

УДК 81'374.82:004.89

КОГНИТИВНЫЙ АСПЕКТ РАЗРАБОТКИ ЛИНГВИСТИЧЕСКОЙ ОНТОЛОГИИ ПО ФИЗИКЕ МАГНИТНЫХ ЯВЛЕНИЙ

Потапова Е.В.

*Таврический национальный университет им. В.И.Вернадского, Симферополь, Украина
E-mail: helen1pota@rambler.ru*

В статье представлен опыт разработки классификации связей между понятиями на примере конкретной области – Физики магнитных явлений. Разработка подходов к формализации проблемно-специфических связей в предметных областях позволит повысить когнитивную эффективность онтологий в сфере искусственного интеллекта.

Ключевые слова: лингвистическая онтология, проблемно-специфические связи, формализация знаний.

Постановка проблемы. Когнитивно-информационная составляющая в постановке и решении новых научных и образовательных задач стала особенно актуальна еще в 90-х годах [1] и на данный момент представляет основу разработки интеллектуальных систем (ИС) анализа текста на естественном языке (ЕЯ) и систем извлечения знаний. Ключевым пунктом систем семантического анализа является эффективная словарная поддержка. Поэтому основная проблема в создании ИС – это проблема создания реально работающего понятийного словаря включающего связи между понятиями и нюансы текстового отображения понятий – т.е. онтологии включающей в себя помимо формального каркаса некий лингвистический компонент. Поиск решения этой проблемы привел к возникновению лингвистических онтологий или в другой интерпретации – онтолингвистических систем [2, 3].

Согласно [4], под онтологиями понимают систему явной концептуализации предметной области (ПрО), то есть формального представления предметной области:

$$O = \langle T, R, F \rangle \quad (1)$$

— это упорядоченная тройка конечных множеств, где T — термины ПрО, которую описывает онтология O ; R — отношения между терминами заданной ПрО; F — функции интерпретации, заданные на терминах и/или отношениях онтологии O .

Формальный подход к построению онтологий ПрО был дополнен лингвистическим, где онтологии выстраиваются на основе корпуса текстов, а термины явно связаны с их дефинициями. Главной характеристикой лингвистических онтологий является то, что понятия онтологии связаны со значениями (“are bound to the semantics”) языковых выражений (слов, именных групп и т.п.) [5]

Лингвистическую компоненту онтологии согласно [2] формально описывает структура вида $LS (\Psi^L, G^L)$, где Ψ^L — множество объектов, G^L — множество атрибутов объектов. Множество атрибутов $G^L = G^N \times G^S \times G^G \times G^C \times G^D$, где G^N — множество лексических параметров; G^S — множество синтаксических параметров (параметры правил синтаксического согласования составляющих); G^G — множество грамматических категорий; G^C — множество композиционных параметров (параметры пра-

вил образования многословных терминов, правил вариативности элементов); G^D – множество дистантных параметров (расстояние между составляющими термина).

На основании вышеизложенного была разработана концептуальная модель лингвистической онтологии по Физике магнитных явлений (ФМЯ) и лексикографическая параметризация терминов онтологии [6, 7]

Одними из основных требований к построению онтологий ПрО [8] являются требования *формализации* знаний ПрО и *полноты* отображения информации ПрО. Формализация знаний ПрО, выраженных помимо терминов еще и некоторым количеством текста на естественном языке, приводит к некоторой потере информации и частичной десемантизации терминов [9]. Следовательно, появляется проблема обеспечения семантической полноты терминов входящих в онтологию. Очевидно, что необходимо увеличить некоторым образом количество информации, зафиксированное в онтологии. Достигается это, как правило, [5, 3, 10, 11] двумя путями:

1) за счет увеличения количества понятий в онтологии, т.е. в число терминов попадают слова, имеющие общезыковое значение и сопутствующие терминам в текстах;

3) за счет введения подробной классификации понятий, отношений между понятиями.

