What's new with ontologies?

IVOA Interop meeting
Victoria, 2006/05/18
UCD/Semantics WG

Ontology (1)

- Formal representation of a domain
  - concepts
  - properties
  - Instances
- Based on Description Logics (math.)
- Common vocabulary for information sharing (between humans, and also machines)
- Attach labels to concepts (link with natural language or other knowledge bases)

Ontology (2)

- Primary concepts vs defined concepts
  - Better to have defined concepts: set of necessary and sufficient conditions – requires lot of work, but allows advanced reasoning

- Simple example for object types:
  - AstrObject
    - CompositeObject
      - MultipleStar
        - BinaryStar
  - Astrobject
    - CompositeObject (hasComponent)
      - MultipleStar (CompositeObject && hasComponent>=2 Star)
        - BinaryStar (CompositeObject && hasComponent=2 Star)

Ontology (3)

- ## Open world assumption
    - Unless you explicitly specify the disjunction of two concepts, the reasoner won't exclude the possibility that they can be equivalent
    - Example:
        - Star (disjoint: Galaxy)
            - VariableStar
            - AGB*
        - Galaxy (disjoint: Star)
            - RadioGalaxy
            - Seyfert
    - It is good practice to use disjoints between siblings when relevant

Ontology (4)

- ## Caution with properties!
  - A star has an effective temperature of 6000K
  - What property do we have for the concept Star ?
  - The property is not the value of the temperature, but rather hasEffectiveTemperature

- ## Tools and standards:
  - Protégé for edition
  - OWL for storage/exchange
  - Racer, Fact++, … for reasoning

Concepts and vocabulary

- Concepts can have multiple subsumption relations (not a simple 'tree')

- The naming of the concepts does not carry meaning (only convenient for edition)

- A plain text description can be associated to a concept: rdfs:comment

- Additional properties can be added, to do a mapping between a concept and other knowledge bases
  - e.g. DoubleStar (concept in the ontology)
    - simbad:hasSimbadNumericCode = "12.13.00.0"
    - ivoa:hasVOConcept = "stars.multiple.binary"

Ontologies in astronomy

- Don't try to directly model Astronomy at large !

- Focus on well-defined problems

  - Ontology for UCDs – richer description of relationships between UCD1+ words (collaboration between CDS and LORIA, MDA project)

  - Ontology of SIMBAD astronomical object types (INAF+CDS in VOTech DS5)

  - IAU thesaurus -> ontology (collab. CDS and IRIT)

Ontologies in astronomy

- The construction of the ontology must be driven by use cases

- Examples for the ontology of astronomical object types:
  - Start from the SIMBAD list of object types
  - Refactor, add intermediate concepts
  - Try to have defined concepts

SIMBAD object types

- Hierarchical classification of astronomical object classes
- Few entries (<150) : well focused
- Related to nomenclature

```
12.13.12.3:      HMXB       HXB      High Mass X-ray Binary
13.00.00.0:  Neb            Neb      Nebula of unknown nature
13.01.00.0:    PartofCloud  PoC      Part of Cloud
13.02.00.0:    PN?          PN?      Possible Planetary Nebula
13.03.00.0:    ComGlob      CGb      Cometary Globule
13.06.00.0:    EmObj        EmO      Emission Object
13.07.00.0:    HH           HH       Herbig-Haro Object
13.08.00.0:    Cloud        Cld      Cloud of unknown nature
13.08.03.0:      GalNeb     GNe      Galactic Nebula
13.08.04.0:      BrNeb      BNe      Bright Nebula
13.08.06.0:      DkNeb      DNe      Dark Nebula
13.08.07.0:      RfNeb      RNe      Reflection Nebula
13.08.11.0:      HI         HI       HI (neutral) region
13.08.12.0:      MolCld     MoC      Molecular Cloud
13.08.13.0:      HVCld      HVC      High-velocity Cloud
13.09.00.0:    HII          HII      HII (ionized) region
13.10.00.0:    PN           PN       Planetary Nebula
13.11.00.0:    HIshell      sh       HI shell
13.12.00.0:    SNR?         SR?      SuperNova Remnant Candidate
13.13.00.0:    SNR          SNR      SuperNova Remnant
14.00.00.0:  Star           *        Star
14.01.00.0:    *inCl        *iC      Star in Cluster
14.02.00.0:    *inNeb       *iN      Star in Nebula
14.03.00.0:    *inAssoc     *iA      Star in Association
14.04.00.0:    *in**        *i*      Star in double system
14.05.00.0:    V*?          V*?      Star suspected of Variability
14.06.00.0:    Pec*         Pe*      Peculiar Star
14.06.01.0:      HB*        HB*      Horizontal Branch Star
14.06.02.0:      YSO        Y*O      Young Stellar Object
14.06.05.0:      Em*        Em*      Emission-line Star
14.06.05.3:        Be*      Be*      Be Star
14.06.12.0:      AGB*       AG*      Asymptotic Giant Branch Star
14.06.15.0:      post-AGB*  pA*      Post-AGB Star
14.06.16.0:      WD*        WD*      White Dwarf
```

Use cases

- ## Help queries in the registry
  - Queries on <subject> relative to astronomical object types (label registry entries with ontology concepts)
  - Broaden or refine queries by finding subsuming or subsumed concepts

- ## SIMBAD update
  - Check concepts consistency during updates
  - e.g. add a stellar component to an instance of DoubleStar yields inconsistency -> MultipleStar

- ## Allow new ways to query SIMBAD
  - List all globular clusters of M31
  - Find binary stars with extrasolar planets