

# Экспансия онтологий: «онтологически базированные» информационные системы

Л. А. Калиниченко<sup>1</sup>

<sup>1</sup>Институт проблем информатики РАН  
Россия, г. Москва, 117333, ул. Вавилова, 44/2  
leonidk@synth.ipi.ac.ru

**Аннотация.** В статье дан краткий анализ состояния работ в области «онтологически базированных» систем доступа к данным и их возможного влияния на развитие информационных систем и баз данных. Обсуждены вопросы соотношения онтологического и концептуального моделирования и соответствующих языковых средств. Дан анализ развития языков на дескриптивных логиках, ориентированных на использование в качестве средств концептуального моделирования в контексте баз данных и информационных систем. Дан краткий обзор известных экспериментальных результатов создания «онтологически базированных» систем доступа к данным.

**Ключевые слова:** онтологии, онтологически базированные системы, концептуальное моделирование, дескриптивные логики, информационные системы.

## 1 Введение

В информатике под онтологией принято понимать формальное представление знаний в виде множества понятий некоторой предметной области и связей между такими понятиями. Такое представление используется для проведения рассуждений относительно сущностей предметной области, а также для ее описания. Таким образом, онтология представляет собой согласованный в сообществе словарь, который можно использовать для моделирования предметной области. Онтологии используются в различных областях информатики, включая: искусственный интеллект, Семантический Веб, системную инженеррию, инженеррию софтвера, биомедицинскую информатику, библиотечные

науки, информационные архитектуры и др., как форма представления знаний о пространстве дискурса. Современные онтологии обладают большим структурным подобием, независимо от конкретного языка представления. Так, большинство онтологий определяют индивиды (экземпляры), классы (понятия), атрибуты и отношения. Примеры онтологий, разработанных для различных предметных областей (таких, как, например, культурное наследие, лингвистика, науки о Земле, геномика, систематика белков, фармацевтика, структура растений, образование, биология, архитектура предприятий, биомедицина, и др.) можно найти в [http://en.wikipedia.org/wiki/Ontology\\_\(information\\_science\)](http://en.wikipedia.org/wiki/Ontology_(information_science)). Эти примеры, а также развитие разнообразных языков онтологического моделирования свидетельствуют о практической востребованности онтологий для структуризации и представления знаний в различных предметных областях.

Последние годы отмечены активизацией экспансии онтологий в область информационных систем и баз данных. В основе такого распространения лежит простая идея – использовать онтологии как концептуальные схемы баз данных. В последнее время онтологические языки и системы довольно часто используются для представления и поддержки концептуальных схем над (реляционными) базами данных. Такая конструкция позволяет использовать аксиомы концептуальных схем и средства онтологических машин вывода при интерпретации запросов к базам данных. Этот подход практически используется при создании относительно небольших приложений, обычно в научных предметных областях. Подобная деятельность сопровождалась введением ряда новых терминов – таких, как Онтологически Базированные Информационные Системы (Ontology Based Information Systems – OBIS), Онтологически Базированный Доступ к Данным (Ontology Based Data Access – OBDA), Онтологически Базированная Интеграция Данных (Ontology Based Data Integration - OBDI) [6,8,9,12].

Целью настоящей статьи является анализ состояния работ в области «онтологически базированных» систем и их влияния на развитие информационных систем и баз данных.

Статья организована следующим образом. В разделе 2 обсуждаются вопросы соотношения онтологического и концептуального моделирования и соответствующих языковых средств, В разделе 3 дан анализ развития языков на дескриптивных логиках, применяемых или специально разработанных для использования в качестве средств концептуального моделирования в контексте баз данных и информационных систем. В четвертом разделе приводится краткий обзор известных экспериментальных результатов создания «онтологически базированных» систем доступа к данным. В заключении подводится итог выполненному в статье анализу.

## 2 Онтологии как концептуальные схемы

Прежде всего интересно проследить, как происходила эволюция определения понятия «онтология» в информатике. В начале 90-х годов прошлого века под воздействием усилий, направленных на создание стандартов интероперабельности компонентов информационных систем, онтологии были идентифицированы в качестве необходимого компонента систем, основанных на знаниях. Согласно определению Т.Грубера, данному в это время и широко используемому до сих пор, онтология есть спецификация концептуализации [17], или, иными словами, онтология представляет собой формальное определение понятий и их отношений для некоторой предметной области.

