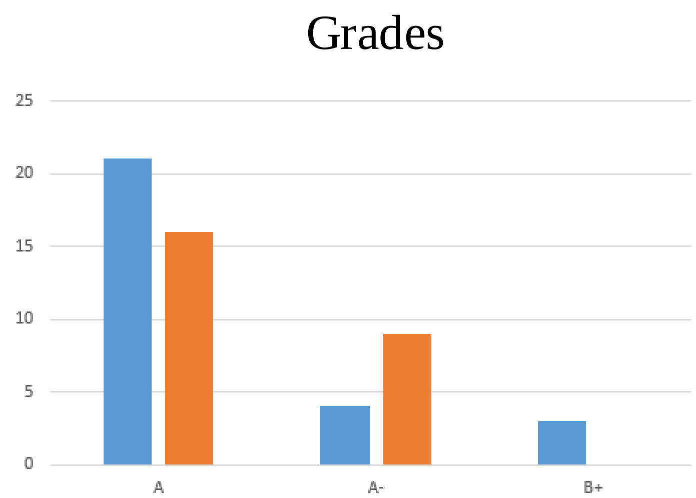
Sentiment Analysis And Other Non-explicit Semantics Of Text

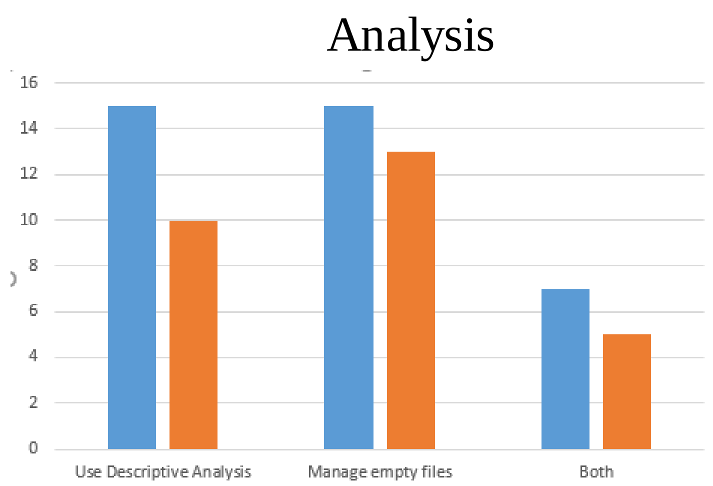
Lu Xiao
lxiao04@syr.edu
213 Hinds Hall

adopted some materials developed in previous courses by Nancy McCracken, Liz Liddy and others; and some instructor resources for the book “Speech and Language Processing” by Daniel Jurafsky and James H. Martin

HOMEWORK 2



Morning session
Afternoon session



Descriptive analysis: the student explained max and min Number of Awards by NSF Organization, distribution of awards amounts, max and min number of sentences, etc.

HOMEWORK 2 - NOTES

- Adding comments to your code is a good programing practice
- Use of tables, figures, and graphs to summarize your results is a great practice commonly used in Industry and Academia
- Remember to cite your sources

Affect, Feeling, Emotion, Sentiment, and Opinion Detection (Munezero, Montero, Sutinen, & Pajunen, 2014)

- A major limitation in the automatic detection of affect, feelings, emotions, sentiments, and opinions in text: the lack of proper differentiation between these subjective terms and understanding of how they relate to one another

- The differences between these five subjective terms

MUNEZERO, MYRIAM & SUERO MONTERO, CALKIN & SUTINEN, ERKKI & PAJUNEN, JOHN. "**ARE THEY DIFFERENT? AFFECT, FEELING, EMOTION, SENTIMENT, AND OPINION DETECTION IN TEXT**" AFFECTIVE COMPUTING, IEEE TRANSACTIONS ON. 5. 101-111 (2014)

Affect, Feeling, Emotion, Sentiment, and Opinion Detection (Munezero, Montero, Sutinen, & Pajunen, 2014)

Affect:

- Non-conscious
- difficult to conceptualize in language
- What can be detected from text is the conscious expression of affect, which we found to be feelings and emotions
- body's way of preparing itself for action in a given circumstance by adding a quantitative dimension of intensity to the quality of an experience

Feeling:

- Conscious phenomena
- Can be detected from text
- sensation that has been checked against previous experiences and labeled
- personal and biographical because every person has a distinct set of previous sensations from which to draw from when interpreting and labelling their feelings

MUNEZERO, MYRIAM & SUERO MONTERO, CALKIN & SUTINEN, ERKKI & PAJUNEN, JOHN. "**ARE THEY DIFFERENT? AFFECT, FEELING, EMOTION, SENTIMENT, AND OPINION DETECTION IN TEXT**" AFFECTIVE COMPUTING, IEEE TRANSACTIONS ON. 5. 101-111 (2014)

Affect, Feeling, Emotion, Sentiment, and Opinion Detection (Munezero, Montero, Sutinen, & Pajunen, 2014)

Emotion:

- complex psychological phenomena
- What we are able to detect is the written conscious experience of five factors (appraisals, feelings, physiological reactions, expressive behavior, and readiness to act in a certain way), which constitute emotions.
- use of any words to convey emotions is influenced by culture
- emotions are in fact expressions of affect
- affective reactions are the conscious representation of affect, which in language is expressed through feelings and emotions.
- social processes and cultural norms play a significant role in specifying how we express emotions

