

9. Пройдакова Е.В., Федоткин М.А. Определение условий существования стационарного распределения выходных потоков в системе с циклическим управлением // Автоматика и телемеханика. – 2008. – № 6. – С. 96–106.
10. Mirasol N.M. The output of an $M|G|\infty$ queueing system is Poisson // Operations Research. – 1963. – V. 11. – P. 282–284.
11. Neuts M.F. A versatile Markovian arrival process // Journal of Appl. Prob. – 1979. – V. 16. – P. 764–779.
12. Дудин А.Н., Клименок В.И. Системы массового обслуживания с коррелированными потоками. – Минск: БГУ, 2000. – 175 с.
13. Назаров А.А., Моисеева С.П. Метод асимптотического анализа в теории массового обслуживания. – Томск: Изд-во НТЛ, 2006. – 112 с.
14. Лопухова С.В., Назаров А.А. Исследование МАР-потока методом асимптотического анализа N -го порядка // Вестник ТГУ. Серия Информатика. Кибернетика. Математика. – 2006. – № 293. – С. 110–115.
15. Лапатын И.Л., Назаров А.А. Исследование выходящего потока системы $МАР|M|\infty$ в условиях растущего времени обслуживания // Теория вероятностей, математическая статистика и их приложения: сб. науч. ст. – Вып. 2. – Минск: РИВШ, 2009. – С. 76–79.
16. Кениг Д., Штойян Д. Методы теории массового обслуживания: Пер. с нем. / Под ред. Г.П. Климова. – М.: Радио и связь, 1981. – 128 с.

Поступила 26.10.2009 г.

УДК 004.65

РАЗРАБОТКА МОДЕЛИ ПРЕДСТАВЛЕНИЯ МЕТАОПИСАНИЙ ИНФОРМАЦИОННОЙ БАЗЫ ЕДИНОЙ ИНФОРМАЦИОННОЙ СРЕДЫ ВУЗА

В.В. Осипова, И.Л. Чудинов

Томский политехнический университет

E-mail: vikosi@tpu.ru

Предложена модель представления системного описания структуры и семантики данных информационной базы единой информационной среды вуза на примере ТПУ, предназначенная для функционирования универсальных процедурно-ориентированных средств по работе с объектами базы данных в терминах предметной области.

Ключевые слова:

Структурные метаданные, семантические метаданные, универсальные процедурно-ориентированные средства, словарь данных, единая информационная среда.

Key words:

Structural metadata, semantic metadata, universal procedure-oriented tools, data dictionary, integrated information environment.

В сфере информатизации деятельности вузов в течение последних десятилетий доминирует подход, предполагающий ввод в действие информационно-программных комплексов (ИПК) для функциональных подразделений («Кадры сотрудников», «Абитуриент», «Аспирантура», «Планирование учебного процесса» и др.). С расширением состава ИПК и накоплением информации в базах данных возрастает потребность в доступе к информации различных подсистем, прежде всего, у руководителей различных уровней. Однако, каждый ИПК, как правило, имеет ограниченный набор процедур и специфику интерфейса, что затрудняет или делает невозможным удовлетворение упомянутых информационных потребностей пользователей. Кроме того, в процессе эксплуатации специализированных подсистем появляется потребность в формировании новых видов отчетов (запросов на чтение), которые могут иметь разовый характер, для автоматизации формирования которых модификация существующих ИПК часто оказывается нерациональной.

Перечисленные выше факторы делают актуальным использование наряду со специализированными предметно-ориентированными ИПК универсальных процедурно-ориентированных инструментов, которые обеспечивали бы единообразный пользовательский интерфейс, автоматически настраиваемый на любые таблицы информационной базы (ИБ), и типовый набор стандартных функций (навигация, чтение, поиск и селекция, агрегация, добавление, изменение и удаление данных). Универсальность в значительной степени обеспечивают любые программы-оболочки, которые используют метаданные словаря баз данных (БД) и реализуют интерфейс для построения SQL-запросов. Однако все эти средства ориентированы на профессиональных разработчиков, требуют специальной квалификации и знания структуры БД и поэтому плохо подходят для использования конечными пользователями. В соответствии с концепцией единой информационной среды (ЕИС) ТПУ создается универсальная информационно-справочная система, предназначенная для отображения содер-

жимого ИБ ЕИС, и универсальная система актуализации данных — для редактирования содержимого ИБ.

Необходимым условием функционирования любых универсальных средств является наличие специального системного описания структуры и семантики данных ИБ ЕИС, представленного в виде метаданных. Основным положением, определяющим структуру системного описания, является идея о том, что конечные пользователи должны работать не с множеством таблиц и столбцов БД, смысл и назначение которых зачастую известны только разработчикам, а с множеством взаимосвязанных типов объектов предметной области. То есть пользователи при работе с универсальными средствами должны иметь возможность использовать хорошо понятные им русскоязычные термины, обозначающие объекты предметной области, их свойства и связи.

