

ПРОЕКТ ONTOWIKI

И.Л. Кафтаников, Н.В. Пименова

ONTOWIKI PROJECT

I.L. Kaftannikov, N.V. Pimenova

Описан проект OntoWiki, определяющий создание, поддержку и развитие динамически развивающейся неограниченной семантической сети представлений и свойств онтологического описания предметных областей. Приведено формальное описание онтологической семантической сети OntoWiki. Описано визуальное отображение OntoWiki в виде сети или дерева.

Ключевые слова: онтологическая семантическая сеть, граф, объект, слот, визуализация.

OntoWiki project identifying the creation, support and development of dynamically developing unlimited semantic network of images and properties of ontological description of subject areas is described in the article. Formal description of ontological semantic network OntoWiki is given. Visual image of OntoWiki in the form of a net or a tree is described.

Keywords: ontological semantic network, graphic chart, object, slot, visualization.

Проект OntoWiki определяет создание, поддержку и развитие динамически развивающейся неограниченной семантической сети представлений и свойств онтологического описания предметных областей, основанной на принципах создания Wiki-энциклопедии и семиотических моделей [1].

Основная идея OntoWiki заключается в том, что семантическая сеть формируется множеством пользователей (с прохождением определенной системы модерации). Каждый пользователь может добавить некоторое слово или совокупность нескольких слов, определяющих понятие, объект, процесс (то есть что-то, выделяемое в естественном языке).

Поскольку семантическая сеть есть граф, то узлами графа являются объекты, ассоциированные с понятиями. Первично OntoWiki может быть представлена множеством исходных понятий, определяемых множеством соответствующих терминов, являющихся заголовками словарных статей одного или нескольких словарей. Каждое понятие может иметь описание (математическое, нечеткое, визуальное и т. п. в соответствии с типом описания), причем типы описания не фиксированы, могут быть дополнены или модифицированы. Если использовать для описания узла фрейм, то он мо-

жет дополняться слотами и могут модифицироваться типы слотов.

Кроме этого, для более точного отображения в системе OntoWiki картины реального мира и его ментального отражения каждое понятие (объект) может быть отнесено к различным предметным областям ПрО. При этом следует отметить, что само понятие предметной области не имеет достаточно формального определения [2] и отнесение объектов к конкретной ПрО производится аксиоматически. Формально онтологическая семантическая сеть OntoWiki может быть определена следующим образом:

$$OW = \langle \{I(O), S, P\} \rangle, \quad (1)$$

где $I(O)$ – множество объектов сети, имеющих имена I ; S – множество связей объектов сети; P – множество предметных областей.

В данном определении следует уточнить свойства имен. Имя в нотации Бэкуса – Наура есть

$$i = v|iv|i j, \quad (2)$$

то есть имена могут состоять из нескольких слов, обозначающих некоторый объект. Кроме этого, поскольку один и тот же объект в разных языках имеет различное вербальное отображение (алфа-

Кафтаников Игорь Леопольдович – канд. техн. наук, зав. кафедрой электронно-вычислительных машин, Южно-Уральский государственный университет; kil@chel.com.ru

Пименова Надежда Владимировна – старший преподаватель кафедры электронно-вычислительных машин, Южно-Уральский государственный университет; info@comp.susu.ac.ru

Kaftannikov Igor Leopoldovich – Candidate of Science (Engineering), head of the Department of Electronic Computers of South Ural State University; kil@chel.com.ru

Pimenova Nadezhda Vladimirovna – senior lecturer of the Department of Electronic Computers, South Ural State University; info@comp.susu.ac.ru

вит, синтаксис), необходимо решить вопрос об отображении языка в OntoWiki. Представление этого отображения может быть реализовано либо в виде свойства каждой сущности, включенной в сеть (объекты, связи, ПрО), либо в формате, аналогичном формату предметной области. Второй подход представляется более адекватным решаемой задаче. Тогда

$$OW = \{OWz\}, OWz = \langle \{I(O), S, P\} \rangle, \quad (3)$$

где z – конкретный естественный язык и OntoWiki есть множество семантических сетей, каждая из которых описана на своем собственном языке. Первая версия проекта реализуется как моноязыковая, но предусматриваются механизмы многоязычия.

Описания объектов или связей должны иметь форматы, хорошо поддающиеся визуализации, и в целом могут быть разделены на два подмножества:

- связи фиксированного (формализованного) типа – заполняются по предлагаемому разработчиками шаблону;
- связи нечеткого (произвольного) типа – имеют формат комментария, в последующем могут быть переведены модераторами в фиксированный формат.

