

CS 181: NATURAL LANGUAGE PROCESSING

Lecture 18: Lexical Semantics

KIM BRUCE
POMONA COLLEGE
SPRING 2008

Disclaimer: Slide contents borrowed from many sources on web!

WHAT'S THE PROBLEM?

- Compositional Semantics
- Based on meanings of words
- Many words have several meanings:
 - bank, dish, bass, ...
- Disambiguation: Choose correct sense from those in a sense “inventory” for word.
- Discrimination: Divide different uses without regard to existing inventories.

FINDING MEANING

- Lexeme is pairing of word with its meaning
- Lexicon is list of lexemes.
- Lexemes represented by lemma
 - Usually form found in dictionary
 - Represent different wordforms (e.g. plural)
- Words have different meanings or senses
 - Often correspond to different items in definitions.

RELATIONS BETWEEN MEANINGS

- If meanings relatively unrelated, called *homonymy*.
 - Ex., bank of river, savings bank
- If meanings are related, called *polysemy*.
 - Ex. bank as institution, bank building
 - Ex. blood bank and savings bank

RELATIONS BETWEEN MEANINGS

- Use of one aspect of concept to refer to other aspects or entire concept *metonymy*.
 - I really love Jane Austen -- *books*
 - The chicken was overcooked.
 - Plums have beautiful flowers.
- Humans generally cope well, but very hard for computers!

MORE FUN HEADLINES

- DRUNK GETS NINE YEARS IN VIOLIN CASE
- FARMER BILL DIES IN HOUSE
- PROSTITUTES APPEAL TO POPE
- STOLEN PAINTING FOUND BY TREE
- RED TAPE HOLDS UP NEW BRIDGE
- DEER KILL 300,000
- RESIDENTS CAN DROP OFF TREES
- INCLUDE CHILDREN WHEN BAKING COOKIES
- MINERS REFUSE TO WORK AFTER DEATH

BANKING ON TROUBLE

- The fisherman jumped off the bank and into the water.
- The bank down the street was robbed!
- Back in the day, we had an entire bank of computers devoted to this problem.
- The bank in that road is entirely too steep and is really dangerous.
- The plane took a bank to the left, and then headed off towards the mountains.

WHO CARES?

- Machine translation
 - Is a “bill” a bird jaw or invoice
- Information retrieval
 - Find all web pages about “cricket”.
 - Sport or insect
- Question answering
 - When is the bill due?
- Knowledge Acquisition
 - US Air flies to Albany airport (which one)

REPRESENTATIONS

- Dictionary (*entries for chair*)
 - a separate seat for one person, typically with a back and four legs.
 - the person in charge of a meeting or organization (used as a neutral alternative to chairman or chairwoman) : *the deputy chair of the Supreme Soviet.*
 - a professorship : *he held a chair in physics.*
 - a particular seat in an orchestra : [as adj., in combination] *she was fourth-chair trumpet.*
 - (the chair) short for electric chair .

REPRESENTATIONS

- Translation into second language
 - chaise
 - directeur
- Context
 - Sit on a chair
 - Take a seat on this chair
 - The chair of the CS department
 - The chair of the committee

APPROACHES TO DISAMBIGUATION

- Knowledge-based using dictionary, thesaurus
- Supervised based on labeled training set (including word senses)
- Unsupervised (w/features, but not word senses)

ALL WORDS DISAMBIGUATION

- Sometimes easier to classify discourse as a whole and use to limit senses:
 - He *put* his *suit* over the *back* of the *chair*.
 - The court ruled that the *suit* should be *put back* in the hands of the *chair* of the mediating board.
- Semantic field is collection of related words
 - semantic field of CLOTHING includes
 - shirt, pants, sewing machine, needle, etc.

INDIVIDUAL WORDS

- Use machine learning to classify
- Build feature vectors of relevant linguistic features (usually window of words around the occurrence)

RELATIONS BETWEEN SENSES

- *Synonymy* and *antonymy*
 - course/class, hot/cold
- Synonymous if freely substitutable
 - but *too* strict.
 - Even big/large fails (*big sister*)
- Antonyms if binary opposites on same scale
 - big/little, dark/light

HYPONYM/HYPERNYM

- Prefer subordinate/superordinate
 - v hyponym of w, if v more specific
 - w hypernym of v, if w more general
 - car < vehicle, square < rectangle
- Useful in classification, taxonomy
- Meronym/holonym: part-whole
 - wheel/car, foot/person

LEXICAL RELATIONS

Hypernym	more general	animal is a hypernym of dog
Hyponym	more specific	dog is a hyponym of animal
Meronym	part of	door is a meronym of house
Holonym	has part	house is a holonym of door
Synonym	similar meaning	car is a synonym of automobile
Antonym	opposite meaning	like is an antonym of dislike
Entailment	necessary action	step is an entailment of walk

WORDNET

- Originally English only
 - Database of word senses
- Create “synsets” for all languages of the world.
 - <http://www.globalwordnet.org/>
- Four databases in English
 - 117,097 nouns
 - 11,488 verbs
 - 22,141 adjectives
 - 4,601 adverbs

