



Концептуально-онтологические аспекты множественности предка в информационных моделях «сущность-связь»

© 2024, В.В. Миронов ✉, К.В. Миронов

Уфимский университет науки и технологий, Уфа, Россия

Аннотация

Проектирование информационных моделей при разработке автоматизированных систем включает построение моделей «сущность-связь», задающих классы сущностей и межклассовых отношений типа «предок-потомок», которые служат концептуально-онтологической основой для последующего создания баз данных. Рассматриваются аспекты отражения в моделях «сущность-связь» семантических ограничений предметной области, накладываемых на отношения между сущностями. Выделяются особые отношения между классами сущностей, названные отношениями множественного предка (МП), в которых для одного экземпляра сущности-потомка возможно несколько экземпляров сущности-предка. Анализируются возможные семантические ограничения, возникающие в этих условиях, и порождаемые ими аномалии. На основе введенных понятий линии восходящего родства и МП-предиката строится формализация МП-целостности. Приводятся примеры задания формальных МП-ограничений: положительных (требующих совпадения экземпляров предка), отрицательных (требующих несовпадения экземпляров предка), смешанных. Исследуется взаимное влияние нескольких МП-ограничений с частично пересекающимися линиями восходящего родства. Оценивается возможность реализации МП-ограничений в реляционных моделях баз данных.

Ключевые слова: информационная модель; модель «сущность-связь»; классы и экземпляры; онтологии в информатике; концептуальная схема; иерархия «предок-потомок».

Цитирование: Миронов В.В., Миронов К.В. Концептуально-онтологические аспекты множественности предка в информационных моделях «сущность-связь». *Онтология проектирования*. 2024. Т.14, №4(54). С.493-503. DOI:10.18287/2223-9537-2024-14-4-493-503.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Введение

В проектировании автоматизированных систем важное место занимают информационные модели (ИМ), характеризующие используемые данные в создаваемой системе. ИМ разрабатываются на различных уровнях абстракции: концептуальные модели ориентированы на выявление информационных потребностей системы; логические (внутренние) – на реализацию баз данных (БД), обслуживающих эти потребности. ИМ концептуального уровня абстракции отражают онтологические аспекты создаваемой системы и должны позволять разработчикам представлять знания о предметной области (ПрО) в виде концептуальной схемы: классы сущностей, их атрибуты, связи, правила и ограничения [1, 2].

Распространение получил класс ИМ «сущность-связь» (*Entity-Relationship Model – ER-модель*), как концептуальная основа логических моделей БД [3], ориентированных на реляционные и объектно-ориентированные БД. Эти модели положены в основу промышленных стандартов (*IDEFIx*, *UML* и др.) и широко применяются в настоящее время [4–7].

ER-модели наглядно задают классы сущностей и связи между ними, т.е. структурные ограничения целостности данных. Вместе с тем они не отражают некоторые важные особенности ПрО, в отличие от онтологических моделей [7, 8]. Ввиду высокой понятности, лако-

ничности и лёгкой реализуемости *ER*-моделей замена их на онтологическую модель применительно к БД считается нецелесообразной [9]. *ER*-модели являются неотъемлемым средством проектирования БД [10, 11].

Целью данной работы является исследование отношений между классами сущностей и обусловленных ими ограничений целостности в *ER*-моделях при проектировании БД. Указанные отношения связаны с межклассовыми отношениями типа «предок-потомок» [12], названными здесь отношениями множественного предка (МП).

1 Отношения множественного предка

Предметом исследования являются межклассовые отношения типа «предок-потомок», которые непосредственно не отражаются в *ER*-моделях.



Рисунок 1 – Простой пример МП-отношения

Простейший случай подобных отношений может возникать в моделях, содержащих два класса сущностей с двумя связями между ними (см., например, рисунок 1). Здесь класс Организация задаёт множество экземпляров – конкретных организаций, а класс Проект – множество конкретных проектов. Две связи типа «один-ко-многим» – Заказчик (заказывает) и Исполнитель (исполняет) – задают отношения «предок-потомок» между сущностями. Конкретная организация может быть заказчиком и исполнителем нескольких проектов,

но каждый проект имеет одну организацию-заказчика и одну организацию-исполнителя. Пусть при этом имеет место ограничение: одна и та же организация не может быть одновременно заказчиком и исполнителем одного и того же проекта. Данное ограничение никак не отражается традиционными средствами в рассматриваемой модели, т.е. ничто не препятствует, чтобы некоторый экземпляр Проекта ссылался на один и тот же экземпляр Организации и как на заказчика, и как на исполнителя. Налицо «нарушение целостности» – аномалия, потенциально допускающая несогласованность и требующая введения дополнительных семантических пояснений к *ER*-модели. Это обстоятельство отражено на схеме с помощью пунктирной линии, концевой кружок которой задаёт направление от класса-предка к классу-потомку. Это ограничение целостности можно назвать МП-ограничением, которое может возникать в *ER*-моделях, имеющих следующие структурные особенности (МП-отношения):

- имеется два класса сущностей, находящихся в отношении «предок-потомок»;
- имеется две или более восходящих связей, ведущих от потомка к предку.

Опыт проектирования *ER*-моделей БД свидетельствует, что МП-отношения, сопровождаемые МП-ограничениями, часто возникают в ходе построения глобальной *ER*-модели путём композиции локальных моделей.

