

## ИНВАРИАНТНАЯ К ПРЕДМЕТНЫМ ОБЛАСТЯМ СХЕМА БАЗЫ ДАННЫХ И ЕЕ ОТЛИЧИТЕЛЬНЫЕ ОСОБЕННОСТИ

### Введение

Результаты исследований состояния информатизации в различных компаниях, организациях, учреждениях свидетельствуют о том, что многие из них владеют определенными информационными системами организационного управления (ИСОУ). При этом для решения возникающих новых задач, как правило, связанных с расширением деятельности и, соответственно, рассматриваемыми предметными областями (ПрО), они хотят иметь более функциональные, с улучшенными характеристиками качества информационные системы (ИС), требующие меньших затрат по сопровождению. В этих условиях востребованными становятся проекты: по разработке новых ИСОУ и их интеграции с существующими информационными системами; разработке новых ИСОУ с целью замены существующих ИС; модернизации существующих ИСОУ. Суть данных проектов заключается в проведении процедур реинжиниринга существующих ИС и их основного функционального компонента – базы данных (БД).

Одним из важных требований, предъявляемых к реинжинирингу существующих ИСОУ и их БД, является своевременность завершения соответствующих проектов в рамках запланированного бюджета с заданными характеристиками качества, которое, как показывают результаты анализа IT-проектов, проведенного международными экспертами, к сожалению, не всегда выполняется. Более 60 % проектов были провалены или завершены с опозданием, причем с гораздо большими затратами, чем планировалось [1, 2, 3].

Таким образом, имеет место проблема, связанная 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 ИСОУ указанная ограниченность возможностей обусловлена ориентацией традиционной методологии их проектирования на итерационную, достаточно трудоемкую процедуру создания уникальной концептуальной модели, логической и физической схем при разработке новой БД либо на существенное их преобразование при модернизации. Что часто приводит к расходованию значительных, не всегда прогнозируемых, временных и финансовых ресурсов. В результате возникает объективная необходимость пересмотра существующих подходов, методологий и технологий создания, модернизации баз данных, которые позволят избавиться от необходимости затратной политики выполнения лишних работ при реинжиниринге БД ИСОУ, в том числе и на этапе физического проектирования РБД.

Стремление избежать излишних затрат, свойственных традиционной методологии создания РБД, не только на этапах концептуального и логического проектирования, предлагаемые решения для которых достаточно подробно излагаются в работах [4 – 8], но и на этапе физического проектирования, актуализировало задачу разработки инвариантной к предметным областям схемы БД.

### Основные отношения схемы базы данных

Создание инвариантной к ПрО схемы БД стало возможным благодаря разработанной модели данных с универсальным базисом отношений [7, 8]. Данная схема является отображением универсального базиса отношений и множества ограничений целостности модели [7, 8], которая в свою очередь является отображением модели данных «объект-событие» [4 – 6].

Основными разработанными и отличающимися элементами инвариантной к ПрО схемы БД как описания содержания, структуры и ограничений целостности, используемых для создания и поддержки БД, являются: состав, структура базовых отношений; реализации ограничений целостности; средства, обеспечивающие безопасность БД.

Для лучшего понимания принципиальных отличительных особенностей базовых отношений (таблиц) инвариантной к предметным областям схемы БД вначале целесообразно отметить, что в схемах, проектируемых по традиционной технологии РБД, каждое базовое отношение, как правило, ассоциируется с конкретной сущностью моделируемой ПрО, а атрибут (столбец) этого отношения – со свойством сущности. При этом информация о базовых отношениях с указанием их имен, имен атрибутов, связанная с описанием моделируемой ПрО, а также типах данных атрибутов, связанных с типами данных конкретной СУБД, на платформе которой реализуется БД, размещается в системном каталоге (словаре) СУБД (рис. 1). Поэтому при таком построении схемы БД различные изменения в ПрО, даже незначительные, вызывают необходимость модернизации словаря СУБД, то есть приводят либо к незначительной, либо значительной модификации существующей схемы, вплоть до разработки новой. В процессе же осуществления последовательного отображения: модель «объект-событие» [4 – 6] – модель данных с универсальным базисом отношений [7, 8] – инвариантная к ПрО схема БД, создается набор фиксированных базовых отношений с их именами и именами атрибутов (заголовками столбцов), непосредственно не ассоциируемых с конкретными сущностями рассматриваемой ПрО и их свойствами.

Это объясняется тем, что семантические понятия/концепты (semantic concepts) модели «объект-событие», связываемые с именами соответствующих базовых отношений инвариантной к ПрО схемы БД ( $R^{sh}$ ), являющихся отображением отношений  $\mathfrak{R}$  модели с универсальным базисом отношений:  $M_{umd} = \langle \mathfrak{R}, Pr, L \rangle$ , где  $Pr$  – множество ограничений целостности;  $L$  – язык манипулирования данными, – есть общие понятия-категории для всех ПрО. А имена атрибутов базовых отношений либо ассоциируются с именами множеств, соотносимых с такими семантическими понятиями, либо с множествами соответствующих им идентификаторов, либо с некоторыми ограничениями им присущими (рис. 2).