MUNEZERO, MYRIAM & SUERO MONTERO, CALKIN & SUTINEN, ERKKI & PAJUNEN, JOHN. "**ARE THEY DIFFERENT? AFFECT, FEELING, EMOTION, SENTIMENT, AND OPINION DETECTION IN TEXT**" AFFECTIVE COMPUTING, IEEE TRANSACTIONS ON. 5. 101-111 (2014)

Affect, Feeling, Emotion, Sentiment, and Opinion Detection (Munezero, Montero, Sutinen, & Pajunen, 2014)

Emotion (Cont.):

- Attempts of detecting emotions have been made using lexical resources such as **WordNetAffect** to mark the presence of emotion bearing words in text or **Affective Norms for English Words (ANEW)** to mark words with normative emotional ratings
- commonsense knowledge bases such as **EmotiNet** have been used to capture emotions in texts that do not have emotion-bearing words naturally
- the choice of words and their intended meaning are personally, contextually, culturally, and socially dependent

Ex: *That ride is wicked.*

the sentence has an unpleasant connotation in some communities; however, the sentence would be associated with pleasantness in some communities such as the United Kingdom.

MUNEZERO, MYRIAM & SUERO MONTERO, CALKIN & SUTINEN, ERKKI & PAJUNEN, JOHN. "**ARE THEY DIFFERENT? AFFECT, FEELING, EMOTION, SENTIMENT, AND OPINION DETECTION IN TEXT**." AFFECTIVE COMPUTING, IEEE TRANSACTIONS ON. 5. 101-111 (2014)

Affect, Feeling, Emotion, Sentiment, and Opinion Detection (Munezero, Montero, Sutinen, & Pajunen, 2014)

Sentiments:

- enduring emotional dispositions that have developed over time about particular objects
- Conclusions about sentiments in text have to be performed for a period.
- Examples of sentiments include romantic love, parental love, loyalty, friendship, patriotism, hate, as well as more transient, acute emotional responses, to social losses (sorrow, envy) and gains (pride, gratitude)
- emotions can be mapped into two broad categories: like and dislike

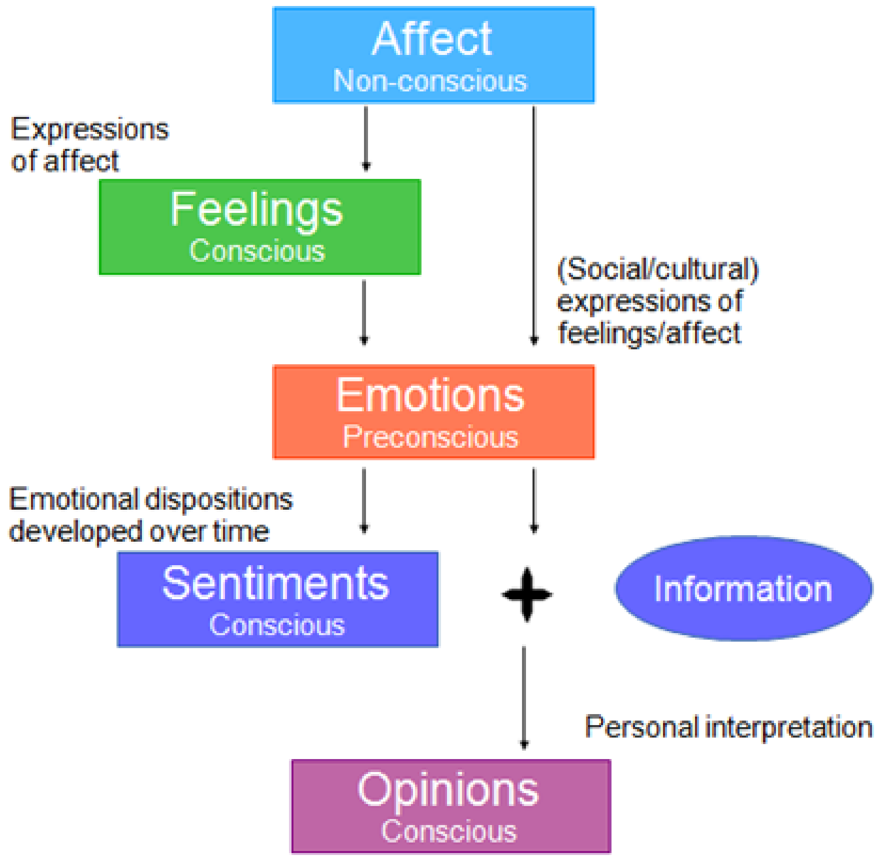
Affect, Feeling, Emotion, Sentiment, and Opinion Detection (Munezero, Montero, Sutinen, & Pajunen, 2014)

Opinions:

- personal interpretations of information, which may or may not be associated with an emotion or sentiment.
- Judgement based on grounds insufficient to produce certainty. Thus, opinions leave room for error in thought.
- an opinion consists of the following four parts: topic, opinion holder, claim, and sentiment. That is, for each opinion, there is a holder who believes a claim about a topic and then associates a positive, negative, or neutral sentiment with the claim.

