

# WordNet and SUMO

## An Introduction

Prashanth Kamle

Department of Computer Science and Engineering  
Indian Institute of Technology, Bombay

May 20, 2009

# Language Definitions

- Vocabulary - a set of pairs  $(f, s)$
- Form  $f$  - a string over a finite alphabet
- Sense  $s$  - an element from a given set of meanings
- Word - a form with a sense
  - ▶ polysemous - a word that has more than one sense
  - ▶ synonymous - two words that share at least one sense in common
- Dictionary - an alphabetical list of words

## Language Definitions (contd)

- Syntactic categories - Verb, Noun etc
- Semantic contexts - The set of contexts in which a particular  $f$  can be used to express a particular  $s$
- Morphology - is defined in terms of a set of relations between word forms. Eg: inflectional, derivational, and compound
- Lexical semantics - is defined in terms of a set of relations between word senses
- Definition - The semantic relations into which a word enters determine the definition of that word

# WordNet

- WordNet is an online lexical database designed for use under program control
- In WordNet, a form is represented by a string of ASCII characters, and a sense is represented by the set of synonyms that have that sense
- WordNet respects the syntactic categories noun, verb, adjective, adverb - the open-class words

# Semantic relations in Wordnet

- **Synonymy** is WordNet's basic relation, because WordNet uses sets of synonyms (synsets) to represent word senses
- **Antonymy** (opposing-name) is also a symmetric semantic relation between word forms
- **Hyponymy** (sub-name) and its inverse, **hypernymy** (super-name), are transitive relations between synsets. These organize the meanings of nouns into a hierarchical structure
- **Meronymy** (part-name) and its inverse, **holonymy** (whole-name) distinguish component parts
- **Troponymy** (manner-name) is for verbs what hyponymy is for nouns

# Morphology in Wordnet

- Inflectional morphology for each syntactic category is accommodated by the interface to the WordNet database
  - ▶ If information is requested for “went”, the system will return what it knows about the verb “go”
- Derivational and compound morphology are entered into the database without explicit recognition of morphological relations
  - ▶ “interpret”, “interpreter”, “misinterpret”, “interpretation”, “reinterpretation”, “interpretive”, “interpretative” and “interpretive dancing” are all distinct words

# WordNet

Word to search for:

Display Options:

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations

## Noun

- **S: (n) table, tabular array** (a set of data arranged in rows and columns) *"see table 1"*
  - [direct hyponym / full hyponym](#)
  - [member meronym](#)
    - **S: (n) row** (a linear array of numbers, letters, or symbols side by side)
    - **S: (n) column** (a vertical array of numbers or other information) *"he added a column of numbers"*
  - [direct hypernym / inherited hypernym / sister term](#)
    - **S: (n) array** (an orderly arrangement) *"an array of troops in battle order"*
      - **S: (n) arrangement** (an orderly grouping (of things or persons) considered as a unit; the result of arranging) *"a flower arrangement"*
        - **S: (n) group, grouping** (any number of entities (members) considered as a unit)
        - **S: (n) abstraction, abstract entity** (a general concept formed by extracting common features from specific examples)
        - **S: (n) entity** (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))
  - [derivationally related form](#)
    - **W: (v) tabulate** [Related to: [table](#)] (arrange or enter in tabular form)
    - **W: (adj) tabular** [Related to: [table](#)] (of or pertaining to or arranged in table form)
- **S: (n) table** (a piece of furniture having a smooth flat top that is usually supported by one or more vertical legs) *"it was a sturdy table"*
- **S: (n) table** (a piece of furniture with tableware for a meal laid out on it) *"I reserved a table at my favorite restaurant"*
- **S: (n) mesa, table** (flat tableland with steep edges) *"the tribe was relatively safe on the mesa but they had to descend into the valley for water"*
- **S: (n) table** (a company of people assembled at a table for a meal or game) *"he entertained the whole table with his witty remarks"*
- **S: (n) board, table** (food or meals in general) *"she sets a fine table"; "room and board"*

## Verb

- **S: (v) postpone, prorogue, hold over, put over, table, shelve, set back, defer, remit, put off** (hold back to a later time) *"let's postpone the exam"*
- **S: (v) table, tabularize, tabularise, tabulate** (arrange or enter in tabular form)

[WordNet home page](#)

# Ontology

- An Ontology is a formal representation of a set of concepts within a domain and the relationships between those concepts
- Ontologies are used to reason about the properties of a domain and may be used to define the domain
- **Upper ontologies** are domain-independent ontologies, intended to be reused and extended for particular domain to form a **domain ontology**



## Suggested Upper Merged Ontology - SUMO

- SUMO and its domain ontologies form the largest formal public ontology in existence today
- SUMO was created by merging publicly available ontological content into a single, comprehensive, and cohesive structure
- Collection of well-defined and well-documented concepts, interconnected into a semantic network and accompanied by a number of axioms
- SUMO is intended as a domain-independent substrate for designing domain ontologies

# Axioms

- Reflect common-sense notions that are generally recognized among the concepts
- Help to constraint interpretation of concepts
- Provide guidelines for automated reasoning systems
- Represented in simplified KIF in SUMO

