

Intelligent Internet Technologies

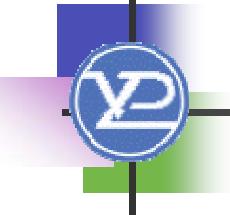
Lecture 17.

Introduction to Ontologies

Alexandra V. Vitko

KNURE, AI Department, alexandra_vitko@yahoo.com





Introduction

- Digitally stored information sources are growing
- Accessing, finding and summarizing information becomes more and more difficult
- Main problem: Gap between

meaning
of information

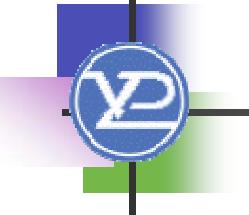
vs.

actual storage
of information



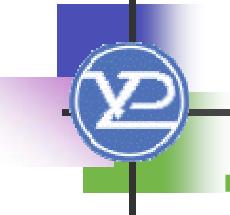
Motivation

- WWW is the largest database of the world:
However: Search engines are restricted to simple keyword based techniques. Interpretation of information contained in web documents is left to the human user!
- Information Integration:
The integration of heterogeneous data (e.g. for building a data warehouse) is a challenging task on account of synonyms and homonyms
- Knowledge Management
„People can't share knowledge if they do not speak a common language“ [*T. Davenport, Working Knowledge*]



Semantic Web

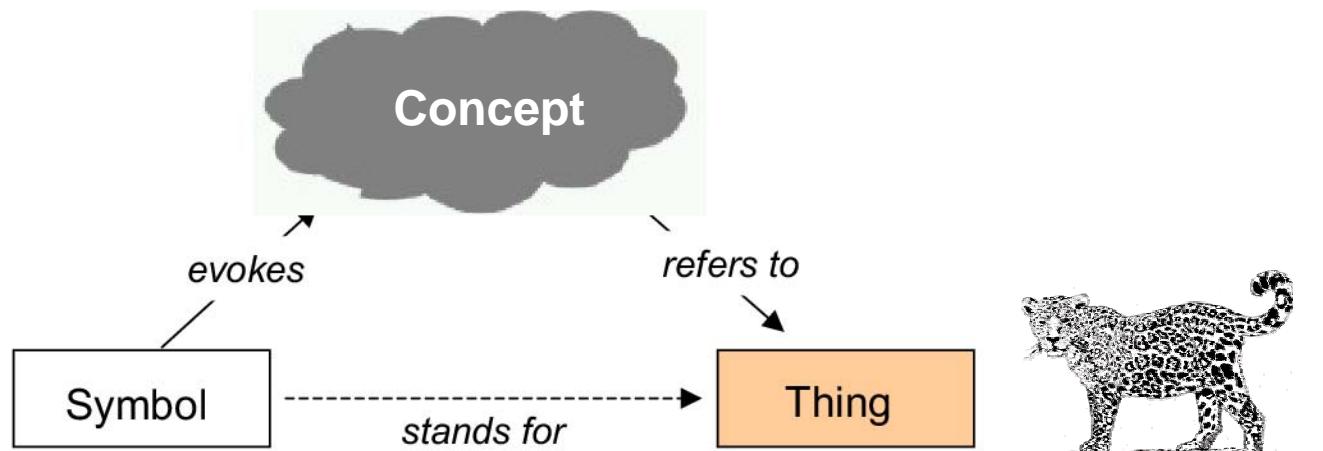
- **Ontologies** are considered as a key component to build the Semantic Web.
- **Semantic Web** – a new vision of the Web.
- The Semantic Web is the new challenge of web using. Its **goal** is to bring structure and semantics to the content of Web pages, creating an environment where software agents, roaming from page to page, can carry out sophisticated tasks for users.



Origin and History

- Humans require words (or at least symbols) to communicate efficiently. The mapping of words to things is only indirect possible. We do it by creating concepts that refer to things.
- The relation between symbols and things has been described in the form of the meaning triangle:

“Jaguar”