Этапы работы над онтологией по ФМЯ докладывались на конференциях и были опубликованы [12, 6, 7]. В ходе работ были проанализированы оба пути развития и сделаны следующие выводы:

1) Увеличение предметной онтологии за счет слов имеющих общезыковое значение не должно носить глобальный характер, т.к. в этом случае онтология перестанет быть предметной, и начнет приближаться к общезыковой онтологии, а это чревато другими проблемами – например колоссальным увеличением объема. Вместо этого, общезыковые понятия, часто употребляемые с тем или иным термином ПрО, сохранялись в библиотеке текстовых входов для каждого термина.

2) Предварительный анализ прикладной области и публикаций по данному вопросу [5, 13, 14] показал, что разработка подробной классификации отношений между понятиями менее подвержена субъективизму конкретной группы экспертов, нежели введение подробной классификации понятии ПрО. Таким образом, можно сделать вывод о том, что посредством конкретизации ассоциативных отношений можно включить в онтологию некоторый (достаточно существенный) объем дополнительной информации о предметной области.

ЦЕЛЬ РАБОТЫ И МЕТОДИКА ВЫПОЛНЕНИЯ

Таким образом, целью следующего этапа работы над онтологией по ФМЯ оказалась разработка подробной классификации отношений между понятиями в предметной области «Физика магнитных явлений». Данная работа является продолжением предыдущего проекта («Словарь терминов по ФМЯ»). Материалом и средством разработки онтологии является инструментальная лексикографическая система по ФМЯ [7]. Словарь составлен на основе нескольких печатных источников [15, 16] и включает в себя более 900 терминов.

Работа выполнялась в несколько этапов. На первом этапе были проанализированы типы связей, установленные стандартом онтологического инжиниринга IDEF5, OWL, а также типы связей принятые в других онтологиях. Затем на материале прикладной области группой экспертов были выделены проблемно-специфические типы связей характерные для ФМЯ. Полученная информация была объединена в классификацию связей для онтологии ФМЯ.

Обзор публикаций по проектированию и поддержанию онтологий и интеллектуальных систем, основанных на них [5, 8, 10, 13] показал, что, как правило, в онтологиях устанавливаются 5 - 6 типов отношений: «часть-целое», «класс-подкласс», «жесткая ассоциация», «не жесткая ассоциация», «несовместимость», «является свойством», «историческая зависимость». К числу «стандартных» (иерархических) связей относят [17, 18]: «часть-целое», «класс-подкласс».

Наиболее подробная разработка данного вопроса содержится в статье Шалфеевой «Классификация структурных свойств онтологий» [14], где помимо прочих свойств подробно рассматривается классификация проблемно-специфических (или «топологических») отношений между понятиями в онтологиях.

Наряду с определением иерархических связей, для создания онтологии необходимо выделение проблемно-специфических связей характерных именно для данной ПрО. С одной стороны наличие стандартных иерархических связей позволяет конструировать онтологии с возможностью объединения и повторного использования (reusable ontology), с другой стороны для того чтобы интеллектуальная система, построенная на данной онтологии работала адекватно, необходима достаточно подробная фиксация проблемно-специфических связей между понятиями в выбранной ПрО. Существует определенный баланс между подробностью определения проблемно-специфических связей и субъективностью конкретной группы экспертов. Таким образом, разработка и внедрение в онтологию подробной системы проблемно-специфических связей может существенно повысить когнитивную эффективность системы анализа ЕЯ текста и извлечения знаний.

РЕЗУЛЬТАТЫ

В результате работы была разработана классификация связей между понятиями ФМЯ. Классификация содержит три больших раздела: Таксономические связи, Композиционные связи и Проблемно-специфические связи.

К числу «стандартных» связей, принятых в онтологии ФМЯ, относятся:

Таксономические связи

- Гипероним – Гипоним (род-вид)

Магнитная структура –

>Соизмеримая магнитная структура

>Несоизмеримая магнитная структура

- Тождество (синонимия)

Спиральная магнитная структура = Винтовая магнитная структура

Композиционные связи

- Целое - Компоненты

магнитная структура – магнитная подрешетка

■ Вещество – Ингредиенты

алин – железо; - никель; - алюминий; - медь.

