NLP

Introduction to NLP

Sentiment Lexicons

Sentiment Lexicons

- SentiWordNet
  - http://sentiwordnet.isti.cnr.it/
- General Inquirer
  - 2,000 positive words and 2,000 negative words
  - http://www.wjh.harvard.edu/~inquirer/
- LIWC
  - http://www.liwc.net/
- MPQA subjectivity lexicon
  - http://www.cs.pitt.edu/mpqa/subj_lexicon.html

General Inquirer

- Annotations
  - Strong Power Weak Submit Active Passive Pleasur Pain Feel Arousal EMOT Virtue Vice Ovrst Undrst Academ Doctrin Econ@ Exch ECON Exprsv Legal Milit Polit@ POLIT Relig Role COLL Work Ritual SocRel Race Kin@ MALE Female Nonadlt HU ANI PLACE Social Region Route Aquatic Land Sky Object Tool Food Vehicle BldgPt ComnObj NatObj BodyPt ComForm COM Say Need Goal Try Means Persist Complet Fail NatrPro Begin Vary Increas Decreas Finish Stay Rise Exert Fetch Travel Fall Think Know Causal Ought Perceiv Compare Eval@ EVAL Solve Abs@ ABS Quality Quan NUMB ORD CARD FREQ DIST Time@ TIME Space POS DIM Rel COLOR Self Our You Name Yes No Negate Intrj IAV DAV SV IPadj IndAdj PowGain PowLoss PowEnds PowAren PowCon PowCoop PowAuPt PowPt PowDoct PowAuth PowOth PowTot RcEthic RcRelig RcGain RcLoss RcEnds RcTot RspGain RspLoss RspOth RspTot AffGain AffLoss AffPt AffOth AffTot WltPt WltTran WltOth WltTot WlbGain WlbLoss WlbPhys WlbPsyc WlbPt WlbTot EnlGain EnlLoss EnlEnds EnlPt EnlOth EnlTot SklAsth SklPt SklOth SklTot TrnGain TrnLoss TranLw MeansLw EndsLw ArenaLw PtLw Nation Anomie NegAff PosAff SureLw If NotLw TimeSpc
- http://www.webuse.umd.edu:9090/tags/
  - Positive: able, accolade, accuracy, adept, adequate…
  - Negative: addiction, adversity, adultery, affliction, aggressive…

Dictionary-based Methods

- Start from known seeds
  - e.g., happy, angry
- Expand using WordNet
  - synonyms
  - hypernyms
- Random-walk based methods
  - words with known polarity as absorbing boundary

Automatic Extraction of Sentiment Words

- Semi-supervised methods

Vasileios Hatzivassiloglou and Kathleen R. McKeown. 1997. Predicting the Semantic Orientation of Adjectives. ACL, 174–181

Molistic

- NACLO problem (2007)

Imagine that you have heard these sentences:

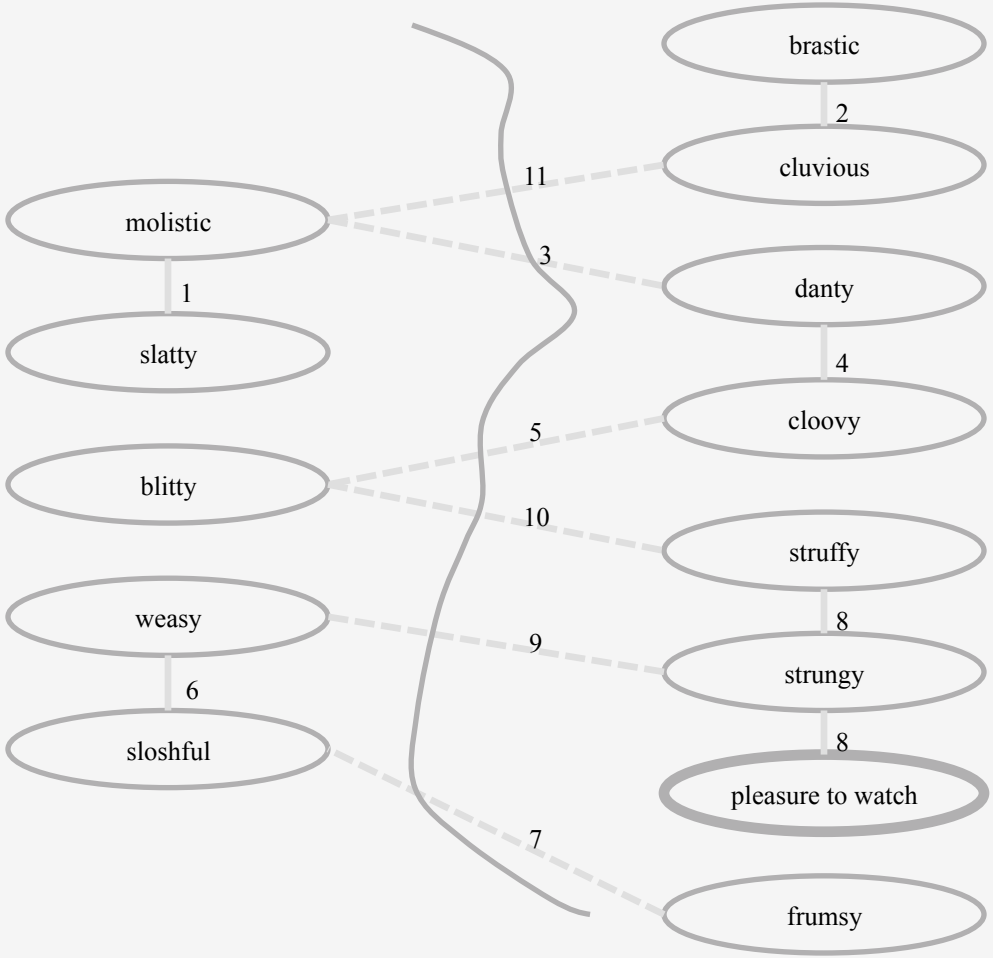
Jane is molistic and slatty.
Jennifer is cluvious and brastic.
Molly and Kyle are slatty but danty.
The teacher is danty and cloovy.
Mary is blitty but cloovy.
Jeremiah is not only sloshful but also weasy.
Even though frumsy, Jim is sloshful.
Strungy and struffy, Diane was a pleasure to watch.
Even though weasy, John is strungy.
Carla is blitty but struffy.
The salespeople were cluvious and not slatty.

A1. Then which of the following would you be likely to hear?

a. Meredith is blitty and brastic.
b. The singer was not only molistic but also cluvious.
c. May found a dog that was danty but sloshful.

A2. What quality or qualities would you be looking for in a person?

a. blitty
b. weasy
c. sloshful


brastic
2
cluvious
11
molistic
3
danty
1
slatty
4
cloovy
5
blitty
10
struffy
8
weasy
9
strungy
6
8
sloshful
pleasure to watch
7
frumsy

PMI (Turney)

- PMI=pointwise mutual information
- Check how often a given unlabeled word appears with a known positive word ("excellent")
- Same for a known negative word ("poor")

$$\text{PMI}(word_1, word_2) = \log_2 \frac{hits(word_1 \text{ NEAR } word_2)}{hits(word_1)hits(word_2)}$$

Dataset

- http://www.cs.jhu.edu/~mdredze/datasets/sentiment/

NLP