Informatica Biomedica

lezione19

Alberto Paoluzzi Mauro Ceccanti

www.dia.uniroma3.it/ paoluzzi/web/did/biomed/

Informatica e Automazione, "Roma Tre" — Medicina Clinica, "La Sapienza"

May 17, 2010

Fonti essenziali:

- Chapt. 6 of I.J. Kalet, *Principles of Biomedical Informatics*, Academic Press, 2009.
- From Wikipedia, the free encyclopedia

Informatica Biomedica: Lezione 19

Semantic Nets and Ontologies (1/3)

In large data stores such as GenBank, the sequences are often thought of as the primary data, and are the subject of much computation (sequence comparisions, sequence searching for small patterns like motifs).

- However, there is a lot of knowledge in the annotations, the text information that accompanies each sequence.
- This is generally not in an easily computable form, but it is computationally powerful material.

Semantic Nets and Ontologies (2/3)

For reasoning about cellular metabolism we need:

- information about which protein(s) a gene encodes,
- and what the function(s) and pathways in which it participates

Projects like GO are making these pieces of knowledge the basic elements of a computational scheme

Semantic Nets and Ontologies (3/3)

- http://www.geneontology.org/GO.doc.shtmlannotation
- The Ontologies

Much effort in bioinformatics is now being expended to extract this symbolic knowledge from the text of published articles and from data resources that have this knowledge embedded as text

Semantic Nets and Ontologies

We need a language that incorporates the idea of classification of entities, particularly hierarchical classification systems that are usually referred to as ontologies.

from Wikipedia

Ontology

(from the Greek $\nu\tau$: of being and $\lambda\gamma\iota\alpha$, -logia: science, study, theory) is the philosophical study of the nature of being, existence or reality in general, as well as the basic categories of being and their relations.

Traditionally listed as a part of the major branch of philosophy known as metaphysics, ontology deals with questions concerning what entities exist or can be said to exist, and how such entities can be grouped, related within a hierarchy, and subdivided according to similarities and differences.

Computer Science definition

- In computer science and information science, an ontology is a formal representation of the knowledge by a set of concepts within a domain and the relationships between those concepts.
- It is used to reason about the properties of that domain, and may be used to describe the domain.

An ontology provides a shared vocabulary, which can be used to model a domain — say, via:

- the type of concepts (class) and/or objects (instances) that exist
- their properties and relations

Frame systems

A really simple frame implementation can be used to illustrate essential ideas.

- While frames can be used to represent data, the emphasis here is on frame systems as knowledge representation languages.
- They are an example of a variety of knowledge representation systems called slot and filler systems,
- where the basic structural idea is to group together names of attributes and their values (of course the values can be other frames).

Frames for knowledge representation

In the field of Artificial Intelligence, a frame is a data structure introduced by Marvin Minsky in the 1970s that can be used for knowledge representation.
Such data structure is used to divide knowledge into substructures by representing stereotyped situations

- Like many other knowledge representation systems and languages, frames are an attempt to resemble the way human beings are storing knowledge.
- It seems like we are storing our knowledge in rather large chunks, and that different chunks are highly interconnected.
- In frame-based knowledge representations, knowledge describing a particular concept is organized as a frame.
- The frame usually contains a name and a set of slots.
- The slots describe the frame with attribute-value pairs [slotname, value]

Structure (anatomy) and function (physiology)

Anatomy is essential to clinical practice as well as biomedical research.

- Structure (anatomy) and function (physiology) are the dual framework for understanding
  1. health
  2. disease
  3. diagnosis
  4. and treatment.
- They are also the essential framework for describing biological knowledge from molecular biology to ecology and evolution.

The Foundational Model of Anatomy Ontology (FMA)

UW Foundational Model of Anatomy

The Foundational Model of Anatomy Ontology (FMA) is an evolving computer-based knowledge source for biomedical informatics.