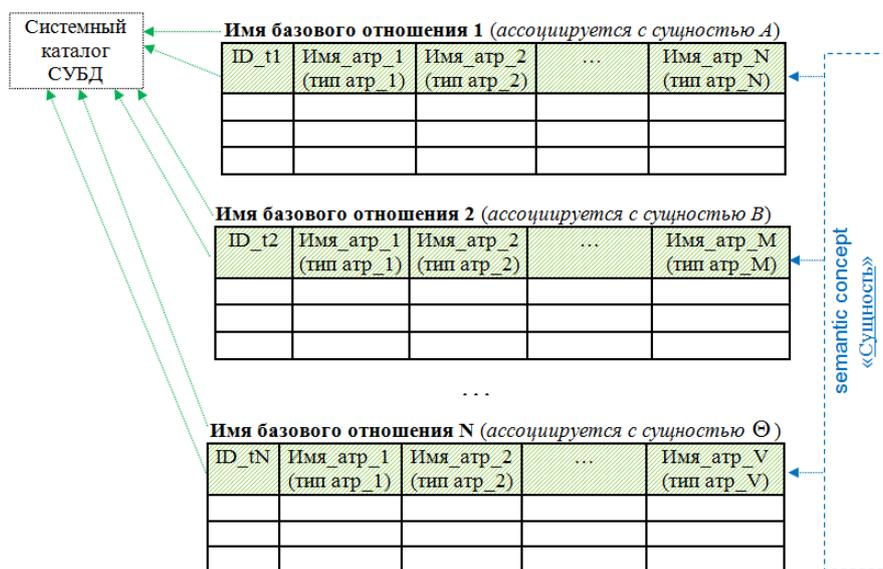


Рис. 1. Базовые отношения схемы БД, разрабатываемой по традиционной технологии, и их связь с сущностями моделируемой ПрО

Поэтому, когда такая информация вместе с указанием типа данных атрибутов, принятых в СУБД, на платформе которой реализуется инвариантная к ПрО схема БД, помещается в системный каталог СУБД, то различные изменения в предметной области не вызывают необходимость внесения в него корректив.

