@MIE 2011

Medical University of Graz

Institute for Medical Informatics, Statistics and Documentation

Tutorial-style Workshop: Ontological Realism for Biomedical Ontologies and Electronic Health Records

Ontology authoring and assessment

Stefan Schulz

Objectives

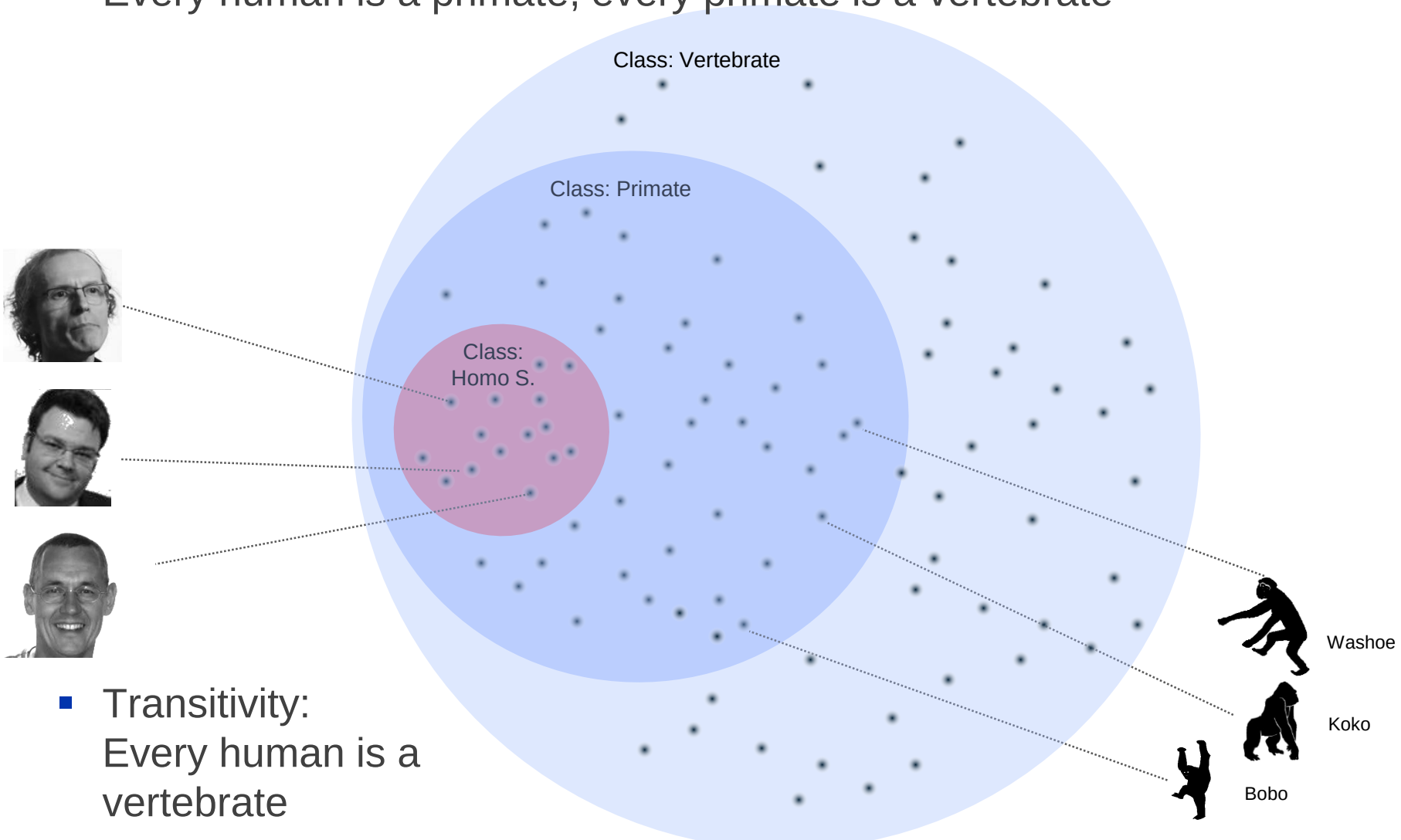
- To demonstrate the heuristic value of the realist approach for authoring OWL ontologies by
  - Presenting four common patterns in OWL ontologies
  - Demonstrating common misconceptions and mistakes using typical examples
  - Introducing methods of double-checking the correctness of ontology axioms
  - Making use of automated reasoning
- To raise awareness of the limitations of what can be expressed by formal ontologies