■ Набор - Члены_из_набора

тензор констант обменного взаимодействия – константа обменного взаимодействия

Для определения типов **Проблемно-специфических связей** было признано целесообразным ввести вначале типизацию понятий ПрО, а именно сопоставить всем терминам словаря пять типов понятий : Сущность/Актор, Процесс, Явление, Свойство, Характеристика.

Сущность – элементарный объект, актор Процесса(Явления), среда реализации Процесса (Явления). Примеры: (парамагнитные) атомы, с некомпенсированными спинами (электрон, ядро)

Явление – то, в чем проявляется, обнаруживается Сущность. Пример: ферромагнетизм. В книге философов Шанжё Ж-П. и Конн А.Материя и мышление.(1990), где затрагиваются некоторые вопросы связанные с искусственным интеллектом, теорией информации и физикой, явление в физическом смысле определяется следующим образом: «... физическое явление есть результат опыта, который, если точно оговорить исходные данные, сможет с идентичным результатом воспроизвести далее экспериментатор в другой лаборатории».

Процесс - последовательная смена состояния Сущности во времени и в пространстве. Примеры: магнитное упорядочение, фазовый переход. Граница между Процессами и Явлениями достаточно условна, т.к. некоторые понятия можно справедливо отнести к обеим категориям.

Свойство (в философии, математике и логике) — атрибут предмета (объекта). Например, о красном предмете говорится, что он обладает свойством красноты.

Свойство – пример: Свойство вещества –*Намагниченность*.

Характеристика (величина, направление и.т.п. – можно измерить) – это количественное описание некоторого Свойства объекта.

Характеристика – пример: Характеристика вещества – *Вектор намагниченности* (значения заряда, спиновых и орбитальных магнитных моментов, величины температуры и энергии).

Таким образом, на основе введенных типов понятий можно определить Проблемно-специфические связи как связи, показывающие «пути» физических взаимодействий между Сущностями, в результате которых возникают соответствующие Процессы и Явления, а также связи между Явлениями, Процессами, Сущностями и их Свойствами и Характеристиками.

На основании такого подхода были определены следующие типы **Проблемно-специфических связей**:

■ связи Сущностей с Процессами и Явлениями

√ Сущность является первичным (основным) источником Процесса(Явления)

электрон – обменное взаимодействие

√ Сущность участвует в Процессе (Явлении)

магнитная подрешетка– слабый ферромагнетизм

- √ Сущность является средой реализации Явления (Процесса)
магнетик - моноклинная магнитоупругость
- причинно-следственные связи (между Процессами и Явлениями)
 - √ Процесс(Явление) является предпосылкой другого Процесса(Явления)
Намагниченность – термомагнитные эффекты
 - √ Процесс(Явление) влияет на другой Процесс(Явление)
неоднородное обменное взаимодействие – обменная спиновая волна
 - √ Процесс(Явление) зависит от Процесса(Явления) / $A = f(B)$ (строгая зависимость по формуле)
магнитоупругая энергия – вектор намагниченности
- временные связи - корреляция прохождения Процессов (Явлений)
 - √ Процесс1 коррелирует с Процесс 2
Намагничивание -магнитное насыщение
- пространственные связи - связи Сущностей, обусловленные взаимным пространственным расположением
 - √ Сущность 1 располагается внутри Сущность 2
Частная петля гистерезиса –полная петля гистерезиса
 - √ Сущность 1 пересекается с Сущность 2
магнитная ось – плоскость симметрии кристалла
- связи между Сущностями (Явлениями, Процессами) и их Свойствами и Характеристиками.
 - √ Сущность(Процесс, Явление) имеет Свойство
магнетик – намагниченность
 - √ Сущность(Процесс, Явление) имеет Характеристику
*магнитная анизотропия – энергия магнитной анизотропии
электрон – спиновый момент*

Разработанная классификация связей была внедрена в Инструментальную лексикографическую систему по ФМЯ [7], где была проведена работа по установлению связей между понятиями в ряде разделов физики магнитных явлений и получены соответствующие схемы связей – карты терминов, которые являются взаимосвязанными фрагментами онтологической схемы предметной области.