На рисунке 2 эта ситуация иллюстрируется на простом примере. Здесь локальная модель «а» задаёт отношение «один-ко-многим» между сущностями Спец (специальность) и Студент: на каждой конкретной специальности может обучаться несколько студентов, но каждый конкретный студент обучается на одной специальности. Аналогичным образом локальная модель «б» утверждает: каждая специальность может включать несколько предметов, но каждый предмет относится к одной специальности. Локальная модель «в» задаёт отношение «многие-ко-многим» между сущностями Студент и Предмет с помощью сущности-пересечения Сдача: каждый конкретный студент может сдать несколько конкретных предметов, а каждый конкретный предмет может быть сдан несколькими конкретными студентами. Экземпляр сущности Сдача соответствует паре «конкретный студент – конкретный предмет». Модель «г» на рисунке 2 представляет собой композицию моделей «а», «б» и «в», содержащую все отношения между сущностями Спец, Студент, Предмет, Сдача.

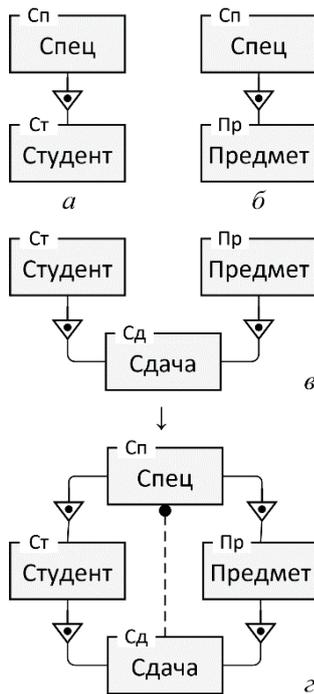


Рисунок 2 – Возникновение МП-отношения в результате композиции моделей

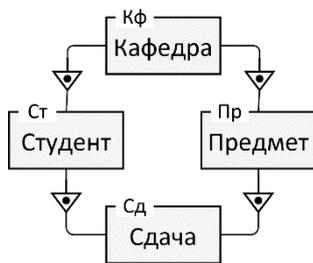


Рисунок 3 – МП-отношение без МП-ограничения

Модель «2» не может учесть те отношения, которые выходят за пределы локальных моделей и относятся к композиции в целом (целое не есть простое объединение частей). Здесь никак не отражено то естественное ограничение, что студент должен сдавать предметы только своей специальности. Если взять конкретный экземпляр сущности Сдача, то для него, согласно модели, существует единственный экземпляр-предок в сущности Студент, для которого существует единственный экземпляр-предок в сущности Спец. Для того же самого экземпляра сущности Сдача существует единственный экземпляр-предок в сущности Предмет, для которого существует единственный экземпляр-предок в сущности Спец. При этом ничто в модели не гарантирует, что в обоих случаях это будет один и тот же экземпляр сущности Спец. Т.е. складывается аномалия, потенциально допускающая несогласованность и требующая введения соответствующих семантических ограничений. На рисунке она отмечена пунктирной линией с концевым кружком. Тёмная заливка кружка (в отличие от светлой на рисунке 1) указывает на то, что экземпляры предка, соответствующие одному потомку, должны совпадать.

Необходимость различать МП-отношение и МП-ограничение обусловлена тем, что не всегда первое влечёт за собой второе. На рисунке 3 приведена иллюстрация МП-отношения, не сопровождающегося МП-аномалией. Здесь класс Кафедра включает множество экземпляров-кафедр, за каждой кафедрой может быть закреплено много студентов и много предметов. При этом студенты изучают и сдают предметы, которые закреплены за разными кафедрами. Таким образом, здесь не накладываются ограничения на экземпляры предков одного потомка, т.е. в отличие от примера на рисунке 2 не требуется, чтобы для одной сдачи кафедра сдавшего студента всегда совпадала (или, наоборот, не совпадала) с кафедрой сданного предмета.

Таким образом, МП-целостность характеризуется наличием МП-отношения и МП-ограничения. МП-ограничение – это

некоторое требование, предъявляемое со стороны ПрО к множеству экземпляров предков каждого экземпляра потомка в МП-отношении. В простых случаях – это требование совпадения или несовпадения экземпляров предков. Возможны и более сложные требования.

2 Формализация

Пусть ER -модель $ERM = \{E, R\}$ – множество сущностей E и бинарных связей R типа «предок-потомок». Другие виды связей («многие-ко-многим»; n -арные, $n > 2$; и др.) могут быть представлены с помощью бинарных связей.

Линия восходящего родства (ЛВР)

$$L = E_D \rightarrow E_1 \rightarrow \dots \rightarrow E_m \rightarrow E_A \quad (1)$$

представляет собой последовательность сущностей $E_D, E_1, \dots, E_m, E_A$, где $m \geq 0$, в которой сущность E_1 является предком сущности E_D , а последующая сущность E_{i+1} является предком

предшествующей сущности E_i (обозначается как $E_i \rightarrow E_{i+1}$). Для $m = 0$ выражение (1) имеет вид $L = E_D \rightarrow E_A$. Сущности E_D и E_A , соответственно сущность-потомок и сущность-предок в ЛВР L , можно записать так:

$$E_D = _L, \quad E_A = \wedge L, \quad (2)$$

где символы « $_$ » и « \wedge » обозначают операции извлечения из ЛВР потомка и предка.

Если сущность E_{i+1} является «множественным» предком сущности E_i (через несколько параллельных связей типа «предок-потомок»), то в записи $E_i \rightarrow E_{i+1}$ требуется уточнение:

$$E_i \xrightarrow{R_i} E_{i+1}, \quad (3)$$

где R_i – идентификатор используемой в ЛВР связи «предок-потомок».

