LING 581: Advanced Computational Linguistics

Lecture Notes
April 10th
Administrivia

• My sincere apologies for canceling last week's classes due to sickness
WordNet Homework Exercise

• Question 1:
  • Find the shortest distance links between *star* and *telescope*
  • (Make sure you have the right word sense)
  • How many are there?

Noun

1. S: (n) star#1 ((astronomy) a celestial body of hot gases that radiates energy derived from thermonuclear reactions in the interior)
2. S: (n) ace#3, adept#1, champion#4, sensation#2, maven#1, mavin#1, virtuoso#1, genius#3, hotshot#1, star#2, superstar#1, whiz#1, whizz#1, wizard#1, wiz#1 (someone who is dazzlingly skilled in any field)
3. S: (n) star#3 (any celestial body visible (as a point of light) from the Earth at night)
   • S: (n) star#4, principal#3, lead#8 (an actor who plays a principal role)
   • S: (n) star#5 (a plane figure with 5 or more points; often used as an emblem)
   • S: (n) headline#1, star#6 (a performer who receives prominent billing)
   • S: (n) asterisk#1, star#7 (a star-shaped character * used in printing)
   • S: (n) star topology#1, star#8 (the topology of a network whose components are connected to a hub)
WordNet Homework Exercise

• But only one sense for *telescope*:

<table>
<thead>
<tr>
<th>Noun</th>
</tr>
</thead>
</table>
| • *S: (n) telescope*
| • *scope* (a magnifier of images of distant objects) |
WordNet Homework Exercise

• Use bfs4.perl:

$ perl bfs4.perl star#n#1 telescope#n#1
Not found (distance 7 and 100000 nodes explored)
$ perl bfs4.perl star#n#1 telescope#n#1 1200000
Max set to: 1200000
Found at distance 8 (232993 nodes explored)
telescope#n#1 holo finder#n#3 hypo optical_device#n#1 hypo device#n#1 hypo instrumentality#n#3 hypo artifact#n#1 ants natural_object#n#1 hype celestial_body#n#1 hype star#n#1
Found at distance 8 (233025 nodes explored)
telescope#n#1 holo prism#n#2 hypo optical_device#n#1 hypo device#n#1 hypo instrumentality#n#3 hypo artifact#n#1 ants natural_object#n#1 hype celestial_body#n#1 hype star#n#1
Found at distance 8 (438947 nodes explored)
telescope#n#1 holo finder#n#3 hypo optical_device#n#1 hypo device#n#1 hypo instrumentality#n#3 hypo artifact#n#1 ants natural_object#n#1 hype constellation#n#2 holo star#n#1
Found at distance 8 (438979 nodes explored)
telescope#n#1 holo prism#n#2 hypo optical_device#n#1 hypo device#n#1 hypo instrumentality#n#3 hypo artifact#n#1 ants natural_object#n#1 hype constellation#n#2 holo star#n#1
All minimal solutions found
WordNet Homework Exercise

• Use bfs4.perl again:

$ perl bfs4.perl star#n#3 telescope#n#1
Not found (distance 8 and 100002 nodes explored)

$ perl bfs4.perl star#n#3 telescope#n#1 200000
Max set to: 200000

Found at distance 8 (166291 nodes explored)

  telescope#n#1 holo finder#n#3 hypo optical_device#n#1 hypo device#n#1 hypo instrumentality#n#3 hypo artifact#n#1
  ants natural_object#n#1 hype celestial_body#n#1 hype star#n#3

$ perl bfs4.perl star#n#3 telescope#n#1 300000
Max set to: 300000

Found at distance 8 (166291 nodes explored)

  telescope#n#1 holo prism#n#2 hypo optical_device#n#1 hypo device#n#1 hypo instrumentality#n#3 hypo artifact#n#1
  ants natural_object#n#1 hype celestial_body#n#1 hype star#n#3

All minimal solutions found
WordNet Homework Exercise

• Graph (*borrowed from Benjamin Mullins*):

```
  artifact#n#1 --> instrumentality#n#2 --> device#n#1
  |         |                        |
  | hyponym |                        |
  |         |                        |
  |         | natural_object#n#1    |

very abstract concepts

semantic opposition

"has part"
```
WordNet Homework Exercise

• Generative Lexicon (Pustejovsky, 1995)
• Qualia structure:
  • Form(al) role: hypernym
  • Constitutive role: meronym
  • Telic role: purpose
  • Agentive role: creator

What is a telescope and what is it used for?

