Does Medical Image Simulation Require Formal Ontologies?

VIP – Virtual Imaging Workshop
Lyon, France, Dec 14th, 2012
What are (formal) Ontologies?
What are (formal) Ontologies?

- **Computer science view**
  - Ontologies are purpose-oriented formal models of meaning (conceptualizations)

- **Cognitive / linguistic view**
  - Ontologies are concept systems or systems of semantic reference (no clear distinction from thesauri)
  - Also adopted by parts of the Semantic Web community

- **Philosophy view (scientific realism)**
  - Ontology is the study of what there is
  - Formal Ontologies give precise mathematical formulations of the properties and relations of certain entities.

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:
Formal ontology in a nutshell

- Formal ontology = logic based ontology
- Description logics: subset of first order logic
- Common standard: OWL (Ontology Web Language)
- Ontologies are taxonomies of classes
- Ontologies can define classes in terms of (Aristotelian) definitions

**Subclass** (aka is-a):

- Primate subClassOf Vertebrate

**Equivalence**:

- Vertebrate equivalentTo Animal and hasPart some Vertebra

\[ \forall a : A(a) \rightarrow B(a) \]

\[ X \text{ equivalent to } Y \text{ and some } r \text{ Z iff } \forall x : X(x) \leftrightarrow Y(x) \land \exists z : C(z) \land r(x, z) \]
Class: Primate

Class: Homo S.

Class: Vertebrate

Tristan
Bernard
Stefan

Washoe
Koko
Bobo

Transitivity: Every human is a primate, every primate is a vertebrate.

Human subClassOf Primate
\[ \forall a: A(a) \rightarrow B(a) \iff \neg \exists a: A(a) \land \neg B(a) \]

Test:
• there is no neoplasms that is not an oncology
• there is no prostate that is not a neoplasm
• there is no oncology that is not a clinical medicine
Doing taxonomy right

∀a: A(a) → B(a) ⇔ ¬∃a: A(a) ∧ ¬B(a)

Test:
- oncology is an instance of a medical discipline
- there is no prostate neoplasm that is not a neoplasm

Labelling!

http://en.wikipedia.org/wiki/OntoClean
Formal ontology in a nutshell

- Formal ontology = logic based ontology
- Description logics: subset of first order logic
- Common standard: OWL (Ontology Web Language)
- Ontologies are taxonomies of classes
- Ontologies can define classes in terms of (Aristotelian) definitions

**Subclass (aka is-a):**

\[ A \text{ subClassOf } B \iff \forall a: A(a) \rightarrow B(a) \]

**Equivalence:**

\[ X \text{ equivalent to } Y \text{ and some } r \text{ Z } \iff \forall x: X(x) \leftrightarrow Y(x) \land \exists z: C(z) \land r(x,z) \]

**Primate subClassOf Vertebrate**

**Vertebrate**

**Homo Sapiens**

**Vertebrate equivalentTo** Animal and hasPart some Vertebra
Aristotelian Definitions:

*Genus proximum and differentia specifica*

Viral Infection

Viral Hepatitis located in Liver

\[ \text{ViralHepatitis} \text{ equivalentTo} \text{ViralInfection} \text{ and locatedIn Liver} \]

Test:
- There is no viral hepatitis that is not located in a liver
- There is no viral hepatitis that is not a viral infection
Ontological Commitment

- Which are the instances?
- Does the label tell us what is meant?
- Is there an implicit context?

Test:
- There is no neoplasm in both lungs that is not a neoplasm in the left lung
- There is no varicose vein in the lower limb that is not a chronic peripheral venous insufficiency
- There is no patient with neoplasm in both lungs that is not a patient with the neoplasm in the left lung
- There is no patient with varicose lower limb veins that is not a patient with a chronic peripheral venous insufficiency
Upper level ontologies: partition of the domain into disjoint and exhaustive categories

- Upper level ontologies enforce a strict categorization
- Constraints on upper-level categories
- Upper level ontology for the biomedical domain **BioTop**

http://purl.org/biotop
How formal ontologies challenge human cognition

- Built around taxonomies of classes
  - ATTENTION: our intuitive way of hierarchically organize terms is not strictly taxonomic
- State what is true for all individual members of a class (instances of a type)
- Requires to distinguish between classes and individuals
  - ATTENTION: human language is often misleading, e.g. Lyon is a big city vs. The liver is a big organ
- Individuals commit to upper-level categories
  - ATTENTION: our thinking fuses mutually dependent entities that belong to different categories, e.g. Cancer (growth process vs. mass of malignant tissue)
- Upper level categories should be made explicit
  - Explicit upper level ontology – common understanding
  - Implicit upper level ontology of each of us – misunderstanding
What formal ontology is not

- Ontology $\neq$ Knowledge representation
  - "There are very few interesting items of knowledge that are truly ontological in this strict sense" (Alan Rector)
  - antinomy: ὄντος (being) vs. ἐπιστήμη (knowledge)

- Ontology is not appropriate for
  - Default knowledge
    - "The hand has 5 fingers" (unless otherwise stated)
  - Probabilistic knowledge
    - Mesothelioma is a rare cancer
  - Contingent knowledge
    - Aspirin prevents myocardial infarction
    - Jaundice is a typical symptom of hepatitis
Why formal ontology at all??