ЛВР как функция экземпляров потомка. Если все связи, образующие ЛВР L , не допускают «сирот», то каждому экземпляру e_D сущности-потомка E_D всегда соответствует один и только один экземпляр e_A сущности-предка E_A , т.е. имеет место функциональная зависимость

$$\wedge L (e_D) = e_A, \quad e_D \in E_D, \quad e_A \in E_A. \quad (4)$$

«Сиротабельные» связи. Пусть на ЛВР присутствуют связи, допускающие «сиротство», т.е. такие экземпляры сущностей потомков, у которых нет соответствующих экземпляров сущностей предков. В этом случае выражение (1) будет записано так:

$$L = E_D \rightarrow E_1 \rightarrow \dots \rightarrow E_k \rightarrow E_{k+1} \rightarrow \dots \rightarrow E_m \rightarrow E_A, \quad (5)$$

где $E_k \rightarrow E_{k+1}$ соответствует «сиротабельной» связи, которая находится ближе всего к сущности-потомку E_D . На экземпляре-сироте происходит «обрыв» экземпляра ЛВР. Это означает отсутствие экземпляра e_A сущности-предка E_A . В таких случаях можно дополнить множества экземпляров сущностей E_{k+1}, \dots, E_m, E_A фиктивными null-экземплярами, полагая, что в случае «обрыва» экземпляра ЛВР он продолжается цепочкой null-экземпляров, т.е. в этом случае

$$\wedge L (e_D) = \text{null}. \quad (6)$$

Для учёта подобных случаев вводятся функции-предикаты $\text{ifnull}(e)$ и $\text{ifnotnull}(e)$, у которых результат «истина», если e соответственно является или не является null-экземпляром, и «ложь» в противном случае.

МП-отношение – это совокупность двух или более различных ЛВР L_1, L_2, \dots , имеющих одинаковых потомков и одинаковых предков, т.е. ведущих разными путями от общего потомка к общему предку:

$$M = \{ L_1, L_2, \dots \}, \quad _L_1 = _L_2 = \dots, \quad \wedge L_1 = \wedge L_2 = \dots. \quad (7)$$

МП-ограничение – это ограничение, заданное на МП-отношении M в виде

$$\forall e_D, \quad e_D \in E_D, \quad \text{Pr}(M) = f(\wedge L_1(e_D), \wedge L_2(e_D), \dots), \quad (8)$$

где \forall – квантор всеобщности; $\text{Pr}(M)$ – предикат МП-целостности (МП-предикат), т.е. некоторая логическая функция f , задающая условие, которое должно выполняться для обеспечения целостности. Таким образом, МП-ограничение накладывается на совместные значения экземпляров предка, полученные для одного экземпляра потомка через различные ЛВР.

МП⁺ / МП-ограничение (положительное / отрицательное МП-ограничения) – частные случаи, когда МП-предикат $\text{Pr}(M)$ задан соответственно в виде равенства или в виде неравенства экземпляров предка:

$$\text{Pr}(M)^+ : \forall (L_i \in M, L_j \in M, i \neq j) \wedge L_i = \wedge L_j, \quad (9)$$

$$\text{Pr}(M)^- : \forall (L_i \in M, L_j \in M, i \neq j) \wedge L_i \neq \wedge L_j, \quad (10)$$

т.е. МП⁺-ограничение для каждого экземпляра потомка требует одних и тех же экземпляров предка по всем ЛВР, а МП-ограничение – различных. В реальных ситуациях могут требоваться более сложные логические условия (смешанные МП-ограничения).

Чтобы формализовать МП-целостность, нужно задать: множество ЛВР (7), составляющих МП-отношение; МП-предикат (8), определяющий МП-ограничение для МП-отношения.

3 Примеры

3.1 Положительные МП-ограничения в отсутствие сирот

На рисунке 4 представлен пример задания положительного МП-ограничения, основанного на двух четырёхзвенных ЛВР. Модель отражает результаты сдачи студентами предметов. Каждая сдача соответствует одному студенту, который прикреплен к одной студенческой группе одной специальности. С другой стороны, каждая сдача соответствует одному предмету, который соответствует одному циклу, также одной специальности. Представленные в модели связи «один-ко-многим» не допускают сирот (тёмные кружки в символах связи). Отношение МП-целостности задано штриховой линией, идущей от потомка к предку. Тёмный кружок, указывающий на предка, означает положительную МП-целостность. Таким образом, МП⁺-целостность Сд^Сп обеспечивается по двум ЛВР:

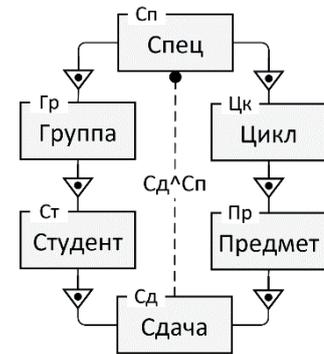


Рисунок 4 – Пример МП⁺-целостности

$$Сд^Сп = \{ L_1 = Сд \rightarrow Ст \rightarrow Gr \rightarrow Sp, L_2 = Сд \rightarrow Пр \rightarrow Цк \rightarrow Sp \}, Pr(Сд^Сп): \wedge L_1 = \wedge L_2, \quad (11)$$

т.е. требуется, чтобы для каждой сдачи совпадали специальности сдавшего студента и сданного предмета. В данном случае задание МП-целостности на графической модели (рисунок 4) исчерпывающе характеризует МП-ограничение: в модели имеется всего две ЛВР, и формальное представление (11) не несёт новой информации по сравнению со схемой. В более сложных случаях графическая нотация бывает недостаточна для задания деталей МП-целостности и требуется дополнительная формальная спецификация.

