

УДК 004.652

МОДЕЛЬ ДАННЫХ «ОБЪЕКТ-СОБЫТИЕ»: ТРЕБОВАНИЯ И СИНТЕЗ МОДЕЛИ

Виталий Есин

Харьковский национальный университет имени В.Н. Каразина, пл. Свободы, 4, г. Харьков, 61022, Украина
y.i.yesin@karazin.ua

Рецензент: Сергей Кавун, доктор экономических наук, к.т.н., проф., Харьковский институт банковского дела УБД НБУ, пр. Победы, 55, г. Харьков, 61174, Украина.
kavserg@gmail.com

Поступила в июне 2017

***Аннотация.** Исходя из необходимости нахождения новых решений актуальной задачи своевременного создания, модернизации в рамках запланированного бюджета баз данных, обладающих требуемыми качествами, формулируются требования, предъявляемые к разрабатываемой модели данных. В соответствии со сформулированными требованиями и общим подходом к проблеме семантического моделирования, предложенным К. Дейтом, синтезируется модель данных, получившая название – «объект-событие».*

Ключевые слова: модель данных, семантическое (концептуальное) моделирование, база данных.

1 Введение

Развитие технологии баз данных (БД) привело к созданию весьма мощных и удобных в эксплуатации информационных систем (ИС), возрастание роли которых в современных условиях является объективной реальностью, обусловленной необходимостью предоставления своевременной и достоверной информации, как одного из наиболее важного ресурса общества, для принятия оптимального решения практически во всех сферах деятельности человека. Анализ ключевых проблем, современного состояния и развития технологий баз данных показал, что основные происходящие сегодня изменения в данной области, обусловленные различными факторами, в том числе, повышенным интересом к организации хранения и обработке больших объемов структурированных и неструктурированных данных, располагающихся во всевозможных источниках, потребностью постоянного совершенствования действующих ИС, путем их адаптации к изменениям внешней среды вне зависимости от конкретной области применения, расширяющимся требованиями потребителей информационных продуктов, вызваны необходимостью решения новых научно-практических задач. Одной из таких актуальных задач является задача адаптации действующих БД информационных систем организационного управления (ИСОУ) к новым условиям функционирования.

Устойчивая тенденция необходимости постоянного совершенствования действующих БД ИСОУ, путем их адаптации к изменениям в предметной области и требованиям бизнес-процессов, ведет к росту востребованности проектов модернизации, интеграции, замены существующих систем. А именно проектов по: разработке новых БД ИСОУ и их интеграции с существующими БД информационных систем; разработке новых БД ИСОУ с целью замены существующих БД информационных систем; модернизации существующих БД ИСОУ, суть которых – реинжиниринг действующих баз данных информационных систем. С реинжинирингом БД ИСОУ в работе связываются такие достаточно широко используемые в различных источниках понятия, как: эволюция баз данных информационных систем, миграция, модернизация, реструктуризация, прямой, обратный инжиниринг и т. д. Деятельность, соотносимая с этими понятиями, подразумевает ее рассмотрение либо как одну из форм, либо как подпроцесс процесса реинжиниринга БД ИСОУ. При этом одним из важных требований, предъявляемых к процессу реинжиниринга существующих БД ИСОУ, является своевременность завершения соответствующих проектов в рамках запланированного бюджета с заданными характеристиками качества, которое, к сожалению, как показывают результаты анали-

за IT-проектов, проведенного международными организациями экспертов, не всегда выполняется. Достаточно большое число проектов (более 60%) были провалены или завершены с опозданием, причем с гораздо большими затратами, чем планировалось [1,2].

Налицо существование нерешенной проблемы, связанной с необходимостью своевременного создания, модернизации в рамках запланированного бюджета информационных систем, обладающих требуемыми качествами, и ограниченностью возможностей существующих методов проектирования. В отношении реляционных баз данных (РБД), как получивших наибольшее распространение в ИС рассматриваемого класса, указанная ограниченность возможностей обусловлена ориентацией традиционной методологии их проектирования, используемой при реинжиниринге БД ИСОУ, на итерационную, достаточно сложную и трудоемкую процедуру создания уникальных концептуальной модели, логической и физической схем при разработке новой БД, либо на существенное их преобразование при модернизации. Что часто влечет за собой значительные, не всегда прогнозируемые объективные затраты временных и финансовых ресурсов. В результате возникает объективная необходимость в пересмотре существующих подходов, методологий и технологий реинжиниринга баз данных. А именно потребность в проведении исследований, направленных на создание таких универсальных моделей и методов обеспечения адаптируемости реляционных баз данных к изменениям в предметной области (ПрО), которые позволят избавиться от необходимости затратной политики выполнения лишних работ при реинжиниринге БД ИСОУ.

Проведенный анализ известных «расширенных» моделей (термин, введенный К. Дейтом [3], для обозначения моделей, используемых при семантическом моделировании): ERM (*entity-relationship model*), EERM (*enhanced entity-relationship model*), HERM (*higher-order entity-relationship model*) [4-6], ORM (*object role modeling*) [7-12], объектной [13-16], семантической бинарной [17,18], семантических сетевых [19,20], инфологической [17, 21-23], онтологической [24-27], а также реляционной модели [28,29], систем доступа к данным, основанным на онтологиях [30], позволил сделать предположение о возможности синтеза таких моделей и методов, если соответствующим образом интегрировать в них подходы и решения, присущие перечисленным выше моделям и системам.

Для доказательства (проверки) справедливости этой гипотезы в первую очередь необходимой стала разработка одной из таких ключевых, востребованных моделей – модели класса «расширенных».

