

УДК 81.322(075.8)

ТЕХНОЛОГИЯ ГЕНЕРАЦИИ ОНТОЛОГИИ БАЗЫ ДАННЫХ С ПОМОЩЬЮ ОНТОЛОГИЧЕСКОГО ТОЛКОВОГО СЛОВАРЯ

И.Н. Габдрахманов, В.Н. Кучуганов, Д.В. Медведев, М.Н. Мокроусов, Н.В. Соболева

*Ижевский государственный технический университет им. М.Т. Калашикова
kuchuganov@istu.ru*

Аннотация

В статье рассматриваются возможности использования ранее созданных реляционных баз данных для формирования онтологии предметной области с целью её применения при построении интерфейсов, основанных на знаниях. Разрабатываемая онтология предметной области включает концепты, описывающие свойства, предметы, отношения, процессы, ситуации, сюжеты. За основу сценария построения онтологии предлагается взять отображение реляционной базы данных с использованием известных алгоритмов получения RDF-дампа базы данных и автоматизированного метода извлечения знаний об отношениях между понятиями на основе анализа схемы базы данных с использованием онтологического толкового словаря. Приводится описание базового набора инструментов интерфейса менеджера, и даются примеры применения инструментария.

Ключевые слова: онтология, база данных, извлечение знаний, RDF, R2RML, онтологический словарь.

Введение

Базы данных (БД) современных информационных систем (ИС) хранят огромные объемы фактов, доступ к которым реализуется на основе соответствующего программного обеспечения, использующего технологии транзакционного доступа и технологии аналитической обработки информации (On-line analytical processing - OLAP). В рамках транзакционных систем, как правило, реализуются типовые сценарии обработки информации, обеспечивающие сбор и хранение информации по текущей ситуации в предметной области (ПрО) и поддержку принятия управленческих решений с учётом сложившейся ситуации. Технологии OLAP используются для решения задач прогнозирования и стратегического планирования, позволяют извлекать знания из данных, но не являются инструментом поддержки принятия решений в повседневной практике менеджеров. Модели ПрО и знания, скрытые в базах данных, в случае их явного представления на уровне интерфейса менеджера позволяют решать, как минимум, следующие задачи: принимать нестандартные решения в нестандартных ситуациях; передавать накопленный опыт; объединять знания из разных источников.

Учитывая современные тенденции в разработке интеллектуальных систем, в качестве модели знаний системы поддержки принятия решений целесообразно использовать онтологию ПрО. В качестве первого шага по созданию онтологии рассматривается автоматизированное извлечение знаний из существующих реляционных баз данных.

Предлагаемое решение основано на методах отображения реляционных баз данных в онтологии, которые применяются в рамках концепции семантического веба, и методах “Knowledge Discovery in Databases” (KDD) [1]. В дополнение к известным алгоритмам KDD предлагается сценарий извлечения знаний о ПрО, основанный на интерактивном взаимодействии с экспертом ПрО и методах реинжиниринга реляционных БД. В специализированный интерфейс разработчика онтологии на основе БД встроены онтологический толковый словарь (ОТС) терминов и определений ПрО.

1 Структура онтологии знаний о предметной области базы данных

Рассмотрим задачу извлечения знаний о ПрО из физической модели данных и представления их в онтологии более подробно, для чего определим некоторую целевую метаонтологию, в рамках которой предполагается построить онтологию предметной области на основании существующей реляционной БД.

В структуре метаонтологии [2] рассматриваются следующие категории концептов: свойства; предметы; процессы; отношения; ситуации; сюжеты. Структура описания каждой категории концептов предопределена и представляет собой некоторый набор атрибутов, в том числе системных.

Концепт-свойство описывается именем, типом значения и методом измерения, если он есть. Допустимыми типами атрибутов могут быть как классические типы данных, так и ссылки на концепты или экземпляры разных категорий, что позволяет описывать отношения между понятиями ПрО. Экземпляры концептов данной категории используются в качестве атрибутов концептов других категорий и, по сути, являются пользовательскими типами данных. Концепты-свойства образуют иерархии, которые позволяют судить о мере близости понятий, например, при осуществлении логического вывода на неточных и неполных данных. В терминологии RDF (Resource Description Framework) концепты-свойства соответствуют предикатам.

Концепты-предметы описываются именем, списком собственных атрибутов и могут иметь состав и геометрическую модель. Атрибут предмета наследует описание некоторого экземпляра концепта-свойства и может иметь дополнительные ограничения в виде указания диапазона, значения по умолчанию и уточнённого экземпляра метода вычисления. В RDF концептам-предметам соответствуют субъекты триплетов.

Концепты-процессы имеют имя, собственные атрибуты, среди которых входные и выходные параметры, а также вычислительная модель. В RDF процесс может рассматриваться как субъект. С точки зрения менеджера в категорию «процессы» должны попасть те понятия, атрибуты которых описывают, в общем случае, агента (исполнителя), бенефицианта (заказчика), реципиента (приёмника действия), предмет воздействия, сцену действия, инструменты, коагентов, эффект и т.д. Другими словами, процессы описываются атрибутами, отвечающими на вопросы: «что?»; «как?»; «где?»; «кто?»; «когда?»; «почему?» и т.д. (по аналогии с моделью Захмана [3]).