3.2 Положительные МП-ограничения в случае допустимости сирот

На рисунке 5 представлен пример задания МП⁺-целостности при наличии связи, допускающей сирот (светлый кружок в символе связи). В отличие от модели на рисунке 4 здесь допустимы студенты, которые (временно) не прикреплены к студенческой группе (например, находятся в академическом отпуске). При этом каждый студент относится только к одной специальности (дополнительная связь от сущности Спец к сущности Студент). В этой модели присутствуют два перекрывающихся отношения положительной МП-целостности: Ст^Сп и Сд^Сп. МП-целостность Ст^Сп задаёт согласованность специальности обучения студента и специальности группы, к которой прикреплен студент. Она базируется на двух ЛВР:

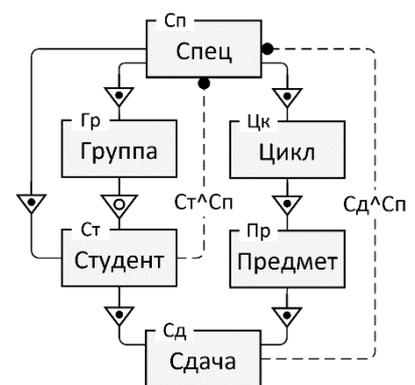


Рисунок 5 – Пример сложной МП⁺-целостности

$$Ст^Сп = \{ L_1 = Ст \rightarrow Sp, L_2 = Ст \rightarrow Gr \rightarrow Sp \}. \quad (12)$$

Предикат Pr(Ст^Сп) должен быть истинным только в двух случаях:

- если студент ст₀ не прикреплен ни к какой к группе;

- если студент $ст_0$ прикреплен к группе, у которой специальность группы $\wedge L_2(ст_0)$ совпадает со специальностью студента $\wedge L_1(ст_0)$.

Таким образом, предикат имеет вид

$$Pr(Ст \wedge Сп): \text{ifnotnull}(\wedge L_{rp}) \Rightarrow (\wedge L_1 = \wedge L_2), \quad (13)$$

где $L_{rp} = Ст \rightarrow \circ Гр$, а « \Rightarrow » обозначает операцию импликации, или в эквивалентной записи:

$$Pr(Ст \wedge Сп): \text{ifnull}(\wedge L_{rp}) \mid (\wedge L_1 = \wedge L_2), \quad (14)$$

где « \mid » обозначает операцию дизъюнкции («или»).

3.3 Пересекающиеся МП-отношения

МП-целостность $Сд \wedge Сп$ (рисунок 5) задаёт согласованность специальностей по различным ЛВР от потомка Сдача к предку Спец. В модели содержится три ЛВР, относящихся к этому МП-отношению:

$$L_3 = Сд \rightarrow \bullet Пр \rightarrow \bullet Цк \rightarrow \bullet Сп, \quad L_4 = Сд \rightarrow \bullet Ст \rightarrow \bullet Сп, \quad L_5 = Сд \rightarrow \bullet Ст \rightarrow \circ Гр \rightarrow \circ Сп. \quad (15)$$

Таким образом, рассматриваемая МП-целостность базируется на трёх ЛВР:

$$Сд \wedge Сп = \{L_3, L_4, L_5\}. \quad (16)$$

МП-предикат $Pr(Сд \wedge Сп)$ должен проверять равенство экземпляров предков, полученных вдоль этих трёх ЛВР, и должен быть истинным в следующих случаях:

- если в сдаче $сд$ сдавший студент не прикреплен к группе и при этом обучается на специальности $\wedge L_4(сд)$, совпадающей со специальностью сданного предмета $\wedge L_3(сд)$;
- если в сдаче $сд$ сдавший студент прикреплен к группе, при этом его специальность $\wedge L_4(сд)$ совпадает со специальностями группы $\wedge L_5(сд)$ и сданного предмета $\wedge L_3(сд)$.

Следует учесть, что ЛВР L_4 включает ЛВР L_1 , ЛВР L_5 включает ЛВР L_2 и при этом соблюдается МП-целостность $Ст \wedge Сп$, основанная на ЛВР L_1 и L_2 . В этих условиях для достижения целостности $Сд \wedge Сп$ достаточно потребовать, чтобы в сдаче $сд$ специальность сдавшего студента совпадала со специальностью сданного предмета, т.е. достаточно контролировать две ЛВР – L_3 и L_4 :

$$Сд \wedge Сп = \{L_3, L_4\}, \quad Pr(Сд \wedge Сп): \wedge L_3 = \wedge L_4. \quad (17)$$

Таким образом, наличие пересекающихся ЛВР в положительных МП-отношениях может упростить МП-предикаты.

В примере на рисунке 5 ЛВР (17), относящиеся к $Сд \wedge Сп$, полностью включают ЛВР (12), относящиеся к $Ст \wedge Сп$. В этих условиях может показаться, что МП-целостность $Сд \wedge Сп$, основанная на (17), автоматически влечёт за собой МП-целостность $Ст \wedge Сп$, что неверно. Это было бы так, если бы у каждого студента обязательно была хотя бы одна сдача (связь Студент–Сдача не допускала бы предков без потомков, т.е. была бы тёмная заливка у треугольника в символе связи). Однако в этой модели допустимы студенты, не имеющие сдач (светлая заливка в треугольнике связи), поэтому такие студенты выпадают из-под контроля со стороны МП-целостности $Сд \wedge Сп$.

3.4 Отрицательные МП-ограничения

Особенности обеспечения отрицательной МП-целостности, требующей различных экземпляров МП-предков для одного МП-потомка по различным ЛВР, представлены примером, показанном на рисунке 6. Модель отражает две роли участия людей (сущность Персона) в диссертации на соискание учёной степени кандидата наук (сущность Диссертация), а именно: автор-диссертант (сущность Автор) и научный руководитель (сущность Научрук). У конкретной диссертации должен быть единственный автор и единственный научный руково-

датель. Конкретная персона может быть автором или научным руководителем. Бинарная связь типа «может быть», или «условная», является частным случаем связи типа «один-многим»: одному экземпляру сущности-предка может соответствовать ноль или один экземпляр сущности-потомка.

