Ontological Foundations for Biomedical Sciences

Stefan Schulz, Kornél Markó, Udo Hahn

Department of Medical Informatics
University Hospital Freiburg (Germany)

Text Knowledge Engineering Lab
University of Jena (Germany)
The World of Life Sciences...

...requires sophisticated organization
Top-Level Division

**Occurrents:**
(Changes of) states of affairs of the physical world:

*Examples: process, state, event, disease, procedure...*

**Continuants:**
Entities of the physical world (*"Biomedical Structure"*):

*Examples: body, organ, tissue, molecule,..*

*disjoint*

*depend on*
Representation of Continuants in Bio-ontologies. What exists?

- Human Anatomy
  - Foundational Model of Anatomy (FMA)
  - Portions of SNOMED, OpenGalen, MeSH

- Other Organisms
  - Open Biological Ontologies (OBO)
    - Mouse (developmental stages), Zebrafish, Drosophila, ...
  - UMLS Semantic Network

- Species-Independent
  - Gene Ontology: Cellular Component branch

Size: 14 (UMLS SN) – $10^3$ (Adult Mouse) – $10^5$ (FMA)
Deficiencies of existing Bio-Ontologies

- Redundancy
- Synonymy
- Ambiguity
- Underspecification
Redundancy

Mouse (embryonal stage TS11, source: MGI)
- cardiovascular system
  -- heart
  -- cardiogenic plate

Mouse (embryonal stage TS18, source: MGI)
- cardiovascular system
  -- heart
  -- atrio-ventricular canal
  -- atrium
  -- bulboventricular groove
  -- bulbus cordis
  -- endocardial cushion tissue
  -- mesentery
  -- outflow tract
  -- pericardium
  -- primitive ventricle
  -- sinus venosus

Mouse (embryonal stage TS26, source: MGI)
- cardiovascular system
  -- heart
  -- aortic sinus
  -- atrio-ventricular canal
  -- atrio-ventricular cushion tissue
  -- atrium
  -- bulbar cushion
  -- endocardial cushion tissue
  -- endocardial tissue
  -- mesentery
  -- pericardium
  -- trabeculae carneae
  -- valve
  -- ventricle

Drosophila (adult, source: FlyBase)
- circulatory system
  -- heart
  -- heart muscle
  -- adult aortic funnel
  -- adult ostia
  -- dorsal diaphragm
  -- heart chamber
  -- terminal opening

Zebrafish (adult, source: ZFIN)
- cardiovascular system
  -- heart
  -- atrium
  -- bulbus arteriosus
  -- hypobranchial vessels
  -- sinus venosus
  -- ventricle

Human, Adult, (source: FMA)
- cardiovascular system
  -- heart
  -- wall of heart
  -- right atrium
  -- left atrium
  -- right ventricle
  -- left ventricle
  -- right side of heart
  -- left side of heart
  -- fibrous skeleton of heart
  -- papillary muscle
  -- cardiac valve
  -- tricuspid valve
  -- mitral valve
  -- aortic valve
  -- pulmonary valve
  -- interatrial septum
  -- (...)
Synonymy

- “Motor Neuron instance-of Neuron” (FlyBase)
- “Motor Neuron narrower Neuron” (MeSH)
- “Motor Neuron subclass-of Neuron” (FMA, OpenGALEN)
Ambiguity, Underspecification

“Cell has-part Axon” (Gene Ontology)
- Do cells without axons exist?
- Do axons without cells exist?

“Neuron has-part Axon” (FMA)
- Does every neuron have an axon?
Ambiguity, Underspecification

■ “Cell has-part Axon” (Gene Ontology)
  ■ Do cells without axons exist?
  ■ Do axons without cells exist?

■ “Neuron has-part Axon” (FMA)
  ■ Does every neuron have an axon?

“Keep in mind that part_of means can be a part of, not is always a part of “

“The part_of relationship (...) is usually necessarily is_part”
GO Editorial Style Guide, Jan 2004

“A part_of B if and only if: for any instance x of A there is some instance y of B which is such that x stands to y in the instance-level part relation, and vice versa”
Rosse & Smith MEDINFO 2004
Semantic framework for biological structure...

- Foundational Relations
- General Attributes
- Theories
Semantic framework for biological structure...

