Do you feel what I feel?

The Social Aspects of Emotions in Twitter Conversations

Suin Kim, JinYeong Bak, Yohan Jo and Alice Oh
{suin.kim, jy.bak, yohan.jo}@kaist.ac.kr, alice.oh@kaist.edu
Users & Information Lab., KAIST

Dec 5, 2011

Submitted to CHI 2012, accepted at NIPS 2011 CSS workshop

Index

- Introduction and Background
- Topic modeling 101
- Twitter data collection
- Social aspects of sentiments and emotions
- Sentiment patterns in conversations
- Conclusion & future work

Our Work

- A computational framework for analyzing the social aspects of sentiments and emotions
- Research Question: Is there a regular pattern of sentiment and emotion transitions in Twitter conversations?
- Research approach
  - Emotion discovery using a probabilistic topic model
  - Analysis of sentiment and emotion transition pattern
  - Identification of major topics that cause emotion influence

Background:
Sentiment, Emotion, and Mood

- Sentiment, emotion, and mood have different meaning in NLP field
- **Sentiment**: "Attitude toward something"
  - Positive / Neutral / Negative
- **Emotion**: "Human feeling"
  - Plutchik's wheel of emotion: Joy-Sadness, Trust-Disgust, Fear-Anger, ...
  - Parrott's tree-structured list of emotions: Joy, Love, Anger, Surprise, Sadness, Fear
- **Mood**: "A temporary state of mind"
  - Profile of mood states (POMS): Tension-Anxiety, Depression-Dejection, ...

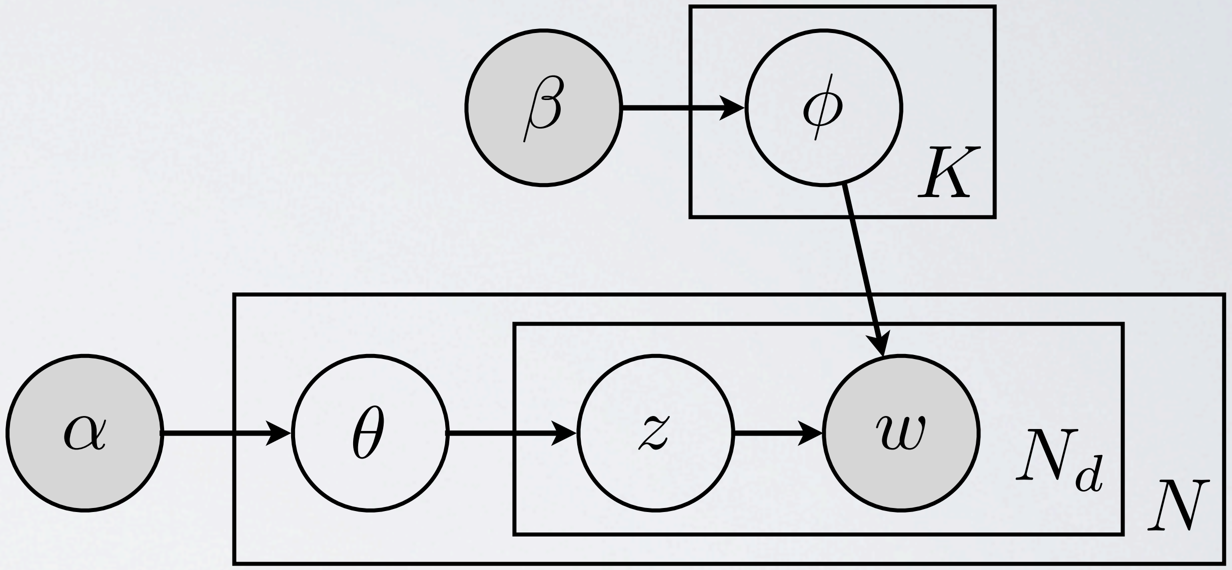
Background: Prior Work

- Twitter has become a good resource as a social media
  - (D-N-Mizil, 2011) studied aspects of human conversations from Twitter
  - (Bollen, 2011) performed sentiment analysis on Twitter using POMS
  - (Ritter, 2010) presented unsupervised model of Twitter conversation
- Prior emotion researches has mostly used naive approach
  - (Kamvar, 2011) used regular expression to discover emotion
  - (Bollen, 2011) used POMS scoring to discover mood
- Ours is the first work to look at emotions in Twitter conversation

Topic Modeling 101: LDA

- Generative process of LDA

> 1. $\phi_k \sim \text{Dirichlet}(\beta)$.
> For each document d,
> 2. $\theta_d \sim \text{Dirichlet}(\alpha)$.
>   For each word w_{di} in document d,
>   2.1. $z_{di} \sim \text{Multinomial}(\theta_d)$.
>   2.2. $w_{di} \sim \text{Multinomial}(\phi_{z_{di}})$.