- Formal definitions create maximum consensus on the meaning of terms
  - Ontologies as standards
  - Reusable terms and axioms
- Formal axioms encode statements about what is considered to be universally true in a domain
  - in contrast to knowledge proper
- Formal axioms permit logic-based reasoning
  - Consistency checking
  - New entailments
  - Equivalence of syntactically heterogeneous expressions can be computed: semantically interoperable systems
Ontologies in life sciences and health care
Ontologies in life sciences and health care

- Bottom-up ontology development:
  OBO (Open biomedical Ontologies) Foundry

- Top-down ontology development:
  SNOMED CT
  (Systematized Nomenclature of Medicine – Clinical terms)
<table>
<thead>
<tr>
<th>RELATION TO TIME</th>
<th>CONTINUANT</th>
<th>OCCURRENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRANULARITY</td>
<td>INDEPENDENT</td>
<td>DEPENDENT</td>
</tr>
<tr>
<td>ORGAN AND ORGANISM</td>
<td>Organism (NCBI Taxonomy)</td>
<td>Anatomical Entity (FMA, CARO)</td>
</tr>
<tr>
<td>CELL AND CELLULAR COMPONENT</td>
<td>Cell (CL)</td>
<td>Cellular Component (FMA, GO)</td>
</tr>
<tr>
<td>MOLECULE</td>
<td>Molecule (ChEBI, SO, RnaO, PrO)</td>
<td>Molecular Function (GO)</td>
</tr>
</tbody>
</table>

The OBO Foundry

- Collaborative bottom up initiative, driven by the success of the Gene Ontology
- Rooted in upper ontologies (BFO + RO)
- Goal of creating a suite of orthogonal interoperable reference ontologies in the biomedical domain
- Moving from semi-formal OBO syntax to OWL-DL
- Cross-ontology definitional axioms:
  - *Calcitonin secreting cell* (Cell Ontology) can be defined as a *Secretory cell* which secretes *Calcitonin* (ChEBI)
  - *Heart development* (Gene Ontology) can be defined as a *Developmental process* which has *Heart* (FMA) as participant

SNOMED CT (Systematized Nomenclature of Medicine - Clinical Terms)

SNOMED CT

SNOMED CT is considered to be the most comprehensive, multilingual clinical healthcare terminology in the world. Each year, avoidable deaths and injuries occur because of poor communication between healthcare practitioners, or because busy practitioners forget or neglect to follow their own criteria for best practices. The delivery of a standard clinical terminology for use across the world’s health information systems can therefore make a significant contribution towards improving the quality and safety of healthcare.

SNOMED CT aims to contribute to the improvement of patient care through underpinning the development of systems to accurately record health care encounters and to deliver decision support to health care providers. Ultimately, patients will benefit from the use of SNOMED CT to more clearly describe and accurately record their care, in building and facilitating better communication and interoperability in electronic health record exchange, and in creating systems that support health care decision making.

SNOMED CT intellectual property rights were transferred to the SNOMED SDO® in the formal creation of the IHTSDO.

SNOMED CT was originally created by the College of American Pathologists by combining SNOMED RT and a computer based nomenclature and classification known as Clinical Terms Version 3, formerly known as Read Codes Version 3, which was created on behalf of the UK Department of Health and is Crown copyright.
SNOMED CT - clinical terminology with ontological foundations

- Terminology for clinical data covering diseases, findings, procedures, organisms, substances etc.
- 311,000 concepts, connected by 1,360,000 relational expressions
- Definitions with DL axioms
- Promoted as an international terminological standard
SNOMED CT: Terminology + Ontology

Concepts (representational units)

Current Concept:
Malignant tumor of breast (disorder)

Parent(s):
(Select a parent to make it the "Current Concept".)
Malignant neoplasm of thorax (disorder)
Neoplasm of breast (disorder)

