
К АНАЛИЗУ ЕСТЕСТВЕННО-ЯЗЫКОВЫХ ОБЪЕКТОВ

Александр Палагин, Сергей Кривый, Виталий Величко, Николай Петренко

Аннотация: Рассматриваются проблемы анализа естественно-языковых объектов (ЕЯО) с точки зрения их представления и обработки в памяти компьютера. Предложена формализация задачи анализа ЕЯО и приведен пример формализованного представления ЕЯО предметной области.

Ключевые слова: термины предметной области, формальная логическая система, онтология.

ACM Classification Keywords: I.2.4 Knowledge Representation Formalisms and Methods - Representation languages, I.2.7 Natural Language Processing - Language models

Conference: The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009

Введение

Проблема извлечения знаний из естественно-языковых объектов, их представление и обработка относятся к области искусственного интеллекта. Этой чрезвычайно важной проблемой занимались многие выдающиеся ученые (Аристотель, Лейбниц, Тарский и др.). В общей проблеме анализа и понимания ЕЯО, во всем многообразии работ [1-5], методов, методологий и терминологии явно просматриваются следующие задачи:

- синтактико-семантический анализ ЕЯО,
- анализ ЕЯО с целью выбора подходящего представления в памяти компьютера и дальнейшей эффективной обработки,
- разработка методов извлечения знаний из ЕЯО.

Рассмотрим эти задачи более детально.

Проблема анализа ЕЯО естественным образом связана с **операционной и денотационной семантикой** языка. В системах искусственного интеллекта и, в частности, в системах анализа ЕЯО преобладает операционная семантика, так как очевидно преобладание интереса к средствам представления знаний (**как?**), а не к их содержанию (**что?**), которое подлежит формализации.

Системы (языки) представления знаний определяют конструкции, поддерживаемые компьютером и обеспечивающие простоту организации процедур, и как следствие – максимальную унификацию языковых средств.

На информационно-логическом уровне преследуется совершенно иная цель - обеспечение отражения содержания объектов реального мира в языковых конструкциях, согласующихся с восприятием, свойственным человеку. Для пользователя такой системы (схемы) компьютерное представление неудобно, поскольку он общается (или, по крайней мере, должен) с системой на естественном языке.

Однако, для системного аналитика, использующего язык формализации для построения модели предметной области (ПдО) и создающего для этого ориентированные на эту ПдО словари понятийных единиц и на их основе - интеллектуальные системы, наполненные конкретным знанием, компьютерное представление играет ключевую роль. Аналитику приходится иметь дело со словами (терминами) и конструкциями какой-то части естественного языка, именно им давать толкование в терминах языка формализации, устанавливать между ними информационно значимые связи. Поэтому, в случае

расхождения языка формализации со строем мышления специалиста, реализация системы обработки данных в конкретном смысловом материале может стать слишком сложной или вообще невозможной для решения проблемой.

Существует много языков представления знаний, которые формализуют ЕЯО фрагментарно, а иногда и недостаточно четко. Поэтому наиболее приемлемым на сегодняшний день являются языки логического типа, в которых понятие "связь по смыслу" формализуется с максимальной полнотой. Это подтверждается и практикой. Одним из главных вопросов при этом является вопрос об уровне абстрактности зафиксированных категорий логического языка. Если нас интересует лишь общая схема логического следования, то достаточно рассматривать лишь исчисление предикатов первого порядка с правилом резолюций в качестве правила вывода, не конкретизируя ничего, кроме списка аксиом и логических констант. Однако, при более богатом содержанием ПдО должен быть решен вопрос о типах термовых констант и виде нелогических аксиом (собственных аксиом), используемых при формализации знаний об эмпирических объектах.

Из всего сказанного следует проблема **разработки языка логического типа для представления знаний, содержащихся в тексте**. Очевидно, что этот язык должен быть расширяемым в сторону способности более полного описания объектов, способности делать предсказания (гипотезы), учитывающие контекст.

Формальная постановка задачи анализа ЕЯО

Пусть $T = t_1, t_2 \dots t_n$ естественно-языковой текст в алфавите X , т. е. $T \in L(Y)$, где $L(Y)$ - язык над алфавитом Y , а $t_i \in T$ - предложения, $i = 1, 2, \dots, n$.

