Learning Representations of Social Meaning

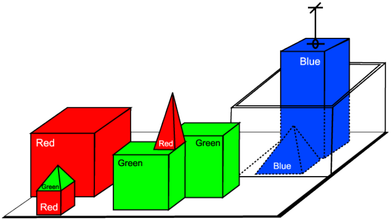
Jacob Eisenstein
@jacobeisenstein

Georgia Institute of Technology

August 3, 2017

What does social meaning mean?

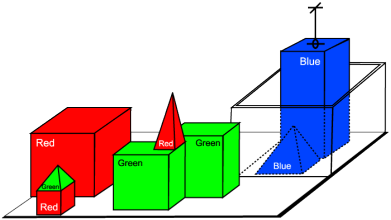
Pick up the big red block.



What does social meaning mean?

Pick up the big red block.

- Would you please pick up the big red block?
- That red block needs picked up.
- Hurry up and grab that block.
- Could you, like, pick up the red block?



What is social meaning about?

Idealized model: saying the "same thing" differently, depending on (and conveying information about):

- the speaker/writer

  dialect, gender, race, "latent user attributes", ...

- the communicative situation

  register, formality, politeness, accommodation, ...

- attitude towards the talk and topic

  sentiment, subjectivity, irony, ...

The department store study

- **Linguistic variable**: (r) in fourth floor
- **Social variable**: class (department stores Klein's, Macy's, and Sak's)
- **Situational variable**: feigned misunderstanding

(Labov, 1972)

The department store study

Use of local New York "r-less" variable decreases with

- socioeconomic status
- emphasis.

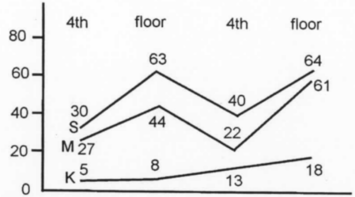


Figure 13.2: Percentage of all (r – 1) by store for four positions (S = Saks, M = Macy's, K = Kleins)

The department store study

Use of local New York "r-less" variable decreases with

- socioeconomic status
- emphasis.

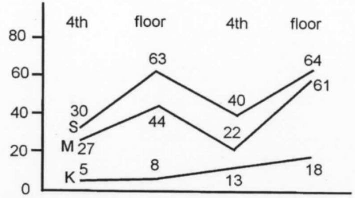


Figure 13.2: Percentage of all (r – 1) by store for four positions (S = Saks, M = Macy's, K = Kleins)

Geographical, socioeconomic, and stylistic variation are all linked!

Disentangling social variables

How to assess the impact of multiple social variables?

- **Linguistic variable**: on versus other pronouns in Montreal French
- **Social variables**: age, gender, education, socioeconomic status
- Model:

$$\log \frac{p}{1-p} = \beta_0 x_0 + \beta_1 x_1 + \ldots + \beta_n x_n \qquad (1)$$

Disentangling social variables

How to assess the impact of multiple social variables?

- **Linguistic variable**: on versus other pronouns in Montreal French
- **Social variables**: age, gender, education, socioeconomic status
- Model:

$$\log \frac{p}{1-p} = \beta_0 x_0 + \beta_1 x_1 + \ldots + \beta_n x_n \qquad (1)$$

(Sankoff & Labov, 1979)

Computational models of social meaning

- What is the role of computation in understanding social meaning in language?
- Why do we need representation learning?
- Is social meaning relevant to "core tasks" in NLP?

What is social meaning about?

Idealized model: saying the "same thing" differently, depending on (and conveying information about):

- the speaker/writer

  dialect, gender, race, "latent user attributes", ...

- the communicative situation

  register, formality, politeness, accommodation, ...

- attitude towards the talk and topic

  sentiment, subjectivity, irony, ...

What is social meaning about?

Idealized model: saying the "same thing" differently, depending on (and conveying information about):

- **the speaker/writer**

  dialect, gender, race, "latent user attributes", . . .

