Computing with Affective Lexicons

Affective, Sentimental, and Connotative Meaning in the Lexicon
Affective meaning

• Drawing on literatures in
  • affective computing (Picard 95)
  • linguistic subjectivity (Wiebe and colleagues)
  • social psychology (Pennebaker and colleagues)

• Can we model the lexical semantics relevant to:
  • sentiment
  • emotion
  • personality
  • mood
  • attitudes
Why compute affective meaning?

- Detecting:
  - sentiment towards politicians, products, countries, ideas
  - frustration of callers to a help line
  - stress in drivers or pilots
  - depression and other medical conditions
  - confusion in students talking to e-tutors
  - emotions in novels (e.g., for studying groups that are feared over time)

- Could we generate:
  - emotions or moods for literacy tutors in the children’s storybook domain
  - emotions or moods for computer games
  - personalities for dialogue systems to match the user
Connotation in the lexicon

- Words have connotation as well as sense
- Can we build lexical resources that represent these connotations?
- And use them in these computational tasks?
Scherer’s typology of affective states

**Emotion**: relatively brief episode of synchronized response of all or most organismic subsystems in response to the evaluation of an event as being of major significance

- angry, sad, joyful, fearful, ashamed, proud, desperate

**Mood**: diffuse affect state …change in subjective feeling, of low intensity but relatively long duration, often without apparent cause

- cheerful, gloomy, irritable, listless, depressed, buoyant

**Interpersonal stance**: affective stance taken toward another person in a specific interaction, coloring the interpersonal exchange

- distant, cold, warm, supportive, contemptuous

**Attitudes**: relatively enduring, affectively colored beliefs, preferences predispositions towards objects or persons

- liking, loving, hating, valuing, desiring

**Personality traits**: emotionally laden, stable personality dispositions and behavior tendencies, typical for a person

- nervous, anxious, reckless, morose, hostile, envious, jealous
Computing with Affective Lexicons

Sentiment Lexicons
Scherer’s typology of affective states

**Emotion**: relatively brief episode of synchronized response of all or most organismic subsystems in response to the evaluation of an event as being of major significance
  - angry, sad, joyful, fearful, ashamed, proud, desperate

**Mood**: diffuse affect state ...change in subjective feeling, of low intensity but relatively long duration, often without apparent cause
  - cheerful, gloomy, irritable, listless, depressed, buoyant

**Interpersonal stance**: affective stance taken toward another person in a specific interaction, coloring the interpersonal exchange
  - distant, cold, warm, supportive, contemptuous

**Attitudes**: relatively enduring, affectively colored beliefs, preferences predispositions towards objects or persons
  - liking, loving, hating, valuing, desiring

**Personality traits**: emotionally laden, stable personality dispositions and behavior tendencies, typical for a person
  - nervous, anxious, reckless, morose, hostile, envious, jealous
The General Inquirer


• Home page: http://www.wjh.harvard.edu/~inquirer
• List of Categories: http://www.wjh.harvard.edu/~inquirer/homcat.htm
• Spreadsheet: http://www.wjh.harvard.edu/~inquirer/inquirerbasic.xls
• Categories:
  • Positiv (1915 words) and Negativ (2291 words)
  • Strong vs Weak, Active vs Passive, Overstated versus Understated
  • Pleasure, Pain, Virtue, Vice, Motivation, Cognitive Orientation, etc
• Free for Research Use
LIWC (Linguistic Inquiry and Word Count)


- 2300 words, >70 classes
- **Affective Processes**
  - negative emotion (*bad, weird, hate, problem, tough*)
  - positive emotion (*love, nice, sweet*)
- **Cognitive Processes**
  - Tentative (*maybe, perhaps, guess*), Inhibition (*block, constraint*)
- **Pronouns, Negation** (*no, never*), **Quantifiers** (*few, many*)
- $30 or $90 fee
MPQA Subjectivity Cues Lexicon


• Home page: http://www.cs.pitt.edu/mpqa/subj_lexicon.html
• 6885 words from 8221 lemmas
  • 2718 positive
  • 4912 negative
• Each word annotated for intensity (strong, weak)
• GNU GPL
Bing Liu Opinion Lexicon


- Bing Liu's Page on Opinion Mining
- http://www.cs.uic.edu/~liub/FBS/opinion-lexicon-English.rar

- 6786 words
  - 2006 positive
  - 4783 negative
SentiWordNet

Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010 SENTIWORDNET 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. LREC-2010