- it is concerned with the representation of classes or types and relationships necessary for the symbolic representation of the phenotypic structure of the human body in a form that is understandable to humans and is also navigable, parseable and interpretable by machine-based systems
- Specifically, the FMA is a domain ontology that represents a coherent body of explicit declarative knowledge about human anatomy.
- Its ontological framework can be applied and extended to all other species.

Purpose of FMA (1/2)

The UW Foundational Model of Anatomy ontology makes available anatomical information in symbolic (non-graphical) form to knowledge modelers and other developers of applications for

- education,
- clinical medicine,
- electronic health record,
- biomedical research and
- all areas of health care delivery and management

Purpose of FMA (2/2)

- The intent is to assure - through the FMA - consistency, and ultimately standards, in anatomical class representation
- Thus, the FMA is a biomedical informatics resource for developing the anatomy content of applications that target specific user groups
- the FMA as such, is not designed as an end-user application for anatomy students, teachers or any other particular user group

Why Foundational? (1/2)

Anatomy provides the foundation for the other biomedical sciences or information domains

- For this reason, anatomy is the first subject taught in health related educational and training programs
- Manifestations of health and disease can be thought of as properties (i.e., attributes) of anatomical structures ranging in size from
  1. biological molecules to
  2. cells,
  3. tissues,
  4. organs,
  5. organ systems and
  6. body parts

Why Foundational? (2/2)

- Therefore, anatomy is fundamental to all biomedical sciences, and the classes or types represented in the Foundational Model of Anatomy ontology generalize to all biomedical domains
- In other words, it is not possible to represent or describe the content domains of other, non-anatomical, biomedical disciplines without explicitly or implicitly referring to anatomical entities
- For example, the circulation must take for granted the existence of the heart and blood vessels, and the same is true for gastritis and the stomach as well as for dementia and the brain
- This means that anatomy is foundational to non-anatomical biomedical disciplines because they reuse anatomical classes

The Components of the FMA

The Foundational Model of Anatomy ontology has four interrelated components:

1. Anatomy taxonomy (At),
2. Anatomical Structural Abstraction (ASA),
3. Anatomical Transformation Abstraction (ATA),
4. Metaknowledge (Mk),

Thus, the Foundational Model of Anatomy ontology may be represented by the abstraction:

FMA = (At, ASA, ATA, Mk)

Anatomy taxonomy (At),

Anatomy
classifies anatomical entities according to the characteristics they share (genus) and by which they can be distinguished from one another (differentia).

This is implemented as a class hierarchy

Anatomical Structural Abstraction (ASA),

Anatomical Structural Abstraction
specifies the part-whole and spatial relationships that exist between the entities represented in At;

- anatomy is not simply about classification of entity types, but about relationships between entities.
- These are structural relationships: so the second component of FMA is a large collection of structural relationships.
- These structural relationships express such ideas as containment, constituent parts, connectivity, etc.

Anatomical Transformation Abstraction (ATA),

Anatomical Transformation Abstraction
specifies the morphological transformation of the entities represented in At during prenatal development and the postnatal life cycle

- The FMA is designed to accommodate things that change with time, in order to describe embryological development.
- This is the motivation for the modeling of anatomical transformation.
- This part of the FMA is more a work in progress than the AT and the ASA.

Metaknowledge (Mk),

Metaknowledge
specifies the principles, rules and definitions according to which classes and relationships in the other three components of FMA are represented.

- The fourth component of the FMA is the meta-knowledge that consists of rules and principles that other components are required to follow.
- These are anatomical axioms

Levels of detail

The FMA includes entries for every element of human anatomy from the body, progressing in levels of detail down to the cellular and subcellular levels

