

ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ ОНТОЛОГИЙ И СЕМАНТИЧЕСКИХ СРЕДСТВ В ПРОЦЕССЕ ОБУЧЕНИЯ

Ярослав Игоревич Поршнев,

Надежда Геннадьевич Силич,

аспиранты Санкт-Петербургского государственного университета информационных технологий механики и оптики

В СТАТЬЕ РАССМАТРИВАЕТСЯ ВОПРОС БОЛЕЕ ЭФФЕКТИВНОГО ПОИСКА И ИСПОЛЬЗОВАНИЯ УЧЕБНЫХ МАТЕРИАЛОВ В СИСТЕМАХ ЭЛЕКТРОННОГО ОБУЧЕНИЯ. ОДИН ИЗ ВАРИАНТОВ РЕШЕНИЯ ЭТОГО ВОПРОСА — ИСПОЛЬЗОВАНИЕ ОНТОЛОГИЙ И СРЕДСТВ СЕМАНТИЧЕСКОЙ СЕТИ ПРИ ПРОЕКТИРОВАНИИ СРЕДСТВ ДИСТАНЦИОННОГО ОБУЧЕНИЯ. ИХ ИСПОЛЬЗОВАНИЕ ПОЗВОЛЯЕТ УВЕЛИЧИТЬ РОЛЬ КОМПЬЮТЕРА ПРИ ОЦЕНКЕ И ОБРАБОТКЕ ИНФОРМАЦИИ В ПРОЦЕССЕ ФОРМИРОВАНИЯ ПРОГРАММ ОБУЧЕНИЯ.

Текущее состояние сети Интернет характеризуется слабой структурированностью данных и практически отсутствием их взаимосвязи. Несмотря на наличие всевозможных способов поиска информации, её извлечения и доставки, отыскать нужную информацию с каждым днём становится всё труднее. Современные web-технологии поиска основаны на полнотекстовом поиске. Все поисковые запросы обслуживаются на основе индекса, содержащего некоторые описания вхождений слов известных данной поисковой системе документов. При этом возникают различные проблемы — выбора того, что следует индексировать, обеспечения равноправного индексирования всего информационного пространства, а также решения, в контексте каких поисковых запросов следует выдавать ту или иную информацию.

Следует подчеркнуть, что в настоящее время компьютеры принимают довольно ограниченное участие в формировании и обработке информации в сети Интернет. Функции компьютеров, в основном, сводятся к хранению, отображению и поиску информации. В то же время создание информации, её оценку, классификацию

и актуализацию продолжает выполнять человек. Как включить компьютер в эти процессы?

Если компьютер пока нельзя научить понимать человеческий язык, то нужно использовать язык, который был бы понятен компьютеру. То есть в идеальном варианте вся информация в Интернете должна размещаться на двух языках: на человеческом языке для человека и на компьютерном языке для понимания компьютера. Эта задача должна быть решена в семантической сети. Слово «семантическая» в данном случае означает «осмысленная», «понятная».

Таким образом, семантическая сеть (Semantic Web) — это концепция сети, в которой каждый ресурс на человеческом языке был бы снабжён описанием, понятным компьютеру. Учитывая всё возрастающее использование компьютерных технологий и сети Интернет в процессе обучения, вопрос применения семантической сети в средствах электронного обучения является на данный момент весьма актуальным и привлекает внимание ряда международных исследовательских проектов.

Рассмотрим, в чём может быть полезно использование средств семантической сети и онтологий в процессе дистанционного обучения. Используя эти средства, мы можем улучшить взаимодействие между различными компонентами обучающих систем и более гибко подбирать сценарии обучения для каждого учащегося. В статье рассмотрены некоторые существующие принципы и разработки в данной области и приведены варианты дальнейшего развития.

Одна из проблем в процессе электронного обучения (e-learning) — необходимость выбора подходящих учебных материалов для гибкого создания учебной программы, наиболее подходящей конкретному учащемуся. Эта проблема связана с определением и представлением уже имеющихся у человека знаний и желаемого уровня знаний и их использованием для формулирования личного учебного плана. Существует несколько подходов к решению этих вопросов, наиболее важные из них — использование стандартов (таких, как набор спецификаций электронного обучения от IMS) и использование онтологий и средств семантической сети для описания и классификации предметной области¹.