Ontology engineering based on ontological realism and description logics

- Most current biomedical ontology projects commit to a simple variant of description logics (OWL-EL++). Theoretical background is set theory.
- Principal types of axioms in OWL ontologies
  - Taxonomies (is-a hierarchies):
    Every homo sapiens is a primate, every primate is a vertebrate
  - Aristotelian class definitions (Genus + Differentia)
    Viral hepatitis is equivalent to hepatitis that is caused by some virus population
  - Partonomies (part-of hierarchies)
    Every liver is part of some digestive system and every digestive system is part of some organism
  - Disjoint partitions
    Nothing is both a human and a chimpanzee

Taxonomies

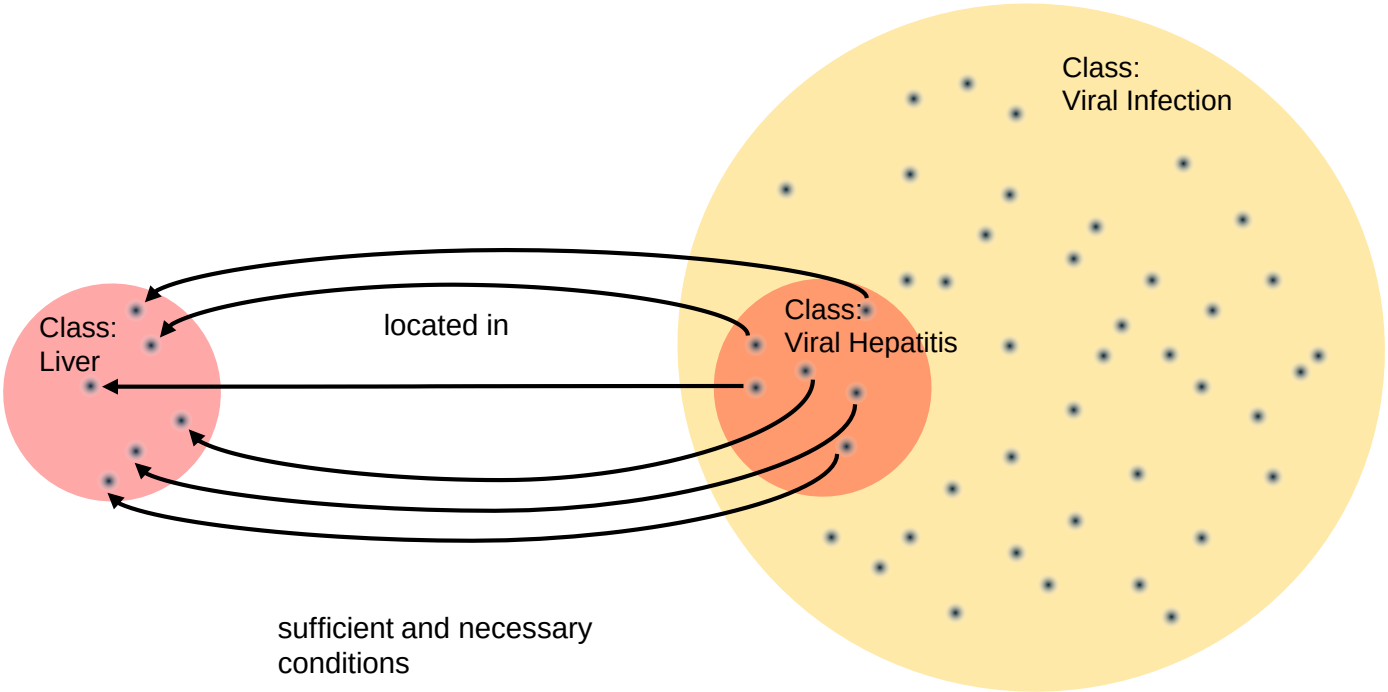
- Every human is a primate, every primate is a vertebrate



