Words

CS395 GAI
Spring, 2005
Overview

• Some basic properties of words
• COMLEX lexicon
• WordNet
• Cyc NL semantics
Words as building blocks

• Phrases are built out of words
  – “Hungry?”
  – “Very hungry”
  – “Hungry enough to eat a horse”

• The basic word being modified is the head of the phrase

• Syntax is about what combinations of kinds of words make sense
Words as compound entities
Kinds of words

• Open-class words
  – The stuff out of which the world is made
  – Nouns, Verbs, adjectives, adverbs
  – Can be extended as the universe of discourse expands

• Closed-class words
  – The linguistic equivalent of logical connectives or programming language primitives
  – Determiners, prepositions, quantifiers, …
  – Relatively fixed, providing part of the structure of the language
COMLEX Lexicon

• 38,000 head words
  – 21,000 nouns, 8,000 adjectives, 6,000 verbs

• Features represent their properties
  – Nouns: 9 features, 9 possible complements
  – Verbs: 5 features, 92 complements
Examples

(verb :orth "build"
  :subc ((np) (np-for-np) (part-np :adval ("up")))
  :TAGS ((TAG :BYTE-NUMBER 6918276 :SOURCE "brown" :LABEL (NP))
           (TAG :BYTE-NUMBER 6914461 :SOURCE "brown" :LABEL (NP))
           (TAG :BYTE-NUMBER 6858039 :SOURCE "brown" :LABEL (NP))))
(noun :orth "assertion"
  :subc ((noun-that-s) (noun-be-that-s)))
(adverb :orth "exceedingly"
  :modif ((PRE-COMPARATIVE) (PRE-QUANT) (PRE-ADJ) (PRE-ADV))
  :features ((DEGREE-ADV)))
(adjective :orth "above-mentioned"
  :features ((apreq) (attributive)))
(verb :orth "abbreviate"
  :subc ((np-pp :pval ("to"))
         (np) (np-np-pred) (np-as-np))
  :features ((vveryving :pastpart t)))
(noun :orth "Prof." :features ((ntitle)))
COMLEX in the KB

• (definitionInDictionary COMLEX31Lexicon <word> <parser-lex-definitions>)
WordNet

• Lexical database for English
  – Informed by psycholinguistics
  – Created by team led by George Miller
  – Project started in 1985

• Covers open-class words
Synsets

• = synonym set
• Fundamental unit of WordNet
  – Each synset represents one lexical concept
• Concepts partially pinned down by examples
• Synsets linked by a variety of relationships
Example: Car

The noun car has 5 senses (first 3 from tagged texts)

1. (598) car, auto, automobile, machine, motorcar -- (4-wheeled motor vehicle; usually propelled by an internal combustion engine; "he needs a car to get to work")
2. (24) car, railcar, railway car, railroad car -- (a wheeled vehicle adapted to the rails of railroad; "three cars had jumped the rails")
3. (1) cable car, car -- (a conveyance for passengers or freight on a cable railway; "they took a cable car to the top of the mountain")
4. car, gondola -- (car suspended from an airship and carrying personnel and cargo and power plant)
5. car, elevator car -- (where passengers ride up and down; "the car was on the top floor")
## WordNet Relations

<table>
<thead>
<tr>
<th>Semantic relation</th>
<th>Description</th>
<th>Part of speech</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synonym</td>
<td>A concept that means exactly or nearly the same as another. WordNet considers immediate hypernyms to be synonyms.</td>
<td>N V Adj Adv</td>
<td>{ sofa, couch, lounge } are all synonyms of one another. { seat } is the immediate hypernym of the synset.</td>
</tr>
<tr>
<td>Antonym</td>
<td>A concept opposite in meaning to another.</td>
<td>N V Adj Adv</td>
<td>{ love } is the antonym of { hate, detest }.</td>
</tr>
<tr>
<td>Hypernym</td>
<td>A concept whose meaning denotes a superordinate.</td>
<td>N V</td>
<td>A { feline, felid } is a hypernym of { cat, true cat }.</td>
</tr>
<tr>
<td>Hyponym</td>
<td>A concept whose meaning denotes a subordinate.</td>
<td>N V</td>
<td>A { wildcat } is a hyponym of { cat, true cat }.</td>
</tr>
<tr>
<td>Substance meronym</td>
<td>A concept that is a substance of another concept.</td>
<td>N V</td>
<td>A { snowflake, flake } is substance of { snow }.</td>
</tr>
</tbody>
</table>
## WordNet Relations

