

## COMP 210: Lecture 26: Ontologies

Trevor Bench-Capon  
Room 215, Ashton Building

<http://www.csc.liv.ac.uk/~tbc/COMP210>

## Ontologies

- A problem we identified with rule based expert systems, was the lack of rigorous semantics for their representation paradigm.
- This made it difficult to
  - Verify and validate the systems
  - Share knowledge
  - Reuse knowledge
  - Acquire knowledge methodically

## Ontologies

- Because building a large knowledge base is an expensive effort, it is important
  - To get it right
  - To be able to make as much use of it as possible
- Ontologies address these problems by providing **a formalisation of a conceptualisation of a domain** (Thomas Gruber)

## Ontologies

- Ontologies are supposed to
  - Provide a common understanding of the domain
    - To share between people and software agents
  - Record design decisions
    - To make assumptions explicit
    - To facilitate merging and re-use
- A common well defined vocabulary for writing rules
- An ontology for a KBS serves many of the purposes of a Data Dictionary for a DB

## Winston's ZOOKEEPER

- Z1: mammal(X):-hair(X) .
- Z2: mammal(X):-givesMilk(X) .
- Z3: bird(X):-feathers(X) .
- Z4 bird(X):-flies(X),  
laysEggs(X).
- Z5: carnivore(X):- mammal(X),  
eats(X,meat).
- Z6: carnivore(X):- mammal(X),  
teeth(X,pointed),  
has(X,claws)
- eyes(X,forwardPointing).
- Z7: ungulate(X):- mammal(X),  
has(X,hoofs).
- Z8: ungulate(X):- mammal(X),  
chewsCurd(X).
- Z9: cheetah(X):- carnivore(X) ,  
colour(X,tawny),  
spots(X,dark) .
- Z10: tiger(X):- carnivore(X),  
colour(X,tawny),  
stripes(X,black).
- Z11: giraffe(X):-ungulate(X),  
legs(X,long), neck(X,long),  
colour(X,tawny),  
spots(X,dark).
- Z12: zebra(X):- ungulate(X),  
colour(X,white),  
stripes(X,black).
- Z13: ostrich(X):-bird(X),  
not flies(X), legs(X,long),  
neck(X,long).
- colour(X,blackandwhite).
- Z14: penguin(X):-bird(X),  
swims(X),  
not flies(X),  
colour(X,blackandwhite).
- Z15: albatross(X):- bird(X),  
flies(X,well).

## ZOOKEEPER

- Works well enough
- Winston liked it enough to use it as a running example in his text book
- But: is it a satisfactory knowledge base?
  - No coherent strategy
  - Not very extensible

## Problems

- Animals divided into mammals and birds:
  - Standard zoological taxonomy – but is this **exhaustive** and **disjoint**?
- Mammals are divided in carnivores (what they eat) and ungulates (have hooves, what their feet are like).
  - Why **this** distinction? **Exhaustive**? **Disjoint**?
- Birds are not divided. **Why not**?
- Stripes are black and spots are dark. **Same**?
- Colour can be white or blackandwhite. **Pattern**?
- Albatross flies well. What **counts as** well?
- Lays eggs?
  - **Observable**? Only applies to **females**?

## Consider the 20 predicates

- Some present alternatives, enabling grouping
  - skin covering {hair,feathers}
  - colour{white, tawny,black and white} **inconsistencies**
  - markings{spots,stripes}
  - movesBy{swims,flies} **incompleteness**
  - feet{hoofs,claws}
- Sometimes only one option is given
  - teeth{pointed,?}
  - eats{meat,?}
  - legs{long,?}
  - neck{long,?} **gaps**
  - stripes{black,?}
  - spots{dark,?}
  - files{well,?}
  - eyes{point forward,?}

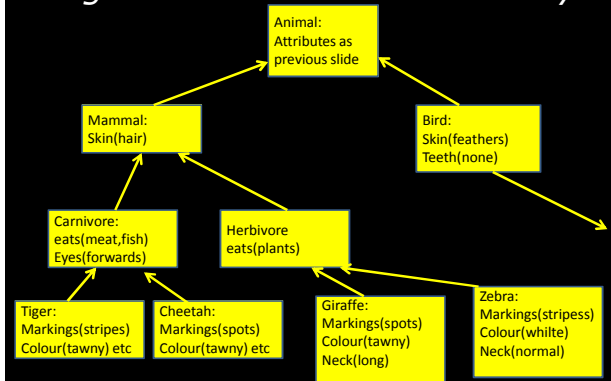
## Let's devise a vocabulary for the domain

- What **attributes** do we want?
- For these attributes: what **values** are **possible**?
- Will form the basis of entity-attribute-value triples to use in our rules

## So: (added values in yellow)

- skin covering {hair, feathers}
- colour{white, tawny, black, **other**}
- markings{spots, stripes, **irregular**}
- movesBy{swims, flies, **neither**}
- feet{hoofs, claws, **toes**}
- teeth{pointed, **rounded, none**}
- eats{meat, **plants, fish**}
- legs{long, **normal**}
- neck{long, **normal**}
- **markingColour**{dark, **light**} %replaces spots and stripes
- eyes{point forward, **sideways**}

## Organise into an Is-A hierarchy



## Strict Specialisation

- The slot fillers are **possible values** not **defaults**
- As we go down the hierarchy we can only
  - **Add attributes**
    - E.g. A soldier has a rank a person does not
  - **Remove values**
    - Soldier has rank {private, sergeant, captain, general}
    - Officer has rank {captain, general}

## Formalisation

- Strict Specialisation enables us to give a formal description
- Can use predicate calculus. E.g.