- the communicative situation

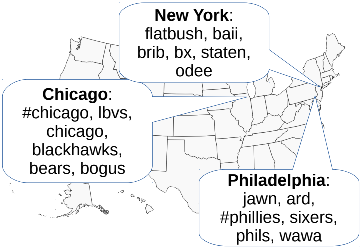
  register, formality, politeness, accommodation, . . .

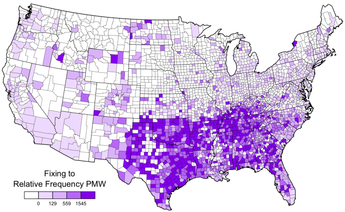
- attitude towards the talk and topic

  sentiment, subjectivity, irony, . . .

Social meaning from social media

Large-scale social media data has sparked a renaissance in geographical dialectology.



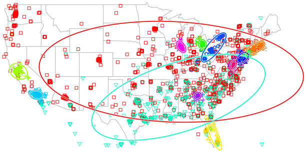


(Eisenstein et al., 2014; Rahimi et al., 2017)

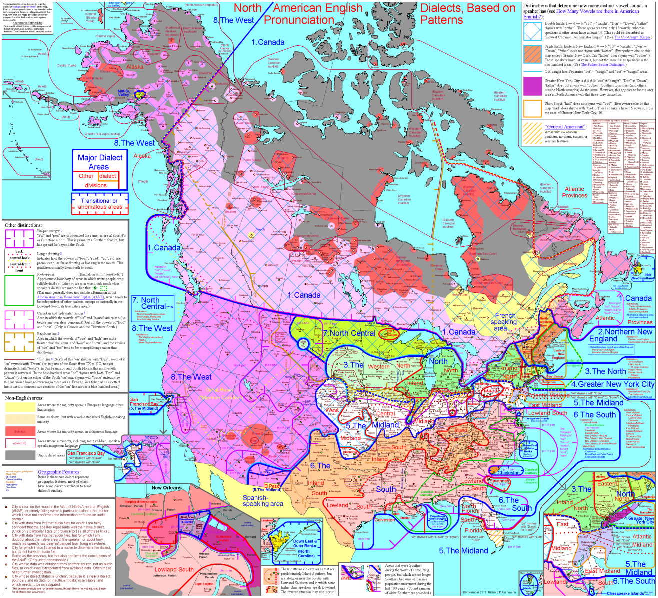
(Grieve et al., 2017)

Representations for geographical dialects

- Language variation generally does not align with geo-political units like cities, states, and nations (Heeringa & Nerbonne, 2001).
- Representation learning for dialect regions:
  - Gaussians (Eisenstein et al., 2010)
  - kd-trees (Roller et al., 2012)

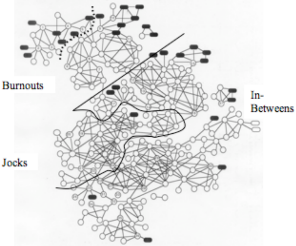


The real picture?



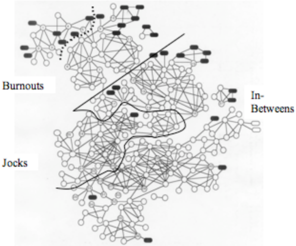
Social variables

- "Demographic" variables: race, gender, age (e.g., Argamon et al., 2007; Eisenstein et al., 2011; Nguyen et al., 2013)
- Eckert (2000): individuals have agency in creating social meanings! Sociolinguistics should focus on **locally-defined** social categories.



Social variables

- "Demographic" variables: race, gender, age (e.g., Argamon et al., 2007; Eisenstein et al., 2011; Nguyen et al., 2013)
- Eckert (2000): individuals have agency in creating social meanings! Sociolinguistics should focus on **locally-defined** social categories.



Can representation learning discover locally-defined social variables from text and metadata?

Gender identity and lexical variation in social media (Bamman, Eisenstein, and Schnoebelen, 2014)


- How to model gendered language differences without gender binary?

