

ИНФОРМАТИКА И ПРОГРАММИРОВАНИЕ

УДК 004.652; 681.3.016

А.М. Бабанов, А.С. Скачкова

ГРАФИЧЕСКАЯ НОТАЦИЯ ДЛЯ МОДЕЛИ «СУЩНОСТЬ – СВЯЗЬ – ОТОБРАЖЕНИЕ»

Статья посвящена описанию графической нотации для семантической модели данных «Сущность – Связь – Отображение». В ней рассматриваются графические примитивы, используемые для представления элементов модели.

Ключевые слова: *семантическая модель, графические примитивы, графическая нотация.*

Любая семантическая модель, будучи ориентированной на человека, не может обойтись без удобного и наглядного графического языка определения схем. Настоящая публикация представляет собой описание новой графической нотации для модели «Сущность – Связь – Отображение». Поскольку эта модель является преемницей модели «Сущность – Связь» (ER-модели), ее графический язык сохраняет многие традиционные конструкции нотаций Чена и Баркера, а также вводит новые, специфические именно для модели «Сущность – Связь – Отображение», графические элементы. Подобный подход позволяет проектировщику в основном использовать знакомые средства и иногда по необходимости вводить новые графические примитивы для более полного отражения семантики предметной области в схеме данных.

Назначение графической нотации

Семантическое моделирование используется в технологии БД при разработке схемы данных. Оно избавляет проектировщика от необходимости работать с моделью данных СУБД на этапе анализа предметной области. Обладая набором близких для человека понятий, семантические модели помогают осуществлять первоначальную формализацию. От того, насколько всесторонне можно описать реальный мир в понятиях модели, зависит полнота схемы данных. Одной из наиболее мощных в этом плане является модель «Сущность – Связь – Отображение» (в дальнейшем ERM-модель от ее англоязычного написания «Entity – Relationship – Mapping Model») [1].

При работе со схемой немаловажным фактором является удобство инструментов проектирования, которые должны быть интуитивно понятными, но не уменьшать при этом выразительные возможности модели. Обычно для этих целей используются графические языки. Удобная графическая нотация и соответствующее средство разработки могут значительно облегчить работу со схемой. Таким образом, для модели «Сущность – Связь – Отображение» необходима графическая нотация, которая и будет описана в данной статье.

Графическая нотация для понятий ER-модели

Так как ERM-модель является развитием ER-модели Чена [1], ее графическая нотация включает в себя графическую нотацию Чена (с некоторыми изменениями, улучшающими ее выразительные способности). Классический вариант нотации Чена [2] представлен на рис. 1.

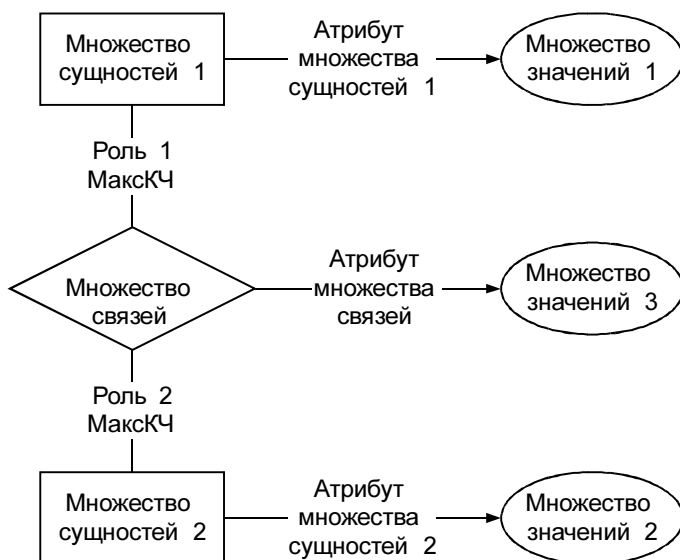


Рис. 1. Графическая нотация Чена

ER-схема в этой нотации представляет собой граф, вершинами которого являются множества сущностей, множества связей, множества значений, обозначаемые поименованными графическими примитивами: прямоугольником, ромбом и овалом соответственно. Роли множеств сущностей во множествах связей обозначаются неориентированными ребрами, соединяющими эти множества. По необходимости они помечаются именами ролей, а также максимальными кардинальными числами (МаксКЧ) соответствующих отображений. Ориентированные дуги представляют атрибуты. Они выходят из вершины множества сущностей или множества связей, входят в вершину множества значений и помечаются именами атрибутов.