2 Требования, предъявляемые к модели

В технологиях баз данных концептуальное (*семантическое*) моделирование ПрО, как отмечается в работе [31], решает две важные задачи: во-первых, обеспечивает представление ПрО на таком уровне абстракции, благодаря которому оно становится достаточно выразительным для аналитиков, разработчиков, экспертов ПрО, программистов и конечных пользователей; во-вторых, дает возможность проектировщикам БД специфицировать ее структурную и поведенческую организацию в виде, независимом от технических особенностей СУБД.

При этом возможности отображения семантики ПрО в концептуальных моделях в значительной мере связаны со свойствами выразительных средств, используемых для их представления: какие понятия принимаются в качестве предопределенных, что представляет собой атомарный факт, каковы используемые для описания ПрО механизмы абстракций, описания поведения сущностей и т.д. [31]. В случае расхождения языка формализации со складом мышления специалиста, реализация системы обработки данных может стать слишком сложной или вообще неразрешимой проблемой [24]. С другой стороны, решая проблему разностороннего и многоуровневого представления данных, необходимо не только учитывать взгляды пользователей, но и аспекты дальнейшей компьютерной организации и управления данными.

Таким образом, на основании результатов анализа существующих достижений в области

семантического моделирования и перспективных направлений ее развития, исходя из необходимости нахождения решений обозначенной проблемы, связанной с необходимостью своевременного создания, модернизации в рамках запланированного бюджета баз данных, обладающих требуемыми качествами, опираясь на приведенные выше заключения, были сформулированы требования, предъявляемые к разрабатываемой модели:

а) модель должна обладать достаточной общностью с тем, чтобы ее средства позволяли обеспечивать прозрачное для всех участников проекта адекватное, комплексное представление данных моделируемой ПрО, их структуры и ограничений целостности;

б) разрыв между создаваемой моделью и реляционной моделью данных, на основе которой в дальнейшем реализовывается БД ИСОУ, приспособленная к динамичным изменениям предметных областей, не должен быть большим. Более того состав и структура набора формальных объектов создаваемой модели и реляционной модели данных должны быть близкими, чтобы при отображении концептуальной схемы ПрО в даталогическую среду можно было легко воспользоваться заранее известными правилами преобразования структур, ограничений целостности из одной модели в другую и не утратить семантики ПрО;

в) возможность комплексного использования модели, как на соответствующем этапе проектирования БД (*в качестве инструмента концептуального моделирования ПрО*), так и на стадии функционирования РБД ИСОУ, как основы пользовательских интерфейсов.

3 Синтез модели «объект-событие»

Синтез модели, удовлетворяющей сформулированным требованиям, проведем в соответствии с общим подходом, изложенным К. Дейтом [3], и включающим четыре основных этапа:

1) выявление некоторого множества *semantic concepts** (*словосочетание, применяемое автором в оригинальном тексте своей монографии [36]*), которые будут использоваться для описания «реального мира» (* – в переводной монографии [3] словосочетание *semantic concepts* определяется как множество семантических концепций или понятий, хотя корректнее было бы вместо термина концепция использовать термин концепт; поэтому далее по тексту будет использоваться либо непереводаемый оригинал автора, либо термины концепт, понятие, базовое понятие);

2) определение набора формальных объектов, которые могут использоваться для представления описанных понятий;

3) определение формальных правил поддержки целостности данных;

4) определение формальных операторов.

Известно, что информация о реальном мире (рассматриваемой ПрО) дается через восприятие. При этом само восприятие достаточно сложно и состоит из множества взаимосвязанных фактов [17]. Системный подход к познанию ориентирует аналитика на рассмотрение любой ПрО с позиций закономерностей системного целого и взаимодействия составляющих его частей, исходя из многоуровневой иерархической организации любой сущности, когда все объекты, процессы и явления с одной стороны уместно рассматривать как множество более мелких подмножеств (*признаков, деталей*), а с другой – любые объекты разумно рассматривать как элементы более высоких классов обобщений [24]. Если исходить из того, что предлагаемая «расширенная» модель должна обеспечивать возможность правильного представления наших восприятий, то при выборе способа представления элементов ПрО, целесообразно руководствоваться принципами, в соответствии с которыми такое представление человек, в первую очередь, строит для себя на естественном языке. В виду того, что все процессы реального мира протекают в пространстве и времени (в большинстве случаев моделируемая ПрО представляет собой динамическую систему, состоящую из определенной последовательности состояний, конечность множества которых, определяется в каждый момент времени), человек, как правило, отмечает в [17], описывает элементарные факты на естественном языке в терминах объектов, свойств объекта (*связей объектов*), значений свойств и

времени наступления события. Поэтому описание моделируемой ПрО целесообразно осуществлять именно в этих терминах, несколько расширив и дифференцировав их набор, который в итоге будет включать следующие понятия (выделенные курсивом):

- *объект*, связываемый с различными выделенными частями моделируемой ПрО, *свойство (характеристика) объекта* (при этом абстракция обобщения позволяет соотнести множество объектов (подклассов) и их свойств с одним общим *классом*);
 - *событие* (определенного *класса*), происходящее с объектом, *свойство (характеристика) события*, *время наступления события*;
 - *значение характеристики объекта и события*,
- выделяя при необходимости в качестве особых характеристик объектов изменяемые во времени признаки (*параметры*, относящиеся к определенным классам), а также некоторые другие данные (содержательное описание, графическое изображение, аудио-, видеoinформацию).

Эти понятия являются неформальными *semantic concepts* создаваемой модели, как ее неотъемлемый элемент, согласно общему подходу к разработке «расширенных» моделей, предложенному К. Дейтом. Более подробно, с формализацией, эти базовые понятия будут рассмотрены ниже.

