

Тема №1 Краткое введение в ОНТОЛОГИЧЕСКИЙ анализ

1. Основные определения
2. Примеры онтологий
3. Решаемые задачи
4. Построение онтологий
5. Типы онтологий

Основные определения

Два значения **онтологии**:

- Философская дисциплина изучает наиболее общие характеристики бытия и сущностей
- Онтология – артефакт, структура, описывающая значения элементов некоторой системы

Основные определения

Онтология как артефакт:

Неформально, онтология представляет собой некоторое описание взгляда на мир применительно к конкретной области интересов.

Это описание состоит из терминов и правил использования этих терминов, ограничивающих их значения в рамках конкретной области.

Основные определения

На формальном уровне, онтология это система, состоящая из **набора понятий** и **набора утверждений** об этих понятиях, на основе которых можно строить классы, объекты, отношения, функции и теории.

Основные компоненты:

- Классы или понятия
- Отношения
- Функции
- Аксиомы
- Примеры

Основные определения

Онтология – спецификация концептуализации (Gruber)

Концептуализация – структура реальности, независимо от

- Словаря
- Конкретной ситуации

Пример:

Кубики на столе: концептуализация - набор возможных положений, но не конкретное расположение.

Примеры онтологий

Онтологией могут быть:

- Глоссарий
- Простая таксономия
- Тезаурус
- Понятийная структура с произвольным набором отношений
- Структура с аксиоматикой

Примеры онтологий

Языки для описания онтологий:

- Ontobroker
- CycL
- Description Logics
- RDF/RDFS
- OWL

Примеры онтологий

Онтология СҮС

- Самый амбициозный проект
- Начат в 1984
- 1 млн. утверждений “common sense”
- Микротеории: пространство, время, причинность
- Онтология 3 тысяч понятий верхнего уровня – в открытом доступе
- www.cyc.com

Примеры онтологий

Лингвистические онтологии:

The main characteristic of this kind of ontologies is that they are bound to the semantics of grammatical units (words, nominal groups, etc)

Основной источник понятий в онтологии
— значения языковых единиц

Примеры:

WordNet, Mikrokosmos, Sensus, PyTез

Примеры онтологий

WordNet

Реляционное описание лексики английского языка

Иерархическая сеть понятий (synset)

Каждое слово относится к одному
или нескольким понятиям

Автор: George Miller
(50-е годы статья «Магическое число 7»)

Версия 1.6:
95 тысяч понятий, около 130 тысяч слов и понятий

Примеры онтологий

EuroWordNet

Интерлингва:
английский WordNet

Первоначально:
испанский, итальянский, голландский

Далее:
немецкий, французский, чешский, эстонский

Известны попытки создать свои структурные ресурсы на базе WordNet: японский, болгарский, румынский, шведский и др.

Решаемые задачи

Лингвистические онтологии и информационный поиск:

- Электронные коллекции разнообразных текстов
- Возможности систем автоматической обработки текста для анализа релевантности контекста ограничены
- Нет возможности подробно проанализировать контекст упоминания понятия в тексте.
- Онтологии специального типа?

Решаемые задачи

Онтологии и вопросно-ответные системы:

- Система ищет в сверхбольшой текстовой коллекции
- Сравнение систем в соревновании TREC и CLEF
- Практическая актуальность: поиск в Интернет не по краткому запросу, а по развернутому вопросу

Постановка задачи:

60-е годы: поиск в специальных базах знаний

Сейчас: поиск в громадных текстовых массивах

Построение онтологий

- Тексты
- Набор словосочетаний: автоматическое извлечение терминов
- Выделение понятия
- Отношения между понятиями:
 - Извлечение из текстов по шаблонам
 - Статистические методы
 - Методы на основе синтаксической структуры

Построение онтологий

Как использовать созданные онтологии?

- Слияние онтологий
- Использование общих онтологий для эффективного создания онтологий в конкретных предметных областях
- Semantic web: одна (или несколько) онтология верхнего уровня, к которой достраиваются специализированные онтологии

Типы онтологий

Типы
онтологий
по
О. Лассилла,
Д. МакГинесс



Типы онтологий

Классификация по цели создания



Типы онтологий

Классификация по содержанию



Типы онтологий

Онтологии представления знаний

The Frame Ontology and the OKBC Ontology
(<http://ontolingua.stanford.edu>)

Model and Syntax Specification. W3C Recommendation.
<http://www.w3.org/TR/REC-rdf-syntax/>

RDF and RDF Schema knowledge representation ontologies
(<http://www.w3.org/1999/02/22-rdf-syntax-ns>
<http://www.w3.org/2000/01/rdf-schema>)

OIL knowledge representation ontology
(<http://www.ontoknowledge.org/oil/rdf-schema/2000/11/10-oil-standard>)

DAML+OIL knowledge representation ontology
(<http://www.daml.org/2001/03/daml+oil>)

Типы онтологий

Характеристики онтологий верхнего уровня

Набор основных отношений:

- Иерархические
 - класс-подкласс
 - часть-целое
- Ассоциативные

Типичные концепты и принципы разделения:

- Сущность
- Явление
- Объект
- Процесс
- Роль

Нацеленность на многократное повторное использование
онтологии

Типы онтологий

Характеристики онтологий предметных областей

Существенные отличия от ОВУ:

- Охват конкретной области знаний (авиация, медицина, культура),
- Круг решаемых задач и вопросов, на которые онтология отвечает, ограничен выбранной областью,
- Большой по сравнению с ОВУ объем онтологии,
- Наличие отношений специфичных для конкретной области.

