

E-mail: rodzina@mopevm.tsure.ru.  
347928, г. Таганрог, пер. Некрасовский, 44.  
Тел.: 8(8634)371-673.  
Кафедра математического обеспечения и применения ЭВМ; старший преподаватель.

**Rodzina Olga Nikolaevna**

Taganrog Institute of Technology – Federal State-Owned Educational Establishment of Higher Vocational Education “Southern Federal University”.

E-mail: rodzina@mopevm.tsure.ru.  
44, Nekrasovskiy, Taganrog, 347928, Russia.  
Phone: 8(8634)371-673.

The Department of Computer Aided Design; associated professor; senior lecturer.

УДК 519.712.2

**С.Н. Щеглов**

**ОНТОЛОГИЧЕСКИЙ ПОДХОД И ЕГО ИСПОЛЬЗОВАНИЕ  
В СИСТЕМАХ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ\***

*В статье рассматривается онтологический подход и его использование в системах представления знаний. Основная роль в описании знаний отводится онтологиям, которые используются при проектировании баз знаний, создании экспертных систем и систем поддержки принятия решений, разработке сред, ориентированных на совместное использование информации несколькими пользователями, и разработке различных поисковых систем. Рассматриваются вопросы, связанные с понятием онтологии, представлены виды онтологий, основные задачи, решаемые с помощью онтологии.*

*Данные; знания; предметная область; естественный язык; сетевые модели; базы общих знаний (common knowledge); экспертные системы; системы управления; интеллектуальные системы.*

**S.N. Shcheglov**

**THE ONTOLOGIC APPROACH AND ITS USE IN SYSTEMS OF  
REPRESENTATION OF KNOWLEDGE**

*In article the ontologic approach and its use in systems of representation of knowledge is considered. The basic role in the description of knowledge is taken away ontologies which are used at designing of knowledge bases, creation of expert systems and systems of support of decision-making, working out of the environments focused on sharing of the information by several users, and working out of various search systems. The questions connected with concept онтологий are considered, kinds онтологий, the primary goals solved with the help ontologies are presented.*

*Data; knowledge; a subject domain; a natural language; network models; bases of the general knowledge (common knowledge); expert systems; control systems; intellectual systems.*

**Введение.** Развитие Интернет привело к появлению значительного количества разрозненных источников и хранилищ информации (баз данных и знаний, отдельных информационных ресурсов и т.п.), характеризующихся различными способами представления, форматами и языками описания информации. Для того, чтобы свести разнообразно представленную информацию к общепонятному виду и

---

\* Работа выполнена при поддержке: РФФИ (грант № 09-07-00318), г/б № 2.1.2.1652.

таким образом организовать ее совместное и многократное использование, позаботились методы стандартизации представления информации.

В настоящее время в сфере информационного менеджмента наметился переход от управления *данными*, характеризующими количественную сторону информационных процессов, к управлению *знаниями*, как качественной составляющей этих процессов. Если данные являются набором разрозненных фактов и значений, представленных в различной форме, то знания – это множество отношений (ограничений, функций, правил) в семантической модели некоторой проблемной области.

Основная роль в описании знаний отводится онтологиям, которые используются при проектировании баз знаний, создании экспертных систем и систем поддержки принятия решений, разработке сред, ориентированных на совместное использование информации несколькими пользователями, и разработке различных поисковых систем. В настоящее время онтологии используются в электронной коммерции [1-3] как средства обмена информацией; в медицине как средства классификации [4] и основы для построения экспертных систем [5]; в планировании и проектировании [6] как средства, предлагающие оптимальные решения для стандартного набора задач, и других областях. В области инженерии знаний онтологии являются средством формирования систем управления знаниями [7]. В области информационных систем онтологии играют важную роль, объединяя основные компоненты любой информационной системы: информационные ресурсы, интерфейс пользователя и прикладные программы [8]. Особенно широкое распространение онтологии получили благодаря развитию распределенных систем искусственного интеллекта, реализуемых на основе технологии агентов.

**Понятие онтологии.** Термин «онтология» в ИИ употребляется в контексте с такими понятиями, как концептуализация знания, модели знаний, системы, основанные на знаниях.

Под концептуализацией понимается процесс перехода от представления ПрО на ОЕЯ (или ЕЯ) к точной спецификации этого описания на некотором формальном языке, ориентированном на компьютерное представление. Концептуализация также трактуется как результат подобного процесса, т.е. описание множества понятий (концептов) ПрО, знаний о них и связях (отношениях) между ними. Онтология – это формально представленные на базе концептуализации знания о ПрО.