Еще в 1998 г. Н.Гуарино дал анализ возможных подходов создания информационных систем (ИС), движимых онтологиями, в котором предначертал варианты использования онтологий в процессе проектирования ИС и в процессе их функционирования [19]. Лейтмотивом этой работы является подчеркивание независимости спецификаций, заданных онтологиями, от каких-либо аспектов реализации. К сожалению, при этом не была прослежена параллель с методами концептуального моделирования, семантически интероперабельных систем, систем интеграции информационных ресурсов и др., в которых непременно вводятся уровни спецификаций, не зависящие от реализации. Следует заметить, что такой стиль первооткрывателей характерен для многих работ в области онтологий.

Показательно, насколько в последнее время развилось понятие «онтология» в сравнении с ранним определением Т.Грубера. Представительной является разделяемая многими точка зрения все того же Т.Грубера [18], которую можно характеризовать следующими высказываниями. В контексте информатики, онтология определяет множество примитивов представления, при помощи которых осуществляется моделирование домена знаний или дискурса. Примитивами представления обычно являются классы (или множества), атрибуты (или свойства), и связи (или отношения) между экземплярами классов. В частности, в контексте систем баз данных онтологии можно рассматривать как уровень абстракции моделей данных<sup>1</sup>, выраженных в иерархических и реляционных моделях, однако ориентированных на моделирование знаний об индивидах, их атрибутах и их связях с другими индивидами<sup>2</sup>. Говорят, что онтологии находятся на «семантическом» уровне, в то время как схемы баз данных являются моделями данных «логического» или «физического» уровня<sup>3</sup>. Ввиду их независимости от моделей данных нижележащих уровней, онтологии используются для

<sup>1</sup> В данном случае имеются в виду схемы баз данных

<sup>2</sup> В контексте разговора об уровне абстракции данные часто превращаются в знания

<sup>3</sup> Т.Грубер намеренно забывает о концептуальных схемах баз данных

интеграции неоднородных баз данных, обеспечения интероперабельности различных систем, а также для спецификации интерфейсов независимых сервисов, основанных на знаниях<sup>1</sup>.

Можно встретить также определения, подобные тому, что недавно было дано Д. Кальванезе [10]. Онтология является схемой представления, определяющей формальную концептуализацию предметной области. При этом спецификация онтологии обычно охватывает два различных уровня. Интенциональный уровень специфицирует множество концептуальных элементов и правил для определения концептуальных структур предметной области<sup>2</sup>. Экстенциональный уровень задает множество экземпляров интенциональных элементов<sup>3</sup>.

Утверждается, что онтологии могут также использоваться как спецификации метауровня для описания модельных категорий, экземплярами которых являются концептуальные элементы. Наряду с анализом расширения содержания самого понятия «онтология» и расширения состава функций, приписываемых онтологиям, интересно также рассмотреть изменение определяемого онтологическим сообществом соотношения онтологических языков и различных традиционных языков информатики.

Собственно, онтологические языки относительно просты. Для представления элементов интенционального уровня нужны средства представления понятий, свойств понятий, связей между понятиями и их свойствами, аксиом. Аксиомы представляют собой логические формулы, которые выражают условия интенционального уровня, которым должны удовлетворять элементы экстенционалов. Представимы такие виды аксиом как выражение отношения «быть подклассом», отношения эквивалентности классов или их композиций, отношения непересечения классов, выражения условий (restrictions) кардинальности, и др.

На экстенциональном уровне представляются индивиды и факты. Индивид, или объект, представляется как экземпляр экстенционала понятия. Факты представляются как связи между экземплярами понятий или свойств.

Нетрудно видеть, что столь простые языки (или их подмножества) естественно отождествлять с подмножествами различных языков, используемых в тех или иных областях информатики. Собственно, такое отождествление нетрудно обнаружить в публикациях онтологического

---

<sup>1</sup> Эта фраза удивительна: как будто не существуют не менее развитые (чем онтологические) языки для интеграции баз данных, концептуального моделирования, обеспечения семантической интероперабельности, спецификации посредников, позволяющие реализовать перечисленные функции

<sup>2</sup> В дедуктивных базах данных этот уровень давно принято называть интенциональной базой данных (IDB)

<sup>3</sup> В дедуктивных базах данных этот уровень давно принято называть экстенциональной базой данных (EDB)

сообщества. Прежде всего, имеется стремление рассматривать онтологии как диаграммы (представляемые в виде семантических сетей, схем модели сущность – связь, диаграмм классов языка UML) [8]. Пример представления онтологии на UML приведен на рис. 1.