Концепты-отношения используются для отображения в онтологии бинарных отношений между понятиями предметной области. Структура описания позволяет задать участников отношения и описать атрибуты отношения, в том числе, возможный метод вычисления. В RDF отношениям будут соответствовать предикаты, но с условием, что объект в соответствующем триплете будет иметь значение в виде IRI (Internationalized Resource Identifier).

Ситуация - это совокупность состояний (значений атрибутов) предметов, процессов и отношений в некоторый момент или отрезок времени. Как и в RDF, ситуация описывается графом.

Сюжет - это упорядоченная во времени последовательность ситуаций, имеющих пересекающиеся множества предметов и(или) процессов. Сюжет может быть представлен в виде сценария, сетевого графика, диаграммы Ганта и т.п.

Для построения графовых моделей процессов, ситуаций, сюжетов используется специальный графический язык, основанный на падежной грамматике Ч. Филлмора и идеях других школ и направлений порождающей семантики [4]. Графический язык основан на плексграмматике [5], где символы грамматических конструкций имеют N точек примыкания (рисунок 1).

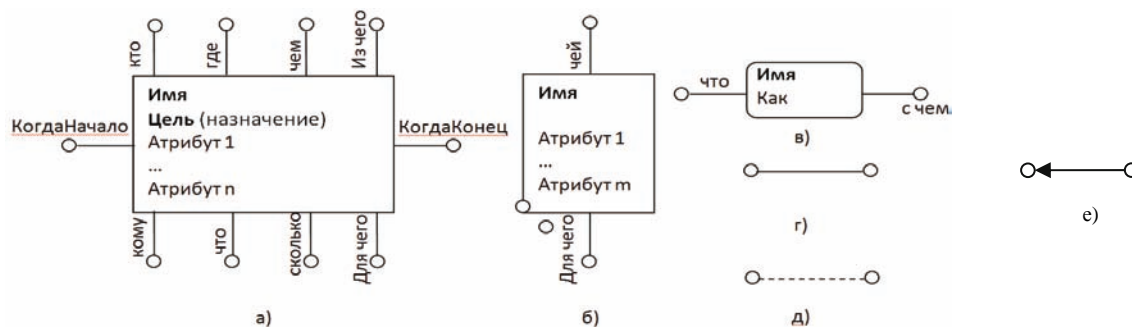


Рисунок 1 – Символы графического языка семантических моделей: а) процесс; б) предмет; в) отношение; г) связь; д) ссылка; е) возврат

Процессы в точках примыкания соединяются с элементами, отвечающими на вопросы: (кто), {из чего}, {чем}, (где), (как), (что), (сколько), (кому), (для чего), (когда), где фигурные скобки обозначают, что несколько символов одного сорта может быть связано с данным символом. Например, для процесса может потребоваться несколько наименований материалов («из чего»), инструментов («чем»).

Предмет характеризуется свойствами, различными по физической природе («какой») и имеет несколько предназначений («для чего») – процессов, в которых он может быть использован.

Свойство связывается с одной сущностью, которую оно характеризует.

Отношение связывает формулой две или более сущности.

Такой графический язык удобен не только в режиме визуального семантического моделирования и контроля корректности знаний, но и для компьютерной интерпретации, поскольку правила плекс-грамматики рекурсивно собирают модель сценария из деталей – основных и производных символов (более подробно показано в [6]).

Для всех категорий концептов поддерживается возможность построения отношений «класс-подкласс» и «целое-часть» без представления этих отношений в явном виде в качестве концепта-отношения.

2 Онтологический толковый словарь

Извлечение знаний из БД должно основываться на понимании семантики ПрО. В качестве первичного источника семантических знаний предлагается использовать онтологический толковый словарь [7].

Онтологический толковый словарь (ОТС) – это толковый словарь, в котором каждое слово или устойчивое словосочетание имеет несколько определений (толкований) таких, что каждое ключевое слово определения содержит гиперссылку на определение в этом же словаре и, возможно, на понятия в онтологической модели ПрО. Для хранения словарных статей, как и в онтологии, используется родовидовая таксономия. Каждое определение в ОТС обладает атрибутами, в некоторых случаях – составом, граф-схемой, пиктограммой (изображением), а также вариантами использования.

Для отображения вариантов толкования используется текстовое описание и, возможно, графовая модель на основе плекс-грамматики (см. рисунок 1).

ОТС содержит понятия ПрО, рассматриваемые с точки зрения их смысла (по аналогии с толковыми словарями естественного языка). ОТС служит для устранения семантической неоднозначности как в ходе общения между специалистами, так и при человеко-машинном диалоге, поскольку представляет собой не просто тезаурус ПрО, но и человеко-машинный

словарь, интерпретирующий термины ограниченного естественного языка в конструкции, используемые для обработки информации.

3 Технология построения онтологии на основе базы данных

БД как источник знаний представляет собой множество структурированных данных, накопленных за определённый период времени, и множество метаданных, содержащих информацию об объектах БД (таблицах, индексах, схеме данных и т.д.).

Структура и состав метаданных определяются особенностями СУБД и теми методами проектирования БД, которые были использованы в каждом конкретном случае.

