OBO Format and Common Logic

History

- 2000/2001
  - OBO format 1.0 created
    - simple
    - graph-oriented
    - cvs-able
- 2004-2006
  - OBO format 1.2 and related efforts
    - http://www.geneontology.org/GO.format.obo-1_2.shtml
    - backwards and forwards incompatible with 1.0
    - Extended to allow simple genus-differentia logical definitions
    - Mappings to OWL
- 2009
  - OBO format 1.3 and CL
    - Backwards and forwards compatible with 1.2
    - Minor extensions
      - relation intersections, unions, compositions
    - http://www.geneontology.org/GO.format.obo-1_3.shtml
    - Formal semantics specified in terms of CL

ISO Common Logic

- Standard for first-order logic (FOL)
- Family of syntaxes
  - CLIF
  - CL-XML
- A CL *text* is a collection of *sentences* (axioms)
  - Sentences can be
    - atomic; e.g. part_of(nucleus,cell)
    - boolean; not has_part(mammalian_erythrocyte,nucleus)
    - quantified; e.g. all nnucleus(n) → exist c,tpart_of(n,c) ^ cell(c)
- Why do we need this when we have OWL/OWL2?
  - RO uses n-ary relations
  - Difficulties defining relations in OWL
  - Type level relations
  - FOL has been around for over a century

OBOF <-> CL

- OBOF tags mapped to CL predicates
  - id: ?r
  - is_transitive: true ➔transitive(?r)
- CL predicates defined in a set of axioms called obolog
  - transitive(*rel*) ∧ *rel*(X, Y) ∧ *rel*(Y , Z) → *rel*(X, Z)
  - transitive(*rel*) ∧ *rel*(x, y, t) ∧ *rel*(y, z , t) → *rel*(x, z , t)

- Every .obo file is a CL text
- This mapping is invisible and not directly to the majority of users
  - .obo becomes a simple surface syntax for CL
- Most OBO files use a subset including only atomic sentences
  - Exceptions: relation definitions

Examples

- http://www.fruitfly.org/~cjm/ro/ro.html
- http://www.fruitfly.org/~cjm/ro/ro-gaz.html
- http://www.fruitfly.org/~cjm/ro/ro-development.html

- SO genome interval relations
  - x starts y *iff* $\alpha(x) = \alpha(y) \wedge \omega(x) < \omega(y)$

Mapping to OWL

- OBOF mapping to OWL now specified in terms of CL
- 3 different mappings
  - OWL-Full, with type-level relations
  - Two DL mappings depending on treatment of time

Reasoning

- Relation reasoning
  - Theorem provers over CL
- Ontology reasoning
  - DL reasoners over OWL translation
- Ontology + data reasoning
  - Datalog subset