Knowledge Representation & Processing

Yizheng Zhao[1,2]

1. National Key Laboratory for Novel Software Technology
2. School of Artificial Intelligence, Nanjing University

Some Basics of Week 3

Limitations of Knowledge Representation

The prospect of representing everything in the world is daunting

- can't write a complete description of everything
- focus on certain aspects of the world — physical objects, events, beliefs, time, etc.
  - e.g., that occur in many different domains — Nanjing University, shopping on the Internet, etc.
- leave placeholders where new knowledge can fit in
  - e.g., define what it means to be a physical object, say a tomato, and the details of different types of objects — fruits, robots, books or whatever — can be filled in later

Certain aspects of the world are hard to capture in formal languages

- most generalizations have exceptions or hold only to a degree
  - e.g., "tomatoes are red" is a useful rule, but some tomatoes are green, yellow or orange
- the ability to handle exceptions/uncertainty is extremely important

Two Types of Ontologies

General-purpose ontology

- applicable in more or less any special-purpose domain (with the addition of domain specific axioms)
- no representational issue can be finessed/brushed under the carpet
- has so far had only limited success — none of top AI applications make use of a shared ontology
- social/political considerations can make it difficult for completing parties to agree on an ontology
  - e.g., that occur in many different domains — Nanjing University, shopping on the Internet, etc.

Special-purpose ontology

- "every ontology is a treaty — a social agreement — among people with some common motive in sharing"
- top AI applications use special-purpose knowledge engineering

Creation of Ontologies

Existing ontologies have been created along four routes:

- by a team of trained ontologist/logicians, who architect the ontology and write axioms; the CYC system was mostly built this way (Lenat and Guha, 1990).
- by importing categories, attributes, and values from an existing database or databases; DBpedia was built by importing structured facts from Wikipedia (Bizer et al., 2007).
- by parsing text documents and extracting information from them; TEXTRUNNER was built by reading a large corpus of Web pages (Banko and Etzioni, 2008).
- by enticing unskilled amateurs to enter commonsense knowledge; the OPENMIND system was built by volunteers who proposed facts in English (Singh et al., 2002; Chklovski and Gil, 2005).

Three Building Blocks of Ontologies

Individuals

- often a physical/conceptual object, e.g., ronaldo, wolffy, 74-E5-0B-3B-8B-42
- the most basic element of the world, because interaction with the world takes place at the level of individual objects

Concepts/Classes

- The organization of individuals into categories is a vital part of knowledge representation
- much reasoning takes place at the level of categories
  - e.g., a shopper would normally have the goal of buying a computer, rather than a specific computer such as 74-E5-0B-3B-8B-42

Roles/ObjectProperties

- categorising an object, e.g., 74-E5-0B-3B-8B-42 is a computer
- describing properties of an object, e.g., 74-E5-0B-3B-8B-42 has a screen
- relating two objects, e.g., 74-E5-0B-3B-8B-42 has the screen XXX