Representing Natural Kinds by Spatial Inclusion and Containment

Stefan Schulz
Medical Informatics, Freiburg University, Germany
http://www.imbi.uni-freiburg.de/medinf

Udo Hahn
Language and Information Engineering (JULIE) Lab, Jena University, Germany
Ontologies of Biological Structure

- Representation of Physical Parts of Organisms
- Large terminological repositories in biology exist and grow ...
  - Foundational Model of Anatomy: Human
    sig.biostr.washington.edu/projects/fm/
  - Open Biological Ontologies (OBO): Mouse, Fly, Fish, Worm, Fungi, ...
    obo.sourceforge.net
  - Gene Ontology (GO) cellular component: Species independent
    www.geneontology.org/
Canonic Representation of Mereotopological Structure

- \( \text{rel } (\text{Class}_1, \text{Class}_2) \), e.g.
  \( \text{part-of } (\text{CellNucleus}, \text{Cell}) \)
Canonic Representation of Mereotopological Structure

- \( \text{rel} (\text{Class}_1, \text{Class}_2), \text{ e.g.} \)
  
  \( \text{part-of} (\text{CellNucleus}, \text{Cell}) \)

Open questions:

- What is the meaning of mereotopological relations in Biology
- Class level reading of mereotopological relations, how to interpret?
Canonic Representation of Mereotopological Structure

- \( \text{rel } (\text{Class}_1, \text{Class}_2) \), e.g.
  - \( \text{part-of } (\text{CellNucleus, Cell}) \)

Open questions:

- What is the meaning of mereotopological relations in Biology
- Class level reading of mereotopological relations, how to interpret?
Fuzziness of Mereotopological Relations in Biology

- a Cell Membrane has-part a Cell
- a Mitochondrion part-of a Cell
- some Glucose part-of a Cell
- some Virus DNA part-of a Cell
Fuzziness of Mereotopological Relations in Biology

Parthood

- a Cell Membrane
  - has-part
  - part-of
  - a Cell

- a Mitochondrion
  - has-location
  - location-of
  - a Cell

- some Glucose
  - contains
  - containee-of
  - some Virus DNA

- a Cell
  - has
  - part
  - has-location

Containment

- a Cell Membrane
  - has-part
  - part-of
  - a Cell

- a Mitochondrion
  - has-location
  - location-of
  - a Cell

- some Glucose
  - contains
  - containee-of
  - some Virus DNA

- a Cell
  - has
  - part
  - has-location
Canonic Representation of Mereotopological Structure

- $\text{rel (Class}_1, \text{Class}_2)$, e.g. $\text{part-of (CellNucleus, Cell)}$

Open questions:
- What is the meaning of mereotopological relations in Biology
- Class level reading of mereotopological relations, how to interpret?
Class level reading of mereotopological relations

$R(C_1, C_2)$ expression, e.g. $Includes(\text{Cell}, \text{CellNucleus})$

Alternative, conflicting interpretations:

- expresses ontological dependencies
  - “each cell includes a cell nucleus” and / or
  - “each cell nucleus is included in a cell”

- permits possible relations
  - “a cell \textit{may} include a cell nucleus”

- reject any assertion not sanctioned by a $R(C_1, C_2)$ expression:
  - “there is a cell nucleus which included in a protein molecule”
Instance and Class level reading of mereotopological relations (I)

**Instance level:**

containee-of, includes: transitive, reflexive, antisymmetric

\[ \text{containee-of}(x,y) \leftrightarrow \text{includes}(y,x) \]

**Class level:**

- class A is a **specific containee** of class B:

\[
SC(A,B) = \text{def} \quad \forall x: \text{instance-of}(x, A) \rightarrow \\
\exists y: \text{instance-of}(y, B) \land \text{containee-of}(y,x)
\]

- class B is a **specific includer** of class A:

\[
SI(B,A) = \text{def} \quad \forall y: \text{instance-of}(y, B) \rightarrow \\
\exists x: \text{instance-of}(x, A) \land \text{includes}(y,x)
\]
**Instance and Class level reading of mereotopological relations (II)**