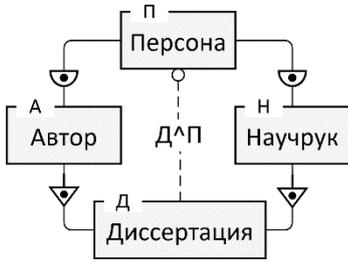


Рисунок 6 – Пример отрицательной МР-целостности

В используемой графической нотации эта связь изображена в виде полукруга, направленного к потомку. МП-целостность $Д^П$ является отрицательной, о чём свидетельствует светлый кружок, указывающий направление к МП-предку. Эта целостность требует, чтобы для каждой диссертации её автор и научный руководитель были разными персонами:

$$Д^П = \{L_1 = Д \rightarrow А \rightarrow П, L_2 = Д \rightarrow Н \rightarrow П\}, Pr(Д^П): \wedge L_1 \neq \wedge L_2. (18)$$

Если предположить, что у диссертации не обязательно должен быть научный руководитель, т.е. связь от Н к Д является «сиротабельной», допускающей сирот ($L_H = Д \rightarrow Н$), то МП-целостность будет выглядеть так:

$$Д^П = \{L_1 = Д \rightarrow А \rightarrow П, L_2 = Д \rightarrow Н \rightarrow П\}, Pr(Д^П): ifnull(\wedge L_H) \mid \wedge L_1 \neq \wedge L_2. (19)$$

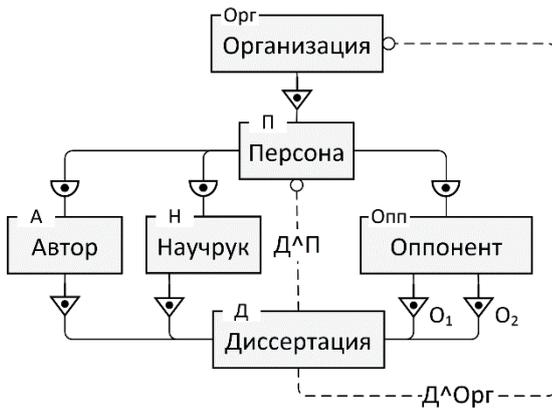


Рисунок 7 – Пример сложной отрицательной МР-целостности

На рисунке 7 представлен более сложный пример, содержащий два взаимосвязанных ограничения отрицательной МП-целостности. В отличие от модели на рисунке 6 здесь дополнительно присутствует Оппонент, как класс Персоны, а также сущность Организация, как предок Персоны. У диссертации должно быть два оппонента, что обеспечивается связями O_1 и O_2 . Каждая Организация может содержать несколько аффилированных Персон, каждая Персона должна быть аффилирована с одной Организацией.

МП-ограничение $Д^П$ в данном случае требует, чтобы для одной и той же диссертации её автор, научный руководитель и оба оппонента были разными персонами:

$$Д^П = \left\{ \begin{array}{l} L_1 = Д \rightarrow А \rightarrow П \\ L_2 = Д \rightarrow Н \rightarrow П \\ L_3 = Д \xrightarrow{O_1} Опп \rightarrow П \\ L_4 = Д \xrightarrow{O_2} Опп \rightarrow П \end{array} \right\}, Pr(Д^П): \forall (L_i, L_j, i, j = \overline{1,4}, i \neq j) \wedge L_i \neq \wedge L_j. (20)$$

Для предиката $Pr(Д^П)$ требуется выполнение следующих условий:

$$Pr(Д^П): \wedge L_1 \neq (\wedge L_2, \wedge L_3, \wedge L_4), \wedge L_2 \neq (\wedge L_3, \wedge L_4), \wedge L_3 \neq \wedge L_4. (21)$$

Для МП-ограничения $Д^Орг$ требуется, чтобы оппоненты диссертации были из разных организаций, не совпадающих с организациями автора и научного руководителя (при этом автор и руководитель могут быть из одной организации). Формально это требование можно записать так:

$$Д^Орг = \left\{ \begin{array}{l} L_5 = Д \rightarrow А \rightarrow П \rightarrow Орг \\ L_6 = Д \rightarrow Н \rightarrow П \rightarrow Орг \\ L_7 = Д \xrightarrow{O_1} Опп \rightarrow П \rightarrow Орг \\ L_8 = Д \xrightarrow{O_2} Опп \rightarrow П \rightarrow Орг \end{array} \right\}, Pr(Д^Орг): \wedge L_7 \neq \wedge L_8, \forall (L_i, L_j)_{\substack{i=(5,6) \\ j=(7,8)}} \wedge L_i \neq \wedge L_j. (22)$$

Для предиката $Pr(D \wedge Org)$ требуется выполнение следующих условий:

$$Pr(D \wedge Org): \wedge L_7 \neq \wedge L_8, \wedge L_5 \neq (\wedge L_7, \wedge L_8), \wedge L_6 \neq (\wedge L_7, \wedge L_8). \quad (23)$$

Следует учесть взаимодействие ограничений $D \wedge П$ и $D \wedge Org$, а именно, что ЛВР, входящие в $D \wedge П$, являются частью ЛВР $D \wedge Org$. Любые экземпляры персоны заведомо различны, если они аффилированы с разными экземплярами организации. Поэтому выполнение требования различной аффилиации оппонентов $\wedge L_7 \neq \wedge L_8$ в (23) автоматически влечёт за собой выполнение требования различных персон оппонентов $\wedge L_3 \neq \wedge L_4$ в (21). Аналогичным образом выполнение требований $\wedge L_5 \neq (\wedge L_7, \wedge L_8)$ и $\wedge L_6 \neq (\wedge L_7, \wedge L_8)$ автоматически влечёт за собой выполнение требований $\wedge L_1 \neq (\wedge L_3, \wedge L_4)$ и $\wedge L_2 \neq (\wedge L_3, \wedge L_4)$.