- We infer latent variables (z, θ, φ) which maximize the model probability
- Probability of generating word w from topic k: $P(w|\varphi_k)$
- In ASUM, one tweet is constrained to be generated from one topic
  Topics related to seed words are found

Topic Modeling 101: ASUM

- A topic is represented as a pmf over the vocabulary
- We use aspect and sentiment unification model (ASUM, Jo 2011)
  - Extension of the latent Dirichlet allocation (LDA)

Sentiment seed words

good, nice, correct, excellent, positive, fortunate, correct, superior, amazing, attractive, best, ...	bad, nasty, poor, negative, inferior, hate, unfortunate, annoying, complain, disappointed, ...

amazon.com

Reviews on electronic devices

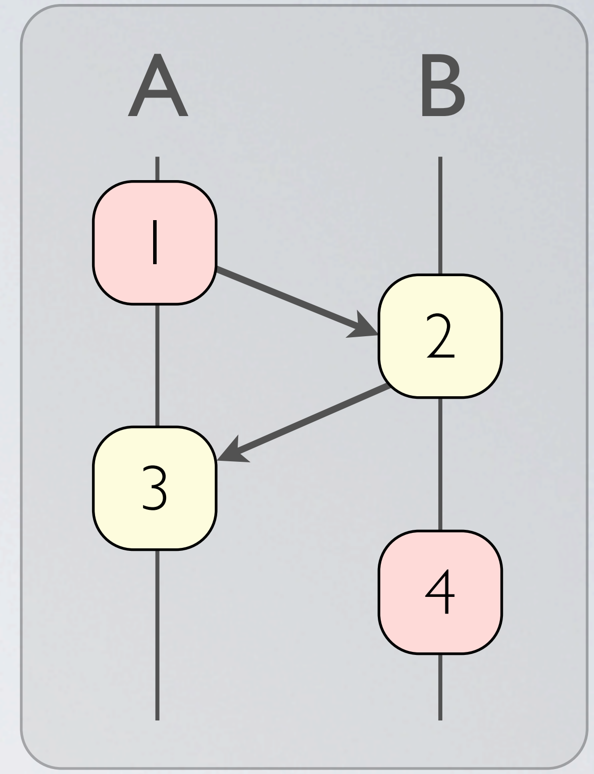
ASUM

Positive	worth money penny extra well every price dollar	screen color bright clear video display crisp great	easy light carry weight lightweight suction small around
Negative	money save waste not_buy away spend not_worth stay	fingerprint glossy magnet screen show finger finish print	screen font point size notebook ssd key shoot