Literally the simplest possible thing

K-means clustering on Twitter timelines

hubs blogged recipe fabric
kidd hubs xo =]
wyd #oomf lmbo shyt
n_ggas wyd finna shyt
#nhl #bruins #mlb knicks

Literally the simplest possible thing

K-means clustering on Twitter timelines

	% Women
hubs blogged recipe fabric	90%
kidd hubs xo =]	80%
wyd #oomf lmbo shyt	60%
n_ggas wyd finna shyt	26%
#nhl #bruins #mlb knicks	11%

- Do women use more standard language?
- Is men's writing more "informational"?

Literally the simplest possible thing

K-means clustering on Twitter timelines

	% Women
hubs blogged recipe fabric	90%
kidd hubs xo =]	80%
wyd #oomf lmbo shyt	60%
n_ggas wyd finna shyt	26%
#nhl #bruins #mlb knicks	11%

- **Do women use more standard language?**
- Is men's writing more "informational"?

Literally the simplest possible thing

K-means clustering on Twitter timelines

	% Women
hubs blogged recipe fabric	90%
kidd hubs xo =]	80%
wyd #oomf lmbo shyt	60%
n_ggas wyd finna shyt	26%
#nhl #bruins #mlb knicks	11%

- **Do women use more standard language?**
- Is men's writing more "informational"?

Literally the simplest possible thing

K-means clustering on Twitter timelines

	% Women
hubs blogged recipe fabric	90%
kidd hubs xo =]	80%
wyd #oomf lmbo shyt	60%
n_ggas wyd finna shyt	26%
#nhl #bruins #mlb knicks	11%

- Do women use more standard language?
- **Is men's writing more "informational"?**

Literally the simplest possible thing

K-means clustering on Twitter timelines

	% Women
hubs **blogged recipe fabric**	90%
kidd hubs xo =]	80%
wyd #oomf lmbo shyt	60%
n_ggas wyd finna shyt	26%
#nhl #bruins #mlb knicks	11%

- Do women use more standard language?
- **Is men's writing more "informational"?**

What is social meaning about?

Idealized model: saying the “same thing” differently, depending on (and conveying information about):

- **the speaker/writer**

  dialect, gender, race, “latent user attributes”, . . .

- the communicative situation

  register, formality, politeness, accommodation, . . .

- attitude towards the talk and topic

  sentiment, subjectivity, irony, . . .

What is social meaning about?

Idealized model: saying the "same thing" differently, depending on (and conveying information about):

- the speaker/writer

  dialect, gender, race, "latent user attributes", ...

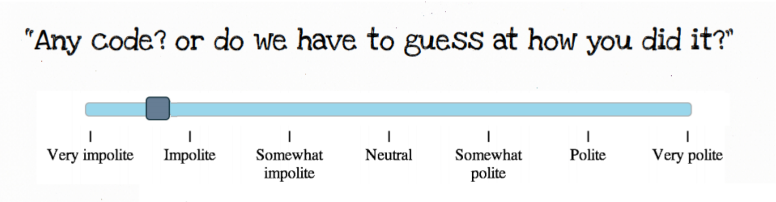
- **the communicative situation**

  register, formality, politeness, accommodation, ...

- attitude towards the talk and topic

  sentiment, subjectivity, irony, ...

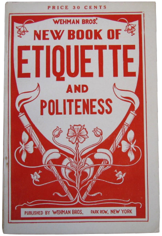
Social-situational variation



- Ask crowdworkers to annotate the extent of some social variable in text:
  - politeness (Danescu-Niculescu-Mizil et al., 2013)
  - formality (Pavlick & Tetreault, 2016)
- Train a supervised system, examine features.

Construct validity

- Politeness and formality are constructs with strong theoretical support (Brown, 1987; Heylighen & Dewaele, 1999)
- Crowdsourced annotations capture *folk linguistic* intuitions about how these constructs work (and that's okay)



But... are there other forms of social meaning? What are the principle linguistic dimensions of social-situational variation?