- Foundational Relations
- General Attributes
- Theories
## Bio-ontologies

<table>
<thead>
<tr>
<th>Occurrents</th>
<th>Continuants</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Changes of) states of affairs of the physical world:</td>
<td>Entities of the physical world</td>
</tr>
<tr>
<td><strong>Universals</strong> (Concepts, Classes of Individuals)</td>
<td>Life, Appendectomy, Mitosis</td>
</tr>
<tr>
<td><strong>Particulars</strong> (Concrete Objects in the world)</td>
<td>My Life, Appendectomy of Patient #123, this Mitosis</td>
</tr>
<tr>
<td></td>
<td>Hand, Blood, Cell, Tree</td>
</tr>
<tr>
<td></td>
<td>My Hand, Blood Sample #12345, this Cell, the Maple Tree in front of the house #xyz</td>
</tr>
</tbody>
</table>

Four disjoint partitions
Some Foundational Relations between Biological Continuants

<table>
<thead>
<tr>
<th>Rel(x,y)</th>
<th>Universals</th>
<th>Particulars</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>Universals</td>
<td>Particulars</td>
</tr>
<tr>
<td></td>
<td>Is-A</td>
<td>part-of, has-location has-branch, bounds, connects has-developmental-form</td>
</tr>
<tr>
<td></td>
<td>Instance-of</td>
<td></td>
</tr>
</tbody>
</table>
Some Foundational Relations between Biological Continuants

<table>
<thead>
<tr>
<th>Rel(x,y)</th>
<th>Universals</th>
<th>Particulars</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Universals</td>
<td>Is-A</td>
<td></td>
</tr>
<tr>
<td>Particulars</td>
<td>Instance-of</td>
<td>part-of, has-location, has-branch, bounds, connects, has-developmental-form</td>
</tr>
</tbody>
</table>
Some Foundational Relations between Biological Continuants

<table>
<thead>
<tr>
<th>Rel(x,y)</th>
<th>x</th>
<th>Universals</th>
<th>Particulars</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><strong>Is-A, Part-Of, Has-Location Bounds, Has-Branch, Connects Has-Developmental-Form</strong></td>
<td><strong>part-of, has-location has-branch, bounds, connects has-developmental-form</strong></td>
</tr>
<tr>
<td>Universals</td>
<td>x</td>
<td>Universals</td>
<td>Particulars</td>
</tr>
<tr>
<td>Particulars</td>
<td>y</td>
<td>Instance-of</td>
<td></td>
</tr>
</tbody>
</table>

This table illustrates the foundational relations between biological continuants, focusing on the universals and particulars that define these relationships.
From Instance-to-Instance relations to Class-to-Class Relations

$A, B$ are classes, $\text{inst-of} = \text{class membership}$

$\text{rel} : \text{relation between instances}$ $\text{Rel} : \text{relation between classes}$

$$\text{Rel} (A, B) =_{\text{def}} \exists x : \text{inst-of} (x, A) \land \text{inst-of} (y, B) \land \text{rel} (x, y) \quad \text{OR}$$

$\text{cf.}$

Rosse & Smith (MEDINFO 2004)
From Instance-to-Instance relations to Class-to-Class Relations

$A, B$ are classes, $\text{inst-of} = \text{class membership}$

$\text{rel: relation between instances}$ $\text{Rel: relation between classes}$

$\text{Rel } (A, B) = \text{def}$

1. $\exists x: \text{inst-of} (x, A) \land \text{inst-of} (y, B) \land \text{rel} (x, y)$
2. $\forall x: \text{inst-of}(x, A) \rightarrow \exists y: \text{inst-of} (y, B) \land \text{rel} (x, y)$

$\text{cf.}$
$\text{Schulz & Hahn (KR 2004, ECAI 2004)}$
$\text{Rosse & Smith (MEDINFO 2004)}$
From Instance-to-Instance relations to Class-to-Class Relations

$A, B$ are classes, $\text{inst-of} = \text{class membership}$

$\text{rel}: \text{relation between instances}$

$\text{Rel}: \text{relation between classes}$

$\text{Rel} (A, B) =_{\text{def}}$

1. $\exists x: \text{inst-of} (x, A) \land \text{inst-of} (y, B) \land \text{rel} (x, y)$
2. $\forall x: \text{inst-of} (x, A) \rightarrow \exists y: \text{inst-of} (y, B) \land \text{rel} (x, y)$
3. $\forall y: \text{inst-of} (y, B) \rightarrow \exists x: \text{inst-of} (x, A) \land \text{rel} (x, y)$

$\text{cf.}$

Rosse & Smith (MEDINFO 2004)
Semantic framework for biological structure...

- Foundational Relations

- General Attributes

- Theories
General Attributes (mutually disjoint classes)

- Dimensionality: Point, 1-D, 2-D, 3-D
General Attributes (mutually disjoint classes)

- **Dimensionality**: Point, 1-D, 2-D, 3-D

- **Solids vs. hollow spaces, vs. Boundaries**
General Attributes (mutually disjoint classes)

- Dimensionality: Point, 1-D, 2-D, 3-D
- Solids vs. hollow spaces, vs. Boundaries
- Collections vs. Masses vs. Count Objects

cf. Schulz & Hahn, FOIS 01
Semantic framework for biological structure...