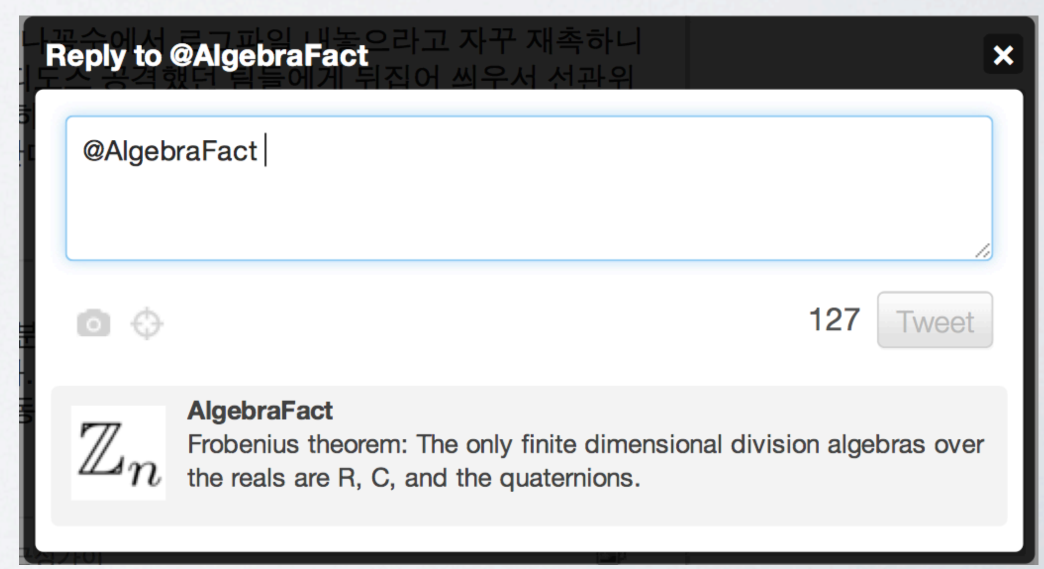
Topic Modeling 101: ASUM

- Input to ASUM
  - Seed words for each sentiment
  - Number of topics to be discovered for each sentiment
  - Dirichlet hyperparameters
- Output of ASUM
  - Discovered topics for each sentiment
  - P(topic|tweet)
  - Topic and sentiment distribution over conversation

Data Collection



- We call a Twitter conversation a ***chain***
  - ***chain***: A sequence of replies between two users
  - We only consider users replying to each other's tweets using the ***reply*** button
- Reply button: Low recall but high precision of finding real conversations



Data Collection

Twitter Gardenhose
Random sample of all public tweets (~1%)