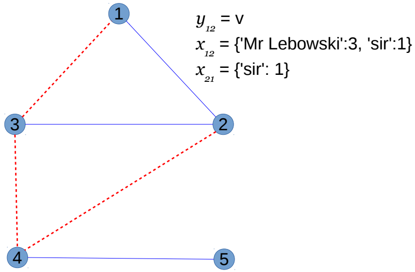
You're Mr. Lebowski, I'm the Dude: Inducing Address Term Formality in Signed Social Networks (Krishnan & Eisenstein, 2015).

Vinodh Krishnan

- How do address terms like dude and Mr. Lebowski convey social meaning?
- What social configurations are likely or unlikely?

Setting

- We observe address terms $\mathbf{x}$ but not relation types y.
- Inferring y gives a labeling over edges.
- Estimating $P(\mathbf{x} \mid y)$ gives the distribution over address terms for each edge type.



Complete model specification

$$P(\boldsymbol{y}, \boldsymbol{x} \mid G; \Theta, \boldsymbol{\beta}, \boldsymbol{\eta}) = P(\boldsymbol{x} \mid \boldsymbol{y}; \Theta) P(\boldsymbol{y} \mid G; \boldsymbol{\beta}, \boldsymbol{\eta})$$

Bayesian inference answers several questions:

1. What is the social meaning of each dyad?
2. How is social meaning expressed in address terms?
3. Are there structural regularities across networks?

Complete model specification

$$P(\mathbf{y}, \mathbf{x} \mid G; \Theta, \boldsymbol{\beta}, \boldsymbol{\eta}) = P(\mathbf{x} \mid \mathbf{y}; \Theta) P(\mathbf{y} \mid G; \boldsymbol{\beta}, \boldsymbol{\eta})$$

Bayesian inference answers several questions:

1. What is the social meaning of each dyad?
2. How is social meaning expressed in address terms?
3. Are there structural regularities across networks?

Complete model specification

$$P(\boldsymbol{y}, \boldsymbol{x} \mid G; \Theta, \boldsymbol{\beta}, \boldsymbol{\eta}) = P(\boldsymbol{x} \mid \boldsymbol{y}; \Theta) P(\boldsymbol{y} \mid G; \boldsymbol{\beta}, \boldsymbol{\eta})$$

Bayesian inference answers several questions:

1. What is the social meaning of each dyad?
2. How is social meaning expressed in address terms?
3. Are there structural regularities across networks?

Complete model specification

$$P(\boldsymbol{y}, \boldsymbol{x} \mid G; \Theta, \boldsymbol{\beta}, \boldsymbol{\eta}) = P(\boldsymbol{x} \mid \boldsymbol{y}; \Theta) P(\boldsymbol{y} \mid G; \boldsymbol{\beta}, \boldsymbol{\eta})$$

Bayesian inference answers several questions:

1. What is the social meaning of each dyad?
2. How is social meaning expressed in address terms?
3. Are there structural regularities across networks?

Forms of address

Names such as Barack, Barack Hussein Obama.

Titles such as Ms., Dr., Private, Reverend.
Titles can be used for address either by preceding a name (e.g., Colonel Kurtz), or in isolation (e.g., Yes, Colonel.).

Placeholder names such as dude, bro, brother, sweetie, cousin, and asshole.
These terms can be used for address only in isolation.

Lexicon induction: titles

- Find terms that frequently appear at the beginning of address spans containing the character's name.
- We manually filter out 17 of 34 candidates, obtaining:

agent, aunt, captain, colonel, commander, cousin, deputy, detective, dr, herr, inspector, judge, lord, master, mayor, miss, mister, miz, monsieur, mr, mrs, ms, professor, queen, reverend, sergeant, uncle