- Home page: [http://sentiwordnet.isti.cnr.it/](http://sentiwordnet.isti.cnr.it/)
- All WordNet synsets automatically annotated for degrees of positivity, negativity, and neutrality/objectiveness
- \[\text{estimable}(J,3)\] “may be computed or estimated”
  
  \[
  \begin{array}{ccc}
  \text{Pos} & 0 & \text{Neg} & 0 & \text{Obj} & 1 \\
  \end{array}
  \]

- \[\text{estimable}(J,1)\] “deserving of respect or high regard”
  
  \[
  \begin{array}{ccc}
  \text{Pos} & .75 & \text{Neg} & 0 & \text{Obj} & .25 \\
  \end{array}
  \]
Scherer’s typology of affective states

**Emotion**: relatively brief episode of synchronized response of all or most organismic subsystems in response to the evaluation of an event as being of major significance

- angry, sad, joyful, fearful, ashamed, proud, desperate

**Mood**: diffuse affect state...change in subjective feeling, of low intensity but relatively long duration, often without apparent cause

- cheerful, gloomy, irritable, listless, depressed, buoyant

**Interpersonal stance**: affective stance taken toward another person in a specific interaction, coloring the interpersonal exchange

- distant, cold, warm, supportive, contemptuous

**Attitudes**: relatively enduring, affectively colored beliefs, preferences predispositions towards objects or persons

- liking, loving, hating, valuing, desiring

**Personality traits**: emotionally laden, stable personality dispositions and behavior tendencies, typical for a person

- nervous, anxious, reckless, morose, hostile, envious, jealous
Two families of theories of emotion

• Atomic basic emotions
  • A finite list of 6 or 8, from which others are generated

• Dimensions of emotion
  • Valence (positive negative)
  • Arousal (strong, weak)
  • Control
Ekman’s 6 basic emotions:
Surprise, happiness, anger, fear, disgust, sadness
Valence/Arousal Dimensions

- **High arousal, low pleasure**: anger
- **High arousal, high pleasure**: excitement
- **Low arousal, low pleasure**: sadness
- **Low arousal, high pleasure**: relaxation
Atomic units vs. Dimensions

**Distinctive**
- Emotions are units.
- Limited number of basic emotions.
- Basic emotions are innate and universal

**Dimensional**
- Emotions are dimensions.
- Limited # of labels but unlimited number of emotions.
- Emotions are culturally learned.

Adapted from Julia Braverman
One emotion lexicon from each paradigm!

1. 8 basic emotions:
   • NRC Word-Emotion Association Lexicon (Mohammad and Turney 2011)

2. Dimensions of valence/arousal/dominance
   • Warriner, A. B., Kuperman, V., and Brysbaert, M. (2013)

• Both built using Amazon Mechanical Turk
Plutchick’s wheel of emotion

• 8 basic emotions
• in four opposing pairs:
  • joy–sadness
  • anger–fear
  • trust–disgust
  • anticipation–surprise
NRC Word-Emotion Association Lexicon

Mohammad and Turney 2011

- 10,000 words chosen mainly from earlier lexicons
- Labeled by Amazon Mechanical Turk
- 5 Turkers per hit
- Give Turkers an idea of the relevant sense of the word
- Result:

```
amazingly  anger  0
amazingly  anticipation  0
amazingly  disgust  0
amazingly  fear  0
amazingly  joy  1
amazingly  sadness  0
amazingly  surprise  1
amazingly  trust  0
amazingly  negative  0
amazingly  positive  1
```
The AMT Hit

Prompt word: startle

Q1. Which word is closest in meaning (most related) to startle?
   - automobile
   - shake
   - honesty
   - entertain

Q2. How positive (good, praising) is the word startle?
   - startle is not positive
   - startle is weakly positive
   - startle is moderately positive
   - startle is strongly positive

Q3. How negative (bad, criticizing) is the word startle?
   - startle is not negative
   - startle is weakly negative
   - startle is moderately negative
   - startle is strongly negative

Q4. How much is startle associated with the emotion joy? (For example, happy and fun are strongly associated with joy.)
   - startle is not associated with joy
   - startle is weakly associated with joy
   - startle is moderately associated with joy
   - startle is strongly associated with joy

Q5. How much is startle associated with the emotion sadness? (For example, failure and heartbreak are strongly associated with sadness.)
   - startle is not associated with sadness
   - startle is weakly associated with sadness
   - startle is moderately associated with sadness
   - startle is strongly associated with sadness

Q6. How much is startle associated with the emotion fear? (For example, horror and scary are strongly associated with fear.)
   - Similar choices as in 4 and 5 above

Q7. How much is startle associated with the emotion anger? (For example, rage and shouting are strongly associated with anger.)
   - Similar choices as in 4 and 5 above

Q8. How much is startle associated with the emotion trust? (For example, faith and integrity are strongly associated with trust.)
   - Similar choices as in 4 and 5 above

Q9. How much is startle associated with the emotion disgust? (For example, gross and cruelty are strongly associated with disgust.)
   - Similar choices as in 4 and 5 above

...
Lexicon of valence, arousal, and dominance

- Supplementary data: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License.