WORDNET.PRINCETON.EDU “CHAIR”

- Noun
 - S: (n) chair (a seat for one person, with a support for the back) "he put his coat over the back of the chair and sat down"
 - S: (n) professorship, chair (the position of professor) "he was awarded an endowed chair in economics"
 - S: (n) president, chairman, chairwoman, chair, chairperson (the officer who presides at the meetings of an organization) "address your remarks to the chairperson"
 - S: (n) electric chair, chair, death chair, hot seat (an instrument of execution by electrocution; resembles an ordinary seat for one person) "the murderer was sentenced to die in the chair"
 - S: (n) chair (a particular seat in an orchestra) "he is second chair violin"

WORDNET.PRINCETON.EDU “CHAIR”

⌘ Verb

- ⌘ S: (v) chair, chairman (act or preside as chair, as of an academic department in a university) "She chaired the department for many years"
- ⌘ S: (v) moderate, chair, lead (preside over) "John moderated the discussion"

WORDNET.PRINCETON.EDU “CHAIR”

⌘ Noun

- ⌘ S: (n) chair (a seat for one person, with a support for the back) "he put his coat over the back of the chair and sat down"
- ⌘ direct hyponym / full hyponym
 - ⌘ S: (n) armchair (chair with a support on each side for arms)
 - ⌘ S: (n) barber chair (a large fixed adjustable chair in which barbers seat their customers)
 - ⌘ S: (n) chair of state (a ceremonial chair for an exalted or powerful person)
 - ⌘ S: (n) chaise longue, chaise, daybed (a long chair; for reclining)
 - ⌘ S: (n) Eames chair (a chair designed by Charles Eames; originally made of molded plywood; seat and back shaped to fit the human body)
 - ⌘ S: (n) fighting chair (a fixed chair from which a saltwater angler can fight a hooked fish)
 - ⌘ S: (n) folding chair (a chair that can be folded flat for storage)

WORDNET.PRINCETON.EDU “CHAIR”

- ⌘ S: (n) chair (a seat for one person, with a support for the back) "he put his coat over the back of the chair and sat down"
- ⌘ direct hyponym / full hyponym
 - ⌘ S: (n) highchair, feeding chair (a chair for feeding a very young child; has four long legs and a footrest and a detachable tray)
 - ⌘ S: (n) ladder-back, ladder-back chair (a chair with a ladder-back)
 - ⌘ S: (n) lawn chair, garden chair (chair left outside for use on a lawn or in a garden)
 - ⌘ S: (n) rocking chair, rocker (a chair mounted on rockers)
 - ⌘ S: (n) straight chair, side chair (a straight-backed chair without arms)
 - ⌘ S: (n) swivel chair (a chair that swivels on its base)
 - ⌘ S: (n) tablet-armed chair (a chair with an arm that has been widened for writing)
 - ⌘ S: (n) wheelchair (a movable chair mounted on large wheels; for invalids or those who cannot walk; frequently propelled by the occupant)

WORDNET.PRINCETON.EDU “CHAIR”

- ⌘ S: (n) chair (a seat for one person, with a support for the back) "he put his coat over the back of the chair and sat down"
- ⌘ part meronym
 - ⌘ S: (n) back, backrest (a support that you can lean against while sitting) "the back of the dental chair was adjustable"
 - ⌘ S: (n) leg (one of the supports for a piece of furniture)
- ⌘ direct hypernym / inherited hypernym / sister term
 - ⌘ S: (n) seat (furniture that is designed for sitting on) "there were not enough seats for all the guests"

IN PYTHON

```
from nltk import wordnet
chair = wordnet.N['chair']
for synset in chair:
    print synset
    print 'gloss: ',synset.gloss
```

```
-----
{noun: chair}
gloss: a seat for one person, with a support for the back; "he put his coat over the back of the
chair and sat down"
{noun: professorship, chair}
gloss: the position of professor; "he was awarded an endowed chair in economics"
{noun: president, chairman, chairwoman, chair, chairperson}
gloss: the officer who presides at the meetings of an organization; "address your remarks to the
chairperson"
{noun: electric_chair, chair, death_chair, hot_seat}
gloss: an instrument of execution by electrocution; resembles an ordinary seat for one person;
"the murderer was sentenced to die in the chair"
{noun: chair}
gloss: a particular seat in an orchestra; "he is second chair violin"
```

WORD RELATIONS

```
>> for concept in chair[0][wordnet.HYPONYM][:10]:
...     print concept
```

```
{noun: armchair}
{noun: barber_chair}
{noun: chair_of_state}
{noun: chaise_longue, chaise, daybed}
{noun: Eames_chair}
{noun: fighting_chair}
{noun: folding_chair}
{noun: highchair, feeding_chair}
{noun: ladder-back, ladder-back_chair}
{noun: lawn_chair, garden_chair}
```