Эти обстоятельства упрощают предикат $Pr(D \wedge П)$:

$$Pr(D \wedge П): \wedge L_1 \neq \wedge L_2, \quad (24)$$

т.е. при наличии предиката $Pr(D \wedge Org)$ предикат $Pr(D \wedge П)$ должен контролировать лишь то, что экземпляры автора и научного руководителя одной диссертации соответствуют различным экземплярам персоны.

Таким образом, наличие пересекающихся ЛВР в МП-отношениях может упростить ограничения. Например, если дополнительно предположить, что автор и руководитель одной диссертации обязательно должны быть из разных организаций, то ограничения $D \wedge Org$ автоматически обеспечат все ограничения $D \wedge П$ и надобность в $D \wedge П$ отпала бы вовсе.

3.5 Расщепление МП-ограничений

Если некоторая МП-целостность M имеет предикат в виде конъюнкции двух или более логических выражений $Pr(M) = (P_1, P_2, \dots)$, то её можно представить в виде эквивалентной совокупности двух или более МП-целостностей, заданных на том же МП-отношении.

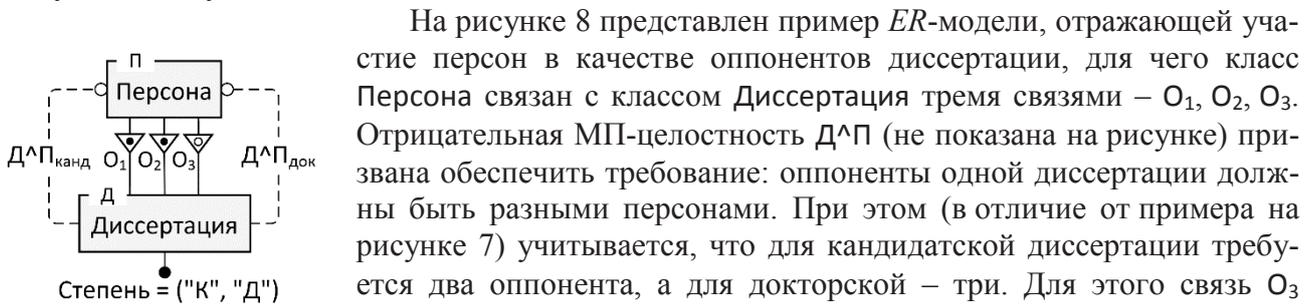


Рисунок 8 – Пример расщеплённого МП-ограничения

На рисунке 8 представлен пример *ER*-модели, отражающей участие персон в качестве оппонентов диссертации, для чего класс Персона связан с классом Диссертация тремя связями – O_1, O_2, O_3 . Отрицательная МП-целостность $D \wedge П$ (не показана на рисунке) призвана обеспечить требование: оппоненты одной диссертации должны быть разными персонами. При этом (в отличие от примера на рисунке 7) учитывается, что для кандидатской диссертации требуется два оппонента, а для докторской – три. Для этого связь O_3 «третий оппонент» допускает сирот (светлый кружочек). Вид конкретной диссертации («кандидатская» или «докторская») задаётся значением атрибута Степень.

МП-ограничение $D \wedge П$ расщеплено на два МП-ограничения для случаев кандидатской и докторской диссертаций соответственно:

$D \wedge П_{канд}$ контролирует кандидатские диссертации, проверяя несовпадение первого и второго оппонентов и отсутствие третьего:

$$D \wedge П_{канд} = \{L_1 = D \xrightarrow{O_1} П, L_2 = D \xrightarrow{O_2} П, L_3 = D \xrightarrow{O_3} П \}, \quad (24)$$

$$Pr(D \wedge П_{канд}): (_L_1 \cdot Степень = "К") \Rightarrow (\wedge L_1 \neq \wedge L_2, isnull(\wedge L_3)).$$

$D \wedge П_{док}$ контролирует докторские диссертации, проверяя присутствие третьего и несовпадение первого, второго и третьего оппонентов:

$$D \wedge П_{док} = \{L_1 = D \xrightarrow{O_1} П, L_2 = D \xrightarrow{O_2} П, L_3 = D \xrightarrow{O_3} П \}, \quad (25)$$

$$Pr(D \wedge П_{док}): (_L_1 \cdot Степень = "Д") \Rightarrow (isnotnull(\wedge L_3), \wedge L_1 \neq \wedge L_2, \wedge L_1 \neq \wedge L_3, \wedge L_2 \neq \wedge L_3).$$

Таким образом, расщепление МП-ограничений в некоторых случаях может сделать представление МП-целостности в *ER*-модели более понятным.

Заключение

Введено понятие МР-целостности, возникающее при наличии в *ER*-модели определённых структурных особенностей – МП-отношений, – когда имеется несколько цепочек связей от некоторого класса-потомка к некоторому классу-предку. На МП-отношениях могут задаваться МП-ограничения, отражающие особенности ПрО, которые накладываются на множество экземпляров предков каждого экземпляра потомка. Введена графическая нотация для обозначения МП-целостности на *ER*-диаграмме.

