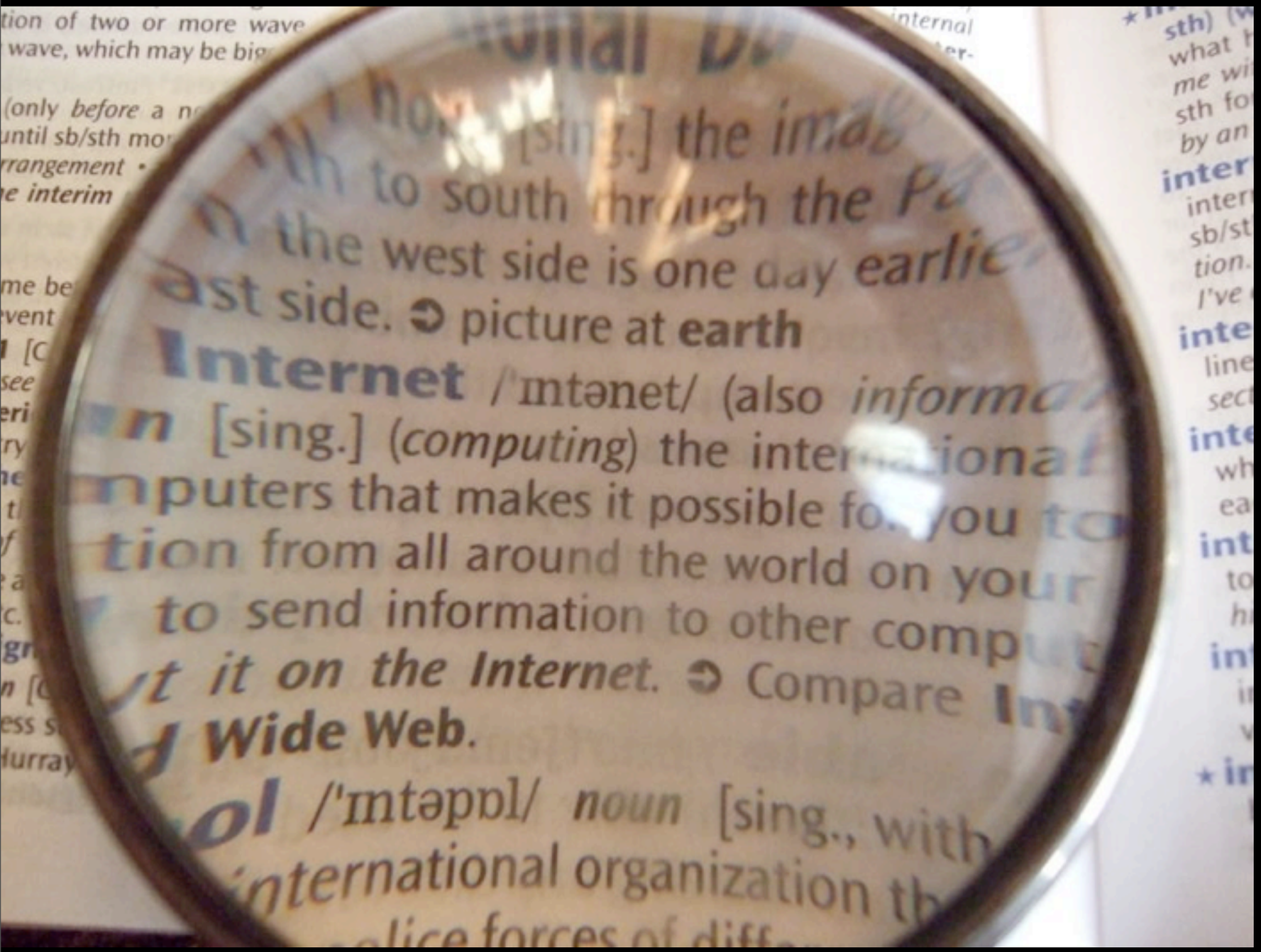
Lexical Semantics with WordNet + Python



NATHAN SCHNEIDER

PyLing Meeting
2013-10-21

Overview

- What is WordNet?
  - lexicon of English words, word meanings, and meaning relationships
  - from Princeton University
  - computational, publicly available
  - large: 117,000 synsets (entries)
- How is it organized?
- Discussion: applications & limitations

http://wordnet.princeton.edu/

Christiane Fellbaum, editor. 1998. *WordNet: an electronic lexical database.* MIT Press, Cambridge, MA.

shower

Websites for browsing WordNet

- http://www.visuwords.com/
- http://www.golovchenko.org/cgi-bin/wnsearch
- http://wordnetweb.princeton.edu/perl/webwn

WordNet: 4 parts of speech

noun (n) verb (v)

