

УДК 004.822

ИНСТРУМЕНТАЛЬНЫЕ СРЕДСТВА СОЗДАНИЯ БАЗ ЗНАНИЙ НА ОСНОВЕ СИСТЕМЫ ОНТОЛОГИЙ

Болотова В.А., Григорьев А.В.

Донецкий национальный технический университет г.Донецк

Кафедра прикладной математики и информатики

E-mail: vi4ka20119891@rambler.ru

Аннотация

Болотова В.А., Григорьев А.В. Инструментальные средства создания баз знаний на основе системы онтологии. В настоящее время использование онтологий в сфере искусственного интеллекта значительно расширяется. Своевременное оценивание свойств онтологий важно при их создании и сопровождении. В статье предлагается классификация онтологий и их свойств, которые позволяют производить оценивание в процессе создания онтологий.

1. Общая постановка проблемы.

Сегодня искусственный интеллект – это обширная область исследований и разработок интеллектуальных систем, предназначенных для работы в трудно формализуемых областях деятельности человека.

В настоящее время принято выделять несколько направлений развития искусственного интеллекта: Одно из направлений связано с разработкой интеллектуальных систем, основанных на знаниях. В данном направлении занимаются разработкой моделей представления знаний, созданием баз знаний.

Одним из перспективных применений методов данного направления является создание интеллектуальных обучающих систем (ИОС). ИОС предназначены для автоматизации и индивидуализации обучения.

Традиционно выделяют четыре модели представления знаний, используемых при построении баз знаний систем основанных на знаниях:

1. Логические модели;
2. Продукционные модели;
3. Фреймовые модели;
4. Семантические сети.

Однако на данный момент особый интерес у исследователей искусственного интеллекта вызывают онтологии. Онтологии могут использоваться для представления знаний о понятиях предметной области и предположительных отношениях между ними, для описания содержания веб-страниц. Кроме того онтологии можно использовать при построении баз знаний не только ИОС, но также любых других приложений.

В исследовании онтологий вопросы возникают с первых шагов. Так, до сих пор не существует единого определения для понятия онтология. Само понятие *онтология* происходит от др.-греч. «онтос» – сущее, «логос» – учение, понятие, т.е. это раздел философии, изучающий бытие.

Википедия определяет *онтологию* (в информатике) как попытку всеобъемлющей и детальной формализации некоторой области знаний с помощью концептуальной схемы [1]. Под концептуальной схемой подразумевается *набор понятий + информация о понятиях* (свойства, отношения, ограничения, аксиомы и утверждения о понятиях, необходимых для описания процессов решения задач в избранной предметной области).

Среди специалистов, занимающихся проблемами компьютерной лингвистики, наиболее устоявшимся (классическим) считается определение онтологии, данное Губертом: «*онтология* – это спецификация концептуализации» [2]. Так же известен ряд расширенных определений Губерта, среди которых можно выделить такие:

1. *Онтология* – это эксплицитная спецификация концептуализации, где в качестве концептуализации выступает описание множества объектов предметной области и связей между ними;
2. *Онтология* – это знания, формально представленные на базе концептуализации. Формально онтология состоит из терминов, организованных в таксономию, их определений и атрибутов, а также связанных с ними аксиом и правил вывода;
3. *Онтология* – формальная спецификация разделяемой концептуализации, которая имеет место в некотором контексте предметной области;
4. *Онтология* – база знаний, описывающая факты, которые предполагаются всегда истинными в рамках определенного общества на основе общепринятого смысла используемого словаря.

На сегодняшний день под онтологией можно понимать:

1. надежный семантический базис в определении содержания;
2. общую логическую теорию, которая состоит из словаря и набора утверждений на некотором языке логики;
3. основу для коммуникации между людьми и компьютерными агентами.

Таким образом, в настоящее время не существует общепринятого определения.

Считается, что онтология представляет собой базу знаний, описывающую факты, которые предполагаются всегда истинными в рамках определенного сообщества на основе общепринятого смысла используемого словаря. База знаний же может описывать факты и утверждения, истинность которых зависит от состояния переменных внешней среды. В данном вопросе пока нет полной ясности.

Цель работы.

Для систематического изучения свойств онтологий и их применения на практике необходима такая классификация свойств онтологий, которая охватывает свойства самых разных онтологий, не ограничивая исследователей в обнаружении новых свойств. Поэтому целью работы является разработка классификации известных из литературы свойств онтологий, удобной для систематического оценивания онтологий на практике.

Такая классификация необходима для выбора наиболее полного набора свойств при оценивании произвольной онтологии в зависимости от «типа» онтологии и от преследуемых целей. Такая классификация может быть положена в основу единого подхода к оцениванию и определения в однозначных терминах свойств онтологий, а затем и к объективному оцениванию самих онтологий.

2. Постановка задач исследования.

Постановка задачи.