Candidates
List of users as a candidate of dyads

Collect
Public tweets from candidate list

Identify
Conversational partners

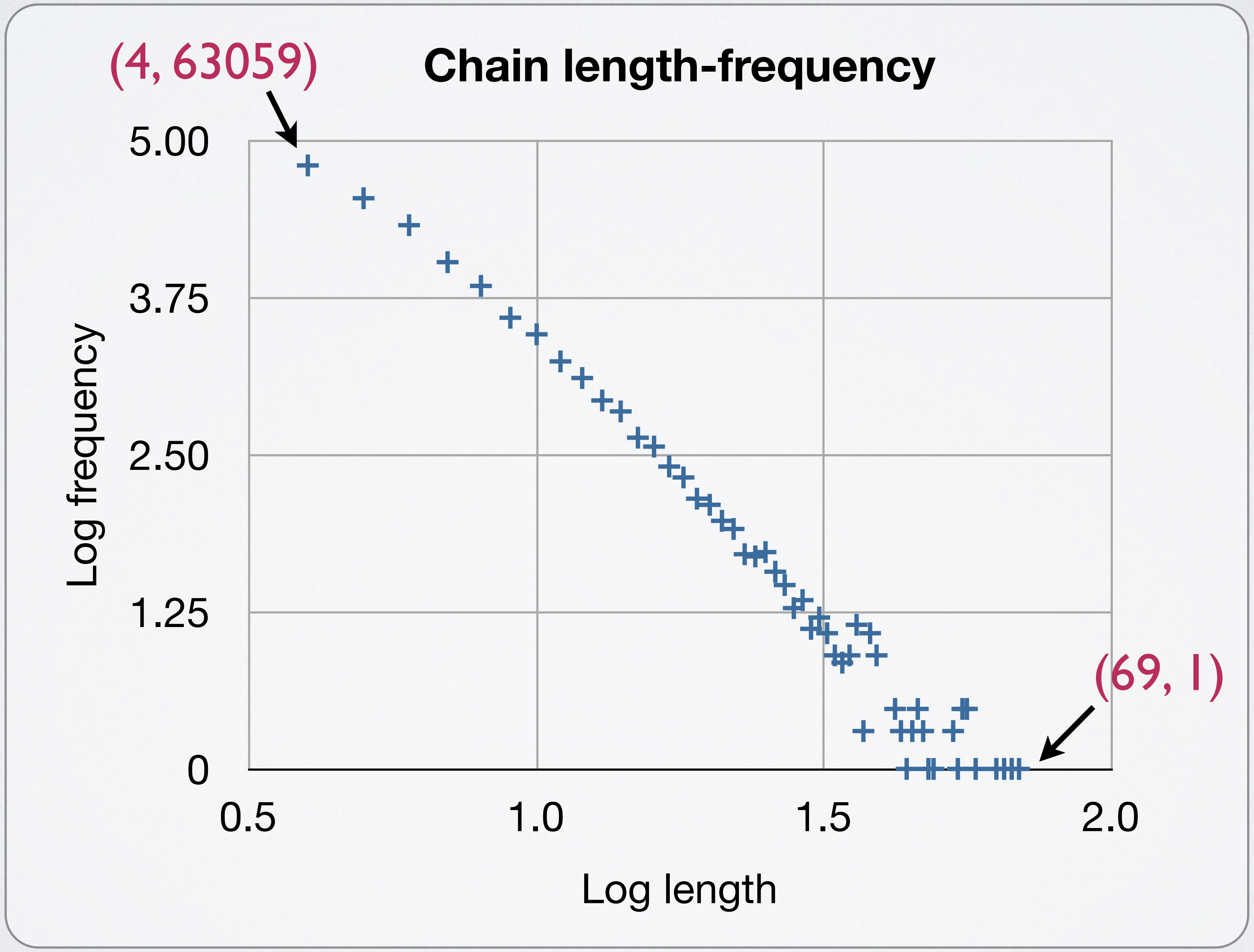
Data Collection

- Filter out tweets from non-English speakers
- Keep only chains of length >=4
- Data statistics

	#Users	#Dyads	#Chains	#Tweets	Tweets per chain
Gathered	136,730	222,024	1,668,308	4,464,456	2.68
Filtered	**62,952**	**77,850**	**153,054**	**871,544**	**5.69**

Data Collection

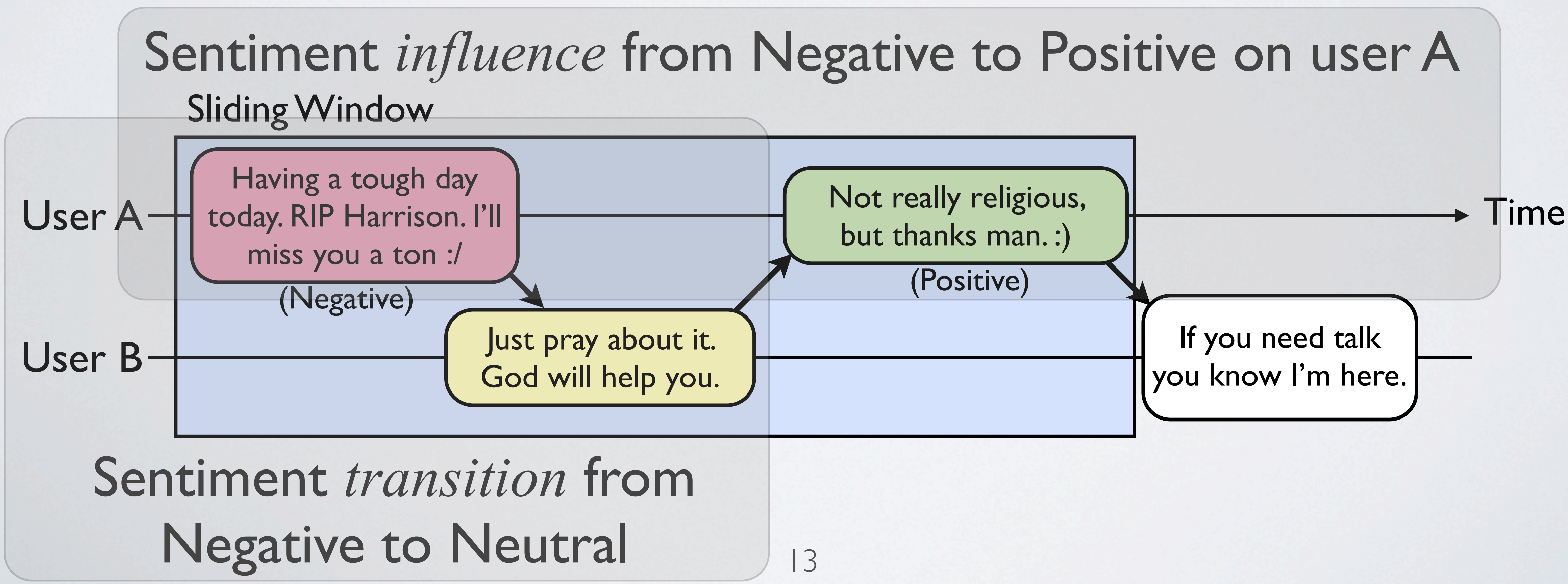
- Chain length follows the Power Law



Log base is 10

Social Aspects of Sentiments

- Questions:
  - How sentiments in a tweet lead to the sentiments in the reply? (Sentiment *transition*)
  - How sentiments in a tweet influence the partner's sentiment? (Sentiment *influence*)

