Sentiment and Opinion

Sep13, 2012

Analysis of Social Media Seminar

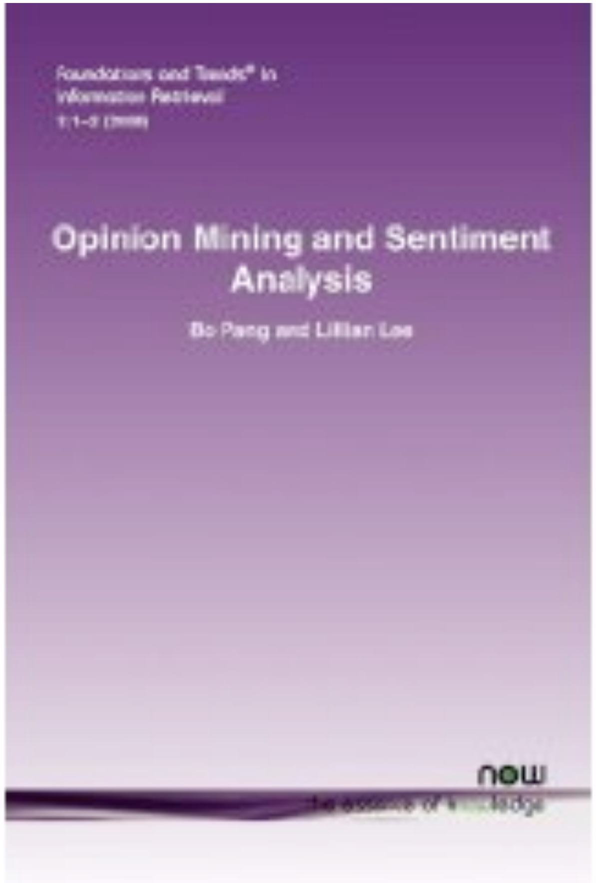
William Cohen

First assignment: due Tuesday

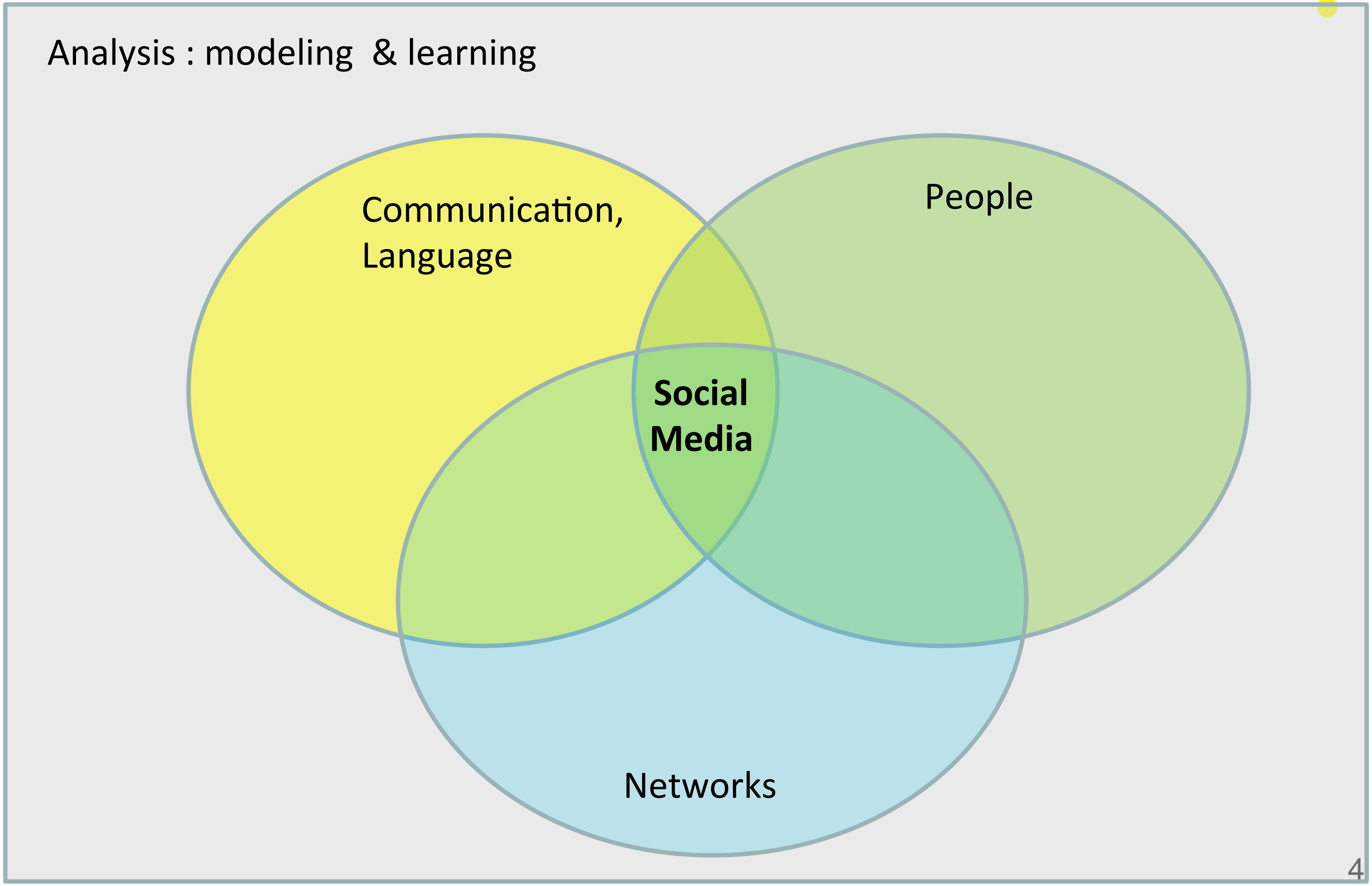
- Go to your user page
  - Your real name & a link to your home page
  - Preferably a picture
  - Who you are and what you hope to get out of the class (Let me know if you' re just auditing)
  - One sentence about what you might want to do for a project