Каждое предложение $t_i \in T$, в свою очередь, имеет структуру $t_i = t_{i_1}, t_{i_2} \dots t_{i_m}$, где t_{i_j} содержательно означают грамматические единицы, из которых построено предложение t_i . Если $t_{i_j} \in t_i$, то $C_L(t_{i_j}) = t_{i_1} \dots t_{i_{j-1}}$ и $C_R(t_{i_j}) = t_{i_{j+1}} \dots t_{i_m}$ будем называть левым и правым контекстом слова t_{i_j} соответственно.

С текстом T свяжем такие объекты:

- S – словарь языка $L(Y)$, где содержатся слова t_{i_j} со своими определителями- a_i и τ_i ;
- $\gamma \subseteq T \times S$ – отношение, определяющее возможные значения и типы слова в словаре S ;
- $M = (D, \Pi)$ – предметная модель, на которой интерпретируется текст T из предметной области D ;
- $\phi \subseteq T \times M$ – отношение интерпретации текста T на модели $M = (D, \Pi)$.

Сигнатура предикатов $\Pi = \{\pi_1^{k_1}, \dots, \pi_r^{k_r}\}$ содержит атомарные предикаты, из которых можно строить сложные формулы. Сейчас мы не будем фиксировать эту сигнатуру, поскольку она зависит от предметной модели M . Поскольку модель не уточняется, то и сигнатуру уточнить нельзя. Заметим только, что каждый атомарный предикат имеет тип.

Определим теперь правила вычисления отношений γ и ϕ .

Отношение γ имеет достаточно простой способ вычисления:

$$\gamma(t_{i_j}) = \{(a_1, \tau_1), (a_2, \tau_2), \dots, (a_s, \tau_s)\},$$

где a_i – возможные значения слова t_{ij} , τ_i – его возможные типы (например, часть речи). Может случиться, что $\gamma(t_{ij}) = \emptyset$. В этом случае значение этого слова считается неопределенным и требует пополнения словаря S .

Отношение ϕ определяется следующим образом:

$$\phi(T) = \phi(t_1) \dots \phi(t_n),$$

где $\phi(t_i) = \{ \phi(\gamma(t_i)\gamma(C_R(t_i))), \phi(\gamma(C_L(t_i))\gamma(t_i)\gamma(C_R(t_i))), \dots, \phi(\gamma(C_L(t_i))\gamma(t_i)) \}$,

$$\phi(\gamma(t_{ij})) = \gamma(\phi(t_{ij})); \quad \phi(\gamma(C_L(t_{ij}))) = C_L(\phi(\gamma(t_{ij}))); \quad \phi(\gamma(\pi_r^k(p_1, \dots, p_k))) = \gamma(\phi(\pi_r^k))(\phi(\gamma(p_1), \dots, \gamma(p_k))),$$

где $\gamma(\phi(\pi_r^k))$ имя предиката, тип которого согласован с аргументами $\gamma(p_1), \dots, \gamma(p_k)$. Из этой формальной постановки проблемы анализа ЕЯО следует, что основные задачи анализа сводятся к таким:

- конкретизировать предметную модель M ; эта задача основная в связи с тем, что предметная модель является по существу базой знаний (конкретизация состоит в том, чтобы определиться с формальным логическим языком, правилами вывода, аксиоматикой и т.п.);
- показать вычислимость отношений γ и ϕ на предметной модели M ;
- построить алгоритмы вычисления отношений γ и ϕ ;
- при вычислении отношений γ и ϕ контролировать соответствие типов аргументов и предикатов;
- определить взаимодействие алгоритмов вычисления γ и ϕ с системами синтаксического и семантического анализа текста.

Рассмотрим пример использования введенного формализма для анализа ЕЯ текста. Задача состоит в том, чтобы получить формальное представление текста в полуавтоматическом режиме взаимодействия с пользователем – экспертом предметной области, при минимальном его участии в процессе анализа.

Использование предложенного формализма для анализа ЕЯО

Словарь S , со словами t_{ij} и их определителями a_i , τ_i создается на основе результатов лексикографических исследований [4] и, по-возможности, является универсальным, независимым от предметной области. В качестве примера ЕЯ текста выберем небольшой фрагмент из предметной области "сверхтвердые материалы" [6]. Слева от предложений указан номер строки в фрагменте текста.