WORD RELATIONS

```
pprint.pprint(
    wordnet.N['chair'][0].tree(wordnet.HYPERNYM))
```

closure gives flat version

↓

```
[{noun: chair},
 [{noun: seat},
  [{noun: furniture, piece_of_furniture, article_of_furniture},
   [{noun: furnishing},
    [{noun: instrumentality, instrumentation},
     [{noun: artifact, artefact},
      [{noun: whole, unit},
       [{noun: object, physical_object},
        [{noun: physical_entity},
         [{noun: entity}]]]]]]]]]]]
```

PATH SIMILARITY

Measurement 0-1 of distance of shortest path in tree.
1 if same word, -1 if not connected

```
>>> wordnet.N['chair'][0].path_similarity(
    wordnet.N['highchair'][0])
0.5

>>> wordnet.N['chair'][0].path_similarity(
    wordnet.N['leg'][0])
0.0625
```

THEMATIC ROLES

✻ Similarity of roles in sentences

Thematic Role	Definition
AGENT	The volitional causer of an event
EXPERIENCER	The experiencer of an event
FORCE	The non-volitional causer of the event
THEME	The participant most directly affected by an event
RESULT	The end product of an event
CONTENT	The proposition or content of a propositional event
INSTRUMENT	An instrument used in an event
BENEFICIARY	The beneficiary of an event
SOURCE	The origin of the object of a transfer event
GOAL	The destination of an object of a transfer event

EXAMPLES

Thematic Role	Example
AGENT	<i>The waiter</i> spilled the soup.
EXPERIENCER	<i>John</i> has a headache.
FORCE	<i>The wind</i> blows debris from the mall into our yards.
THEME	Only after Benjamin Franklin broke <i>the ice</i> ...
RESULT	The French government has built a <i>regulation-size baseball diamond</i> ...
CONTENT	Mona asked " <i>You met Mary Ann at a supermarket</i> "?
INSTRUMENT	He turned to poaching catfish, stunning them <i>with a shocking device</i> ...
BENEFICIARY	Whenever Ann Callahan makes hotel reservations <i>for her boss</i> ...
SOURCE	I flew <i>in from Boston</i> .
GOAL	I drove <i>to Portland</i> .

ROLES

- ❁ If can identify roles then can answer questions.
- ❁ Verbs allow thematic roles in different posns
 - ❁ John threw the ball to Mary
 - ❁ John threw Mary the ball
 - ❁ The ball was thrown to Mary by John.
- ❁ Certain verbs (e.g. throwing, future having) allow such alternations

PROBLEMS

- Not easy to label.
- E.g., instruments:
 - The cook opened the jar with the new gadget.
 - The new gadget opened the jar.
 - Lynn ate her dinner with a fork.
 - *The fork ate the dinner
- Generalized semantic roles:
 - Proto-agent, proto-patient (fuzzier)

LEXICAL RESOURCES

- Proposition Bank (PropBank)
 - Sentences annotated w/ semantic roles.
 - All (English/Chinese) sentences of Penn Treebank
- FrameNet
 - frame specific roles -- common across (some) verbs.

PROPBANK

- How does a verb relate to its arguments?
- 1 million words of WSJ annotated
- Each verb given specific set of roles
 - Arg0 is proto-agent
 - Arg1 is proto-patient
 - Rest specific to verb sense.

EXAMPLE

- cover (smear, put over)
 - Arguments
 - arg0: causer of covering
 - arg1: thing covered
 - arg2: covered with
 - Example:
 - John covered the bread with peanut butter

FRAMENET

- Frames are situation types
- Change-position-on-a-scale frame:
 - This frame consists of words that indicate the change of an Item's position on a scale (the Attribute) from a starting point (Initial value) to an end point (Final value).
- Look up "increase" at <http://framenet.icsi.berkeley.edu>

CORE ROLES

- ATTRIBUTE:
 - The ATTRIBUTE is a scalar property that the ITEM possesses.
- DIFFERENCE
 - The distance by which an ITEM changes its position on the scale.
- FINAL STATE
 - A description that presents the ITEM's state after the change in the ATTRIBUTE's value as an independent predication.
- FINAL VALUE
 - The position on the scale where the Item ends up.

CORE ROLES

- INITIAL STATE
 - A description that presents the ITEM's state before the change in the ATTRIBUTE's value as an independent predication.
- INITIAL VALUE
 - The initial position on the scale from which the ITEM moves away.
- ITEM
 - The entity that has a position on the scale.
- VALUE RANGE
 - A portion of the scale, typically identified by its end points, along which the values of the ATTRIBUTE fluctuate.

CORE ROLES

- Frame works with “rise”, “increase”, “fell”, “dropped”, etc.

SELECTIONAL RESTRICTIONS

- Thematic roles place restrictions on arguments
 - John drank two glasses of juice.
 - $\exists e, x, y. \text{Drinking}(e) \wedge \text{Agent}(e, x) \wedge \text{Theme}(e, y) \wedge \text{Isa}(y, \text{DrinkableThing})$
- Can use WordNet w/hypernym to determine juice is drinkable.
- Restrict roles to synsets.

ANY QUESTIONS?