Lexicon induction: placeholder names

- Find terms that are frequently tagged as the unique element in an address span.
- After manually filter out 41 of 96 candidates, we obtain:

asshole, babe, baby, boss, boy, bro, bud, buddy, cocksucker, convict, cousin, cowboy, cunt, dad, darling, dear, detective, doll, dude, dummy, father, fella, gal, ho, hon, honey, kid, lad, lady, lover, ma, madam, madame, man, mate, mister, mon, moron, motherfucker, pal, papa, partner, peanut, pet, pilgrim, pop, president, punk, shithead, sir, sire, son, sonny, sport, sucker, sugar, sweetheart, sweetie, tiger

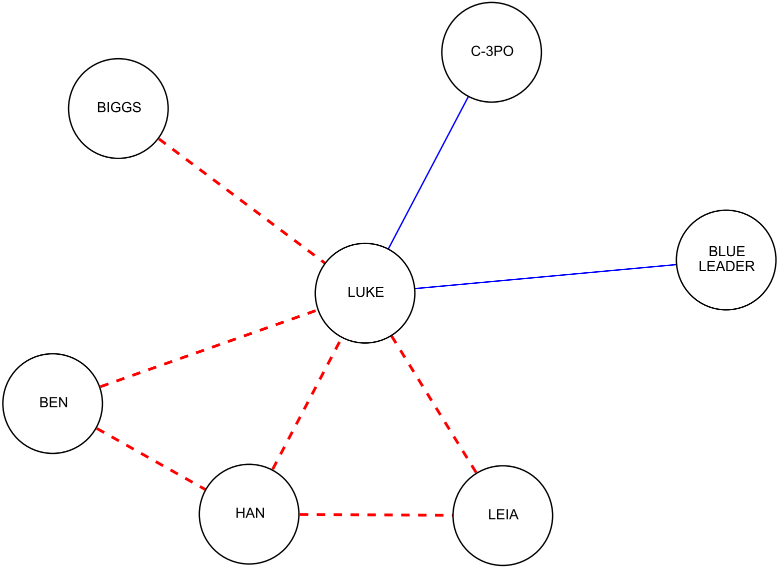
Cluster coherence

V-cluster	T-cluster
sir	FIRSTNAME
mr+LASTNAME	man
mr+FIRSTNAME	baby
mr	honey
miss+LASTNAME	darling
son	sweetheart
mister+FIRSTNAME	buddy
mrs	sweetie

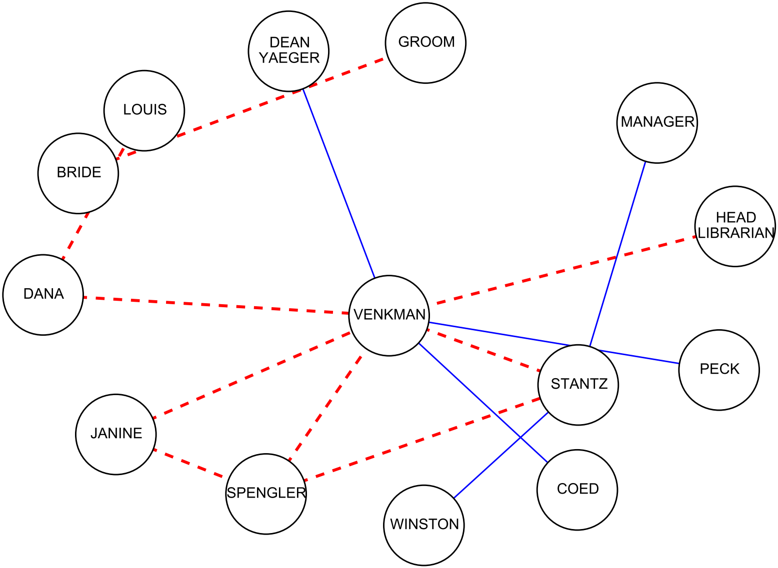
On an intrusion task

- Raters found the intruder in 73% of cases for the full model (M4).
- ... versus 52% in the text-only model (M1).

