ и зададим на нем бинарное отношение \prec :

def
 $e \prec \mathbf{E}_i \Leftrightarrow e \in \mathbf{E}_i$, то есть множество \mathbf{E}_i , $i = 1, 2, \dots, n$ (тип сущности подкласс) содержит элемент e множества \mathbf{E} (тип сущности суперкласс).

Ограничение непересечения можно уточнить посредством функциональности отношения \prec : если отношение \prec функционально, то ограничение непересечения является непересекающим. Если же отношение \prec не функционально, то это ограничение является пересекающим.

Аналогично для ограничения участия: если $\pi_1^2(\prec) = \mathbf{E}$, то ограничение участия обязательное, в случае $\pi_1^2(\prec) \subseteq \mathbf{E}$ ограничения участия необязательное.

В табл. 2 показаны все виды ограничений, которые накладываются на тип сущности суперкласс и его типы сущностей подкласс, а также дано уточнения этих ограничений.

Табл. 2 – Уточнения ограничений, которые накладываются на тип сущности суперкласс и его типы сущностей подкласс

	\prec функциональное	\prec не функциональное
$\pi_1^2(\prec) = \mathbf{E}$	Обязательное и непересекающее ограничение (типы сущностей подкласс образуют разбиение типа сущности суперкласс)	Обязательное и пересекающее ограничение (типы сущностей подкласс образуют покрытие типа сущности суперкласс)
$\pi_1^2(\prec) \subseteq \mathbf{E}$	Необязательное и непересекающее ограничение (типы сущностей подкласс образуют разбиение типа сущности суперкласс или его собственного подмножества сущностей)	Необязательное и пересекающее ограничение (типы сущностей подкласс образуют покрытие типа сущности суперкласс или его собственного подмножества сущностей)

Ограничения, которые накладываются на тип сущности суперкласс и его типы сущностей подкласс, применяются для каждого способа группировки сущностей типа сущности суперкласс. Очевидно, что нет смысла применять данные ограничения к типам сущностей подкласс разных логически независимых способов группировки сущностей типа сущности суперкласс, так как в данном случае ограничения не несут полезную информацию.

Рассматривая тип сущности суперкласс и его типы сущностей подкласс, интересно рассмотреть способы, которыми они могут задаваться. В литературе (смотри, например, [Elmasri, Navathe, 2004, часть IV, с. 103]) предлагается способ *определяющего предиката* (*defining predicate*) и *определяющего атрибута* (*defining*

attribute). То есть для каждого фиксированного способа группировки сущностей типа сущности суперкласс выделяют некоторые определяющие атрибуты, а для каждого типа сущности подкласс некоторый определяющий предикат, заданный на этих определяющих атрибутах (точнее на значениях атрибутов). Для того, чтобы определить принадлежность сущности типа сущности суперкласс соответствующему типу сущности подкласс необходимо проанализировать значение предиката данного типа сущности подкласс: если значение истинно, то принадлежность имеет место, если ложно, то не имеет.

Возникает задача: можно ли задать некоторое отношение на множестве, интерпретирующем тип сущности суперкласс, и по нему получить множества, интерпретирующие типы сущностей подкласс.

Данная задача имеет решение для частного случая – для типа сущности суперкласс и его типов сущностей подкласс с обязательным и непересекающим видом ограничений, которое на них накладывается. Как отмечалось ранее, данный вид ограничений можно уточнить с помощью разбиения. Основываясь на классических теоретико-множественных результатах, отношение эквивалентности позволяет однозначно разбивать множество, на котором задано отношение, на непересекающие классы, содержащие эквивалентные между собой элементы (смотри, например, [Шрейдер, 1971]). Эти классы эквивалентности и интерпретируют типы сущностей подкласс.

При этом следует отметить, что при построении типов сущностей подкласс некоторого типа сущности суперкласс проектант в модели явно указывает количество соответствующих типов сущностей подкласс. В то время как отношение эквивалентности, в общем случае, не позволяет задавать количество необходимых классов эквивалентности (точнее говоря, оценить мощность фактор-множества).

Перейдем от частного случая к более общему, то есть снимем условие транзитивности и обобщим понятие эквивалентности до понятия толерантности [Шрейдер, 1971]. Хорошо известно, что между отношениями толерантности и покрытиями множества существует тесная связь (напомним, что с помощью покрытия выше уточнялся тип сущности суперкласс и его типы сущностей подкласс с обязательным и пересекающим видами ограничений). Однако на этом пути возникает проблема при построении классов толерантности (элементов покрытия): если по покрытию отношение толерантности строится однозначно, то обратная задача построения покрытия по отношению толерантности имеет, вообще говоря, несколько решений [Шрейдер, 1971]. Таким образом, при задании покрытий отношениями толерантности нужны дополнительные средства.

Типы связи суперкласс/подкласс и *isa*

Тип связи суперкласс/подкласс (superclass/subclass relationship type) – это тип связи между типом сущности суперкласс и его типами сущностей подкласс.