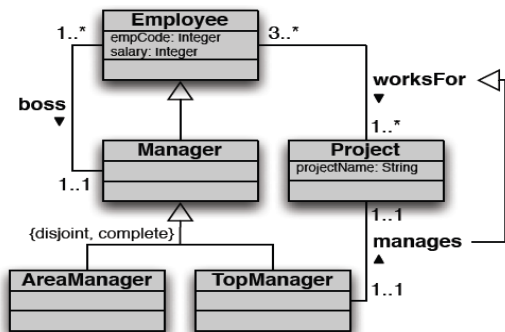


Рис.1. Представление онтологии на UML

Точка зрения онтологического сообщества на соотношение онтологических языков с другими языками выражается следующим образом [10]:

- Онтологические языки по отношению к языкам представления знаний: онтологии являются схемами представления знаний;
- Онтологические языки по отношению к логике: логика это инструмент для придания семантики онтологическим языкам;
- Онтологические языки по отношению к концептуальным моделям данных: концептуальные схемы являются онтологиями специального вида, пригодными для концептуализации конкретной логической модели данных;
- Онтологические языки по отношению к языкам программирования: определения классов в программах являются специальными онтологиями, служащими для концептуализации конкретных структур, используемых в вычислениях.

Эта схема отражает стремление позиционировать онтологические языки среди различных языков информатики, одновременно определяя для последних их место среди онтологических языков<sup>1</sup>. Часто в онтологическом сообществе к классу онтологических языков относят языки, которые никогда таковыми не были. Вот, например, одна из классификаций онтологических языков [10]:

#### 1. Графовые:

- Семантические сети;

<sup>1</sup> Этот список нетрудно продолжить. Скажем, схемы баз данных – это онтологии специального вида для концептуализации структур баз данных, и т.д.

- Концептуальные графы;
  - Диаграммы классов UML, схемы модели сущность – связь;
2. Фреймовые:
    - Системы на фреймах;
    - ОКВС, XOL;
  3. Логические:
    - Deskриптивные логики (например, SHOIQ, DLR, DL-Lite, OWL);
    - Языки правил (например, RuleML, LP/Prolog, F-Logic);
    - Логики первого порядка (KIF);
    - Неклассические логики (например, немонотонные, вероятностные логики).

Попытаемся разобраться, что нового дает такая экспансия онтологической терминологии. Может создаться впечатление, что онтологии открывают дорогу новым направлениям после десятилетий развития и исследований в области концептуальных моделей данных, дедуктивных баз данных, систем интеграции неоднородных баз данных, семантической интероперабельности, и пр. Но ведь в таких областях необходимый теоретический фундамент и конкретные высокоуровневые модели, языки давно созданы. Фактически предпринимается попытка открытия заново уже известных результатов и объявления их достижениями исследований в области онтологий. В действительности же речь идет об исследованиях возможностей онтологических языков в различных областях информатики, в том числе для концептуального моделирования в области информационных систем и баз данных. При этом базы данных, их схемы, информационные системы, концептуальные схемы не перестают оставаться таковыми при использовании тех или иных языковых средств. Что же нового привносят в теорию и практику баз данных и информационных систем исследования в области онтологических языков?

Онтологические модели, рассматриваемые в публикациях, посвященных их использованию в базах данных и информационных системах (БД и ИС), основаны на логике предикатов первого порядка, чаще всего на ее подмножествах – дескриптивных логиках. По существу, в контексте БД и ИС онтологические языки (в частности, языки на дескриптивных логиках) играют роль моделей данных, и не более. Таким образом, следует сосредоточиться на анализе особенностей языков на дескриптивных логиках и новизны, привносимой ими в контекст БД и ИС (что могут дать модели данных на дескриптивных логиках в сравнении с реляционными, объектными и другими моделями данных).

Принципиально важным является соотношение концептуального и онтологического моделирования. Подробное рассмотрение этого вопроса дано в [23]. Здесь достаточно будет ограничиться лишь несколькими основными соображениями.

Концептуальное моделирование реализует абстрактное, семантическое моделирование предметной области (определение классов объектов предметной области, их взаимосвязей, ограничений), независящее от реализации, и служащее в качестве средства порождения эталонной спецификации, отражающей консенсус в сообществе, включающем разработчиков, пользователей ИС, и, собственно, самих ИС.

Концептуальные схемы применяются также в качестве глобальных схем при интеграции информационных ресурсов (баз данных), в процессе проектирования ИС и в процессе интерпретации запросов.

Концептуальная схема определяет структуру предметной области, тогда как онтология должна быть ориентирована главным образом на определение используемых в предметной области понятий. Существенно, что концептуальные схемы БД и ИС, помимо описания классов объектов предметной области и ограничений, содержат описания поведения объектов (методов, функций, процессов), чего онтологии не содержат. Однако, в ряде публикаций Веба онтология буквально отождествляется с концептуальной схемой: "In computer science, an ontology is the attempt to formulate an exhaustive and rigorous conceptual schema within a given domain, a typically hierarchical data structure containing all the relevant entities and their relationships and rules (theorems, regulations) within that domain".

