

Lecture notes by Edward Loper

Course: CIS 630 (Lexical Semantics)

Professor: Martha Plamer

Institution: University of Pennsylvania

1 Karin Kipper: Word Sense Disambiguation

Comparison of 3 different approaches to word sense disambiguation..

1.1 Definition of the task

map from a word to a set of candidates. The set of candidates are a finite pre-enumerated set. dependant on context – what are we trying to accomplish? Fine or coarse grained?

1.2 Approaches

- acquisition from explicit knowledge source (wordnet etc)
- acquisition from training
- automatic acquisition

Different possible knowledge sources:

- POS
- local collocations
- global context
- syntactic relations
- etc..

1.3 Approach 1: Exemplar based

Learns from examples. For each example found, store features of that example:

- POS tag or surrounding 3 words
- morph. root
- keywords in sentence
- local collocations in specific positions
- which verb is used

Find features that are statistically significant (occur with the word sense more than chance)..

Define distance between 2 examples as the sum of diffs btwn features

Got accuracy of 89% on disambiguation of the noun "interest"

Try with 191 words: 121 nouns, 70 verbs... Got 54% and 67% in brown and WSJ corpora.

Advantages of approach/testing

- uses several knowledge sources
- very fine-grained
- evaluated on many words

Disadvantages

- sparse data
- supervised

- hard to switch domains – bigger corpus gives more examples
- difficult to tune constants
- features not weighted in any way

1.4 Approach 2: Dependency Grammar

Look at the local context of word in a parse tree.

Build local contexts.. eg:

```
|| "the facility employed a good person"
|| - facility(subj employ head)
|| - person(adjn good mod)
|| - person(obj employ head)
```

Get local contexts for a word, find all other words in the same context, compare each other word's sense to possible sense of word, pick most likely one..

Evaluated against SEMCOR. Only disambiguated nouns. Grouped different WordNet senses. 3 measures of correctness: strict correctness, similarity ≥ 0.27 ..

Advantages

- uses parser
- uses raw text
- easier to switch domains (training set not annotated)
- fewer sparse data problems..

Disadvantages

- Ignores local context, type of WordNet relations (sister/daughter)
- weak verb contexts
- 'one sense per discourse' for related senses
- grouped WordNet senses
- evaluate with 'similar enough' measures..

1.5 Approach 3: Decision Lists

if-then analysis used for classification. Use local collocations.

Start with raw text, put in some collocations.. Find new collocations, iterate..??

Disambiguated 12 nouns with 2 sentences on 460 million words..

gets 96.5% accuracy

Advantages

- uses untagged corpus
- easy to port to new domains
- simple
- only 1 knowledge source

Disadvantages

- need very large corpus
- very coarse-grained sense distinctions
- homonyms
- 'one sense per discourse' might not always be true..

2 Representing Verbs: Semantic Templates

2.1 Syntactic Structure

|| x, <y PUT_{loc} z>

PUT stands for a class of verbs..

2.2 Semantic Templates (Lexical-Conceptual Structures)

What should they look like? How are they built up?

∃ mapping between syntactic structure and semantic templates.

Decomposition of predicates..

We are interested in verb classes: classes that share certain semantic components.

candidate for LCS of "the ice melted":

```
|| [y BECOME <melted>]
||
||     y: argument
||     BECOME: semantic primitive
||     <melted>: constant (what is particular to the verb)
```

This structure is also used for "dried," "exploded," etc. (unaccusatives, change of state verbs)

Pinker '89: children learning verbs. postulated ≈ 12 LCS's

Denominal verbs of putting

- I butter the verb, I paint the wall: you place butter/paint on some thing.

```
|| LCS: [x CAUSE [<BUTTER> BECOME Ploc Z]]
||
||     x: agent
||     <BUTTER>: y argument, constant?
||     CAUSE: primitive
||     BECOME: primitive
```

- I pocket the money, I shelf the book: you put something in your pocket.

```
|| LCS: [x CAUSE [y BECOME P-<POCKET>]]
||
||     x: agent
||     y: argument
||     <POCKET>: location
||     CAUSE: primitive
||     BECOME: primitive
```

Aspect

Vendler: There are 4 basic types of aspect:

- state: be tall, live, etc. (*imperative)
 - stage (be quiet)

- individual (be tall, be smart)
- activity: write, sing, dance, etc. (can say "... for an hour")
- accomplishment: build, destroy, etc. (can say "... in an hour")
- achievements: recognize (can say "... at noon")

Most verbs are associated with a type, but can be changed..