- Transitivity:
  Every human is a vertebrate

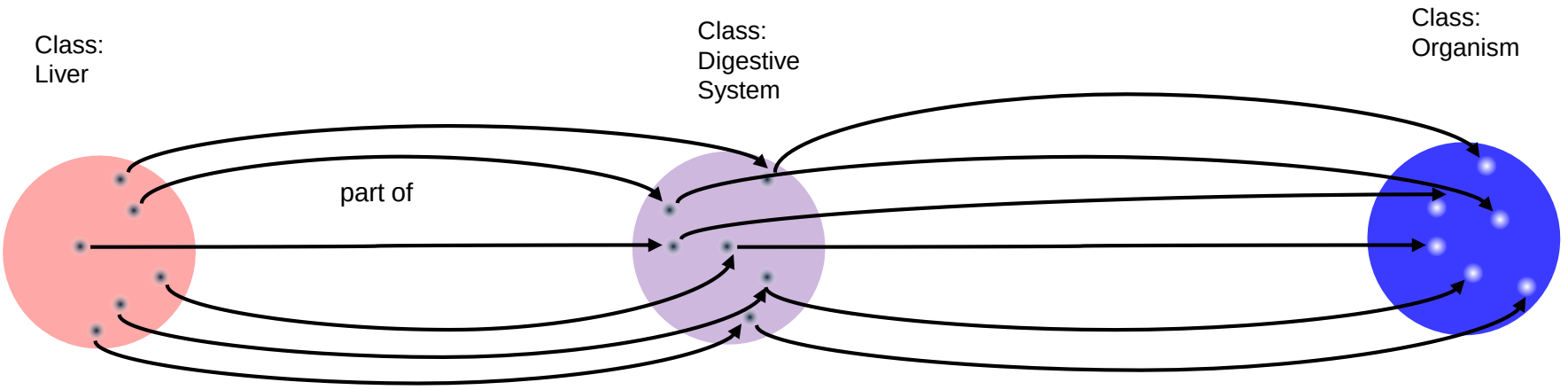
Aristotelian definitions

- *Viral Hepatitis is equivalent to Viral infection that is located in some Liver*