- Ratings for 14,000 words for emotional dimensions:
  - **valence** (the pleasantness of the stimulus)
  - **arousal** (the intensity of emotion provoked by the stimulus)
  - **dominance** (the degree of control exerted by the stimulus)
Lexicon of valence, arousal, and dominance

- **valence** (the pleasantness of the stimulus)
  - 9: happy, pleased, satisfied, contented, hopeful
  - 1: unhappy, annoyed, unsatisfied, melancholic, despaired, or bored
- **arousal** (the intensity of emotion provoked by the stimulus)
  - 9: stimulated, excited, frenzied, jittery, wide-awake, or aroused
  - 1: relaxed, calm, sluggish, dull, sleepy, or unaroused;
- **dominance** (the degree of control exerted by the stimulus)
  - 9: in control, influential, important, dominant, autonomous, or controlling
  - 1: controlled, influenced, cared-for, awed, submissive, or guided
- Again produced by AMT
Lexicon of valence, arousal, and dominance: Examples

<table>
<thead>
<tr>
<th>Valence</th>
<th>Arousal</th>
<th>Dominance</th>
</tr>
</thead>
<tbody>
<tr>
<td>vacation</td>
<td>8.53</td>
<td>rampage</td>
</tr>
<tr>
<td>happy</td>
<td>8.47</td>
<td>tornado</td>
</tr>
<tr>
<td>whistle</td>
<td>5.7</td>
<td>zucchini</td>
</tr>
<tr>
<td>conscious</td>
<td>5.53</td>
<td>dressy</td>
</tr>
<tr>
<td>torture</td>
<td>1.4</td>
<td>dull</td>
</tr>
</tbody>
</table>
Concreteness versus abstractness

- The degree to which the concept denoted by a word refers to a perceptible entity.
  - Do concrete and abstract words differ in connotation?
  - Storage and retrieval?
  - Bilingual processing?
  - Relevant for embodied view of cognition (Barsalou 1999 inter alia)
    - Do concrete words activate brain regions involved in relevant perception
  - Supplementary data: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License.
  - 37,058 English words and 2,896 two-word expressions ("zebra crossing" and "zoom in"),
  - Rating from 1 (abstract) to 5 (concrete)
  - Calibrator words:
    - shirt, infinity, gas, grasshopper, marriage, kick, polite, whistle, theory, and sugar
Concreteness versus abstractness

- Supplementary data: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License.
- Some example ratings from the final dataset of 40,000 words and phrases
  
  banana 5  
bathrobe 5  
bagel 5  
brisk 2.5  
badass 2.5  
basically 1.32  
belief 1.19  
although 1.07
# Perceptual Strength Norms

## Connell and Lynott norms

<table>
<thead>
<tr>
<th>Word</th>
<th>Perceptual strength</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Auditory</td>
<td>Gustatory</td>
<td>Haptic</td>
<td>Olfactory</td>
<td>Visual</td>
<td>Concreteness</td>
<td>Imageability</td>
</tr>
<tr>
<td>soap</td>
<td>0.35</td>
<td>1.29</td>
<td>4.12</td>
<td>4.00</td>
<td>4.06</td>
<td>589</td>
<td>600</td>
</tr>
<tr>
<td>noisy</td>
<td>4.95</td>
<td>0.05</td>
<td>0.29</td>
<td>0.05</td>
<td>1.67</td>
<td>293</td>
<td>138</td>
</tr>
<tr>
<td>atom</td>
<td>1.00</td>
<td>0.63</td>
<td>0.94</td>
<td>0.50</td>
<td>1.38</td>
<td>481</td>
<td>499</td>
</tr>
<tr>
<td>republic</td>
<td>0.53</td>
<td>0.67</td>
<td>0.27</td>
<td>0.07</td>
<td>1.79</td>
<td>376</td>
<td>356</td>
</tr>
</tbody>
</table>