Важными средствами определения ограничений целостности в ER-модели являются понятия «зависимость по существованию» (Е-зависимость) и «зависимость по идентификации» (ID-зависимость). Чен предлагает в таком случае прямоугольник, соответствующий зависимому множеству сущностей, очерчивать двойной линией, множеству связей добавлять пометку «Е» или «ID», а стрелкой на ребре роли указывать на зависимое множество сущностей. В нашей графической нотации предлагается изображать зависимость ориентированной дугой, направленной в сторону зависимого множества сущностей с соответствующей пометкой на дуге. На рис. 2 приведены оба способа представления зависимостей (слева – Чена, справа – для ERM-модели). Как видим, новая нотация уменьшает количество обозначений, снимая при этом неоднозначность представления (у Чена в этом случае неясно, куда направлена какая зависимость).

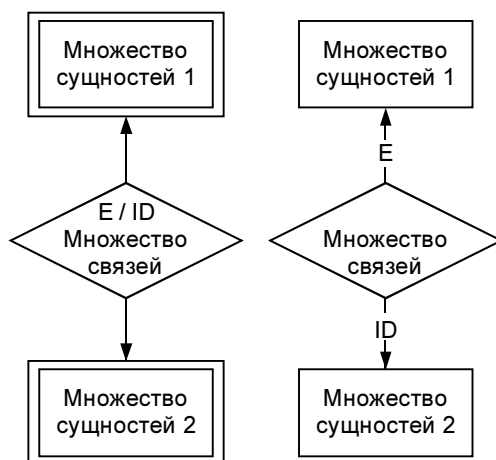


Рис. 2. Представление зависимостей
(слева – у Чена, справа – в ERM-модели)

В нотации Чена явно не определяются агрегатные множества значений, иногда используемые в качестве области значений атрибутивных отображений. В нашей модели они играют важную роль, и поэтому требуется их явное представление на схеме. Сама операция агрегации (фактически являющаяся Декартовым произведением) представляется кружком с символом « $\times$ », к которому подходят ориентированные дуги от множеств-операндов, а из него выходит единственная дуга к множеству-результату (рис. 3).

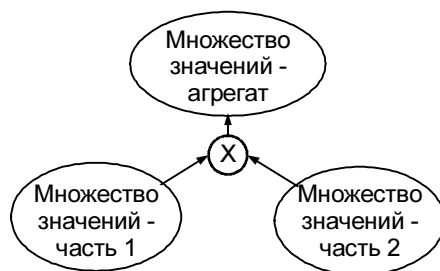


Рис. 3. Представление агрегации множеств значений

Для графического представления ER-схем часто используется нотация Баркера [3]. В нашей графической нотации используется ее сокращенная форма задания атрибутов в виде списка атрибутов внутри графического примитива, обозначающего множество сущностей или множество связей, так как это показано на рис. 4.

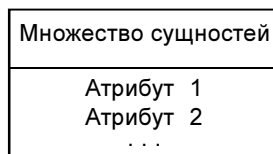


Рис. 4. Представление атрибутов в нотации Баркера

Графическая нотация для понятий EER-модели

Для понятий EER-модели (Enhanced Entity-Relationship Model – расширенная модель «Сущность – Связь») «специализация» и «категоризация» в нашей графической нотации предлагается обозначение, близкое к представлению соответствующих фактов в UML – треугольник, с вершиной которого соединяется суперкласс специализации, а с основанием – подклассы. Кроме того, на диаграмме можно выразить два ограничения целостности, связанных со специализацией: ограничение непересечения и ограничение участия. Ограничение непересечения делит специализации на пересекающиеся и непересекающиеся, ограничение участия – на полные и частичные. Пересекающиеся специализации помечаются двумя перекрещивающимися отрезками внутри треугольника, непересекающиеся – двумя параллельными отрезками. Если специализация является полной, то граница треугольника двойная, иначе – одинарная. Все варианты специализаций показаны на рис. 5.

