Ontology Languages

Frank Wolter

Department of Computer Science

University of Liverpool

About The Module

- **These slides and other material** for this module are available at the module site

  `http://cgi.csc.liv.ac.uk/~frank/teaching/comp08/comp321.html`

  Recommended readings:

- **Steffen Staab et al.** *Handbook of Ontologies*. Springer, 2004
- **Franz Baader et al.** *The Description Logic Handbook*. Cambridge University Press, 2007
- **Pascal Hitzler et al.** *Foundations of Semantic Web Technologies*. CRC, 2009
- **Serge Abiteboul et al.** *Web Data Management*. Cambridge University Press, 2011
- **Dean Allemang et al.** *Semantic Web for the Working Ontologist*. Morgan Kaufman, 2008

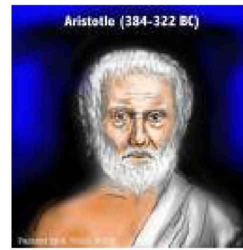
Ontology: Origins and History

Ontology in Philosophy

$\boldsymbol{o\nu\tau o\lambda o\gamma i\alpha}$

a philosophical discipline — a branch of philosophy that deals with the nature and the organisation of reality

- Science of Being (Aristotle, Metaphysics, IV, 1)
- Tries to answer the questions:
  - What characterises being?
  - Eventually, what is being?
- How should things be classified?



Ontologies in Computer Science

"A formal, explicit specification of a shared conceptualization"

[based on Tom Gruber 1993]

A bit more concrete:

- Assume a domain of interest such as Football, Medicine, a University, a Department within a University, a retail company, etc.
- An ontology provides a model of such a domain that is shared by a group of users and is given in a formal language.
- Nowadays, such a model is mostly given by introducing a common vocabulary that provides names for classes (concepts), properties of classes, and attributes of classes. It also defines relationships between classes, properties, and attributes.

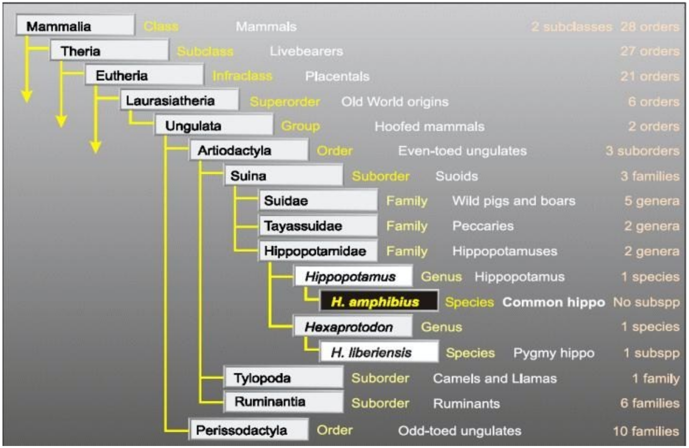
Ontology examples: Lightweight taxonomies

Taxonomy is the practice and science of classification.

The word comes from the Greek $\tau\alpha\xi\iota\varsigma$ ('order') and $\nu o\mu o\varsigma$ ('law' or 'science')

- Linnaean taxonomy: a classification of living things

  (Carl Linnaeus, 1707–1778, 'father of modern taxonomy')