Некоторые авторы данный тип связи называют типом связи класс/подкласс (*class/subclass relationship type*) [Elmasri, Navathe, 2004, часть IV, с. 87].

Данный тип связи имеет такие особенности:

- он является $(n + 1)$ -арним типом связи (то есть это тип связи между одним типом сущности суперкласс и его n типами сущностей подкласс), где $n \geq 1$;
- топ простой кардинальности всех типов сущностей подкласс в этом типе связи суперкласс/подкласс обязательное (по терминологии работы [Буй, Сильвейструк, 2006] степень участия каждого типа сущности подкласс в типе связи суперкласс/подкласс полная (mandatory)).

Тип связи isa (isa relationship type) – это тип связи суперкласс/подкласс между типом сущности суперкласс и его типом сущности подкласс.

Данный тип связи имеет такие особенности:

- он является бинарным типом связи „один к одному” (1:1);
- топ простой кардинальности типа сущности подкласс в этом типе связи обязательное.

Следуя работе [Буй, Сильвейструк, 2006], бинарный тип связи уточняется посредством бинарного отношения, причем бинарный тип связи будет вида „один к одному”, когда соответствующее отношение и обратное к нему отношение функциональное. Тип связи isa можно уточнить посредством тождественного отношения, в данном случае вида $\Delta_{E_1} = \{ \langle e, e \rangle \mid e \in E_1 \}$ (диагональ на множестве E_1). Тривиально, что тождественное отношение и обратное к нему (которое совпадает с исходным отношением) являются функциональными, то есть тип связи isa – это действительно бинарный тип связи „один к одному”.

Отметим, что в доступной русскоязычной литературе нет четко очерченных данных типов связей [Гарсиа-Молина, Ульман, Уидом, 2004; Дейт, 1998; Коннолли, Бегг, Страчан, 2003; Крэнке, 2003], причем в [Коннолли, Бегг, Страчан, 2003] употребляется термин „тип связи суперкласс/подкласс” для типа связи isa.

Выводы

Типы сущностей суперкласс и подкласс целесообразно вводить в модель „сущность-связь”:

- во-первых, они позволяют не описывать несколько раз аналогичные сущности, благодаря чему экономится время проектировщика, а диаграммы сущностей и связей становятся более удобными для восприятия;
- во-вторых, они позволяют вводить в модель значительный объем семантической информации в форме, приемлемой для пользователей модели.

Литература

- [Elmasri, Navathe, 2004] Elmasri R., Navathe S. Fundamentals database systems, 4-th edition. – Pearson: Addison-Wesley, 2004. – 1030 p.
- [Smith, Smith, 1977] Smith, J., and Smith, D. Database abstractions: Aggregation and generalization // ACM Transactions on Database Systems. – 1977. – P. 105-133.
- [Буй, Сильвейструк, 2006] Буй Д.Б., Сильвейструк Л.Н. Формализация структурных ограничений в модели „сущность-связь” // Proceedings of the XII-th International Conference „Knowledge-Dialogue-Solution” (June 20-25, 2006, Varna, Bulgaria). – Sofia. – 2006. – P. 223-229.
- [Буй, Сильвейструк, 2007] Буй Д., Сильвейструк Л. Модель „сущность-связь”: роли, сильные и слабые типы сущностей и типы связей // Proceedings of the XIII-th International Conference „Knowledge-Dialogue-Solution” (June 18-24, 2007, Varna, Bulgaria). – Sofia. – 2007. – Vol. 1. – P. 316-322
- [Гарсиа-Молина, Ульман, Уидом, 2004] Гарсиа-Молина Г., Ульман Дж., Уидом Дж. Системы баз данных. Полный курс.: пер. с англ. – Москва: Издательский дом „Вильямс”, 2004. – 1088 с.
- [Дейт, 1998]. Дейт Дж. Введение в системы баз данных.: пер. с англ. – Киев: „Диалектика”, 1998.– 784 с.
- [Коннолли, Бегг, Страчан, 2003] Коннолли Т., Бегг К., Страчан А. Базы данных: проектирование, реализация и сопровождение. Теория и практика, 3-е изд.: пер. с англ. – Москва: Издательский дом „Вильямс”, 2003.– 1440 с.
- [Крэнке, 2003] Крэнке Д. Теория и практика построения баз данных. 8-е изд. Санкт-Петербург: „Питер”, 2003. – 800 с.
- [Шрейдер, 1971] Шрейдер Ю.А. Равенство, сходство, порядок – Москва: Наука, 1971. – 255 с.

Информация об авторах

Буй Дмитрий – заведующий лабораторией проблем программирования, Киевский национальный университет имени Тараса Шевченко, факультет кибернетики: Украина, Киев, 03680, пр. Глушкова 2, корп.6; e-mail: buy@unicyb.kiev.ua

Сильвейструк Людмила – аспирантка, Киевский национальный университет имени Тараса Шевченко, факультет кибернетики: Украина, Киев, 03680, пр. Глушкова 2, корп.6; e-mail: slm-klm@rambler.ru