<table>
<thead>
<tr>
<th>Semantic relation</th>
<th>Description</th>
<th>Part of speech</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part meronym</td>
<td>A concept that is a part of another concept.</td>
<td>NADA</td>
<td>A { crystal, watch crystal, watch glass } is a part of a { watch, ticker }.</td>
</tr>
<tr>
<td>Member meronym</td>
<td>A concept that is a member of another concept.</td>
<td>NADA</td>
<td>An { associate } is a member of an { association }.</td>
</tr>
<tr>
<td>Substance of holonym</td>
<td>A concept that has another concept as a substance.</td>
<td>NADA</td>
<td>A { tear, teardrop } has { water, H2O } as a substance.</td>
</tr>
<tr>
<td>Part of holonym</td>
<td>A concept that has another concept as a part.</td>
<td>NADA</td>
<td>A { school system } has a { school, schoolhouse } as a part.</td>
</tr>
<tr>
<td>Member of holonym</td>
<td>A concept that has another concept as a member.</td>
<td>NADA</td>
<td>{ organized crime, gangland, gangdom } has { gang, pack, ring, mob } as a member.</td>
</tr>
<tr>
<td>Attribute</td>
<td>An adjective that is the value of a noun.</td>
<td>NADA</td>
<td>{ fast (vs. slow) } is a value of { speed, swiftness, fastness }.</td>
</tr>
</tbody>
</table>
## WordNet Relations

<table>
<thead>
<tr>
<th>Semantic relation</th>
<th>Description</th>
<th>Part of speech</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cause to</td>
<td>A verb that is the cause of a result.</td>
<td>×</td>
<td>{give} is the cause of the result {have, have got, hold}</td>
</tr>
<tr>
<td>Entailment</td>
<td>A verb that involves unavoidably a result.</td>
<td>×</td>
<td>To {die, decease, perish, go, exit, pass away, expire} involves unavoidably to {leave, leave behind}.</td>
</tr>
<tr>
<td>Troponym</td>
<td>A verb that is a particular way to do another.</td>
<td>×</td>
<td>To {samba} is a particular way to {dance, trip the light fantastic}.</td>
</tr>
<tr>
<td>Pertainym</td>
<td>An adjective or adverb that relates to a noun.</td>
<td>×</td>
<td>{criminal} relates to {crime}.</td>
</tr>
<tr>
<td>Attribute</td>
<td>An adjective that is the value of a noun.</td>
<td>×</td>
<td>{fast (vs. slow)} is a value of {speed, swiftness, fastness}.</td>
</tr>
<tr>
<td>Value</td>
<td>A noun that has an adjective for a value.</td>
<td>×</td>
<td>{weight} has {light (vs. heavy)} as a value.</td>
</tr>
</tbody>
</table>
WordNet information in the KB

• (synonymousExternalConcept <concept> WordNet-1997Version <synset-id> )
Semantic Information in KB

• Instances of SemTransPred
  – (nounSemTrans <lex-entry> <sense-id> <frame> <semtrans-expressions>)
  – (verbSemTrans <lex-entry> <sense-id> <frame> <semtrans-expressions>)
Acknowledgements

• Some materials provided by
  – Prof. Dan Moldovan (U Texas Dallas)
  – Profs. Ray Larson & Marc Davis (UC Berkeley)