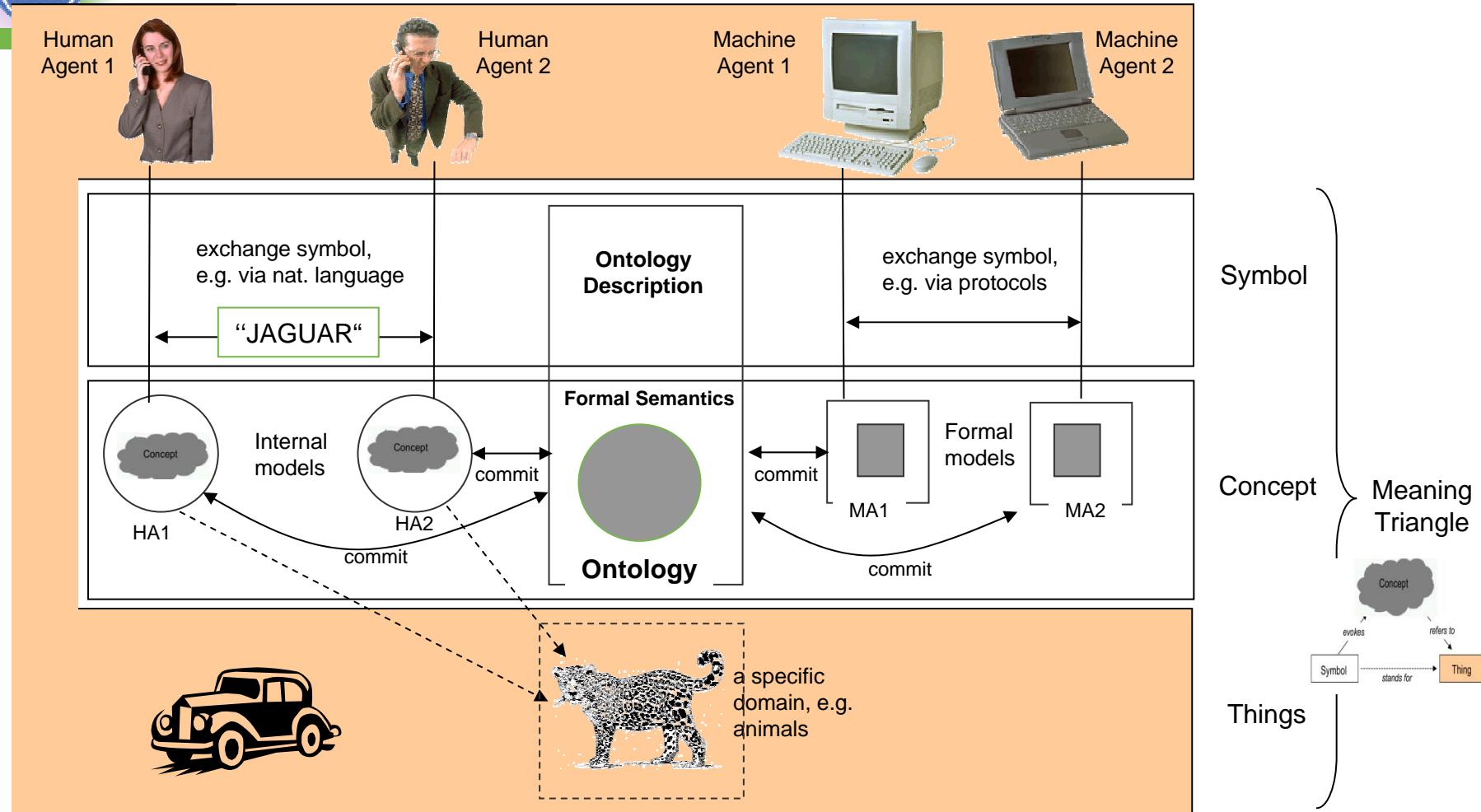


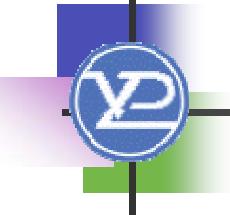
[Ogden, Richards, 1923]





Human and machine communication

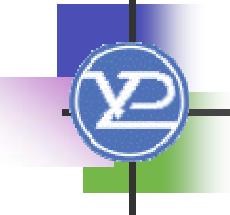




Application areas

Ontologies have become a popular research topic in various communities:

- Knowledge engineering
- Natural language processing
- Knowledge management
- Information integration
- Information retrieval
- Digital libraries
- WWW applications (Semantic Web)
- Electronic commerce
- Multi-agent systems



What is an Ontology?

- **Ontology**
philosophical discipline, branch of philosophy that
deals with the nature and the organization of reality

- **Science of Being (Aristotle, Metaphysics, IV, 1)**

- **Tries to answer the questions:**
What is being?
What are the features common to all beings?