Announcements

- Office hours
  - William: 1-2pm Fri

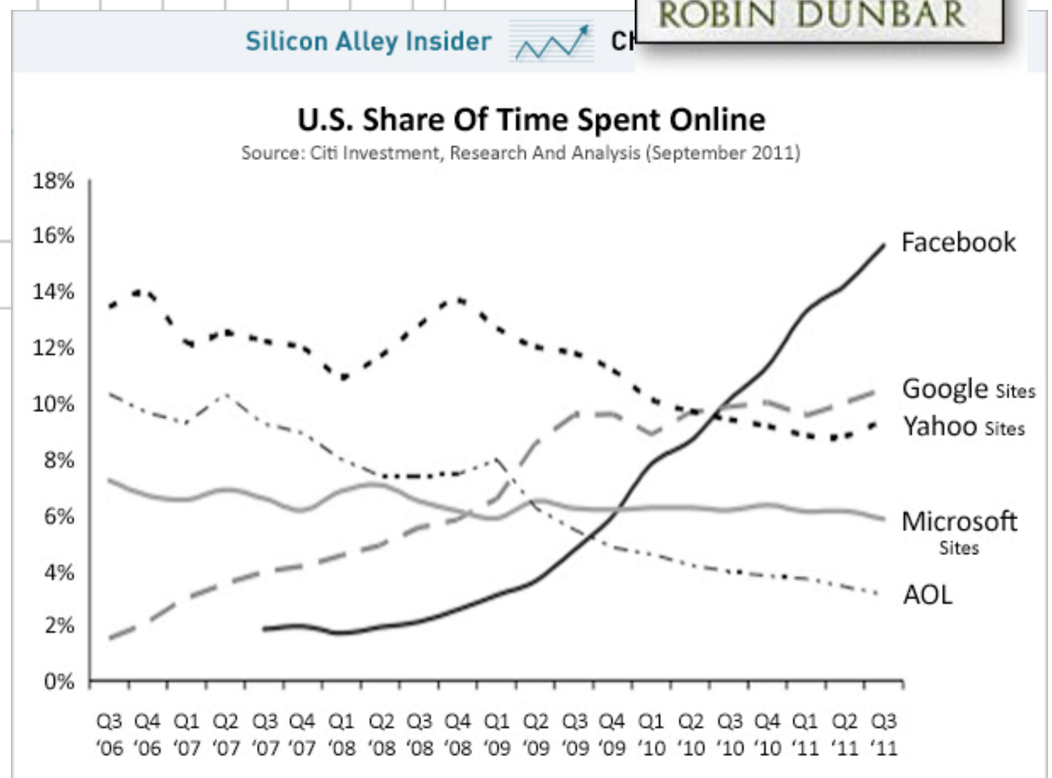


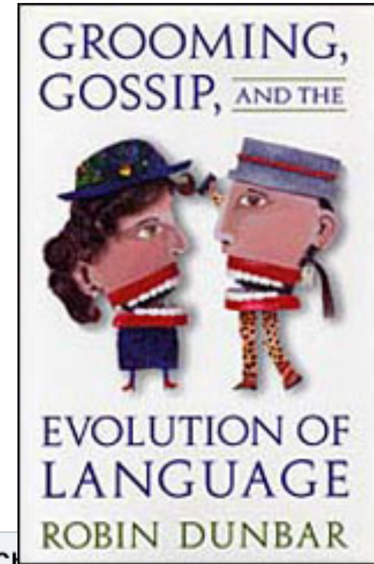
Motivations

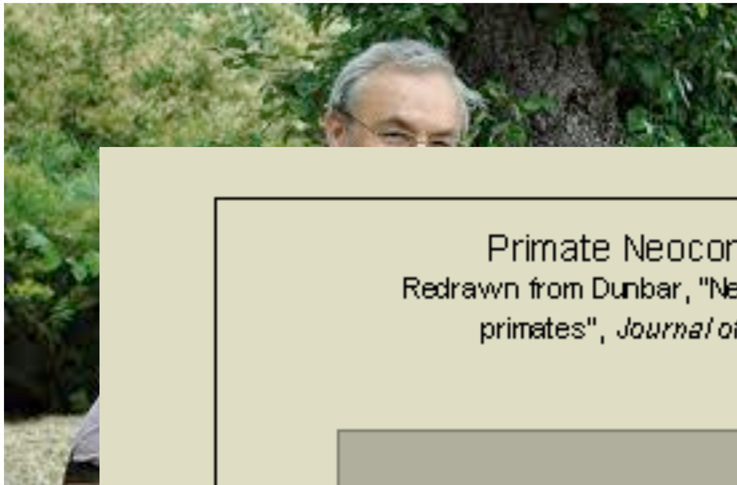


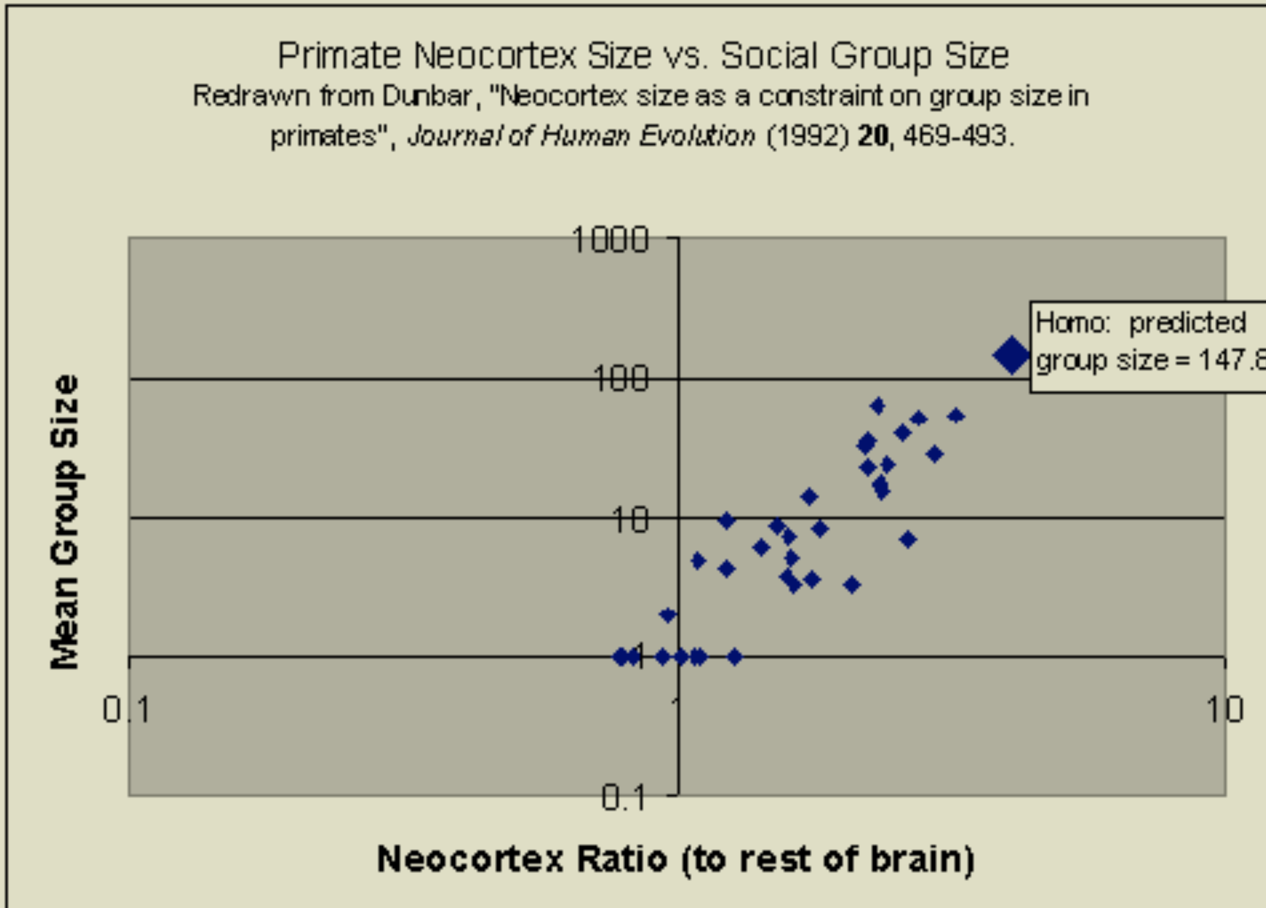
MAY 2011


Silicon Alley Insider
Chart of the Day
Top 10 U.S. Web Brands By Total Minutes
Facebook
Yahoo!
Google
AOL
MSN / WindowsLive / Bing
YouTube
eBay
Apple
EA
Microsoft
Billions of Minutes
0
5
10
15
20


