Ontologies - A Very General Overview

Marian F Ursu
Department of Computing, Goldsmiths College
February 2003

Overview

• context and motivation for ontologies
• definitions
• ontologies for humans, ontologies for machines
• examples: WordNet and CYC; investigation
• stating/expressing an ontology
• reasoning within an ontology
• levels of abstraction; types of ontologies
• ontologies and domain descriptions
• conjecture: ontologies are necessarily required
• ontologies, meta-data, meta-data standard, RDF

Ontologies - Motivation

• truism: many agents can understand each other only
  if they share a common language (syntax and semantics)

Ontology - Definitions

• a shared and common/unitary understanding of a
domain that can be communicated between people
and heterogeneous software systems (Fensel, 2001)

• a shared and common language (or a standard
language) for a specific domain
  • vocabulary
  • syntax / grammar

Preamble

• the meaning of the term "ontology" in computing is different its meaning in philosophy
  • there is a connection, though
• initially, understand the term "ontology" in computing as a "standard vocabulary"
Ontologies - Examples

• in medicine …
• in computing …
• in the financial sector …
• in the context of local governmental councils …
• in museums …
• level of abstraction

Ontologies – Who Uses Them?

• humans
  • informal
• software agents
  • machine-readable (interpretable) - formal

Ontologies - Examples

• WordNet
  • informal
  • thesaurus for natural language terms, explained in natural language
• CYC
  • formal
  • theories for common sense knowledge

From Informal to Formal

• there is a continuum between informal and formal
  • informally (sic!): the better the quality of the reasoning/processing that a software agent can perform within an ontology, the more formal the ontology is
  • the degree of formality of an ontology increases with its expressive power regarded from the point of view of software agents

highly informal  O1  O2  rigorously formal

WordNet

• explore and discuss its features
CYC - Applications

• applications currently available or in development
  - Integration of Heterogeneous Databases
  - Knowledge-Enhanced Retrieval of Captioned Information
  - Guided Integration of Structured Terminology (GIST)
  - Distributed AI
  - WWW Information Retrieval
• potential applications
  - Online brokering of goods and services
  - "Smart" interfaces
  - Intelligent character simulation for games
  - Enhanced virtual reality
  - Improved machine translation
  - Improved speech recognition
  - Sophisticated user modelling
  - Semiotic data mining

Informal Ontology - Language

• example: WordNet (simplified)
  - some pre-defined categories:
    • Noun: name; explanation
    • Verb: name; explanation
    • Adjective: name; explanation
    • Adverb: name; explanation
  - some pre-defined relationships
    • Synonym
    • Hapenym (X is a kind of …)
    • Hyponym (… is a kind of X)
• pre-defined terms are “understood” by the system
WordNet Representation - Relational DB

<table>
<thead>
<tr>
<th>Name</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>seminar</td>
<td>any meeting for an exchange of ideas</td>
</tr>
<tr>
<td>seminar</td>
<td>a course offered for a small group of students</td>
</tr>
<tr>
<td>conference</td>
<td>a pre-arranged meeting for consultation or discussion</td>
</tr>
<tr>
<td>meeting</td>
<td></td>
</tr>
<tr>
<td>gathering</td>
<td></td>
</tr>
<tr>
<td>social group</td>
<td></td>
</tr>
<tr>
<td>group</td>
<td></td>
</tr>
</tbody>
</table>

hypernyms

<table>
<thead>
<tr>
<th>Name</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>seminar</td>
<td></td>
</tr>
<tr>
<td>conference</td>
<td></td>
</tr>
<tr>
<td>meeting</td>
<td></td>
</tr>
<tr>
<td>gathering</td>
<td></td>
</tr>
<tr>
<td>social group</td>
<td></td>
</tr>
<tr>
<td>group</td>
<td></td>
</tr>
</tbody>
</table>

nouns

<table>
<thead>
<tr>
<th>Name</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>seminar</td>
<td></td>
</tr>
<tr>
<td>conference</td>
<td></td>
</tr>
<tr>
<td>meeting</td>
<td></td>
</tr>
<tr>
<td>gathering</td>
<td></td>
</tr>
<tr>
<td>social group</td>
<td></td>
</tr>
<tr>
<td>group</td>
<td></td>
</tr>
</tbody>
</table>

synonyms

<table>
<thead>
<tr>
<th>Name</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>seminar</td>
<td></td>
</tr>
<tr>
<td>conference</td>
<td></td>
</tr>
<tr>
<td>meeting</td>
<td></td>
</tr>
<tr>
<td>gathering</td>
<td></td>
</tr>
<tr>
<td>social group</td>
<td></td>
</tr>
<tr>
<td>group</td>
<td></td>
</tr>
</tbody>
</table>

WordNet - Reasoning Capabilities

- find all the explanations for a given noun (name)
- find all the level one synonyms for a certain noun
- find all the level one hypernyms for a certain noun
- find all the hypernyms for a certain noun
- ...
- quite poor

Formal Ontology - Formal Language

- declarative
- a subset of natural language, but much better structured, to the extent that machines can make inferences — can reason — with its expressions
  - to infer/reason means to deduce new expressions from existing ones
  - domain categories and relationships cab be introduced by means of the language