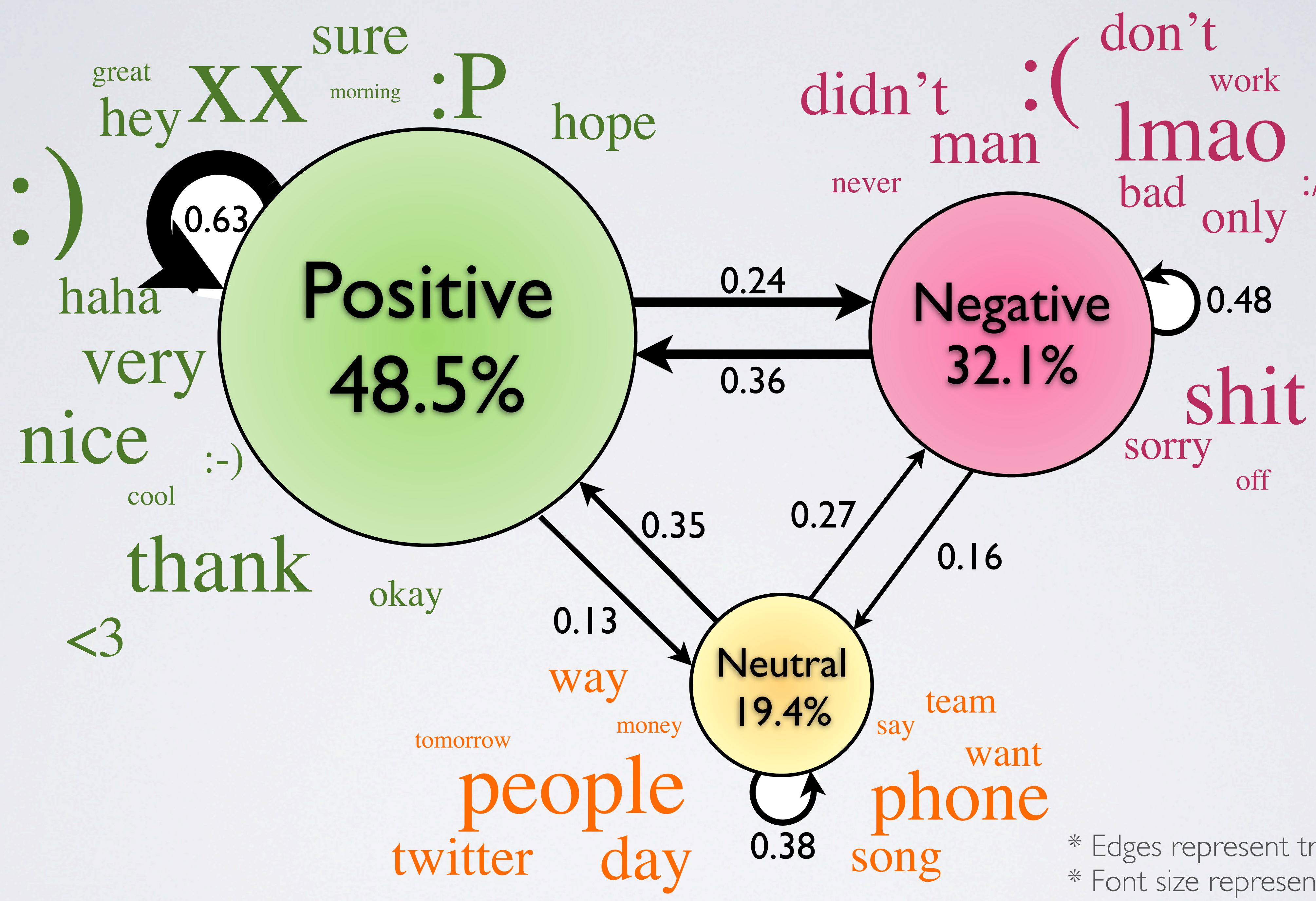


Social Aspects of Sentiments

- We analyze our corpus with 3 sentiments, for 30 topics each
  - Positive, Negative, Neutral
- Emoticons as a seed words
  - (Read 2005) (Pak 2010) used emoticons as a sentiment label

Positive	`:) :D XD (: :-) =) :)) :] =D =] :-D 8) :p ...`
Negative	`:( :/ ): :'( :S =/ :-( =( :-/ :(( ]: :\ /: :$ :I :C ...`
Neutral	No seed words

Sentiment Transition



* Edges represent transition probability
* Font size represents appearance freq.