Следует заметить, что основополагающими *semantic concepts*, давшими название создаваемой модели, являются базовые понятия «объект» и «событие». Объект как нечто, представляющее интерес, может существовать независимо от того, определены или нет его свойства и связи с другими объектами. Объект возникает, когда субъект начинает проявлять к нему интерес, и исчезает, когда этот интерес утрачивается. При этом единственное свойство, с которым следует соотносить существование объекта – это время его возникновения, исчезновения и изменения, связанное с событиями с ним происходящими [17].

В соответствии с определением модели данных, как совокупности правил описания и структурирования данных, допустимых операций над ними и видов ограничений целостности, которым они должны удовлетворять, модель «объект-событие» (\mathfrak{M}) в формализованном виде можно представить как кортеж:

$$\mathfrak{M} = \langle S(\mathfrak{A}, \mathbb{R}, \mathbb{F}), P, L \rangle, \quad (1)$$

где $S(\mathfrak{A}, \mathbb{R}, \mathbb{F})$ – множество правил описания и структурирования данных ПрО; \mathfrak{A} – множество базовых понятий модели (некоторые из которых приведены в таблице 1); \mathbb{R} – множество отношений между базовыми понятиями модели; \mathbb{F} – множество функций интерпретации, заданных на базовых понятиях (глоссарий, составленный для множества понятий \mathfrak{A}) и отношениях; P – множество ограничений целостности; L – язык модели данных.

Таблица 1 – Перечень базовых понятий модели

<i>Базовое понятие</i>	<i>Определение</i>	<i>Условное обозначение</i>
Раздел	– некоторая выделенная и уникально поименованная часть предметной области.	Раздел
Класс объектов	– совокупность типов объектов, объединяющих экземпляры объектов, выделенные по нескольким значительным качественным признакам, и идентифицируемая именем.	КлассО
Тип объектов	– совокупность схожих по нескольким значительным качественным признакам экземпляров объектов, идентифицируемая именем.	ТипО
Экземпляр объекта (<i>объект</i>)	– однозначно идентифицируемый объект из набора объектов, принадлежащих некоторому типу и классу объектов.	ЭкзО

Продолжение таблицы 1

<i>Базовое понятие</i>	<i>Определение</i>	<i>Условное обозначение</i>
Характеристика типа объектов	– один поименованный признак (качество, свойство) из всей совокупности признаков, описывающих тип объектов определенного класса.	ТипХОп
Фактическая характеристика объекта	– один поименованный признак (качество, свойство) из всей совокупности признаков, описывающих экземпляры объектов определенного класса.	ТипХОф
Значение характеристики объекта	– значение, присвоенное характеристике экземпляра объекта.	ЗначХО
Класс событий	– совокупность событий (экземпляров событий), выделенных по некоторым качественным признакам, которые могут происходить с экземплярами объектов определенного класса в некоторый момент или интервал времени, и идентифицируемая именем.	КлассС
Событие (экземпляр события)	– факт или действие, которое происходит (произошло, будет происходить) с некоторым объектом в определенный момент или интервал времени. Идентифицируется временем и объектом, принадлежит некоторому классу событий. С одним экземпляром объекта в один и тот же момент (интервал) времени может происходить только одно событие одного класса (при допустимости нескольких событий разных классов).	ЭкзС
Характеристика события	– один поименованный признак (качество, свойство) из всей совокупности признаков, описывающих событие определенного класса.	ТипХС
Значение характеристики события	– значение, присвоенное характеристике экземпляра события, которое произошло с конкретным экземпляром объекта.	ЗначХС
Класс параметров объектов	– совокупность характеристик параметров объектов, выделенных по некоторым качественным признакам, идентифицируемая именем.	КлассПО
Характеристика параметра объекта	– изменяемый во времени один поименованный признак (качество) из всей совокупности признаков, описывающих экземпляры объектов определенного класса.	ТипХПО
Единица физической величины	– символическое обозначение единиц физической величины.	ЕдИзмер
Документ	– структурированные или неструктурированные данные, необходимые для дополнения, детализации описания существенных свойств (признаков, качеств), связываемых с основными базовыми понятиями модели.	Документ
Папка документов	– поименованная совокупность документов, выделенных по каким-либо признакам.	Папка

Каждое из базовых понятий (далее в работе часто вместо их полных названий используются условные обозначения) множества \mathcal{A} :

$$\mathcal{A} = \{\text{Раздел, КлассО, ТипО, ЭкзО, ТипХОф, ТипХОп, ЗначХО, КлассПО, ТипХПО, ТипЗнПО, ЗначХПО, КлассС, ЭкзС, ТипХС, ЗначХС, Документ, Папка, ЕдИзмер}\} \quad (2)$$

соотносится с одноименным множеством (рассматривается как имя некоторого множества), являющимся основным компонентом некоторого формального объекта, требующего определения в соответствии с указанным выше подходом к синтезу модели. Например, базовое понятие «Раздел» соотносится с множеством, имеющим имя *Раздел*, базовое понятие «КлассО» соотносится с множеством, имеющим имя *КлассО*, и т. д.

При описании моделируемой ПрО каждое из этих множеств будет содержать в качестве элементов конкретные разделы, классы объектов, событий, параметров объектов, типы объектов, соответствующие характеристики и т. д., имена (*названия*) которых будут либо заимствованы из словарей терминов рассматриваемой ПрО, либо получены в результате неформального соглашения между разработчиками и пользователями. Тогда множество отношений между базовыми понятиями модели «объект-событие» \mathbb{R} – есть множество отношений между одноименными множествами, соотносимыми с этими понятиями, и их элементами, как математических структур (формальных объектов), позволяющих описывать свойства различных элементов моделируемой ПрО и их взаимосвязи (взаимодействия).