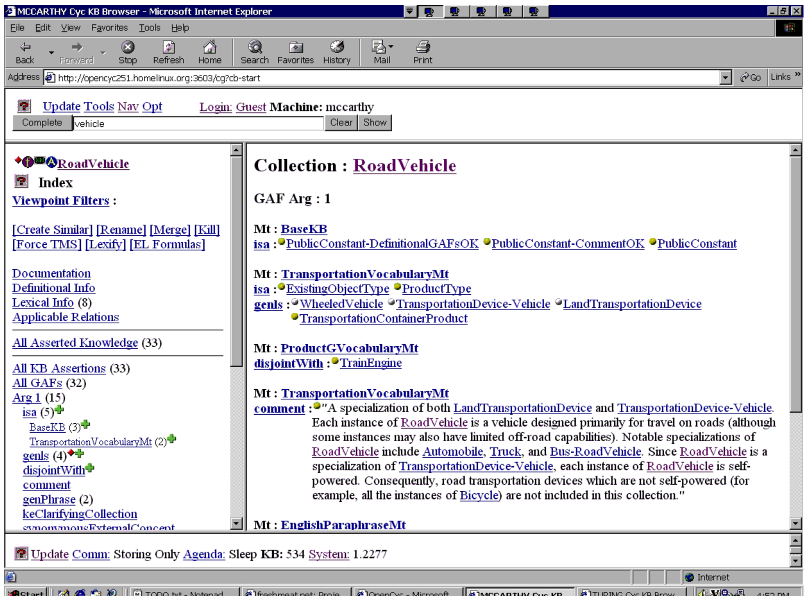


- Yahoo! Web Directory (http://dir.yahoo.com/)
- Open Directory Project 590,000 categories (http://dmoz.org/)
- Amazon product catalog

Ontology examples: Heavy weight

- Cyc, upper ontology for all of human consensus reality (started in 1994)
  `http://www.opencyc.org/`

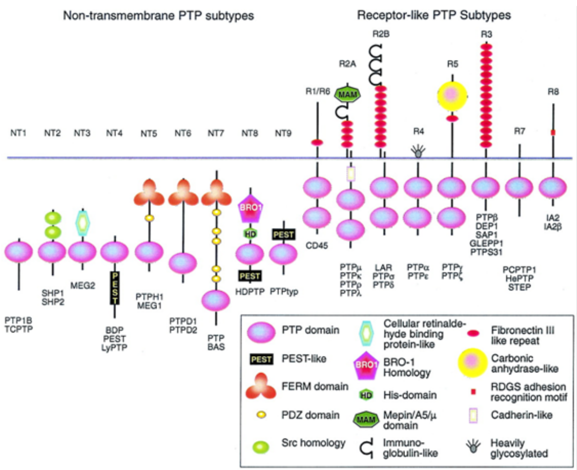
  the world's largest and most complete general knowledge base and commonsense reasoning engine

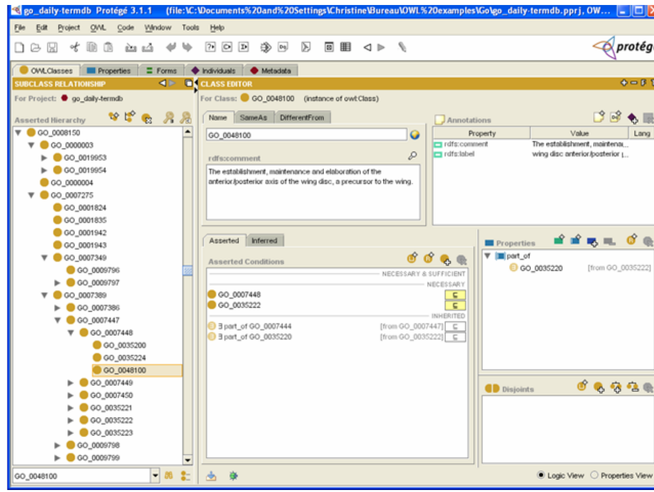


Ontology examples: e-Science

- Open Biomedical Ontologies Consortium: GO `http://www.geneontology.org/`, MGED `http://mged.sourceforge.net/ontologies/MGEDontology.php`, ...