Silicon Alley Insider
U.S. Share Of Time Spent Online
Source: Citi Investment, Research And Analysis (September 2011)
18%
16%
14%
12%
10%
8%
6%
4%
2%
0%
Facebook
Google Sites
Yahoo Sites
Microsoft Sites
AOL
Q3 '06
Q4 '06
Q1 '07
Q2 '07
Q3 '07
Q4 '07
Q1 '08
Q2 '08
Q3 '08
Q4 '08
Q1 '09
Q2 '09
Q3 '09
Q4 '09
Q1 '10
Q2 '10
Q3 '10
Q4 '10
Q1 '11
Q2 '11
Q3 '11


GROOMING, GOSSIP, AND THE
EVOLUTION OF LANGUAGE
ROBIN DUNBAR




Primate Neocortex Size vs. Social Group Size
Redrawn from Dunbar, "Neocortex size as a constraint on group size in primates", *Journal of Human Evolution* (1992) **20**, 469-493.
1000
100
10
1
0.1
0.1
10
Homo: predicted group size = 147.8
Mean Group Size
Neocortex Ratio (to rest of brain)

memeorandum
Political Web, Page A1

TOP ITEMS:

Wall Street Journal:

Romney Offends the Pundits — Doesn't he know he's not supposed to debate foreign policy? — Tuesday's assaults on the U.S. Embassies in Benghazi and Cairo have injected foreign policy into the Presidential campaign, but suddenly the parsons of the press corps are offended by the debate.

Discussion: ABCNEWS, Althouse, Hugh Hewitt's TownHall Blog, First Read, Le·gal In·sur·rec· tion, Weekly Standard and Instapundit

RELATED:

Josh Rogin / Foreign Policy:

Inside the public relations disaster at the Cairo embassy — One staffer at the U.S. Embassy in Cairo was responsible for the statement and tweets Tuesday that have become grist for the presidential campaign, and that staffer ignored explicit State Department instructions not to issue the statement …

Discussion: Wall Street Journal, Hot Air, American Spectator, americanthinker.com, Outside the Beltway, The Hinterland Gazette, GILL REPORT, Election 2012, NationalJournal.com, Washington Wire, Talking Points Memo, The Atlantic Online, Weekly Standard and First Read

Gail Collins / New York Times:

Mitt's Major Meltdown — Mitt Romney broke our deal. — Perhaps he didn't know he'd made it, although, really, I thought it was pretty clear. — He could do anything he wanted during this campaign as long as he sent out signals that once he got in the White House he was not likely to be truly crazy.

Discussion: The Hinterland Gazette, Balloon Juice, Connecting.the.Dots and Politico

New York Times:

Turmoil Spreads to U.S. Embassy in Yemen — SANA, Yemen — Turmoil in the Arab world linked to a contentious video denigrating the Prophet Muhammad spread on Thursday to Yemen, where hundreds of protesters stormed the United States Embassy, two days after assailants killed the American ambassador …

Discussion: Washington Post, IHT Rendezvous, Gawker, This Just In, Right Wing News, Scared Monkeys, New York Magazine, Salon, Jammie Wearing Fools, NO QUARTER, Towleroad News #gay and Outside the Beltway, more at Mediagazer »

Some sentences expressing "opinion" or something a lot like opinion

– **Wow**, this is my 4th Olympus camera.
– Most voters **believe** that he's not going to raise their taxes.
- The United States **fears** a spill-over from the anti-terrorist campaign.
- "We foresaw electoral fraud but not **daylight robbery**," Tsvangirai said.

Content cheerfully pilfered from this 250+slide tutorial: EUROLAN SUMMER SCHOOL 2007, Semantics, Opinion and Sentiment in Text, July 23-August 3, University of Iaşi, Romania http://www.cs.pitt.edu/~wiebe/tutorialsExtendedTalks.html

Manual and Automatic Subjectivity and Sentiment Analysis

Jan Wiebe
Josef Ruppenhofer
Swapna Somasundaran
University of Pittsburgh

And an invited talk from Lillian Lee at ICWSM 2008

A Tempest

Or, on the flood of interest in:
sentiment analysis,
opinion mining,
and the computational treatment of subjective language

Lillian Lee
Cornell University
http://www.cs.cornell.edu/home/llee

"Romance should never begin with sentiment. It should begin with science and end with a settlement."

— Oscar Wilde, *An Ideal Husband*

Specific motivation: “Opinion Question Answering”

Q: What is the international reaction to the reelection of Robert Mugabe as President of Zimbabwe?

A: African observers generally approved of his victory while Western Governments denounced it.

More motivations

- **Product review mining:** What features of the ThinkPad T43 do customers like and which do they dislike?
- **Review classification:** Is a review positive or negative toward the movie?
- **Tracking sentiments toward topics over time:** Is anger ratcheting up or cooling down?
- **Etc.**

[These are all ways to *summarize* one sort of content that is common on blogs, bboards, newsgroups, etc. –W]

More motivations

People search for and are affected by online opinions.

TripAdvisor, Rotten Tomatoes, Yelp, ...
Amazon, eBay, YouTube...
blogs, Q&A and discussion sites, ...

According to a Comscore '07 report and an '08 Pew survey:
60% of US residents have done online product research.
15% do so on a typical day.

73%-87% of US readers of online reviews of services say the reviews were significant influences. (more on economics later)

But 58% of US internet users report that online information was missing, impossible to find, confusing, and/or overwhelming.

Creating technologies that find and analyze reviews would answer a tremendous information need.

In 2006, 31% of US residents used the internet for gathering or sharing political information (60M+ people).

- Major reason?
  28%: to get perspectives from *within* their community.
  34%: to get perspectives from *outside* it.

- 28% said that most sites they use *share* their point of view.
  29% said that most *challenge* their point of view.

[Rainie and Horrigan Pew survey, '07]

Can't we just look for words like "great" or "terrible"?
Yes, but ...

- This laptop is a great deal.
- A great deal of media attention surrounded the release of the new laptop.
- This laptop is a great deal ... and I've got a nice bridge you might be interested in.

- 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.

- She ran the gamut of emotions from A to B. [Dorothy Parker, describing Katharine Hepburn]
- Read the book. [Bob Bland]

Some early work on opinion

A source of information about semantic orientation: **conjunction**

Web Results **1** - **10** of about **762,000** for **"was very nice and"**.

The Homestay Experience - Cultural Kaleidoscope 2006
My host's home **was very nice and** comfortable. I got to try all types of food; Malaysian, Chinese, Indonesian and I loved it all. My host's parents were very **...**
www.gardenschool.edu.my/studentportal/aec/Kaleidoscope06/experience.asp - 10k - Cached - Similar pages - Note this

PriceGrabber User Rating for Watch Your Budget - PriceGrabber.com
Reviews, Camera I purchased **was very nice and** a bargain. There was a problem with shipping, but was resolved quickly. Buy with confidence from this vendor. **...**
www.pricegrabber.com/rating_getreview.php/retid=5821 - Similar pages - Note this

Testimonials
"Everybody **was very nice and** service was as fast as they possibly could. **...** "Staff member who helped me **was very nice and** easy to talk to." **...**
www.sa.psu.edu/uhs/news/testimonials.cfm - 22k - Cached - Similar pages - Note this

Naxos Villages - Naxos Town or Chora Reviews: Very nice and very **...**
-Did you enjoy the trip to Naxos Town: Yes it **was very nice and** very scenic -In order to get to the village were there enough signs in order to find it: It **...**

Hatzivassiloglou & McKeown 1997

1. Build training set: label all adj. with frequency > 20; test agreement with human annotators
2. Extract all conjoined adjectives

Web Results **1** - **10** of about **762,000** for **"was very nice and"**.