В модели «объект-событие» было определено конечное множество таких отношений. Среди них:

– математические отношения, позволяющие формально описывать, так называемые, классификационные отношения (отношения классификации) [24,32], определяющие тип взаимодействия между элементами предметной области. А именно, отношения, фиксирующие связи между элементами «владельцами» и «подчиненными» соответствующих множеств *Раздел*, *КлассО*, *КлассС*, *КлассПО*, *ЭкзО*, *Папка* (данный тип взаимосвязи между элементами множеств, обобщающий типы классификационных отношений «род-вид», «класс-подкласс», «целое-часть», назван в модели – «владелец-подчиненный»):

$$R \subseteq \text{Раздел} \times \text{Раздел} = \text{Раздел}^2 = \{(P_i, P_j) \mid P_i, P_j \in \text{Раздел}\}, \quad (3)$$

где R – бинарное отношение на множестве *Раздел*; роли упорядоченных элементов кортежей отношения R распределяются следующим образом: P_i – «подчиненный», P_j – «владелец» ($i, j \in Ix$; $Ix = \{1, \dots, |\text{Раздел}|\}$ – некоторое множество индексов; $|\text{Раздел}|$ – мощность множества *Раздел*).

Аналогичные формы записи для остальных отношений этой группы:

$$C \subseteq \text{КлассО} \times \text{КлассО} = \text{КлассО}^2 = \{(Кл_{o_k}, Кл_{o_l}) \mid Кл_{o_k}, Кл_{o_l} \in \text{КлассО}\}; \quad (4)$$

$$O \subseteq \text{ЭкзО} \times \text{ЭкзО} = \text{ЭкзО}^2 = \{(Экз_{o_r}, Экз_{o_s}) \mid Экз_{o_r}, Экз_{o_s} \in \text{ЭкзО}\} \quad (5)$$

и т. д.;

– отношения, позволяющие формально описывать, так называемые признаковые [24,32] отношения, приписывающие различные качественные признаки понятиям, которые используются для обозначения элементов моделируемой ПрО. Ниже приведены некоторые n -арные отношения с типами взаимодействия между определенными множествами, соответствующими одноименным базовым понятиям, и их элементами, названными в модели – «иметь характеристику», «иметь документ», «иметь значение характеристики»:

- «иметь характеристику», например, характеристику события (для класса событий), которая «имеет меру»:

$$X \subseteq \text{ТипХС} \times \text{ЕдИзмер} \times \text{КлассС} = \{(T_{xc}, Ед_{\phi}, Кл_c) \mid T_{xc} \in \text{ТипХС} \wedge Ед_{\phi} \in \text{ЕдИзмер} \wedge Кл_c \in \text{КлассС}\}, \quad (6)$$

- «иметь документ» (для основных базовых понятий модели):

$$D \subseteq \text{Документ} \times \text{Папка} \times \mathcal{A} = \{(d, f, u) \mid d \in \text{Документ} \wedge f \in \text{Папка} \wedge u \in \mathcal{A}\}; \quad (7)$$

- «иметь значение характеристики», например, для фактической характеристики объекта:

$$Z_{H_\Phi} \subseteq \text{ЗначХО} \times \text{ЕдИзмер} \times \text{ТипХОф} \times \text{ЭкзО} = \{(Z_{H_{хоф}}, E_{д_\Phi}, T_{хоф}, Экз_о \mid \text{Зн}_{хоф} \in \text{ЗначХО} \wedge E_{д_\Phi} \in \text{ЕдИзмер} \wedge T_{хоф} \in \text{ТипХОф} \wedge Экз_о \in \text{ЭкзО}\}; \quad (8)$$

– математическое отношение ЭкзС , позволяющее формально описывать факты или действия, которые происходят (*произошли или произойдут*) с некоторыми объектами в определенный момент или интервал времени, и определяющее тип взаимосвязи между соответствующими базовыми понятиями, названный в модели – «иметь событие»:

$$\text{ЭкзС} \subseteq \text{КлассС} \times \text{ВремяНС} \times \text{ВремяКС} \times \text{ЭкзО} = \{(Кл_с, Вр_{нс}, Вр_{кс}, Экз_о) \mid \text{Вр}_{нс} \in \text{ВремяНС} \wedge \text{Вр}_{кс} \in \text{ВремяКС} \wedge Кл_с \in \text{КлассС} \wedge Экз_о \in \text{ЭкзО}\}, \quad (9)$$

где ВремяНС – множество времен начала событий; ВремяКС – множество времен окончания событий (допускается отсутствие – неопределенное значение (*null*) элемента $\text{Вр}_{кс} \in \text{ВремяКС}$).

Множество функций \mathbb{F} . В модели «объект-событие» кроме декларативных функций интерпретации, приведенных в виде глоссария, составленного для множества понятий \mathcal{Q} (см.табл. 1), были определены следующие виды функциональных отношений, как специальный вид отношений:

– отношение, определяющее тип взаимодействия между соответствующими базовыми понятиями модели «образуют тип»:

$$f : \text{ЭкзО} \times \text{ТипХОн} \rightarrow \text{ТипО} \Leftrightarrow \text{ЭкзО} \times \text{ТипХОн} \xrightarrow{f} \text{ТипО}, \quad (10)$$

где для любой упорядоченной пары $(Экз_о, T_{хон})$ из $\text{ЭкзО} \times \text{ТипХОн}$ ($Экз_о \in \text{ЭкзО}$; $T_{хон} \in \text{ТипХОн}$) существует не более одного элемента $T_о \in \text{ТипО}$, такого, что $(Экз_о, T_{хон}, T_о) \in f$, тогда: $T_о = f(Экз_о, T_{хон})$.