- 1 ПОЛУЧЕНИЕ И СВОЙСТВА ПОЛИКРИСТАЛЛИЧЕСКИХ МАТЕРИАЛОВ НА ОСНОВЕ АЛМАЗА
- 2 1.Материалы, синтезированные из неалмазного углерода
- 3 Поликристаллические материалы на основе алмаза по своим физико-механическим свойствам
- 4 могут быть близкими к монокристаллам, а по некоторым — и превосходить их. Так,
- 5 большинство алмазных поликристаллов обладает изотропией свойств, отличается высокой
- 6 износостойкостью и превосходит монокристаллы по трещиностойкости (ударной вязкости).
- 7 По принятой в настоящее время классификации поликристаллические сверхтвердые
- 8 материалы (ПСТМ) делятся на четыре группы.
- 9 I группа — ПСТМ, получаемые при высоких давлениях и температурах путем превращения
- 10 графита в алмаз в присутствии специальных сплавов-растворителей углерода. К этой
- 11 группе относятся, например, поликристаллы типа баллас (АСБ) и карбонадо (АСПК).
- 12 Следует отметить, что в месторождениях алмазов встречаются природные
- 13 поликристаллы подобных типов.
- 14 II группа — ПСТМ, получаемые путем спекания алмазных порошков в условиях высоких

- 15 статических давлений и температур.
 16 III группа — ПСТМ, получаемые при высоких статических давлениях и температурах на
 17 подложке из твердого сплава или стали, а также ПСТМ в твердосплавной или металлической
 18 обойме.
 19 IV группа — ПСТМ, получаемые методом спекания алмазных порошков при более низких
 20 давлениях (до 1, 9 ГПа) в присутствии связующего материала.

Уточним сигнатуры предикатов предметной модели. С помощью поверхностного синтактико-семантического анализа [7] для всех слов из приведенного примера найдем $\gamma(t_i)$, используя словарь S .

Для каждого предложения текста строятся деревья зависимостей, в которых определяются синтактико-семантические отношения между словами. Далее автоматически выделяются из текста термины предметной области D в соответствии с заданными шаблонами, например: аббревиатура, существительное, прилагательное_существительное, существительное_прилагательное_существительное и типами синтактико-семантических отношений: объектное, принадлежность (между двумя существительными), определительное (между прилагательным и существительным), однородные слова. В качестве терминов, используемых в формальном представлении текста, автоматически выбираются преимущественно многословные словосочетания, в которых значения существительных уточняются с помощью связанных прилагательных.

Приведем фрагмент списка терминов, состоящего из отдельных слов и словосочетаний: *свойства поликристаллических материалов, алмаз, материал, неалмазный углерод, поликристаллический материал, физико-механическое свойство, алмазный поликристалл, поликристалл, изотропия свойств, высокая износостойкость, трещиностойкость, ударная вязкость, поликристаллический сверхтвердый материал, ПСТМ, I группа ПСТМ, II группа ПСТМ, III группа ПСТМ, IV группа ПСТМ, высокое давление, давление, высокая температура, температура, превращение графита, графит, сплав-растворитель углерода, баллас, АСБ, карбонадо, АСПК, природный поликристалл, алмазный порошок, спекание алмазного порошка, высокое статическое давление, подложка из твердого сплава, подложка из стали, металлическая обойма, твердосплавная обойма, низкое давление, связующий материал.*

Используя словарь S , в предметной области D определяются отношения ϕ для t_i с грамматическими характеристиками: глагол, существительное, причастие, деепричастие. Для этого используется информация из лексикографической базы данных. Для более точного выбора значения ϕ учитываются $C_R(t_i)$ и $C_L(t_i)$. Под правым и левым контекстом слова понимаются слова непосредственно связанные с t_i синтактико-семантическими отношениями. Так для словосочетаний *получаемые путем* (строки 9,14 примера), *получаемые методом* (строка 19) отношение ϕ принимает значение *способ получения*, для *получаемые при*, *получаемые в присутствии* (строки 9,10) – $\phi = \{\text{условие получения}\}$, для *синтезированные из* (строка 2) – $\phi = \{\text{состав}\}$. В словосочетании *получаемые на подложке* (строки 16,17) предлог *на* в правом контексте слова *получаемые* указывает на предложный падеж, т.е. отношения места, поэтому для *получаемые на* $\phi = \{\text{способ получения}\}$. Для слова *делятся* (строка 8) характеристикой (a_i) является использование его в описании отношений "класс-подкласс" (гипонимия) или "часть-целое" (меронимия). Конкретизируем значение *делятся*, используя выделенные синтактико-семантические отношения, $C_R(t_i)$ и $C_L(t_i)$. Найдя в правом контексте слова *делятся* связанные слова – *на группы* (для группы $a_i = \{\text{"класс-подкласс"}\}$), определяем конкретное значение $\phi = \{\text{"класс-подкласс"}\}$.