Самым распространенным на данный момент является определение, предложенное в [2], согласно которому онтология есть точная (выраженная формальными средствами) спецификация концептуализации. С этой точки зрения каждая БЗ, система, основанная на знаниях, или программный агент явно или неявно базируются на некоторой концептуализации. Множества понятий и отношений между ними отражаются в словаре. Таким образом, считается, что основу онтологии составляют множества представленных в ней терминов.

Формально онтология состоит из терминов, организованных в таксономию, их определений и атрибутов, а также связанных с ними аксиом и правил вывода.

Онтология является не абсолютной (единственной) спецификацией концептуализации ПрО, а зависит от целей ее создания, т.е. задач, при решении которых планируется ее применять. Независимо от вида онтологии она должна включать словарь терминов и некоторые спецификации их значений, что позволяет ограничивать возможные интерпретации терминов и отражать взаимосвязь понятий данной ПрО. При таком подходе онтология похожа на известное понятие *тезауруса*.

Выделяют шесть интерпретаций понятия «онтология» [8]:

- 1) неформальная концептуальная система (представление концептуализации);

- 2) формальный взгляд на семантику;
- 3) спецификация концептуализации;
- 4) представление концептуальной системы через логическую теорию;
- 5) словарь, используемый логической теорией;
- 6) метауровневая спецификация логической теории.

Согласно первой интерпретации онтология является концептуальной системой, которую можно использовать в качестве фундамента некоторой БЗ. Согласно второй интерпретации онтология, на основе которой построена БЗ, служит для формального представления семантики входящих в нее терминов. Согласно третьей интерпретации онтология – это БЗ особого типа, отчуждаемая от ее разработчиков и позволяющая однозначно понимать представленные в ней знания. Интерпретации 4–6 трактуют онтологию как специальную грамматическую систему.

В неформальной трактовке онтология представляет собой описание некоторой ПрО. Это описание состоит из терминов и правил их использования, ограничивающих их значения в этой ПрО. На формальном уровне онтология — это система, состоящая из набора понятий и набора утверждений об этих понятиях, на основе которых можно строить классы, объекты, отношения, функции и теории.

На метауровне онтология является *разновидностью сетевой модели знаний о ПрО*. Эта модель может быть статической или динамической. Во втором случае говорят об онтологии как о модели мира, которая может представлять состояния моделируемой ПрО во времени.

**Виды онтологии.** Основной областью применения онтологии является интеграция информации. Онтологии связывают два существенных аспекта [1-7]:

- ◆ определяют формальную семантику информации, позволяя обработку информации компьютером,
- ◆ определяют семантику реального мира, позволяя, основываясь на общей терминологии, связывать информацию, представленную в виде, требуемом для компьютерной обработки, с информацией, представленной в удобной для восприятия человеком форме.

Для любых двух абстрактных систем, пользующихся одним и тем же словом, не существует гарантии, что они смогут правильно использовать одну и ту же информацию, пока они не примут одну и ту же концептуализацию. Две концептуальные модели систем могут пользоваться пересекающимися онтологиями, в то время, как сами не будут пересекаться. Это означает, что восходящий метод интеграции систем, основанный на интеграции многих частных онтологий, может не работать, особенно, если частные онтологии ориентированы только на концептуальные связи, значимые для конкретного контекста. Следовательно, более удобно опираться на одну онтологию верхнего уровня вместо того, чтобы полагаться на договоренности, основанные на пересечении различных онтологий [1-5].

Однако разработка общей достаточной онтологии верхнего уровня является серьезной проблемой, которую пока решить не удалось [6]. Сейчас принято придерживаться принципа интеграции частных онтологий в более общие, основываясь на разрабатываемых методах стандартизации таких, как специальные языки моделирования и средства трансляции.

Онтологии определенного вида могут использоваться для поддержки нескольких альтернативных теорий, относящихся к одной и той же предметной области:

- ◆ *общие (родовые) онтологии*, затрагивающие общие, фундаментальные аспекты концептуализации такие, как «часть», «причина», «участие», «представление».