- Foundational Relations
- General Attributes
- Theories
Theories

- A set of formal axioms which describe a restricted (local) domain.
- Four orthogonal theories for Biological Structure
  - Granularity
  - Species
  - Development
  - Canonicity
Theories

- A set of formal axioms which describe a restricted (local) domain.

- Four orthogonal theories for Biological Structure
  - Granularity
  - Species
  - Development
  - Canonicity
Granularity

- Level of detail (molecular, cellular, tissue, organ)
- Change in Granularity level may be non-monotonous
  - Change of sortal restrictions:
    - 3-D $\rightarrow$ 2-D boundary
    - Count concept $\rightarrow$ Mass concept
  - Change of relational attributions:
    - disconnected $\rightarrow$ connected
Theories

- A set of formal axioms which describe a restricted (local) domain.
- Four orthogonal theories for Biological Structure
  - Granularity
  - Species
  - Development
  - Canonicity
Linnean Taxonomy of Species

Deuterostomia
  Anthrospoda
    Onychophora (velvet worms)
    Tardigrada (water bears)
    Nemastomata (roundworms)
    Nematomorpha (horsehair worms)
    Kinorhyncha
    Lorichia
    Priapulida (penis worms)
    ? Platynematida (flatworms, tapeworms, flukes)
    ? Chaetognatha (arrow worms)
    ? Gastrotricha
    ? Rotifera (rotifers)
    ? Gnathostomulida (jaw worms)
    ? Micragnostoza
    ? Cyclophora
    ? Mesozoza
  Annelida (bristleworms, ragworms, earthworms, leeches and their allies)
    sipuncula (peanut worms)
    mollusca (snails, clams, squids, etc.)
    nemertea (ribbon worms)
    bryozoa (moss animals)
    Entoprocta (kamptozoans)
  Echinodermata
    Callichordata
      Urochordata (tunicates)
      Cephalochordata
        Yunnanozoan
        Cranial
        Euprotreta
        Vertebrata
        Hypercota
          ? Eucoronta
          Pteraspidomorpha
            ? Ectodonta
            ? Thelodonti
            ? Anaspida
            Galeaspida
            ? Pituriaspida
            Osteostraci
            Gnathostomata
            ? Sarcopterygi
            Actinopterygi
            Acanthodii
            Chondrichthyes
            Placodermi
            ? Coelacanthimorpha
            Porolepimorpha
            Ditoriiformes
            Rhizodontimorpha
            Osteolepimorpha
            Terrestrial Vertebrata

THE TREE OF LIFE WEB PROJECT
http://tolweb.org
Linnean Taxonomy of Species
Linnean Taxonomy of Species
Species

Introduction of axioms at the highest common level
Theories

- A set of formal axioms which describe a restricted (local) domain.
- Four orthogonal theories for Biological Structure
  - Granularity
  - Species
  - Development
  - Canonicity
Development

- Represents time-dependent "snapshots" from the life cycle of an organism, e.g., zygote, embryo, fetus, child, adult
- Development stages are species-dependent e.g. metamorphosis
Theories

- A set of formal axioms which describe a restricted (local) domain.
- Four orthogonal theories for Biological Structure
  - Granularity
  - Species
  - Development
  - Canonicity
Canonicity

Degrees of “Wellformedness” of Biological Structure:

- Canonic structure
Canonicity

Degrees of “Wellformedness” of Biological Structure:

- Canonic structure
- Structural Variations
Canonicity

- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure
  - Structural Variations
  - Pathological Structure

- Congenital
- Acquired
Canonicity

- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure
  - Structural Variations
  - Pathological Structure
  - Lethal Structure
Canonicity

Degrees of “Wellformedness” of Biological Structure:
- Canonic structure
- Structural Variations
- Pathological Structure
- Lethal Structure
- Derivates of biological structure
Canonicity

- Five canonicity levels: each level introduces axioms valid for higher levels

<table>
<thead>
<tr>
<th>Level</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theory</td>
<td>any amount of matter, if of biological origin</td>
<td>any living or dead organism</td>
<td>any living organism</td>
<td>living organism without pathologic modifications</td>
<td>ideal organism</td>
</tr>
<tr>
<td>Set of Axioms</td>
<td>$n_1$</td>
<td>$n_2$</td>
<td>$n_3$</td>
<td>$n_4$</td>
<td>$n_5$</td>
</tr>
</tbody>
</table>
## Examples

<table>
<thead>
<tr>
<th>Granularity</th>
<th>low</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>general</td>
<td></td>
<td></td>
</tr>
<tr>
<td>specific</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Species</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>embryo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>adult</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Development</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>high</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Coverage: Foundational Model of Anatomy