– отношение, определяющее тип взаимодействия между соответствующими базовыми понятиями модели *Раздел* и *КлассО*. Это отношение агрегации – «целое-часть» («*Раздел* включает *КлассО*»):

$$\rho : \text{КлассО} \rightarrow \text{Раздел} \Leftrightarrow \text{КлассО} \xrightarrow{\rho} \text{Раздел}. \quad (11)$$

Использование именно отношения «целое-часть» в данном случае обусловлено его примечательной особенностью – такие отношения могут быть установлены между сущностями различных семантических типов: физическими объектами, процессами и действиями, географическими регионами, свойствами и состояниями, коллекциями и множествами, абстрактными сущностями и т.д. [33]. Это соответственно в определенной степени упрощает описание некоторых элементов ПрО;

– отношения, определяющие тип взаимодействия между соответствующими базовыми понятиями модели «образуют класс» (например, класс объектов):

$$\alpha : \Psi \rightarrow \text{КлассО} \Leftrightarrow \Psi \xrightarrow{\alpha} \text{КлассО}, \quad (12)$$

где $\Psi \subseteq \text{ТипО} \times \text{ТипХОф}$ и т. д.

Представленные множества базовых понятий модели «объект-событие» (как множества определенных «полезных» *semantic concepts*), отношений \mathbb{R} и функций \mathbb{F} (как набор формальных объектов), определяют правила описания и структурирования данных ПрО.

Полученный набор формальных объектов, как математических структур, несложно реализовать в рамках реляционной модели, которая в своем большинстве случаев [34], вполне до-

статочна для моделирования различных ПрО. Это позволяет выполнить предъявляемое к модели требование по обеспечению близости (согласованности) состава и структуры ее набора формальных объектов, с базисом отношений реляционной модели данных, на основе которой реализуется БД ИСОУ, адаптированная к изменениям предметных областей.

Множество ограничений целостности Р. Как известно [17], ограничения вводятся в модели данных в целях повышения их семантической и расширения возможностей поддержки целостности данных. В модели, в соответствии с классификацией, приводимой в различных авторитетных источниках [17,35], специфицируются следующие типы ограничений целостности данных: - явные (семантические ограничения целостности); - неявные (внутренние, поддерживаются самой структурой модели данных). Ряд явных ограничений целостности в модели относится к характеристикам объектов, событий, параметров объектов. Прежде всего, это множества допустимых значений для соответствующих характеристик (ограничения на допустимые значения):

$$\begin{aligned} D &= \{D_1, D_2, D_3\}; D_1 = \{D_1^1, \dots, D_1^{K_1}\}; D_2 = \{D_2^1, \dots, D_2^{K_2}\}; D_3 = \{D_3^1, \dots, D_3^{K_3}\}; \\ D_1^i &= \text{dom}(T_{xo}^{i\text{cнuc}}); i = 1 \dots K_1; T_{xo}^{i\text{cнuc}} \in \text{TunXO}^{\text{cнuc}} \subseteq (\text{TunXO}\phi \cup \text{TunXOn}); \\ D_2^j &= \text{dom}(T_{xc}^{j\text{cнuc}}); j = 1 \dots K_2; T_{xc}^{j\text{cнuc}} \in \text{TunXC}^{\text{cнuc}} \subseteq \text{TunXC}; \\ D_3^k &= \text{dom}(T_{xno}^{k\text{cнuc}}); k = 1 \dots K_3; T_{xno}^{k\text{cнuc}} \in \text{TunXΠO}^{\text{cнuc}} \subseteq \text{TunXΠO}, \end{aligned} \quad (13)$$

где $\text{dom}(T_{xo}^{i\text{cнuc}})$, $\text{dom}(T_{xc}^{j\text{cнuc}})$, $\text{dom}(T_{xno}^{k\text{cнuc}})$ – домены соответствующих характеристик объектов ($T_{xo}^{i\text{cнuc}}$), событий ($T_{xc}^{j\text{cнuc}}$), параметров объектов ($T_{xno}^{k\text{cнuc}}$) моделируемой ПрО, принадлежащих к так называемому перечисляемому (списочному) типу (значения для них выбираются из заранее сформированного списка); K_1, K_2, K_3 – число соответствующих характеристик объектов, событий, параметров объектов списочного типа.

Следует отметить, что в достаточно многих ПрО, (автопром, авиапром, военно-промышленный комплекс), необходимо учитывать ограничения, накладываемые на конкретные экземпляры объектов. Например, комплектующие к некоторым изделиям (объектам) должны пройти процедуру соответствующей приемки, сертификации. После чего только они (при соответствующей идентификации), а никакие другие элементы, могут быть установлены в нужное изделие. Человек может понять нечто только в конкретном контексте [17]. Поэтому определение множества допустимых экземпляров объектов, задаваемых в модели в виде домена: $\text{dom}(\text{Экз}_o^{\text{оэп}})$, является также неотъемлемой и востребованной частью описания ПрО ($\text{Экз}O_o \in \text{dom}(\text{Экз}_o^{\text{оэп}}) \subseteq \text{Экз}O$), т.к. впоследствии позволяет уменьшить количество потенциальных ошибок.

Следующим ограничением, выражаемым в модели, является ограничение, накладываемое на используемые единицы физических величин характеристик объектов, событий, параметров объектов рассматриваемой ПрО, в виде задания домена: $\text{dom}(E\delta_\phi)$.