Из главной страницы «Словаря терминов» (Рис.1) имеется доступ к отдельным страницам словарных статей терминов. Словарные статьи терминов связаны гиперссылками с другими словарными статьями и в то же время словарная статья термина связана гиперссылкой с Картой термина. Сами по себе Карты терминов тоже связаны между собой гиперссылками в соответствии со смыслом термина и его связями с другими терминами, которые были определены экспертом. Таким образом, по мере определения связей между терминами будет сформирована общая онтологическая схема предметной области, состоящая из отдельных Карт терминов.

На данном этапе существует 352 карты терминов. Карты терминов генерируются Инструментальной лексикографической системой в автоматическом режиме в формате HTML. На рис.2 приведена карта термина «магнитная подрешетка». Дизайн изображения обусловлен тем, что изначально Электронный словарь по ФМЯ с

онтологическими картами терминов проектировался в качестве дополнительного пособия в учебном процессе.

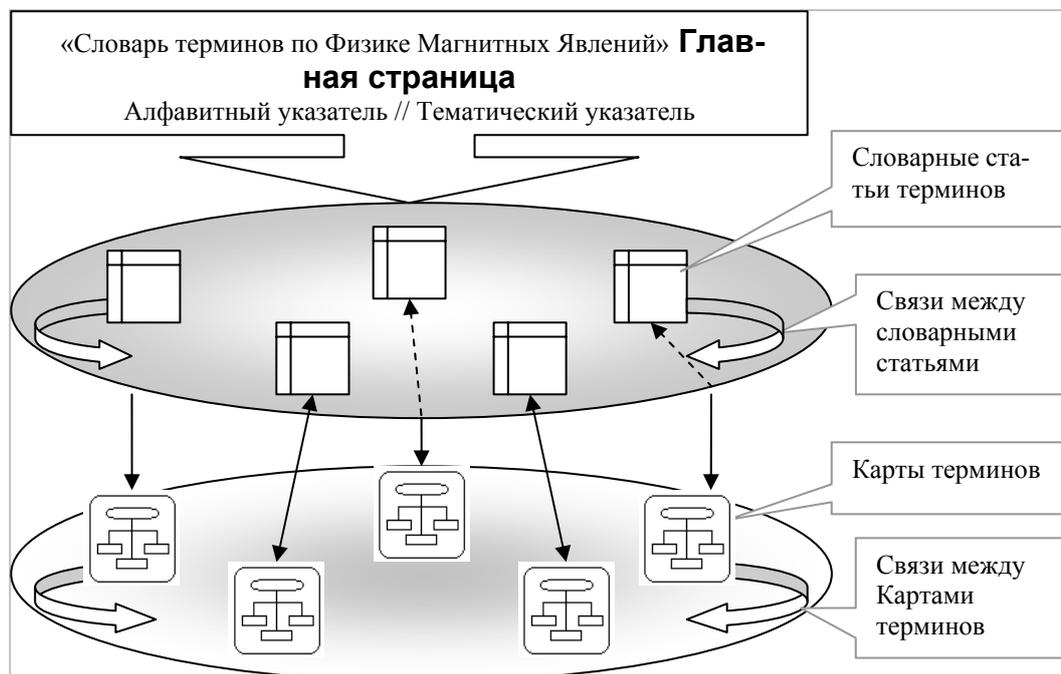


Рис. 1. Схема навигации по словарным статьям и картам терминов.