Consider "eat": as intransitive, it's an activity. "eat an apple" is an accomplishment. "eat apples" is an activity.

James Pustiovsky: in Levin and Pinker book..

Emmon Bach: imperfect paradox, distinguishes activities from accomplishments and achievements. "John is running" entails "John has run." But "John is building a house" does not entail "John has built a house."—tells you whether the event has an endpoint..

David Dowty: word meaning and Montague grammar '79.. "almost test" that distinguishes activities from accomplishments. "John almost swam" implies he didn't swim at all. "John almost built a house" implies he MAY have started building a house.

Levin & Malka Rappoport Hovav: Extended verb senses: building up templates

She sweeps:

|| [x ACT <sweep>]

she swept the floor:

|| [x ACT on y BY <sweep>]

She swept the floor clean:

|| [x CAUSE [y BECOME <clean>] [BY <sweep>]]

?

Um.. See the Levin & Rapoport paper..??

Combine elementary templates into more complex ones. How is this constrained?

<http://www.neci.jn.nec.com/homepages/sandway/pappi/projects.html>

Another proposal:

Have primary and secondary templates, and restrictions on how they can be combined.

2.3 Data: English resultatives

- AP resultatives: expressing final state
 - John wiped the table clean
 - John wiped the table clean/free of crumbs
 - PP resultatives
 - The thief tore the painting [into shreds]

|| * the thief tore the painting ruined

|| * john wiped the table into a shiny surface

Dejectivals and construction/destructions don't subcat for either.

|| chris cleaned the table

|| * chris cleaned the table sparkling

|| * chris cleaned the table into a brilliant surface

|| also, wrote

A small class (hammered) can take both.

Try to build semantic templates.. Try to cover aspect, sub-cat, verb-specific semantics

primary/secondary templates check-off system licenses secondary predicates

secondary templates:

```
|| y be<state>  
|| y become y'
```

Variables in secondary template are only licensed if they also occur in the primary template..

check web site for more on this.. sounds interesting.

3 Extracting a Computational Lexicon

Use on-line dictionaries for various tasks like pp attachment..

'90s: dictionaries start changing to SGML etc. where relational knowledge is encoded..

MindNet..

3.1 Source Materials

Dictionaries (many at LDC).. LDOCE (Longman's dictionary of contemporary english) defines words using ~2500 base words. Grammar codes, countability, etc.

MindNet uses: american heritage 3rd edition, LDOCE, Microsoft Encarta..

Difficult to separate dictionary & encyclopedic knowledge

3.2 Extraction Strategies

Extract POS, etc.. need robust extraction, even for explicitly marked-up entries..

Use definitions as bags of words to derive similarity measures between words..

Use regexps to derive metonymy etc... easier with parse tree..

4 Erwin Talk on Historical Syntax

Questions:

- Where do words come from?
- How does meaning change?
- How does syntax associated with words change?

4.1 Word Origins

- borrowed words
- 'new' words: blends, shortenings, affixation, etc.

Restrictions on new words

Some types more productive than others (N, V, adj...).. Must be able to tell category from context.

4.2 Meaning Change

- generalize from specific instances (Xerox, orange)
- restriction/narrowing: girl comes from "gurle"=young person. "skyline" originally meant horizon
- combination: broadcast meant scatter seed over a field, came to mean diffuse media through space..
- reinterpretations: woeful, awesome, hopeful..

