

Рычков А.А., Щелкунов М.Л., Коренберг В.М.
СЕМАНТИЧЕСКАЯ БАЗА ЗНАНИЙ КАК ОСНОВА УЧЕБНО-МЕТОДИЧЕСКОГО КОМПЛЕКСА

lixus@inbox.ru

ГОУ ВПО УГТУ-УПИ

г. Екатеринбург

В статье рассматривается возможность применения семантической базы знаний как основы для создания учебно-методического комплекса.

In this article considered an ability of semantic knowledge base usage as a basis of educational methodic system.

Одна из наиболее серьезных проблем современных электронных учебно-методических комплексов – практически полное отсутствие интеллектуальной обработки данных со стороны компьютера, которая особенно важна для реализации функций поиска и навигации. В настоящее время поиск, как правило, ведется по соответствию описания или текстового содержания указанному в запросе шаблону. В конечном счете, результат во многом зависит от того, насколько качественно пользователь составил запрос системе.

Допустим, пользователь системы вводит в строке запроса «битва под Полтавой». С точки зрения обычной поисковой системы требуется найти полное вхождение заданной строки либо найти все три слова по отдельности в тексте. Кроме того, система может отбросить предлог «под» как часто встречающееся слово. В результате пользователь получит некоторый набор текстов, в которых так или иначе упоминается историческое событие, данные о котором он хочет получить. Далеко не всегда можно гарантировать, что этот результат удовлетворит пользователя.

Семантические системы предлагают совершенно иной подход к организации данных. В настоящее время большинство из существующих реализаций таких систем являются экспериментальными, ведется процесс разработки стандартов в данной области. Кроме того, подавляющее большинство из доступных источников по теории семантических моделей являются зарубежные, поэтому четкой терминологии в данной области (зачастую термины являются неточной калькой с английского языка) нет.

При разработке семантической базы знаний в рамках данной работы использовалась авторская концепция, сходная той, которая применяется в системах на основе языка семантического представления данных в сети Интернет – Ontology Web Language (OWL). В рамках этой концепции создается набор библиотек, каждая из которых содержит некоторое множество объектов, объединенных общей тематикой.

Объекты делятся на материальные и информационные. В роли материальных объектов могут выступать личности, предметы окружающего мира, события и т.д. Информационные объекты – это те объекты, которые содержат описания для материальных объектов. В роли информационного объекта может выступать книга, статья, журнал. Каждый объект обладает некоторым набором

свойств, часть из которых являются идентифицирующими его свойствами (то есть однозначно определяющими его по отношению к другим объектам). Кроме этого, каждый материальный объект содержит ссылки на информационные объекты, в которых находится его описание. Следует отметить, что свойства объекта принадлежат не ему самому, а библиотеке, в которой он находится. Сам объект содержит только список свойств из библиотеки, которые он использует, а также значения этих свойств. Значения могут относиться к следующим типам данных: целочисленный тип, число с плавающей точкой, строка, дата, ссылка на другой объект. В принципе, пользователь может определять свои типы данных в зависимости от поставленных задач.

Между объектами могут существовать следующие виды связей:

1. Обобщение;
2. Использование в качестве значения свойства;
3. Ссылка на объект (материальный объект ссылается на информационный).

При использовании связей типа обобщение объекты, обладающие некоторым общим набором свойств, объединяются в группы, для каждой из которых определяется объект, называемый типовым. Такой объект обладает набором свойств, которые наследуют от него другие объекты, для которых он является типовым. В свою очередь типовые объекты тоже могут объединяться в группы и для них выделяться свои типовые объекты.

При использовании объекта в качестве значения свойства можно выделить несколько вариантов связей, среди которых можно выделить *ассоциацию*, *агрегацию*, и *зависимость*. При использовании отношения *ассоциации* мы можем сказать, что первый объект так или иначе ассоциирован со вторым. Отношение *агрегации* имеет место между несколькими объектами в том случае, если один из объектов представляет собой некоторую сущность, включающую в себя в качестве составных частей другие сущности. Отношение *зависимости* указывает некоторое семантическое отношение между двумя элементами модели или двумя множествами таких элементов, которое не является отношением ассоциации, обобщения или агрегации.

Следует отметить, что семантические связи между объектами могут существовать как в рамках одной библиотеки, так и между различными библиотеками.

За счет связей множество объектов образуют семантическую сеть, которая может быть использована для интеллектуального поиска данных внутри системы. Пользователь вводит в строке запроса «битва под Полтавой». В системе содержится набор тематических библиотек, одна из которых называется, допустим, «История». Библиотека «История» содержит объект «битва под Полтавой», который и возвращает поисковая система.