Granularity: Low                                     High
Species: General                                    Specific
Development: Embryo                                 Adult
Canonicity: Low                                     High
Coverage: Gene Ontology

Granularity:
- low
- high

Species:
- general
- specific

Development:
- embryo
- adult

Canonicity:
- low
- high
Examples

Connects (RightVentricle, Left Ventricle)

- Granularity = normal
- Species = mammal
- Development = adult
- Canonicity = 4-5

false

- Granularity = any
- Species = vertebrate
- Development = early embryo
- Canonicity = any

true

Is-A (Membrane, 3-D object)

- Granularity = normal
- Species = any
- Development = any
- Canonicity = any

true

- Granularity = lowest
- Species = any
- Development = any
- Canonicity = any

false
Conclusion

Integration of bio-ontologies requires

- Uncontroversial semantics of relations and attributes
- Clear commitment to theories, such as granularity, species, development and canonicity

Redundancy can be avoided

- Encoding axioms at the highest common level in the species taxonomy (e.g. vertebrates, arthropods, primates) and benefit from inheritance in subsumption hierarchies
...requires sophisticated organization

- Formalization and Standardization of Clinical Terminologies
- Basis for the Annotation of Genes and Gene Products
- Semantic reference for scientific communication
- Machine-supported reasoning and decision-support

Bio-ontologies!
## Upper level classification of entities

<table>
<thead>
<tr>
<th>Continuants (physical objects, ...)</th>
<th>Individuals (concrete objects)</th>
<th>Universals (Concepts, Classes of Individuals)</th>
</tr>
</thead>
<tbody>
<tr>
<td>my left hand</td>
<td>Hand, Blood, Cell</td>
<td></td>
</tr>
<tr>
<td>a blood sample</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a concrete cell</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Occurrents (events, processes, actions...)</th>
<th>Individuals (concrete objects)</th>
<th>Universals (Concepts, Classes of Individuals)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peter’s diabetes</td>
<td></td>
<td>Diabetes mellitus</td>
</tr>
<tr>
<td>appendectomy of Patient #12345</td>
<td></td>
<td>Appendectomy</td>
</tr>
</tbody>
</table>
Circulatory Routes
Systemic Arteries
above the diaphragm
Mereotopological Quiz

• Cranial Cavity *has-location* Head
  Is Cranial Cavity *part-of* Head? 🙄declspec.epic

• Brain *has-location* Cranial Cavity
  Is Brain *part of* Cranial Cavity? 🙄declspec.epic

• Glioblastoma *has-location* Brain
  Glioblastoma *part-of* Brain? 🙄declspec.epic

• Brain metastasis *has-location* Brain
  Brain metastasis *part-of* Brain? 🙄declspec.epic

• Embryo *has-location* Uterus
  Embryo *part-of* Uterus? 🙄declspec.epic

Images from: Sobotta CD-ROM
transitive closure by taxonomic subsumption
Subtheories of an Ontology of Biological Structure

1. Taxonomy
2. Mereology
3. Topology

Mereology

Hand

Thumb

Thumbnail

"is-a"

part-of"connection"
Subtheories of an Ontology of Biological Structure

- Taxonomy
- Mereology
- Topology

Canonical relationships

(Schulz et al. AMIA 2000)

Topological Primitives:
Structure of Talk

- Introduction
- Foundational Relations
- Foundational Attributes
- Theories
  - Granularity
  - Species
  - Development
  - “Canonicity”
The World of Life Sciences...
Generalized Representation of Living Systems: Top Level

Biological Entities

Biological Occurrents

process, state, event, ...

dependence

Biological Continuants

organism, organ, tissue, cell, molecule, ..
Ontological Account for Biological Continuants

- Foundational Relations
- Foundational Attributes
- Theories
  - Granularity
  - Species
  - Development
  - “Canonicity”
Granularity

- **Taxonomic:** degree of specialization

- **Mereologic:** degree of dissection
  
  \{molecular level, cellular level, tissue level, organ level, population level\}
Change in Granularity level may be non-monotonous

- Change of sortal restrictions:
  - 3-D $\rightarrow$ 2-D boundary
  - Count concept $\rightarrow$ Mass concept

- Change of relational attributions:
  - disconnected $\rightarrow$ connected
Canonicity

- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure
Canonicity

- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure
  - Structural Variations
Canonicity

- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure
  - Structural Variations
  - Pathological Structure
Canonicity

- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure
  - Structural Variations
  - Pathological Structure
  - Lethal Structure
Canonicity

- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure
  - Structural Variations
  - Pathological Structure
  - Lethal Structure
  - Derivates of biological structure