На основе сформированного списка терминов, уточненных значений ϕ и построенных деревьев зависимостей для предложений текста, определяются элементы множества D : ГРУППА

ПОЛИКРИСТАЛЛИЧЕСКИХ СВЕРХТВЕРДЫХ МАТЕРИАЛОВ, СПОСОБ ПОЛУЧЕНИЯ, УСЛОВИЕ ПОЛУЧЕНИЯ, СОСТАВ, СВОЙСТВА. В приведенном примере в множество D попали элементы, характеризующие иерархические и атрибутивные отношения. Это обусловлено видом ЕЯО – научная проза и значениями интерпретации отношения ϕ , определенными для слов из словаря S . Сформированный список элементов D уточняется экспертом предметной области.

Из построенных деревьев зависимостей фраз текста получим множество предметных констант и их семантические значения: ГРУППА ПОЛИКРИСТАЛЛИЧЕСКИХ СВЕРХТВЕРДЫХ МАТЕРИАЛОВ = {I группа, II группа, III группа, IV группа}; СПОСОБ ПОЛУЧЕНИЯ = {спекание алмазного порошка, превращения графита, металлическая обойма, твердосплавная обойма, подложка из твердого сплава, подложка из стали}; СВОЙСТВА = {высокая износостойкость, изотропия свойств}; УСЛОВИЕ ПОЛУЧЕНИЯ = {низкое давление, высокое давление, сплав-растворитель углерода, высокое статическое давление, высокая температура, связующий материал}; СОСТАВ = {неалмазный углерод}.

На следующем шаге заменим указательное местоимение *этой* (строка 10 примера), связанное со словом *группа*, на его значение в соответствии с алгоритмом, приведенным в [7]. Из предыдущего предложения (строка 9) можно получить однозначное значение данного местоимения – *I группа ПСТМ*.

В приведенном примере имеется синонимия терминов: ПСТМ, поликристаллические материалы на основе алмаза, поликристаллические сверхтвердые материалы, алмазные поликристаллы; трещиностойкость – ударная вязкость; баллас – АСБ; карбонадо – АСПК. Для задания отношения синонимии требуется просмотр списка возможных синонимов экспертом предметной области. Список возможных синонимов формируется автоматически, используя простые правила, например, аббревиатура в скобках может быть синонимом однословного или многословного термина, стоящего в предложении впереди аббревиатуры. После уточнения отношения синонимии любой из терминов синонимов может быть использован в дальнейших операциях с текстом. В строке 13 встречается неопределенное словосочетание *подобных типов*. Выделим данное словосочетание и предъявим его эксперту для уточнения значения слова *подобных*.

На подмножествах D определим отношения-предикаты: отношение «ГРУППА ПСТМ – СПОСОБ ПОЛУЧЕНИЯ» – $R_{г-сп} \subseteq \text{ГРУППА ПСТМ} \times \text{СПОСОБ ПОЛУЧЕНИЯ}$.

Семантическое значение $R_{г-сп}$ в данном примере следующее:

$$R_{г-сп} = \{(I \text{ группа ПСТМ, превращение графита}), \\ (II \text{ группа ПСТМ, спекание алмазного порошка}), \\ (III \text{ группа ПСТМ, подложка из твердого сплава, подложка из стали, металлическая обойма, твердосплавная обойма}), \\ (IV \text{ группа ПСТМ, спекание алмазного порошка})\}$$

Поверхностный формальный анализ полученных отношений показывает их семантическую неоднородность. Так для III группы ПСТМ в терминах, обозначающих способ получения, не содержатся отглагольные существительные, в отличие от других предметных констант СПОСОБА ПОЛУЧЕНИЯ. Это говорит о том, что или модель M неполна и необходимо пополнять сигнатуру предикатов Π , или пополнять отсутствующими понятиями множество D . Кроме того, такой формальный анализ позволяет обратить внимание эксперта на точность формулировок текста предметной области.