MUNEZERO, MYRIAM & SUERO MONTERO, CALKIN & SUTINEN, ERKKI & PAJUNEN, JOHN. "**ARE THEY DIFFERENT? AFFECT, FEELING, EMOTION, SENTIMENT, AND OPINION DETECTION IN TEXT**" AFFECTIVE COMPUTING, IEEE TRANSACTIONS ON. 5. 101-111 (2014)

Affect, Feeling, Emotion, Sentiment, and Opinion Detection (Munezero, Montero, Sutinen, & Pajunen, 2014)



MUNEZERO, MYRIAM & SUERO MONTERO, CALKIN & SUTINEN, ERKKI & PAJUNEN, JOHN. "**ARE THEY DIFFERENT? AFFECT, FEELING, EMOTION, SENTIMENT, AND OPINION DETECTION IN TEXT**". AFFECTIVE COMPUTING, IEEE TRANSACTIONS ON. 5. 101-111 (2014)

Why Sentiment Analysis?

- *Movie*: is this review positive or negative?
- *Products*: what do people think about the new iPhone?
- *Customer relations:* what to people say about your company?
- *Public sentiment*: how is consumer confidence? Is despair increasing?
- *Politics*: what do people think about this candidate or issue?
- *Prediction*: predict election outcomes or market trends from sentiment

Types Of Tasks:

- Positive or negative movie review?
  -  unbelievably disappointing
  -  Full of zany characters and richly applied satire, and some great plot twists
  -  this is the greatest screwball comedy ever filmed
  -  It was pathetic. The worst part about it was the boxing scenes.

Types Of Tasks: Google Product Search

HP Officejet 6500A Plus e-All-in-One Color Ink-jet - Fax / copier / printer / scanner

$89 online, **$100** nearby ★★★★☆ 377 reviews

September 2010 - Printer - HP - Inkjet - Office - Copier - Color - Scanner - Fax - 250 sh

Reviews

Summary - Based on 377 reviews

1 star	2	3	4 stars	5 stars

What people are saying

ease of use	"This was very easy to setup to four computers."
value	"Appreciate good quality at a fair price."
setup	"Overall pretty easy setup."
customer service	"I DO like honest tech support people."
size	"Pretty Paper weight."
mode	"Photos were fair on the high quality mode."
colors	"Full color prints came out with great quality."

TYPES OF TASKS: BING SHOPPING

HP Officejet 6500A E710N Multifunction Printer

Product summary Find best price **Customer reviews** Specifications Related items

$121.53 - $242.39 (14 stores)

☐ Compare

Average rating ★★★½☆ (144)

★★★★★	(55)
★★★★☆	(54)
★★★☆☆	(10)
★★☆☆☆	(6)
★☆☆☆☆	(23)
☆☆☆☆☆	(0)

Most mentioned

Performance	(57)
Ease of Use	(43)
Print Speed	(39)
Connectivity	(31)

More ▼

Show reviews by source

- Best Buy (140)
- CNET (5)
- Amazon.com (3)

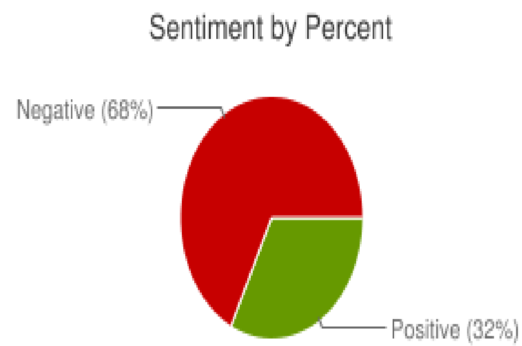
Target Sentiment On Twitter

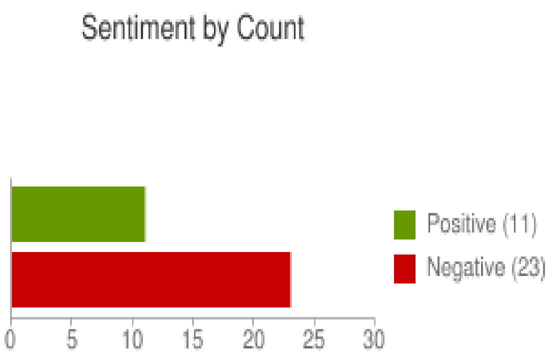
- [Twitter Sentiment App]()
- Alec Go, Richa Bhayani, Lei Huang. 2009. Twitter Sentiment Classification using Distant Supervision

Type in a word and we'll highlight the good and the bad

"united airlines" Search [Save this search]()

Sentiment analysis for "united airlines"





jljacobson: OMG... Could **@United airlines** have worse customer service? W8g now 15 minutes on hold 4 questions about a flight 2DAY that need a human.
Posted 2 hours ago

12345clumsy6789: I hate **United Airlines** Ceiling!!! Fukn impossible to get my conduit in this damn mess! ?
Posted 2 hours ago

EMLandPRGbelgiu: EML/PRG fly with Q8 **united airlines** and 24seven to an exotic destination. http://t.co/Z9QloAjF
Posted 2 hours ago

CountAdam: FANTASTIC customer service from **United Airlines** at XNA today. Is tweet more, but cell phones off now!
Posted 4 hours ago

Terminology

- For the more subtle aspects of the semantics of text:
- Affective aspects of text is that which is "influenced by or resulting from emotions"
  - One aspect of non-factual aspects of text
- Subjective aspects of text
  "The **linguistic** expression of somebody's **opinions**, **sentiments**, emotions, evaluations, beliefs, speculations ***(private states)***"
  - A private state is not open to objective observation or verification
  - Subjectivity analysis would classify parts of text as to whether it was subjective or objective