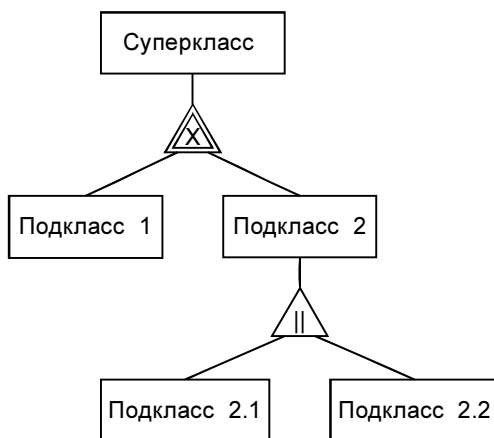


Рис. 5. Представление специализаций

Категоризации тоже обозначаются треугольником. Только подкласс категоризации соединяется с вершиной треугольника, а суперклассы – с его основанием. Треугольник помечается символом « $\cup$ » (знак теоретико-множественной операции объединения). Ограничение участия, которое имеет смысл и для категоризации, обозначается так же, как и для специализации, – двойной или одинарной границей треугольника. На рис. 6 изображена полная категоризация.

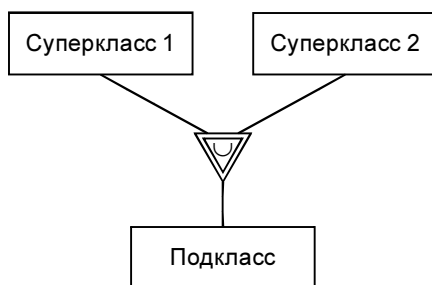


Рис. 6. Представление категоризации

Графическая нотация для понятий ERM-модели

Основное нововведение модели «Сущность – Связь – Отображение» – понятие «отображение». Для его обозначения добавляется новый вид вершин графа – стрелка. Имя отображения пишется внутри стрелки, под ним в скобках указываются количественные характеристики отображения (глобальное минимальное кардинальное число (ГМКЧ), реальное минимальное кардинальное число (РМКЧ), максимальное кардинальное число (МаксКЧ)), причем РМКЧ можно опускать. Множества, формирующие область определения отображения (ООО), соединяются ребрами с началом стрелки, а множества, формирующие область значений отображения (ОЗО), – с концом стрелки. Ребра помечаются именами ролей образов и прообразов, как на рис. 7.

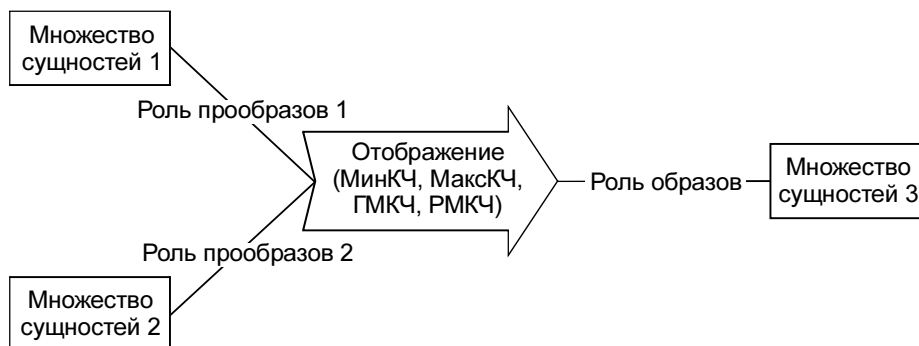


Рис. 7. Представление отображения

Реляционные отображения определяются множествами связей. По-разному распределяя n множеств сущностей, участвующих во множестве связей степени n , по двум полюсам – ООО и ОЗО, можно получить $2^n - 2$ реляционных отображений между множествами сущностей. Чтобы не перегружать схему, на диаграмме обычно показываются только самые важные из них – отображения, определяющие роль, и отображения, определяемые ролью. Их количество равно $2n$. Причем они соединяются, с одной стороны, с множеством сущностей, играющих эту роль, а с другой – с множеством связей. Множество связей в этом случае следует понимать, как Декартово произведение остальных множеств прообразов и образов соответственно. На рис. 8 показаны реляционные отображения, порождаемые множеством связей степени 3. В этом случае все реляционные отображения являются либо отображениями, определяющими роль, либо отображениями, определяемыми ролью. Действительно, $2^n - 2 = 2n = 6$.