# An axiom

```
(=>
  (instance ?DRIVE Driving)
  (exists (?VEHICLE)
    (and
      (instance ?VEHICLE Vehicle)
      (patient ?DRIVE ?VEHICLE))))
```

“If there is an instance of Driving, there is a Vehicle that participates in that action”

# Concepts

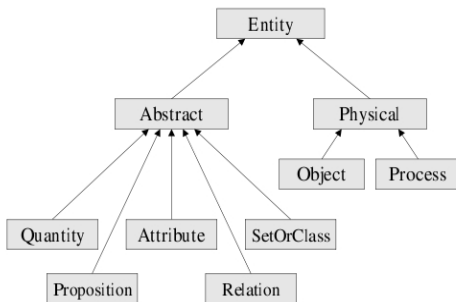


Figure: Top level concepts in SUMO



# Browser - contd.

## Related WordNet synsets

**animal, animate being, beast, brute, creature, fauna**

*a living organism characterized by voluntary movement*



**Animalia, kingdom Animalia, animal kingdom**

*taxonomic kingdom comprising all living or extinct animals*



**animal order**

*the order of animals*



**animal**

*of the nature of or characteristic of or derived from an animal or animals; "the animal kingdom"; "animal instincts"; "animal fats"*



**zoological**

*of or relating to animals or animal groups; "zoological garden"*



**animate**

*relating to animal life as distinct from plant life; "animate life"*



See more related synsets on a separate page.

# Browser - contd.

## Axioms (9)

If **lang** is an instance of **animal language** and **proc** is an agent of **agent** and **lang** is an instrument for **proc**, then **agent** is an instance of **animal** and **agent** is not an instance of **human**.

```
(=>
  (and
    (instance ?LANG AnimalLanguage)
    (agent ?PROC ?AGENT)
    (instrument ?PROC ?LANG))
  (and
    (instance ?AGENT Animal)
    (not
      (instance ?AGENT Human))))
```

If **process** is an instance of **psychological process**, then there exists animal **animal** so that **animal** experiences **process**.

```
(=>
  (instance ?PROCESS PsychologicalProcess)
  (exists
    (?ANIMAL)
    (and
      (instance ?ANIMAL Animal)
      (experiencer ?PROCESS ?ANIMAL))))
```

If **act** is an instance of **surgery** and **animal** is a patient of **act**, then there exists cutting **subact** so that **animal** is an instance of **animal** and **cutting** is a patient of **animal** and **subact** is a subpr

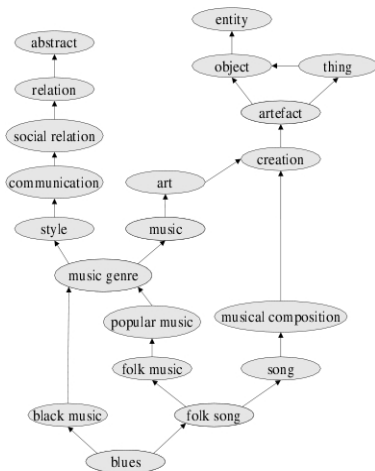
```
(=>
  (and
    (instance ?ACT Surgery)
    (patient ?ACT ?ANIMAL))
  (exists
    (?SUBACT)
    (and
      (instance ?SUBACT Cutting)
      (instance ?ANIMAL Animal)
      (patient ?ANIMAL ?CUTTING)
      (subProcess ?SUBACT ?ACT))))
```

# APIs

- Browser is distributed along with the source code
- Libraries can access SUMO and WordNet files programmatically
- Inferencing, using Vampire theorem prover



# Wordnet and ambiguity



# SUMO vs Wordnet

- SUMO and WordNet define conceptualizations (simplifications) of our world
- WordNet with the chief purpose to map these conceptualizations into natural language terms
- SUMO with the purpose to organize them into a logical structure

# Mapping

- Mapping enriches WordNet database files by tagging each synset with the corresponding SUMO concept
- The WordNet synset may be declared as equivalent to the SUMO term, as subsumed by it, or as an instance of it

Note: WordNet hypernymy/hyponymy relations cover subclass-superclass relations as well as class-instance relations. In SUMO, these relations are strictly distinguished

# Conclusion

- SUMO is intended as a domain-independent substrate for designing domain ontologies
- The WordNet lexicon provides a link between formal content expressed in SUMO and natural language
- SUMO Browser facilitates browsing the content of SUMO, WordNet, and eventual user-defined domain ontologies in a user-friendly way

## References

- 1 Miller, G.A. (1995) *Wordnet: A Lexical Database for English*. Communications of the ACM archive Volume 38 , Issue 11 (November 1995) Pages: 39 - 41
- 2 Sevchenko, M. (2003). *Online Presentation of an Upper Ontology*. In Proceedings of Znalosti 2003, Ostrava, Czech Republic, February 19-21, 2003
- 3 Niles, I., and Pease, A. (2001). *Towards a Standard Upper Ontology*. In Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001), Chris Welty and Barry Smith, eds, Ogunquit, Maine, October 17-19, 2001
- 4 Pease, A. *An introductory tutorial on KIF and SUMO*. Presented at Ontolog, September 6, 2007.
- 5 SUMO Ontology portal, <http://www.ontologyportal.org/>