Отношение «ГРУППА ПСТМ – УСЛОВИЕ ПОЛУЧЕНИЯ» – $R_{г-уп} \subseteq \text{ГРУППА ПСТМ} \times \text{УСЛОВИЕ ПОЛУЧЕНИЯ}$

$$R_{г-уп} = \{(I \text{ группа ПСТМ, высокое давление, высокая температура, сплав-растворитель углерода}), \\ (II \text{ группа ПСТМ, высокое статическое давление, высокая температура}), \\ (III \text{ группа ПСТМ, высокое статическое давление, высокая температура}), \\ (IV \text{ группа ПСТМ, низкое давление, связующий материал})\}$$

Эти и другие отношения можно получить, исходя из области D , используя операции реляционной алгебры. Аналогичным способом можно получить и другие интересующие пользователя отношения-предикаты, т.е. полностью построить сигнатуру исходного множества предикатов Π на области D .

Покажем использование, полученных в результате автоматического анализа текста, множества предметных констант и отношений-предикатов для построения онтологии предметной области.

Практически в любой ПдО знания делятся на "статическую" и "динамическую" составляющие, а соответствующая онтология ПдО представляется онтологией объектов и онтологией процессов. Под онтологией объектов будем понимать кортеж множеств $O = \langle X, R, F \rangle$, где X – конечное множество концептов (понятий) заданной предметной области, R – конечное множество отношений между концептами X , F – конечное множество функций интерпретации, заданных на множествах X и/или R . Причем для приведенного примера достаточно рассмотреть частный случай, когда множество F тождественно множеству аксиом A , представляющих истинные высказывания о соответствующих понятиях X . Аналогичную структуру примем и для онтологии процессов ПдО.

Разделим список терминов, учитывая элементы множества D , на понятия-объекты, свойства понятий-объектов, понятия-процессы и свойства понятий-процессов. Результат такого деления приведен в Таблице 1.

Таблица 1.

Понятия-объекты	Свойства понятий-объектов	Понятия-процессы	Свойства понятий-процессов
1.1. Поликристаллический сверхтвёрдый материал	1. Свойства поликристаллических материалов	1.1. Способ получения ПСТМ 1.2. Процесс получения ПСТМ	1. Условие получения
1.2. Поликристаллический материал			2. Высокая температура
1.3. Алмазный поликристалл	2. Физико-механическое свойство	2. Процесс получения ПСТМ I группы	3. Высокое давление
1.4. ПСТМ			4. Высокое статическое давление
2. Температура	3. Высокая износостойкость	3. Процесс получения ПСТМ II группы	
3. Давление			
4. Алмаз	4. Изотропия свойств	4. Процесс получения ПСТМ III группы	
5. Алмазный порошок			
6. Природный поликристалл	5.1. Трещиностойкость	5. Процесс получения ПСТМ IV группы	
7. Неалмазный углерод			
8. Графит	5.2. Ударная вязкость	6. Спекание алмазного порошка	
9. I группа ПСТМ			
10. II группа ПСТМ			
11. III группа ПСТМ			
12. IV группа ПСТМ			
13.1. Баллас			
13.2. АСБ			
14.1. Карбонадо			
14.2. АСПК			
15. Подложка из твёрдого сплава			
16. Подложка из стали			
17. Твёрдосплавная обойма			
18. Металлическая обойма			
19. Сплав-растворитель углерода			
20. Связующий материал			

Множество A сформируем из свойств понятий-объектов и свойств понятий-процессов. Полученные аксиомы приведены ниже.

Аксиомы для понятий-объектов (П-О).

1. Для П-О 1: $A_1^1 = \langle \text{ПСТМ имеют физико-механические свойства} \rangle$;

$A_2^1 = \langle \text{ПСТМ имеют свойство высокой износостойкости} \rangle$;

$A_3^1 = \langle \text{Физико-механические свойства ПСТМ могут обладать характеристикой изотропии} \rangle$;

$A_4^1 = \langle \text{ПСТМ обладают свойством ударной вязкости со значением, большим чем значение у монокристаллов} \rangle$.