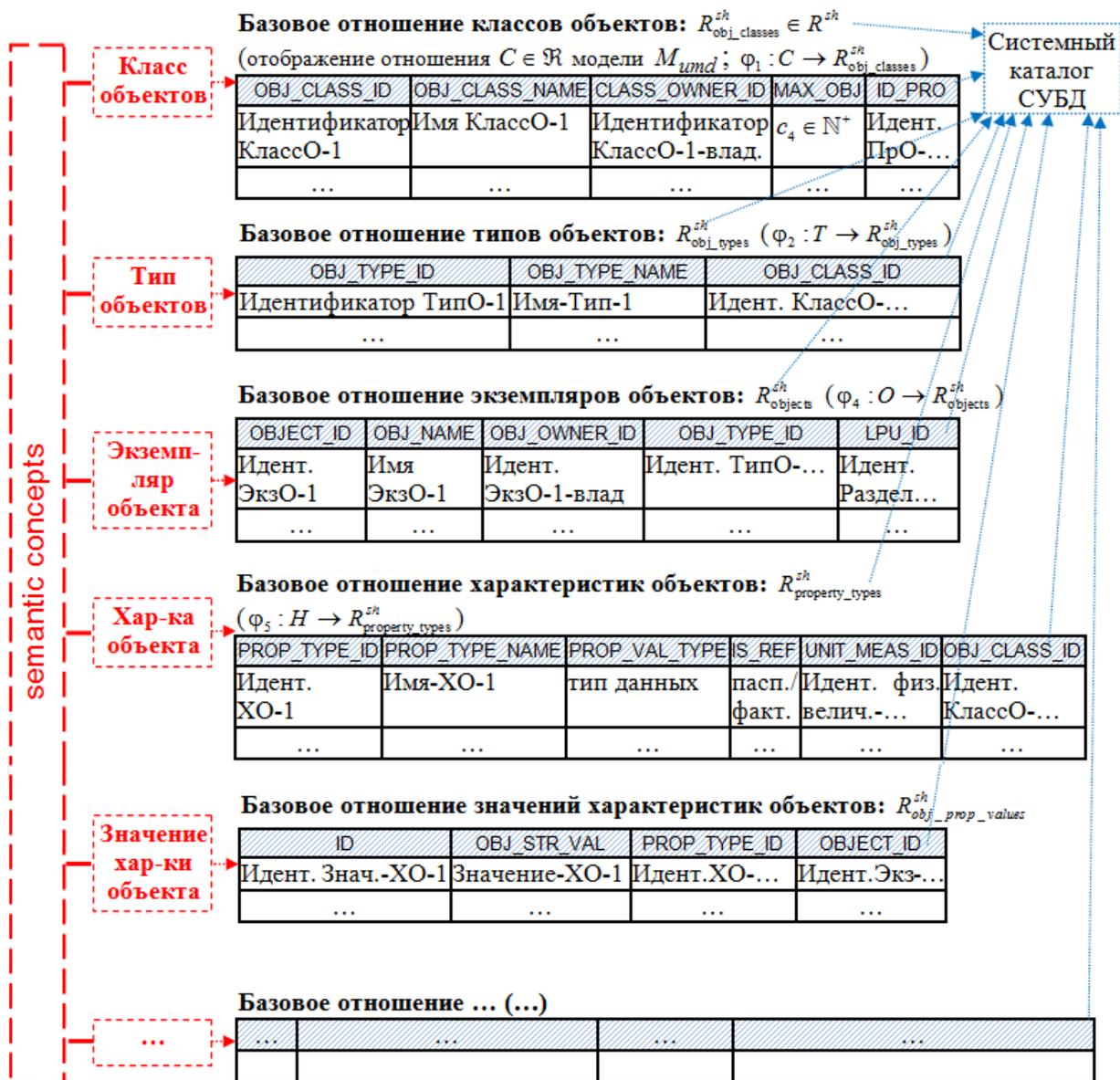


Рис. 2. Базовые отношения инвариантной к Про схемы БД и их связь с семантическими концептами модели «объект-событие»

Описание моделируемой Про (сущностей, их свойств, связей), трансформированное при традиционной разработке РБД в набор именованных базовых отношений с атрибутами и связями между отношениями, в случае использования инвариантной к Про схемы БД отображается в область изменяемых данных фиксированных базовых отношений схемы, ассоциируемых с метаданными Про. Эта область определяется как область проблемных метаданных (выделена горизонтальными линиями в правых таблицах рис. 3, с их соответствием в левых таблицах, являющихся отображением моделируемой Про, выполненном при традиционной разработке РБД). В ней содержатся метаданные предметной области, используемые при определении ее данных. При этом данные моделируемой Про помещаются также в область изменяемых данных, но других отношений инвариантной к Про схемы БД, ассоциируемых с экземплярами объектов, событий, значений характеристик, конкретных документов. Эта область определяется как область данных (выделена точками рис. 3).

Такая организация данных обеспечивает совместное хранение метаданных и данных моделируемой Про в области изменяемых данных соответствующих базовых отношений инвариантной к Про схемы БД, в отличие от раздельного хранения метаданных Про в словаре СУБД и самих данных собственно в базе данных (в базовых отношениях ее схемы) при

традиционном построении РБД. В результате появляется гибкость – возможность описывать ПрО, в том числе и происходящие в ней изменения, на уровне проблемных метаданных, без изменения собственно структуры базовых отношений схемы БД. Это позволяет созданную базу данных достаточно просто адаптировать к изменениям в ПрО при стабильной структуре отношений схемы БД. Заранее неограниченное многообразие элементов ПрО распределяется по фиксированному набору базовых отношений инвариантной к ПрО схемы БД, обеспечивая при этом возможность одновременного хранения и использования данных различных существенно отличающихся ПрО.