  Used, e.g., for 'in silico' investigations relating theory and data

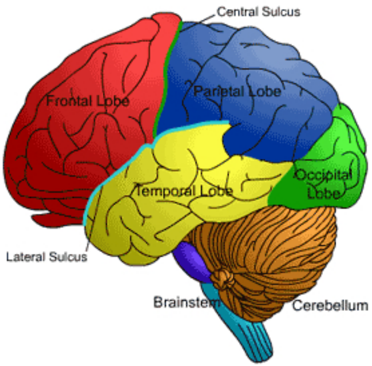


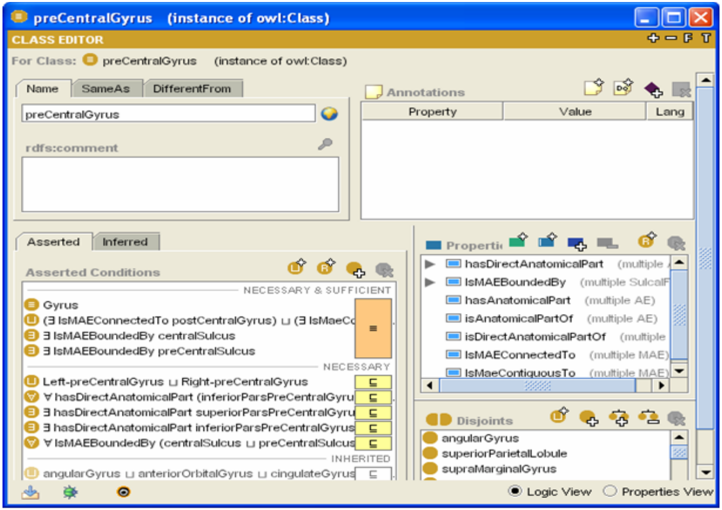


Ontology examples: Medicine

- Terminologies such as Snomed CT, NCI, Galen and FMA

Used, e.g., for semi-automated annotation of MRI images

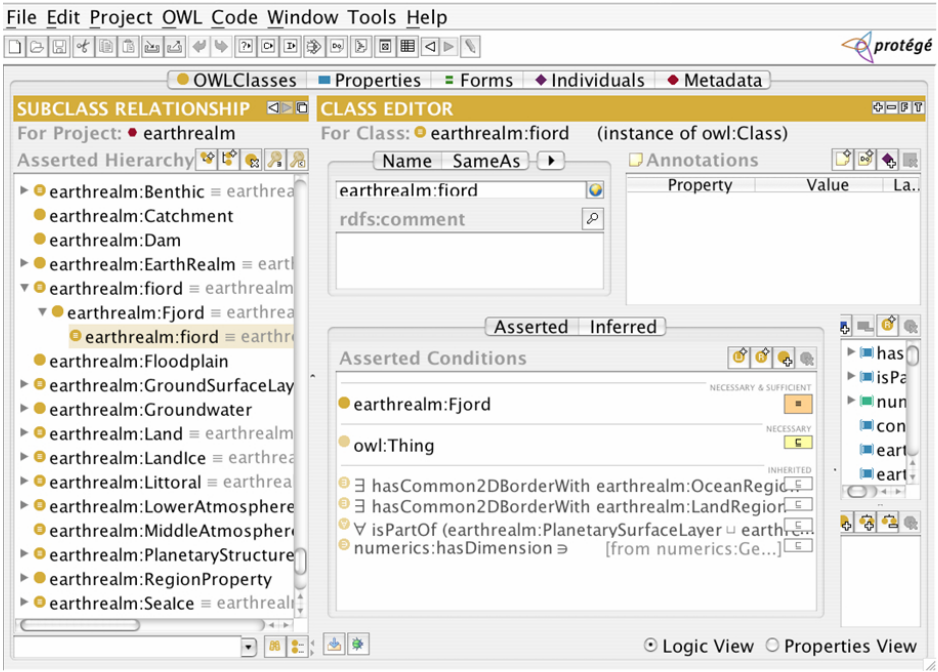




Clinicians use different terms that mean the same thing: 'heart attack,' 'myocardial infarction,' and 'MI' may mean the same thing to a cardiologist, but to a computer, they are all different.