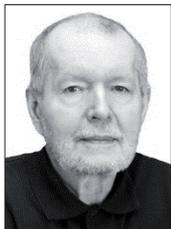
Выполнено формальное описание МП-целостности в символьной нотации на основе введённой системы взаимоувязанных определений: ЛВР; «сиротабельная» связь; МП-отношение; МП-ограничение (в т.ч. частные случаи – положительное и отрицательное МП-ограничения); МП-предикат. Это позволяет задавать формальные концептуально-онтологические спецификации правил обеспечения МП-целостности, дополняющие и уточняющие графическую нотацию.

На примерах *ER*-моделей продемонстрированы особенности задания МП-целостности в графической и в символьной нотациях. Рассмотрены примеры положительных и отрицательных МП-ограничений. Показано взаимное влияние нескольких МП-ограничений с пересекающимися ЛВР, учёт которого позволяет в ряде случаев упростить МП-предикаты. Показано, что расщепление МП-ограничений со сложными МП-предикатами даёт возможность сделать представление МП-целостности более понятным.

Список источников

- [1] *Jardine D.A.* The ANSI/SPARC DBMS Model. North-Holland Pub. Co., 1977. 225 p.
- [2] *Мартин Дж.* Организация баз данных в вычислительных системах. Пер. с англ. под ред. А.Л. Щерца. М.: Мир, 1978. 616 с.
- [3] *Peter Pin-Shan Chen.* The Entity-Relationship Model — toward a unified view of data // ACM Transactions on Database Systems. 1976, vol.1. P.9–36. DOI 10.1145/320434.320440.
- [4] *Storey V.C. et al.* An ontology-based expert system for database design // Data and Knowledge Engineering, vol.28, no.1. P.31-46, 1998. DOI 10.1016/S0169-023X(98)00012-3.
- [5] *Sugumar V., Storey V.* Supporting database designers in entity-relationship modeling: An ontology-based approach. ICIS 2003 Proceedings. 2003. P.59-71. <https://aisel.aisnet.org/icis2003/6>.
- [6] *Chujai P., Kerdprasop N., Kerdprasop K.* On transforming the ER model to ontology using Protégé OWL tool. *International Journal of Computer Theory and Engineering*. 2014. Т.6. №6. С.484.
- [7] *Семенова В.А., Смирнов С.В.* Механизм нормализации эмпирического контекста в онтологическом анализе данных // СИИТ. 2021. Т.3, №3(7). С.45-52. DOI:10.54708/26585014_2021_33745. EDN QXRTXB.
- [8] *Боргест Н.М.* Системный и онтологический анализы: схожесть и различие понятий // *Онтология проектирования*. 2024. Т.14, №1(51). С.9-28. DOI: 10.18287/2223-9537-2024-14-1-9-28. EDN KRGWSR.
- [9] *Perakath C. Benjamin et al.* IDEF5 Method Report. Prepared for Armstrong Laboratory AL/HRGA. Knowledge Based Systems Inc. Sept. 21, 1994. 187 p. <https://online-pmo.com/wp-content/Education/Idef5.pdf>.
- [10] *Антонов А.А., Быков А.Н., Чернышев С.А.* Обзор существующих способов формирования онтологии предметной области при моделировании // *Международный журнал информационных технологий и энергоэффективности*. 2021. Т.6, №4(22). С.12-17. EDN WMBSEO.
- [11] *Богданова Д.Р., Шахмаматова Г.Р., Ниязгулов А.М.* Структура информационного хранилища системы поддержки принятия клинических решений. *Онтология проектирования*. 2024. Т.14, №2(52). С.270-278. DOI 10.18287/2223-9537-2024-14-2-270-278. EDN CAYCVK.
- [12] *Миронов В.В., Гусаренко А.С., Тугузбаев Г.А.* Персонализированные шаблоны конструкторских документов в учебном проектировании: онтологические аспекты и ситуационно-ориентированная реализация. *Онтология проектирования*. 2023. Т.13, №3(49). С.333-351. DOI:10.18287/2223-9537-2023-13-3-333-351.

Сведения об авторах



Миронов Валерий Викторович, 1952 г.р., профессор кафедры автоматизированных систем управления Уфимского университета науки и технологий. Диплом радиофизика (Воронежский университет, 1975). Д-р техн. наук (1995). Труды в области ситуационного управления, иерархических моделей и баз данных. ORCID: 0000-0002-0550-4676. AuthorID (РИНЦ): 691759. Author ID (Scopus): 57192962687. Researcher ID (WoS): AAB-9377-2022. mironov@list.ru ✉

Миронов Константин Валерьевич, 1991 г.р., доцент кафедры вычислительной техники и защиты информации Уфимского университета науки и технологий. Диплом специалиста по защите информации (Уфимский государственный авиационный технический университет, 2012). Степень PhD Технического университета Вены (2015). Труды в области робототехники и искусственного интеллекта, обработки и защиты информации. ORCID: 0000-0002-4828-1345. AuthorID (РИНЦ): 939814. Author ID (Scopus): 56732791500. Researcher ID (WoS): AAB-9377-2022. mironovconst@gmail.com



Поступила в редакцию 30.07.2024, после рецензирования 16.09.2024. Принята к публикации 8.10.2024.