A telescope is an instrument that is used to view distant objects. If you want to look at the planets, you can use a telescope. The higher the magnification on the telescope, the better your view will be. Galileo is often credited with the invention of the telescope, but this is incorrect.
WordNet Homework Exercise

• But it's not so easy to collect this data:

The main **purpose** of a **telescope** is to gather light, i.e. to collect and focus photons. We can think of a **telescope** then as a "light bucket" - the bigger the bucket, the more photons a **telescope** can collect.

**The Purpose of a Telescopes**
www.astro.cornell.edu/academics/courses/astro201/purpose.htm
WordNet Homework Exercise

• But it's not so easy to collect this data:

What is the main function of a telescope?

The main purpose of astronomical telescope is to make objects from outer space appear as bright, contrasty and large as possible. That defines its three main function: light gathering, resolution and magnification.

Main functions of a telescope - Amateur Telescope Optics
www.telescope-optics.net/functions.htm

People also ask

What is a telescope and what is it used for?

An optical telescope which uses lenses is known as a refracting telescope or a refractor; one which uses a mirror is known as a reflecting telescope or a reflector. Besides optical telescopes, astronomers also use telescopes that focus radio waves, X-rays, and other forms of electromagnetic radiation.

telescope facts, information, pictures | Encyclopedia.com articles about ...
WordNet Homework Exercise

What was the original purpose of the telescope?

It is quite a common belief that Italian astronomer Galileo Galilei invented the telescope but this is not strictly true. The earliest workings towards the design of the **refracting telescope** were made by German-Dutch lensmaker Hans Lippershey (above) in 1608.  
Feb 7, 2013

Did Galileo invent the telescope? | - Space Answers
https://www.spaceanswers.com/astronomy/did-galileo-invent-the-telescope/

Why is the telescope so important?

Astronomers use **telescopes** because they're much better than our eyes. Here are a few reasons: 1. **Telescopes** see lots of colors - **telescopes** can collect light that our eyes are unable to: radio, microwave, infrared, ultraviolet, x-rays and gamma rays.  
Jul 18, 2015

Why do astronomers use telescopes? (Beginner) - Curious About ...
curious.astro.cornell.edu/.../telescopes/.../749-why-do-astronomers-use-telescopes-begin...
Appendix: modifying the bfs code

- `querySense()`: map `word` into `word#pos#sense_num`

```perl
perl senses.perl star
star#n#1 star#n#2 star#n#3 star#n#4 star#n#5 star#n#6 star#n#7 star#n#8
star#v#1 star#v#2 star#v#3
star#a#1
```

```perl
use WordNet::QueryData;
my $wn = WordNet::QueryData->new( noLoad => 1);
my @r = $wn->validForms($ARGV[0]);
foreach $wpos (@r) {
  @r2 = $wn->querySense($wpos);
  print "@r2 \n";
}
```
Connection paths and Similarity

```python
>>> from nltk.corpus import wordnet as wn
>>> star1 = wn.synset('star.n.01')
>>> star3 = wn.synset('star.n.03')
>>> telescope1 = wn.synset('telescope.n.01')
>>> star1.path_similarity(telescope1)
0.09090909090909091
>>> star3.path_similarity(telescope1)
0.09090909090909091
>>> star3.path_similarity(star1)
0.3333333333333333
>>> star3.lch_similarity(star1)
2.538973871058276
>>> star3.lch_similarity(telescope1)
1.2396908869280152
>>> star1.lch_similarity(telescope1)
1.2396908869280152
>>> star1.wup_similarity(star3)
0.8571428571428571
>>> star1.wup_similarity(telescope1)
0.4444444444444444
>>> star3.wup_similarity(telescope1)
0.4444444444444444
and others...
```
Perl WordNet::Similarity

- `synset1.wup_similarity(synset2)`: Wu-Palmer Similarity: Return a score denoting how similar two word senses are, based on the depth of the two senses in the taxonomy and that of their Least Common Subsumer (most specific ancestor node). Note that at this time the scores given do not always agree with those given by Pedersen's Perl implementation of Wordnet Similarity.