The Homestay Experience - Cultural Kaleidoscope 2006
My host's home **was very nice and** comfortable. I got to try all types of food; Malaysian, Chinese, Indonesian and I loved it all. My host's parents were very **...**
www.gardenschool.edu.my/studentportal/aec/Kaleidoscope06/experience.asp - 10k -
Cached - Similar pages - Note this

PriceGrabber User Rating for Watch Your Budget - PriceGrabber.com
Reviews, Camera I purchased **was very nice and** a bargain. There was a problem with shipping, but was resolved quickly. Buy with confidence from this vendor. **...**
www.pricegrabber.com/rating_getreview.php/retid=5821 - Similar pages - Note this

Testimonials
"Everybody **was very nice and** service was as fast as they possibly could. **...** "Staff member who helped me **was very nice and** easy to talk to." **...**
www.sa.psu.edu/uhs/news/testimonials.cfm - 22k - Cached - Similar pages - Note this

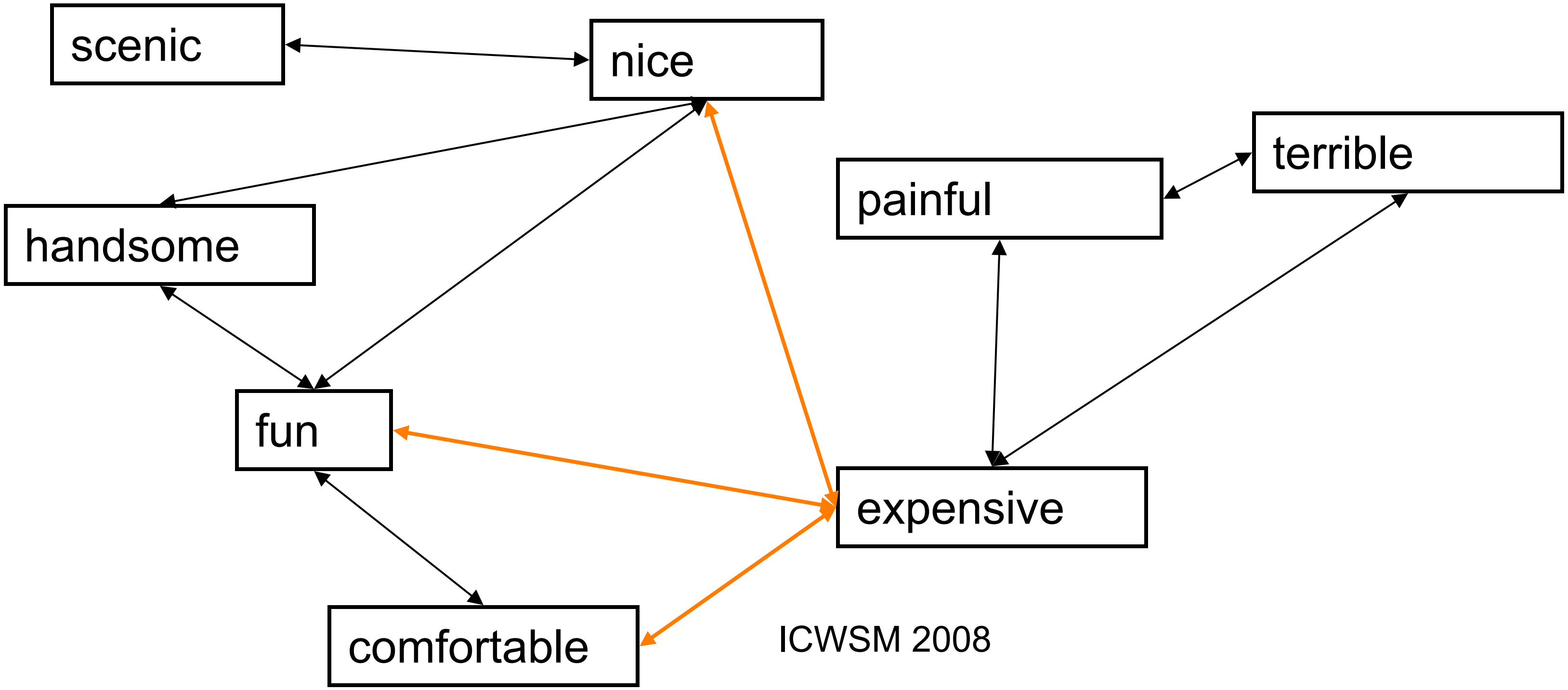
Naxos Villages - Naxos Town or Chora Reviews: Very nice and very **...**
-Did you enjoy the trip to Naxos Town: Yes it **was very nice and** very scenic. -In order to get to the village were there enough signs in order to find it: It **...**