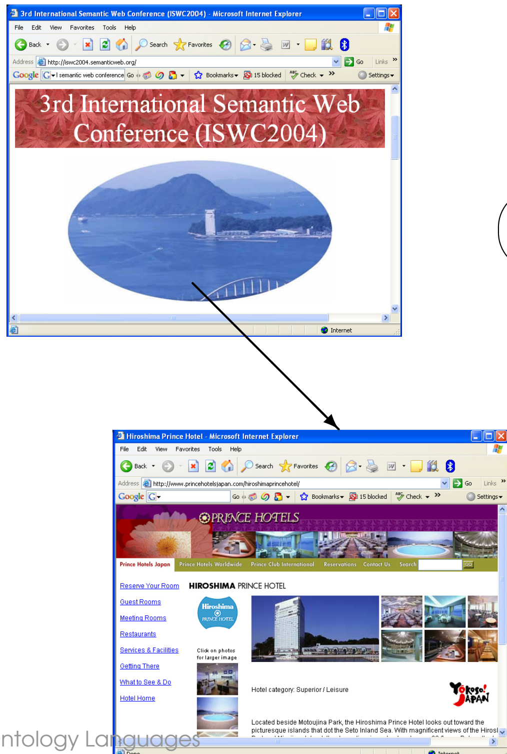
Ontology examples: organising complex information

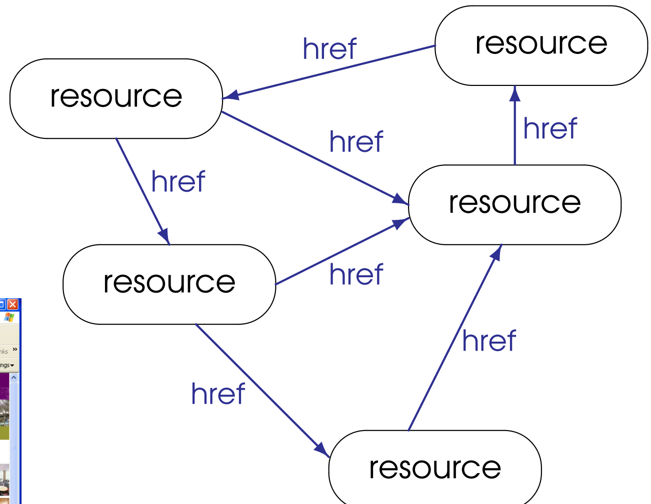
- E.g., UN-FAO, NASA, Ordnance Survey, General Motors, Lockheed Martin, . . .



Ontology example application: Semantic Web

Where we are Today: the ***Syntactic*** Web





The Syntactic Web is...

A place where **computers** do the **presentation** (easy)
and **people** do the **linking** and **interpreting** (hard).

- A hypermedia, a digital library

  A library of documents (called web pages) interconnected by links

- A database, an application platform

  A common portal to applications accessible through web pages, and presenting their results as web pages

- A platform for multimedia

  BBC Radio 4 anywhere in the world.

- A naming scheme

  Unique identity for those documents

Impossible (?) using the Syntactic Web...

- Complex queries involving **background knowledge**

  Find information about "animals that use sonar but are neither bats nor dolphins" (e.g., Barn Owl)

- Locating information in **data repositories**

  Travel enquiries
  Prices of goods and services
  Results of human genome experiments

- Finding and using "**web services**"