Child(ren):
(N-16) (Select a child to make it the "Current Concept".)
Carcinoma of breast (disorder)
Familial cancer of breast (disorder)
Hormone receptor positive malignant neoplasm of breast (disorder)
Local recurrence of malignant tumor of breast (disorder)
Malignant lymphoma of breast (disorder)
Malignant melanoma of breast (disorder)
Malignant neoplasm of axillary tail of breast (disorder)
Malignant neoplasm of breast lower inner quadrant (disorder)
Malignant neoplasm of breast lower outer quadrant (disorder)
Malignant neoplasm of breast upper inner quadrant (disorder)
Malignant neoplasm of breast upper outer quadrant (disorder)
Malignant neoplasm of female breast (disorder)
Malignant neoplasm of male breast (disorder)
Primary malignant neoplasm of breast (disorder)
Sarcoma of breast (disorder)

Current Concept:
Fully Specified Name: Malignant tumor of breast (disorder)
ConceptId: 254837009

Defining Relationships:
Is a
Malignant neoplasm of thorax (disorder)
Neoplasm of breast (disorder)

Group 1
Associated morphology
Malignant neoplasm of primary, secondary, or uncertain origin (morphologic abnormality)
Breast structure (body structure)

Finding site
This concept is fully defined.

Qualifiers:
View Qualifying Characteristics and Facts

Descriptions (Synonyms):
Fully Specified Name: Malignant tumor of breast (disorder)
Preferred: Malignant tumor of breast [379661016]
Synonym: Breast cancer [379662011]
Synonym: CA - Breast cancer [379663018]
Preferred: Malignant tumour of breast [379664012]

Related Concepts:
- All "Is a" antecedents -
- All descendents/subtypes -
- Related concepts demo -

http://viw2.vetmed.vt.edu/sct/menu.cfm
Bioportal – repository for biomedical ontologies

Term Search
Search for a term in multiple ontologies

myocardium
advanced options

Myocardium - NCI Thesaurus
http://ncit.bioontology.org/ni/CI/thesaurus.owl#Myocardium
The stratified muscle tissue of the heart enveloped by the epicardium and the endocardium.
details - visualize - 2 more from this ontology

myocardium - Mouse adult gross anatomy
http://purl.obolibrary.org/obo/MA_0000164
details - visualize - 1 more from this ontology

myocardium - Radlex
http://purl.obolibrary.org/obo/RD/1398
details - visualize - 2 more from this ontology

Myocardium - Medical Subject Headings (MeSH)
http://purl.obolibrary.org/obo/WF/0000236
The muscle tissue of the heart. It is composed of striated, involuntary muscle cells (MYOCYTES, CARDIAC) connected to form the contractile pump to generate blood flow.
details - visualize - 1 more from this ontology

myocardium - Experimental Factor Ontology
http://www.ebi.ac.uk/efo/efo/0000919
Muscle layer of organ which has as its parts the myocardium proper and the conducting system of the heart.
details - visualize - 2 more from this ontology

Myocardium - Logical Observation Identifier Names and Codes
http://purl.obolibrary.org/obo/NCI/LON1987 &
details - visualize - 6 more from this ontology

Myocardium - Foundational Model of Anatomy
http://fma.bioontology.org/fma#Myocardium
Muscle layer of organ which has as its parts the myocardium proper and the conducting system of the heart.
details - visualize - 1 more from this ontology

Myocardium - Cell line ontology
http://purl.obolibrary.org/obo/FA/C_9462
details - visualize - 6 more from this ontology

Myocardium - Read Codes, Clinical Terms Version 3 (CTV3)
http://purl.obolibrary.org/obo/C3D/C000188

details - visualize - 3 more from this ontology

Submit your ontology to
http://bioportal.bioontology.org
Formal ontologies and beyond...

Medical Subject Headings (MeSH)

Details

- **Ontology ID:** 1351
- **BioPortal PURL:** http://purl.bioontology.org/ontology/MeSH
- **Status:** RRF
- **Categories:** Health
- **Groups:** Unified Medical Language System
- **Contact:** Stuart Nelson, M.D., nelson@nih.gov
- **Home Page:** http://www.nlm.nih.gov/mesh/meshhome.html
- **Publications Page:** http://www.nlm.nih.gov/mesh/meshhome.html
- **Documentation Page:** http://www.nlm.nih.gov/mesh/meshhome.html
- **Description:** Medical Subject Headings (MeSH)/National Library of Medicine; 2011
- **License Information:** This ontology is made available via the UMLS. Users of all UMLS ontologies must abide by the terms of the UMLS license, available at https://uts.nlm.nih.gov/license.html