Partonomies

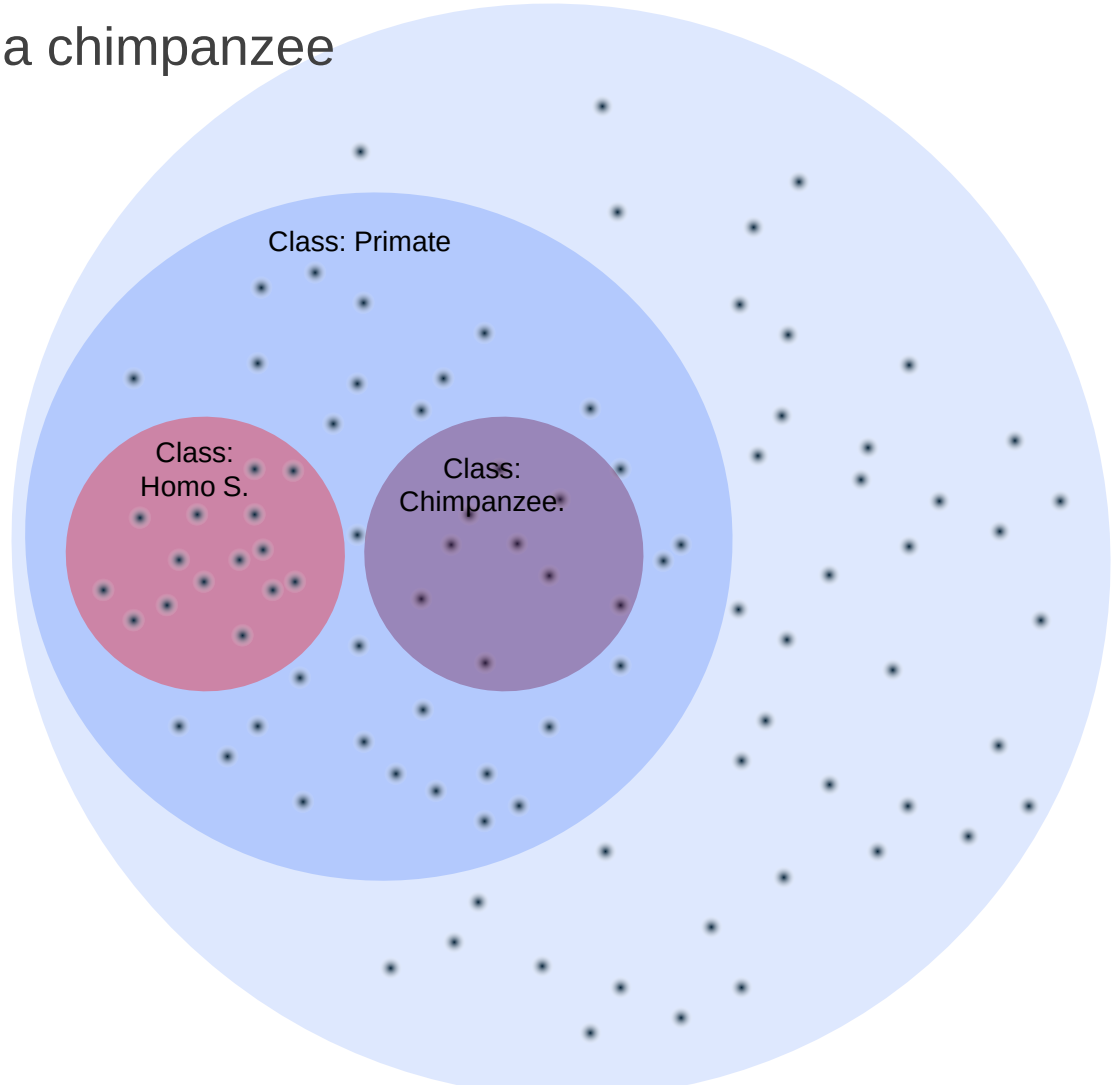
- *Every liver is part of some digestive system and every digestive system is part of some organism*



- *Transitivity: Every liver is part of some organism*

Disjoint partitions

- No human is a chimpanzee



Rephrasing ontology axioms for checking plausibility

- Taxonomies (is-a hierarchies):
  Every homo sapiens is a primate, every primate is a vertebrate
  - *Is there an individual human which is not a primate (vertebrate) ?*
- Aristotelian class definitions (Genus + Differentia)
  Viral hepatitis is equivalent to hepatitis that is caused by some virus population
  - *Is there any individual disease of the type viral hepatitis which is not caused by some virus population ?*
  - *Is there an individual disease which is not caused by any virus population which is of the type viral hepatitis ?*
- Partonomies (part-of hierarchies)
  Every liver is part of some digestive system and every digestive system is part of some organism
  - *Is there some liver that is not part of some digestive system (organism)*
- Disjoint partitions
  Nothing is both a human and a chimpanzee
  - *Is there some individual that is a member of the class human and is also a member of the class chimpanzee*

Typical errors (I): Taxonomy

- *Clinical Medicine*
- - *Oncology*
- - - *Cancer*
- - - - *Lung*
- - - - *Breast*
- - - - *Prostate*
- - - - *Colon*

- *Geography*
- - *Countries*
- - - *BRIC Countries*
- - - - *Brazil*
- - - - *Russia*
- - - - *India*
- - - - *China*

Is there an individual which is member of X but not of Y ?