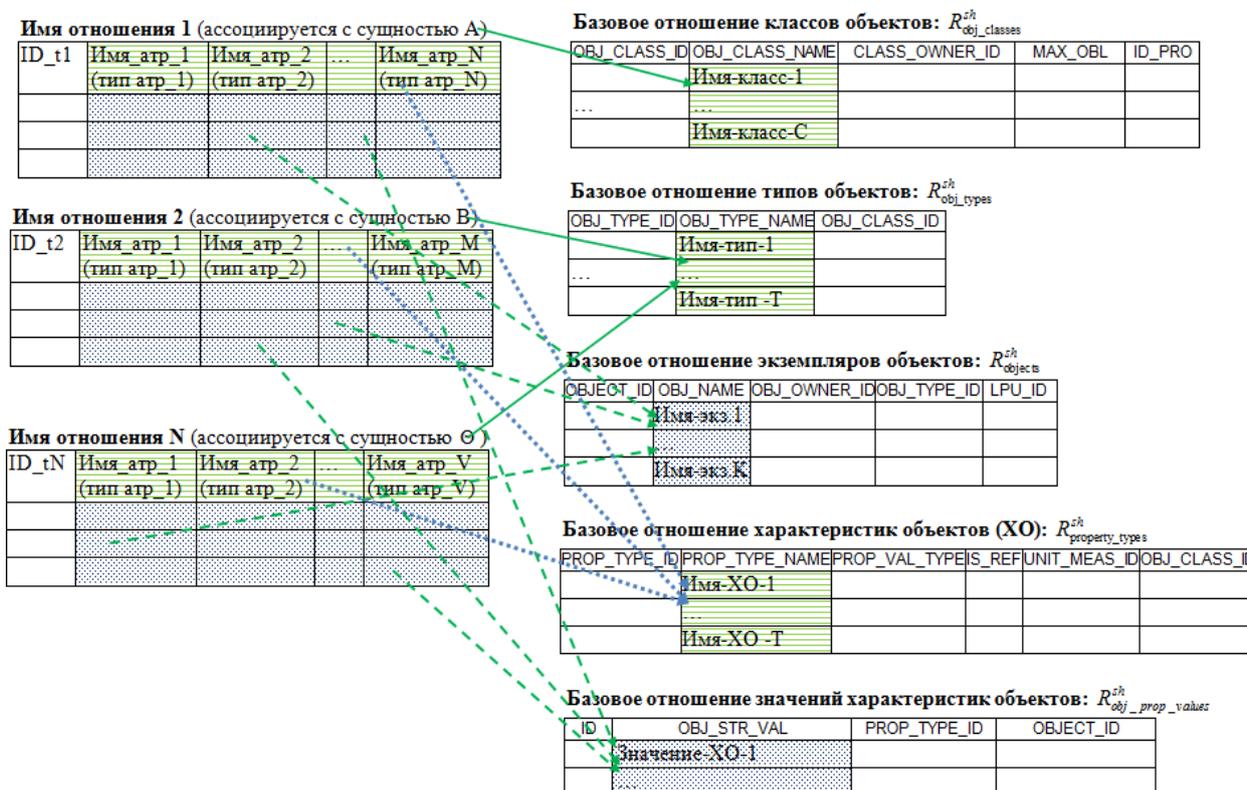


Рис. 3. Схематическое представление трансформации описания предметной области в инвариантную к ПрО схему БД и в схему РБД, спроектированную по традиционной технологии

Ниже приведены результаты отображения универсального базиса отношений  $\mathfrak{R}$  модели  $M_{umld}$  в базовые отношения разрабатываемой схемы ( $\varphi: \mathfrak{R} \rightarrow R^{sh}$ ), представленные в виде диаграммы в нотации IDEF1X (рис. 4).

Полученные базовые отношения инвариантной к ПрО схемы БД имеют принципиальные отличия в назначении, структуре, месте хранения описания метаданных моделируемой ПрО относительно создаваемых отношений при традиционной технологии проектирования реляционных БД. Их число, структура, в отличие от структуры и числа базовых отношений схем, разрабатываемых по традиционной технологии реляционных БД, не зависят от набора данных, они инвариантны к рассматриваемой ПрО.

### Обеспечение поддержки ограничений целостности данных

С целью предотвратить появление в базе противоречивых данных с использованием средств декларативной и процедурной поддержки ограничений целостности были определены реализации ограничений целостности  $Pr^{sh}$  в создаваемой схеме как некоторый набор шаблонов конструкций операторов (программный код), реализующий ограничение, полученный в результате отображения множеств ограничений целостности  $Pr$ , которые специ-

фицируются в модели  $M_{umd}$  во множество ограничений целостности  $Pr^{sh}$  ( $\gamma: Pr \rightarrow Pr^{sh}$ ), реализуемых в предлагаемой схеме.