Вместе с тем, онтологическая модель предметной области задает определения понятий, которыми могут быть аннотированы соответствующие определения концептуальной схемы – вот пример того, где онтологии могли бы проявить себя оригинальным образом.

### **3 Развитие языков на дескриптивных логиках в контексте БД и ИС**

В этом разделе дан краткий анализ развития языков на дескриптивных логиках, имеющих отношение к контексту БД и ИС. Этот анализ будет подразделен на две части. В первой рассматриваются языки на дескриптивных логиках, созданные до включения языков на дескриптивных логиках в стек W3C. Во второй части основное внимание будет сосредоточено на соотношении языков на дескриптивных логиках (и других языков) в стеке W3C.

С конца 80-х годов 20 века проводились исследования и разработки в области дескриптивных логик, ориентированных на использование в контексте баз данных. Одним из первых в этом ряду является язык CLASSIC [3], ориентированный на описание структуры объектов базы данных не только в терминах их отношений с другими объектами, но также и интенционально. Язык обладает способностью вывода, например, определения классов, к которым данный объект принадлежит. Цели и

результаты создания языка обсуждались еще в то время в терминах, близких к тем, что можно встретить сегодня в работах по «онтологически базированному» доступу к базам данных. Конечно, при этом не упоминался термин «онтология». Несколько позже стали развиваться языки, ориентированные на поддержку определений понятий предметных областей, на терминологическую интеграцию. К этому ряду систем можно отнести LOOM [24], GRAIL [26], SIMS [2], OIL [13]. Эти языки поддерживают практически все операции формирования понятий. LOOM обеспечивает также поддержку рассуждений над A-Box и T-Box.

Заметной вехой этого периода является появление языков, объединяющих возможности дескриптивных логик с программами на правилах. Язык AL-log [4,12] представляет собой комбинацию простой дескриптивной логики с языком Datalog. Язык CARIN,[14] расширяет дескриптивную логику Хорновскими правилами без функций. Работы по интеграции дескриптивных логик с рассуждениями на правилах выявили необходимость ограничения выразительной способности терминологической части языка для сохранения разрешимости. Интересно, что уже в это время дескриптивные логики использовались в экспериментах по интеграции данных. Так, язык CARIN и дескриптивная логика DLR на n-арных отношениях использовались в работах по интеграции реляционных баз данных [14].

Этот период был отмечен также работами по развитым языкам представления информации, основанных на логике первого порядка, объектных и фреймовых моделях. Такими языками являются онтологический язык Ontolingua [16], язык СИНТЕЗ [20], ориентированный на обеспечение семантической интероперабельности и интеграции неоднородных информационных ресурсов, логико-фреймовый язык F-Logic [21], язык ОКВС [11] для обеспечения интероперабельности баз знаний. В сравнении с языками на дескриптивных логиках, языки на основе логики первого порядка и объектно-фреймовых моделей обладают большим разнообразием возможностей определения типов, классов, логических утверждений.



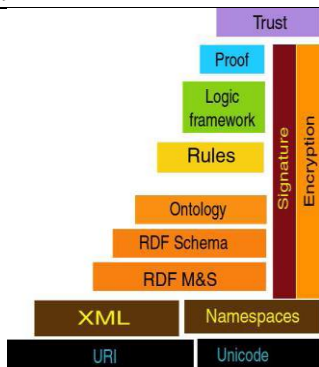


Рис. 2. Одностековая архитектура W3C

По мере развития деятельности W3C в области Семантического Веба все больше появляется работ в этом контексте по онтологическим языкам на дескриптивных логиках. Языки, признанные W3C в качестве технологической основы Семантического Веба, организуются в виде Стэка Семантического Веба (рис. 2). Этот стек определяет иерархию языков, в которой каждый уровень опирается на средства нижележащих уровней. Он также демонстрирует развитие Семантического Веба как расширения классического гипертекстового Веба. Основные языки Семантического Веба, имеющие непосредственное отношение к онтологическим спецификациям, включают: 1) RDF – простой язык для представления данных, ссылающихся на объекты («ресурсы») и их связи. Модель данных, основанная на RDF, представима в синтаксисе XML; 2) язык RDF Schema, расширяющий RDF и позволяющий определять схемы свойств и классов ресурсов, представимых в RDF; 3) семейство языков OWL (OWL 2) для представления онтологий, имеющих формальную семантику, а также основанную на RDF/XML сериализацию.