Typical errors (I): Taxonomy

- *Clinical Medicine*
- - *Oncology*
- - - *Cancer*
- - - - *Lung*
- - - - *Breast*
- - - - *Prostate*
- - - - *Colon*

Correction:

- *Clinical Disease*
- - *Oncologic Disease*
- - - *Cancer*
- - - - *Lung Cancer*
- - - - *Breast Cancer*
- - - - *Prostate Cancer*
- - - - *Colon Cancer*

Is there an individual which is member of X but not of Y ?

Typical errors (I): Taxonomy

- *Geography*
- - *Countries*
- - - *BRIC Countries*
- - - - *Brazil*
- - - - *Russia*
- - - - *India*
- - - - *China*

Correction:

- *Geographical Entity*
- - *Country*
- - - *BRIC Country*

(Members: Brazil, Russia, India, China*)*

Is there an individual which is member of X but not of Y ?

Typical errors (II): Existential quantification

Painkiller equivalentTo *Chemical* and
treats some *Pain*

UreterCarcinoma subclassOf
mayHaveFinding some *Pain*

Is there any individual member of X which is unrelated to any member of Y ?

Typical errors (II): Existential quantification

Painkiller equivalentTo *Chemical* and **treats** some *Pain*

Painkiller equivalentTo *Chemical* and **bearerOf** some (*Disposition* and **realizedBy** only *TreatingPain)*

UreterCarcinoma subclassOf **mayHaveFinding** some *Pain*

UreterCarcinoma subClassOf **bearerOf** some (*Disposition* and **realizedBy** only Pain)

Is there any individual member of X which is unrelated to any member of Y ?

Typical errors (III): Direction of quantification

- *AntibodyProducingCell* subclassOf
  partOf some *LymphoidTissue*

Is there any individual member of X which is unrelated to any member of Y ?

Typical errors (III): Direction of quantification

- *AntibodyProducingCell* subclassOf
  partOf some *LymphoidTissue*

- *LymphoidTissue* subclassOf
  hasPart some *AntibodyProducingCell*

Is there any individual member of X which is unrelated to any member of Y ?

Typical errors (IV): Distributive statements

- *Calcium-Activated_Chloride_Channel-2* subClassOf
  Gene_Product_Expressed_In_Tissue some *Lung* and
  Gene_Product_Expressed_In_Tissue some *Mammary_Gland* and
  Gene_Product_Expressed_In_Tissue some *Trachea*

Is there any individual member of the channel which is expressed in all lung, mammary gland, and trachea?

Typical errors (IV): Distributive statements

- *Calcium-Activated_Chloride_Channel-2* subClassOf
  Gene_Product_Expressed_In_Tissue some *Lung* and
  Gene_Product_Expressed_In_Tissue some *Mammary_Gland* and
  Gene_Product_Expressed_In_Tissue some *Trachea*

- *Calcium-Activated_Chloride_Channel-2* subClassOf
  Gene_Product_Expressed_In_Tissue some
  (*Lung* or *Mammary_Gland* or *Trachea)*

Is there any individual member of the channel which is expressed in all lung, mammary gland, and trachea?