Работа предполагает прояснение ситуации с определением понятий «онтология» и «свойство онтологии», анализ существующих классификаций свойств и обсуждение возможности их использования для систематического оценивания онтологий и их моделей на практике. Представить классификацию свойств, нацеленную на охват и упорядочение всех известных из литературы свойств, а также на «накопление» ранее не обсуждавшихся, но присущих онтологиям свойств. С учетом наличия в структуре онтологий различных видов связей детализировать классификацию различных структурных свойств.

Проблематика онтологий.

В настоящее время актуальной является задача формирования концептуальных «прозрачных» представлений для слабо-структурированных предметных областей. Это

связано с информационным бумом и лавинообразным ростом «пустой» или зашумленной информации.

Сегодня ведущей парадигмой структурирования информационных потоков являются онтологии, или иерархические концептуальные структуры, которые формируются аналитиком на основе изучения и структурирования потоков информации, документов, протоколов извлеченных знаний и других источников.

Онтологический инжиниринг (ОИ) развивает основные положения инженерии знаний – науки о моделях и методах извлечения, структурирования и формализации знаний.

Собственно инженерия знаний – это ветвь искусственного интеллекта, в то время как ОИ охватывает более широкий круг приложений - от систем управления знаниями до дистанционного обучения.

Онтологический инжиниринг делает первые шаги, поэтому каждый аналитик идет методом проб и ошибок, создавая сложнейшие онтологические структуры, отражающие лабиринты профессиональных знаний в различных прикладных областях.

Система VpWin есть средство создания ряда предметно-ориентированных онтологий, отличающихся графическим интерфейсом и типом используемых отношений.

Потребности в разработке онтологий.

- для совместного использования людьми или программными агентами общего понимания структуры информации;
 - для возможности повторного использования знаний в предметной области;
 - для того чтобы сделать допущения в предметной области явными;
 - для отделения знаний в предметной области от оперативных знаний;
 - для анализа знаний в предметной области.
- другими словами онтологии позволяют представить новые понятия так, что они становятся пригодными для машинной обработки. С помощью них можно «перекинуть мостик» между новыми понятиями, с которыми система еще не встречалась, и описаниями уже известных классов, отношений, свойств и объектов реального мира.

Основные вопросы, которые подлежат решению.

1. Разработка проблемно-ориентированного графического редактора по разработке системы онтологии, что отвечает многоуровневой модели пространства и времени.
2. Разработка системы ввода семантических продукций над множеством онтологий.

Критерии построения онтологий.

Цель создания онтологий – обеспечить поддержку деятельности по накоплению, разделению и повторному использованию знаний. Исходя из этой цели, в [8] введен ряд критериев, которым должна отвечать онтология:

1. прозрачность – онтология должна быть связной, т.е она должна позволять делать выводы, которые согласуются с исходными определениями понятий. По крайней мере, определяемые аксиомы должны быть логически согласованы между собой.
2. расширяемость – онтология должна быть разработана с возможностью использования разделяемого и пополняемого словаря.
3. независимость от синтаксиса – концептуализация должна быть специфицирована на уровне знания максимально независимо от представления понятий на уровне символов.
4. минимальный базис при высокой выразительности – онтология должна вводить минимальный базовый набор понятий, но их должно быть достаточно, чтобы моделировать мир в требуемых целях и создавать сложные сцены.

3. Решение задачи и результаты исследований.

На тему «как построить онтологию» существует множество рекомендаций, руководств, советов, подсказок и указаний. Однако до сих пор не предложено общей

методологии построения онтологий. Некоторые шаги в этом направлении были предприняты в следующих работах:

1. IDEF5: язык и метод для описания онтологий [8];
2. статья Грубера, в которой приведено обсуждение принципов описания онтологий [9];
3. статья Скуса, посвященная вопросам согласования, как пути для совместного использования и интеграции онтологий, разработанных разными группами [10];
4. статьи Гомез-Переза, посвященные технологиям совместного использования знаний [11].

Указанные выше работы приводят дальнейшие шаги в развитии методологии, но не предлагают базовой (skeletal) методологии. Попытаемся вывести такую базовую методологию, опираясь на работы [12] и [13]. Такая методология должна включать ограниченное число стадий и быть применимой к построению любой онтологии.

В работе [13] авторы считают, что комплексная базовая методология для разработки онтологий должна включать следующие стадии:

1. Определение цели.
2. Построение онтологии.
3. Зафиксировать онтологию.
4. Закодировать онтологию.
5. Провести интеграцию существующих онтологий.
6. Оценить.
7. Задokumentировать.

Дополнительно для каждой стадии должны быть описаны технология, методика, принципы и рекомендации, например, типы связей между стадиями (рекомендуемый порядок, чередование, входы/выходы).

Выводы

Существуют различные толкования самого понятия онтологии. В данной работе под онтологией понимается структурная спецификация некоторой предметной области, ее концептуальное описание в виде формализованного представления, которое включает словарь терминов предметной области и логические выражения, описывающие взаимосвязи этих понятий.