Существуют естественные источники готовых метаданных, такие как репозиторий Oracle Designer (структура ИБ ЕИС ТПУ разрабатывается с использованием Oracle Designer) и словарь данных БД Oracle (ИБ ЕИС представляет собой БД под управлением сервера Oracle). Так в словаре данных, в котором представлены описания всех объектов БД (таблиц, атрибутов, ограничений целостности и др.), единственным источником хранения семантического описания этих объектов являются комментарии. Репозиторий Oracle Designer предоставляет значительно более богатые возможности по представлению семантики данных, что обусловлено возможностью построения ER-диаграмм с помощью Entity Relationship Diagrammer, входящего в состав Oracle Designer. Фактически описательные возможности обеспечены и ограничены возможностями модели данных Barker ER, лежащей в основе этого инструмента [1]. Несмотря на все положительные свойства Barker ER как языка концептуального описания предметной области и проектирования БД, эта модель данных обладает рядом ограничений с точки зрения ее использования «как есть» в качестве языка определения семантических метаданных в рамках системного описания ИБ ЕИС:

- ограниченные возможности описания категориальных отношений (только супертип-подтип, причем подтип может иметь не более одного супертипа);
- недостаточная типизация сущностей и отношений между ними, обусловленная универсальным характером модели.

Недостаточная семантическая выразительность репозитория Oracle Designer и словаря данных обусловила необходимость разработки собственной модели представления системного описания ИБ ЕИС — системного справочника (СС) — и реализующей ее БД метаданных. Важной особенностью этой модели является то, что она ориентирована не на процесс прямого проектирования БД — от концептуального описания предметной области к по-

строению физической структуры, а на обратную задачу — надстраивание семантического слоя над уже существующей БД. В связи с этим СС включает в себя 2 основные части: структурную (физическую) и семантическую.

Структурное описание СС опирается на реляционную модель данных и позволяет описывать структуру таблиц и представлений реляционной БД: их полные системные имена, имена и домены составляющих их атрибутов, декларативные ограничения целостности — т. е. фактически повторяет основные элементы моделей словаря данных [2]. Домен атрибутов рассматривается как множество всех допустимых значений атрибута, а также формируется иерархия доменов в зависимости от указания точности значений; используется разделение доменов на системные (соответствие с типами данных на сервере Oracle) и прикладные (формируются в зависимости от источника заполнения значений, например, из последовательности). Структурные метаданные используются для динамической генерации SQL-запросов к ИБ ЕИС.

Семантическое описание СС используется для динамического построения «дружественного» интерфейса пользователя в универсальных процедурно-ориентированных средствах, опирается на разновидности ER-модели и представляет собой описание сущностей (типов объектов предметной области) и отношений между ними на языке, приближенном к естественному. Под сущностью понимается что-либо, что может быть идентифицировано: объект (реальный или воображаемый), свойство объекта или отношение между объектами. В зависимости от того, какое из выше упомянутых понятий представляет собой сущность, она может быть: а) стержневой (соответствует объекту); б) характеристической (соответствует свойству, некой стороной описания); в) ассоциативной (соответствует связи).

Между сущностями могут быть определены отношения (связи). В отношении каждая из двух сущностей играет роль старшей или зависимой. При этом одной сущности может соответствовать таблица или представление ИБ ЕИС в терминах предметной области, т. е. отношения между сущностями на семантическом уровне представляют собой связи между таблицами в БД на физическом уровне (атрибуты участия сущности в отношении являются атрибутами таблицы, с которой сущность ассоциирована).

Понятие отношения ограничивается простыми отношениями между двумя сущностями (одной из которых может быть ассоциативная сущность, по сути представляющая отношение N:M), поэтому все типы отношений, используемых в СС, сводятся к абстракциям 1:1 и 1:M, даже если изначальная семантика отношения предполагает N:M. Возможные максимальные мощности (1 или M) указываются для каждого типа роли участия сущности в отношении. Каждое отношение относится к одному

из типов отношений, которые соответствуют основным, устоявшимся семантическим категориям отношений между объектами [3].

Далее приведена классификация типов отношений СС (в скобках указаны названия ролей, которые сущности играют в отношении; буквой «с» помечена старшая роль, «з» — зависимая):

- Комментарий (комментирующий объект (с, 1) — комментируемый объект (з, М)) — отношение между двумя стержневыми сущностями или между стержневой (комментирующей) и характеристической или ассоциативной (комментируемой) сущностями.
- Характеристика (объект (с, 1) — характеристика (з, М)) — отношение между стержневой и характеристической сущностями.
- Агрегация (целое (с, 1) — часть (з, М)) — отношение между стержневыми сущностями.
- Иерархия (подчинения, владения) (главный (с, 1) — подчиненный (з, М)) — отношение между стержневыми сущностями.
- Обобщение (супертип (с, 1) — подтип (з, 1)) — отношение между стержневыми сущностями.
- Участие в факте (измерение (с, 1) — факт (з, М)) — отношение между стержневой и ассоциативной сущностями.