- ◆ *промежуточные онтологии*, содержащие общие понятия и отношения, характерные для конкретной предметной области. В идеальном случае, они используются в качестве интерфейса между онтологиями предметных областей и общими онтологиями, но могут выступать как онтологии верхнего уровня для описания знаний конкретной ПрО;
- ◆ *онтологии верхнего уровня*, являющиеся конкретным назначением понятий общих и промежуточных онтологий. Это уникальный модуль, находящийся над онтологиями предметной области, или являющийся самостоятельной, независимой от предметной области теорией;
- ◆ *онтологии предметных областей*, содержащие понятия определенной области знаний или входящих в нее областей;
- ◆ *онтологии задач*, описывающие определенные задачи или деятельность, используя словарь, введенный в общих, промежуточных онтологиях и онтологиях предметных областей;
- ◆ *онтологии-приложения*, являющиеся специализацией онтологий предметных областей и онтологий задач и опирающиеся на определения, характерные для конкретного приложения.

Описание онтологии в терминологии словаря обычно представляется в форме, предлагаемой теорией логики первого порядка, где слова из словаря выступают в качестве унарных или бинарных имен предикатов, называемых соответственно *понятиями* и *отношениями*. В самом простом случае онтология описывает иерархию связанных понятий, в более сложных случаях, добавляются подходящие аксиомы, выражающие связи между понятиями и ограничивающие их интерпретацию [8].

В [8,9] предлагается следующая схема интеграции онтологий (рис. 1):

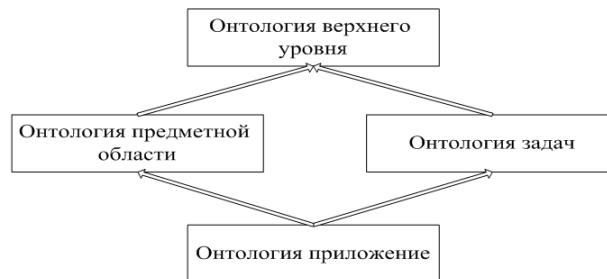


Рис. 1. Многоуровневая схема отношений между онтологиями

Представленные на рисунке (см. рис. 1) онтологии рассматриваются следующим образом.

*Онтологии верхнего уровня* описывают общие понятия такие, как пространство, время, материя, объект, событие, действие и т.п., которые не зависят от конкретной задачи в ПрО.

*Онтологии предметной области* состоят из объектов и связей между ними, описанных в терминологии конкретной предметной области.

*Онтологии задач* ориентированы на решение конкретных проблем и включают все понятия, необходимые для описания процесса логического вывода, от самых абстрактных понятий, относящихся к схеме вывода, до более специальных, характерных для отдельных методов [4].

*Онтологии-приложения* описывают понятия, зависящие, как от конкретной проблемной области, так и от задачи. Эти понятия часто характерны для обеих со-

единяемых онтологий и соответствуют ролям, играемым объектами предметной области при выполнении определенной деятельности.

Онтология предметной области (или ее часть) может образовывать концептуальные связи с несколькими ПрО. Одно и то же множество объектов с одним и тем же набором свойств может использоваться в различных ПрО. Для каждой ПрО характерна определенная модель поведения объектов предметной области. Знания предметной области вместе с моделью, описывающей их поведение в рамках определенных в онтологии задач функций, образуют частную модель ПрО. Программная реализация данной модели является онтологией-приложением. Онтологию-приложение можно рассматривать как усеченную онтологию ПрО применительно к включенным в онтологию-приложение задачам.

**Основные задачи, решаемые с помощью онтологий.** Сфера применения онтологий быстро расширяется. Коротко охарактеризуем основные задачи, в которых опыт применения онтологий насчитывает от 5 до 10 лет.

**1. Создание и использование БЗ.** Практическое формирование БЗ началось с развитием онтологического подхода. Одним из ключевых аспектов технологии БЗ является создание алгоритмических и программных средств построения онтологий и работы с ними.

Онтологии позволяют формировать модели ПрО, интегрируя декларативные описания и определения (в том числе тексты на предметном языке экспертов). Выделяют следующие основные требования экспертов в прикладных областях к средствам построения онтологий [9]:

- ◆ близость языка, которым оперирует система, основанная на знаниях, к профессиональному языку специалиста ПрО;
- ◆ возможность использования введенных знаний для решения большинства (в пределе всех) предметных задач, а не формирование БЗ заново каждый раз для постановки и решения новой задачи;
- ◆ открытость языка, т.е. возможность включения в него новых языковых конструкций, которые появляются в данной ПрО.
- ◆ Со стороны программистов, реализующих НИТ, также предъявляется ряд требований к языкам представления знаний в рамках онтологий:
- ◆ универсальность языка, т.е. возможность представления знаний разного типа независимо от ПрО;
- ◆ унифицированность языковых конструкций, обеспечивающая возможность разным НАС обмениваться знаниями;
- ◆ наличие эффективных алгоритмов разбора языковых выражений.