2. Для П-О (категории) 2: $A_1^2 = \langle \text{Температура может иметь неопределённое значение "высокая"} \rangle$.
3. Для П-О (категории) 3: $A_1^3 = \langle \text{Давление может иметь неопределённое значение "высокое"} \rangle$;
 $A_2^3 = \langle \text{Давление может иметь неопределённое значение "высокое статическое"} \rangle$;
 $A_3^3 = \langle \text{Давление может иметь неопределённое значение "низкое"} \rangle$.

Аксиомы для понятий-процессов (П-П).

1. Для П-П 2: $A_1^2 = \langle \text{Процесс получения ПСТМ I группы протекает при высокой температуре} \rangle$;
 $A_2^2 = \langle \text{Процесс получения ПСТМ I группы протекает при высоком давлении} \rangle$.
2. Для П-П 3: $A_1^3 = \langle \text{Процесс получения ПСТМ II группы протекает при высокой температуре} \rangle$;
 $A_2^3 = \langle \text{Процесс получения ПСТМ II группы протекает при высоком статическом давлении} \rangle$.
3. Для П-П 4: $A_1^4 = \langle \text{Процесс получения ПСТМ III группы протекает при высокой температуре} \rangle$;
 $A_2^4 = \langle \text{Процесс получения ПСТМ III группы протекает при высоком статическом давлении} \rangle$.
4. Для П-П 5: $A_1^5 = \langle \text{Процесс получения ПСТМ IV группы протекает при низком давлении} \rangle$.

Выделенные понятия-объекты, понятия-процессы, отношения ϕ и R могут быть представлены как фрагмент онтологии, построенной на основе анализа ЕЯО. На рисунке 1 представлен фрагмент онтологии ПдО "Сверхтвёрдые материалы", а точнее – её раздела "Поликристаллические сверхтвёрдые материалы". Он построен автоматизированным способом на основе знаний, содержащихся в приведенном выше ЕЯО.

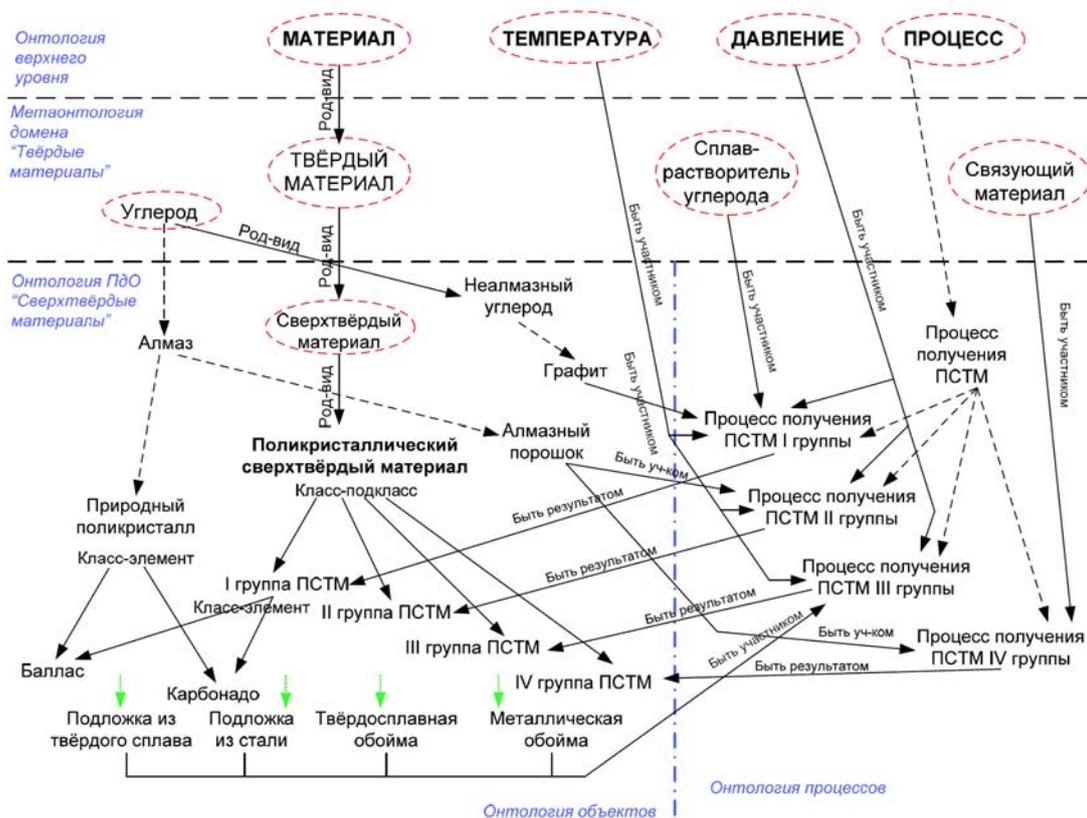


Рисунок 1. Фрагмент онтологии ПдО "Сверхтвёрдые материалы"