Reviews

- **Add your review**

**Review by Steschu on 12/14/2012**

MeSH is not an ontology. It has never claimed to be one. Its concepts are not classes in the sense of OWL. Its hierarchical links are not subclass relations. If you interpret them as such you get strange inferences such as “Every thumb is a hand”. This would do injustice to MeSH, which is a great resource, which fulfills its goals without subscribing to OWL semantics.
Alternative to formal ontologies:
INFORMAL terminologies / thesauri
Alternative to formal ontologies: INFORMAL terminologies / thesauri

- Group together words / terms according to similarity in meaning
- Basic relations:
  - Synonymy
  - Broader / Narrower (ordering relations)
- Concept = Group of (quasi)synonyms
- Multiple hierarchies
- Mainly designed for retrieval
- Text definitions / explanations (scope notes) if required
- No formal semantics
<table>
<thead>
<tr>
<th>MeSH Heading</th>
<th>Breast Neoplasms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree Number</td>
<td>C04.588.180</td>
</tr>
<tr>
<td>Tree Number</td>
<td>C17.800.090.500</td>
</tr>
<tr>
<td>Annotation</td>
<td>human only; BREAST NEOPLASMS, MALE is also available; for animal use MAMMARY NEOPLASMS, ANIMAL 24.6+; coordinate IM with histological type of neoplasm (IM)</td>
</tr>
<tr>
<td>Scope Note</td>
<td>Tumors or cancer of the human BREAST.</td>
</tr>
<tr>
<td>Entry Term</td>
<td>Breast Cancer</td>
</tr>
<tr>
<td>Entry Term</td>
<td>Breast Tumors</td>
</tr>
<tr>
<td>Entry Term</td>
<td>Cancer of Breast</td>
</tr>
<tr>
<td>Entry Term</td>
<td>Cancer of the Breast</td>
</tr>
<tr>
<td>Entry Term</td>
<td>Human Mammary Carcinoma</td>
</tr>
<tr>
<td>Entry Term</td>
<td>Mammary Carcinoma, Human</td>
</tr>
<tr>
<td>Entry Term</td>
<td>Mammary Neoplasm, Human</td>
</tr>
<tr>
<td>Entry Term</td>
<td>Mammary Neoplasms, Human</td>
</tr>
<tr>
<td>Entry Term</td>
<td>Neoplasms, Breast</td>
</tr>
<tr>
<td>Entry Term</td>
<td>Tumors, Breast</td>
</tr>
<tr>
<td>Allowable Qualifiers</td>
<td>BL BS CF CH CI CL CN CO DH DI DT EC EH EM EN EP ET GE HI IM ME MI MO NU PA PC PP PS PX RA</td>
</tr>
<tr>
<td>Entry Version</td>
<td>BREAST NEOPL</td>
</tr>
</tbody>
</table>
Example: Medical Subject Headings (MeSH)

Literature search: monoclonal antibodies and cancer therapy

MeSH terms
"Nontologies"

- Use OWL syntax, which should not be interpreted according to description logics semantics
- Formal reasoning would lead to incorrect entailments
- Examples: NCI thesaurus, Radlex
- Many other ontologies contain problematic axioms that contradict the intended meaning
- Example (NCI thesaurus):
  
  \[
  \text{Calcium-Activated}\_\text{Chloride}\_\text{Channel-2} \sqsubseteq \text{Gene}\_\text{Product}\_\text{Expressed}\_\text{In}\_\text{Tissue} \text{some } \text{Lung} \text{ and } \text{Gene}\_\text{Product}\_\text{Expressed}\_\text{In}\_\text{Tissue} \text{some } \text{Mammary}\_\text{Gland} \text{ and } \text{Gene}\_\text{Product}\_\text{Expressed}\_\text{In}\_\text{Tissue} \text{some } \text{Trachea} \\
  \]
  
  \[
  \text{Ureter}\_\text{Small}\_\text{Cell}\_\text{Carcinoma} \sqsubseteq \text{Disease}\_\text{May}\_\text{Have}\_\text{Finding} \text{some } \text{Pain} \\
  \]
(N)ontologies of interest for imaging

- **Radlex**
  - 24800 classes covering anatomy, procedures, diseases, substances, devices, relevant for radiologic imaging
  - all classes are also individuals (punning)
  - Relations ('is a', 'part of') asserted are at the level of individuals
  - On classes no other axioms than subclass axioms

- **Foundational model of anatomy**
  - Complete model: Protégé Frames (no formal semantics)
    - Parts of it available as OWL
  - All assertions at class level
  - Logical entailments only true for "canonical" anatomy

- **SNOMED CT**
Challenges of "correct" ontology for image representation including simulation