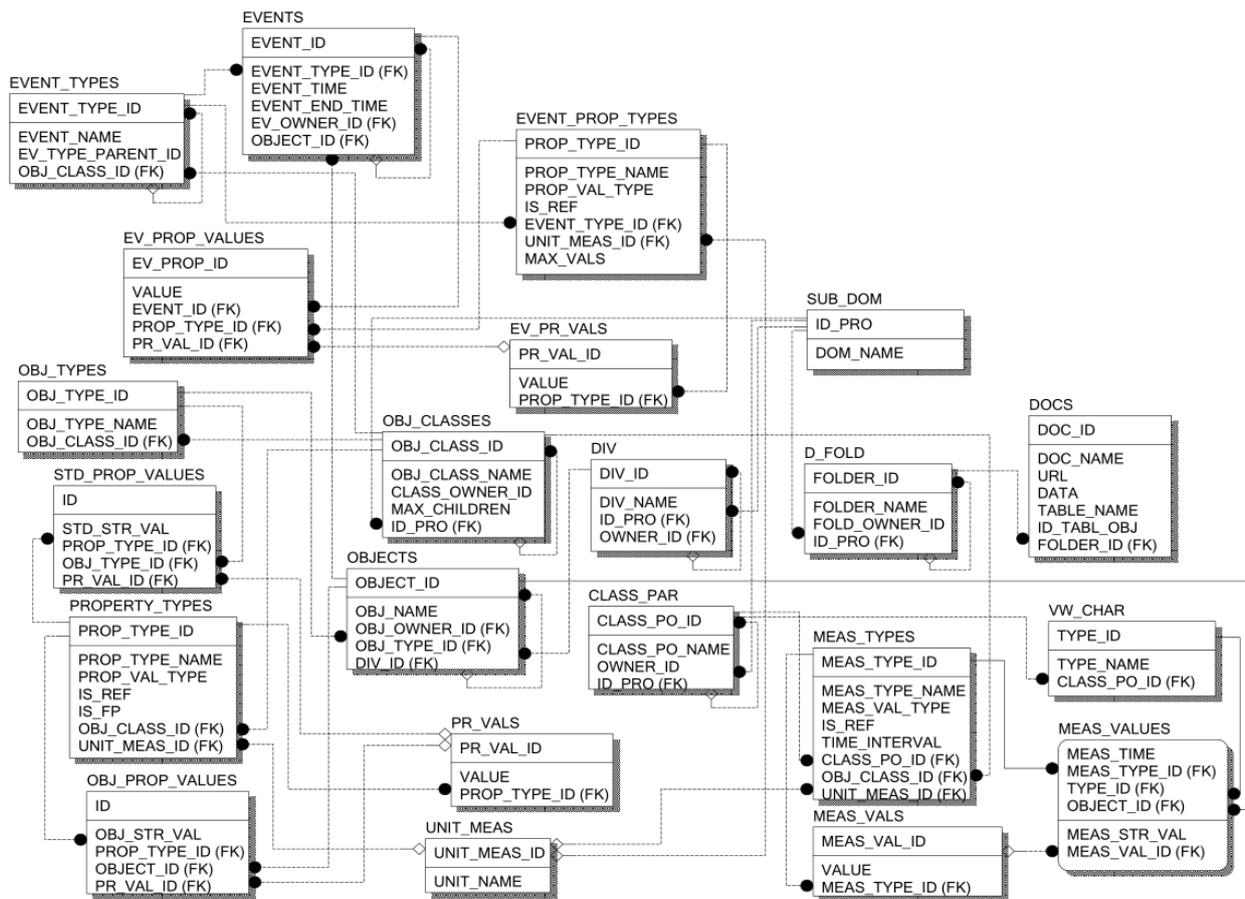


Рис. 4. Диаграмма основных базовых отношений инвариантной к ПрО схемы БД

Декларативная поддержка ограничений целостности заключается в определении ограничений средствами DDL (Data Definition Language) языка SQL. Средства декларативной поддержки целостности использовались при создании базовых отношений инвариантной к ПрО схемы БД для определения таких типов ограничений, как целостность сущностей, ссылочная целостность, обязательные данные (*not null*), ограничения доменов.

Как известно, целостность сущностей связывается в первую очередь с уникальностью и несократимостью первичных ключей [3, 9]. При создании основных базовых отношений схемы, представленных в виде основных строк кода языка DDL (как результат отображения  $\varphi: \mathcal{R} \rightarrow \mathcal{R}^{sh}$ ), подобные требования поддержки целостности данных были определены для всех отношений схемы с помощью конструкций стандарта ISO SQL *primary key*, *unique* операторов создания (изменения) базовых отношений. Формально этот факт можно представить как результат отображений:

$$\gamma_{PK} : Pr_{PK} \rightarrow Pr_{constr\_primary\_key}^{sh} ; \gamma_{UK} : Pr_{UK} \rightarrow Pr_{constr\_unique}^{sh}$$

В результате отображения  $\gamma_{FK} : Pr_{FK} \rightarrow Pr_{constr\_foreign\_key}^{sh}$  (применения конструкции *foreign key* операторов *alter (create) table*) для обеспечения ссылочной целостности были определены внешние ключи отношений схемы и стратегии действий при удалении данных.

Для ускорения доступа к данным в соответствии с рекомендациями, приведенными в работах [3, 11, 12], на внешних ключах отношений предлагаемой схемы были определены

дополнительные индексы (с учетом, что на первичных и уникальных ключах СУБД, как правило, автоматически создает уникальный индекс).

В результате отображения  $\gamma_{not\_null} : Pr_{not\_null} \rightarrow Pr_{constr_{not\_null}}^{sh}$  (применения спецификатора *not null* в операторах *alter (create) table*) были заданы ограничения, запрещающие присваивание соответствующим атрибутам неопределенных значений *null*.

С использованием отображения множества ограничений целостности модели данных с универсальным базисом отношений  $M_{umd}$  в инвариантной к ПрО схеме БД были определены ограничения для доменов атрибутов признаков, типов данных характеристик объектов, событий, параметров объектов и некоторые другие (как результат отображения  $\gamma_{dom} : Pr_{dom} \rightarrow Pr_{constr_{check}}^{sh}$  – применения конструкции *check* оператора *alter table*).

Однако не все ограничения можно реализовать с помощью декларативной поддержки целостности. Поэтому наряду со средствами этого способа реализации ограничений целостности широкое применение нашли средства процедурной поддержки – триггеры, хранимые процедуры, функции, механизмы которых в течение нескольких последних лет во многих коммерческих СУБД были существенно расширены [10].