Процесс построения онтологии ПдО (в том числе и автоматизированным способом) предполагает этап начального развития онтологии. На этом этапе инженер по знаниям, совместно с экспертом в данной

ПдО, определяет список категорий онтологии верхнего уровня и метаонтологии домена "Твёрдые материалы", которые непосредственно связаны с концептами ПдО отношениями "выше-ниже". Далее эти категории структурируются и составляют начальное развитие (совместно с базовыми концептами ПдО) онтологического графа (ОГ). На рисунке вершины ОГ, соответствующие категориям, заключены в эллипсы с пунктирными линиями. Ребра ОГ, выполненные пунктирными линиями, в фрагменте текста в явном виде не указаны и вводятся как наиболее вероятные инженером по знаниям. При последующей обработке других текстов их следует уточнить (подтвердить или опровергнуть). На Рис.1 ОГ ПдО разделён на онтологии объектов и процессов штрихпунктирной линией.

Выводы

Данная работа посвящена проектированию автоматизированных систем приобретения и обработки знаний с применением онтологий предметных областей. Полученные в результате автоматического анализа текста множества предметных констант, построенные отношения-предикаты являются исходными данными для построения онтологий предметной области. Формальное представление ЕЯО позволяет облегчить работу эксперта по составлению или уточнению онтологии. Дальнейшее направление исследований связано с уточнением семантики языка логического типа для представления знаний с целью выполнения различных операций логического вывода (получения новых знаний) на основе анализа ЕЯО и использованием в определении семантических отношений результатов работы системы анализа ЕЯО на основе вхождения слов[8].

Благодарности

Статья частично финансирована из проекта **ITHEA XXI** Института Информационных теории и Приложений FOI ITHEA и Консорциума FOI Bulgaria (www.ithea.org, www.foibg.com).

Литература

1. Кулик Б.А. Логика естественных рассуждений. С.-П.: Невский диалект.-2001.-127 с.
2. Рубашкин В.Ш. Представление и анализ смысла в интеллектуальных информационных системах.-М.:Наука.-1989.-191с.
3. Тейз А., Грибомон П., Юлен Г. И др. Логический подход к искусственному интеллекту: От модальной логики к логике баз данных.-М.: Мир.-1998.-492с.
4. Широков В.А. Феноменологія лексикографічних систем. К.:Наукова думка.-2004.-327с.
5. Апресян Ю.Д. и др. Лингвистический процессор для сложных информационных систем.-М.: Наука.-1992.-287с.
6. Новые материалы. Колл. авторов. Под научной редакцией Ю.С. Карабасова. – М.: МИСИС, 2002 - 736 с.
7. Палагін О.В., Світла С.Ю., Петренко М.Г., Величко В.Ю. Про один підхід до аналізу та розуміння природномовних об'єктів. Комп'ютерні засоби, мережі та системи. -2008, №7. с.128-137.
8. Палагин А.В., Кривый С.Л., Бибииков Д.С. Обработка предложений естественного языка с использованием словарей и частоты появления слов 10 стр. (в печати).

Информация об авторах

Палагин Александр Васильевич – Ин-т кибернетики им. В.М. Глушкова НАН Украины, Киев-187 ГСП, 03680, просп. акад. Глушкова, 40, e-mail: palagin_a@ukr.net

Кривый Сергей Лукьянович – Киевский национальный университет им. Т.Г. Шевченко, Киев, ГСП, 01601, ул. Владимирская, 64, e-mail: krivoi@i.com.ua

Величко Виталий Юрьевич – Ин-т кибернетики им. В.М. Глушкова НАН Украины, Киев-187 ГСП, 03680, просп. акад. Глушкова, 40, e-mail: vitaly@aduis.kiev.ua

Петренко Николай Григорьевич – Ин-т кибернетики им. В.М. Глушкова НАН Украины, Киев-187 ГСП, 03680, просп. акад. Глушкова, 40, e-mail: petrnq@ukr.net