Семантическая сеть или семантический веб — новая концепция развития сети Интернет, принятая и продвигаемая Консорциумом сети (World Wide Web). Автор идеи — Тим Бернерс-Ли — впервые использовал данный термин в статье в 2001 году. Семантическая сеть — это надстройка над существующей сетью, которая призвана сделать текстовую информацию более понятной для компьютеров. Целью этого проекта является внедрение в Web таких технологий, которые позволят существенно повысить уровень интеграции информации, обеспечить развитую машинную обработку данных, дать возможность выдавать более адекватные ответы на поисковые запросы. Семантическая сеть предполагает наличие у любой информации, находящейся в сети, точного смысла, который нельзя было бы перепутать даже в случае совпадения фраз или слов, встреченных в разных контекстах. Фактически это означает, что любая информация связывается с некоторым неотделимым от неё контекстом².

В семантической сети можно находить и объединять данные из самых различных источников, а также использовать правила логического вывода для оценки ценности и качества найденных источников, преобразовывать результаты в пригодную для анализа форму. Фундаментом семантической сети являются три основные технологии:

- 1) спецификация XML (eXtensible Markup Language), позволяющая определить синтаксис и структуру документов;
- 2) система онтологий, позволяющая определять термины и отношения между ними;
- 3) механизм описания ресурсов RDF (Resource Description Framework), обеспечивающий модель кодирования для значений, определённых в онтологии.

В семантической сети используются также другие технологии и концепции, в частности, универсальные идентификаторы ресурсов URI (Uniform Resource Identifier), цифровые подписи, системы логического вывода. Онтологии определяют термины, используемые для описания и представления знаний той или иной предметной области, в частности — базовых понятий этой области и связей между ними. Можно сказать, что основная цель онтологий заключается в представлении смысла понятий, используемых в конструкциях RDF (фактически, RDF — это язык общения программных систем, работающих в среде Интернет, а онтологии составляют его словарь-тезаурус). Онтология необходима для людей, для приложений систем баз данных

¹ **Todorova C., Stefanov K.** Selection and use of domain ontologies in Learning Networks for Lifelong Competence Development. [Электронный ресурс]: Proceedings of International Workshop in Learning Networks for Lifelong Competence Development, TENCompetence Conference. March 30th-31st, Sofia, Bulgaria Режим доступа: <http://dspace.ou.nl/bitstream/1820/762/1/Paper25.pdf>, свободный.

² **Мальков М.В.** Технологии семантической сети и дистанционное обучение [Электронный ресурс]: Электрон. журн. Образовательные технологии, 2007. Режим доступа: http://www.naukapro.ru/ot2007/1_003.htm, свободный.

³ **Мальков М.В.** Онтологии в учебном процессе [Электронный ресурс]: Электрон. журн. Образовательные технологии, 2007. Режим доступа: http://www.naukapro.ru/ot2007/3_003.htm, свободный.

и различных других информационных систем, которые совместно используют специфическую информацию в какой-либо предметной области.

Онтологии включают доступные для компьютерной обработки определения основных понятий предметной области и связи между ними. Они обеспечивают возможности повторного применения знаний, могут быть использованы для поиска информационных ресурсов в Интернет и управления знаниями в этой среде³.

Онтология определяет общее соглашение о семантике конкретной области и способствует установлению корректных связей между значениями элементов области, тем самым создавая условия для их совместного использования. Онтологии используются для поддержки автоматизированного обмена данными и для интеграции приложений. Механизмы поиска также применяют онтологии для выборки страниц с синтаксически различными, но семантически одинаковыми словами.

Базовая составляющая онтологии определяет классы объектов и взаимодействие между этими классами. Ключевыми здесь являются понятия подкласса, суперкласса и наследования. Например, класс «Студент» — подкласс класса «Человек» (а «Человек» — суперкласс класса «Студент»), поскольку любой объект, являющийся студентом, — человек. Средствами семантического программирования классы определяются как подмножества наследственно-конечной надстройки, обладающие специальными качествами.

Онтологии используются как механизм выражения и распределения знаний для определения общего словарного запаса и поддержки интеллектуальных запросов в разнообразных базах данных. Онтологии и метаданные описывают организацию и содержание ресурсов.

Онтология кодирует объекты и свойства в понятном для компьютера формате. Конечно, за описанием объектов и их свойств должна стоять простая и понятная логика. С другой стороны, эта логика должна иметь строгое определение и корректную семантику, что позволяет делать автоматическую обработку знаний, заложенных в онтологию.