Scherer Typology Of Affective States

- **Emotion**: brief organically synchronized … evaluation of a major event
  - *angry, sad, joyful, fearful, ashamed, proud, elated*
- **Mood**: diffuse non-caused low-intensity long-duration change in subjective feeling
  - *cheerful, gloomy, irritable, listless, depressed, buoyant*
- **Interpersonal stances**: affective stance toward another person in a specific interaction
  - *friendly, flirtatious, distant, cold, warm, supportive, contemptuous*
- **Attitudes**: enduring, affectively colored beliefs, dispositions towards objects or persons
  - *liking, loving, hating, valuing, desiring*
- **Personality traits**: stable personality dispositions and typical behavior tendencies
  - *nervous, anxious, reckless, morose, hostile, jealous*

Sentiment Analysis

- **Emotion**: brief organically synchronized … evaluation of a major event
  - *angry, sad, joyful, fearful, ashamed, proud, elated*
- **Mood**: diffuse non-caused low-intensity long-duration change in subjective feeling
  - *cheerful, gloomy, irritable, listless, depressed, buoyant*
- **Interpersonal stances**: affective stance toward another person in a specific interaction
  - *friendly, flirtatious, distant, cold, warm, supportive, contemptuous*
- **Attitudes: enduring, affectively colored beliefs, dispositions towards objects or persons**
  - *liking, loving, hating, valuing, desiring*
- **Personality traits**: stable personality dispositions and typical behavior tendencies
  - *nervous, anxious, reckless, morose, hostile, jealous*

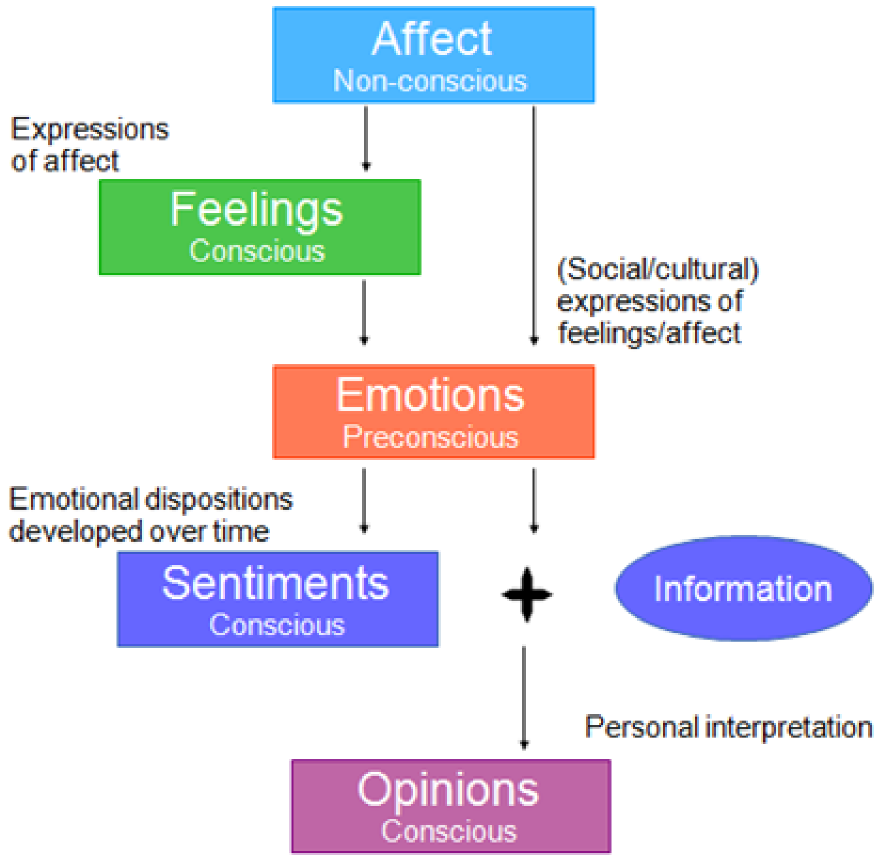
Sentiment Analysis

- Sentiment analysis is the detection of attitudes
  "enduring, affectively colored beliefs, dispositions towards objects or persons"
  - **Holder (source)** of attitude
  - **Target (aspect)** of attitude
  - **Type** of attitude
    - From a set of types
      - Like, love, hate, value, desire, etc.
    - Or (more commonly) simple weighted polarity:
      - Positive, negative, neutral, together with strength
  - **Text** containing the attitude
    - Sentence or entire document

Sentiment Analysis Task Levels

- **Simplest task is polarity**:
  - Is the attitude of this text positive or negative?
    - Negative / positive attitude of reporter / blogger
    - Favorable / unfavorable review of a product
    - Right / left political leaning of speaker
- **More complex:**
  - Rank the attitude of this text from 1 to 5
    - Sometimes called strength or intensity
- **Advanced:**
  - Detect the target, source or complex attitude type
  - May also be referred to as
    opinion extraction, opinion mining, or sentiment mining

Affect, Feeling, Emotion, Sentiment, and Opinion Detection (Munezero, Montero, Sutinen, & Pajunen, 2014)



MUNEZERO, MYRIAM & SUERO MONTERO, CALKIN & SUTINEN, ERKKI & PAJUNEN, JOHN. "**ARE THEY DIFFERENT? AFFECT, FEELING, EMOTION, SENTIMENT, AND OPINION DETECTION IN TEXT**." AFFECTIVE COMPUTING, IEEE TRANSACTIONS ON. 5. 101-111 (2014)

What's The Problem?