Идеальная реляционная БД – это нормализованная БД, построенная на основании семантической модели данных ПрО (ER-диаграммы). Семантическая модель данных – это хороший источник для построения онтологии, т.к. содержит знания о сущностях ПрО и отношениях между ними. Проблема в том, что большинство существующих проектов БД не документирует эту часть проекта на естественном для менеджера языке, или в документации к ИС семантическая модель данных вообще отсутствует, а сама БД далека от идеала. В таких условиях решение задачи проектирования онтологии по физической модели требует использования методов интеллектуального анализа, основанных на знаниях технологии проектирования реляционных БД, шаблонах концептов онтологии и знания терминов подмножества естественного языка для рассматриваемой ПрО.

Задачу проектирования онтологии по БД необходимо отличать от задачи отображения БД в онтологию. В первом случае речь идёт о выявлении понятий, которые явно не описываются в физической модели данных, таких как, например, процессы, ситуации и сюжеты. Во втором случае таблицы и связи между ними напрямую переносятся в существующую онтологию, что позволяет видеть схему БД на уровне онтологии и работать с БД из онтологии. В то же время в этих задачах есть общая цель: преобразовать реляционную модель данных в объектную модель.

Отображение БД в онтологию можно использовать в качестве одного из инструментов общего сценария извлечения знаний из БД. В настоящее время наиболее известны две платформы, предназначенные для преобразования реляционных данных в RDF-формат или в OWL-онтологии: D2RQ [8]; Virtuoso [9]. В рамках W3C разрабатывается спецификация языка отображения баз данных “R2RML: RDB to RDF Mapping Language” [10].

Преобразование основано на следующих особенностях реляционных БД:

- как правило, каждый объект ПрО представлен отдельной таблицей БД;
- каждое свойство объекта представлено столбцом таблицы;
- каждый экземпляр объекта идентифицируется первичным ключом и представлен строкой таблицы;
- отношения между объектами задаются с помощью внешних ключей.

Такое преобразование позволяет хранить описание структуры данных в формате онтологии, что обеспечивает доступ к данным из модели знаний, но не позволяет классифицировать понятия ПрО с учётом их семантики. Только эксперт может определить, какому понятию онтологии соответствует та или иная таблица БД или некоторый фрагмент схемы данных.

Решение задачи автоматизированного извлечения семантических понятий из физической модели данных (схемы данных) в онтологию предполагает учёт следующих обстоятельств:

- как правило, отсутствие семантики (с точки зрения ПрО) в используемых идентификаторах;
- отсутствие явных классификаций;
- использование суррогатных первичных ключей;

- бизнес-правила «защиты» в программный код;
- другое.

Задача становится ещё более сложной, если для рассматриваемой ПрО не существует готовой онтологии или её шаблона. Подход, основанный на извлечении знаний из БД, позволяет сформировать некоторый прототип онтологии ПрО, назовём его онтологией на основе БД, который в дальнейшем может быть уточнён. Процедура проектирования такой онтологии предполагает непосредственное участие менеджера, который в этом случае будет рассматриваться как эксперт предметной области.

На рисунке 2 представлены основные этапы проектирования онтологии на основе БД.

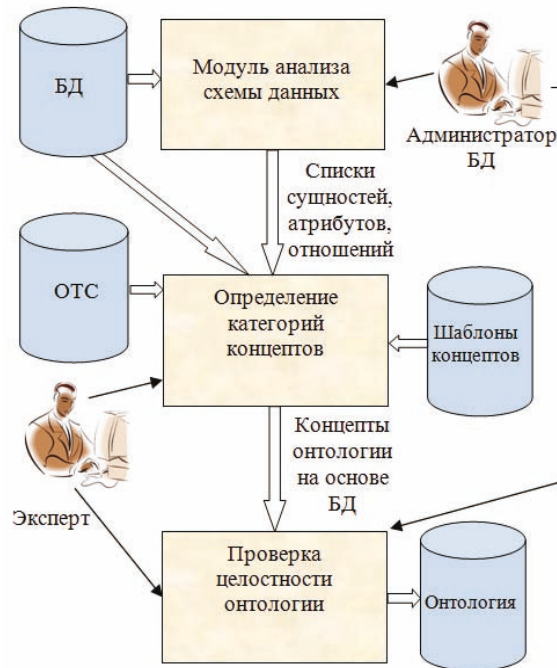


Рисунок 2 – Основные этапы проектирования онтологии на основе базы данных

Формирование списков сущностей, атрибутов и отношений может быть реализовано на основе алгоритмов отображения БД, адаптированных для ситуации отсутствия готовой онтологии ПрО. На этом этапе можно не рассматривать лингвистические аспекты отображения, связанные с именами физических объектов БД.

Рассмотрим более подробно этап определения категорий концептов.

Для решения задачи определения имени понятия, под которым это понятие будет представлено в онтологии, используются автоматизированные методы, основанные на сопоставлении терминов из ОТС и образцов экземпляров понятий, извлекаемых непосредственно из БД по запросу эксперта. При наличии в схеме БД описаний полей таблиц на естественном языке возможен подход [11], основанный на сопоставлении терминов из ОТС и описаний из БД.