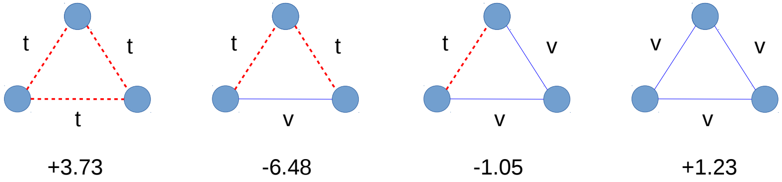
Star Wars



Ghostbusters



Network features



What is social meaning about?

Idealized model: saying the “same thing” differently, depending on (and conveying information about):

- the speaker/writer

  dialect, gender, race, “latent user attributes”, ...

- **the communicative situation**

  register, formality, politeness, accommodation, ...

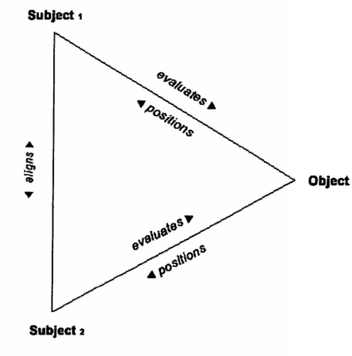
- attitude towards the talk and topic

  sentiment, subjectivity, irony, ...

A Multidimensional Lexicon for Interpersonal Stancetaking (Pavalanathan, Fitzpatrick, Kiesling, and Eisenstein, 2017)

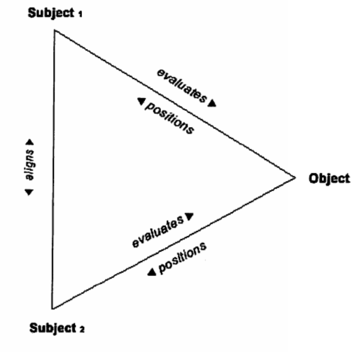


- What interactional and linguistic patterns distinguish online communities?
- How do automatically-identified linguistic styles relate to annotated constructs such as formality and politeness?



The **stance triangle** is proposed as a representation for interactional meaning (Du Bois, 2007)

- Alignment (or not) between speakers
- Evaluation of (and positioning w.r.t.) the object of discussion and the talk itself



The **stance triangle** is proposed as a representation for interactional meaning (Du Bois, 2007)

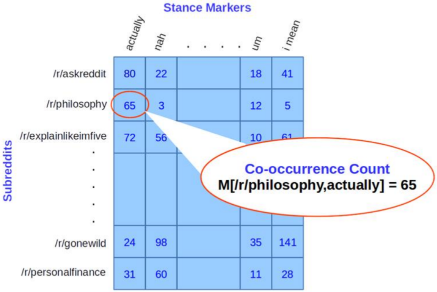
- Alignment (or not) between speakers
- Evaluation of (and positioning w.r.t.) the object of discussion and the talk itself

Hard to annotate! (Fitzpatrick et al 2014)

Learning representations of interpersonal stance

Hypotheses:

- Online communities differ in their characteristic stances
- Stance is revealed (in part) by a lexicon of **stance markers**



STANCE LEXICON: SEED LIST

- List of markers from Biber and Finegan (1989) 448 seed markers
  - certainty adverbs *actually, of course, in fact*
  - affect markers *amazing, thankful, sadly*
  - hedges *kind of, maybe, something like*
- Dialog act markers from Jurafsky et al. (1998) 74 seed markers
  - *oh yeah, nah, wow*

STANCE LEXICON: EXPANSION

Similar words from word embeddings trained on Reddit comments

- Seed stance markers

Seed term	Expanded terms
significantly	considerably, substantially, dramatically
certainly	surely, frankly, definitely
incredibly	extremely, unbelievably, exceptionally

228 additional markers

- Seed dialog acts

Seed term	Expanded terms
nope	nah, yup, nevermind
great	fantastic, terrific, excellent