Атрибутивные отображения на схеме могут быть представлены, как и другие отображения, – стрелкой, кроме того, как было описано выше, сохраняется возможность изображать атрибутивные отображения в нотациях Баркера и Чена. Таким образом, атрибуты на схеме могут быть показаны тремя способами, которые представлены на рис. 9.

Отношения между отображениями типа «следствие» и «эквивалентность», а также операции, позволяющие образовывать производные отображения, помогают представить важные с точки зрения семантики законы предметной области. Операции на схеме изображаются маленькими кружками с соответствующими знаками. На рис. 10 приведены обозначения операций.

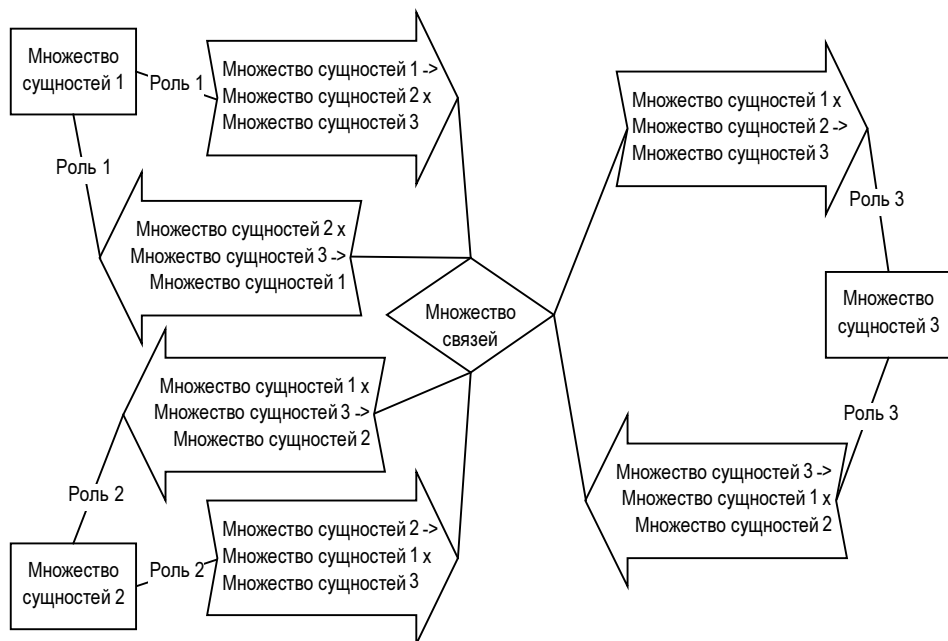


Рис. 8. Представление реляционных отображений

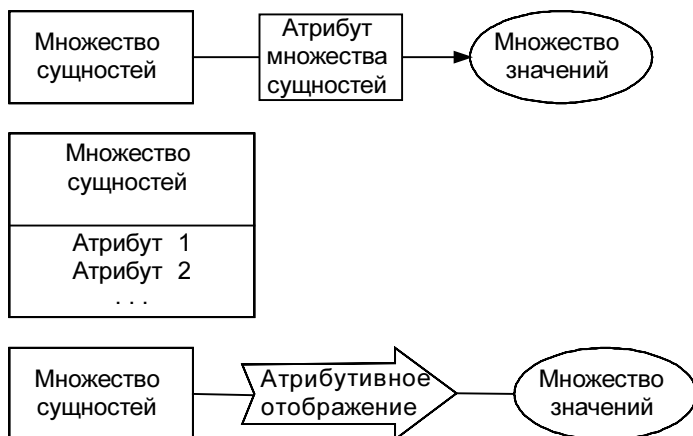


Рис. 9. Варианты представления атрибутов



Рис. 10. Представление операций над отображениями (объединение, пересечение, разность, агрегация, композиция, проекция и инверсия соответственно)

К кружкам, представляющим операции, на схеме подходят ориентированные дуги от отображений-операндов, а из кружка выходит единственная дуга в отображение-результат (см. рис. 11, *а* и 11, *б*). Если операция является некоммутативной, ребра помечаются порядковыми номерами операндов, как на рис. 11, *б*.