Модель «объект-событие» позволяет представить ограничение по существованию, заключающееся в том, что для существования элемента в отношении S_1 необходимо, чтобы он был связан с элементом в отношении S_2 ($S_1 \rightarrow S_2$: каждый элемент S_1 отображен в один элемент S_2). Например, в модели без конкретного класса объекта, не может быть его типов, характеристик, экземпляров и событий с ними происходящими и т.д. При удалении класса объектов, удаляются все его типы, характеристики, экземпляры, события и все «подчиненные» ему классы объектов, а также их типы, характеристики, экземпляры, события и т.д. Аналогично для классов событий и параметров объектов, разделов, экземпляров объектов и событий.

Как внутренние ограничения следует рассматривать требования возможности выражения принадлежности элементов моделируемой ПрО к множествам отношений модели «объект-

событие», а также к множеству значений с помощью предикатов. При этом сами предикаты могут быть заданы явно, как в случае выражения ограничений:

– на максимальное количество экземпляров объектов для определенного класса объектов:

$$C = \{(Kl_{o_k}, Kl_{o_l}, c_{obj}) \mid Kl_{o_k}, Kl_{o_l} \in \text{Класс}O \wedge c_{obj} \in \mathbb{N}^+\}, \quad (14)$$

где c_{obj} – элемент упорядоченной тройки отношения (14), определенный на множестве натуральных положительных чисел \mathbb{N}^+ , ограничивающий число экземпляров объектов ($|\text{Экз}O| \leq c_{obj}$) для класса объектов Kl_{o_k} ;

– на максимальное количество значений, которые могут быть присвоены определенной характеристике события для экземпляра события заданного класса:

$$X = \{(T_{xc}, Ed_{\phi}, Kl_c, x_{ce}) \mid T_{xc} \in \text{Tun}XC \wedge Ed_{\phi} \in \text{EdИзмер} \wedge Kl_c \in \text{Класс}C \wedge x_{ce} \in \mathbb{N}^+\}, \quad (15)$$

где x_{ce} – элемент упорядоченной четверки отношения (15), определенный на множестве натуральных положительных чисел \mathbb{N}^+ , ограничивающий количество значений, которые могут быть присвоены характеристике события T_{xc} для экземпляра события класса Kl_c и т.д.

При этом выделение в отдельную компоненту базового понятия «событие», напрямую связанного со временем, усиливает контроль за непротиворечивостью данных – ограничением их целостности, обеспечивая сохранение сведений о свойствах или связях, которые либо являются на текущий момент достоверными – состояние ПрО, либо их утратили.

Уникальная идентификация элементов множеств, соотнесенных с одноименными базовыми понятиями модели, достигается уникальностью именования, в том числе: для элементов множеств *Раздел*, *КлассО*, *КлассС*, *КлассПО*, *Папка*, *ЭкзО* – уникальностью соответствующих имен в иерархиях «владелец-подчиненный» (иерархических имен) в рамках рассматриваемой ПрО; для элементов множеств *TunXOf*, *TunXOn*, *TunXC*, *TunXIO*, *Tun3nIO* – уникальностью имен характеристик конкретных классов и типов объектов, классов событий и параметров объектов соответственно.

Ссылочная целостность, предполагающая обязательное наличие объекта, на который ссылается другой объект, в модели обеспечивается благодаря существующим типам взаимодействия между элементами множеств *Раздел*, *КлассО*, *КлассС*, *КлассПО*, *Папка*, *ЭкзО* – «владелец-подчиненный» (выражения (3)-(5) и им подобные).

Основными операциями в модели «объект-событие» являются: операция создания (вставки) элементов множеств, одноименных базовым понятиям модели \mathcal{A} , и их связей между собой в соответствии с \mathbb{R} и F ; операция изменения указанных элементов множеств, одноименных базовым понятиям модели; операция удаления указанных элементов множеств, одноименных базовым понятиям модели, а также зависимых от них (в соответствии с \mathbb{R} и F) элементов других множеств; операция выборки в соответствии с указанными условиями элементов из множеств, одноименных базовым понятиям модели.

Информационные запросы могут быть выражены в понятиях операций над множествами. Однако для неподготовленного специалиста теоретико-множественная модель является непростой для ее восприятия и понимания. Поэтому для достижения простоты и гибкости практического использования модели «объект-событие» был разработан специальный язык L – язык модели данных (ЯМД), сочетающий ясность и простоту использования его операторов для участников проекта, а также полноту представления данных моделируемой ПрО.

Более детально язык модели данных, близкий к некоторому подмножеству естественного языка, а также его возможности рассмотрены в работах [37,38].

4 Выводы

1. В результате анализа достижений в области семантического моделирования и перспективных направлений его развития, а также исходя из необходимости нахождения решений проблемы, связанной с потребностью своевременного создания и модернизации, в рамках запланированного бюджета, баз данных, обладающих требуемыми качествами, сформулированы требования, предъявляемые к разрабатываемой модели.

2. В соответствии с представленными требованиями синтезирована модель «объект-событие», которая содержит определенные подходы и решения, апробированные в известных «расширенных» моделях: ERM, EERM, HERM, ORM, объектной, семантических сетевых, онтологической, инфологической, основанной на логике предикатов с расширенной поддержкой концепции времени.

3 В процессе создания модели «объект-событие» были определены: множество базовых понятий модели, которые используются для описания реального мира; множество отношений между базовыми понятиями модели и множество функций интерпретации, заданных на базовых понятиях и отношениях (как набор формальных объектов, необходимых для представления базовых понятий модели); множество ограничений целостности; множество допустимых операций над данными и специальный язык модели данных, сочетающий в себе понятность и простоту использования его операторов для участников проекта, а также полноту представления данных моделируемой ПрО.