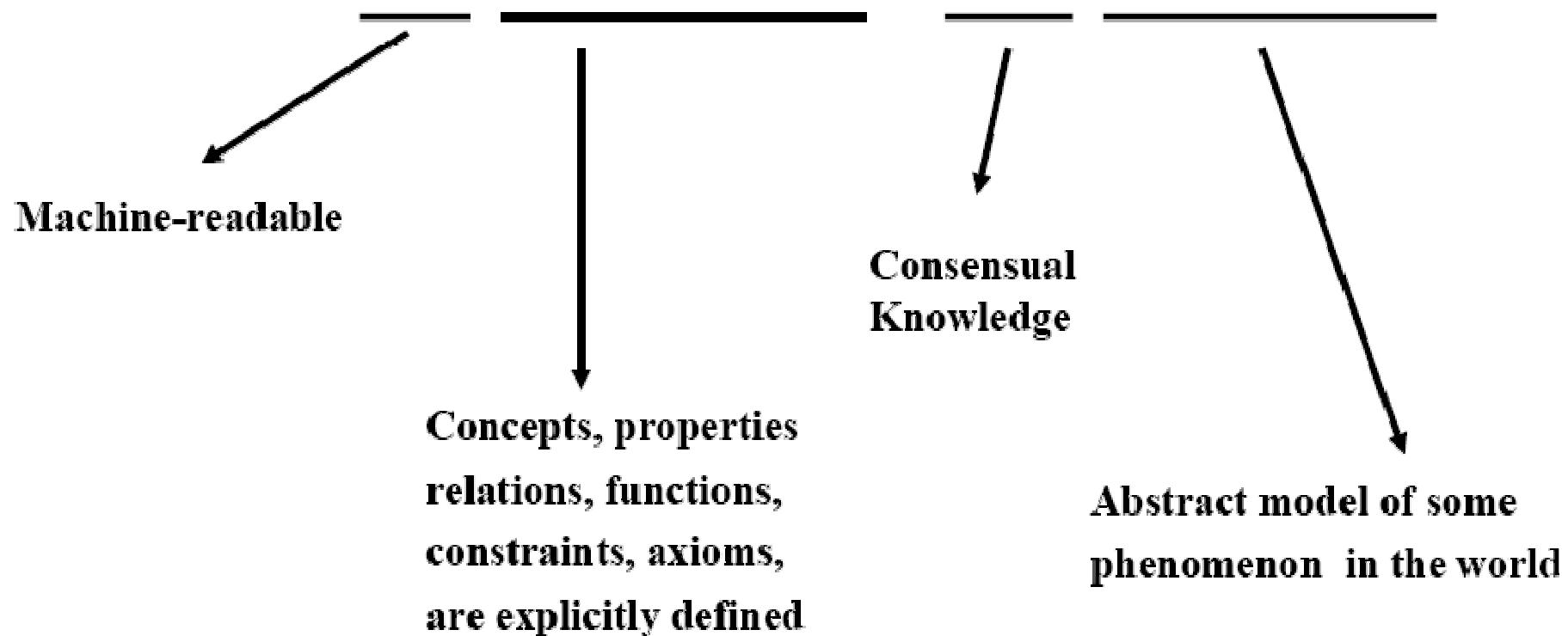
What are ontologies in computer science?

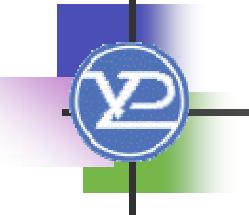
- An ontology is an explicit specification of a conceptualization [Gruber, 93]
- An ontology is a shared understanding of some domain of interest. [Uschold, Gruninger, 96]
- There are many definitions, an ontology (in our sense) is
 - a sound semantic basis to define meaning
 - general logical theory constituted by a vocabulary and set of statements in some logical language
 - foundation for communication between human and machine agents



Definition of Ontology

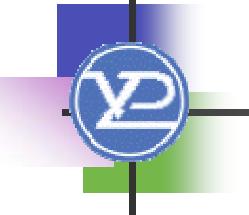
“An ontology is a formal, explicit specification of a shared conceptualization”