Для выявления сущностей, которые соответствуют, например, такой категории концептов как процесс, производится сопоставление шаблонов описания концепта и сущности по следующему сценарию:

- для анализируемой сущности выявляются отношения с другими сущностями (подсхема БД);
- подсхема БД сопоставляется с шаблонами концептов;

- эксперт оценивает соответствие роли атрибута или отношения, приписываемой на основании шаблона, его представлениям о реальных процессах ПрО;
 - в процессе анализа возможна работа с экземплярами данных.
- Рассмотрим фрагмент схемы БД (рисунок 3), предназначенной для хранения учётных данных по поставкам товарно-материальных ценностей (ТМЦ).

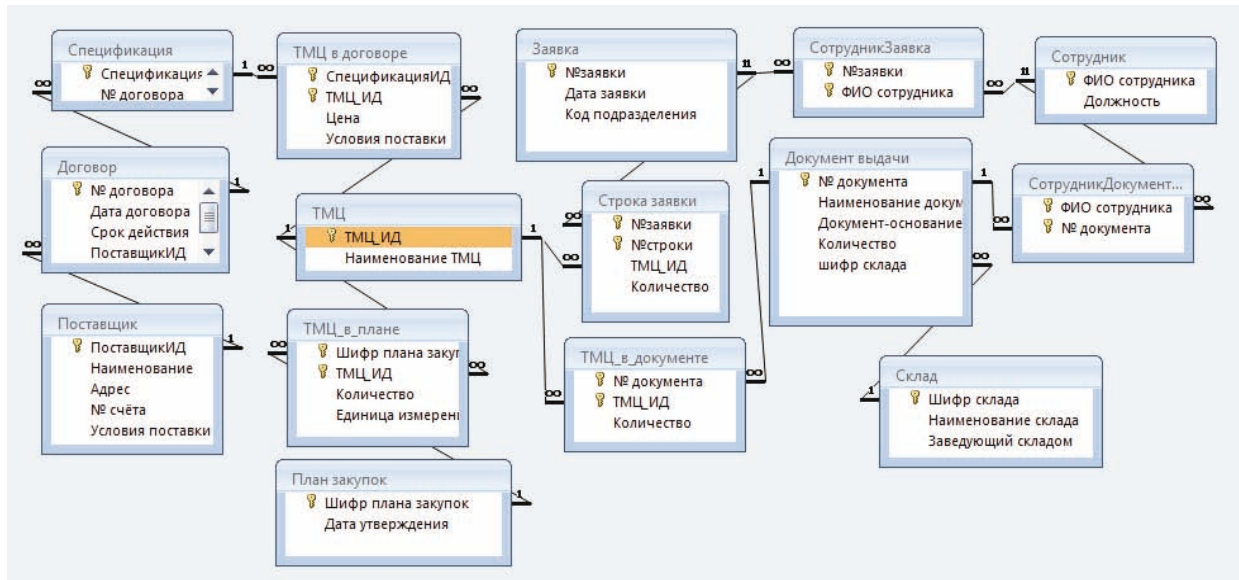


Рисунок 3 – Фрагмент схемы базы данных «Закупка товарно-материальных ценностей»

В результате применения методов обратного проектирования БД, основанных на правилах построения реляционной модели данных, можно восстановить семантическую модель данных в виде сущностей предметной области и бинарных связей между ними. Эксперт может определить смысл связей и дать им соответствующие названия (рисунок 4) с учётом собственных представлений о ПрО и подсказок, извлекаемых из ОТС.

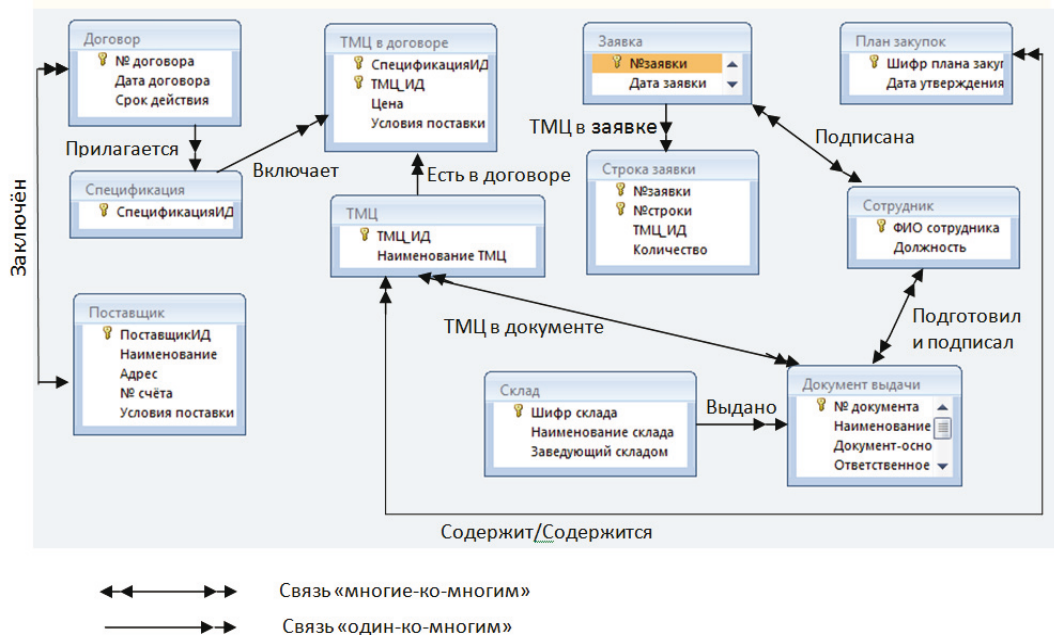


Рисунок 4 – Фрагмент семантической модели данных «Закупка товарно-материальных ценностей»

Следующий шаг предполагает распознавание основных категорий концептов. Наиболее сложной проблемой является выявление процессов, так как они явно в семантической модели данных не представлены.