Typical errors (V): confusion or real objects or processes with information objects

- *ThumbAbsent* subClassOf *FindingOfThumb* and **hasFindingSite** some *ThumbStructure*
- *BiopsyPlanned* **associatedProcedure** *some Biopsy*
- *PresumedViralAgent* subClassOf *Virus*
- *BorderOfHeart* subClassOf **partOf** *Heart*

Checking ontology correctness after classification

- Description logics reasoner (e.g. HermiT) computes new entailments from the asserted axioms:

  Example:

  AmputationOfTheFoot equivalentTo *rg* some
  *(***method** some *Amputation and* **procedureSiteDirect** some *FootStructure)*
  AmputationOfToe equivalentTo rg some
  *(***method** *some Amputation* and **procedureSiteDirect** some *ToeStructure)*
  ToeStructure subClassOf *FootStructure*

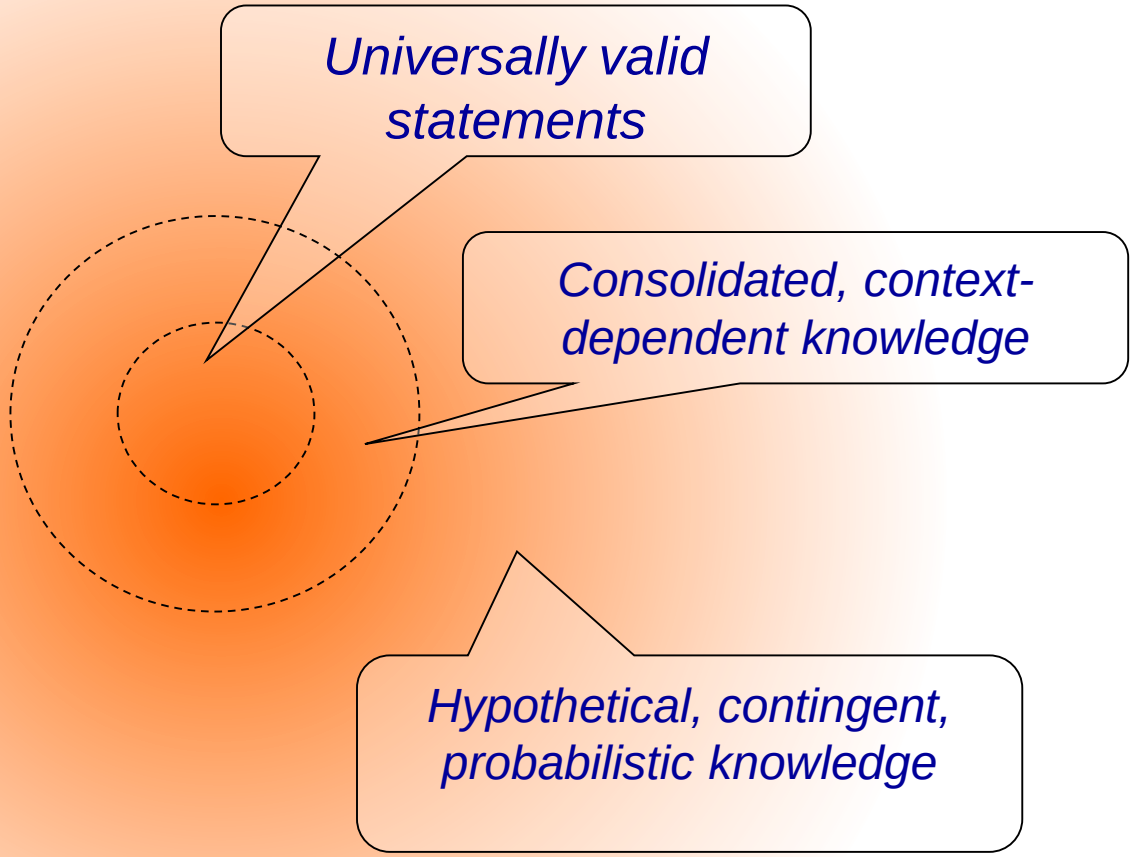
  AmputationOfToe subclassOf *AmputationOfTheFoot*

Is there an individual which is an amputation of a toe but not an amputation of the foot?