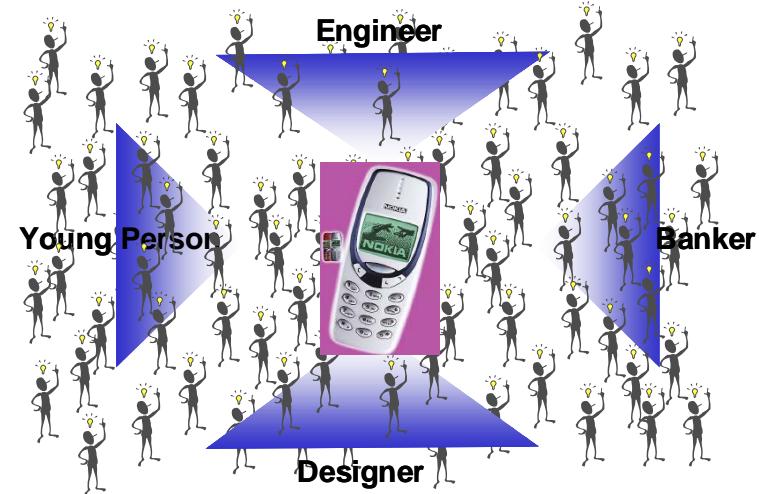
Why develop an ontology?

- To share common understanding of the structure of information among people or software agents.
- To enable reuse of domain knowledge
- To make domain assumptions explicit
 - Easier to change domain assumptions
 - Easier to understand and update legacy data
- To separate domain knowledge from the operational knowledge
- To analyze domain knowledge



Ontologies: made for sharing

- Interoperating resources, be it by people or systems, requires a consistent shared understanding of what the information contained means
- “*... people [and machines] can't share knowledge if they don't speak a common language*” [Davenport]

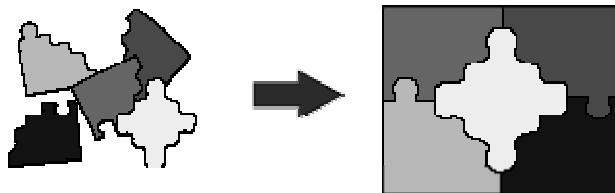


Disparate modeling paradigms, languages and software tools limit
=> Interoperability
=> Knowledge sharing & reuse

Reuse and Sharing

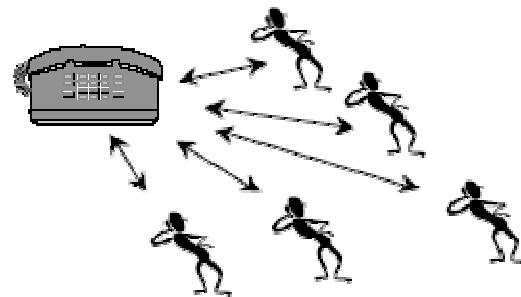
Reuse means to build new applications
assembling components

already built



Sharing is when different
applications use the same

resources



Advantages:

- Less money
- Less time
- Less resources

Sharing info ⇒ Sharing meaning

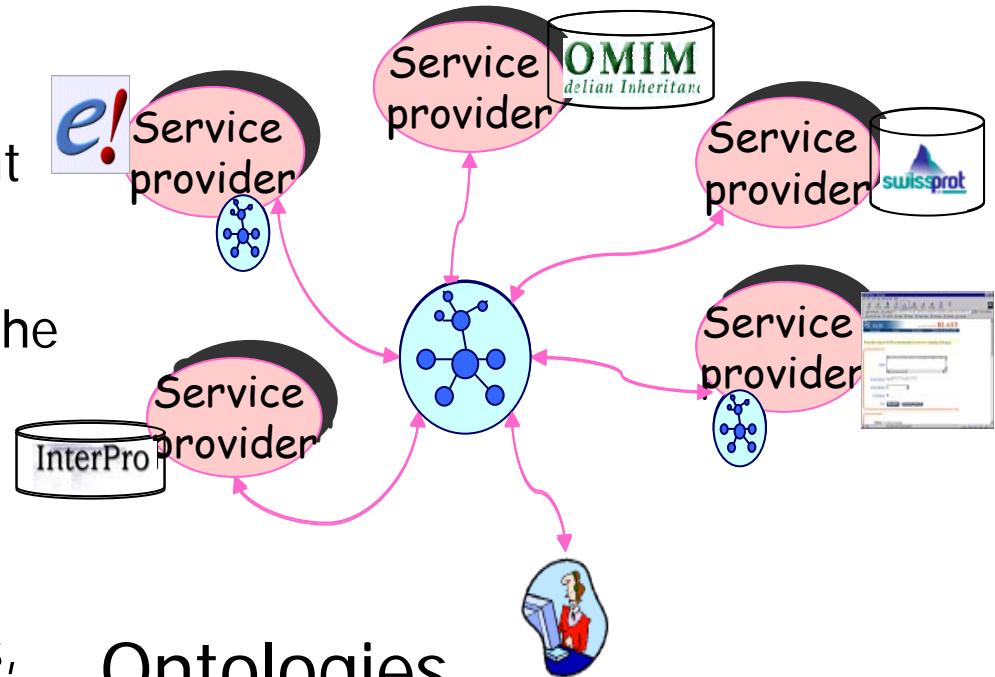


Metadata

- Data describing the content and meaning of resources and services.
- But everyone must speak the same language...