Каждая словарная статья в ОТС сопровождается соответствующим плекс-символом и ссылками на другие словарные статьи, описывающие те сущности, которые могут быть присоединены в точках примыкания этих символов. В результате, для терминов, обозначающих составные (сложные) работы, могут быть построены семантические модели сценариев (СеМС). На рисунке 5 показан пример СеМС для понятия «Закупка ТМЦ». В качестве терминальных символов здесь использованы символы графического языка (см. рисунок 1) и символы перекрестков из стандарта IDEF3.

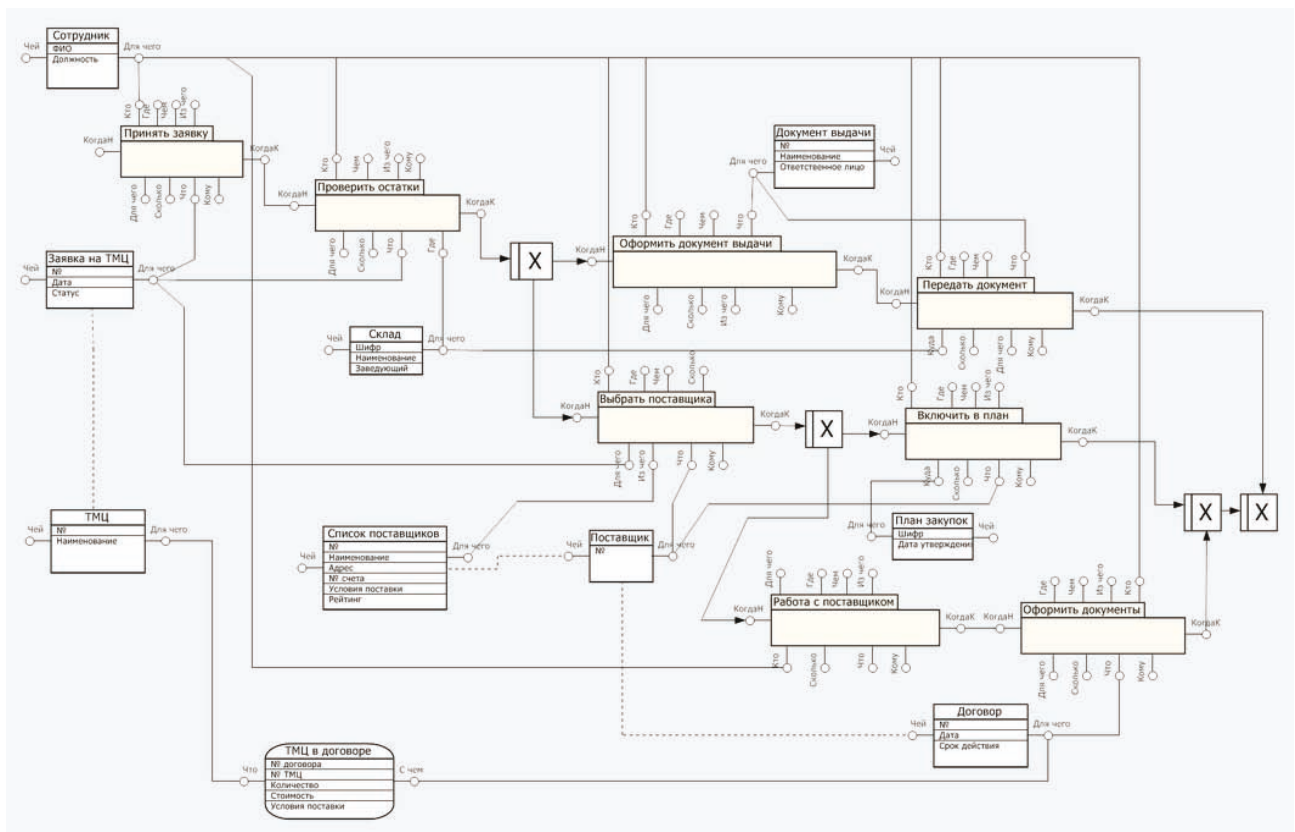


Рисунок 5 – Семантическая модель сценария «Закупка»

ОТС позволяет построить несколько вариантов модели. Предпочтение отдаётся варианту, максимально охватывающему термины, извлекаемые из БД. Окончательное решение о включении соответствующей модели в онтологию принимает эксперт.

Между графическими символами в СеМС и БД устанавливается связь, позволяющая в дальнейшем извлекать экземпляры понятий путём выбора соответствующего понятия непосредственно на схеме.