В качестве идентифицирующих свойств объекта «битва под Полтавой» можно выделить:

1. Дата (8 июля 1709 года);
2. Место (г. Полтава);

3. Участвующие стороны (Российская империя и Швеция);
4. Командующие войсками (Карл XII и Петр I);
5. Силы сторон;
6. Потери каждой стороны;
7. Итоги (победа русских).

Причем в качестве значения свойства «место» выступает не просто строка «Полтава», а ссылка на объект «Полтава» (город). Последний может находиться как в библиотеке «История», так и в другой библиотеке, которая может называться, допустим, «География». То же самое можно сказать и в отношении значений для свойств «участвующие стороны» и «командующие войсками».

Так как приведенным выше набором идентифицирующих свойств может обладать не только объект «битва под Полтавой», но и любой другой объект, описывающий историческое событие – битву, то мы можем выделить для всех таких объектов типовой объект «битва». В свою очередь, вполне очевидно, что типовой объект «битва» является частным случаем для типового объекта «историческое событие», которое имеет меньший набор свойств (например, только дата и место). С точки зрения объектно-ориентированного программирования мы бы могли сказать, что объект «битва под Полтавой» является экземпляром класса «битва», который, в свою очередь, порожден от класса «историческое событие».

Следует отметить, что в семантических системах для названия объекта и свойства может быть несколько синонимов. Объект «битва» может иметь для своего наименования синоним «сражение». В результате если пользователь введет в строке поиска «сражение под Полтавой», то он получит абсолютно идентичный результат, как если бы он ввел в строке поиска «битва под Полтавой». То же самое в случае, если он введет в качестве строки запроса «Полтавская битва». Система «знает», что слово «Полтавская» образовано от названия города «Полтава», поэтому результатом запроса будет все тот же объект «битва под Полтавой».

Благодаря своей организации данных семантические базы знаний идеально подходят в качестве основы для создания учебно-методического комплекса, обеспечивая их мощными возможностями интеллектуального поиска информации. Следует отметить, что в семантическом виде мы можем хранить данные любого характера, вплоть до учета успеваемости обучающихся, так что теоретически на основе семантической системы мы можем построить полноценный комплекс автоматизации учебного процесса.

Существующие в настоящее время технологии семантического представления знаний, как правило, являются экспериментальными, в настоящее время в мире ведется работа по формированию стандартов в данной области.

В рамках данной работы создавалась семантическая база знаний как основа учебно-методического комплекса, главная задача которого – размещение методических материалов в рамках различных дисциплин. Наполнение и управление контентом семантической базы знаний происходит по принципу социальной сети. Разрабатываемая система может работать как в режиме от-

крытого доступа, когда каждый желающий, имеющий подключение к сети Интернет, может стать участником проекта, так и режим закрытого доступа, когда над проектом может работать только ограниченный контингент участников.

По умолчанию каждый пользователь может создавать себе библиотеки, для которых он является администратором. При желании он может делегировать права администратора другому пользователю. Система позволяет организовать чтение и редактирование данных на основе политики групп. Каждый пользователь, желающий организовать совместную работу по наполнению своих библиотек, может создать для этого одну или несколько групп и пригласить для участия в них других участников проекта. Возможные права доступа к библиотеке: чтение, чтение и запись, отсутствие доступа.

Для удобства управления контентом библиотек пользователи могут создавать в соответствии со своими задачами собственные группы библиотек, в которые они могут добавлять свои собственные библиотеки или подключать чужие, если к ним имеется доступ. Кроме этого, он может разрешать другим пользователям подключать и использовать его группы библиотек.

Система обеспечивает два уровня для управления семантическими объектами. На низком уровне пользователи работают непосредственно с семантическими объектами. На высоком уровне им представляется интерфейс, ориентированный на быстрое создание типовых объектов различной тематики. Благодаря наличию высокого уровня управления объектами работать с системой может даже неподготовленный пользователь, который не знаком с семантическими моделями.

Для удобства взаимодействия пользователей существует встроенная система обмена сообщениями.

Основная целевая аудитория системы – преподаватели и учащиеся. Преподаватели выступают в роли основных пользователей, осуществляющих информационное наполнение системы. Учащиеся в этой системе могут выступать как в качестве пассивных участников в режиме «только чтение», так и в качестве полноценных пользователей системы, участвующих в процессе ее информационного наполнения (такой уровень доступа определяется преподавателями).

Техническая основа системы – ASP.NET 2.0 и MS SQL Server 2005 (в будущем планируется поддержка новых версий ASP.NET и SQL Server).

В настоящее время реализован низкий уровень управления данными (редактирование объектов и связей между ними). Реализована система интеллектуального поиска. В дальнейших планах – реализация «высокоуровневого» интерфейса управления данными.

1. Tim Berners-Lee, James Hendler, Ora Lassila. The Semantic Web. Scientific American, May 2007.
2. Miguel Salmeron. Evolution of Knowledge. Scientific American, October 2006.