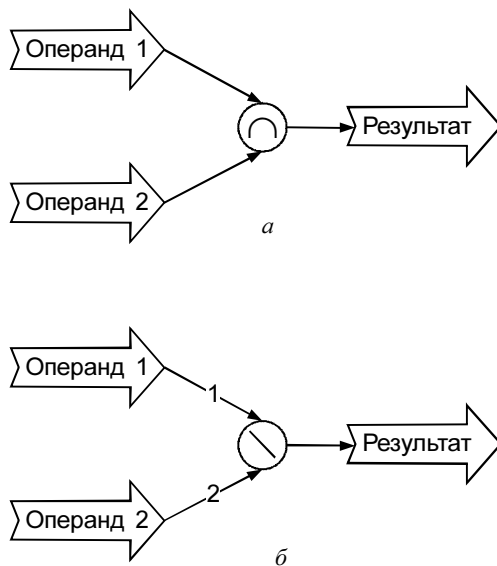


Рис. 11. Представление операций:
а – коммутативной; *б* – некоммутативной

Отношения следствия (рис. 12) и эквивалентности (рис. 13) между отображениями обозначаются однонаправленными и двунаправленными дугами, соединяющими отображения.

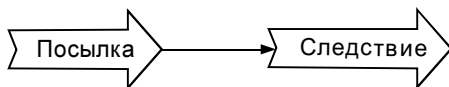


Рис. 12. Представление отношения «следствие»
между отображениями

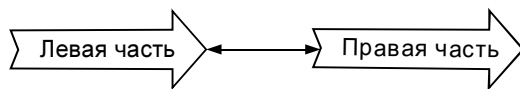


Рис. 13. Представление отношения «эквивалентность»
между отображениями

Введенное в ERM-модели понятие «основание деления специализации» также может быть представлено на схеме. Отображение, являющееся основанием деления для специализации, соединяется с ней пунктирной линией.

Примеры

Для демонстрации возможностей графической нотации приведены примеры схемы предметной области, связанной с ближайшими родственниками. На рис. 14 показана диаграмма множеств сущностей и множеств связей (аналог ER-диаграммы) в обозначениях новой графической нотации.