- Consider classifying a subjective text unit as either positive or negative.
  - Example: The most thoroughly joyless and inept film of the year, and one of the worst of the decade. [Mick LaSalle, describing *Gigli*]
- Can't we just look for words like *great* or *terrible* ?
  - Yes, but …
    - Words have different meanings in different contexts

What's The Problem?

- Subtlety, sarcasm or metaphor:
  - Perfume review in *Perfumes: the Guide*:
    - "If you are reading this because it is your darling fragrance, please wear it at home exclusively, and tape the windows shut."
  - Dorothy Parker on Katherine Hepburn
    - "She runs the gamut of emotions from A to B"

Thwarted Expectations And Ordering Effects

- "This film should be brilliant. It sounds like a great plot, the actors are first grade, and the supporting cast is good as well, and Stallone is attempting to deliver a good performance. However, it **can't hold up**."
- Well as usual Keanu Reeves is nothing special, but surprisingly, the very talented Laurence Fishbourne is **not so good** either, I was surprised.

Domain Adaptation

- Certain sentiment-related indicators seem domain-dependent.
  - .*Read the book*..: good for book reviews, bad for movie reviews
  - .*Unpredictable*.: good for movie plots, bad for a car's steering [Turney '02]
- In general, sentiment classifiers (especially those created via supervised learning) have been shown to often be domain dependent
  - [Turney '02, Engstr ¨om '04, Read 05,Aue & Gamon '05, Blitzer, Dredze & Pereira '07].

Sentiment Detection: Polarity And Intensity

Sentiment Polarity

- Classic Sentiment polarity task from Pang and Lee:
  - Is an IMDB movie review positive or negative?
  - Data: *Polarity Data 2.0: (people indicate polarity of own review)*
    - http://www.cs.cornell.edu/people/pabo/movie-review-data
- Treat as a document classification task
  - Positive, negative, and (possibly) neutral
- Similar but different from topic-based text classification.
  - In topic-based text classification, topic words are important.
  - In sentiment classification, sentiment words are more important, e.g., great, excellent, horrible, bad, worst, etc.

Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? Sentiment Classification using Machine Learning Techniques. EMNLP-2002, 79—86.
Bo Pang and Lillian Lee. 2004. A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts. ACL, 271-278

Treat As A Classification Problem

- Tokenization
  - May be some differences, especially for text from social media
- Feature Extraction
  - The most important part!
- Classification using different classifiers
  - Naïve Bayes
  - MaxEnt
  - SVM

It turns out that MaxEnt and SVM are better than Naïve Bayes at some sentiment domains.

Sentiment Tokenization Issues

- For text from web, deal with HTML and XML markup
- Or Twitter mark-up (names, hash tags)
- Capitalization (preserve for words in all caps)
- Phone numbers, dates
- Emoticons
- Useful code:
  - [Christopher Potts sentiment tokenizer]()

Potts emoticons

```
[<>]?                        # optional hat/brow
[:;=8]                       # eyes
[\-o\*\']?                   # optional nose
[\)\]\(\[dDpP/\:\}\{@\|\\]   # mouth
|                            #### reverse orientation
[\)\]\(\[dDpP/\:\}\{@\|\\]   # mouth
[\-o\*\']?                   # optional nose
[:;=8]                       # eyes
[<>]?                        # optional hat/brow
```

Extracting Features For Sentiment Classification

- Which words to use?
  - Only adjectives
  - All words
    - All words turns out to work better, at least on this data
- Syntax is not used as often
  - Constituent or dependency parses are occasionally used
  - Particularly at phrase level to find dependencies of opinion words
    - Also for finding the scope of negation
  - Can be used to shift the "valence"
    - For negation, intensification and diminution
      - Very good, deeply suspicious
      - Should have been good
      - He is a great actor, *however* this performance . . .
        - However changes the valence of great to be negative

Handling Negation Is Important!

- How to handle negation:
  - I **didn't** like this movie vs I really like this movie
  - Pang and Lee simple approximation to negation:
    - Add NOT_ to every word between negation and following punctuation:

```
didn't like this movie , but I
```

⬇

```
didn't NOT_like NOT_this NOT_movie but I
```

- Negation has both scope and focus
  - These may be represented in more complex structures
  - Details in Wilson "Fine-grained sentiment analysis"

Sentiment Lexicons

- One of the early approaches to sentiment analysis was to just count the words in each document that had either a positive or negative polarity from a (hand-built) sentiment lexicon.
  - This approach usually not very accurate on individual documents, but it's easy because doesn't need training data.
  - May be useful over aggregate collections or to show trends over time.
- Now we use either presence or frequencies of sentiment words as features of the classifier

MPQA Subjectivity Cues Lexicon

- Gives a list of words that have been judged to be weakly or strongly positive, negative or neutral in subjectivity
- Home page: http://www.cs.pitt.edu/mpqa/subj_lexicon.html
- 6885 words from 8221 lemmas
  - 2718 positive, 4912 negative
  - GNU GPL license
  - Examples:

type=weaksubj len=1 word1=abandoned pos1=adj stemmed1=n priorpolarity=negative
type=weaksubj len=1 word1=abandonment pos1=noun stemmed1=n priorpolarity=negative
type=weaksubj len=1 word1=abandon pos1=verb stemmed1=y priorpolarity=negative
type=strongsubj len=1 word1=abase pos1=verb stemmed1=y priorpolarity=negative
type=strongsubj len=1 word1=abasement pos1=anypos stemmed1=y priorpolarity=negative

Theresa Wilson, Janyce Wiebe, and Paul Hoffmann (2005). Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. Proc. of HLT-EMNLP-2005.