- Self-transitions make up the largest proportion: "*sentiment accommodation*"
- Transitions to the positive sentiments are high

Sentiment Influence

Positive → Positive		Positive → Negative		Negative → Positive		Negative → Negative	
Topic 16	**Topic 19**	**Topic 17**	**Topic 38**	**Topic 44**	**Topic 47**	**Topic 72**	**Topic 42**
follow	heart	lmao	:(	hope	today	shit	off
thank	xD	ur	school	better	week	bro	still
please	big	ass	exam	sorry	still	da	bad
followed	kisses	lol	tomorrow	feel	feel	chillin	feel
welcome	much	shit	year	soon	work	nigga	down
Topic 8	**Topic 21**	**Topic 48**	**Topic 54**	**Topic 59**	**Topic 37**	**Topic 49**	**Topic 32**
thank	morning	weather	sleep	home	money	game	ur
:D	day	rain	work	want	buy	lmao	loool
much	hope	:(	still	house	want	shit	lool
very	hello	hot	awake	sleep	had	team	looool
welcome	happy	cold	tired	bed	pay	win	man

Greeting	Teasing Complaint	Sympathy	Swear Mocking

Social Aspects of Emotions

- (Parrott, 2001) described a tree-structured list of emotions:

Primary	Secondary	Tertiary
Love	Affection, Lust, Longing	Adoration, Fondness, Liking, Arousal, Desire, Infatuation, Longing, Compassion, Caring, Compassion
Joy	Cheerfulness, Zest, Contentment, Pride, Optimism, Enthrallment, Relief	Amusement, Bliss, Glee, Jolliness, Zeal, Thrill, Pleasure, Triumph, Hope, Rapture, Relief, Eagerness, Pleasure
Surprise	Surprise	Amazement, Astonishment
Anger	Irritability, Exasperation, Rage, Disgust, Envy, Torment	Aggravation, Agitation, Annoyance, Grouchy, Grumpy, Crosspatch, Frustration, Fury, Wrath, Contempt, Anger
Sadness	Suffering, Sadness, Disappointment, Shame, Neglect, Sympathy	Agony, Anguish, Depression, Dismay, Guilt, Alienation, Pity, Defeatism, Remorse, Displeasure, Gloom, Grief
Fear	Horror, Nervousness	Alarm, Shock, Fear, Horror, Terror, Panic, Hysteria, Anxiety, Suspense, Uneasiness, Worry, Distress, Dread

Social Aspects of Emotions

- Finding emotions from tweets

  1. Run ASUM to discover topics from Twitter conversation data

  2. Expand the emotion lexicons: 139 → 702 words

  3. Find the topics specialize in emotion using emotion lexicons
     "Topic 23 is specialized in *Love* emotion"

  4. Classify tweets from the topic
     "This tweet is generated from topic 23, thus it contains *Love* emotion"

Social Aspects of Emotions

- *Corr* measures correlation between emotion and topic
  - Probability of generating the set of seed words ***w*** from emotion *c*:

$$Corr(c,t) = \frac{\gamma_c}{n_c} \sum_{i=1}^{n_c} P(w_i|\phi_t)$$

- *Spec* measures specialization of a certain emotion for topic t

$$Spec(t) = \frac{\max_c Corr(c,t)}{\sum_c Corr(c,t) - \max_c Corr(c,t)}$$

- Select topics with high *Spec* score

Joy

Topic 31
day
hope
morning
great
weekend

Topic 23
party
home
fun
tomorrow

Topic 40
happy
birthday
th
hahaha
year

Topic 46
thank
much
very
hope

Sadness

Topic 99
hope
sorry
better
feel
soon
aww
okay
hugs
feeling
happened

Topic 86
:'(
miss
him
sorry
want
sad
wish
wanna
much
cry

Love

Topic 1
her
smiles
laughs
eyes
want

Topic 47
school
luck
exam
tomorrow

Topic 3
twitter
justin
follow
selena
fan

Topic 9
pic
hair
look
nice

Fear

Topic 60
:/
school
tomorrow
exam
year
doing
much
day
next
maths

Topic 87
dont
lmao
tell
want
him
her
please
face
who
omg

Anger

Topic 95
eat
want
food
chicken
hungry

Topic 73
she
hate
being
never

Topic 59
she
:(
car
were
off
home
her
phone
house
night

Surprise

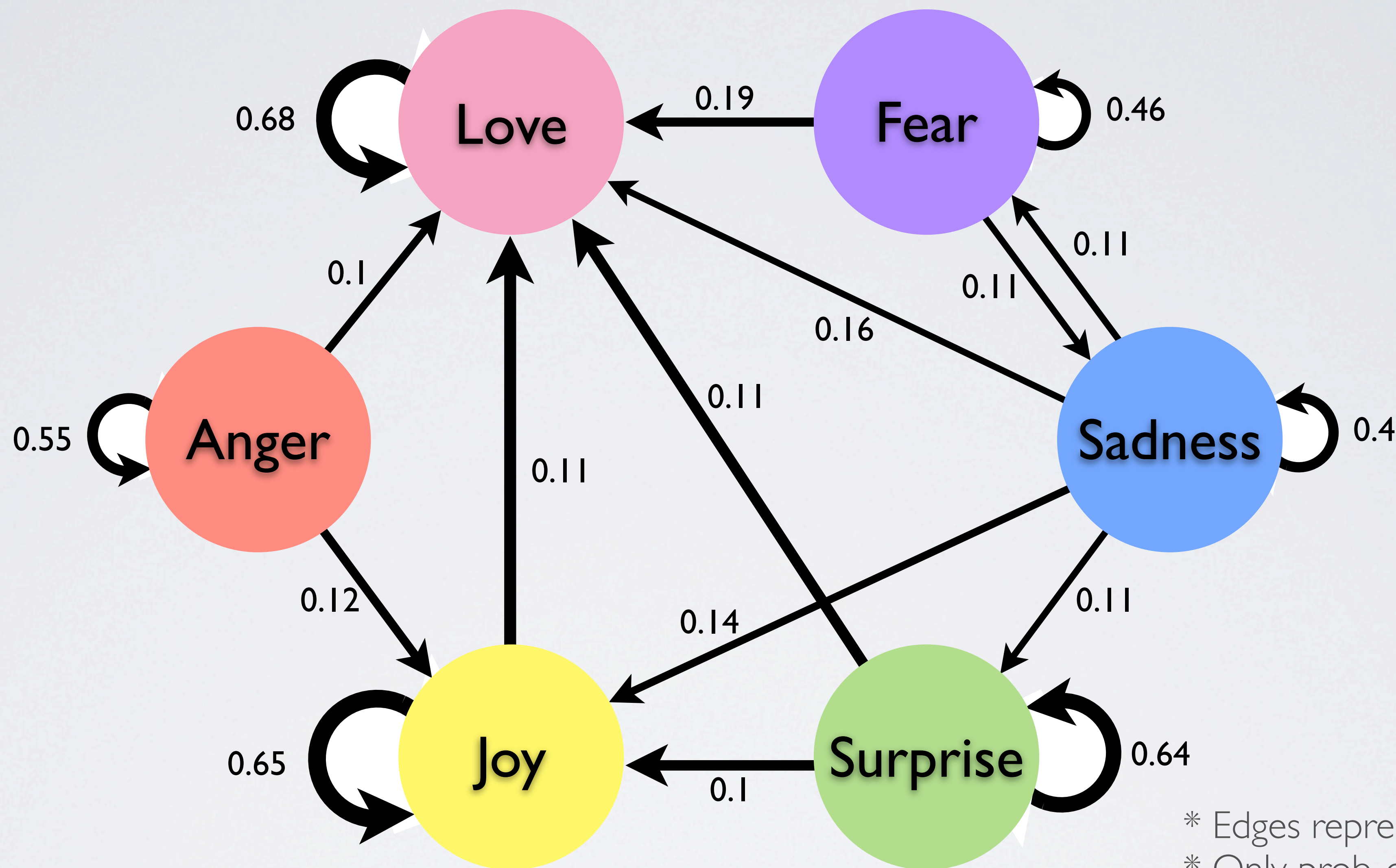
Topic 18
song
album
music
new
her
she
amazing
listening
awesome
great

