Lecture 8: Ontologies

Dr. Roman V Belavkin

BIS4410

Contents

1 Introduction to Ontology

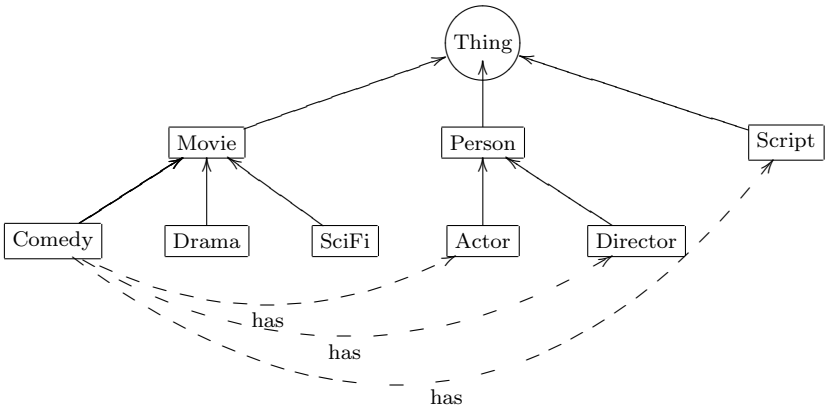
Introduction to Ontology

- There is a discipline *Ontology*, which is the philosophical study of 'being'.
- There are specific ontologies used in knowledge management (KM), knowledge representation and computer science that describe a system of objects or concepts in some domain.

  An ontology is a specification of a conceptualisation.

  (Gruber, 1993)
- Ontologies in KM are used to define properties of and relationships between the concepts.
- Ontologies are used by people (e.g. by experts) and by computers (e.g. in semantic web applications).
- Computer applications of ontologies require standardisation in definition of computer languages for ontologies.

Example: Movie Ontology



Domains of Ontologies

- Ontologies can describe knowledge and concepts at different levels of abstraction:

  Upper ontologies represent common concepts and relations used across a wide range of domains (e.g. OpenCyc, SUMO, WordNet).

  Domain ontologies represent concepts and relations specific to some domain (e.g. Movie Ontology, Gene Ontology).

- Different domain ontologies are often incompatible with each other or with the upper ontologies.

2 Components of Ontologies

Components of Ontologies

- Ontologies usually consist of:

  Individuals or instances of objects.

  Classes or sets of collections of objects.

  Attributes or properties that objects may have.

  Relations ways in which concepts can be related to one another.

- Individuals, classes and attributes together can be considered as the set of all concepts $c_1, \ldots, c_n \in C$.

- Relations are 'links' between pairs of concepts, such as $(c_1, c_3) \in r_1$, $(c_2, c_4) \in r_2$ means c_1 is related to c_3 by relation r_1, and c_2, c_4 by r_2.

- Ontologies may also contain *restrictions* (constraints describing individuals or classes), *axioms* (a priori assertions always assumed to be true) and *events* (changes of attributes or relations).

Individuals

- Individuals are specific instances of the concepts or objects.

  Example 1 (Individuals).
  - In the Movie ontology, individuals can be a specific film (*Sherlock Holmes: A Game of Shadows*), a specific director (*Guy Ritchie*), a specific actor (*Robert Downey*).
  - The film genre (*Action*) is not an individual.

- Individuals represent the ground or atomic level of the ontology.
- An ontology may have no individuals, only classes.

Classes

- Classes, types or categories are sets of individuals.

  Example 2 (Classes). In the Movie ontology, movie genre (e.g. Comedy, Drama), types of person (Actor, Director) are classes.

- Classes can be organised into a hierarchy or taxonomy using the *SubclassOf* relation $\subseteq$.
- All ontologies have at least two classes:

  Thing representing the class of all concepts (i.e. the universe or domain).

  Nothing representing the empty set (a subset of any set).

- As in formal concept analysis (FCA), classes can be characterised by their *Extent* (all elements of the class) or *Intent* (all common attributes within the class).
- FCA algorithms can be used to automatically derive the taxonomy of an ontology.

Subtype relation

Definition 3 (Subtype, Subclass-Of)**.** relation between two concepts A and B corresponds to the subset relation: $B \subseteq A$ means B is a subtype of A.