В исходном состоянии часть понятий связывается разработчиками (по крайней мере, в графическом представлении) друг с другом. Минимально можно просто указывать, что связь есть, но без конкретизации типа связи. Для этого можно использовать механизм экспертного определения наличия взаимосвязей, используемого авторами, когда значения ячеек матрицы взаимосвязей представляются каждым экспертом в бинарном формате, а совокупное значение, заносимое в ячейки матрицы, есть нормированное.

$$A_{ij} = \text{Sum} (Nk(A_{ij})/K,$$

где $Nk(A_{ij})$ – значение, устанавливаемое экспертом, для ячейки A_{ij} ; $Nk(A_{ij}) = (0,1)$; K – число экспертов.

Полученное численное значение оценки наличия связи может использоваться как оценка уровня связности объектов на уровне локальной группы. Уровень связи объектов для расширяемой группы может быть получен в процессе функционирования проекта голосованием по связности

объектов и при соответствующем изменении уровня оценки. При этом возможна модификация оценивания, в том смысле, что каждый оценивающий указывает не только наличие связи, но и предлагает нормированное численное значение «силы» связи, в соответствии со своими представлениями. Это значение можно определить как некий коэффициент ассоциативности (число, отображающее «вспоминаемость» одного связанного объекта при вводе другого). В этом случае уровень связи может быть несимметричным, а сами связи – направленными. Множество типов связи конечно, но расширяемо. Первоначально типы связи задаются (формальные отношения, нечеткие, словесные, просто наличие связи и т. п., что можно придумать или вспомнить) и для них формируются механизмы соответствия (технические средства описания и привязки к понятиям), а также способы извлечения новых знаний. Для нечетких связей необходимо разработать механизмы отображения их в формальные.

Визуальное отображение OntoWiki формируется в виде сети (WiKi-представление) или дерева (Onto-представление), где центром (корнем) является выбранный термин (понятие) с показом ближайших понятий, связанных тем или иным образом с исходным термином. Для связей отображаются только те типы связей, которые заказаны, но при представлении могут быть показаны те объекты, для которых есть другие типы связей, но сами связи не указываются.

Объект (вершина) OntoWiKi в системе $(I(O), \text{ПрО}, Z)$ может быть описан различными способами.

Вариант I: предметно-терминологическая матрица (табл. 1).

Таблицу можно дополнять или модифицировать. Здесь мы имеем принадлежность термина к различным ПрО. Если бы речь шла об одном языке, то на этом можно было бы остановиться, но для международного применения нужна мультязычность. Вариант ее введения – строки в таблице.

Вариант II – контейнер термина вида (табл. 2).

Под слотами понимаются свойства понятия (объекта), причем слоты могут быть различных типов (точными, диапазонными, нечеткими, вплоть до направленности). Эти слоты могут быть просто присоединенными к термину, без какой-либо специфики, а могут и быть, например, осно-

Таблица 1

Коса (русский)		Языки		
		Английский	Русский	Немецкий и т. д.
Предметные области	Прическа	–	коса	
	Заготовка сена	shallow spit	коса	–
	Береговая линия	Scythe	коса	уууууу
	Электромонтаж	xxxxxx	коса	ууууууу
	Еще что-то	xxxxxx	–	уууууууу

Таблица 2

1. Прическа	коса	Слот 1.1	Слот 1.2	Слот 1.3		...			Слот 1.n
2. Заготовка сена	коса	Слот 2.1	Слот 2.2	Слот 2.3					Слот 2.n
3. Береговая линия	коса	Слот i.j							Слот i.n
4. Электро-монтаж	коса	Слот m.1	Слот m.2	Слот m.3					Слот m.n

ванием для установления отношения, действия, к другому понятию (объекту). Таким образом, мы будем структурировать взаимосвязи объектов примерно так, как они определяются в образном представлении человека: связь не просто объект – объект, а через ПрО-слот конкретизацию. Наверное, можно будет найти отношения, для которых такой тип конкретизации будет плохо подходить в нашем сознании, но представляется, что таких отношений будет немного.

Взаимосвязи (отношения, действия) также рассматриваются многосторонне. Для них опреде-

ляются виды, типы, направленность выражение связи, комментарии и т. п.

Литература

1. Поспелов, Д.А. *Ситуационное управление: теория и практика* / Д.А. Поспелов. – М.: Наука, 1985. – 288 с.
2. Востров, Г. *Проблемы моделирования предметных областей в информационных системах* / Г. Востров, Е. Малахов, К. Корнилова. – <http://www.codenet.ru/progr/other/modeling-problems/>

Поступила в редакцию 7 июня 2012 г.