Formal Ontology - Example

**FACTS**

- man(Danny)
- man(Tony)
- man(Marian)
- friend-of(Danny, Tony)
- friend-of(Marian, Danny)
- works-for(Lew-Council, Danny)
- works-for(Gold-Uni, Marian)

**RULES**

- person(x) IF man(x)
- person(x) IF woman(x)
- friend-of(x, y) AND friend-of(y, z) IF
- friend-of(x, y) AND friend-of(y, z) AND
- colleague-of(x, y) IF works-for(x, z) AND
- works-for(y, z)

**INFER**

- FROM fact(a)
- AND rule(b IF a)
- INFERENCE

Formal Ontology - Reasoning Capabilities

- find all the men
- find all the persons
- find all the friends of ‘marian’
- find all the pairs of friends
- find all the friends of ‘marian’ who work for Lewisham Council
- are there any colleagues who do not work for the same institution?
- ...
- quite rich

Formal Ontology - Definition

- a formal explicit specification of a shared and consensual conceptualisation (Fensel, 2001)
  - conceptualisation: abstract model of a part of ‘the world’
  - explicit: the elements of the model and the constraints on their use are explicitly described (as opposed to being implicit in someone’s head or embodied in a piece of software)
  - shared and consensual: used unitarily by a group of users
  - formal: machine-readable/interpretable (accompanied by inference rules)
Ontologies vs Databases - Differences

• an ontology is usually syntactically and semantically richer than a database
  • this aspect relates to technology
• an ontology must be shared and consensual terminology
  • this aspect is beyond technology - it relates to the dynamics and interactions between social systems

Ontology - Ambiguous Term

• a set of abstractions that can be used to describe a part of "the world"
  • is the set of pre-defined categories and relationships of WordNet an ontology?
  • or is WordNet and ontology?

Ontologies - Different Levels of Abstraction

• ontologies have different levels of abstraction
  • the set of pre-defined categories and relationships of WordNet is more generic than WordNet but they are both ontologies
  • the set of pre-defined categories and relationships of WordNet (plus the corresponding syntax) represent an ontology for the representation of thesauri for natural language (any language)
  • WordNet is an ontology for natural language representations

Types of Ontologies - Examples

• domain ontologies
  • vocabularies for medicine, museums, government agencies
• ontologies for general and common sense knowledge
  • e.g. for mathematics (numbers, arithmetic operators), for physical descriptions (space, time) and for common sense knowledge (dangerous, avoid, …)
• representational ontologies
  • domain (understood as above) independent; e.g. the frame ontology (Gruber 1993), for frame based (Object Oriented) representations

Ontology - Most Frequent Use

• given an information domain such as that of museums, financial institutions or government councils, the term ontology is usually associated with the definition of the terms that can be used to describe the respective domain rather than with particular domain representations
  • making a parallel to database systems, an ontology is usually regarded as the database schema rather than the database itself (note that, in database systems, the difference between the schema/intension of the database and the extension of the database is clearly defined)
Skeletal Methodology for Ontology Creation

• identify purpose and scope
• build ontology
  □ capture
    ▪ identify key concepts and relationships
    ▪ produce precise unambiguous textual definitions
  □ code
    ▪ choose/commit to a representational ontology (meta-ontology)
    ▪ choose language
    ▪ write the code
  □ integrate with existing ontologies
• evaluate
• create documentation
• refer to (Uschold and Gruninger, 1996)

Ontologies - Conjectures

• ontologies are needed in the context of any knowledge based or knowledge management system, whenever the issue of knowledge sharing arises
• the more refined and formal (sic!) an ontology is, the better are the capabilities for developing “intelligent” software agents
• ontologies are necessarily required for the next stage in the development of the WWW — the semantic web
  □ current meta-data standards (e.g., Dublin Core and e-GMS) represent only the beginning (will become clearer later)

Ontologies and the Semantic Web

machine 1
static resource committed to ontology O
machine 2
static resource committed to ontology O
machine n
static resource committed to ontology O

the internet
collect data

machine q
software agent committed to ontology O

Ontologies and Meta-Data

• an ontology can be used to express meta-data
• a meta-data standard/class is an (informal) ontology

Ontology / Meta-Data Standard and RDF

• “simple”:
  □ RDF is a domain independent “language” for resource description
    ▪ resource - general meaning (e.g. not restricted to the web)
    ▪ a resource description could be regarded as being equivalent to meta-data
    ▪ RDF is a domain independent language for meta-data descriptions
  □ for a specific domain, RDF should be used together with the domain’s ontology / meta-data standard
    ▪ elements of the vocabulary of an ontology (elements of a meta-data standard) are used to make statements about resources
    ▪ RDF specifies the structure in which such statements should be written
### Summary

- why ontologies
- what are ontologies
- from very informal to rigorously formal
- examples
- representations and reasoning
- types of ontologies
- ontologies and domain descriptions
- the semantic web and ontologies
- ontologies and meta-data standards
- ontologies, meta-data standards and RDF

### References