Conclusion

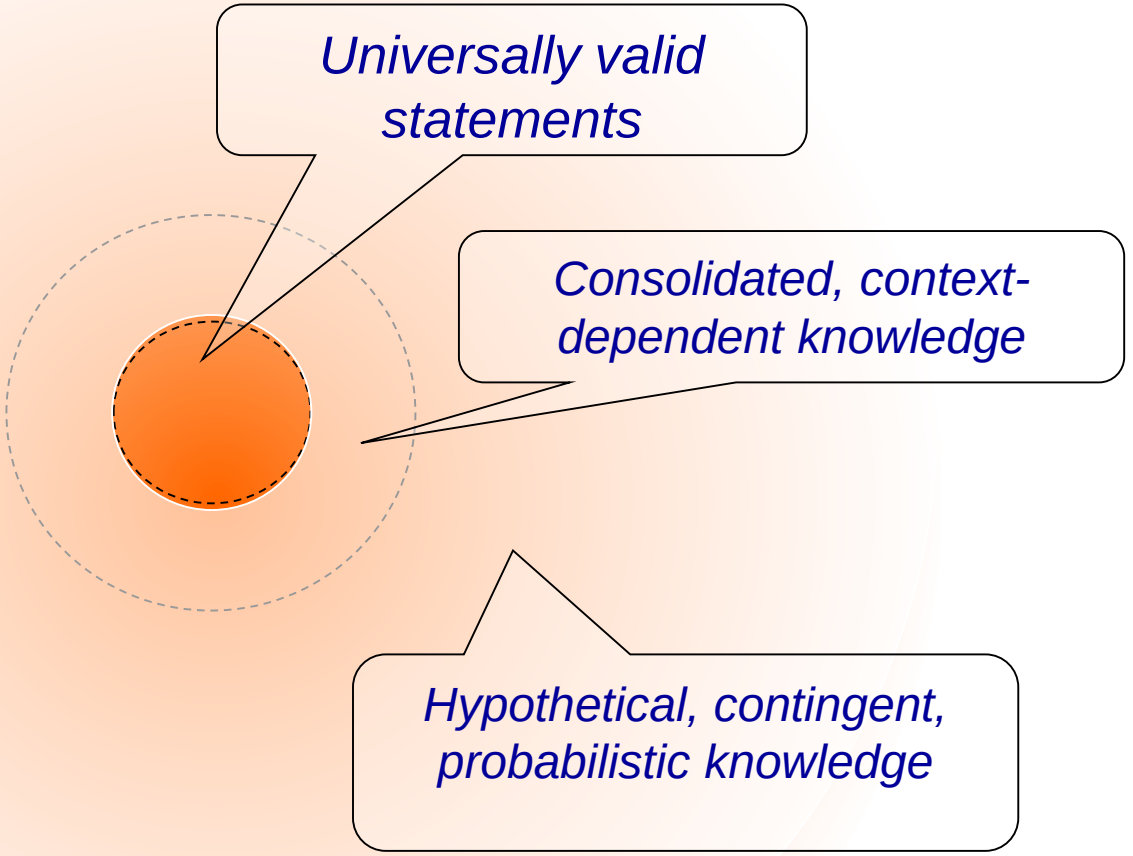
- Ontologies describe classes of real things (individuals), which exist in the world
- Ontologies state what is universally true for all members of a class: Whatever you assert about a class you assert about each of its members
- Most logical axioms used in ontologies are relatively simple, but ontology users and engineers must understand them
- Rephrase the ontology axioms (conversion into negative statements, introduction of individuals) is a useful method to check for correctness and plausibility
- Not only the asserted axioms but also their entailments as computed by DL reasoners should be checked by this method

Ontology $\subset$ Knowledge Representation



Domain Knowledge

Ontologies !