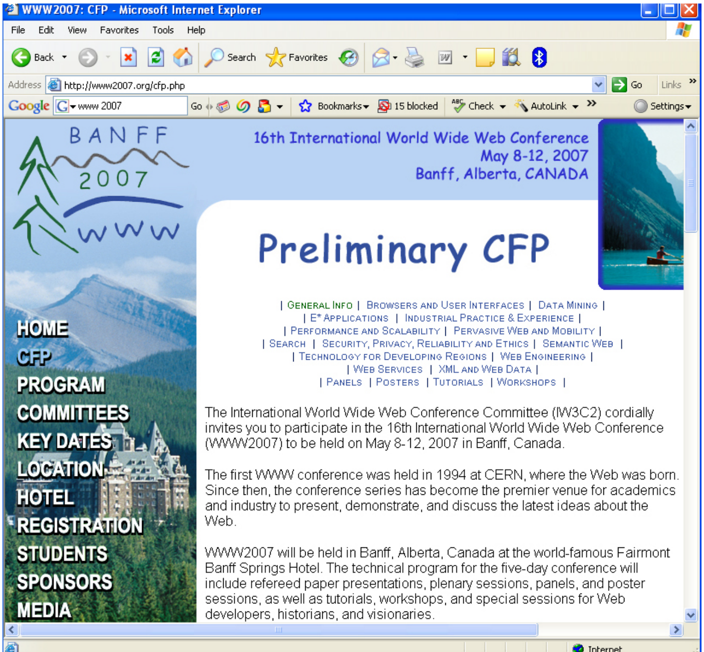
  Visualise surface interactions between two proteins

- Delegating complex tasks to web "**agents**"

  Book me a holiday next weekend somewhere warm, not too far away, and where they speak French or English

What is the Problem?

Consider a typical web page:



- Markup consists of:
  - rendering information (e.g., font size and colour)
  - hyper-links to related content
- Semantic content is accessible to humans but not (easily) to computer...

What information can we see...

WWW2007
The sixteenth International World Wide Web Conference
May 8–12, 2007
Banff, Alberta, Canada