nice and comfortable

nice and scenic

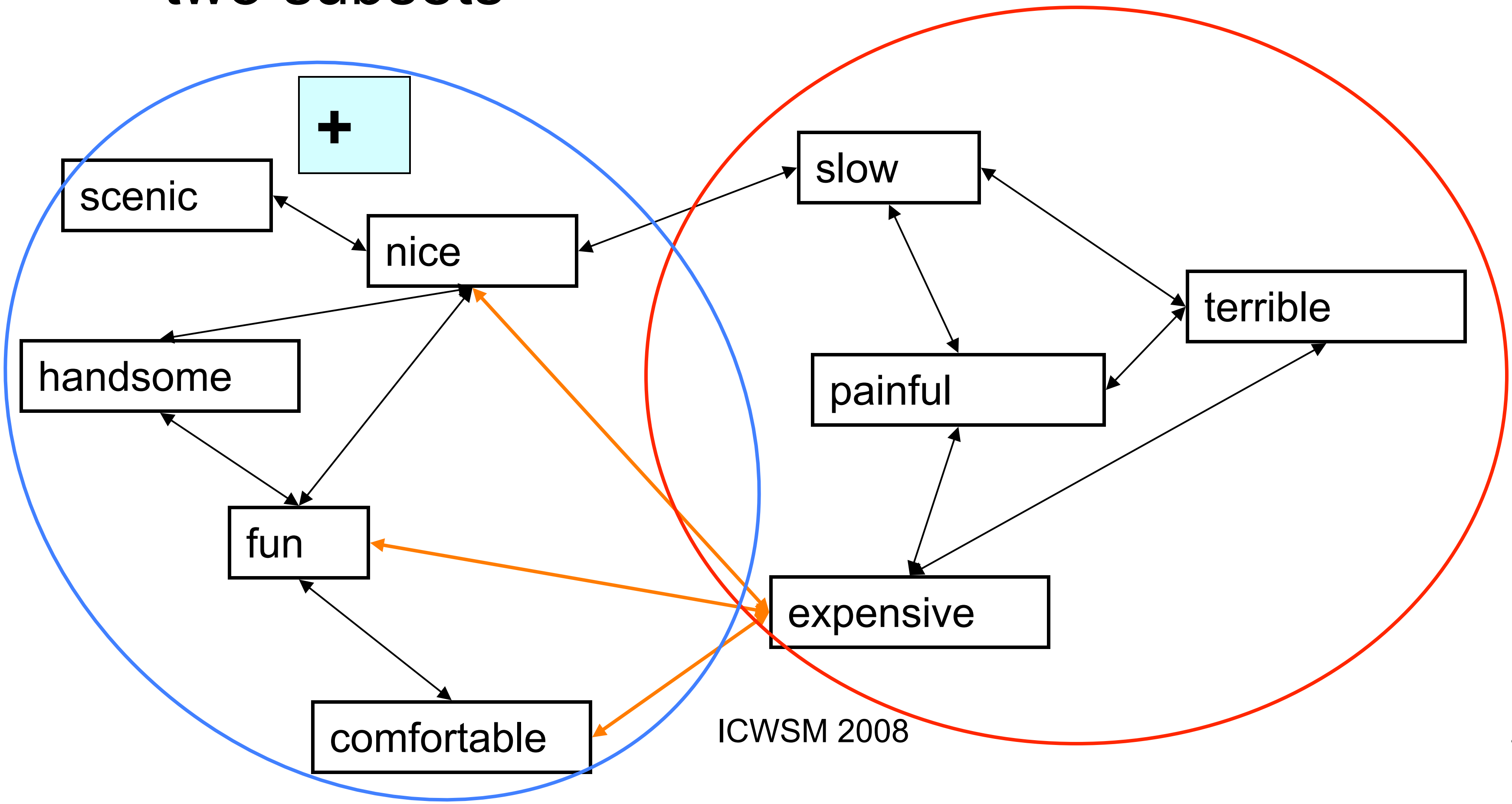
Hatzivassiloglou & McKeown 1997

3. A supervised learning algorithm builds a graph of adjectives linked by the same or different semantic orientation



Hatzivassiloglou & McKeown 1997

4. A clustering algorithm partitions the adjectives into two subsets



Hatzivassiloglou & McKeown 1997

- Specifics:
  - 21M words of POS-tagged WSJ text
  - Hand-labeled 1336 adjectives appearing >20 times ignoring ones like "unpredictable", "cheap",..
  - Good inter-annotator agreement (97% on direction of orientation, 89% on existence of orientation)
  - Used hand-built grammer to find conjunctions using *and, or, but, either-or, neither-nor*
  - Also looked at pairs like *thoughtful, thoughtless:* almost always oppositely oriented, but not frequent.

Hatzivassiloglou & McKeown 1997

Conjunction category	Conjunction types analyzed	% same-orientation (types)	% same-orientation (tokens)	P-Value (for types)
All conjunctions	2,748	77.84%	72.39%	$< 1 \cdot 10^{-16}$
All *and* conjunctions	2,294	81.73%	78.07%	$< 1 \cdot 10^{-16}$
All *or* conjunctions	305	77.05%	60.97%	$< 1 \cdot 10^{-16}$
All *but* conjunctions	214	30.84%	25.94%	$2.09 \cdot 10^{-8}$
All attributive *and* conjunctions	1,077	80.04%	76.82%	$< 1 \cdot 10^{-16}$
All predicative *and* conjunctions	860	84.77%	84.54%	$< 1 \cdot 10^{-16}$
All appositive *and* conjunctions	30	70.00%	63.64%	0.04277

Table 1: Validation of our conjunction hypothesis. The P-value is the probability that similar or more extreme results would have been obtained if same- and different-orientation conjunction types were actually equally distributed.

Hatzivassiloglou & McKeown 1997

- Can you predict if two words are same/different orientation given their participation in these patterns?

Prediction method	Morphology used?	Accuracy on reported same-orientation links	Accuracy on reported different-orientation links	Overall accuracy
Always predict same orientation	No	77.84%	–	77.84%
	Yes	78.18%	97.06%	78.86%
But rule	No	81.81%	69.16%	80.82%
	Yes	82.20%	78.16%	81.75%
Log-linear model	No	81.53%	73.70%	80.97%
	Yes	82.00%	82.44%	82.05%

Table 2: Accuracy of several link prediction models.

Hatzivassiloglou & McKeown 1997

- Given learned *P(orientation(a)=orientation(b))* find ML cluster (heuristically). Mark the most bigger cluster as *positive.* Cross-validate based on #links in the graph.

α	Number of adjectives in test set ($\lvert A_\alpha \rvert$)	Number of links in test set ($\lvert L_\alpha \rvert$)	Average number of links for each adjective	Accuracy	Ratio of average group frequencies
2	730	2,568	7.04	78.08%	1.8699
3	516	2,159	8.37	82.56%	1.9235
4	369	1,742	9.44	87.26%	1.3486
5	236	1,238	10.49	92.37%	1.4040

Table 3: Evaluation of the adjective classification and labeling methods.

Hatzivassiloglou & McKeown 1997

- Later/related work:
  - LIWC, General Inquirer, other hand-built lexicons
  - Turney & Littman, TOIS 2003: Similar performance with 100M word corpus and PMI – higher accuracy better if you allow abstention on 25% of the "hard" cases.
  - Kamps et al, LREC 04: Determine orientation by graph analysis of Wordnet (distance to "good", "bad" in graph determined by synonymy relation)
  - SentiWordNet, Esuli and Sebastiani, LREC 06: Similar to Kamps et al, also using a BOW classifier and WordNet glosses (definitions).

First, some early & influential papers on opinion:

Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia, July 2002, pp. 417-424.

Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews

Peter D. Turney
Institute for Information Technology
National Research Council of Canada
Ottawa, Ontario, Canada, K1A 0R6
peter.turney@nrc.ca

Turney’s paper

- Goal: classify reviews as “positive” or “negative”.
  - Epinions “[not] recommended” as given by authors.
- Method:
  - Find (possibly) meaningful *phrases* from review (e.g., “bright display”, “inspiring lecture”, …)
    - (based on POS patterns, like ADJ NOUN)
  - Estimate “semantic orientation” of each candidate phrase
  - Assign overall orentation of review by averaging orentation of the phrases in the review

Semantic orientation (SO) of phrases

$$PMI(word_1, word_2) = \log_2\left[\frac{p(word_1 \,\&\, word_2)}{p(word_1)\,p(word_2)}\right]$$

$$\begin{aligned}\mathrm{SO}(\mathit{phrase}) = {} & \mathrm{PMI}(\mathit{phrase}, \text{“excellent”}) \\ & - \mathrm{PMI}(\mathit{phrase}, \text{“poor”})\end{aligned}$$

$$= \log_2\left(\frac{\text{hits}(\mathit{phrase}\ \text{NEAR “excellent”})\ \text{hits}(\text{“poor”})}{\text{hits}(\mathit{phrase}\ \text{NEAR “poor”})\ \text{hits}(\text{“excellent”})}\right)$$

Table 2. An example of the processing of a review that the author has classified as *recommended*.[6]

Extracted Phrase	Part-of-Speech Tags	Semantic Orientation
online experience	JJ NN	2.253
low fees	JJ NNS	0.333
local branch	JJ NN	0.421
small part	JJ NN	0.053
online service	JJ NN	2.780
printable version	JJ NN	-0.705
direct deposit	JJ NN	1.288
well other	RB JJ	0.237
inconveniently located	RB VBN	-1.541
other bank	JJ NN	-0.850
true service	JJ NN	-0.732
Average Semantic Orientation		0.322