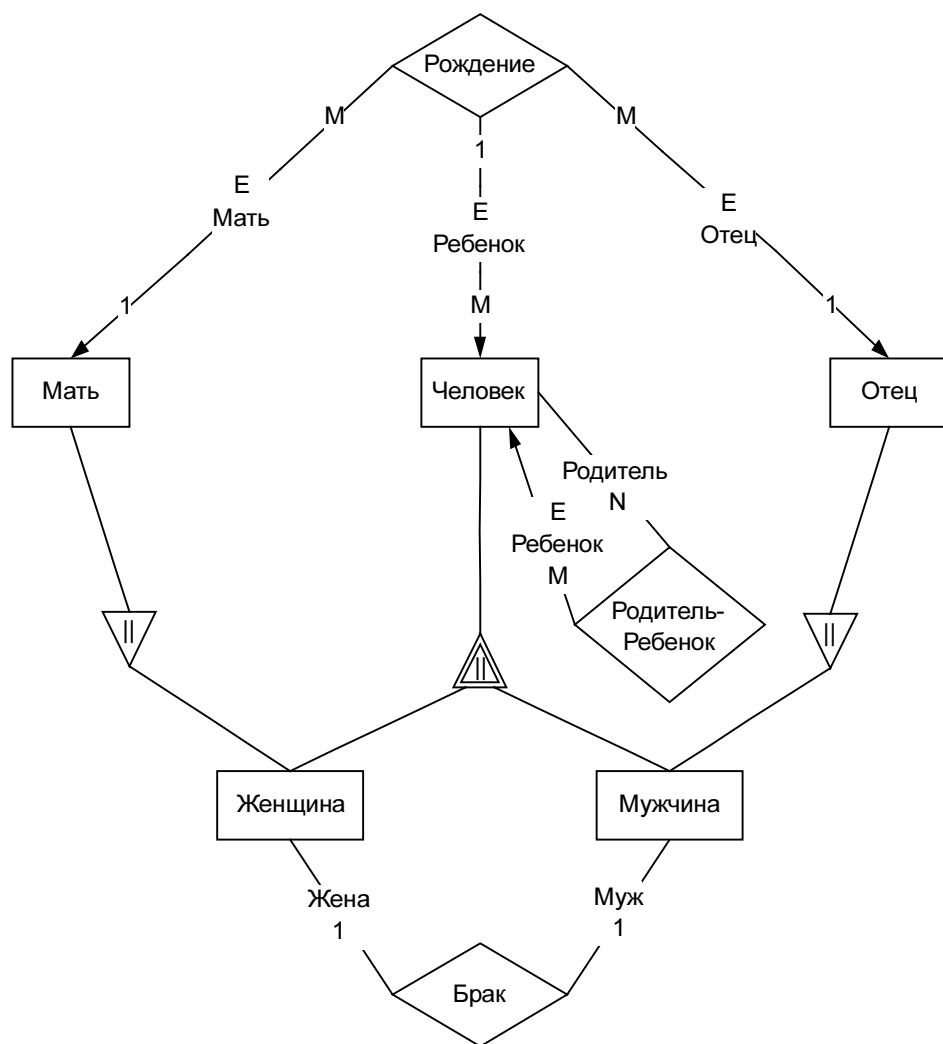


Рис. 14. Диаграмма множеств сущностей и множеств связей

Для уточнения схемы можно создать диаграмму множества связей, на которой будут приведены определяемые им отображения. На рис. 15 показана диаграмма, демонстрирующая все отображения множества связей «Рождение».

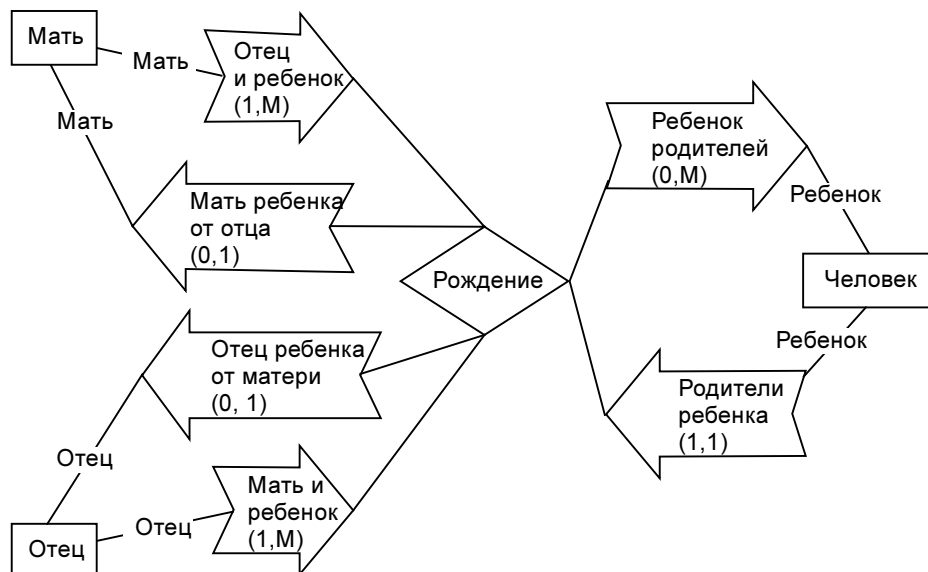


Рис. 15. Диаграмма отображений множества связей «Рождение»

Заключение

Разработанная графическая нотация сочетает в себе понятную и привычную графическую нотацию Чена, удобные элементы нотации Баркера и выразительную мощь конструкций ERM-модели. Благодаря новой графической нотации, проектировщик базы данных может в удобной графической форме строить ERM-схемы произвольной сложности. Он может ограничиться только разработкой ER-схемы, а может отразить закономерности предметной области в более полной форме с использованием аппарата отображений.

В настоящий момент приведенную графическую нотацию уже можно использовать на практике. Для нее создан соответствующий графический редактор (иллюстрации к статье получены с его помощью), который является частью реализуемого на факультете информатики CASE-инструмента. Ведется также разработка средств для автоматической трансформации ERM-схемы в реляционную схему.

ЛИТЕРАТУРА

1. Бабанов А.М. Семантическая модель «Сущность – Связь – Отображение» // Вестник ТГУ. Серия «Управление, вычислительная техника и информатика». 2007. № 1. С. 77 – 91.
2. Chen P.P. The Entity – Relationship Model – Toward a Unified View of Data // ACM Trans. Database Systems. 1976. V. 1. No. 1. P. 9 – 36.
3. Barker R. CASE*Method: Entity-Relationship Modelling. MA.: Addison-Wesley, 1990. 240 p.

Статья представлена кафедрой программной инженерии факультета информатики Томского государственного университета, поступила в научную редакцию 2 июля 2007 г.