В последнее время жизнеспособность одностековой архитектуры (SSA), развиваемой W3C, подвергается сомнению. Основное соображение заключается в том, что едва ли (особенно, принимая во внимание младенческий возраст Семантического Веба) единственная иерархия языков окажется достаточной для всех будущих семантических потребностей Веба [22]. Ведь любая технология, включая технологию создания языков, в конце концов устаревает. Кроме того, никакая технология не может справиться со всеми проблемами. Более реалистичным представляется мультитековая архитектура (MSA), в которой множество стеков могло бы сосуществовать бок о бок друг с другом. Так, под влиянием языков правил, включенных в состав стека W3C в результате реализации проекта RIF (Rule Interchange Format),

произошло раздвоение стека (рис. 2) и превращение его в мультистек MSA (рис. 3).

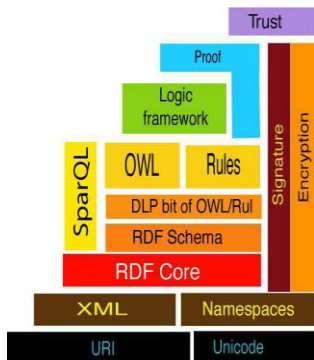


Рис. 3. Мультистековая архитектура W3C

MSA является расширяемым, так что другие стеки могут быть добавлены к нему при необходимости. Каждый уровень стека в мультистеке является синтаксическим и семантическим расширением предыдущего уровня. В этом аспекте интересно посмотреть, что происходит в точке раздвоения стека. Языки правил являются языками логического программирования, исповедующими немонотонные рассуждения в отличие от языков OWL, основанных на дескриптивной логике. Характерным является (неявное) присутствие в MSA (рис. 3) уровня языка программ на дескриптивной логике (Description Logic Programs (DLP)), все операторы которого отображаются в Хорновские правила (подмножество логики первого порядка (рис. 4)) [15].

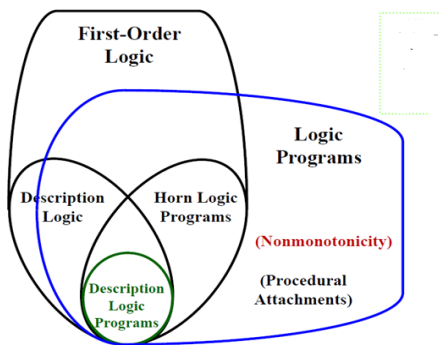


Рис. 4. Соотношение логических языков

Этот уровень должен хотя бы незримо присутствовать в стеке, для обеспечения совместимости в стеке снизу вверх: семантики DLP в стеке OWL и в стеке правил остаются одинаковыми. Вместе с тем, логические программы, основанные на подмножестве логики первого порядка (например, OWL) не поддерживают понятия ограничений целостности и их нарушений. Вместо этого используются условия (restrictions) как утверждения, определяющие желательное состояние среды. Например, если при выводе при условии моногамии оказывается, что у Ивана две жены – Анна и Вера, то OWL приходит к заключению, что Анна и Вера – это одно и то же лицо. В отличие от этого, программы на правилах с немонотонной моделью рассуждений будут трактовать такую ситуацию как противоречие в базе данных.

Более того, имеется мнение, согласно которому утверждение о том, что такие языки на правилах как Datalog являются расширением уровня DLP, неверно. Например, если в базе данных имеется единственный факт *знает(иван, анна)*, то DLP и Datalog дадут разные ответы на запрос, знает ли *иван* в точности одно лицо. Согласно семантике OWL ответ будет “*unknown*”, тогда как согласно семантике языков на правилах ответом будет “*yes*”. Однако в действительности оба ответа верны. Все зависит от того, какой стек выбирает пользователь – стек правил, или стек OWL.

Вместе с тем создание MSA потребует еще значительных усилий. Например, открытым остается вопрос о месте реляционной модели в стеке при рассмотрении языка OWL 2 QL [25]. Можно ли считать этот профиль расширением реляционной модели? Ведь противоречия между ограничениями и условиями (restrictions), о которых речь шла выше, сохраняются. Интересно, что создание профиля OWL 2 QL оказалось возможным благодаря исследованиям, проведенным по семейству дескриптивных логик DL-Lite [7,9].

Целью создания DL-Lite является поддержка базовых средств онтологических языков при сохранении приемлемой сложности рассуждений (в функции рассуждений включаются также ответы на запрос, представляемый как объединение конъюнктивных запросов над экстенциональным уровнем (ABox)). Сложность рассуждений DL-Lite является полиномиальной по размеру TBox, а сложность ответа на запрос оценивается как Logspace от размера ABox.

Существенно, что DL-Lite во время обработки запросов позволяет вести рассуждения в TBox и ABox независимо, так что вначале реализуются рассуждения над TBox, а затем запрос над ABox, может быть реализован с помощью SQL.