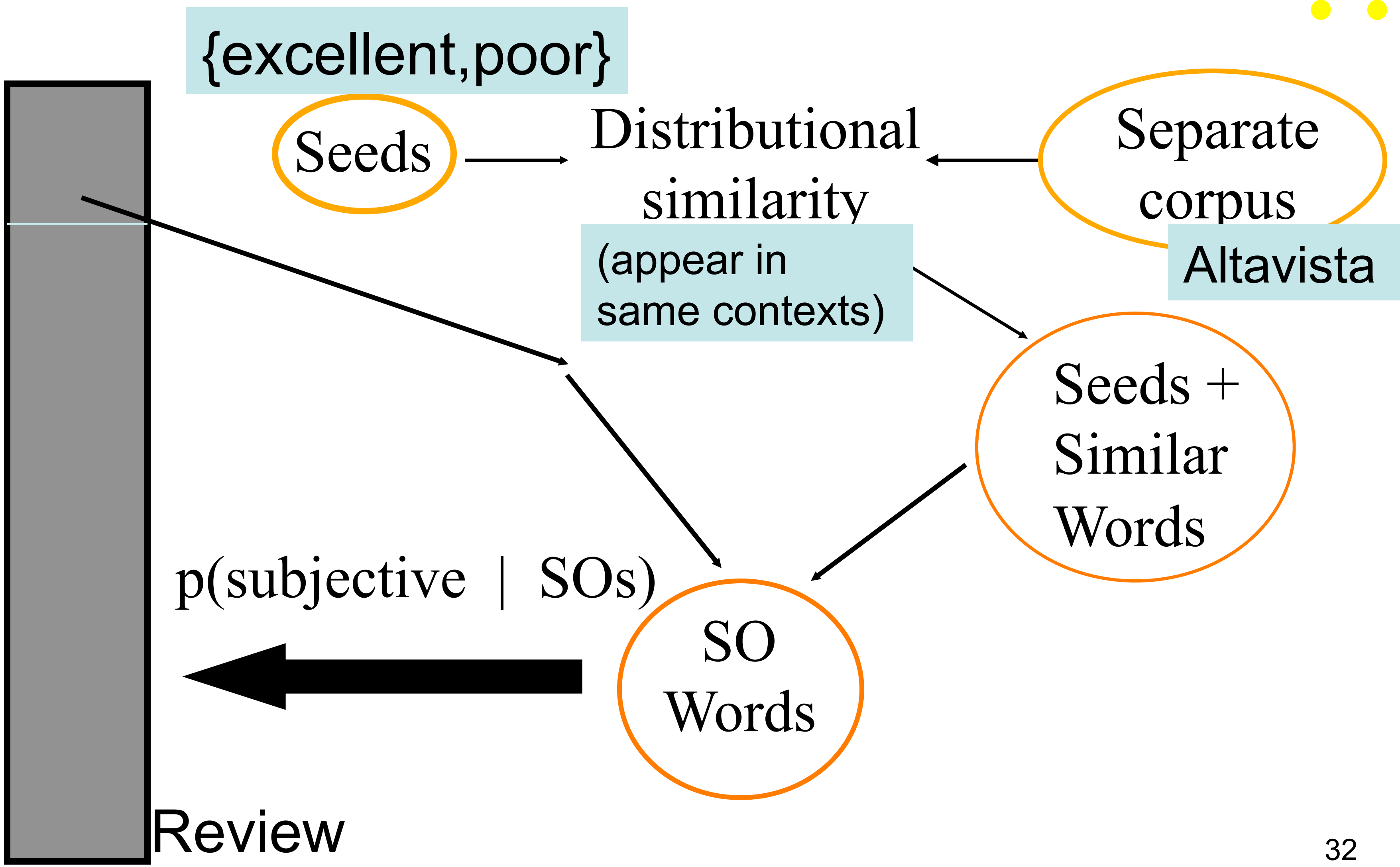
Table 3. An example of the processing of a review that the author has classified as *not recommended*.

Extracted Phrase	Part-of-Speech Tags	Semantic Orientation
little difference	JJ NN	-1.615
clever tricks	JJ NNS	-0.040
programs such	NNS JJ	0.117
possible moment	JJ NN	-0.668
unethical practices	JJ NNS	-8.484
low funds	JJ NNS	-6.843
old man	JJ NN	-2.566
other problems	JJ NNS	-2.748
probably wondering	RB VBG	-1.830
virtual monopoly	JJ NN	-2.050
other bank	JJ NN	-0.850
extra day	JJ NN	-0.286
direct deposits	JJ NNS	5.771
online web	JJ NN	1.936
cool thing	JJ NN	0.395
very handy	RB JJ	1.349
lesser evil	RBR JJ	-2.288
Average Semantic Orientation		-1.218

Table 5. The accuracy of the classification and the correlation of the semantic orientation with the star rating.

Domain of Review	Accuracy	Correlation
Automobiles	84.00 %	0.4618
Honda Accord	83.78 %	0.2721
Volkswagen Jetta	84.21 %	0.6299
Banks	80.00 %	0.6167
Bank of America	78.33 %	0.6423
Washington Mutual	81.67 %	0.5896
Movies	65.83 %	0.3608
The Matrix	66.67 %	0.3811
Pearl Harbor	65.00 %	0.2907
Travel Destinations	70.53 %	0.4155
Cancun	64.41 %	0.4194
Puerto Vallarta	80.56 %	0.1447
All	74.39 %	0.5174

William's picture of Jan's picture of this paper…



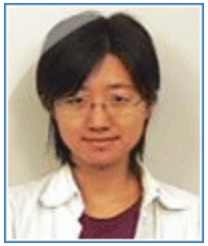
Key ideas in Turney 2002

- Simplification:
  - classify an entire document, not a piece of it. (Many reviews are mixed.)
- Focus on what seems important:
  - Extract semantically oriented words/phrases from the document. (Phrases are less ambiguous than words – eg “Even poor students will learn a lot from this lecture”).
- Bootstrapping/semi-supervised learning:
  - To assess orientation of phrases, use some kind of contextual similarity of phrases

Issues

- Is polarity context-independent?
  - "Unpredictable plot" *vs* "Unpredictable steering"
  - "Read the book"
- Turney's solution:
  - Use "unpredictable plot" as feature, not "unpredictable".
- Other ideas?
  - … ?

Pang et al EMNLP 2002

Thumbs up? Sentiment Classification using Machine Learning Techniques

Bo Pang and **Lillian Lee**
Department of Computer Science
Cornell University
Ithaca, NY 14853 USA
{pabo,llee}@cs.cornell.edu

Shivakumar Vaithyanathan
IBM Almaden Research Center
650 Harry Rd.
San Jose, CA 95120 USA
shiv@almaden.ibm.com

Abstract

We consider the problem of classifying documents not by topic, but by overall senti-

which both labels movie reviews that do not contain explicit rating indicators and normalizes the different rating schemes that individual reviewers use. Sentiment classification would also be helpful in

Methods

- *Movie review* classification as pos/neg.
- Method one: count human-provided polar words (sort of like Turney):
  - Eg, "*love, wonderful, best, great, superb, still, beautiful*" vs "*bad, worst, stupid, waste, boring, ?, !*" *gives* 69% accuracy on 700+/700- movie reviews
- Method two: plain 'ol text classification
  - Eg, Naïve Bayes bag of words: 78.7; SVM-lite "set of words": 82.9 was best result
  - Adding bigrams and/or POS tags doesn't change things much.

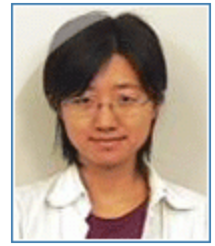
Results

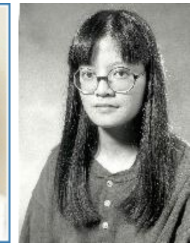
	Features	# of features	frequency or presence?	NB	ME	SVM
(1)	unigrams	16165	freq.	**78.7**	N/A	72.8
(2)	unigrams	”	pres.	81.0	80.4	**82.9**
(3)	unigrams+bigrams	32330	pres.	80.6	80.8	**82.7**
(4)	bigrams	16165	pres.	77.3	**77.4**	77.1
(5)	unigrams+POS	16695	pres.	81.5	80.4	**81.9**
(6)	adjectives	2633	pres.	77.0	**77.7**	75.1
(7)	top 2633 unigrams	2633	pres.	80.3	81.0	**81.4**
(8)	unigrams+position	22430	pres.	81.0	80.1	**81.6**

Intuitions

- Objective descriptions of what *happens* in the movie vs the review author's *opinion about it* are confusing things?
- Start and end of the review have most of the information?
- ---?