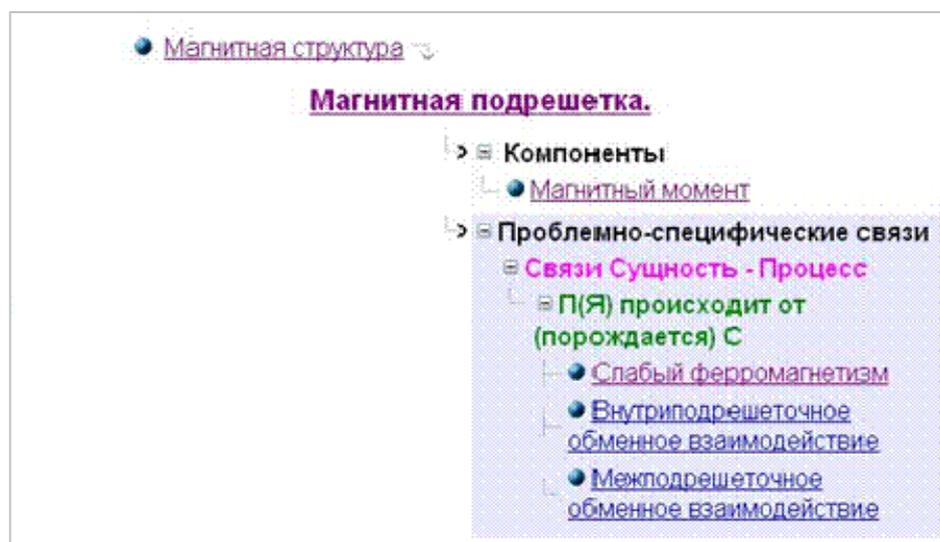


Рис.2. Карта термина – «Магнитная подрешетка»

Выводы и перспектива. Разработанная классификация связей между терминами ПрО делает описание каждого термина онтологии более полным. Количество связей на обработанном участке онтологии увеличились на 40%, что существенно увеличивает информационную полноту системы и повышает ее когнитивную эффективность.

Подробная классификация проблемно-специфических связей позволит с помощью онтологии более точно вычислять ряд "служебных" семантических функций, представляющих важные для анализа текста и обработки знаний внутриязыковые связи, не имеющие прямой предметной интерпретации.

В аспекте повторного использования и интеграции онтологий нужно отметить, что введенная здесь типизация понятий имеет прямые категориальные соответствия в онтологиях верхнего уровня, а установленные типы проблемно-специфических связей приводятся в соответствие с принятыми в других онтологиях ПрО типами связей.

Литература

1. Палагин А.В. К проблеме проектирования системы активации научно-исследовательской деятельности / Палагин А.В.: В кн. "Вопросы когнитивно-информационной поддержки постановки и решения новых научных проблем" Сб. научн. тр./ НАН Украины. Ин-т кибернетики им. В.М.Глушкова. - Киев, 1995, С. 4-16.
2. Невзорова О.А. Онтологическая поддержка методов решения задач семантико-синтаксического анализа текстов./ Невзорова О.А. [Электронный ресурс]. КИИ-2008 Одиннадцатая национальная конференция по искусственному интеллекту с международным участием. г. Дубна 29 сентября – 3 октября, 2008. //Материалы конференции, Режим доступа: www.gai.org/cai-08/files/cai-08_paper_234.doc.
3. Методы и средства автоматизированного проектирования прикладной онтологии /Добров Б.В., Лукашевич Н.В., Невзорова О.А., Федун Б.Е.// Известия РАН. Теория и системы управления. М.: 2004. № 2. С. 58-68.
4. Gruber T.R. A translation approach to portable ontologies: Knowledge Acquisition/ Gruber T.R. V.5(2) — 1993.
5. Fernandez M. A Building a Chemical Ontology Using Methodology and the Ontology Design Environment / Fernandez M., Gomez-Perez A., Pazos J.// IEEE Intelligent Systems, Jan./Feb. Pages 37-46, 1999.
6. Бержанский В.Н. Формирование онтологии на материале терминосистемы физики магнитных явлений/ Бержанский В.Н., Потапова Е.В., Салюк О.Ю. // Прикладна лінгвістика та лінгвістичні технології: MegaLing 2007: сб.науч. тр./ отв.ред. В.А.Широков – К.: Довіра, 2008, С.61-67.
7. «Инструментальная лексикографическая система в области физики магнитных явлений» /Дикарева С.С., Бержанский В.Н., Полулях С.Н., Потапова Е.В. // «Бионика интеллекта». – Харьков:ХНУРЕ, № 2(71), 2009, С. 48-53.
8. Гладун А.Я. Онтологии в корпоративных системах. Часть I и II [Электронный ресурс] / Гладун А.Я., Рогошина Ю.В. // "Корпоративные системы" (№1, 2006) Режим доступа: <http://www.management.com.ua/ims/ims115.html>.
9. Новиков А.И. Семантика текста и ее формализация. / Новиков А.И.: - М.: Наука, 1983. -215 с.
10. Лукашевич Н.В. АвиаОнтология: анализ современного состояния ресурса [Электронный ресурс] / Лукашевич Н.В., Невзорова О.А.// Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог 2007» Режим доступа: <http://www.dialog-21.ru/archive/2007/lukashevich.htm>.
11. Рубашкин В. Ш. Онтологии - проблемы и решения. Точка зрения разработчика. / Рубашкин В.Ш. [Электронный ресурс] Материалы Конференции «Диалог-2007». Режим доступа: <http://www.dialog-21.ru/dialog2007/materials/html/74.htm>.
12. Бержанский В.Н. Моделирование и визуализация когнитивных связей в терминосистеме физики магнитных явлений. / Бержанский В.Н. Полулях, С.Н., Потапова Е.В./ MegaLing 2006. Горизонти прикладної лінгвістики та лінгвістичних технологій. Доповіді міжнародної конференції. Ред.