Логики, включенные в семейство DL-Lite, являются максимальными дескриптивными логиками, обеспечивающими эффективный ответ на запрос над большим числом экземпляров объектов в базе данных.

Семейство DL-Lite образуется следующим образом (рис. 5).

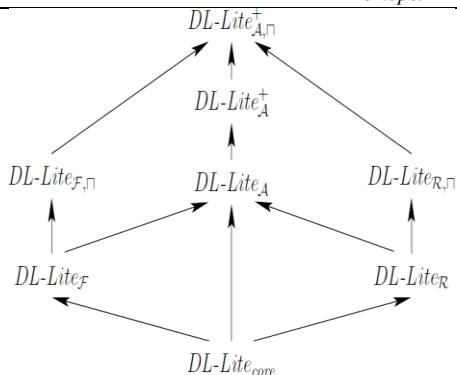


Рис. 5. Семейство дескриптивных логик DL-Lite

$DL-Lite_{core}$  поддерживает утверждения ISA над понятиями, условия непересечения понятий, типизацию ролей, ограничения включения. В  $DL-Lite_F$  добавляются средства выражения условий функциональной зависимости на ролях.  $DL-Lite_R$  расширяется утверждениями ISA над ролями, непересечения ролей. В  $DL-Lite_A$  добавляются возможности совместного использования утверждений включения ролей и утверждений функциональности.

Профиль OWL 2 QL основан на  $DL-Lite_R$  [25]. Существенно, что этот профиль включает всего около трети набора аксиом, поддерживаемых OWL 2 [25]. Вместе с тем, простые дескриптивные логики, подобные  $DL-Lite_R$ , все же годятся для поддержки простых онтологических языков, простых моделей данных (подобных модели Сущность – Связь и подмножеству диаграмм классов UML).

#### 4 Онтологически базированный доступ к данным

«Онтологически базированные» технологии предполагают использование информационных ресурсов на основе концептуализации предметных областей и доступ к данным, опосредованный онтологией (концептуальным взглядом на данные). Собственно, исследовательских работ в этой области известно немного. Применяемые в них подходы рассматриваются в настоящем разделе. Следует сразу заметить, что эти работы применимы только к реляционным базам данных. Собственно, для обеспечения доступа к данным при использовании онтологий требуется 1) определить отображение схем реляционных баз данных в онтологическую (концептуальную) схему; 2) преобразовать конъюнктивный запрос в терминах концептуальной схемы в запрос над реляционной базой данных; 3) реализовать полученный запрос. Рассмотрим последовательно, как это

делается. Рассматриваемый подход вполне приемлем для реализации профиля OWL 2 QL.

Реляционные базы хранят данные, в то время как экземплярами понятий (классов) являются объекты. Основной механизм преодоления такого несоответствия импеданса хорошо известен из области объектных баз данных. При этом уникальный идентификатор объекта продуцируется Сколемовской функцией, параметром которой является значение кортежа базы данных. Отображение реляционной модели данных в концептуальную конструируется как совокупность конъюнктивных запросов над атомарными понятиями, атрибутами, ролевыми атрибутами и соответствующих им запросов на языке SQL. Формально, отображение состоит из утверждений вида  $\phi \rightsquigarrow \psi$ , где  $\phi$  – это произвольный SQL запрос арности  $n > 0$  над базой данных, а  $\psi$  – объединение конъюнктивных запросов арности  $n > 0$  в терминах онтологии  $T$  (например, TBox в DL-Lite<sub>R</sub>).

Для ответа на запрос каждое утверждение отображения  $\phi \rightsquigarrow \psi$  расщепляется на несколько утверждений вида  $\phi \rightsquigarrow p$ , по одному для каждого атома  $p$  в  $\psi$ . Атомы запроса  $q$  унифицируются с атомом правой части отображения, в результате чего образуется объединение конъюнктивных запросов. Затем каждый атом разворачивается, используя левую часть отображения, для получения SQL запроса.

Следует заметить, что перед выполнением описанной выше процедуры запрос дедуктивно расширяется на основе аксиом, заданных в схеме, и машины вывода DL-Lite<sub>R</sub>. Кроме того, согласно онтологической модели, информация о предметной области неполна. Полученные в результате преобразования запросы обеспечивают получение достоверных (certain) ответов, независимо от того, каким образом может быть произведено пополнение неполной информации.

Описанный подход реализован в QuOnto [1] – инструментальном средстве, поддерживающем рассуждения в онтологиях семейства DL-Lite, а также ответы на запросы, представленные в виде UCQ (объединения конъюнктивных запросов). QuOnto допускает подключение внешних реляционных СУБД (таких как Oracle, DB2, IBM Information Integrator, SQL Server, MySQL, и др.).