LIWC (Linguistic Inquiry And Word Count)

- Text analysis software based on dictionaries of word dimensions
- Dimensions can be syntactic
  - Pronouns, past-tense verbs
- Dimensions can be semantic
  - Social words, affect, cognitive mechanisms
- Other categories (http://www.liwc.net/comparedicts.php)

- James Pennebaker, Univ. of Texas at Austin
  - http://www.liwc.net/

- Often used for positive and negative emotion words in opinion mining

ANEW (Affective Norms For English Words)

- From the NIMH Center for the Study of Emotion and Attention at the University of Florida
  - http://csea.phhp.ufl.edu/Media.html
  - See also the paper by Dodds and Danforth on Happiness of Large-Scale Written Expressions
  - Free for research use

- Provides a set of emotional ratings for a large number of words in the English language

The General Inquirer

- Home page: http://www.wjh.harvard.edu/~inquirer
- List of Categories: http://www.wjh.harvard.edu/~inquirer/homecat.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

Philip J. Stone, Dexter C Dunphy, Marshall S. Smith, Daniel M. Ogilvie. 1966. The General Inquirer: A Computer Approach to Content Analysis. MIT Press

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

Minqing Hu and Bing Liu. Mining and Summarizing Customer Reviews. ACM SIGKDD-2004.

SentiWordNet

- Home page: http://sentiwordnet.isti.cnr.it/
- All WordNet synsets automatically annotated for degrees of positivity, negativity, and neutrality/objectiveness
- [estimable(J,3)] "may be computed or estimated"

```
Pos  0   Neg 0   Obj 1
```

- [estimable(J,1)] "deserving of respect or high regard"

```
Pos .75  Neg 0   Obj .25
```

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

Which Sentiment Lexicon To Use?

- An area of active research in the sentiment analysis community
- It is now recognized that the amount of overlap between the lexicons is small!
  - But in general, where there is overlap, the sentiment polarity of the words is in agreement, 2% or less disagreement.
    - Except for SentiWordNet, which disagrees up to 25%
      - Chris Potts, Sentiment Symposium Tutorial
- How to represent features from sentiment words still under research:
  - Frequency of all positive and all negative words
  - Presence of positive or negative words (particularly for twitter)
  - Sum of the positive or negative intensity scores

Build A Sentiment Lexicon

- For some domains, it has been shown that the best lexicon is one built for that domain
- Automatic lexicon building from unlabeled data
  - Bootstrapping
    - Identify a number of seed words of positive and negative polarity
    - Search for text involving those words that also have connecting words, such as "and"
    - Other words that occur with the connecting word are added to the lexicon with the appropriate polarity

Building A Sentiment Lexicon

- Automatic lexicon building from labeled data
  - In some cases, the domain has lots of text that has been labeled with sentiment
  - Twitter
    - Use tweets labeled with sentiment hashtags: #good, #happy, #bad, #sad
    - Use tweets labeled with happy or sad emoticons
  - Collect words from the positive and negative labeled texts and keep the frequent ones as part of a lexicon
  - Using Mutual Information scores or other measures

VADER

VADER is reported to be among the best sentiment prediction tool available for classifying social media texts and online review comments (Hutto & Gilbert, 2014; Ribeiro, Araújo, Gonçalves, Gonçalves, & Benevenuto, 2016)

- Lexicon
- Syntactic rules, e.g., all capitalization, punctuation marks, etc.

Hutto, C.J., Gilbert, E.E.: VADER: a parsimonious rule-based model for sentiment analysis of social media text. In: Eighth International Conference on Weblogs and Social Media (ICWSM-2014), Ann Arbor, MI, June 2014

Ribeiro, F. N., Araújo, M., Gonçalves, P., Gonçalves, M. A., & Benevenuto, F. (2016). SentiBench-a benchmark comparison of state-of-the-practice sentiment analysis methods. EPJ Data Science, 5(1), 23.

Opinion Analysis

Opinion Mining

- The third level of sentiment analysis is sometimes called opinion mining because you are finding sentiment towards aspects or attributes
- Businesses spend a huge amount of money to find consumer sentiments and opinions.
  - Consultants, surveys and focus groups, etc
  - Text in the form of transcripts of interviews or survey responses
- Opinions also available on the web
  - product reviews
  - blogs, discussion groups

Sentence Level Detection

- Sentence level or sub-sentence level detection of subjectivity as a binary classifier
  - Wiebe, many projects
  - Pang and Lee – for movie reviews, first determine which sentences express opinions and then label for opinion polarity
- Clause level opinion strength
  - Wilson, "How mad are you?"
- Detection of sentences with subjectivity or sentiment is important to then find aspects or attributes
  - *The food was great but the service was awful.*