Повторное использование только в рамках предметной области

Типы онтологий

Онтологии электронной коммерции

The United Nations Standard Products and Services Codes (UNSPSC) (<http://www.unspsc.org/>)

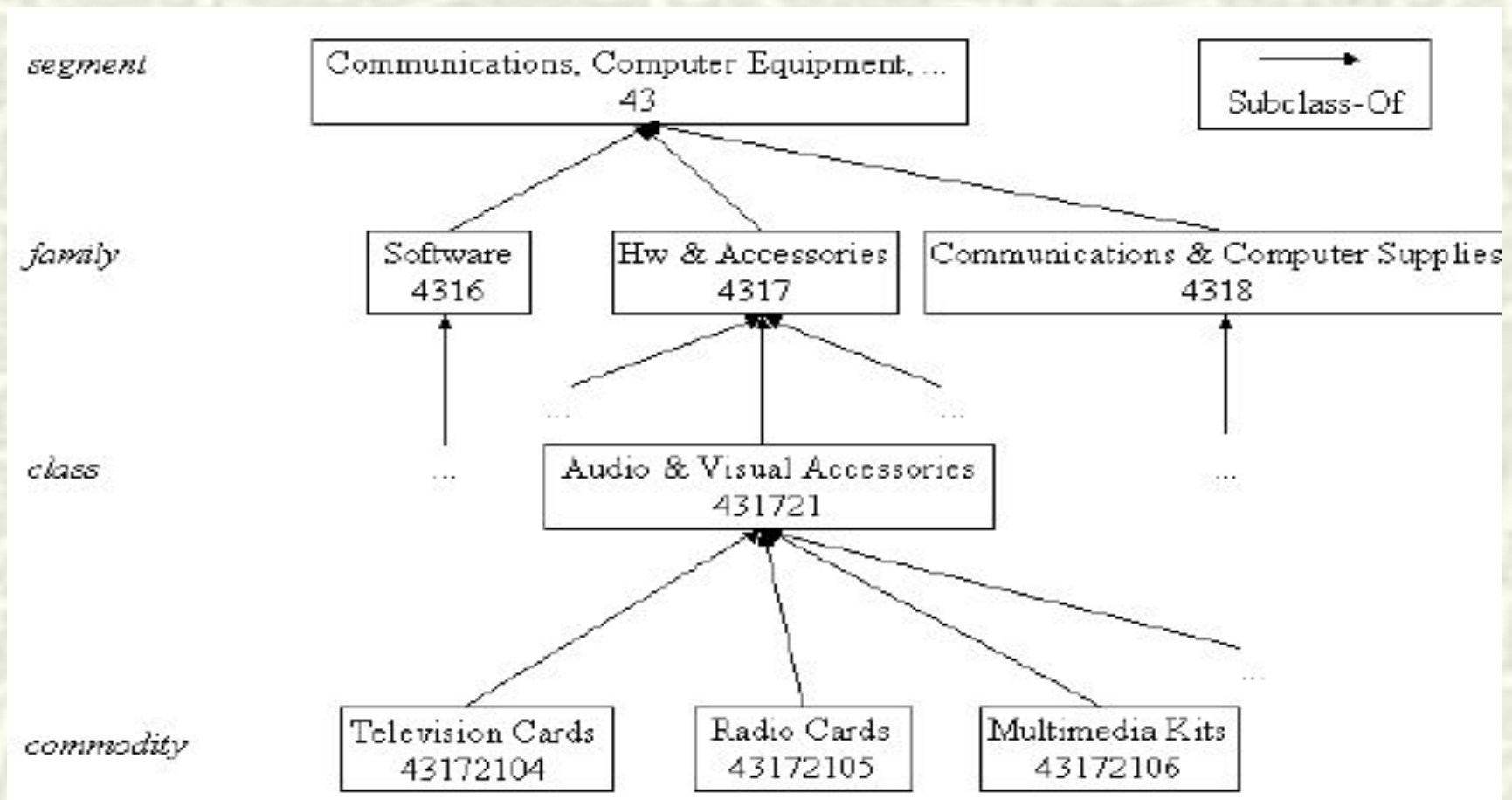
NAICS (North American Industry Classification System)
(<http://www.census.gov/epcd/www/naics.html>)

SCTG (Standard Classification of Transported Goods)
(<http://www.statcan.ca/english/Subjects/Standard/sctg/sctg-menu.htm>)
[E-cl@ss](http://www.e-class.de/) (<http://www.e-class.de/>)

RosettaNet (<http://www.rosettanet.org>)

Типы онтологий

Онтологии электронной коммерции



Типы онтологий

Характеристики прикладных онтологий

Направленность на решение конкретной задачи, обеспечение работы некоторого приложения.

Круг вопросов, на которые онтология отвечает:

- Задается перед ее созданием

- Влияет на процесс концептуализации

- Используется для проверки “компетентности” онтологии

Типы онтологий

Лингвистические онтологии

WordNet (<http://www.hum.uva.nl/~ewn/gwa.htm>)

EuroWordNet (<http://www.hum.uva.nl/~ewn/>)

The Generalized Upper Model

(<http://www.darmstadt.gmd.de/publish/komet/gen-um/newUM.html>)

The Mikrokosmos ontology (<http://crl.nmsu.edu/mikro> [user and password are required])

SENSUS

(<http://www.isi.edu/natural-language/projects/ONTOLOGIES.html>)

Типы онтологий

Характеристики лексических онтологий

Отличительное свойство: фиксация в одном ресурсе (лексикализованных) понятий (слов) вместе с их языковыми свойствами.

Круг задач тесно связан с NLP (Natural Language Processing)

Типы онтологий

Характеристики лексических онтологий

Основной набор отношений:

- Синонимия
- Гипонимия
- Меронимия

Большие и сверхбольшие тезаурусы:

- WordNet
- Тезаурус по общественно-политической тематике