Необходимость применения в инвариантной к ПрО схеме БД триггеров, хранимых процедур и функций обусловлена многими факторами, в том числе невозможностью реализации средствами декларативной поддержки для некоторых отношений сложных ограничений целостности и ссылочной целостности, проверки корректности вводимых данных, а также потребностью автоматического формирования (генерации) первичных ключей. В связи с этим, следующие ограничения целостности были реализованы с помощью средств процедурной поддержки ( $Pr_{constr_{proc}}^{sh}$ ):

– ограничение на возможность изменения занесенных в соответствующие отношения схемы метаданных ПрО (идентификаторов ассоциируемых с определенными базовыми понятиями, рис. 2; чисел максимальных значений в отношении  $R_{event\_prop\_types}^{sh}$  [max\_vals],  $max\_vals \in at(R_{event\_prop\_types}^{sh})$ ); удаления списочных значений соответствующих характеристик из отношений  $R_{pr\_vals}^{sh}$ ,  $R_{ev\_pr\_vals}^{sh}$ ,  $R_{meas\_vals}^{sh}$ , если они присутствуют в отношениях  $R^{sh}$ , ассоциируемых с данными моделируемой ПрО, и т. д.);

– ограничение на возможность ввода новых данных, противоречащих введенным метаданным ПрО (для отношений  $R^{sh}$ , ассоциируемых с данными ПрО);

– реализация ссылочной целостности для отношений схемы  $R^{sh}$ , связанных с отношением  $R_{docs}^{sh}$  (конкретный документ из отношения  $R_{docs}^{sh}$  связывается с конкретным экземпляром соответствующего отношения  $R^{sh}$  (рис. 4));

– ограничение максимального количества экземпляров объектов (из отношения  $R_{objects}^{sh}$ ) для определенного класса объектов (из  $R_{obj\_classes}^{sh}$ );

– ограничение максимального количества значений (из  $R_{ev\_prop\_values}^{sh}$ ), которые могут быть присвоены определенной характеристике события (из  $R_{event\_prop\_types}^{sh}$ ) для экземпляра события (из  $R_{events}^{sh}$ ) заданного класса;

– ограничение на количество событий (из  $R_{events}^{sh}$ ), происходящих с одним экземпляром объекта (из  $R_{objects}^{sh}$ ):

а) в один и тот же момент времени с одним экземпляром объекта не может происходить больше одного события;

б) у одного события, происходящего с одним экземпляром объекта, может быть несколько подчиненных событий с разными экземплярами объектов, происходящими в один момент времени, но у конкретного экземпляра события, которое происходит с экземпляром объекта определенного класса, событие-владелец может быть только одно;

– генерация уникальных значений кодов первичных ключей для отношений схемы  $R^{sh}$ .

На рис. 5 приведена блок-схема алгоритма использования средств декларативной и процедурной поддержки ограничений целостности при разработке объектов инвариантной к ПрО схемы БД.

Следует заметить, что область применения процедур (в том числе триггерных), функций как объектов инвариантной к предметным областям схемы БД не ограничивается задачами поддержки целостности. Они используются также для решения следующих задач: выдачи предупреждающих сообщений об исключительных ситуациях или напоминающих сообщениях о необходимости выполнения некоторых действий; преобразования данных; обеспечения защищенности данных от нежелательного разглашения (нарушения конфиденциальности), искажения, потери или снижения меры доступности и других.

При решении проблемы защиты БД ИСОУ как важнейшего корпоративного ресурса в процессе создания инвариантной к ПрО схемы базы данных были разработаны специальные средства (в виде реализованных объектов схемы) и правила их использования, обеспечивающие: управление доступом к объектам схемы, защиту данных и сокрытие объектов, восстановление неправильно измененных или утраченных данных. Это отдельная и важная тема, которую целесообразно более подробно рассмотреть в рамках другой статьи.

В соответствии со сформулированными факторами успешного проектирования инвариантной к предметным областям схемы БД, в том числе предусматривающими необходимость учета в максимально возможной степени особенностей требуемой обработки данных, характеристик будущих пользователей, возможностей технических и программных средств, на которых планируется реализация БД ИСОУ, в дополнение к приведенным выше объектам схемы были разработаны и другие ее объекты. В их числе: хранимые процедуры, функции, представления (views), некоторые дополнительные базовые отношения. Таким образом, в состав предлагаемой схемы базы данных вошли следующие объекты (компоненты):

– фиксированный набор базовых отношений  $R^{sh}$ , основные из которых приведены на диаграмме (рис. 4);

– виртуальные отношения (представления – views);

– триггеры, хранимые процедуры (в том числе хранимые процедуры интерпретатора специального непроцедурного языка модели данных (ЯМД) [13, 14]), функции, пакеты программ;

– политики безопасности (набор декларативных команд, которые определяют, как и когда следует применять ограничения доступа пользователей к кортежам  $R^{sh}$ );

– последовательности и т. д.