**2. Организация эффективного поиска в БД, информационных каталогах, БЗ.** Совершенствование механизмов поиска по ключевым словам и формальным языкам запросов не избавляет от высокого уровня информационного шума и неполноты получаемых результатов. Идеальным решением было бы формулирование запросов на ЕЯ. Однако приемлемых методов, реализующих такой подход, в настоящее время нет. Использование онтологий позволяет *точнее интерпретировать смысл терминов*, фигурирующих в запросах, а также *дополнять или расширять запрос* понятиями, которые связаны с терминами запроса отношениями род-вид, синоним, часть-целое, ассоциация и др. Подобное расширение запроса преследует цель уменьшения неполноты ответа на него.

**3. Создание систем, реализующих механизмы рассуждений (ЭС, системы управления, интеллектуальные роботы и др.).** Подобные системы активно разрабатываются в последние годы, и потребность в них быстро растет. Прежде всего, к ним относятся ЭС для технической и особенно медицинской диагностики. Обяза-

тельным компонентом таких систем является блок объяснения решения. Объяснения должны раскрывать суть возникающих ситуаций и их причины, а также обосновывать предлагаемые сценарии действий. Наличие объяснений позволяет человеку действовать осмысленно, а не слепо доверять СПР. И при принятии решения, и при объяснении должна учитываться семантика, как отдельных терминов, так и составленных из них высказываний и их композиций. Достижению данной цели способствует использование онтологии. Реализация средств формирования объяснений на базе онтологического подхода позволяет существенно снизить уровень ошибок, обусловленных человеческим фактором. И чем сложнее система, тем актуальнее такой путь.

**4. Организация поиска по смыслу в текстовой информации.** Текстовая информация до сих пор является основой документооборота. Ее объем очень велик, а задачи поиска и систематизации ответственны и сложны. Механизм индексирования текстовых документов весьма трудоемок. К тому же он не решает проблем неполноты и поискового шума. Это особенно проявляется при использовании запросов типа «где» и «как», а также фактографических запросов.

Для организации поиска по смыслу в текстовой информации необходимы методы извлечения семантики из текстовых документов и запросов и сопоставления получаемых семантических представлений. Подобные методы также повысят эффективность автоматического реферирования, аннотирования и классификации документов. Новыми задачами, связанными с извлечением знаний из текста, являются:

- ◆ формирование сообщений на заданную тему;
- ◆ извлечение новых фактов по интересующей теме;
- ◆ реализация виртуального собеседника.

**5. Семантический поиск в Internet.** Одной из центральных проблем Internet является организация эффективного поиска информации. Онтологии позволяют формировать *информационные профили узлов сети* и на этапе предварительного отбора подходящих для поиска узлов отсеивать нерелевантные узлы.

**6. Представление смысла в метаданных об IP.** Современные языки представления метаданных, как правило, строятся на базе языка XML и модели RDF. В рамках данной задачи онтологии применяются при формировании пространств имен, словарей и квалификаторов для обеспечения их единообразных интерпретаций. Заметим, что без RDF (или другой модели такого рода) конструкциям, основанным на XML, недостает семантической выразительности.

В модели RDF используется объектно-ориентированная система классов. Базовая модель включает три типа объектов:

- ◆ ресурс, идентифицируемый URI;
- ◆ свойство, описывающее ресурс (в таком качестве также может выступать ссылка на другой ресурс или описание другого ресурса);
- ◆ утверждение, состоящее из ресурса-субъекта, свойства (ресурса-объекта) и предиката, связывающего их и представляющего значение свойства.

**7. Построение и использование баз общих знаний (common knowledge) для различных интеллектуальных систем.** Человек в процессе рассуждений использует не только знания, ассоциируемые с данной ПрО, но и знания более высокой степени общности. К таким знаниям относятся описания свойств пространства, времени, личности и т.п. *Знания верхнего уровня (common knowledge)* позволяют доопределять модели конкретных предметных ситуаций с учетом взглядов и роли человека. Эти знания представляются в онтологиях верхнего уровня (общих онтологиях, онтологиях общих знаний).

Степень общности отражаемых знаний служит основанием для выделения *трех уровней онтологии*:

- ◆ общих онтологий;
- ◆ предметных онтологий;
- ◆ онтологий задач.

