What Semantics Do We Need for A Semantic Web for Medicine?

How much formality do we need ?

Stefan Schulz

University Medical Center Freiburg, Medical Informatics,   Freiburg, Germany
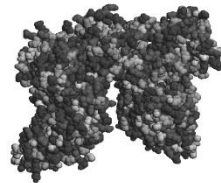
UNIVERSITÄTS
FREIBURG KLINIKUM

Example

- Using Semantic Web standards (OWL-DL)
- Using Biomedical Ontology standards (OBO)
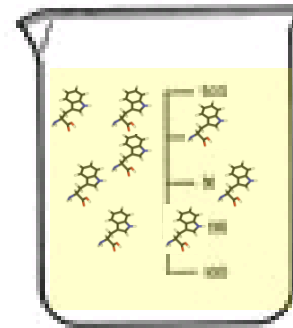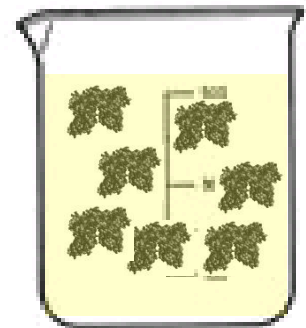- Terminological Inference

Classes

- Amino Acid
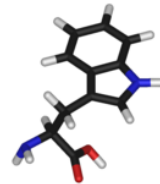
- Protein

- Aminoaciduria

- Proteinuria

Relations (OBO RO)

- *hasPart* / *partOf*

  (parthood in a broad sense):

  relates continuants

- *hasLocation* / *locationOf*

  relates continuants or occurrents with

  continuants

- transitive, reflexive, antisymmetric

Description Logic $\mathcal{EL}^+$

- Subsumption ⊑

- Equivalence ≡

- Existential quantification ∃

- Conjunction ⊓

- transitive roles

Axioms

Protein ⊑ ∃*hasPart.AminoAcid*

Aminoaciduria ≡ *Disorder* ⊓
    ∃*hasLocation.*(*Body* ⊓
        ∃*hasPart.*(*PortionOfUrine* ⊓
           ∃*hasPart.AminoAcid*))

Proteinuria ≡ *Disorder* ⊓
    ∃*hasLocation.*(*Body* ⊓
        ∃*hasPart.*(*PortionOfUrine* ⊓
           ∃*hasPart.Protein*))

Inference

false!

Proteinuria ⊑ *Aminoaciduria*

(since Proteins have Amino Acids as parts, and partOf is transitive)

- Is this error due to formal underspecification ?
- Is *hasPart* not always transitive?

Formal correctness
but ontological sloppyness

AminoAcid: hidden ambiguity:

- AminoAcidSingleMolecule

- AminoAcidResidue

- AminoAcidSingleMoleculeCollection

  - AminoAcidSingleMoleculeCollectionLowConc

  - AminoAcidSingleMoleculeCollectionHighConc

Corrected Axioms

Aminoaciduria ≡ *Disorder* ⊓

  ∃*hasLocation*.(*Body* ⊓

    ∃*hasPart*.(*PortionOfUrine* ⊓

      ∃*hasPart*.*AminoAcidSingleMoleculeCollectionHighConc*))

Proteinuria ≡ *Disorder* ⊓

  ∃*hasLocation*.(*Body* ⊓

    ∃*hasPart*.(*PortionOfUrine* ⊓

      ∃*hasPart*.*ProteinMoleculeCollectionHighConc*))

Two sides of the same coin

Formal Correctness

Ontological Correctness

assures consistency

assures adequacy

Conclusion

- Even little formality must be rooted in a correct ontological foundation to prevent unintended models with inadequate inferences

- If we do not know **exactly** what we are formalizing we cannot rely on machine reasoning. In this case we should give preference to informal, thesaurus-like knowledge representations