Диаграмма компонентов инвариантной к ПрО схемы БД с отношениями зависимости между ними в нотации языка UML приведена на рис. 6.

Следует отметить, что при необходимости, например, для решения некоторых задач, данная схема может дополняться новыми объектами. При этом обязательным условием является недопустимость внесения любых изменений в ее ядро (то есть во все выше рассмотренные элементы схемы). В настоящее время данная схема БД реализована на платформах СУБД Oracle и PostgreSQL.

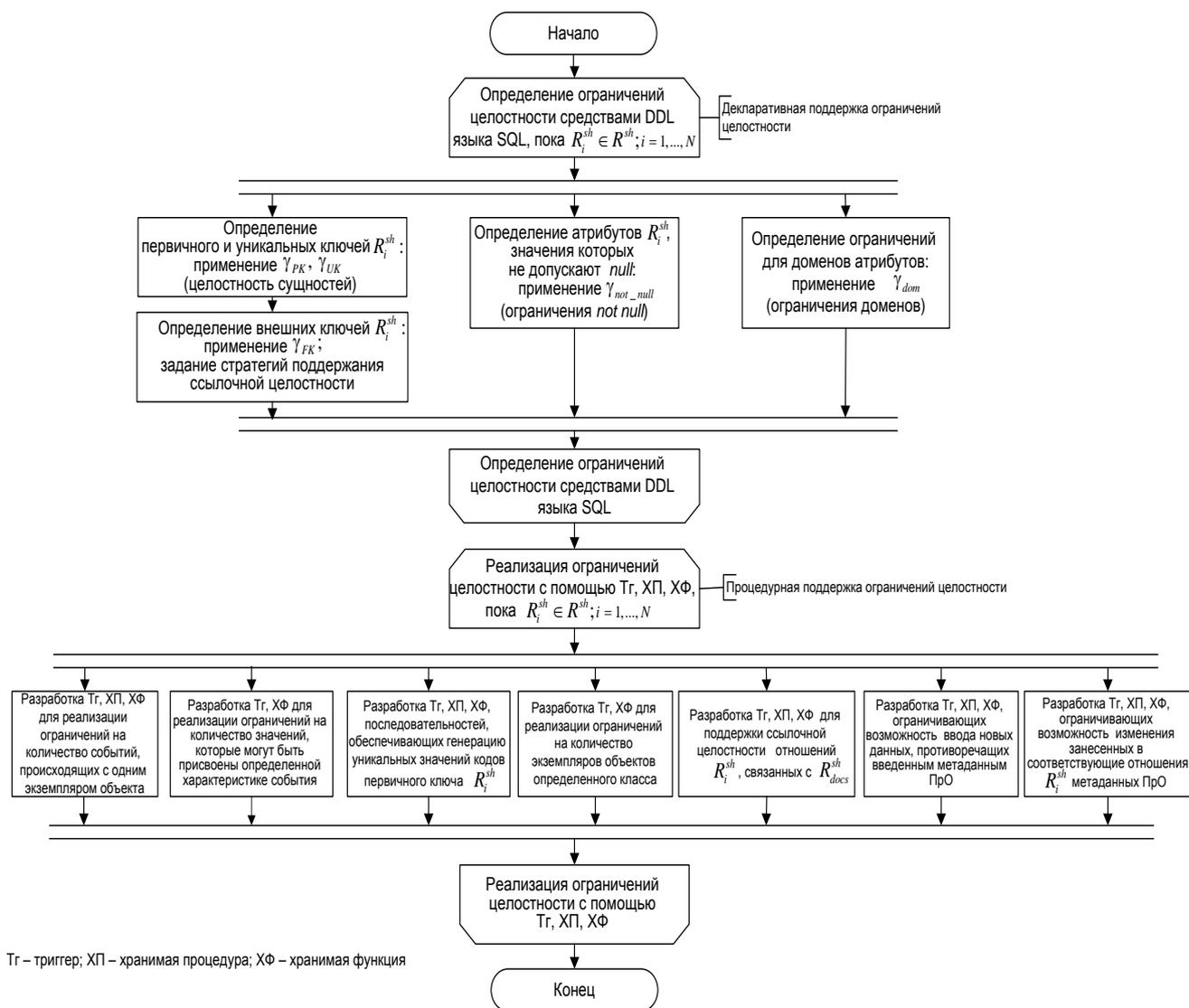


Рис. 5. Блок-схема алгоритма использования средств декларативной и процедурной поддержки ограничений целостности при разработке объектов инвариантной к ПрО схемы БД



Рис. 6. Диаграмма компонентов инвариантной к ПрО схемы БД

### Выводы

1. Число, структура базовых отношений инвариантной к предметным областям схемы БД в отличие от структуры и числа отношений схем, разрабатываемых по традиционной технологии РБД, не зависят от набора данных, они инвариантны к рассматриваемой ПрО.