Terminologies

- Shared and common vocabularies
- For search engines, agents, curators, authors and users
- But everyone must mean the same thing...



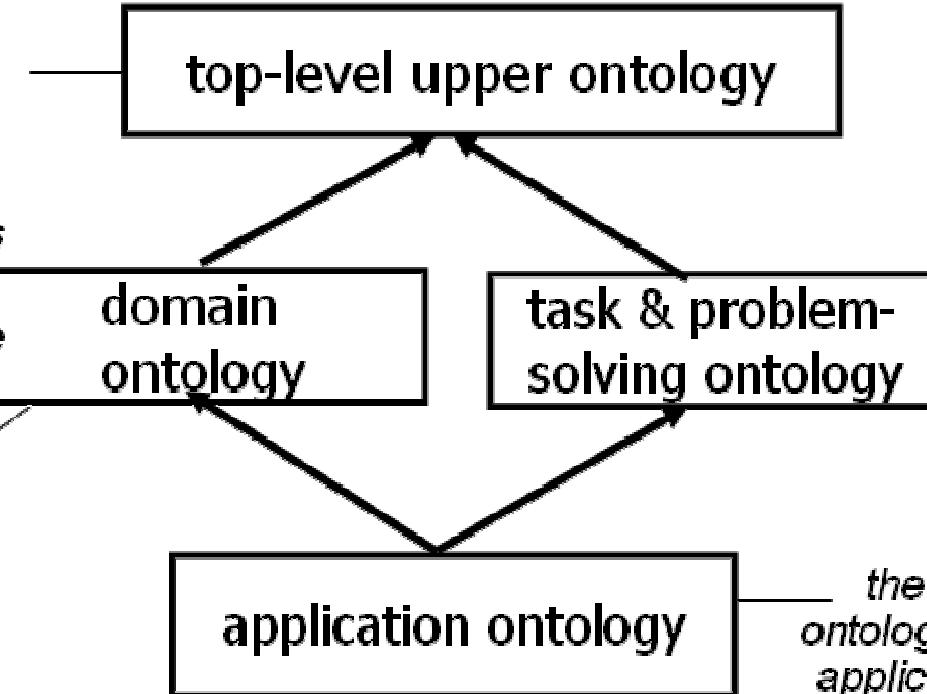
Ontologies

- Shared and common understanding of a domain
- Essential for search, exchange and discovery

Types of Ontologies [Guarino, 98]

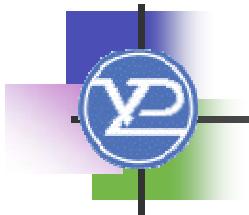
 describe very general concepts like space, time, event, which are independent of a particular problem or domain. It seems reasonable to have unified top-level ontologies for large communities of users.

describe the vocabulary related to a generic domain by specializing the concepts introduced in the top-level ontology.



describe the vocabulary related to a generic task or activity by specializing the top-level ontologies.

the most specific ontologies. Concepts in application ontologies often correspond to roles played by domain entities while performing a certain activity.



Specific ontologies

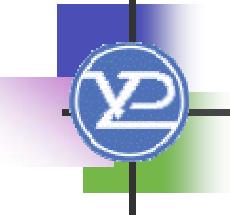
- Domain-oriented
 - Domain-specific
 - Medicine => cardiology => rhythm disorders
 - Domain generalizations
 - components, organs, documents, gene function
- Task-oriented
 - task specific
 - configuration design, instruction, planning, annotation analysis
 - task generalisations
 - problem solving methods



Upper Ontologies

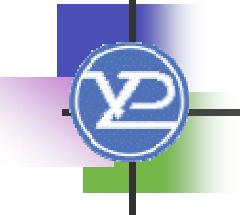
- A.k.a. core, generic or reference
- Common high level concepts
- “Physical”, “Abstract”, “Structure”, “Substance”
- Useful for ontology re-use
- Important when generating or analysing natural language expressions

- Top Level ontologies
- WordNet
- EuroWordNet
- CyC
- SENSUS
- Sowa Top Level
- GUM
- Etc...