Domain Knowledge

Medical University of Graz

HOME

ICBO + FOIS
ICBO
FOIS

ICBO: Call for Papers
ICBO: Organization

FOIS: Call for Papers
FOIS: Organization

Workshops and Tutorials
Early Career Consortium
Keynotes
Important Dates
Author Instructions
Submission
Venue
Sponsors
Registration
Downloads

ICBO 2012: 3rd International Conference on Biomedical Ontology

The use of biomedical ontologies in annotation of both clinical and experimental data is now a common technique in integrative translational research. To be maximally effective, such ontologies must work well together. As they become more more widely used, the coordination problems become ever more urgent. ICBO 2012 addresses these problems. It will bring together representatives of all major communities involved in ontology use and development in biomedical research, health care, and related areas.

ICBO 2012 Chairs:

Ronald Cornet, Robert Stevens, Melanie Courtot, Ludger Jansen, Trish Whetzel, Janna Hastings

FOIS 2012: 7th International Conference on Formal Ontology in Information Systems

The philosophical discipline of Ontology has become practically relevant with the evolution of complex information systems which rely on robust and coherent representations. Such representations and associated reasoning techniques constitute the modern discipline of formal ontology, which is now applied to artificial intelligence, computational linguistics, bioinformatics, GIS, knowledge engineering, information retrieval, and the Semantic Web. FOIS is intended to explore both theoretical issues and concrete applications.

FOIS 2012 Chairs:

Michael Grüninger, Maureen Donnelly, Giancarlo Guizzardi

Local Organizer

Stefan Schulz

Institute for Medical Informatics, Statistics and Documentation

Medical University of Graz

Auenbruggerplatz 2
8036 Graz, Austria

Email:
stefan.schulz [at]
medunigraz.at

News

Jul 21: FOIS call for papers open.
Aug 8: ICBO call for papers open.

Maintainer: Stefan Schulz

http://purl.org/icbofois2012

Appendix

For close-to-language triple statements to logical axioms

Informal Thesauri

- Examples: MeSH, UMLS Metathesaurus, WordNet
- Describe **terms** of a domain
- **Concepts**: represent the meaning of (quasi-) synonymous terms
- Concepts related by (informal) semantic relations
- Linkage of concepts:
  C1 Rel C2

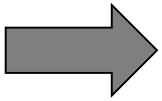
Formal ontologies

- Examples: openGALEN, OBO, SNOMED
- Describe **entities** of a domain
- **Classes**: collection of entities according to their properties
- Axioms state what is universally true for all members of a class
- Logical expressions:
  C1 comp **rel** quant *C2*

The translation of triples into DL statements is ambiguous

Translation of triples

C1 Rel C2



C1 subClassOf **rel** some *C2*

or

C1 subClassOf **rel** only *C2*

or

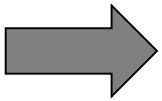
C2 subclassOf inv(**rel**) some *C2*

or…

Translation of groups of triples

C1 Rel C2

C1 Rel C3



C1 subClassOf (**rel** some *C2*) and (**rel** some *C3*)

or

C1 equivalentTo (**rel** some *C2*) and (**rel** some *C3*)

or

C1 equivalentTo (**rel** some ((*C2* or *C3*)))

or …

Ontologies are not exactly made for represent contingent knowledge

- *"Aspirin Treats Headache"*
  "Headache Treated-by Aspirin"
  (seemingly intuitively understandable)
- Translation problem:
  - Not every aspirin tablet treats some headache
  - Not every headache is treated by some aspirin
- Description logics do not allow probabilistic, default, or normative assertions
- Axioms can only state what is true for all members of a class
- Introducing dispositions into ontology possible but not very intuitive ("every aspirin tablet has an inherent disposition which is only realized by treating headache)

Schulz S, Stenzhorn H, Boeker M, Smith B: Strengths and limitations of formal ontologies in the biomedical domain. RECIIS - Electronic Journal in Communication, Information and Innovation in Health, 2009; 3 (1): 31-45:http://dx.doi.org/10.3395/reciis.v3i1.241en
Schulz S, Jansen L: Molecular interactions: On the ambiguity of ordinary statements in biomedical literature. Applied Ontology, 2009; 4 (1): 21-34: http://dx.doi.org/10.3233/AO-2009-0061