Семантические метаданные СС по сути являются частью онтологии вуза и отражают семантику объектов ИБ ЕИС ТПУ [4]. Выявление объектов в виде сущностей и отношений между ними для семантического описания СС осуществляется на основании методики декомпозиции целей и функций, описанной в [5]. Декомпозиционный подход, лежащий в основе этой методики и применяемый в теории системного анализа, позволяет определить структуру предметной области как системы. В соответствии с методикой после формулирования глобальной цели системы определяются основания декомпозиции глобальной цели, которые, как правило, являются неэлементарными, и последующие этапы связаны с дальнейшей декомпозицией полученных объектов с использованием выявленных оснований в качестве соответствующих классификаторов. Для каждой итерации процедуры декомпозиции необходимо определять основные разделы требований и уточнение полученных в результате предыдущих этапов требований с учетом интерпретации оснований декомпозиции.

В связи с тем, что не существует готовых источников представления семантического описания объектов ИБ ЕИС ТПУ, выявление сущностей предметной области производится с помощью описанной выше методики. Ниже приводится пример, иллюстрирующий применение этой методики для определения требований со стороны Министерства образования как системы, оказывающей влияние на вуз, к дипломированным специалистам как конечному продукту (КП) университета.

- Выявление потребности — обеспечение возможности получения высшего образования:
 - Предмет труда:
 - абитуриент, желающий получить высшее образование;
 - документ, подтверждающий полученные им необходимые знания (аттестат, диплом).
 - Средства труда:
 - абитуриент (с помощью своих знаний);
 - денежные средства (для платного обучения).
 - Кадры:
 - приемная комиссия.
- Производство КП — подготовка дипломированного специалиста (обучение населения для получения высшего образования):
 - Предмет труда:
 - студент, получающий высшее образование.
 - Средства труда:
 - преподаватель (с помощью своих знаний);
 - дисциплина;
 - оборудование;
 - учебная литература;
 - аудитория, корпус.
 - Кадры:
 - преподаватель.
- Потребление КП — выпуск специалиста:
 - Предмет труда:
 - специалист, получивший высшее образование;
 - диплом.
 - Средства труда:
 - студент (с помощью своих знаний).
 - Кадры:
 - государственная аттестационная комиссия.

Полученные в результате описания объектов ИБ ЕИС ТПУ и их взаимосвязей представляют собой семантические метаданные СС.

Так как системное описание не является частью словаря данных ИБ ЕИС (и автоматически не актуализируется СУБД), а представляет собой отдельную БД, критическое значение имеет задача поддержания соответствия структурных метаданных СС и словаря данных ИБ ЕИС. Поддержание метаданных в актуальном состоянии осуществляется с помощью приложения по актуализации системного описания, в рамках которой реализуются следующие основные функции:

- автоматическая актуализация системного описания данными из словаря данных ИБ ЕИС и репозитория Oracle;
- прямое редактирование системного описания (вставка, изменение, удаление элементов описания);

- мониторинг актуальности системного описания (проверка соответствия системного описания текущей структуре ИБ ЕИС); собственно устранение выявленных отличий производится в полуавтоматическом режиме (пользователь должен выбрать объект, требующий синхронизации со словарем данных).

Ввиду того, что семантическое описание СС фактически представляет собой надстройку над структурными метаданными, необходимо поддерживать соответствие между этими частями. С одной стороны, семантическое описание должно соответствовать описанию структуры ИБ ЕИС, которая хранится в виде структурных метаданных СС, поэтому сначала происходит актуализация СС из словаря БД, а затем заполняется семантическая часть СС. С другой стороны, в результате семантические описания отсутствуют в текущей структуре ИБ ЕИС в связи с тем, что ИБ ЕИС проектировалась как проблемно-ориентированная БД, а пол-

ное (точное) описание структуры и семантики данных характерно для интегрированной предметно-ориентированной БД. В результате возникает потребность дополнять семантическую часть СС самостоятельно без соответствующего описания в структурной части СС.

Таким образом, представленное в СС системное описание включает концептуальное изложение содержания и смысла информации об объектах ИБ ЕИС ТПУ и позволяет неспециалистам в области ИТ использовать универсальные средства для работы с объектами БД в терминах предметной области. В итоге происходит сокращение расходов на автоматизацию и информационную поддержку простых задач и задач разового характера, а также ускорение процессов ввода в эксплуатацию предметно-ориентированных подсистем, зависящих от данных, для актуализации которых не созданы специализированные средства.

СПИСОК ЛИТЕРАТУРЫ

1. Barker R. Case Method – Entity Relationship Modelling. – Wokingham: Addison Wesley Publishing Co., 1990. – 240 с.
2. Navathe S.B. Evolution of Data Modeling for Databases // Communications of the ACM. – 1992. – V. 35. – № 9. – P. 112–123.
3. Storey V.C. Understanding Semantic Relationships // VLDB Journal. – 1993. – № 2. – P. 455–488.
4. Тузовский А.Ф., Чириков С.В., Ямпольский В.З. Системы управления знаниями (методы и технологии) / под общ. ред. В.З. Ямпольского. – Томск: Изд-во НТЛ, 2005. – 260 с.
5. Основы системного подхода и их приложение к разработке территориальных автоматизированных систем управления / под ред. Ф.И. Перегудова. – Томск: Изд-во Томского ун-та, 1976. – 243 с.

Поступила 06.10.2009 г.