What Is An Ontology

- An ontology is an explicit description of a domain:
 - concepts
 - properties and attributes of concepts
 - constraints on properties and attributes
 - Individuals (*often, but not always*)
- An ontology defines
 - a common vocabulary
 - a shared understanding



Components of an Ontology

Concepts are organized in taxonomies

Relations

$R: C_1 \times C_2 \times \dots \times C_{n-1} \times C_n$

Subclass-of: Concept1 x Concept2
Connected to: Component1 x Component2

Functions

$F: C_1 \times C_2 \times \dots \times C_{n-1} \rightarrow C_n$

Mother-of: Person \rightarrow Women
Price of a used car: Model x Year x Kilometers \rightarrow Price

Instances

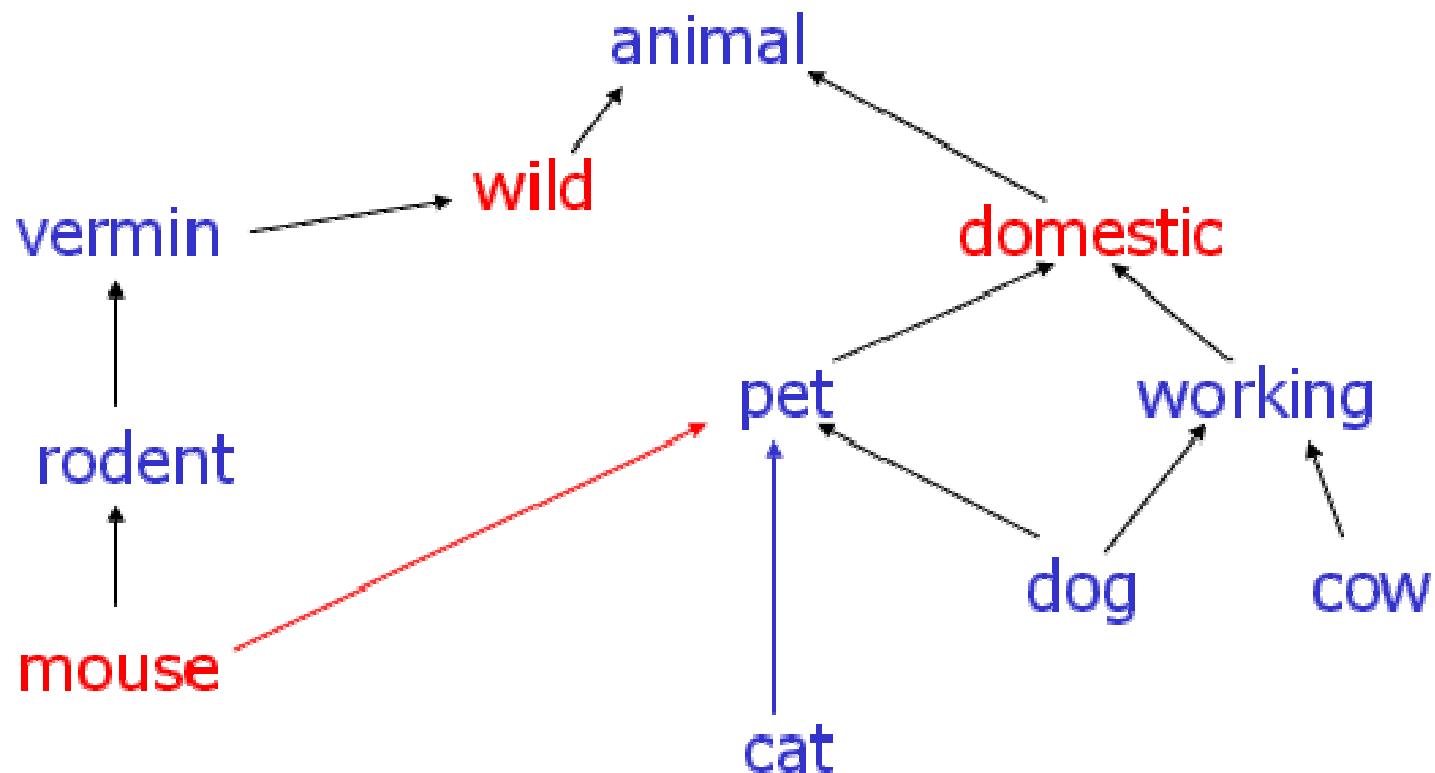
