BioTopLite: An Upper Level Ontology for the Life Sciences
Evolution, Design and Application

Stefan Schulz\textsuperscript{1,2}, Martin Boeker\textsuperscript{2}

\textsuperscript{1} Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Austria
\textsuperscript{2} Institute of Medical Biometry and Medical Informatics, University Medical Center, Freiburg, Germany
Good reasons for highly constrained upper-level ontologies (ULOs)

- Guidance of ontology engineering process
- Standardization of ontology artefacts
- Decrease degrees of freedom  
  → increase in interoperability
- Prevention of design errors
- Enforcement of a coherent upper-level view on the world
  - Example: division continuants – occurrents
  - information entities – real entities
Existing ULOs have shortcomings for use in the biomedical domain

- DOLCE: focussed on cognitive science, atemporal models
- BFO, current version 1.1 has no relations, unintuitive labels, version 2.0 under development, controversial issues
- GFO-Bio: difficult to understand
- UMLS Semantic network, GALEN upper level, Semanticscience Integrated Ontology: overly pragmatic, no principled criteria for upper-level divisions, more language-oriented than philosophy-oriented
Evolution of BioTop since 2006 (I)

- Inspired by GENIA ontology: fixing of issues and broadening of scope
- Design characteristics
  - OWL-DL
  - Mutually exclusive upper-level categories
  - Limited set of relations
  - Highly constrained
  - Understandable naming
- Focus on cell biology and biomolecules
Evolution of BioTop since 2006 (II)

- Use for ontology building in large EU projects:
  - @neurist (neurology, neurosurgery)
  - DebugIT (nosocomial infections)

- Need for increased performance
  - Separation of many biochemistry related classes → ChemTop (no further maintained, due to evolution of ChEBI)

- Alignment with UMLS Semantic Network
  - Addition of medicine-related classes
  - External criterion for scoping

- Alignment with upper-level ontologies
  - BFO, RO, DOLCE
Evolution of BioTop since 2006 (III)

- Creation of "lite" version BTL
- Testing in ontology tutorials
- Basis for Guideline for "Good ontology design"
- Use in experimental ontologies in the SNOMED CT development process
- New requirements, e.g.
  - Causation
  - Inherent ambiguity of medical terms, e.g.
    - "Fracture" structure or process
    - "Allergy" process or disposition
  - Condition equivalentTo
    - MaterialObject or Disposition or Process
BioTopLite -> BioTopLite2
Main principles

- Pragmatic realist view
- Agnostic stance with regard to the existence of universals
- Compulsive use of top-level: domain classes must be placed under them
- Flat hierarchy, no top-level classes like Continuant, Occurrent
- Intuitive naming of classes and relations ('has locus' -> 'is Included in')
- Information object as toplevel class
- Set of relations (object properties) considered as closed. Relational predicates to be reified form process subclasses
- Further reduction of object properties ('processual part of' -> 'part of')
- All instances are considered to be temporally qualified (since ternary relations like 'part of' (a, b, t) not possible)
### BioTopLite2: Characteristics

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classes</td>
<td>53</td>
</tr>
<tr>
<td>Object properties</td>
<td>37</td>
</tr>
<tr>
<td>Logical axioms</td>
<td>240</td>
</tr>
<tr>
<td>- Subclass axioms</td>
<td>172</td>
</tr>
<tr>
<td>- Equivalence axioms</td>
<td>5</td>
</tr>
<tr>
<td>- Disjointness axioms</td>
<td>14</td>
</tr>
<tr>
<td>- CGI</td>
<td>14</td>
</tr>
<tr>
<td>- SubObjectProperty axioms</td>
<td>25</td>
</tr>
<tr>
<td>- Property chains</td>
<td>2</td>
</tr>
</tbody>
</table>
BioTopLite2: Classes

- 'immaterial object'
- 'immaterial three dimensional physical entity'
- 'one dimensional physical entity'
- 'two dimensional physical entity'
- 'zero dimensional physical entity'
- 'information object'
- 'plan'
- 'quality'
- 'canonicity'
- 'object quality'
- 'process quality'
- 'role'
- 'temporal region'
- 'point in time'
- 'time interval'
- 'value region'
BioTopLite2: Object Properties

- 'is condition of'
- 'is included in'
  - 'inheres in'
  - 'is part of'
    - 'is boundary of'
    - 'is component part of'
    - 'is granular part of'
- 'is participant in'
  - 'has life'
  - 'is agent in'
  - 'is outcome of'
  - 'is patient in'
  - 'is preceded by'
  - 'is projection of'
  - 'is referred to at time'
  - 'is represented by'
- precedes
- 'projects onto'
- represents
### BTL2 example: Axioms for "material object"

- 'has component part' **only** 'material object'
- 'has granular part' **only** 'material object'
- 'has life' **only** life
- 'has life' **some** life
- 'has part' **only** ('material object' or 'immaterial object')
- 'is bearer of' **only** (disposition or 'information object' or 'object quality' or function or role)
- 'is included in' **only** 'entity at some time'
- 'is part of' **only** 'material object'
- 'is participant in' **only** process
- 'projects onto' **only** ('immaterial three dimensional physical entity' and 'entity at some time')
- 'projects onto' **some** 'immaterial three dimensional physical entity'
- includes **only** (disposition or function or 'immaterial object' or 'information object' or 'material object' or quality or role)
- includes **only** 'entity at some time'
- **particular**

#### condition

`bClass Of (Anonymous Ancestor)

- 'is represented by' **only**
  - ('entity at some time'
    - and 'information object')

- 'is condition of' **only** situation
- disposition
  - or function
  - or 'material object'
  - or process`
Temporally qualified entities (I)

- Problem: relations between continuant (endurant) objects are time-dependent, e.g.
  - part-of (Heart#1234, John, 20130101)
  - part-of (Heart#1234, Jack, 20130103)

- OWL object properties are only binary

- Proposed solution (implemented in BTL2, currently under discussion for BFO2):
  - Instances of continuant entities are considered to be temporally qualified, such as
    Heart#1234@20130101 rdf:Type Heart
    Heart#1234@20130103 rdf:Type Heart
Temporally qualified entities (II)

- Consequences of temporally qualified continuants at the class (Tbox) level:
  - Class *Entity at some* useful as a means to enforce that instances of time-dependent classes be placed in a temporal context
  - 'Material object' subClassOf
    - 'is included in' only 'Entity at some time'

- Advantage: expression of temporary relatedness:
  - 'Structured biological entity' subClassOf
    - 'at some time' some ('is part of' some *Organism*)
  - 'at some time' relates a temporally qualified continuant with any of its temporally qualified "siblings"
Current use of BTL / BTL2

- SemanticHealthNet EU Network of Excellence: Upper level for information model and clinical terminology (SNOMED CT)
  http://www.semantichealthnet.eu

- CELDA: ontology of cell types, in vitro as well as in vivo, based on species, anatomy, subcellular structures, developmental stages and origin
  http://cellfinder.org/about/ontology

- International Health Terminology Standards Development Organization: in several experimental ontologies (event, condition, episode; observables)
Access to BioTop and BioTopLite

- Website: http://purl.org/biotop
- Mailing list: https://groups.google.com/forum/#!forum/biotop
- Feedback welcome!
Acknowledgements

People who contributed to the development of BioTop or provided important ideas:

Elena Beisswanger, Edward Cheetham, Bruce Goldberg, Udo Hahn, Robert Hoehndorf, Ludger Jansen, Catalina Martínez-Costa, Alan Rector, Daniel Schober, Kent Spackman, Holger Stenzhorn