Scientific article

DOI: 10.18287/2223-9537-2024-14-4-493-503

Conceptual and ontological aspects of the plurality of ancestors in "entity-relationship" information models

© 2024, V.V. Mironov ✉, K.V. Mironov

Ufa University of Science and Technologies, Ufa, Russia

Abstract

The design of information models in the development of automated systems involves constructing entity-relationship models, which define classes of entities and interclass relationships of the ancestor-descendant type, serving as a conceptual and ontological foundation for the subsequent creation of databases. This article explores aspects of representing the semantic constraints of the subject area imposed on relationships between entities in entity-relationship models. A special type of relationship between classes of entities, termed multiple ancestor (MA) relations, is highlighted, where several instances of the ancestor entity may exist for a single instance of the descendant entity. The article analyses the possible semantic constraints that arise in such cases and the anomalies they generate. Based on the introduced concepts of ascending kinship lines and the MA predicate, a formalisation of MA integrity is presented. Examples of formal MA constraints are provided, including positive (requiring the coincidence of ancestor instances), negative (requiring the mismatch of ancestor instances), and mixed. The mutual influence of several MA constraints with partially intersecting lines of ascending kinship is also examined. Finally, the feasibility of implementing MA constraints in relational database models is evaluated.

Keywords: *information model; entity-relationship model; classes and instances; ontology in computer science; conceptual schema; ancestor-descendant hierarchy.*

For citation: *Mironov VV, Mironov KV. Conceptual and ontological aspects of the plurality of ancestors in "entity-relationship" information models [In Russian]. *Ontology of designing*. 2024; 14(4): 493-503. DOI:10.18287/2223-9537-2024-14-4-493-503.*

Conflict of interest: The authors declare no conflict of interest.

List of figures

Figure 1 — A simple example of MA relationship

Figure 2 — The emergence of an MA relation as a result of the composition of models

Figure 3 — MA relation without MA constraint

- Figure 4 — Example of a positive MA integrity
 Figure 5 — Example of a complex positive MA integrity
 Figure 6 — Example of a negative MA integrity
 Figure 7 — Example of a complex negative MA integrity
 Figure 8 — Example of a split MA constraint

References

- [1] **Jardine DA.** The ANSI/SPARC DBMS Model. North-Holland Pub. Co., 1977. 225 p.
- [2] **Martin J.** Computer Data-base Organization [In Russian]. Translated from English edited by A.L. Shchers. Moscow, Mir Publ., 1978. 616 p.
- [3] **Peter Pin-Shan Chen.** The Entity-Relationship Model — toward a unified view of data. *ACM Transactions on Database Systems*. 1976; 1: 9–36. DOI 10.1145/320434.320440.
- [4] **Storey VC. et al.** An ontology-based expert system for database design. *Data and Knowledge Engineering*, 1998; 28(1): 31–46. DOI 10.1016/S0169-023X(98)00012-3.
- [5] **Sugumaran V, Storey V.** Supporting database designers in entity-relationship modeling: An ontology-based approach. *ICIS 2003 Proceedings*. 2003. P.59-71. <https://aisel.aisnet.org/icis2003/6>.
- [6] **Chujai P, Kerdprasop N, Kerdprasop K.** On transforming the ER model to ontology using Protégé OWL tool. *International Journal of Computer Theory and Engineering*. 2014; 6(6): 484.
- [7] **Semenova VA, Smirnov SV.** The mechanism of normalization of the empirical context in the ontological analysis of data [In Russian]. *SIIT*. 2021; 3(7): 45-52. DOI 10.54708/26585014_2021_33745. EDN QXRTXB.
- [8] **Borgest NM.** System and ontological analysis: similarities and differences between the concepts [In Russian]. *Ontology of Designing*. 2024; 14(1): 9-28. DOI 10.18287/2223-9537-2024-14-1-9-28. EDN KRGWSR.
- [9] **Perakath C. Benjamin et al.** IDEF5 Method Report. Prepared for Armstrong Laboratory AL/HRGA. Knowledge Based Systems Inc. Sept. 21, 1994. 187 p. <https://online-pmo.com/wp-content/Education/Idef5.pdf>.
- [10] **Antonov AA, Bykov AN, Chernyshev SA.** A review of existing methods for forming the ontology of a subject area in modeling [In Russian]. *International Journal of Information Technology and Energy Efficiency*. 2021; 6(22): 12-17. EDN WMBSEO.
- [11] **Bogdanova DR, Shakhmametova GR, Niiazgulov AM.** The structure of the CDSS information repository based on the ontological approach [In Russian]. *Ontology of Designing*. 2024; 14(2): 270-278. DOI 10.18287/2223-9537-2024-14-2-270-278. EDN CAYCVK.
- [12] **Mironov VV, Gusarenko AS, Tuguzbaev GA.** Personalized engineering documentation templates in instructional design: ontological aspects and situation-based implementation [In Russian]. *Ontology of Designing*. 2023; 13(3): 333-351. DOI 10.18287/2223-9537-2023-13-3-333-351. EDN GEPSYX.

About the authors

Valery Viktorovich Mironov (b. 1952), Professor at the Department of Automated Control Systems of the Ufa University of Science and Technology. Diploma in Radiophysics (Voronezh University, 1975). Doctor of Technical Sciences (1995). Researches are in the field of situational management, hierarchical models and databases. ORCID: 0000-0002-0550-4676. AuthorID (RSCI): 691759. Author ID (Scopus): 57192962687. Researcher ID (WoS): AAB-9377-2022. mironov@list.ru ✉

Konstantin Valerievich Mironov (b. 1991), Associate Professor at the Department of Computer Engineering and Information Security, of the Ufa University of Science and Technology. Diploma in Information Security (Ufa Aviation Technical University, 2012). PhD from the Technical University of Vienna (2015). Researches are in the field of robotics and artificial intelligence, information processing and information security. ORCID: 0000-0002-4828-1345. AuthorID (RSCI): 939814. Author ID (Scopus): 56732791500. Researcher ID (WoS): AAB-9377-2022. mironovconst@gmail.com

Received July 30, 2024. Revised September 16, 2024. Accepted October 8, 2024.