Elements

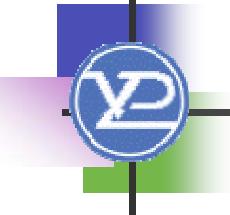
Axioms

Sentences which are always true

Taxonomy remark 1

- The world is not a tree, it's a lattice



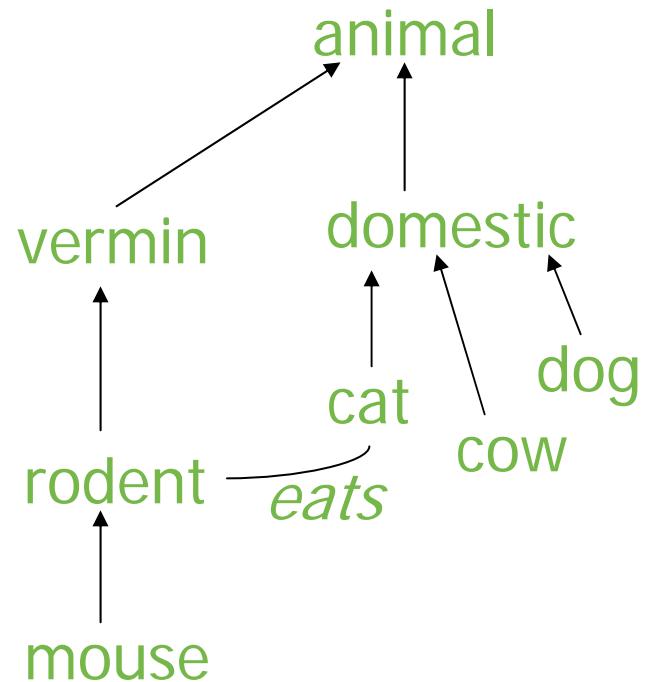


Taxonomy remark 2

- What does the taxonomy mean?
- Concept A is a parent of concept B if every instance of B is also an instance of A
- Superset/subset

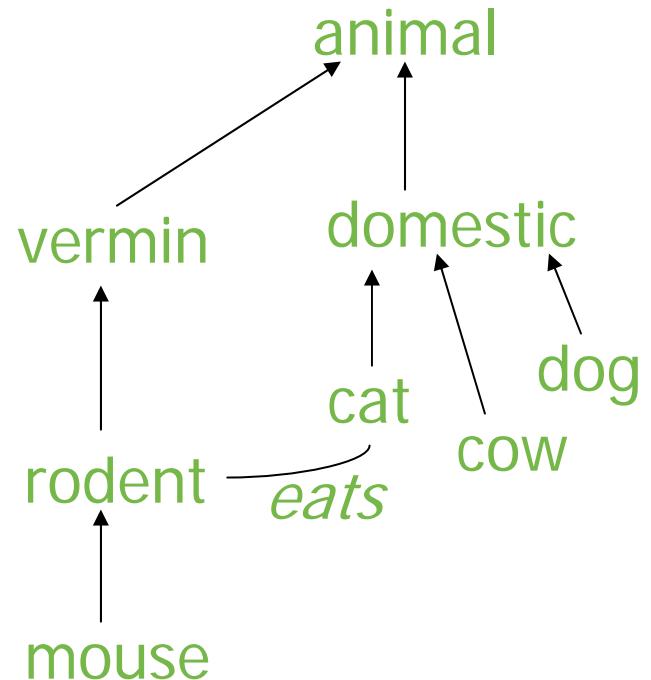
An explicit description of a domain

- Concepts (class, set, type, predicate)
 - cat, mouse...
- Properties of concepts and relationships between them (slot)
 - Taxonomy: generalisation ordering among concepts isA, partOf, subProcess
 - Relationship, Role or Attribute: functionOf, hasActivity location, eats, size