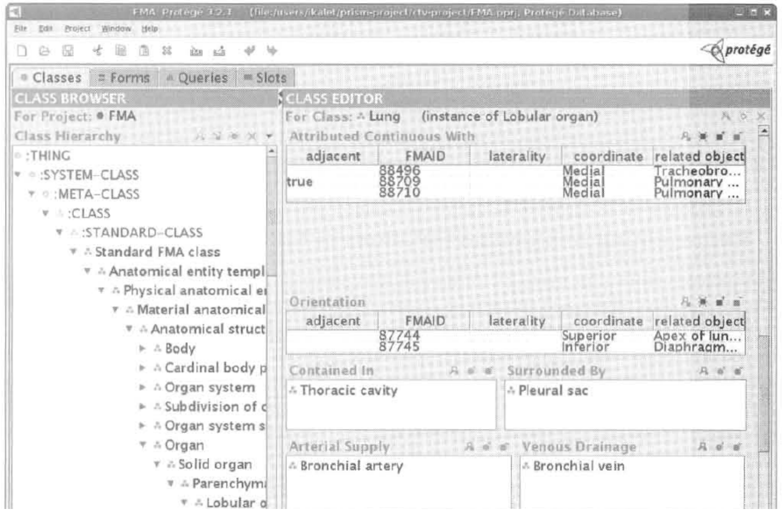
- Most importantly, the FMA represents relations between entities, not only in terms of a superclasses or subsumption hierarchy (class-subclass relationships) but also other relationships such as composition (various part-of relations), spatial relations, and connectivity (e.g., for the blood vessels and lymphatic systems, upstream and downstream connectivity)
- In addition to the general relationships already mentioned, there are specialized relationships that apply only to certain subclasses of anatomical entities

Spatial relationships

- The upstream and downstream relationships hold between entities of the same type, but there are also relationships between different types
- For example, arterial supply, venous drainage, and lymphatic drainage are relationships between types of vessels and types of organs
- In total there are nearly 200 relationship types in the FMA
- The FMA is organized as a class hierarchy, as seen in the Protégé screen capture below

A part of the FMA class hierarchy

A screen capture of a part of the FMA class hierarchy, displayed in the Protégé class browser window.



Metaclasses

In the FMA, every anatomical entity is modeled as a class, because the intent is to model canonical anatomy

- Building class hierarchies is commonplace in biomedical ontologies
- What is not commonplace is the rich model of relationships in the FMA
- These are among the many pieces of information that are associated with the FMA classes

Simple frame model

- A simple model would be to define a metaclass, that is, anatomical entity, as a frame with template slots that would be given values in all the instances
- not all anatomical entities have the same attributes, and so in the model, not all the instances of the anatomical entity class should have the same collection of slots, aside from the question of variations in their values
- For example, it makes sense for every entity to have a name, an ID, synonyms, and a description

For anatomy ontology it does not work

However, anatomy cannot be modeled in this simple way:

- it does not make sense for every entity to have a lymphatic drainage. In particular, a lymphatic chain or lymphatic vessel does not have a lymphatic drainage:
- Some entities participate in upstream and downstream connectivity relations, including blood vessels and lymphatic vessels, but most organs have no such relations
- Therefore, organs such as the heart, lung, muscle, etc should not have slots like *efferent to* or *afferent to*, which express upstream and downstream connectivity
- These must be defined as template slots by additional metac1asses, where needed

Large metaclass-class-instance hierarchy

The result of this modeling need is that the FMA has not only a very large class-subclass hierarchy but also a correspondingly large metaclass-class-instance hierarchy, with many levels

- Because of this need to model metaclass-class relations at many levels, no ordinary object-oriented modeling system will do the job
- A full-featured frame system is necessary

Knowledge bases downloads

- The FME also provides an easily available method of exploring (say, without installig Protégé) the FMA to individuals or groups considering the adoption of the Foundational Model of Anatomy knowledge base
- To download the knowledge base: look at UW-SIG Downloadable Software
- See Web enabled interface to FME

The Foundational Model Explorer

The Foundational Model Explorer (FME) is an Internet based software application developed for viewing the content and organization of the Digital Anatomist Foundational Model of Anatomy ontology (FMA)

- It was developed by the Structural Informatics Group at the University of Washington
- The initial purpose of the FME was to provide a simple and intuitive interface to the FMA for domain experts, in the field of anatomy, participating in the evaluation of the FMA