Pang & Lee EMNLP 2004





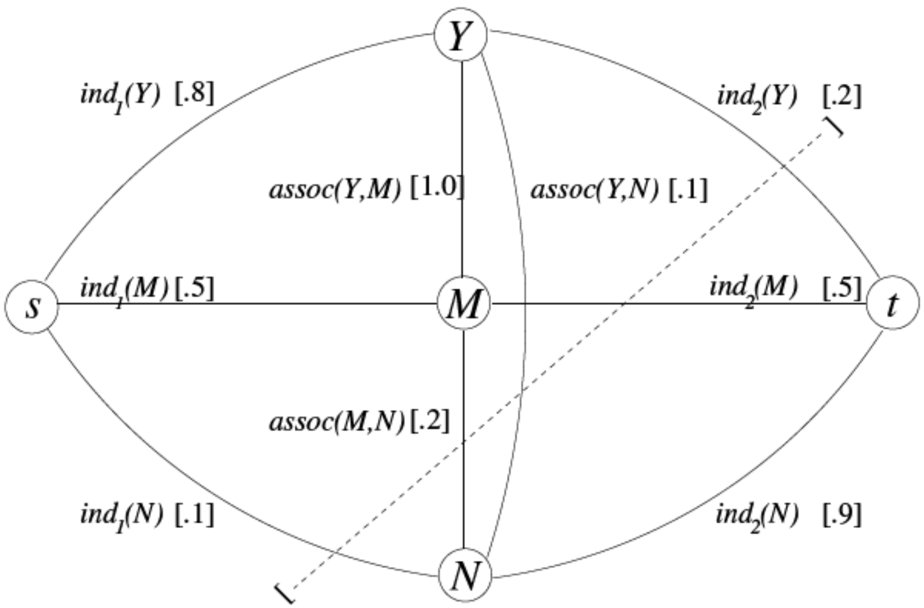
A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts

Bo Pang and **Lillian Lee**
Department of Computer Science
Cornell University
Ithaca, NY 14853-7501
{pabo,llee}@cs.cornell.edu

Pang & Lee EMNLP 2004

- Can you capture the discourse in the document?
  - Expect longish runs of *subjective* text and longish runs of *objective* text.
  - Can you tell which is which?
- Idea:
  - Classify sentences as subjective/objective, based on two corpora: short biased reviews, and IMDB plot summaries.
  - Smooth classifications to promote longish homogeneous sections.
  - Classify polarity based on the *K* "most subjective" sentences

Pang & Lee EMNLP 2004



C_1	Individual penalties	Association penalties	Cost
{Y,M}	$.2 + .5 + .1$	$.1 + .2$	1.1
(none)	$.8 + .5 + .1$	0	1.4
{Y,M,N}	$.2 + .5 + .9$	0	1.6
{Y}	$.2 + .5 + .1$	$1.0 + .1$	1.9
{N}	$.8 + .5 + .9$	$.1 + .2$	2.5
{M}	$.8 + .5 + .1$	$1.0 + .2$	2.6
{Y,N}	$.2 + .5 + .9$	$1.0 + .2$	2.8
{M,N}	$.8 + .5 + .9$	$1.0 + .1$	3.3

Figure 2: Graph for classifying three items. Brackets enclose example values; here, the individual scores happen to be probabilities. Based on *individual* scores alone, we would put Y ("yes") in C_1, N ("no") in C_2, and be undecided about M ("maybe"). But the *association* scores favor cuts that put Y and M in the same class, as shown in the table. Thus, the minimum cut, indicated by the dashed line, places M together with Y in C_1.