Microsoft Excel Worksheet
Computing with Affective Lexicons

Using the lexicons to detect affect
Lexicons for detecting document affect: Simplest unsupervised method

- **Sentiment:**
  - Sum the weights of each positive word in the document
  - Sum the weights of each negative word in the document
  - Choose whichever value (positive or negative) has higher sum

- **Emotion:**
  - Do the same for each emotion lexicon
Lexicons for detecting document affect: Simplest supervised method

- Build a classifier
  - Predict sentiment (or emotion, or personality) given features
  - Use “counts of lexicon categories” as a features
  - Sample features:
    - LIWC category “cognition” had count of 7
    - NRC Emotion category “anticipation” had count of 2
- Baseline
  - Instead use counts of all the words and bigrams in the training set
  - This is hard to beat
  - But only works if the training and test sets are very similar
Computing with Affective Lexicons

Sample affective task: personality detection
Scherer’s typology of affective states

**Emotion**: relatively brief episode of synchronized response of all or most organismic subsystems in response to the evaluation of an event as being of major significance

- angry, sad, joyful, fearful, ashamed, proud, desperate

**Mood**: diffuse affect state ... change in subjective feeling, of low intensity but relatively long duration, often without apparent cause

- cheerful, gloomy, irritable, listless, depressed, buoyant

**Interpersonal stance**: affective stance taken toward another person in a specific interaction, coloring the interpersonal exchange

- distant, cold, warm, supportive, contemptuous

**Attitudes**: relatively enduring, affectively colored beliefs, preferences predispositions towards objects or persons

- liking, loving, hating, valuing, desiring

**Personality traits**: emotionally laden, stable personality dispositions and behavior tendencies, typical for a person

- nervous, anxious, reckless, morose, hostile, envious, jealous
The Big Five Dimensions of Personality

Extraversion vs. Introversion
sociable, assertive, playful vs. aloof, reserved, shy

Emotional stability vs. Neuroticism
calm, unemotional vs. insecure, anxious

Agreeableness vs. Disagreeable
friendly, cooperative vs. antagonistic, faultfinding

Conscientiousness vs. Unconscientious
self-disciplined, organised vs. inefficient, careless

Openness to experience
intellectual, insightful vs. shallow, unimaginative
Various text corpora labeled for personality of author

**Pennebaker, James W., and Laura A. King. 1999.** "Linguistic styles: language use as an individual difference." *Journal of personality and social psychology* 77, no. 6.

- 2,479 essays from psychology students (1.9 million words), “write whatever comes into your mind” for 20 minutes

**Mehl, Matthias R, SD Gosling, JW Pennebaker. 2006.** Personality in its natural habitat: manifestations and implicit folk theories of personality in daily life. *Journal of personality and social psychology* 90 (5), 862

  - Speech from Electronically Activated Recorder (EAR)
  - Random snippets of conversation recorded, transcribed
  - 96 participants, total of 97,468 words and 15,269 utterances

**Schwartz, H. Andrew, Johannes C. Eichstaedt, Margaret L. Kern, Lukasz Dziurzynski, Stephanie M. Ramones, Megha Agrawal, Achal Shah et al. 2013.** "Personality, gender, and age in the language of social media: The open-vocabulary approach." *PloS one* 8, no. 9

  - Facebook
  - 75,000 volunteers
  - 309 million words
  - All took a personality test
## Ears (speech) corpus (Mehl et al.)

<table>
<thead>
<tr>
<th>Introvert</th>
<th>Extravert</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Yeah you would do kilograms. Yeah I see what you’re saying.</td>
<td>- That’s my first yogurt experience here.</td>
</tr>
<tr>
<td>- On Tuesday I have class. I don’t know.</td>
<td>Really watery. Why?</td>
</tr>
<tr>
<td>- I don’t know. A16. Yeah, that is kind of cool.</td>
<td>- Damn. New game.</td>
</tr>
<tr>
<td>- I don’t know. I just can’t wait to be with you and not have to do this every night, you know?</td>
<td>- Oh.</td>
</tr>
<tr>
<td>- Yeah. You don’t know. Is there a bed in there? Well ok just...</td>
<td>- That’s so rude. That.</td>
</tr>
<tr>
<td>- With the Chinese. Get it together.</td>
<td>- Yeah, but he, they like each other.</td>
</tr>
<tr>
<td>- I tried to yell at you through the window.</td>
<td>He likes her.</td>
</tr>
<tr>
<td>Oh. xxxx’s fucking a dumb ass. Look at him. Look at him. I wish we had a camera. He’s fucking brushing his t-shirt with a tooth brush. Get a kick of it. Don’t steal nothing.</td>
<td>- They are going to end up breaking up and he’s going to be like.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unconscientious</th>
</tr>
</thead>
<tbody>
<tr>
<td>- With the Chinese. Get it together.</td>
</tr>
<tr>
<td>- I tried to yell at you through the window.</td>
</tr>
<tr>
<td>Oh. xxxx’s fucking a dumb ass. Look at him. Look at him. I wish we had a camera. He’s fucking brushing his t-shirt with a tooth brush. Get a kick of it. Don’t steal nothing.</td>
</tr>
</tbody>
</table>
### Essays corpus (Pennebaker and King)