Помимо правил плекс-грамматики, на которых основаны шаблоны описания процессов, сюжетов и ситуаций, в процессе синтеза онтологии БД используются продукционные правила «ЕСЛИ–ТО», опирающиеся на метаданные БД и принципы проектирования реляционных БД. Например, «Если сущность среди своих атрибутов не содержит внешних ключей, то она может быть отнесена к категории предметов». Часть правил предполагает анализ не только метаданных, но и самих данных, хранимых в таблицах. В частности, использование суррогатных первичных ключей в БД не даёт ответ на вопрос об атрибутах концептов, зна-

чения которых гарантируют уникальность экземпляров концептов. Поиск естественных ключей при отсутствии соответствующих индексов возможен путём анализа текущего состояния БД. Такого рода анализ может потребовать значительных вычислительных ресурсов и гарантирует результат только на текущий момент времени. Аналогичный подход потребуется при автоматическом поиске отношений «класс–подкласс» на сущностях БД. Для решения подобных задач возможен альтернативный сценарий, основанный на непосредственном участии экспертов в принятии соответствующих решений.

Таким образом, автоматизированная система разработчика онтологической модели на основе БД должна включать:

- графический редактор семантических шаблонов;
- модуль логического вывода, осуществляющий автоматическое извлечение сущностей ПрО и связей между ними;
- модуль автоматизированного определения категорий концептов на основе шаблонов концептов и ОТС;
- модуль визуального отображения онтологии.

На рисунке 6 показан браузер системы управления знаниями KG - Knowledge Guide, (Ижевск, ИжГТУ), где можно проверить и, при необходимости, отредактировать полученную онтологию.

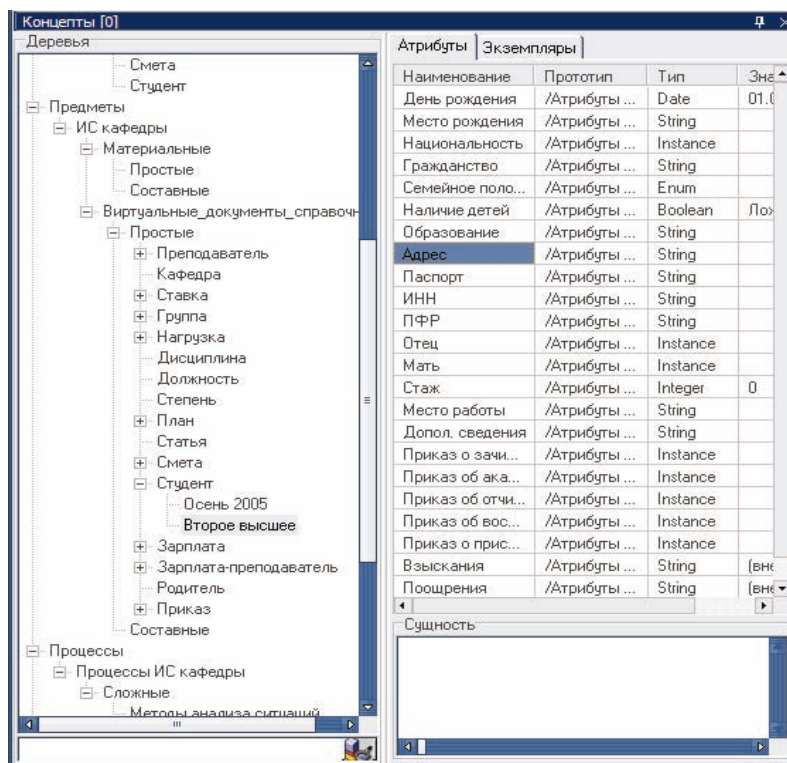


Рисунок 6 – Браузер системы управления знаниями KG

4 Примеры использования онтологии

Используемые для проектирования онтологии инструменты могут быть полезны и на этапе эксплуатации онтологии. Поэтому в состав системы, предоставляющей менеджеру интерфейс, основанный на знаниях, предлагается включить:

- онтологический толковый словарь - ОТС;

- онтологию ПрО;
- библиотеку запросов о состояниях ПрО;
- графический редактор нерегламентированных запросов к источникам данных;
- поисковую машину с использованием технологий гипертекста и гиперграфики;
- модуль логического вывода;
- модуль оценки эффективности решений.

Поисковая машина обеспечивает работу непосредственно с документами. Поисковый запрос позволяет находить фрагменты документа, обращаться к связанным с этим фрагментом статьям ОТС и понятиям онтологии.

Пример интерфейса, основанного на онтологии, представлен на рисунке 7. Здесь менеджеру предоставляется возможность выборки информации из БД с помощью дерева онтологии или путём выделения, например, некоторого помещения непосредственно на планах производственных площадей. Сценарий извлечения данных позволяет использовать различные фильтры.