Вместо реляционной СУБД имеется возможность совместной работы QuOnto с федеративной базой данных – результатом интеграции реляционных баз данных. Такая компоновка QuOnto именуется MASTRO-I [5]. Нужно иметь в виду, что федеративная база данных представлена для QuOnto единственной схемой отношения. Для формирования федеративной базы данных используется подход GAV. Функционирование QuOnto в MASTRO-I ничем не отличается от уже описанного в настоящем разделе.

## 5 Заключение

В статье дан краткий анализ состояния работ в области «онтологически базированных» систем доступа к данным и их возможного влияния на развитие информационных систем и баз данных. Обсуждены вопросы соотношения онтологического и концептуального моделирования и соответствующих языковых средств, Дан анализ развития языков на дескриптивных логиках, ориентированных на использование в качестве средств концептуального моделирования в контексте баз данных и информационных систем. Приведен краткий обзор экспериментальных реализаций «онтологически базированных» систем доступа к данным.

Анализ показывает, что работы по «онтологически базированным» информационным системам скорее воспроизводят некоторые результаты из области дедуктивных баз данных, не добавляя заметной новизны. Фактически исследования в этой области акцентированы на создании онтологических языков для концептуального моделирования, т.е. на создании концептуальных моделей данных на дескриптивных логиках.

Создание семейства дескриптивных логик DL-Lite как максимальных подмножеств средств, обладающих приемлемой для работы с базами данных сложностью, является важным результатом. Отражение этих результатов можно видеть в профиле OWL 2 QL [25], ориентированном на DL-Lite<sub>R</sub>.

С точки зрения выразительности языковых средств для работы с базами данных языки на дескриптивных логиках привносят мало нового. Достаточно заметить, что полученные концептуальные модели на дескриптивных логиках могут быть отображены в существующие модели данных систем интеграции баз данных (например, в язык СИНТЕЗ [20]) с сохранением их семантики. Более того, образованные базы данных со схемами в OWL легко могут быть интегрированы с другими, традиционными, базами данных в системах интеграции баз данных. Например, в системе СИНТЕЗ для интеграции таких OWL-базированных баз данных достаточно создания соответствующего адаптера (отображение OWL в СИНТЕЗ, сохраняющее семантику, построено [27]).

В архитектурном плане работы по «онтологически базированным» системам требуют согласования предположений об открытом мире (OWL) и замкнутом мире (базы данных) на стыке этих сред, а также уточнения функционирования систем при различных трактовках условий (restrictions) в OWL и ограничений целостности в базах данных.

В целом, представляется, что с точки зрения баз данных и информационных систем идея так называемых «онтологически базированных» систем пока не является достаточно продуктивной, так что такие артефакты как «онтологически базированный доступ к данным», или «онтологически базированная интеграция данных» пока являются

лишь отражением терминологической экспансии онтологий в область баз данных и информационных систем. Вместе с тем, выполненные исследования и разработки важны для включения реляционных баз данных в контекст Семантического Веба.

## Литература

1. Acciarri, D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini, M. Palmieri, and R. Rosati. QUONTO: QUerying ONTOlogies. In Proc. of AAAI 2005, pages 1670–1671.
2. Yigal Arens, Chun-Nan Hsu, and Craig A. Knoblock. Query processing in the sims information mediator. In Advanced Planning Technology. AAAI Press, California, USA, 1996.
3. A. Borgida, R. J. Brachman, D. L. McGuinness, L.A. Resnick. CLASSIC: A Structural Data Model for Objects, ACM SIGMOD Record, Volume 18, Issue 2, June 1989
4. Diego Calvanese, Giuseppe De Giacomo, and Maurizio Lenzerini. Description logics for information integration. In Computational Logic: From Logic Programming into the Future, LNCS. Springer-Verlag, 2001.
5. Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Antonella Poggi, Riccardo Rosati. MASTRO-I: Efficient integration of relational data through DL ontologies. In Proceedings of the 2007 International Workshop on Description Logic (DL 2007), CEUR Electronic Workshop Proceedings, 2007
6. Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Antonella Poggi, Riccardo Rosati. Ontology-based database access Proc. of the 15th Italian Conf. on Database Systems (SEBD 2007). 2007.
7. Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Riccardo Rosati. Tractable Reasoning and Efficient Query Answering in Description Logics: The DL-Lite Family. Journal of Automated Reasoning, Volume 39, 2007
8. Diego Calvanese. Ontology-based Data Management Masters Ontology Spring School 2009 September, 2009 <http://ksg-projects.meraka.csir.co.za/krr-projects/events/moss09-1/MOSS-09-OBDM-calvanese-draft.pdf>
9. Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Antonella Poggi, Mariano Rodriguez-Muro, and Riccardo Rosati. Reasoning Ontologies and Databases: The *DL-Lite* Approach. Web 2009, LNCS 5689, pp. 255–356, 2009. Springer-Verlag Berlin Heidelberg 2009
10. Diego Calvanese. Knowledge Bases and Databases. Part 2: Ontology-Based Access to Information. Presentation, 2009,