- \( SC(A,B) \) is not the inverse of \( SI(B,A) \)
- Class A is an **obligatory containee** of class B
  \[ OC(A,B) \iff_{def} SI(B,A) \]
- Class B is an **obligatory includer** of class A
  \[ OI(B,A) \iff_{def} SC(A,B) \]
- \( SC, SI, OC, OI \): transitive, reflexive, antisymmetric
- \( SC \) and \( SI \) propagate via Is-A:
  if A is a specific containee (includer) of B, every subclass of A is a specific containee (includer) of B, too
- \( OC \) and \( OI \) do not propagate via Is-A:
  if A is an obligatory containee (includer) of B, not any subclass of A is an obligatory containee (includer) of B
<table>
<thead>
<tr>
<th>Arg1</th>
<th>Arg2</th>
<th>Cell</th>
<th>Neuron</th>
<th>Axon</th>
<th>Cell Membrane</th>
<th>Organ</th>
<th>Cell Nucleus</th>
<th>Heart</th>
<th>Water</th>
<th>Blood</th>
<th>Insulin</th>
<th>Electron</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell</td>
<td>Is-A⁻¹</td>
<td>OI</td>
<td>SI</td>
<td>SC</td>
<td>SC</td>
<td>SC</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
</tr>
<tr>
<td>Neuron</td>
<td>Is-A</td>
<td>OI</td>
<td>SI</td>
<td>SC</td>
<td>SC</td>
<td>SC</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
</tr>
<tr>
<td>Axon</td>
<td>SC</td>
<td>SC</td>
<td>OI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
</tr>
<tr>
<td>Cell Membrane</td>
<td>SC</td>
<td>OC</td>
<td>OC</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
</tr>
<tr>
<td>Organ</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SC</td>
<td>SC</td>
<td>SC</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
</tr>
<tr>
<td>Cell Nucleus</td>
<td>SC</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>SC</td>
<td>SI</td>
<td>OC</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
</tr>
<tr>
<td>Heart</td>
<td>SI</td>
<td>SI</td>
<td>IS-A</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
</tr>
<tr>
<td>Water</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>SC</td>
<td>SC</td>
<td>SC</td>
<td>SI</td>
<td>SC</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
</tr>
<tr>
<td>Blood</td>
<td>SI</td>
<td>SI</td>
<td>OC</td>
<td>SI</td>
<td>SC</td>
<td>SC</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
</tr>
<tr>
<td>Insulin</td>
<td></td>
<td></td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>SI</td>
</tr>
<tr>
<td>Electron</td>
<td></td>
<td></td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
<td>OC</td>
</tr>
</tbody>
</table>
Extended Taxonomies

- Express mereotopological hierarchies as taxonomies
- Purpose: Better performance in large knowledge bases
- Introduction of reificator nodes:
  - real classes:
    \[ I_s-A(A,B_{SC}) =_{def} SC(A,B) \]
    \[ I_s-A(A,B_{SI}) =_{def} SI(A,B) \]
  - “pseudo classes” (do not capture properties of all instances!):
    \[ I_s-A(A,B_{OC}) =_{def} OC(A,B) \]
    \[ I_s-A(A,B_{OI}) =_{def} OI(A,B) \]

A must be terminal nodes
Extended Taxonomies

Cell

Water

Intracellular Membrane

Organ

Interphase Eukaryotic Cell

Cell Nucleus

Nuclear Membrane
Extended Taxonomies

Cell

Water

Intracellular Membrane

Specific Containees of Cell Nucleus

Cell Nucleus
Nuclear Membrane
Extended Taxonomies

Specific Includers of Cell Nucleus

Interphase Eukaryotic Cell, Tissue, Organ, Blood, Organism (…)

Organ
Interphase Eukaryotic Cell
Cell Nucleus
Nuclear Membrane
Extended Taxonomies

**Obligatory Containees of Cell Nucleus**

- Cell Nucleus
- Nuclear Membrane
- Intracellular Membrane
- Water

**Cell**

- **Organ**
- **Interphase Eukaryotic Cell**
- **Cell Nucleus**
- **Nuclear Membrane**
- **Water**

**Extended Taxonomy Diagrams**
Extended Taxonomies

Obligatory Includers of Cell Nucleus

Interphase Eukaryotic Cell, Cell
Conclusion

- Capturing mereotopological basics in biomedical ontologies: Two recommendations:
  - Create consensus by conflating part-whole and locative relations to one base relation (containee-of / includes)
  - Eliminate ambiguity by explicitly introducing class level relations with a precise semantics

- Using extended taxonomies: Improve reasoning in large knowledge bases

- To do: express other mereotopological relations (overlap, disconnectedness) in terms of class-level predicates
www.imbi.uni-freiburg.de/medinf