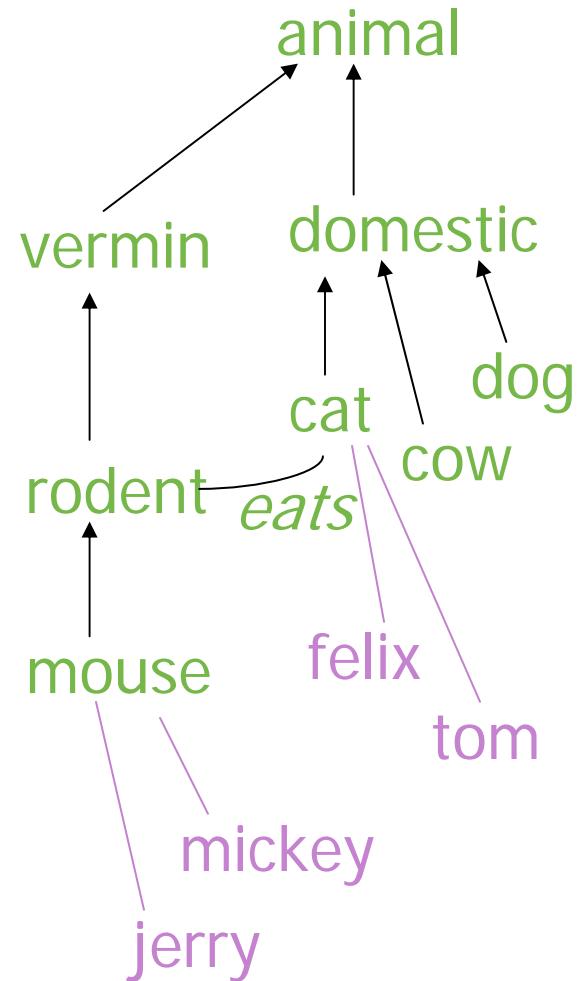
An explicit description of a domain

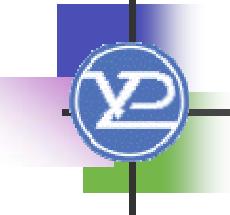
- Constraints or axioms on properties and concepts:
 - value: integer
 - domain: cat
 - cardinality: at most 1
 - range: $0 \leq X \leq 100$
 - oligonucleotides < 20 base pairs
 - cows are larger than dogs
 - cats cannot eat only vegetation
 - cats and dogs are disjoint
- Values or concrete domains
 - integer, strings
 - 20, tryptophan-synthetase



An explicit description of a domain

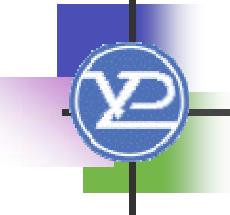
- Individuals or Instances
 - tom, jerry...
- Nominals
 - Concepts that cannot have instances
 - Instances that are used in conceptual definitions
 - ItalianDog = Dog bornIn Italy
- Instances
 - An ontology = concepts+properties+axioms+values +nominals
 - A knowledge base = ontology+instances





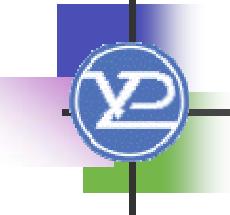
Ontology Examples

- Taxonomies on the Web
 - Yahoo! categories
- Catalogs for on-line shopping
 - Amazon.com product catalog
- Domain-specific standard terminology
 - Unified Medical Language System (UMLS)
 - UNSPSC - terminology for products and services
- General ontologies: Wordnet, Cyc, ...



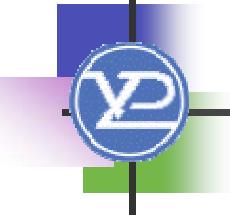
Ontology Merging and Conjoining

- Semantic Web will contain many independently produced ontologies
- Ontology users will typically assemble multiple modular ontologies
- Independently developed ontologies will frequently –
 - Express overlapping knowledge in a common domain
 - Use differing representations and vocabularies
- For such ontologies to be used together –
Their representational differences must be reconciled
- I.e., the ontologies need to be **merged** or **conjoined**
 - Merging produces a new ontology
 - Conjoining produces statements of relationships among terms from multiple ontologies



Standard Ontologies

- Will there be one standard ontology?
 - Almost certainly not
- Will there be some standard ontologies?
 - Almost certainly
 - Some professional domains have standardized vocabularies
 - E.g., medicine, electrical engineering
- Barriers to standardization
 - Obtaining consensus is difficult and slow
 - “Standard” vocabularies may be logically inconsistent with others
 - Many applications require customized vocabularies
 - With application-specific simplifying assumptions and abstractions



Current ontology standardization initiatives

- SUO (SUO consortium proposal) <http://suo.ieee.org/>
- Global WordNet Consortium
- ISO SC4
- eCommerce standards (UCEC, ebXML,...)
- Cultural repositories standards (Harmony, CIDOC)
- CEN/ISSS EC WG (MULECO)
- DAML (especially DAML-S) <http://www.daml.org/>
- W3C Web Ontology Working Group
 - <http://www.w3.org/2001/sw/WebOnt/>
- Projects
 - OntoWeb <http://www.ontoweb.org/>
 - WonderWeb <http://wonderweb.semanticweb.org/>