- 
- <http://www.inf.unibz.it/~calvanese/teaching/08-09-kbdb/lecture-notes/p2-obda-2up.pdf>
11. Vinay K. Chaudhri, Adam Farquhar, Richard Fikes, Peter D. Karp, and James P. Rice. Open knowledge base connectivity (okbc) specification document 2.0.3. Technical report, SRI International and Stanford University (KSL), April 1998
  12. F. Donini, M. Lenzerini, D. Nardi, and A. Schaerf. Al-log: Integrating datalog and description logics. *Journal of Intelligent Information Systems (JIIS)*, 27(1), 1998.
  13. D. Fensel, I. Horrocks, F. Van Harmelen, S. Decker, M. Erdmann, and M. Klein. Oil in a nutshell. In 12th International Conference on Knowledge Engineering and Knowledge Management EKAW2000, France, 2000.
  14. François Goasdoué, V´eronique Lattes, and Marie-Christine Rousset. The use of carin language and algorithms for information integration: The picssel project,. *International Journal of Cooperative Information Systems (IJCIS)*, 9(4):383 – 401, 1999
  15. Benjamin N. Grosz, Ian Horrocks, Raphael Volz, Stefan Decker. Description Logic Programs: Combining Logic Programs with Description Logic. WWW2003, May 20–24, 2003, Budapest, Hungary
  16. Tom Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, 1993
  17. Thomas R. Gruber. Toward Principles for the Design of Ontologies Used for Knowledge Sharing *International Journal Human-Computer Studies* 43, 1995, p.907-928.
  18. T. Gruber (2008). "Ontology". In the *Encyclopedia of Database Systems*, Ling Liu and M. Tamer Özsu (Eds.), Springer-Verlag, 2008.
  19. Nicola Guarino Formal Ontology in Information Systems. Proceedings of FOIS'98, Trento, Italy, 6-8 June 1998. Amsterdam, IOS Press, pp. 3-15.
  20. Калиниченко Л.А. СИНТЕЗ – язык определения, проектирования и программирования сред неоднородных интероперабельных информационных ресурсов. – Москва, ИПИ РАН, 1993, 110 стр.
  21. M. Kifer, G. Lausen, and J. Wu. Logical foundations of object-oriented and frame-based systems. *Journal of the ACM*, 1995
  22. Michael Kifer, Jos de Bruijn, Harold Boley, and Dieter Fensel A Realistic Architecture for the Semantic Web, 2005
  23. Kogalovsky M.R., Kalinichenko L.A. Conceptual and Ontological Modeling in Information Systems. *Programming and Computer Software*. - - Moscow: MAIK Nauka/Interperiodica, Pleiades Publishing Inc., 2009. -- V. 35, N. 5. -- P. 241–256.
  24. Robert M. MacGregor. Using a description classifier to enhance deductive inference. In Proceedings Seventh IEEE Conference on AI Applications, pp. 141–147, 1991
  25. OWL 2 Web Ontology Language: Profiles, W3C, 2009



26. A.L. Rector, S. Bechofer, C.A. Goble, I. Horrocks, W.A. Nowlan, and W.D. Solomon. The grail concept modelling language for medical terminology. *ArtificialIntelligence in Medicine*, 9:139 – 171, 1997.
27. С. А. Ступников, Н. А. Скворцов. Взаимное отображение канонической информационной модели и языка OWL 2 // Труды 12-й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» RCDL'2010. – Казань: Казанский государственный университет, 2010. – С. 392-398
28. H.Wache, T. Voegelé, U. Visser, H. Stuckenschmidt, G. Schuster, H. Neumann and S. Hubner, *Ontology-Based Integration of Information, A Survey of Existing Approaches*, 2001 WS

## **Abstract**

### **Ontological Expansion: Ontologically Based Information Systems**

L.A. Kalinichenko

The paper provides brief analysis of the state of the art in the area of “ontologically based” data access systems and of the possible influence of such works on the development of research in the database and information systems fields. The issues of correlation between ontological and conceptual modeling and between the respective languages are discussed. An analysis of development of the languages based on description logics and oriented on their use as a facility for conceptual modeling in the database context is given in brief. Short overview of experimental results of development of “ontologically based” data access systems is also provided.