2. Разработанная инвариантная к предметным областям схема базы данных позволяет упростить процесс адаптации РБД ИСОУ к изменениям условий функционирования. Это достигается за счет использования созданного набора базовых отношений, имеющих принципиальные отличия в назначении, структуре, месте хранения описания метаданных моделируемой ПрО относительно создаваемых базовых отношений при традиционной технологии проектирования реляционных БД. Предлагаемая организация данных обеспечивает совместное хранение метаданных и данных моделируемой ПрО в области изменяемых данных соответствующих базовых отношений схемы БД, в отличие от раздельного хранения метаданных ПрО в словаре СУБД и самих данных собственно в базе данных (в базовых отношениях ее схемы) при традиционном построении РБД. В результате появляется гибкость – возможность описывать ПрО, в том числе и происходящие в ней изменения, на уровне проблемных метаданных, без изменения собственно структуры базовых отношений схемы БД. При расширении набора объектов, событий, характеристик объектов, событий, параметров объектов моделируемой ПрО в БД не создаются новые базовые отношения, атрибуты, ключи или иные объекты схемы, а просто добавляется новая запись в одно из существующих базовых отношений предлагаемой схемы. Более того, в БД ИСОУ, построенных на основе инвариантной к предметным областям схемы базы данных, можно одновременно хранить данные нескольких ПрО.

3. Предлагаемая инвариантная к предметным областям схема БД при необходимости может быть расширена в интересах потребителя информационного продукта, путем дополнения различными новыми объектами: базовыми и виртуальными отношениями, процедурами, функциями и другими. Обязательным условием при этом является недопустимость любых изменений ядра схемы.

### Список литературы:

1. Chaos Manifesto 2013: Think Big, Act Small online version. The Standish Group [Electronic resource]. Access mode : <http://www.versionone.com/assets/img/files/ChaosManifesto2013.pdf>.
2. Standish Group 2015 Chaos Report – Q&A with Jennifer Lynch [Electronic resource]. Access mode : <https://www.infoq.com/articles/standish-chaos-2015>.

3. Connolly T. M. Database systems: a practical approach to design, implementation, and management. Sixth edition / Thomas M. Connolly, Carolyn E. Begg. Harlow, Essex, England : Pearson Education Limited, 2015. 1329 p.
4. Security and noise immunity of telecommunication systems: new solutions to the codes and signals design problem. Collective monograph. Edited by Sergey G. Rassomakhin, Alexandr A. Kuznetsov. Minden, Nevada, USA : ASC Academic Publishing. 2017. 198 p., Chapter 8, Yesin V.I., Yesina M.V. Means for conceptual modeling of information system databases. P.160-196.
5. Есин В. И. Модель данных «объект-событие»: требования и синтез модели // Computer science and cyber security. International electronic scientific journal. 2017. Issue. 3 (7). P. 33-44. Access mode : <http://periodicals.karazin.ua/cscs/article/view/10003/9527>.
6. Есин В. И. Выразительные средства модели данных «объект-событие» / В. И. Есин // Радиотехника. 2017. Вып. 191. С. 99-112.
7. Есин В. И. Модель данных с универсальной фиксированной структурой / В. И. Есин // Теоретичні та прикладні аспекти побудови програмних систем : матеріали міжнародної наукової конференції, м. Київ, 15-17 грудня 2014 р. Кіровоград : ФО-П Александрова М. В., 2014. С. 112-116.
8. Есин В. И. Универсальная модель данных и ее математические основы / В. И. Есин // Системи обробки інформації. 2011. № 2(92). С.21-24.
9. Дейт К. Дж. Введение в системы баз данных. 8-е изд. ; пер. с англ. / К. Дейт. – Москва : Изд. дом "Вильямс", 2005. 1328 с.
10. Грофф Д. Р. SQL: полное руководство, 3-е изд. : Пер. с англ. / Грофф Д. Р., Вайнберг П. Н., Оппель Э. Дж. – Москва : ООО "И.Д. Вильямс", 2015. – 960 с.
11. Гарсиа-Молина Г. Системы баз данных // Г. Гарсиа-Молина, Д. Д. Ульман, Д. Уидом. Москва : Изд. дом "Вильямс", 2003. 1088 с.
12. Фиайли К. SQL ; пер. с англ. / К. Фиайли. Москва : ДМК Пресс, 2003. – 456 с.
13. Есин В. И. Язык для универсальной модели данных / В. И. Есин, М. В. Есина // Системи обробки інформації. 2011. № 5(95). С.193-197.
14. Есин В. И. Интерпретатор языка для универсальной модели данных / В. И. Есин, М. В. Есина // Наука і техніка Повітряних Сил Збройних Сил України. 2011. № 2(6). С. 140-143.

*Харьковский национальный  
университет имени В.Н.Каразина*

*Поступила в редколлегию 25.02.2018*