4. Применение средств разработанной модели «объект-событие» позволяет:

- обеспечить комплексное представление данных моделируемой ПрО, их структуры и ограничений целостности;
- упростить процесс преобразования концептуальной модели ПрО в схему реляционной БД;
- реализовать возможность использования модели, как на этапе концептуального проектирования БД, так и на стадии функционирования РБД ИСОУ.

Ссылки

- [1] Chaos Manifesto 2013: Think Big, Act Small online version. The Standish Group [Electronic Resource]. – Way of access: <http://www.versionone.com/assets/img/files/ChaosManifesto2013.pdf>. – Title from the screen.
- [2] Standish Group 2015 Chaos Report – Q&A with Jennifer Lynch [Electronic Resource]. – Way of access: <https://www.infoq.com/articles/standish-chaos-2015>. – Title from the screen.
- [3] Deit K. Dzh. Vvedenie v sistemy baz dannykh / K. Dzh. Deit; per. s angl. – [8-e izd.]. – Moskva : Izdatel'skii dom "Vi-l'yams", 2005. – 1328 s.
- [4] Chen P. P. S. The entity-relationship model – toward a unified view of data / P. P. S. Chen // ACM Transactions on Database Systems (TODS). – 1976. – Vol.1. – № 1. – P. 9 – 36.
- [5] Teorey T. J. Database modeling and design: logical design / T. J. Teorey, S. S. Lightstone, T. Nadeau. – Elsevier, 2006. – 282 p.
- [6] Thalheim B. Entity-relationship modeling: foundations of database technology / B. Thalheim. – Berlin; Heidelberg: Springer-Verlag, 2000. – 639 p.
- [7] Halpin T. Business roles and object-role modeling / T. Halpin // DBP&D. – 1996. – № 10. – P. 66–72.
- [8] Halpin T. Conceptual schema and relational database design / T. Halpin. – [2nd edition]. – Sydney, Australia: Prentice-Hall of Australia Pty. Ltd., 1995. – 500 p.
- [9] Halpin T. Entity Relationship modeling from an ORM perspective: Part 1 / T. Halpin // Journal of Conceptual Modeling. – 2000. – № 13. – P. 1 – 10.
- [10] Halpin T. Using object role modeling to design relational databases: Interview / T. Halpin // DBMS. – 1995. – № 8 (9). – 38 p.
- [11] Nijssen G. M. The Entity-Relationship Data Model Considered Harmful / G. M. Nijssen, D. J. Duke, S. M. Twine // Proc. 6th Symposium on Empirical Foundations of Information and Software Sciences. – Atlanta, Ga., 1988. – P. 109–130.
- [12] Nijssen G. M. Conceptual Schema and Relational Database Design: a fact oriented approach / G. M. Nijssen, T. A. Halpin. – Prentice-Hall, Inc., 1989. – 342 p.
- [13] Kalinichenko L. A. Standart sistem upravleniya ob"ektnymi bazami dannykh ODMG-93: kratkii obzor i otsenka sosto-yaniya / L. A. Kalinichenko // SUBD. – 1996. – №1. – S. 102–109.
- [14] Object-Oriented Database System Manifesto / M. Atkinson, F. Bancilhon, D. DeWitt and other // Proc. 1st Int. Conf. Deductive and Object-Oriented Databases. – Kyoto, Japan, 1989. – P. 40–57.