Finding Aspect/Attribute/Target Of Sentiment

- Frequent phrases + rules
  - Find all highly frequent phrases across reviews (“`fish tacos`”)
  - Filter by rules like “occurs right after sentiment word”
    - “...`great fish tacos`” means `fish tacos` a likely aspect

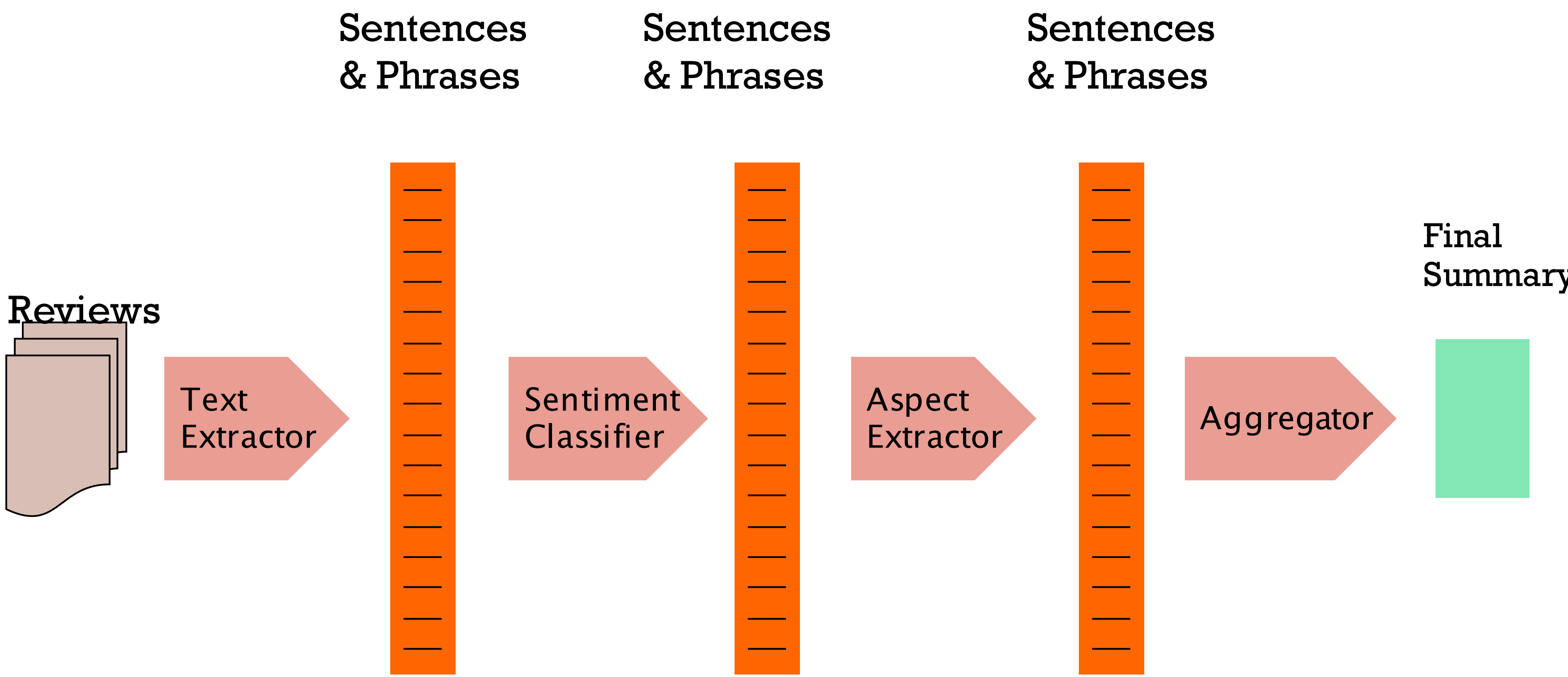
Casino	casino, buffet, pool, resort, beds
Children’s Barber	haircut, job, experience, kids
Greek Restaurant	food, wine, service, appetizer, lamb
Department Store	selection, department, sales, shop, clothing

M. Hu and B. Liu. 2004. Mining and summarizing customer reviews. In Proceedings of KDD.
S. Blair-Goldensohn, K. Hannan, R. McDonald, T. Neylon, G. Reis, and J. Reynar. 2008. Building a Sentiment Summarizer for Local Service Reviews. WWW Workshop.

Finding Aspect/Attribute/Target Of Sentiment

- But the aspect name may not be in the sentence
- Other methods to find aspects:
  - For restaurants/hotels, aspects are well-understood
  - Supervised classification
    - Hand-label a small corpus of restaurant review sentences with aspect
    - food, décor, service, value, NONE
  - Train a classifier to assign an aspect to a sentence
    - "Given this sentence, is the aspect *food, décor, service, value,* or *NONE*"

Putting It All Together: Finding Sentiment/Opinion For Aspects



S. Blair-Goldensohn, K. Hannan, R. McDonald, T. Neylon, G. Reis, and J. Reynar. 2008. Building a Sentiment Summarizer for Local Service Reviews. WWW Workshop

Joint Topic/Sentiment Analysis

- An alternative approach to first finding the aspect or attribute and then the opinion or sentiment is to find them both in the same classification
  - Comparative studies of related products
  - Topics that have various features and attributes
    - Consumers
    - Political areas

Lin, C., & He, Y. (2009, November). Joint sentiment/topic model for sentiment analysis. In *Proceedings of the 18th ACM conference on Information and knowledge management* (pp. 375-384). ACM.