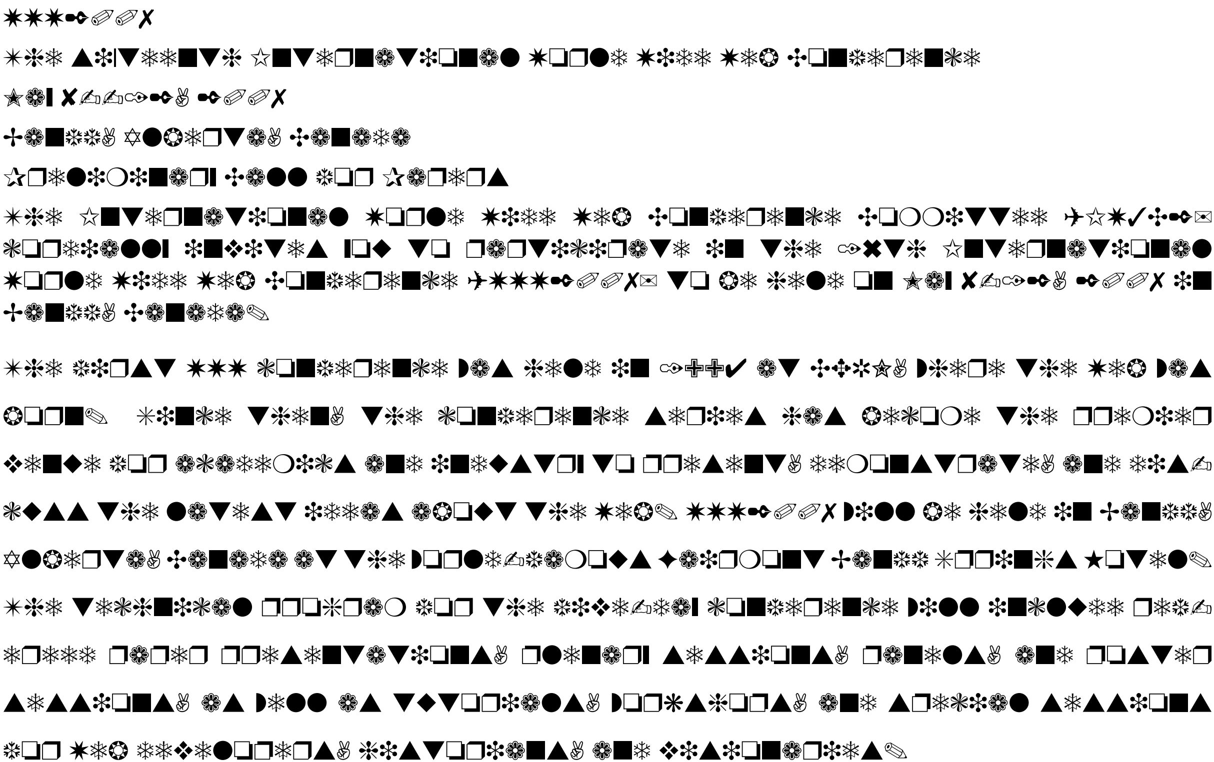
Preliminary Call for Papers

The International World Wide Web Conference Committee (IW3C2) cordially invites you to participate in the 16th International World Wide Web Conference (WWW2007) to be held on May 8-12, 2007 in Banff, Canada.

The first WWW conference was held in 1994 at CERN, where the Web was born. Since then, the conference series has become the premier venue for academics and industry to present, demonstrate, and discuss the latest ideas about the Web.

WWW2007 will be held in Banff, Alberta, Canada at the world-famous Fairmont Banff Springs Hotel. The technical program for the five-day conference will include refereed paper presentations, plenary sessions, panels, and poster sessions, as well as tutorials, workshops, and special sessions for Web developers, historians, and visionaries.

What information can a machine see...



Solution: XML markup with “meaningful” tags?

<name>WWW2007
The sixteenth International World Wide Web Conference</name>
<date>May 8–12, 2007</date>
<location>Banff, Alberta, Canada</location>
<slogan>Preliminary Call for Papers</slogan>

<announcement>The International World Wide Web Conference Committee (IW3C2) cordially invites you to participate in the 16th International World Wide Web Conference (WWW2007) to be held on May 8-12, 2007 in Banff, Canada.</announcement>

<statement>The first WWW conference was held in 1994 at CERN, where the Web was born. Since then, the conference series has become the premier venue for academics and industry to present, demonstrate, and discuss the latest ideas about the Web. WWW2007 will be held in Banff, Alberta, Canada at the world-famous Fairmont Banff Springs Hotel. The technical program for the five-day conference will include refereed paper presentations, plenary sessions, panels, and poster sessions, as well as tutorials, workshops, and special sessions for Web developers, historians, and visionaries.</statement>

But What About...

<name>✶✶✶➼✐✐✗
✳❊❈ ▲❊▎❈❈■▼❊ ☆■▼❈❒■❀▼❊❑■❀● ✶❑❒●❈ ✶❊❈❈ ✶❈✪ ✢❑■❈❈❒❈■❈❈</name>

<date>✯❀▌✘✍✍➾➼✌ ➼✐✐✗</date>

<location>✢❀■❈❈❈✌ ✡●✪❈❒▼❀✌ ✢❀■❀❈❀</location>

<slogan>✰❒❈●❊❍❊■❀❒▌ ✢❀●● ❈❑❒ ✰❀❒❈❒▲</slogan>

<announcement>✳❊❈ ☆■▼❈❒■❀▼❊❑■❀● ✶❑❒●❈ ✶❊❈❈ ✶❈✪ ✢❑■❈❈❒❈■❈❈ ✢❑❍✍ ❍❊▼▼❈❈ ✈☆✶✓✢➼✉ ❈❑❒❈❊❀●●▌ ❊■❖❊▼❈▲ ▌❑◆ ▼❑ ❒❀❒▼❊❈❊❒❀▼❈ ❊■ ▼❊❈ ➾✖▼❊ ☆■▼❈❒■❀▼❊❑■❀● ✶❑❒●❈ ✶❊❈❈ ✶❈✪ ✢❑■❈❈❒❈■❈❈ ✈✶✶✶➼✐✐✗✉ ▼❑ ✪❈ ❊❈●❈ ❑■ ✯❀▌ ✘✍➾➼✌ ➼✐✐✗ ❊■ ✢❀■❈❈✌ ✢❀■❀❈❀✎</announcement>

