

# 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
  - <http://wordnet.princeton.edu/>
- 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

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

# WordNet Relations

Semantic relation	Description	Part of speech				Example
		N	V	Adj	Adv	
<b>Part meronym</b>	A concept that is a part of another concept.	×				A { <i>crystal, watch crystal, watch glass</i> } is a part of a { <i>watch, ticker</i> }.
<b>Member meronym</b>	A concept that is a member of another concept.	×				An { <i>associate</i> } is a member of an { <i>association</i> }.
<b>Substance of holonym</b>	A concept that has another concept as a substance.	×				A { <i>tear, teardrop</i> } has { <i>water, H2O</i> } as a substance.
<b>Part of holonym</b>	A concept that has another concept as a part.	×				A { <i>school system</i> } has a { <i>school, schoolhouse</i> } as a part.
<b>Member of holonym</b>	A concept that has another concept as a member.	×				{ <i>organized crime, gangland, gangdom</i> } has { <i>gang, pack, ring, mob</i> } as a member.
<b>Attribute</b>	An adjective that is the value of a noun.	×				{ <i>fast (vs. slow)</i> } is a value of { <i>speed, swiftness, fastness</i> }

# WordNet Relations

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

# 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)