

УДК 004.7

**МОДЕЛИ ПРЕДМЕТНЫХ ОБЛАСТЕЙ В СИСТЕМАХ  
ОРГАНИЗАЦИОННОГО УПРАВЛЕНИЯ**

А.Г. Додонов, Д.В. Ландэ, В.Г. Пуятин

*Институт проблем регистрации информации НАН Украины*

e-mail: dwlande@gmail.com

Сегодня находят все большее применение знание-ориентированные информационные системы с онтолого-управляемой архитектурой [1]. Они реализуют интегрированную информационную технологию, включающую компьютерную обработку естественно-языковых объектов, описывающих предметные знания: объекты, их свойства и понятия предметной области [2, 3].

В настоящее время уже не хватает традиционных методов и средств обработки информации для решения прикладных задач в специализированных предметных областях, связанных с извлечением предметных знаний и их последующим использованием в этих предметных областях.

Онтологический подход логично интегрируется с системами организационного управления, которые относятся к сложным динамическим системам. Состав и содержание параметров систем организационного управления требует наличия информации о критериях качества и целях управления, структуре и элементном составе объекта управления, а также, и характеристиках предметной области и внешней среды, с которой система должна взаимодействовать.

При формировании предметных онтологий, применяемых к системам организационного управления, решается несколько содержательных задач, среди которых:

- выявление основных понятий и соответствующих им слов и словосочетаний из соответствующей предметной области;
- выявление различных семантических связей между объектами управления и соответствующих им понятий;
- выявление главных понятий в предметной области;
- выявление главных связей.

Построение онтологии в процессе изучения системы организационного управления может рассматриваться как прямая задача, где заранее известна схема сети организационного управления, основные объекты и связи, в соответствии с которыми формируется сеть понятий.

Задача автоматического формирования модели предметной области еще на этапе формирования системы, когда не определены объекты и связи, может рассматриваться как обратная задача, когда сам состав понятий и связей, выбираемых при создании предметной области, определяют дальнейший состав (а иногда и функции) системы.

Именно в решении такой обратной задачи заключается проблема формирования онтологий предметных областей. Автоматическое формирование онтологий базируется на знаниях, заложенных учеными, специалистами, экспертами в таких источниках, как:

- 1) фактографические базы данных (знаний);
- 2) текстовые корпуса;
- 3) ресурсы современных социальных, наукометрических, библиографических сетей.

Конечно, если существуют фактографические базы данных или знаний, то проблема автоматического формирования онтологий на их основе может оказаться несложной. Но при этом задача предварительного формирования таких баз данных и знаний, сама по себе, является проблемной.



Для работы с третьим типом источников предлагается методика построения моделей предметных областей на основе зондирования информационных сетей [7]. Как такая сеть рассматривается сеть понятий, соответствующих тегам сервиса Google Scholar Citations (<http://scholar.google.com/citations>). На интерфейсе этой системы, соответствующем определенному тегу постранично в ранжированном виде отображаются имена ученых, которые отметили свою деятельность этим тегом, а также другие теги, приписанные ими. Множество тегов образуют сеть, производную от биграфа «ученный-теги». Узлы в этой сети соответствуют понятиям, маркированным тегами, а связи – некоторую семантическую связь между ними.

Зондирование опорной модельной сети осуществляется по такому алгоритму:

1. Выбирается определенное количество узлов опорной (зондируемой) сети, определяемых как базовые для новой сети, соответствующей результатам зондирования.
2. Для каждого из рассматриваемых узлов опорной сети определяются смежные с ним узлы ("соседи"), которые добавляются к создаваемой сети с результатами зондирования.
3. От текущего узла опорной сети осуществляется переход к соседнему узлу, имеющему наибольшую степень.
4. Если имеет место "зацикливание" (выбирается узел, к которому уже был осуществлен переход по этому алгоритму), происходит переход к следующему по степени соседнему узлу. Если таких узлов не осталось – осуществляется переход к пункту 2.
5. Если перечень базовых узлов завершен, считается, что сеть, соответствующая результатам зондирования, построена.
6. Формирование базового стартового перечня узлов-понятий и правил отбора «конечных» узлов выполняется экспертами в предметной области.
7. Для построения модели предметной области (в рассматриваемом примере для области искусственного интеллекта) экспертным путем были определены базовые теги на английском языке: *survivability, reliability, dependability, resilience, fault tolerance, availability, safety, durability, security, resiliency* и др.