$(\forall x)(\text{officer}(x) \rightarrow (\text{rank}(x, \text{captain}) \vee \text{rank}(x, \text{general})))$   
 $(\forall x)(\text{isa}(\text{animal}, x) \rightarrow \text{isa}(\text{mammal}, x) \vee \text{isa}(\text{bird}, x))$

- Usual logic to use is, however Description Logic
  - Modules on that in third year

## Current Work on Ontologies

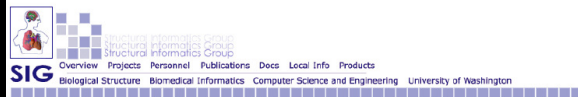
- Many substantial ontologies have been developed.
- Especially in the Medical Domain
- E.g. SNOWMED
  - "SNOMED Clinical Terms (SNOMED CT) is a dynamic, scientifically validated clinical health care terminology and infrastructure that makes health care knowledge more usable and accessible. The SNOMED CT Core terminology provides a common language that enables a consistent way of capturing, sharing and aggregating health data across specialties and sites of care."

## Place of Ontologies in Modern AI

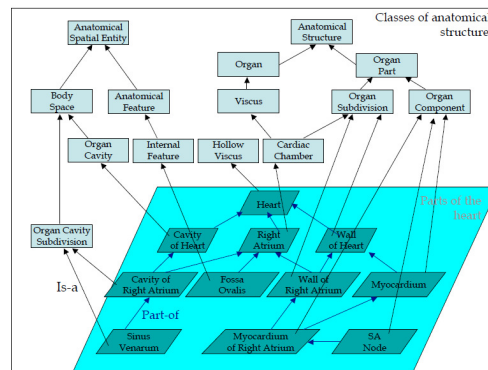
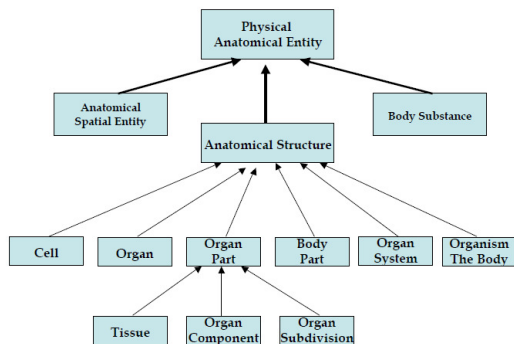
- Specific domain ontologies: but general purpose – supposed to support many applications
  - SNOMED CT (200,000 medical concepts)
  - The Gene Ontology GO (22,700 definitions)
  - NCI cancer ontology (500,000 triples)
  - Foundational Model of Anatomy
  - SUMO Ontology
    - "The goal . . . is to develop a standard upper ontology that will promote data interoperability, information search and retrieval, automated inferencing, and natural language processing."
- Knowledge about everything
  - CYC, OpenCyc
- Semantic Web

## Foundational Model of Anatomy

- Long-term project at University of Washington to create a comprehensive ontology of human anatomy
- 72K concepts, 1.9M relationships
- One of the largest and best developed ontologies in biomedicine



### Top level of the Foundational Model of Anatomy

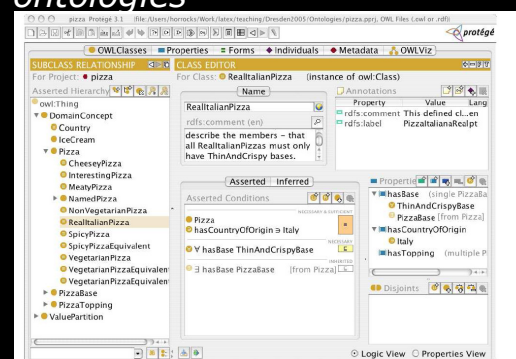




## Ontologies and the Semantic Web

- Annotations must be meaningful for the computer
- Background knowledge is required
- Users do not want define concepts themselves
- Annotate Web pages with concepts defined in ontologies available on the Web
  - Ontologies provide a shared understanding of a domain
  - Improve accuracy of Web searches
  - Web searches will be able to generalise/specialise queries

## Protege – for building ontologies



## Tools and infrastructures

- Editors/environments
  - Oiled, Prote'ge', Swoop, Construct, Ontotrack, . . .
- Reasoning systems
  - Cerebra, FaCT++, Kaon2, Pellet, Racer,
- Design methodologies
- Foundational ontologies, modularisation, etc.

## Summary

- Ontologies began as specialised knowledge bases defining a *common understanding* of terms to be used in rule based systems
- Have now developed to provide repositories of detailed knowledge
- Presentation in FO or description logic has a well-defined semantics
- Ontologies for Semantic Web
  - If you are interested, enrol in COMP318 next year
- Automated reasoning procedures needed