Identifying, Finding and Encoding Semantic Relations

Christiane Fellbaum
Princeton University

Questions

• What kind of semantic knowledge does the NLP community need?
• How to represent semantic knowledge?
• How to expand on our present knowledge?
Assumptions

Need a repository for word form-meaning pairs (lexicon) that serves as a standard for word sense representation, applications and evaluation

Assumptions

- Lexicon has a structure
- Entities, events, properties are labeled (more or less) systematically
- Structure and lexicalization patterns can be captured with semantic and lexical relations
- Relations reflect (dis)similarities among labeled concepts in a fairly systematic way
- Semantic similarity as reflected by relations is useful for WSD
WordNet--the Plus side

- Broad coverage
- Multilingual
- Freely accessible
- Continually enriched both by Princeton and the user community

Limitations of WordNet

- Too sparse: too few relations, too few links
- Few syntagmatic (cross-POS) links
- Links are not weighted
- Many arcs are not directed
  \((\text{dollar} \rightarrow \text{green}, \not\rightarrow \text{green} \rightarrow \text{dollar})\)
- Sense inventory is too fine-grained for current automatic WSD
- Polysemy would be less of a problem if WordNet’s internal connectivity was greater
Sources of WordNet-style relations

- Classical relations (Aristotle)
- Lexical-semantic analysis of entities, events (causation, entailment, troponymy,...)
- Finding examples via lexico-syntactic patterns (Cruse 1986; Hearst 1993)
- Lexico-syntactic patterns presuppose specific relations

WordNet connections based on human judgment

- (Robust) word association norms
- Human judgments of associations among WordNet concepts show many connections not currently encoded (WordNetPlus, Boyd-Graber et al. 2006)

- Can’t all be easily classified or labeled!
Moving away from preconceived relations

- Reconsider: structure of the lexicon
- Which concepts are distinguished and labeled with words?
- Discover systematic differences among concepts/words that can be encoded as relations

Focus: Rigidity

- Important meta-property for distinguishing concepts in ontology
- Rigidity distinguishes Types vs. Roles
  e.g., DOLCE ontology (Guarino and Welty 2002), Generative Lexicon (Pustejovsky 1995)
Rigidity

rigid entities: *dog, orchid, man, shirts,..*
vs.
non-rigid: *pet, houseplant, teacher, laundry,..*

Adjectives

stage-level vs. individual-level (Carlson 1977)

*tall, intelligent, female,..
vs.
married, tired, surprised, ...

time-dependent:
*John is no longer tired/married/*tall/*intelligent*
Rigid and non-rigid terms may be related via shared hypernym

```
  plant
   / \
  orchid violet *houseplant
```

Rigid and non-rigid terms are compatible and not mutually exclusive:

*This is an orchid and a houseplant*
(type, role co-hyponyms)

cf.

*This is an orchid and a violet*
(type co-hyponyms)
Non-rigid properties are defeasible:

*This orchid is not a houseplant* (type)

*This orchid is not a plant* (role)

Type and role nouns noun are semantically similar when sharing a superordinate. A given entity can be labeled with both kinds of nouns. Useful for co-reference resolution. Temporal relations.
Encoding

Role and type nouns can be systematically distinguished and encoded linked to shared superordinates with para(allel) relations (Cruse 1986):

plant
     /\   
    /   \  
orchid houseplant

Relations in the verb lexicon

Lexicalization patterns show systematic, productive encoding of hyponymy (troponymy), causation

Verb classes:
Manner verbs
Change-of-state verbs

Can be distinguished via syntactic criteria
Another relation

Analogous to Type-Role distinction
Distinguish hyponyms (troponyms) from
“purpose” verbs
examples:
*exercise, control, greet, help, punish*
don’t encode manner or change-of state
not productive (?)

“purpose” verbs

move

exercise

run

running is **necessarily** a kind of moving (hyponym)
running is **not necessarily** a kind of exercising
Co-Hyponyms
defeasible/non-defeasible

Run but not {exercise/*move}
Wave but not {greet/*gesture}
Scrub but not {clean/*rub} the table

Fair amount of verb hyponyms in WN are defeasible
But no systematic encoding, distinction
Relation is not always captured by co-hyponymy

• Find verbs expressing purpose
• encode them in WordNet with “parallel” relation, following Cruse’s suggestion for Types and Role
• Problem: such concepts are not systematically encoded
Finding examples via lexico-syntactic patterns

V-ing is (not) V-ing
to V is (not) to V
V-ing is a way of V-ing

Patterns are also valid for regular hyponyms
but few such pairs are extracted (for pragmatic reasons?)

Web examples

**spraying** the action with a little WD-40 is not **cleaning**

**shake hands**, using the right hand, and explain that his is a way of **greeting** one another

**tipping**, leaving a gratuity, is a way of **thanking** people for their service