Pang & Lee EMNLP 2004

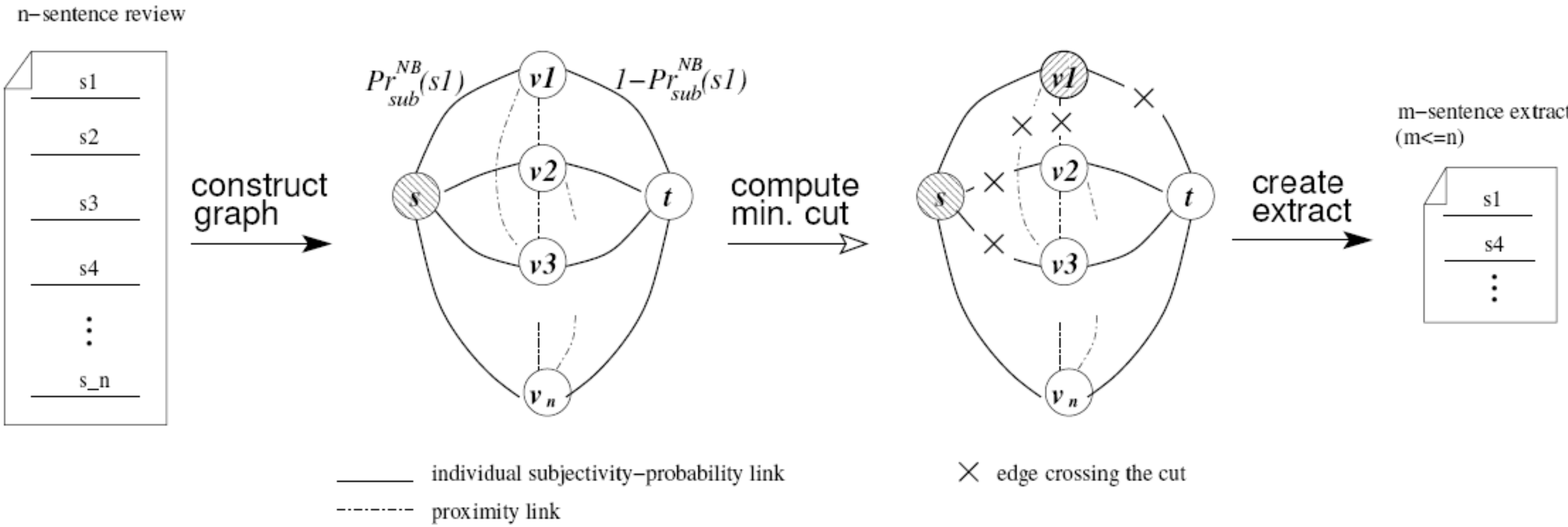
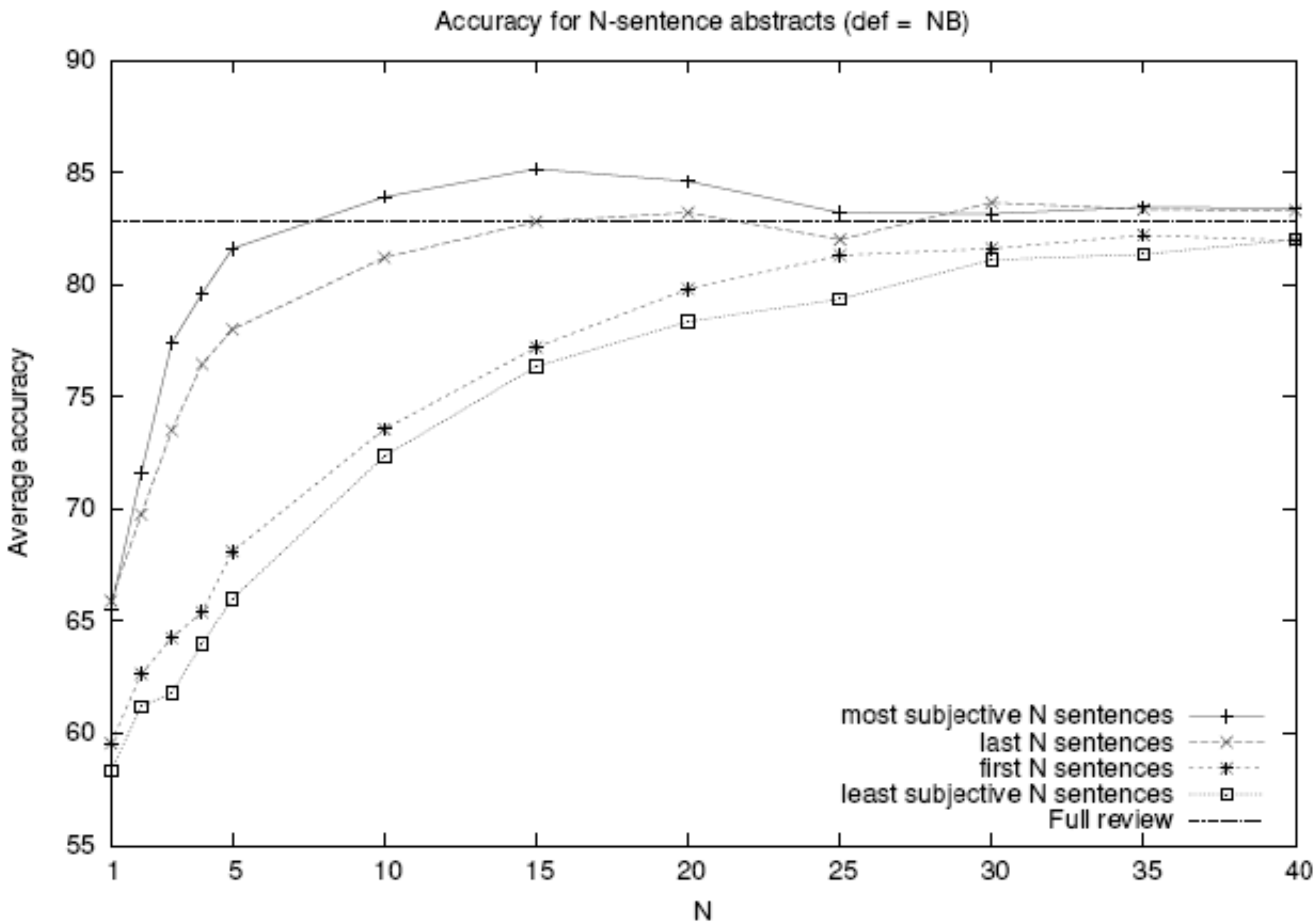
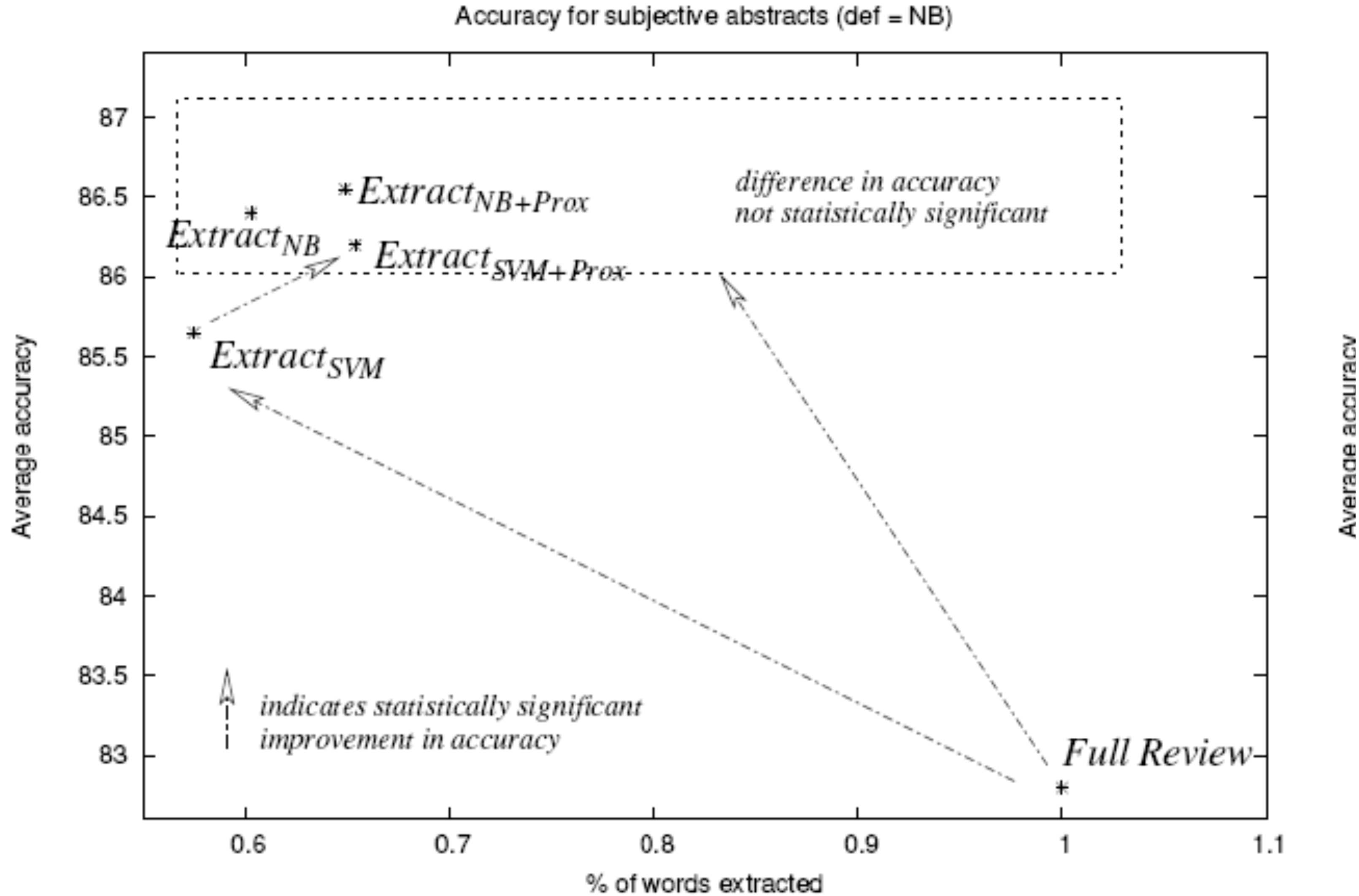


Figure 3: Graph-cut-based creation of subjective extracts.



Accuracy for N-sentence abstracts (def = NB)
Average accuracy
N
90
85
80
75
70
65
60
55
1
5
10
15
20
25
30
35
40
most subjective N sentences
last N sentences
first N sentences
least subjective N sentences
Full review


Accuracy for subjective abstracts (def = NB)
Average accuracy
87
86.5
86
85.5
85
84.5
84
83.5
83
0.6
0.7
0.8
0.9
1
1.1
% of words extracted
Extract$_{NB+Prox}$
Extract$_{NB}$
Extract$_{SVM+Prox}$
Extract$_{SVM}$
difference in accuracy
not statistically significant
indicates statistically significant
improvement in accuracy
Full Review
Average accuracy