Основные задачи, которые успешно решаются на базе онтологий, включают предоставление знаний для вывода информации, соответствующей запросу пользователя, фильтрацию и классификацию информации, индексирование собранной информации, организацию общей терминологии, которой могут пользоваться для коммуникации программные агенты и пользователи.

Наряду с использованием онтологий в семантической сети ведутся исследования по переходу от клиент-серверной архитектуры к мультиагентным системам при разработке средств дистанционного обучения⁴. Рассмотрим вариант архитектуры системы, которая не зависит от платформ и предоставляет доступ обучаемым в любое время, с любого компьютера без потери важной информации, собранной системой о них в их профилях.

Чертой данной архитектуры является реализация распределённости и персонализации



Архитектура мультиагентной системы дистанционного обучения

⁴ Келеберда И.Н., Лесная Н.С., Репка В.Б. Использование мультиагентного онтологического подхода к созданию распределённых систем дистанционного обучения. [Электронный ресурс]: Образовательные технологии и общество. Электрон. журн. Образовательные технологии и общество, апрель 2004. Режим доступа: http://ifets.ieee.org/russian/depositary/v7_i2/html/3.html, свободный.

с помощью мультиагентного онтологического подхода. Распределённость обеспечивается за счёт программных агентов, территориально рассредоточенных на различных компьютерах. Например, персональные агенты создаются для каждого обучаемого на портале дистанционного обучения, агент-координатор осуществляет управление системой на сервере (на котором хранится текущая информация процесса обучения в виде профилей), а агент обучающих ресурсов — доступ к учебным материалам с компьютеров различных поставщиков образовательных услуг (см. рис.).

Алгоритм функционирования системы при индивидуальном подборе учебного материала с использованием онтологических моделей для работы с профилями обучаемого и обучающего ресурса состоит из следующих этапов:

1. Посредством web-интерфейса производится регистрация пользователя в системе. Персональный агент собирает необходимые данные о пользователе и на основе этих данных и онтологической модели обучаемого создаёт профиль с информацией о человеке, упорядоченной в соответствии с используемыми стандартами. Получившийся профиль сохраняется в системе для дальнейшего использования.
2. Агент-координатор, используя профиль обучаемого и онтологические модели обучающих ресурсов, формирует запрос на получение профилей учебных материалов РР.
3. Агент обучающих ресурсов на основании полученного запроса проводит поиск метаданных учебных ресурсов, формирует и передаёт агенту-координатору профили учебных материалов.
4. Агент-координатор производит анализ на основе пересечения профилей обучаемого и учебных материалов и формирует запрос на получение необходимых учебных ресурсов.

5. Агент обучающих ресурсов в соответствии с запросом формирует множество учебных материалов и передаёт их агенту-координатору.

6. На заключительном этапе агент-координатор передаёт персональному агенту учебные материалы для отображения обучаемому.

Для реализации подобных систем и более успешного использования принципов семантической сети помимо развития средств описания онтологий требуется дальнейшее развитие стандартов и спецификаций в области дистанционного обучения, касающихся единообразного представления информации об учебных материалах и обучаемых. Надо отметить, что развитие принципов семантической сети и спецификаций в области электронного обучения осуществляется не отдельно, а в тесном взаимодействии⁵. Так, при разработке онтологий для улучшения взаимодействия и повторного использования часто принимают во внимание существующие стандарты и спецификации. Сами же онтологии позволяют более чётко формализовать предметную область, что помогает в последующем создавать более проработанные стандарты.

Таким образом, в результате проведённого анализа были выделены возможности использования семантической сети и онтологий в области электронного обучения для создания и более гибкого использования учебных ресурсов при формировании учебных программ, учитывающих персональные характеристики пользователей. Применение подобных принципов при проектировании систем дистанционного обучения может обеспечить более высокую степень взаимодействия систем, более быстрый и точный поиск нужных материалов, их повторное использование. □

⁵ Kravcik M. & Gasevic D. Leveraging the Semantic Web for Adaptive Education. [Электронный ресурс]: Journal of Interactive Media in Education (Adaptation and IMS Learning Design. Special Issue, ed. Daniel Burgos), 2007/06. Режим доступа: <http://dspace.ou.nl/bitstream/1820/1080/1/kravcik-2007-06.pdf>, свободный.