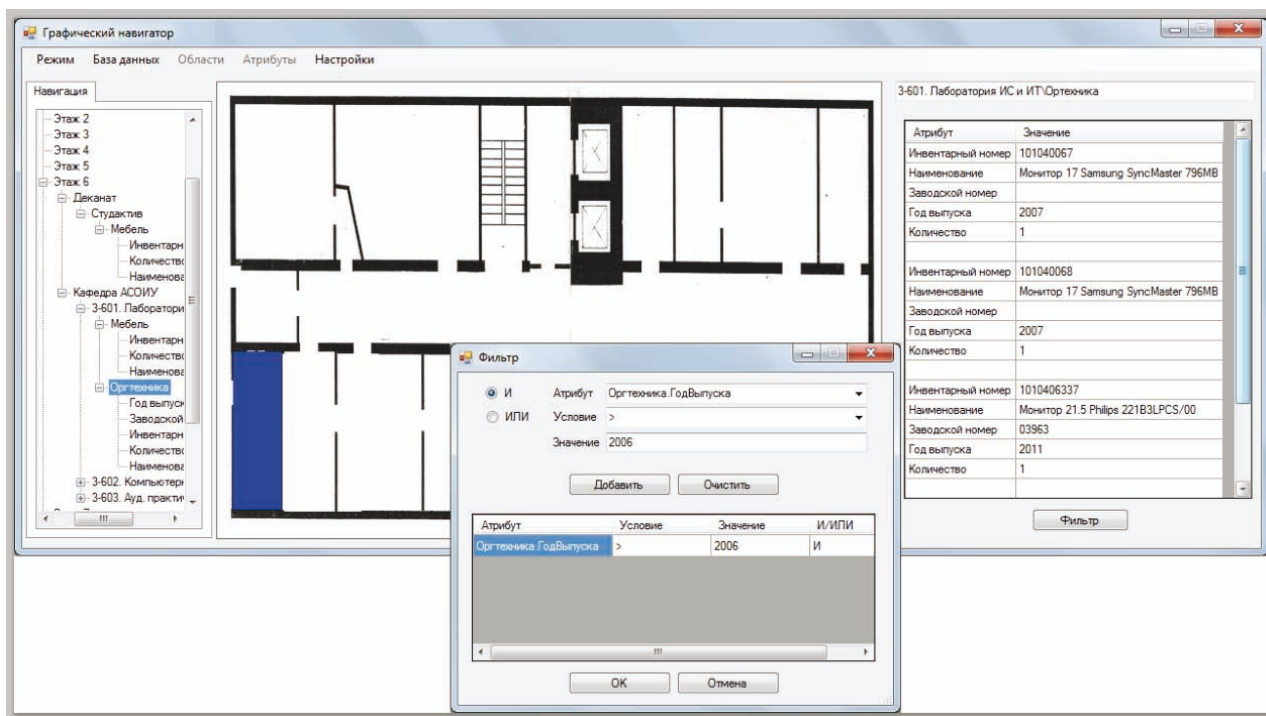


Рисунок 7 – Интерфейс менеджера отдела материально-технического обеспечения

Модуль логического вывода – это часть редактора онтологии, для которой возможны разные варианты реализации в зависимости от того, какие знания и данные будут использованы в качестве средства описания исходной и целевой ситуации. Способ реализации вывода может быть определён из контекста или выбран явно пользователем. В модуле логического вывода предусмотрены следующие варианты решателей задач:

- прямой вызов внешней процедуры для решения поставленной задачи, например, обращение к инструменту OLAP внешнего производителя (сценарий: конвертирование данных и вызов процедуры);
- логический вывод на основе чётких правил продукционного типа;
- нечёткий вывод на правилах;
- «оркестровка» как вариант реализации архитектуры (программные сервисы взаимодействуют друг с другом по заранее определённому сценарию в виде графа решения задачи);

- тренажёр для проведения имитационного моделирования в процессе поиска и анализа решений.

Заключение

Сценарий формирования онтологии, основанный только на системных таблицах БД и самой БД, не позволяет ответить на все вопросы, которые стоят перед разработчиками онтологий в силу того, что любая база данных создаётся с учётом гипотезы закрытости мира [12], согласно которой любая не включённая в базу данных информация считается неверной. В онтологиях же делается предположение о том, что некоторая информация может просто отсутствовать и выводы, которые следуют из сиюминутного снимка базы данных, могут быть просто ошибочными. Таким образом, окончательное решение о структуре онтологии должно приниматься человеком.

Рассмотренный подход основан на использовании инструментов онтологического толкового словаря, которые помогают менеджеру (эксперту) принять решения по структуре онтологии. ОТС играет при этом роль человеко-машинного словаря. С одной стороны, семантическая модель сценария служит менеджеру в качестве когнитивной карты (Mind Map) и инструментария для генерации запросов к базе данных, с другой, благодаря плекс-грамматике графического языка, – система осуществляет контроль корректности получаемой онтологии и ее связей с базой данных.