Disjointedness : when each instance of the supertype A belongs to at most one subtype $B \subseteq A$. Otherwise, if $o \in B$ and $o \in C$ (i.e. o participates belongs to multiple subtypes), the subtypes $B \subseteq A$ and $C \subseteq A$ are overlapping.

Completeness : when each instance of a supertype participates in at least one subtype. Otherwise the subtypes are partial.

Example 4 (Disjointness and Completeness).
- In the Movie ontology, classes Movie and Person are disjoint
- Comedy and SciFi are not (e.g. *The Hitchhiker's Guide to the Galaxy*).

Attributes

- Concepts can be described by the set of common attributes, such as parts of an object.

  Example 5 (Attributes). A movie can be described by the set of 'parts' it has, such as Script, Director, Actors, Music.

- Attributes can be other concepts in their own right (i.e. individuals or classes), but they define the context for other concepts.

Relations

- Relations define how pairs of concepts can be related.

  Example 6 (Relations).
  - Subclasses are related to their superclasses by relation `SubclassOf` or `IS-A`.
  - The fact that a movie has a director or an actor can be expressed by relation `HAS` or `HAS-PART`.

- Relations can have specific properties, such as *symmetry* (i.e. $(a, b) \in r_1$ implies $(b, a) \in r_1$) or *transitivity* (i.e. $(a, b) \in r_2$ and $(b, c) \in r_2$ implies $(a, c) \in r_2$).

  Example 7 (Transitive relations). Relation `IN` is transitive:
  - Hendon is `IN` London, and London is `IN` UK.
  - Hendon is `IN` UK.

Ontology Graph

Definition 8 (Ontology)**.** is a structure $O = \{C, R, A\}$, where

C is a set $\{c_1, \ldots, c_n\}$ of concepts.

R is a set $\{r_1, \ldots, r_m\}$ of binary relations $r_i \subseteq C \times C$ between concepts.

A is a set of axioms.

Definition 9 (Ontology Graph)**.** An ontology can be represented by a (multi)-graph (or labelled graph), in which nodes (vertexes) represent concepts $c_1, \ldots, c_n \in C$, and labelled arrows (edges) represent the relations $r_1, \ldots, r_m$, $r_i \subseteq C \times C$.

3 Ontology Languages and Projects

Ontology Languages

There is a number of formal languages for describing and engineering ontologies, such as

CycL : developed for the Cyc project. Based on First Order Predicate Calculus.

RIF : (Rule Interchange Format) is the language combining ontologies and rules.

OBO : (Open Biomedical Ontologies) used for biological and biomedical ontologies.

OWL : (Web Ontology Language) developed for using ontologies over the WWW. Endorsed by the W3C.

Ontology Projects

- Some published upper ontologies:

  Cyc : A large upper ontology for formal representation of the common sense knowledge. `http://www.cyc.com/`

  WordNet : A lexical database for the English language `http://wordnet.princeton.edu/`

  ThoughtTreasure : A commonsense knowledge base and architecture for natural language processing. `http://www.signiform.com/tt/htm/tt.htm`

- Some domain ontologies

  CContology : Customer Complaint Ontology `http://www.jarrar.info/CContology/`

  Movie Ontology : `http://www.movieontology.org/`

  Music Ontology : `http://musicontology.com/`

  Disease Ontology : `http://diseaseontology.sourceforge.net/`

  Gene Ontology : `http://www.geneontology.org/`

  Plant Ontology : `http://www.plantontology.org/`

Summary

- Ontologies are common specifications of a conceptualisation used for knowledge representation and reasoning.
- There are domain ontologies and upper ontologies.
- Ontologies contain individuals, classes, attributes of concepts and relations between them.
- There are several formal languages for ontologies.
- KM applications take advantage of domain specific ontologies as well as a number of published upper ontologies.
- Ontology engineering is becoming an important area of Knowledge Management, E-commerce and the Semantic Web (Gomez-Perez, Fernandez-Lopez, & Corcho, 2004).

References

Gomez-Perez, A., Fernandez-Lopez, M., & Corcho, O. (2004). *Ontological engineering: With examples from the areas of knowledge management, e-commerce and the semantic web*. Springer.

Gruber, T. R. (1993, June). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199-220.