На рис. 2 приведен пример архитектуры сети понятий предметной области, построенной в соответствии с приведенным алгоритмом по указанным базовым тегам.

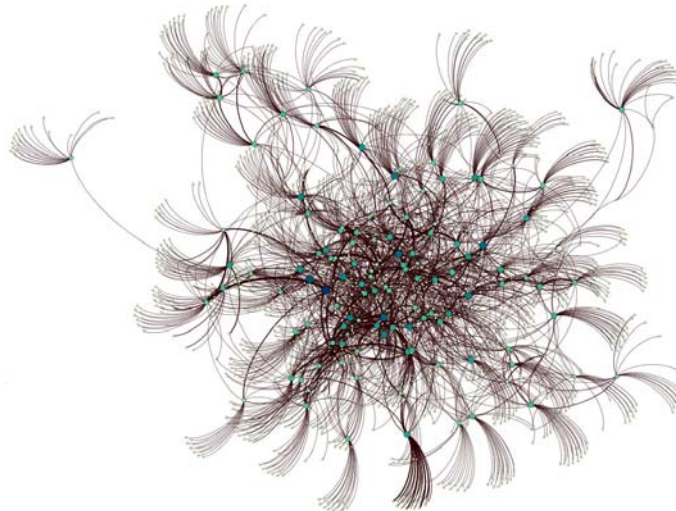


Рис. 2 - Фрагмент архитектуры сети понятий

Построенная сеть понятий оказалась связной. При количестве базовых тегов 20, общее количество узлов-тегов, которые были охвачены алгоритмом, составили 125, а количество нетерминальных узлов – 1450.

Автоматически сформированные терминологические основы онтологий и соответствующие, пусть и простейшие, семантические связи по выбранным свойствам (живучести, надежности) могут использоваться, в частности, при проектировании систем организационного управления в качестве «единого» для всех участников разработки языка предметной области, для обучения, тренингов, организации семантического поиска (организации контекстных подсказок информационно-поисковых систем), навигации пользователей в соответствующих информационных ресурсах.

Список литературы

1. Палагин А.В. Онтологические методы и средства обработки предметных знаний: монография / А.В. Палагин, С.Л. Крывый, Н.Г. Петренко. – Луганск: изд-во ВНУ им. В. Даля, 2012. – 324 с.
2. Ланде Д.В. Елементи комп'ютерної лінгвістики в правовій інформатиці. – К.: НДПП НАПрН України, 2014. – 168 с.
3. Добров Б.В., Соловьев В.Д., Лукашевич Н.В., Иванов В.В. Онтологии и тезаурусы. Модели, инструменты, приложения / Б.В. Добров, В.Д. Соловьев, Н.В. Лукашевич, В.В. Иванов. – М: Бином, 2009. – 173 с.
4. Додонов А.Г. Живучесть информационных систем / А.Г. Додонов, Д.В. Ландэ. – К.: Наук. думка, 2011. – 256 с.
5. Ландэ Д.В. Построение терминологической сети предметной области / Д.В. Ландэ, А.А. Снарский, В.Г. Путятин // Реєстрація, зберігання і обробка даних, 2014. – Т. 14. – № 2. – С. 114-121.
6. Ландэ Д.В. Применение КГГВ-алгоритма для научных текстов / Д.В. Ландэ, А.А. Снарский, Е.В. Ягунова // Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2014): материалы IV междунар. науч.-техн. конф. (Минск 20-22 февраля 2014 года) / – Минск: БГУИР, 2014. – С. 199-204.
7. Ланде Д.В. Мережевий підхід до створення моделі предметної області // Правове реулювання інформаційних відносин та сфери інтелектуальної власності в умовах кіберцивілізації: Матеріали науково-практичної конференції / 26 березня 2015 р., м. Київ. – К.: НДПП НАПрН України, НТУУ "КПІ", 2015. – С. 18-20.