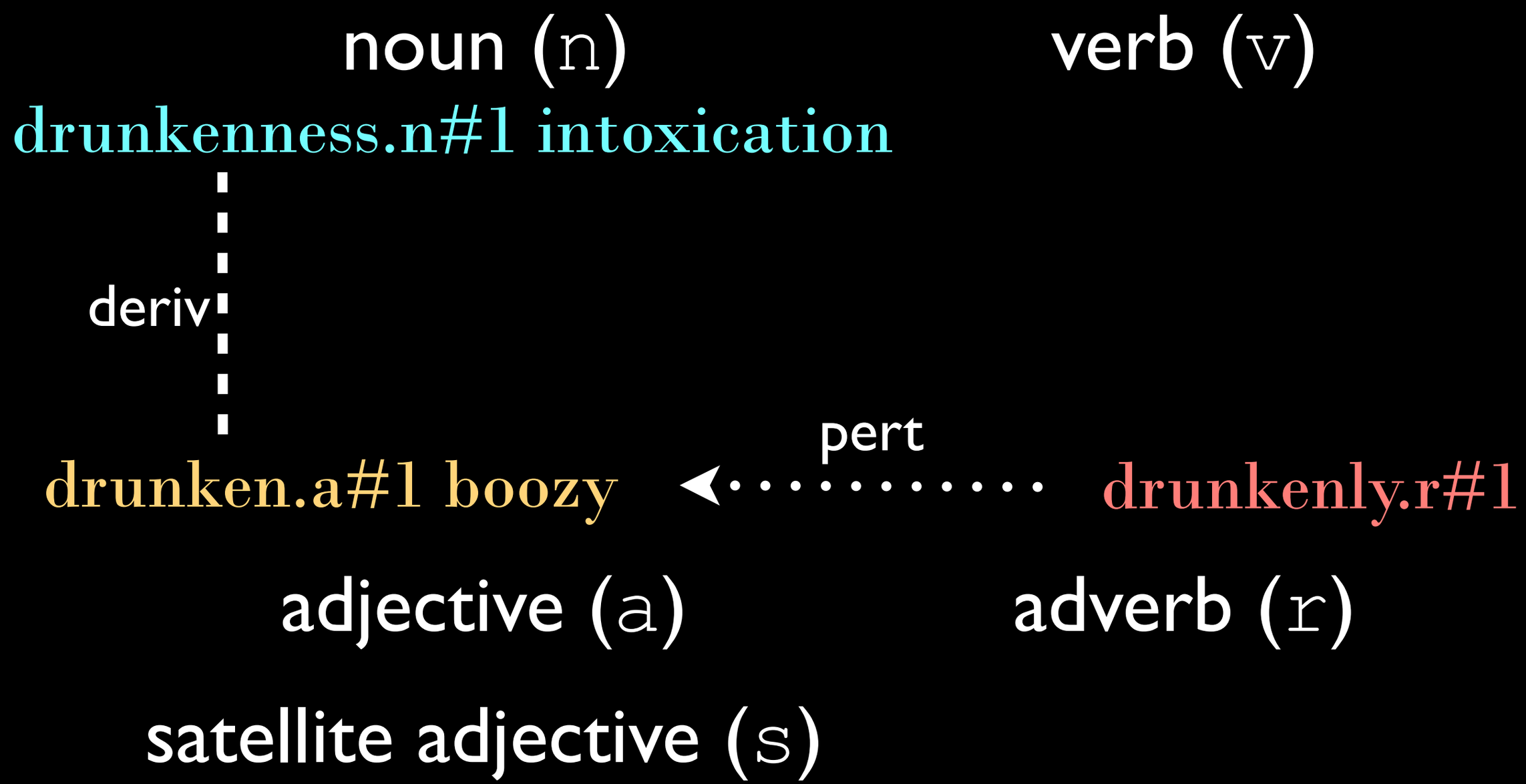
shower.n#1 plumbing fixture - - - - - deriv - - - - - shower.v#3 bathe

shower.n#5 exhibitor - - - - - - deriv - - - - - - show.v#1 demo, present

adjective (a) adverb (r)

satellite adjective (s)

WordNet: 4 parts of speech



Technical details: naming

- **term:** one or more word stems, separated by underscores
- **lexeme:** term + {n,v,a,r}
- **lemma:** sense-disambiguated lexeme
- **synset:** collection of synonymous lemmas. named by the **head lemma**
- **lemmas** can also be named relative to the synset name

NLTK

```
get_together

get_together.n#1
get_together.v#1
get_together.v#2
get_together.v#3
get_together.v#4
```

```
meeting.n.02.get_togeth
assemble.v.03.get_toge
meet.v.02.get_together
collaborate.v.01.get_to
join.v.01.get_together
```

WordNet in NLTK

- ```
  from nltk.corpus import wordnet as wn
  wn.synsets('get_together')
  wn.synset('meeting.n.02').definition
  wn.synset('meeting.n.02').lemmas
  wn.synset('meeting.n.02').hypernyms()
  wn.morphy('meetings', wn.VERB)
  wn.morphy('meetings', wn.NOUN)
  for l in wn.all_lemma_names(): …
  for s in wn.all_synsets(): …
  ```
- http://nltk.org/book/ch02.html#wordnet
- http://nltk.org/api/nltk.corpus.reader.html#module-nltk.corpus.reader.wordnet
- http://nltk.org/api/nltk.corpus.reader.html#module-nltk.corpus.reader.semcor

SemCor: a corpus annotated with WordNet senses
  ```

Discussion

- Applications
- Limitations

SEE SOMETHING SAY SOMETHING
IF SOMETHING DOES NOT LOOK RIGHT, LET US KNOW RIGHT AWAY
CALL PORT AUTHORITY POLICE AT 412.255.1385
PortAuthority.org