**8. Обеспечение общей терминологии для множества специалистов и совместно используемых приложений.** Большинство практических задач относятся не к одной, а к нескольким ПрО. Такие задачи, как правило, решаются в рамках совместной деятельности группы специалистов, имеющих разную предметную подготовку. Члены группы могут взаимодействовать друг с другом с помощью телекоммуникационных технологий. Все это требует формирования общих понятийных пространств, обеспечивающих адекватное понимание информации, которой обмениваются специалисты. Онтологический подход существенно упрощает решение данной проблемы. Очевидно, что в данном случае применяются онтологии, относящиеся ко всем трем перечисленным выше уровням.

**9. Многократное применение БЗ и информационных массивов, представляющих сведения о технических системах на различных стадиях их жизненного цикла.** Данная задача приобрела актуальность в связи с развитием CALS-технологий, в рамках которых жизненный цикл технической системы (технического объекта) рассматривается с единых позиций, начиная с момента выявления потребности и до момента прекращения эксплуатации объекта и его утилизации. Каждый этап жизненного цикла должен быть обеспечен соответствующей информационной моделью. Важно, чтобы эта модель не создавалась каждый раз заново, а передавалась с этапа на этап, доопределяясь и развиваясь на каждом этапе.

Определение онтологии как формального представления ПрО, построенного на базе концептуализации, предполагает выделение ее трех взаимосвязанных компонентов: таксономии терминов, описаний смысла терминов, а также правил их использования и обработки. Таким образом, *модель онтологии O* задает тройка [4]:

$$O = (X, R, \Phi),$$

где  $X$  – конечное множество концептов (понятий, терминов) ПрО, которую представляет онтология;  $R$  – конечное множество отношений между концептами;  $\Phi$  – конечное множество функций интерпретации, заданных на концептах и (или) отношениях. Данная модель является разновидностью сетевой модели знаний.

**Заключение.** В заключение следует отметить, что процесс проектирования интеллектуальных систем требует применения различных моделей знаний, так как различные этапы творческой деятельности конструктора не представляется возможным описать одной моделью. Поэтому при создании интеллектуальных систем необходимо применительно к их архитектуре обеспечить эффективное использование всех имеющихся в системе видов знаний, что возможно при построении её в виде гибридной интеллектуальной системы.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Гаврилова Т.А. Хорошевский В.Ф.* Базы знаний интеллектуальных систем. – СПб: Питер, 2000. – 384 с.
2. *Gruber T.R.* A translation approach to portable ontologies // Knowledge. 1993. – № 5(2). – P. 199-220.
3. *Guarino N., Guaretta P.* Ontology's and Knowledge Bases. Towards a logical Calcification // Towards Very Large Knowledge Bases. – Amsterdam: press, 1995.

4. *Гаврилова Т.* Онтология для изучения инженерии знаний // Труды Международной научно – практической конференции KDS-2001, 2001.
5. *Городецкий В.И., Самойлов В.В., Малов О.А.* Современное состояние технологии извлечения знаний из баз и хранилищ данных. Ч.1 // Новости искусственного интеллекта. – 2002, №3. – С. 3-12.
6. *Городецкий В.И., Самойлов В.В., Малов О.А.* Современное состояние технологии извлечения знаний из баз и хранилищ данных. Ч.II // Новости искусственного интеллекта. – 2002, №4. – С. 3-9.
7. *Башмаков А.И., Башмаков И.А.* Механизмы наследования, выявления и разрешения противоречий в обобщенной модели представления предметной области. Часть I // Техническая кибернетика. – 1994, № 5. – С. 14-27.
8. *Башмаков А.И., Башмаков И.А.* Механизмы наследования, выявления и разрешения противоречий в обобщенной модели представления предметной области. Часть II // Теория и системы управления. 1995, №3. –С. 175-189.
9. *Вениаминов Е.М., Машунина М.Ю.* Принципы построения открытого языка шаблонных выражений в системе представления знания // Научно-техническая информация. Сер. 2. Информационные процессы и системы. – 2000, №7. – С. 10-17.

**Щеглов Сергей Николаевич**

Технологический институт федерального государственного образовательного учреждения высшего профессионального образования «Южный федеральный университет» в г. Таганроге  
E-mail: leo@tsure.ru.  
347928, г. Таганрог, пер. Некрасовский, 44.  
Тел.: 8(8634)371-625.  
Кафедра систем автоматизированного проектирования; доцент.

**Shcheglov Sergey Nikolaevich**

Taganrog Institute of Technology – Federal State-Owned Educational Establishment of Higher Vocational Education “Southern Federal University”.  
E-mail: leo@tsure.ru.  
44, Nekrasovskiy, Taganrog, 347928, Russia.  
Phone: 8(8634)371-625.  
Department of Computer Aided Design; associate professor.