Список источников

- [1] Fayyad, U. From Data Mining to Knowledge Discovery in Databases / U. Fayyad, G. Piatetsky-Shapiro, P. Smyth // AI Magazine. – 1996. - 17(3). – P. 37-54.
- [2] Кучуганов, В.Н. Вербализация реальности и виртуальности. Ассоциативная семантика / В.Н. Кучуганов // – Искусственный интеллект и принятие решений. - 2011. - №1. – С. 95-106.
- [3] John A. Zachman, J. The Zachman Framework™: The Official Concise Definition / J.A. Zachman // <http://test.zachmaninternational.com/index.php/home-article/13#maincol>. (Дата публикации: 2008; дата обращения: 12.01.2012).
- [4] Chomsky, N. Three Factors in Language Design / N. Chomsky // Linguistic Inquiry - Winter 2005. – Vol. 36, N 1. - P. 1-22.
- [5] Feder, J. Languages of Encoded Line Patterns / J. Feder // Presented at Information and Control. – 1968. - P. 230-244.
- [6] Кучуганов, В.Н. Элементы теории ассоциативной семантики / В.Н. Кучуганов // Управление большими системами. Выпуск 40. – М.: ИПУ РАН, 2012. - С. 30-48.
- [7] Кучуганов, В.Н. Автоматизированная система обучения английскому языку на основе онтологического толкового словаря / В.Н. Кучуганов, М.Н. Мокроусов // Труды Международных научно-технических конференций «Интеллектуальные системы» (AIS'07) и «Интеллектуальные САПР» (CAD-2007). Т. 3. – М.: Физматлит, 2007. – С. 343-350.
- [8] The D2RQ Platform v0.7 - Treating Non-RDF Relational Databases as Virtual RDF Graphs. User Manual and Language Specification // <http://www4.wiwiw.fu-berlin.de/bizer/d2rq/spec/> (Дата публикации: 10.08.2010; дата обращения 8.01.2012).
- [9] Mapping Relational Data to RDFwith Virtuoso's RDF Views // <http://virtuoso.openlinksw.com/whitepapers/relational%20rdf%20views%20mapping.html> (Дата публикации: 22.09.2010; дата обращения 8.01.2012).
- [10] R2RML: RDB to RDF Mapping Language. W3C Working Draft 20 September 2011 // <http://www.w3.org/TR/2011/WD-r2rml-20110920/> (Дата публикации: 20.09.2011; дата обращения: 8.01.2012).
- [11] Биряльцев, Е.В. Онтологии реляционных баз данных. Лингвистический аспект / Е.В. Биряльцев, А.М. Гусенков. // <http://www.dialog-21.ru/digests/dialog2007/materials/html/08.htm> (Дата публикации: 2007; дата обращения: 15.12.2011).
- [12] Левшин, Д. Базы данных в Семантической паутине // Д. Левшин // Открытые системы. - 2009. - №7. - <http://www.osp.ru/os/2009/07/10464695/> (Дата публикации: 2009; дата обращения: 14.12.2011).

Сведения об авторах

Габдрахманов Ильшат Накипович, к.т.н., доцент кафедры «Автоматизированные системы обработки информации и управления» Ижевского государственного технического университета имени М.Т. Калашникова. Область научных интересов: искусственный интеллект, интеллектуальное планирование.

Ishat Nakipovich Gabdrakhmanov, PhD, Associate Professor of the Department of Automated Systems of Information Processing and Control at M.T. Kalashnikov Izhevsk State Technical University. Artificial Intelligence and intelligent planning are area of his interests.



Кучуганов Валерий Никонорович, доктор технических наук, профессор. Заведующий кафедрой «Автоматизированные системы обработки информации и управления» Ижевского государственного технического университета имени М.Т. Калашникова. Область интересов: информационные технологии проектирования, исследования и управления.

Valery Nikonorovich Kuchuganov, DSc, Professor. He is the head of the Department of Automated Systems of Information Processing and Control at M.T. Kalashnikov Izhevsk State Technical University. Research interests: information technologies of design, research and management.



Медведев Дмитрий Валерьевич, 1987 года рождения. Окончил Вятско-Полянский механический техникум в 2007 г. по специальности «Автоматизированные системы управления», в 2012 г. Ижевский государственный технический университет по специальности «Автоматизированные системы управления и обработки информации». В настоящее время учится в магистратуре по программе «Информационно-управляющие системы» направления «Информатика и вычислительная техника».

Dmitry Valerievich Medvedev (b. 1987) graduated from the in Vyatskiye Polyani Technical School in 2007. Major is «Automated Management Systems». Graduated from the Izhevsk State Technical University in 2012. Major is «Automated Information Processing and Management». Now he is getting Master's degree in «Information and Management Systems».



Мокроусов Максим Николаевич, 1983 г. рождения. Окончил Ижевский государственный технический университет в 2006 г., к.т.н. (2010). Доцент кафедры «Автоматизированные системы обработки информации и управления». Области научных интересов: автоматизированная обработка естественного языка, семантический анализ текста, онтологии.

Maksim Nikolaevich Mokrousov, born in 1983. He graduated from the Izhevsk State Technical University in 2006, PhD (2010). He is associate professor at the Department of Automated Systems of Information Processing and Control. Research interests: natural language processing, semantic analysis of the text, ontology.

Соболева Наталья Владимировна, 1955 г. рождения. Окончила Ижевский механический институт в 1976 г. Работает старшим преподавателем кафедры «Автоматизированные системы обработки информации и управления». Области научных интересов: базы данных, онтологии.

Natalia Vladimirovna Soboleva, born in 1955. She graduated from the Izhevsk Mechanical Institute in 1976, a senior lecturer at the Department of Automated Systems of Information Processing and Control. Research interests: a database and ontology.