4.3 Why do lexical changes take place?

Self-imposed factors

Remove words because of taboo (remove fik, fok, because the sound like...)

Shifts in word frequency

- zipf

new words start out in zipf's tail.. Semantic shift towards higher frequency..

semantic shift: change frequencies of meanings of words..

changes tend to occur with s-curves.

4.4 Function words

For the most part, new words have to be content words. Function word generation is harder.. How do we get new function words?

Gramaticalization: current function words are changed/extended into functional words (I will eat, I have eaten, etc.).

Copula from pronoun. "Bob he the farmer" ..

5 NN compounds

5.1 Headedness of compounds

In NN compounds (dog house, etc), one word is more important: dog house is a type of dog. In general, in english NN compounds, right element is the head. Head determines/carries gender, number, etc. Within indoeuropean languages, almost all languages have right-headed compounds.

3 Types of compounds:

- union of equals. e.g., "devamanussa": god and man.. when you pluralize, it becomes gods and men. Or washer-dryer.
- doghouse: specifies type of an object.
- redhead, pickpocket: head of the compound is really external/implied. (a pickpocket isn't a type of pocket)

Difference between compounds and collocations?

Some questions:

- How do you identify compounds?
- structure within compounds: [[N1 N2] N3] vs [N1 [N2 N3]]
- determining semantic relations..

6 Construction Grammar..

Semantic representation of 3-place predicates:

- ditransitives: S VP NP NP
- caused-motion: S VP NP PP
- resultatives: S V NP AP

How do we extend this across languages? We don't want to simply enumerate constructions..

Traditionally (GB), semantics is in lexicon, projected onto syntax.. verb subcats 3 arguments.

Consider sentences like: "John sneezed the napkin off the table" Do we want to say that sneeze takes 3 arguments??

Represent meaning as construction.. no strict separation between roles of lexicon and syntax: both are constructions, and both can carry meaning.

polysamy of constructions like ditransitive:

```
|| john gave mary the book
||   vs.
|| john promised mary the book
```

the additional meanings (allowed, etc) appear in multiple constructions:

```
|| John let mary out of the house
```

7 Erwin's talk: Frequency (& meaning) of antonymous adjectives

7.1 One View: Scales

Antonyms give 2 endpoints of a scale, with things between (dry, arid, soggy, wet). Endpoints are called absolute/direct antonyms. Indirect antonyms have non-endpoint members.

7.2 Another View

dry/wet are antonyms, and arid, parched, soggy, moist, etc. are in the synsets of wet and dry.

7.3 Hypothesis 1

antonyms co-occur in sentences more frequently than chance.
results are borne out in experiments..

7.4 Hypothesis II

antonyms have the same selectional restrictions (select the same head nouns).

7.5 Sparse Data

When this experiment was run before, it was run on 1mil words – we get sparse data problems. Try running it on something bigger.

8 Chris's talk

What is interpretation and how does it function? Discourse semantics & lexical semantics.

How is the information encoded & transmitted? Abstract away from the pragmatics..

Assertions: transmit some encoded piece of information..

2 types of info transmitted: info about the world; and interpretation.

Consistent use of new and old words – transmit info about how words are used.

Stanaker(sp?) equates modality with reference?

$$\begin{array}{|c|} \hline | i \quad j \quad k \\ \hline \phi | T \quad T \quad F \\ \hline \end{array}$$

Expand ϕ into its own i,j,k: the proposition can be different in different worlds (this is just incorporating context?)

Leads to 2 traditions:

- discourse representation theory: when you use a definite noun phrase, you create a new pointer.. you refer to the pointer with later references..

interpretation is not truth value, but effect on the world/participants in discussion..
types of inference..

- entailment
- presuppositional implication
- abduction (skip it for now)

assume entailment is fairly lexical..

john is a bachelor \rightarrow john is unmarried..

presuppositions: does "jon believes mary's children are bald" presuppose that mary has children?

holes, plugs, filters. holes let everything through, filters let some thing through, plugs let nothing through (in terms of presuppositions).