- Same terms (e.g. "cardiac motion") can be used for different things
  1. A real cardiac motion in a patient (process)
  2. Part of an image (information entity) that represents a real cardiac motion
  3. A simulation artifact (information entity), which does not refer to any specific cardiac motion
  4. The "concept" cardiac motion (cognitive entity)

- If you prefer 4. or if the distinction between 1. – 4. does not matter, then you shouldn't use formal ontologies
Provision of controlled terms
  - Good text definitions should be available

Hierarchy expansion for retrieval
  - hierarchical links at the level of broader term / narrower terms
    - is-a $\rightarrow$ is narrower than
    - part-of $\rightarrow$ is narrower than

"Hand-crafted" inference rules
  - no use of description logics classifiers

Possible standard SKOS

http://www.w3.org/2004/02/skos/intro
Own experiences with ontologies in large projects

- EU funded projects with multiple partners
  1. @neurIST: Data integration (clinical, genomic, simulation) on cerebral aneurysms
  2. DebugIT: Decision support system for infectious diseases
  3. SemanticHealthNet: Semantic interoperability between heterogeneous semantic representations in the EHR

- Experiences:
  - in 1. and 2. much effort put in formal ontology
  - Mosty used as a controlled vocabulary (1.)
  - DL reasoning only for computing inferred ontology, which then used with production rules
  - 3. Formal foundation seems fundamental to reach the interoperability goal. However, intellectual input considerable and scalability still open
Current state of the art of Applied Ontology as a discipline
Current state of the art of Applied Ontology as a discipline

- Applied Ontology – still emerging discipline
- Prevalence of makeshift ontology artifacts
- Ontology engineering required to be more principled
- Necessary resources
  - Standards (Semantic Web – OWL)
  - Good practice guidelines (e.g. GoodOD Guideline)
  - Quality management
  - Best-of-breed examples
  - Industry-standard tools
    - Editors
    - Reasoners

http://www.iph.uni-rostock.de/GoodOD-Guideline.1299.0.html
Current state of the art of Applied Ontology as a discipline

Guideline on Developing Good Ontologies in the Biomedical Domain with Description Logics

URL: http://www.purl.org/goodod/guideline

Version 1.0
December 2012

Send feedback to:
martin.boeker@uniklinik-freiburg.de
ludger.jansen@uni-rostock.de

Schulz S\textsuperscript{1,*}, Seddig-Raufie D\textsuperscript{1}, Grewe N\textsuperscript{2}, Röhl J\textsuperscript{2}, Schober D\textsuperscript{1}, Boeker M\textsuperscript{1}, Jansen L\textsuperscript{2}

\textsuperscript{1}: Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg
\textsuperscript{2}: Institute of Philosophy, University of Rostock
\textsuperscript{3}: Department of Medical Informatics, University of Graz

11th December 2012
Current state of the art of Applied Ontology as a discipline

- Applied Ontology – still emerging discipline
- Prevalence of makeshift ontology artifacts
- Ontology engineering required to be more principled
- Necessary resources
  - Standards (Semantic Web – OWL)
  - Good practice guidelines (e.g. GoodOD Guideline)
  - Quality management
  - Best-of-breed examples
  - Industry-standard tools
    - Editors
    - Reasoners

http://www.iph.uni-rostock.de/GoodOD-Guideline.1299.0.html
Decision on using formal ontology in life science research projects

- Functional requirements
  - Controlled terminology
  - Query expansion for retrieval
  - Precise definitions of terms
  - Precise classification of domain entities
  - Reasoning to establish semantic equivalence
  - Representation of contingent knowledge
  - Default reasoning
  - Probabilistic reasoning
### Decision matrix

**Thesauri / Ontologies / KR formalism**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Thesauri</th>
<th>Ontology</th>
<th>KR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controlled domain language</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Query expansion for retrieval</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Precise definitions of terms</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Precise classification of domain entities</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Reasoning to establish semantic equivalence</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Representation of contingent knowledge</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Non-monotonic reasoning</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Probabilistic reasoning</td>
<td></td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>
Further readings
Ontology on the Web

- Description Logics: http://dl.kr.org/
- Protégé: http://protege.stanford.edu/
- Bioontology: http://www.bioontology.ch/
- Buffalo Ontology Site: http://ontology.buffalo.edu/smith/
- OBO Foundry: http://obofoundry.org/
- Bioportal: http://bioportal.bioontology.org/
- SNOMED CT: http://www.ihtsdo.org/snomed-ct/
  http://terminology.vetmed.vt.edu/sct/menu.cfm
- CO-ODE (Pizza ontology): http://www.co-ode.org/
- GoodOD Guideline: http://www.iph.uni-rostock.de/GoodOD-Guideline.1299.0.html