Example Results For Opinion Of Aspect

Rooms (3/5 stars, 41 comments)

- (+) The room was clean and everything worked fine – even the water pressure ...
- (+) We went because of the free room and was pleasantly pleased ...
- (-) …the worst hotel I had ever stayed at ...

Service (3/5 stars, 31 comments)

- (+) Upon checking out another couple was checking early due to a problem ...
- (+) Every single hotel staff member treated us great and answered every ...
- (-) The food is cold and the service gives new meaning to SLOW.

Dining (3/5 stars, 18 comments)

- (+) our favorite place to stay in biloxi.the food is great also the service ...
- (+) Offer of free buffet for joining the Play

Feature-based Summary (Hu AND Liu, Kdd-04)

- From reviews, extract a summary:

GREAT Camera., Jun 3, 2004

Reviewer: **jprice174** from Atlanta, Ga.

I did a lot of research last year before I bought this camera... It kinda hurt to leave behind my beloved nikon 35mm SLR, but I was going to Italy, and I needed something smaller, and digital.

The pictures coming out of this camera are amazing. The 'auto' feature takes great pictures most of the time. And with digital, you're not wasting film if the picture doesn't come out. …

….

Feature Based Summary:

Feature1: **picture**

Positive: 12

- The pictures coming out of this camera are amazing.
- Overall this is a good camera with a really good picture clarity.

…

Negative: 2

- The pictures come out hazy if your hands shake even for a moment during the entire process of taking a picture.
- Focusing on a display rack about 20 feet away in a brightly lit room during day time, pictures produced by this camera were blurry and in a shade of orange.

Feature2: **battery life**

…

Detection Of Representative Rationales In Wikipedia Afd Deliberations (Mao, Xiao, & Mercer, 2014)

Input a discussion

Vote, deliberation content, and user name

Divide the discussion into groups

Sentence-to-sentence semantic similarity (SEMILAR)

Divide each group into three subgroups (positive, neutral, negative)

Using **our recursive algorithm** to predict polarity

Extract representative rationales based on the similarity score

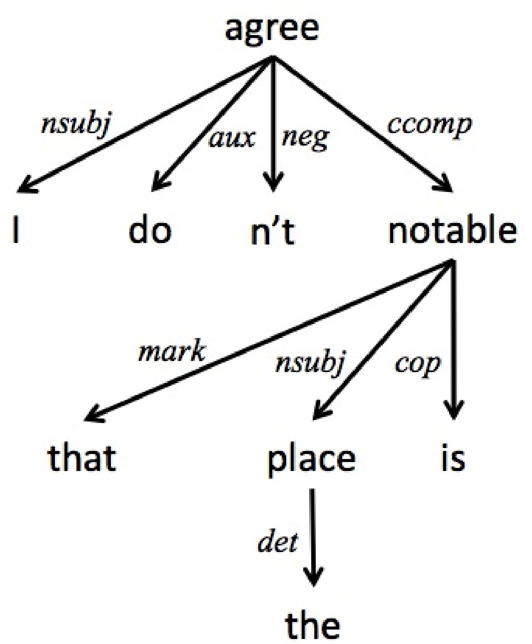
Our Recursive Algorithm to Predict Sentiment Polarity at Sentence Level

-Based on the dependency structure tree

- Take a node as input, and the polarity score for the node as output.

-The algorithm assigns a polarity score to each node in the dependency structure tree

-Integrate five types of negations



Five Negations

- Local negation: A not usually modifies the sentiment word.
  - "The place is *not notable*."
- Predicate negation: using verbs with negative polarity.
  - "I *disagree* that the place is *notable*."
- Subject negation: a subject leads to the negation of its predicate.
  - "*Neither* one of us *agrees* that the place is *notable*."
- Preposition negation: the polarity of the object following the preposition "of" can be changed by the word modified by the preposition.
  - "It is a *violation* of *notability.*"
- Modifier negation: some sentiment word's polarity can be negated by its modifier.
  - "The place is of *indeterminable notability*."

Sentiment Analysis

- Determine the sentiment polarity of a statement in our language context (“notable”)
  - MPQA Subjectivity Lexicon + additional words

Textual Data Input → Stanford Parser → Part-of-speech tagged text → Check modified MPQA subjectivity lexicon to obtain the prior polarity

Dependency relations

(if not in MPQA, marked as ‘non-sentiment’)

MPQA format:

type=strongsubj len=1 word1=aberration pos1=adj stemmed1=n priorpolarity=negative

Experiment And Evaluation On Sentence Polarity Prediction

- Methods
  - Stanford sentiment analysis tool vs. our algorithm

	Stanford sentiment analysis	recursive algorithm without machine learning	recursive algorithm with machine learning
Accuracy (%)	48.73	58.47	60.17

- Accuracy of Stanford sentiment analysis tool classifying movie review
  - 5-category: 45.7%, 2-category: 85.4%

Summary On Sentiment Analysis

- Understanding semantics of less factual aspects of text
- Generally modeled as classification or regression task
  - predict a binary label for polarity
  - or predict an ordinal label for the level of sentiment
- Features:
  - Negation is important
  - Using all (filtered) words works well for some tasks
  - Finding subsets of words may help in other tasks
    - Hand-built polarity lexicons
    - Use seeds and semi-supervised learning to induce lexicons