В.А. Широков, С.С. Дікарева. Україна, Партеніт, 20-26 вересня 2006. Сімферополь: «ДиАйПи». 2006. – С. 137-138.

13. Овдей О.М. Обзор инструментов инженерии онтологий. / Овдей О.М., Проскудина Г.Ю. [Электронный ресурс]// Труды Шестой Всероссийской научной конференции "Электронные библиотеки: перспективные методы и технологии, электронные коллекции", 2004 Режим доступа: <http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2004/part4/op>.

14. Шалфеева Е.А. Классификация структурных свойств онтологий. / Шалфеева Е.А. // «Штучный интеллект» 3'2005, С.67-77.

15. Физика твердого тела. Энциклопедический словарь: в 2 т. / Под ред. В. Г. Барьяхтара. Т. 1.– Киев: Наукова думка, 1996. – 656 с., Т. 2.– Киев: Наукова думка, 1998. – 648 с.

16. Терминологический справочник «Магнетизм и магнитные материалы» / Под ред. Ф. В. Лисовского и Л. И. Антонова. – М.: «Вагриус», 1997. – 240 с.

17. Стандарт онтологического исследования IDEF5 [Электронный ресурс]// Режим доступа: <http://www.idef.com/idef5.html>.

18. OWL Web Ontology Language Guide. [Электронный ресурс]W3C Recommendation 10 Feb 2004. Smith, Welty, McGuinness, eds. // Режим доступа: <http://www.w3.org/TR/owl-guide/>.

Потапова О.В. Когнітивний аспект розробки лінгвістичної онтології по фізиці магнітних явищ / О.В. Потапова // Вчені записки Таврійського національного університету ім. В. І. Вернадського. Серія «Філологія. Соціальні комунікації». – 2011. – Т. 24 (63), № 3. – С. 292-299.

У статті представлено досвід розробки класифікації зв'язків між поняттями на прикладі конкретної області - Фізики магнітних явищ. Розробка підходів до формалізації проблемно-специфічних зв'язків у предметних областях дозволить підвищити когнітивну ефективність онтологій у сфері штучного інтелекту.

Ключові слова: лінгвістична онтологія, проблемно-специфічні зв'язки, формалізація знань.

Potapova E.V. Cognitive aspects of linguistic ontology development on physics of magnetic phenomena / E.V. Potapova // Scientific Notes of Taurida V. I. Vernadsky National University. – Series: Philology. Social communications. – 2011. – Vol. 24 (63), No 3. – P. 292-299.

This article presents the experience of relation classification developing between the concepts in the specific field of science - the physics of magnetic phenomena. Development of approaches to formalize the problem-specific relations in the subject domains can improve the cognitive efficiency of ontology usage in the field of artificial intelligence.

Keywords: linguistic ontology, problem-specific relations, formalization of knowledge.

Поступила в редакцію 01.09.2011 г.