Topic 30
where
live
here
same
awesome

Topic 68
she
old
than
look

Discovered topics with high *Spec* score

Emotion Transition



* Edges represent transition probability
* Only prob. of >= 0.1 are shown

- Self-transitions make up the largest proportion
- Transitions to positive emotions are high

Emotion Influence

Anger → Joy		Sadness → Joy		Surprise	
				→Love	→Anger
Topic 23	**Topic 17**	**Topic 81**	**Topic 99**	**Topic 18**	**Topic 79**
day	eat	better	hope	song	song
night	food	feel	sorry	album	shit
party	cream	hope	better	amazing	new
home	make	pain	feel	listening	lmao
fun	chocolate	sick	soon	great	never
come	drink	:(	hugs		

Consolation
Suggest

Sympathy

- Sympathy tends to change partner's sentiment from negative to positive
- Similar subject but in different sentiment change partner's sentiment differently

Sentiment Patterns in Conversations

- Interlocutors share a common sentiment in most conversations
  - What if two interlocutors have opposing sentiments?
  - What topic influence the overall sentiments of conversation?
  - How can we find such topics and conversations?

Sentiment Patterns in Conversations

- The overall sentiment of an interlocutor u in conversation v:
  - $Senti(u, v) = \frac{p_{u,v} - n_{u,v}}{p_{u,v} + n_{u,v}}$
  - $p_{u,v}$: # positive tweets by u in v, $n_{u,v}$: # negative tweets
- Only conversations with both $|Senti(u_1, v)|$ and $|Senti(u_2, v)|$ greater than 0.5 are considered
- 4% of the corpus showed Pos-Neg pattern

Sentiment Patterns in Conversations

- Emotions in sentiment-opposing conversations
  - Feeling upset - Sympathy, relieve
  - Complaining - Apology
- An example of complaint-apology conversation

A: Sorry to hear about your bags. If you would like us to get someone to contact you DM us your reference and contact number.

B: it's on it's way to manch. If the woman on the check in desk in Miami hadn't been trying to be all smart! Been no problem.

A: Sorry about that. Pleased to hear they located it quickly for you though.

B: mistakes happen.

▪Apology ▪Complaint

Conclusion

- A novel computational framework for analyzing the social aspects of sentiments and emotions
- Answer to ***Do you feel what I feel?*** is ***yes***:
  - Self-transitions account for the largest proportions
- Twitter users tend to feel good on the most of other cases
- *Sympathy*, *apology*, and *complaining* influence emotion

Future Work

- Modeling social behavior of people
  - Comparisons to other media
  - A joint probabilistic model for sentiment-emotion discovery
- Improving Human-Computer Interaction scenarios
  - Computer can understand the sentiment and emotion from text

References

1. Blei, D., Ng, A., and Jordan, M. Latent dirichlet allocation. The Journal of Machine Learning Research 3 (2003), 993–1022.

2. Jo, Y., and Oh, A. Aspect and sentiment unification model for online review analysis. In Proceedings of WSDM (2011).

3. Parrott, W. Emotions in social psychology: essential readings. Psychology Pr, 2001.

4. Plutchik, R. Emotion: A psychoevolutionary synthesis. Harper & Row New York, 1980.

5. Using emoticons to reduce dependency in machine learning techniques for sentiment classification, J Read, 2005.

6. Kamvar, S., and Harris, J. We feel fine and searching the emotional web. In Proceedings of WSDM, ACM (2011), 117–126.

7. Danescu-Niculescu-Mizil, C., Gamon, M., and Dumais, S. Mark my words!: linguistic style accommodation in social media. In Proceedings of WWW (2011).

8. Twitter as a corpus for sentiment analysis and opinion mining, Pak and Paroubek, 2010.

Q&A

Back-up slides

The Social Aspects of Emotions

- Finding emotion in a short text is not a trivial task, also even it is difficult for human.

- Results from Amazon Mechanical Turk:

  - We requested 1,500 tweets to be classified into
    1) 3 sentiments: Positive, Neutral, Negative
    2) 7 Emotions: Love, Joy, Surprise, Anger, Sadness, Fear, not emotional
    by 3 distinct workers for each tweet.

	2/3 agreement	Full agreement
Sentiment	72%	19%
Emotion	91.3%	35.7%

*Proportion.