На волне интереса к онтологиям были созданы инструментальные средства и механизмы, специально ориентированные на широкое применение онтологий в задачах интеллектуального поиска, классификации, выявления несогласованности в данных, моделирования поведения интеллектуальных агентов. Однако даже наличие хорошего инструментального окружения не снимает проблем, связанных с трудностью проектирования и построения самих онтологий, а автоматизация процесса извлечения онтологий, как и в целом, задача извлечения знаний, и по настоящее время не имеют своего эффективного решения. Тем ценнее становятся уже разработанные онтологии и опыт их использования для решения широкого круга задач.

В процессе создания современных интеллектуальных информационных систем зачастую требуется интеграция знаний из разнородных источников и, как следствие, эффективное решение задач, связанных с тиражированием знаний. По-прежнему не имеет своего удовлетворительного решения проблема автоматизации процесса выбора адекватного специфике конкретной проблемной области и принятого в ней стиля рассуждения экспертов средства представления знаний. Поэтому и по сей день актуальны исследования, направленные на разработку такого подхода к представлению и тиражированию знаний, который с одной стороны позволял бы наиболее адекватно учитывать специфику проблемной области, а с другой – представлять и использовать знания в некотором унифицированном виде.

Онтологические модели за время исследований в этой области претерпели значительное развитие. В настоящее время для создания и поддержки онтологий существует целый ряд инструментов, которые помимо общих функций редактирования и просмотра выполняют поддержку документирования онтологий, импорт и экспорт онтологий разных форматов и языков, поддержку графического редактирования, управление библиотеками онтологий и т.д.

Данные инструментарию построения онтологий обладают рядом существенных недостатков. Большинство инструментов хранит свои онтологии в текстовых файлах, что ограничивает размер онтологий, имеют низкую производительность, нужна дополнительная разработка алгоритмов для удобства работы с хранимыми метаданными, обладают избыточностью функций, что затрудняет работу пользователя.

Список литературы

1. Онтология_(информатика) [Электронный ресурс]. Энциклопедия Википедия. – Режим доступа: [http://ru.wikipedia.org/wiki/Онтология_\(информатика\)](http://ru.wikipedia.org/wiki/Онтология_(информатика))
2. Gruber T.R. A translation approach to portable ontologies [Электронный ресурс] / T.R. Gruber // Knowledge Acquisition. 1993. – № 5(2). – С. 199-220. – Режим доступа : <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>
3. Gruber T. Toward Principles for the Design of Ontologies // International Workshop on Formal Ontology, March. 1993. Padova. Italy.
4. Ясницкий Л. Н. Введение в искусственный интеллект. – М.: Издательский центр «Академия», 2005. – 176 с.
5. Клещев А. С., Артемьева И. Л.. Математические модели онтологий предметных областей. Часть 1. Существующие подходы к определению понятия «онтология». // Научно – техническая информация, серия 2 «Информационные процессы и системы», 2001, № 2, с. 20 – 27.
6. Т. А. Гаврилова, В. Ф. Хорошевский. Базы знаний интеллектуальных систем. – С.-Пб.: Питер, 2001
7. Клещев А. С., Артемьева И. Л.. Математические модели онтологии предметной области. Часть 2. Компоненты модели. // Научно – техническая информация, серия 2 «Информационные процессы и системы», 2001, № 3, с. 19 – 28.
8. KBSI. Knowledge based systems incorporated, Technical Report IDEF5 Method Report [Электронный ресурс] AL/HRGA, Wright-Patterson Air Force Base, Ohio, September 1994. – Режим доступа: <http://www.idef.com/pdf/Idef5.pdf>
9. Gruber T. Towards principles for the design of Ontologies used for knowledge sharing // International Journal of Human-Computer Studies. – 1995. – № 43(5/6). – С. 907-928.
10. Skuce D. Conventions for reaching agreement on shared Ontologies / D. Skuce // In Proceeding of the 9th Knowledge Acquisition for Knowledge Based Systems Workshop, 1995 [Электронный ресурс]. – Режим доступа: http://portal.acm.org/ft_gateway.cfm?id=1135849&type=pdf&coll=GUIDE&dl=GUIDE&CFID=43461927&CFTOKEN=47811814.
11. Gomez-Perez A. Evaluation and assessment of knowledge sharing technology // Towards Very Large Knowledge Bases – Knowledge Building and Knowledge Sharing 1995 / A. Gomez-Perez, N. Juristo, J. Pazos ; in N.J. Mars, editor. – Amsterdam : IOS Press, 1995. – P. 289-296.
12. Гладун А.Я. Онтологии в корпоративных системах. Часть I / А.Я. Гладун, Ю.В. Рогущина // Корпоративные системы. – 2006. – № 1. – С. 41-47.
13. Uschold Mike. Towards a Methodology for Building Ontologies [Электронный ресурс]/ Mike Uschold, Martin King. – Режим доступа: <http://www.aiai.ed.ac.uk/project/pub/documents/1995/95-ont-ijcai95-ont-method.ps>