- [15] The object database standard: ODMG – 93 / [Ed. by R. G.G. Cattell]. – Burlington, USA: Morgan Kaufmann Publ., 1994. – 169 p.
- [16] The object data standard: ODMG 3.0. / [Ed. by R.G.G. Cattel, Douglas K. Barry]. – Burlington, USA: Morgan Kauffmann Publ., 2000. – 280 p.
- [17] Tsikritzis D. Modeli dannykh / D. Tsikritzis, F. Likhovski: per. s angl. – Moskva: Finansy i statistika, 1985. – 344 s.
- [18] Abrial J. R. Data semantics / J. R. Abrial // Data Base Management: [Ed. by J. W. Klimbie and K. L. Koffeman]. – North-Holland, Amsterdam, 1974. – P. 1–59.
- [19] Mylopoulos J. Information System Design at the Conceptual Level – the TAXIS Project / J. Mylopoulos, A. Borgida, S. Greenspan, H.K.T. Wong // IEEE Database Engineering Bulletin. – 1984. – Vol. 7. – № 4. – P. 4–9.
- [20] Roussopoulos N. Using semantic networks for data base management / N. Roussopoulos, J. Mylopoulos // Proceedings of the 1st International Conference on Very Large Data Bases. – ACM. – 1975. – P. 144–172.
- [21] Bubenko J. Data Models and their Semantics / J. Bubenko // Data Design. Infotech State of the Report Series. – 1980. – Vol. 8. – № 4. – P. 107–136.
- [22] Langefors B. Infological model and information user views / B. Langefors // Information Systems. – 1980. – № 5. – P.17–32.
- [23] Sundgren B. Conceptual foundation of the infological approach to data bases / B. Sundgren // Data Base Management: [Ed. by J. W. Klimbie and K. L. Koffeman]. – North-Holland, Amsterdam, 1974. – P. 61–96.
- [24] Palagin A. V. Ontologicheskie metody i sredstva obrabotki predmetnykh znaniy: monografiya / A. V. Palagin, S. L. Kryvyi, N. G. Petrenko. – Lugansk: Izd-vo VNU im. V. Dalya, 2012. – 323 s.
- [25] Gruber T. R. A translation approach to portable ontology specifications / T. R. Gruber // Knowledge acquisition. – 1993. – Vol. 5. – № 2. – P. 199 – 220.
- [26] Gruber T. R. Toward principles for the design of ontologies used for know ledge sharing / T. R. Gruber // International journal of human-computer studies. – 1995. – Vol. 43. – № 5. – P. 907–928.
- [27] Guarino N. Formal Ontology and Information Systems / N. Guarino // Formal Ontology in Information Systems: Proceedings of FOIS'98, 6–8 June 1998, Trento, Italy. – Amsterdam: IOS Press, 1998. – P. 3–15.
- [28] Codd E. F. A relational model of data for large shared data banks / E. F. Codd // Communications of the ACM. – 1970. – Vol. 13. – № 6. – P. 377–387.
- [29] Codd E. F. Extending the database relational model to capture more meaning / E. F. Codd // ACM Transactions on Database Systems (TODS). – 1979. – Vol. 4. – № 4. – P. 397–434.
- [30] Kogalovskii M. R. Sistemy dostupa k dannym, osnovannye na ontologiyakh / M. R. Kogalovskii // Programirovanie. – 2012. – № 4. – S. 55–77.
- [31] Kogalovskii M. R. Kontseptual'noe modelirovanie v tekhnologiyakh baz dannykh i ontologicheskie modeli / M. R. Kogalovskii, L. A. Kalinichenko // Ontologicheskoe modelirovanie: trudy simpoziuma. – Moskva : IPI RAN, 2008. – S. 114–148.
- [32] Pospelov D. A. Situatsionnoe upravlenie: teoriya i praktika / D. A. Pospelov. – Moskva : Nauka; Gl. red. fiz.-mat. lit., 1986. – 288 s.
- [33] Lukashevich N. V. Tezaurusy v zadachakh informatsionnogo poiska / N. V. Lukashevich. – Moskva: Izdatel'stvo Moskovskogo universiteta, 2011. – 512 s.
- [34] Kuznetsov S. D. Osnovy baz dannykh / S. D. Kuznetsov. – Moskva: Internet-Universitet Informatsionnykh Tekhnologii; BINOM. Laboratoriya znaniy, 2007. – 484 s.
- [35] Tal'khaim B. Obzor semanticheskikh ogranichenii dlya modelei baz dannykh / B. Tal'khaim [Elektronnyi resurs]. – Rezhim dostupa : [www.intsys.msu.ru/magazine/archive/v3\(3.../thalheim-307-351.pdf](http://www.intsys.msu.ru/magazine/archive/v3(3.../thalheim-307-351.pdf). – Zagl. s ekrana.
- [36] Date C. J. An Introduction to Database Systems / C. J. Date: [8th Edition]. – Pearson: Addison-Wesley, 2004. – XXVII, 983, I – 22 p.
- [37] Esin V. I. Yazyk dlya universal'noi modeli dannykh / V. I. Esin, M. V. Esina // Systemy obrobky informacii'. – 2011. – № 5(95). – S. 193–197.
- [38] Esin V. I. Yazyk opisaniya i manipulirovaniya dannymi, khranyashchimisya v BD s UMD / V. I. Esin, M. V. Esina // Komp'yuternoe modelirovanie v naukoemkikh tekhnologiyakh (KMNT-2010): tezisy dokl. mezhdunar. nauch.-tekhn. konf., 18-21 maya 2010 g. – Khar'kov: Khar'kovskii natsional'nyi universitet im. V.N. Karazina, 2010. – Chast' 2. – S. 104–108.

Reviewer: Sergii Kavun, Doctor of Sciences (Economics), Ph.D. (Engineering), Full Prof., Kharkiv Educational and Research Institute of the University of Banking, Kharkiv, Ukraine.

E-mail: kavserg@gmail.com

Received: June 2017

Authors:

V. Yesin, Doctor of Sciences (Engineering), Department of Information Systems and Technologies Security, V.N. Karazin Kharkiv National University, Kharkiv, Ukraine.

E-mail: v.i.yesin@karazin.ua

The data model "object-event": requirements and synthesis of the model.

Abstract. Requirements imposed to the developed data model are formulated based on the need to find new solutions of the actual problem of timely creation, modernization within the planned budget of databases that have the required qualities. The data model, called "object-event", is synthesized in accordance with the formulated requirements and the general approach to the problem of semantic modeling proposed by C. Date.

Key words: data model, semantic (conceptual) modeling, database.

Рецензент: Сергій Кавун, доктор економічних наук, к.т.н., проф., Харківський інститут банківської справи УБС НБУ, Харків, Україна.

E-mail: kavserg@gmail.com

Надійшло: Червень 2017.

Автори:

Віталій Єсін, д.т.н., каф. безпеки інформаційних систем і технологій, Харківський національний університет імені В.Н. Каразіна, Харків, Україна.

E-mail: v.i.yesin@karazin.ua

Модель даних «об'єкт-подія»: вимоги та синтез моделі.

Анотація. Виходячи з необхідності знаходження нових рішень актуальної задачі своєчасного створення, модернізації в рамках запланованого бюджету баз даних, що володіють необхідними якостями, формулюються вимоги, що висувуються до розроблюваної моделі даних. Відповідно до сформульованих вимог і загального підходу до проблеми семантичного моделювання, запропонованого К. Дейтом, синтезується модель даних, що отримала назву – «об'єкт-подія».

Ключові слова: модель даних, семантичне (концептуальне) моделювання, база даних.