<statement>✳❊❈ ❈❊❒▲▼ ✶✶✶ ❈❑■❈❈❒❈■❈❈ ◗❀▲ ❊❈●❈ ❊■ ➾✚✚✔ ❀▼ ✢✥❊✰✌ ◗❊❈❒❈ ▼❊❈ ✶❈✪ ◗❀▲ ✪❑❒■✎ ✳❊■❈❈ ▼❊❈■✌ ▼❊❈ ❈❑■❈❈❒❈■❈❈ ▲❈❒❊❈▲ ❊❀▲ ✪❈❈❑❍❈ ▼❊❈ ❒❒❈❍❊❈❒ ❖❈■◆❈ ❈❑❒ ❀❈❀❈❈❍❊❀▲ ❀■❈ ❊■❈◆▲▼❒▌ ▼❑ ❒❒❈▲❈■▼ ❈❈❍❑■▲▼❒❀▼❈✌ ❀■❈ ❈❊▲❈◆▲▲ ▼❊❈ ●❀▼❈▲▼ ❊❈❊❀❑◆▼ ▼❊❈ ✶❈✪✎ ✶✶✶➼✐✐✗ ◗❊●● ✪❈ ❊❈●❈ ❊■ ✢❀■❈❈✌ ✡●✪❈❒▼❀✌ ✢❀■❀❈❀ ❀▼ ▼❊❈ ◗❑❒●❈✍❈❀❍❑◆▲ ✦❀❊❒❍❑■▼ ✢❀■❀❈❀ ✳❒❒❊■❈▲ ★❑▼❈●✎✎✎ </statement>

Still the Machine only sees...

Need to add ontologies: agreement on the meaning of annotations!

Aims of this Module

- To introduce logical languages used to build ontologies and terminologies in computer science and information systems.
- To introduce logic-based methodologies for designing and maintaining ontologies.
- To introduce reasoning problems and procedures in standard ontology languages such as description logics.
- To study the trade-off between expressive power and computational complexity of reasoning for ontology languages.
- To study applications of ontologies in bio-informatics, medical informatics, and the semantic web.

Learning Outcomes of this Module

At the conclusion of the course students should:

- Be able to translate natural language to standard ontology languages (such as description logics) and back.
- Understand the formal semantics of standard ontology languages.
- Be able to design ontologies in standard ontology languages.
- Understand and be able to apply reasoning procedures for description logics.
- Be able to apply ontologies in computer science applications.
- Understand how reasoning is used in the design and maintenance of ontologies.

Syllabus

- Introduction to ontologies and terminologies and their applications (2 lectures)
- Introduction to ontology languages: Description Logics, Rule-based languages, first-order logic (7 lectures)
- Introduction to the design and maintenance of ontologies (6 lectures)
- Algorithms for reasoning in description logics and their applications (11 lectures)
- Ontologies and databases (4 lectures)

Requirements for ontology languages

- A well-defined **syntax**.

  This is a necessary condition for machine-processing of ontologies.

- A **formal semantics**; i.e., a precise description of the meaning of the sentences of an ontology.

  One important use of a formal semantics is to define what the correct answers to queries to an ontology are.

  In what follows **answering queries** will also be called **reasoning**.

- efficient automated support for **answering queries** (also called automated reasoning support).

- sufficient **expressive power** to model the domain of interest.

Some ontology languages

Weak languages:

- From conceptual modeling: ER-diagrams, UML-diagrams.
- Schema.org (Microdata, RDFa) for markup of websites
- The Resource Description Framework (RDF) and its extension RDFS with schema vocabulary.

More expressive languages:

- Description logics and the standard OWL (Web Ontology Language);
- Datalog and rule-based languages;
- Conceptual graphs;
- Predicate Logic.