<table>
<thead>
<tr>
<th>Introvert</th>
<th>Extravert</th>
</tr>
</thead>
<tbody>
<tr>
<td>I’ve been waking up on time so far. What has it been, 5 days? Dear me, I’ll never keep it up, being such not a morning person and all. But maybe I’ll adjust, or not. I want internet access in my room, I don’t have it yet, but I will on Wed??? I think. But that ain’t soon enough, cause I got calculus homework [...]</td>
<td>I have some really random thoughts. I want the best things out of life. But I fear that I want too much! What if I fall flat on my face and don’t amount to anything. But I feel like I was born to do BIG things on this earth. But who knows... There is this Persian party today.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neurotic</th>
<th>Emotionally stable</th>
</tr>
</thead>
<tbody>
<tr>
<td>One of my friends just barged in, and I jumped in my seat. This is crazy. I should tell him not to do that again. I’m not that fastidious actually. But certain things annoy me. The things that would annoy me would actually annoy any normal human being, so I know I’m not a freak.</td>
<td>I should excel in this sport because I know how to push my body harder than anyone I know, no matter what the test I always push my body harder than everyone else. I want to be the best no matter what the sport or event. I should also be good at this because I love to ride my bike.</td>
</tr>
</tbody>
</table>
Classifiers

  - Various classifiers, lexicon-based and prosodic features
  - regression and SVM, lexicon-based and all-words
## Sample LIWC Features

LIWC (Linguistic Inquiry and Word Count)


<table>
<thead>
<tr>
<th>Feature</th>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger words</td>
<td>LIWC</td>
<td>hate, kill, pissed</td>
</tr>
<tr>
<td>Metaphysical issues</td>
<td>LIWC</td>
<td>God, heaven, coffin</td>
</tr>
<tr>
<td>Physical state/function</td>
<td>LIWC</td>
<td>ache, breast, sleep</td>
</tr>
<tr>
<td>Inclusive words</td>
<td>LIWC</td>
<td>with, and, include</td>
</tr>
<tr>
<td>Social processes</td>
<td>LIWC</td>
<td>talk, us, friend</td>
</tr>
<tr>
<td>Family members</td>
<td>LIWC</td>
<td>mom, brother, cousin</td>
</tr>
<tr>
<td>Past tense verbs</td>
<td>LIWC</td>
<td>walked, were, had</td>
</tr>
<tr>
<td>References to friends</td>
<td>LIWC</td>
<td>pal, buddy, coworker</td>
</tr>
<tr>
<td>Imagery of words</td>
<td>MRC</td>
<td>Low: future, peace - High: table, car</td>
</tr>
<tr>
<td>Syllables per word</td>
<td>MRC</td>
<td>Low: a - High: uncompromisingly</td>
</tr>
<tr>
<td>Concreteness</td>
<td>MRC</td>
<td>Low: patience, candor - High: ship</td>
</tr>
<tr>
<td>Frequency of use</td>
<td>MRC</td>
<td>Low: duly, nudity - High: he, the</td>
</tr>
</tbody>
</table>
Normalizing LIWC category features
(Schwartz et al 2013, Facebook study)

• Mairesse:
  Raw LIWC counts
• Schwartz et al:
  Normalized per writer:

\[ p(\text{category} \mid \text{subject}) = \frac{\sum_{\text{word} \in \text{category}} \text{freq (word, subject)}}{\sum_{\text{word} \in \text{vocab (subject)}} \text{freq (word, subject)}} \]
Sample results

• Agreeable:
  • +Family, +Home, -Anger, -Swear

• Extravert
  • +Friend, +Religion, +Self

• Conscientiousness:
  • -Swear, -Anger, -NegEmotion,

• Emotional Stability:
  • -NegEmotion, +Sports,

• Openness
  • -Cause, -Space
Decision tree for predicting extraversion in essay corpus (Mairesse et al)