- Online (link doesn't work anymore: used to exist...)
  - http://marimba.d.umn.edu/cgi-bin/similarity/similarity.cgi
Perl WordNet::Similarity

- Perl Module: [https://metacpan.org/pod/WordNet::Similarity](https://metacpan.org/pod/WordNet::Similarity)
sudo cpanm WordNet::Similarity::path
---> Working on WordNet::Similarity::path
Fetching http://www.cpan.org/authors/id/T/TP/TPEDERSE/WordNet-Similarity-2.07.tar.gz ... OK
Configuring WordNet-Similarity-2.05 ... OK
==> Found dependencies: Text::Similarity
---> Working on Text::Similarity
Fetching http://www.cpan.org/authors/id/T/TP/TPEDERSE/Text-Similarity-0.13.tar.gz ... OK
Configuring Text-Similarity-0.13 ... OK
Building and testing Text-Similarity-0.13 ... OK
Successfully installed Text-Similarity-0.13
Building and testing WordNet-Similarity-2.05 ... OK
Successfully installed WordNet-Similarity-2.05
2 distributions installed
WordNet::Similarity

- Perl Module:
  - https://metacpan.org/pod/WordNet::Similarity

```perl
use WordNet::QueryData;
use WordNet::Similarity::path;
my $wn = WordNet::QueryData->new;
my $measure = WordNet::Similarity::path->new ($wn);
my $value = $measure->getRelatedness("car#n#1", "bus#n#2");
my ($error, $errorString) = $measure->getError();
die $errorString if $error;
print "car (sense 1) <-> bus (sense 2) = $value\n";
```
Connection paths and Similarity

- [http://www.nltk.org/howto/wordnet.html](http://www.nltk.org/howto/wordnet.html)

```python
>>> dog = wn.synset('dog.n.01')
>>> cat = wn.synset('cat.n.01')
>>> hit = wn.synset('hit.v.01')
>>> slap = wn.synset('slap.v.01')

>>> dog.path_similarity(cat)
0.2
>>> hit.path_similarity(slap)
0.14285714285714285
>>> dog.lch_similarity(cat)
2.0281482472922856
>>> hit.lch_similarity(slap)
1.3121863889661687
>>> dog.wup_similarity(cat)
0.8571428571428571
>>> wn.wup_similarity(hit, slap)
0.25
```
Connection paths and Similarity

>>> from nltk.corpus import wordnet_ic
>>> brown_ic = wordnet_ic.ic('ic-brown.dat')
>>> dog.res_similarity(cat, brown_ic)
7.911666509036577
>>> hit.res_similarity(slap, brown_ic)
0
>>> dog.jcn_similarity(cat, brown_ic)
0.4497755285516739
>>> hit.jcn_similarity(slap, brown_ic)
0.055239649531940084
>>> dog.lin_similarity(cat, brown_ic)
0.8768009843733973
>>> hit.lin_similarity(slap, brown_ic)
0.0

• Overall, the *ic-brown.dat* file lists every word existing in the Brown corpus and their information content values (which are associated with word frequencies).
Examples from class

```python
>>> from nltk.corpus import wordnet as wn
>>> hit = wn.synset('hit.v.01')
>>> slap = wn.synset('slap.v.01')
>>> punch = wn.synset('punch.v.01')
>>> hit.path_similarity(slap)
0.14285714285714285
>>> slap.path_similarity(punch)
0.2
>>> kill = wn.synset('kill.v.01')
>>> murder = wn.synset('murder.v.01')
>>> kill.path_similarity(murder)
0.5
>>> dog = wn.synset('dog.n.01')
>>> cat = wn.synset('cat.n.01')
>>> dog.path_similarity(cat)
0.2
>>> wife = wn.synset('wife.n.01')
>>> husband = wn.synset('husband.n.01')
>>> wife.path_similarity(husband)
0.3333333333333333
>>> father = wn.synset('father.n.01')
>>> son = wn.synset('son.n.01')
>>> mother = wn.synset('mother.n.01')
>>> daughter = wn.synset('daughter.n.01')
>>> father.path_similarity(son)
0.1
>>> mother.path_similarity(daughter)
0.1
>>> toad = wn.synset('toad.n.01')
>>> frog = wn.synset('frog.n.01')
>>> frog.path_similarity(toad)
1.0
>>> dog.path_similarity(cat)
0.2
```

1. (2) **frog**, toad, toad frog, anuran, batrachian, salientian -- (any of various tailless stout-bodied amphibians with long hind limbs for leaping; semiaquatic and terrestrial species)