112 additional markers

A MULTIDIMENSIONAL LEXICON FOR INTERPERSONAL STANCETAKING

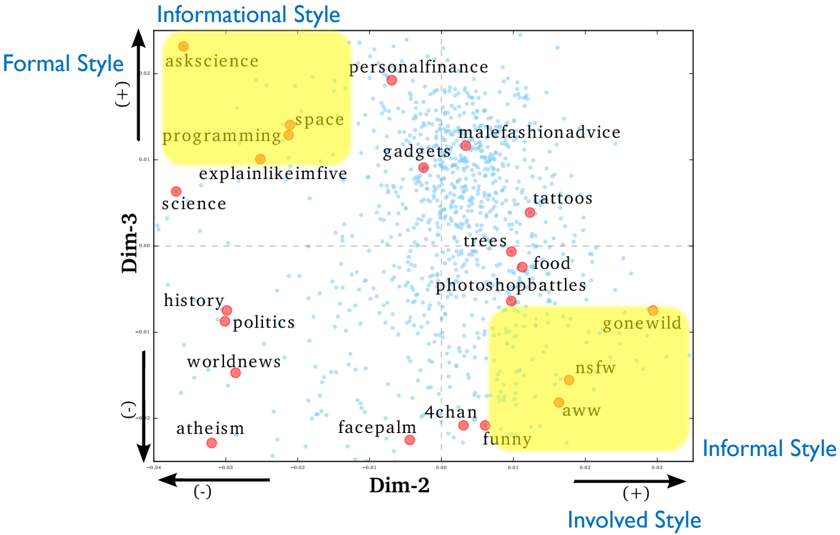
Dim-1	-	beautifully, pleased, thanks, spectacular, delightful
	+	just, even, all, no, so
Dim-2	-	suggests that, demonstrates, conclude, demonstrated, demonstrate
	+	lovely, awww, hehe, aww, haha
Dim-3	-	funnier, hilarious, disturbing, creepy, funny
	+	thanks, ideally, calculate, estimate, calculation
Dim-4	-	phenomenal, bummed, enjoyed, fantastic, disappointing
	+	hello, thx, hehe, aww, hi

A MULTIDIMENSIONAL LEXICON FOR INTERPERSONAL STANCETAKING

Dim-1	-	beautifully, pleased, thanks, spectacular, delightful
	+	just, even, all, no, so
Dim-2	-	suggests that, demonstrates, conclude, demonstrated, demonstrate
	+	lovely, awww, hehe, aww, haha
Dim-3	-	funnier, hilarious, disturbing, creepy, funny
	+	thanks, ideally, calculate, estimate, calculation
Dim-4	-	phenomenal, bummed, enjoyed, fantastic, disappointing
	+	hello, thx, hehe, aww, hi

A MULTIDIMENSIONAL LEXICON FOR INTERPERSONAL STANCETAKING

Dim-1	-	beautifully, pleased, thanks, spectacular, delightful
	+	just, even, all, no, so
Informational vs. Involved Style	-	suggests that, demonstrates, conclude, demonstrated, demonstrate
	+	lovely, awww, hehe, aww, haha
Dim-3	-	funnier, hilarious, disturbing, creepy, funny
	+	thanks, ideally, calculate, estimate, calculation
Dim-4	-	phenomenal, bummed, enjoyed, fantastic, disappointing
	+	hello, thx, hehe, aww, hi


Informational Style
Formal Style
(+)
Dim-3
(-)
askscience
space
programming
explainlikeimfive
science
personalfinance
malefashionadvice
gadgets
tattoos
trees
food
photoshopbattles
history
politics
worldnews
atheism
facepalm
4chan
funny
gonewild
nsfw
aww
Informal Style
Dim-2
(-)
(+)
Involved Style

MODEL VALIDATION: INTRINSIC EVALUATIONS

Word intrusion task

suggests that	demonstrates		conclude	demonstrated

suggests that	demonstrates	awww	conclude	demonstrated

MODEL VALIDATION: INTRINSIC EVALUATIONS

Word intrusion task

suggests that	demonstrates		conclude	demonstrated

suggests that	demonstrates	awww	conclude	demonstrated

intruder term

MODEL VALIDATION: INTRINSIC EVALUATIONS

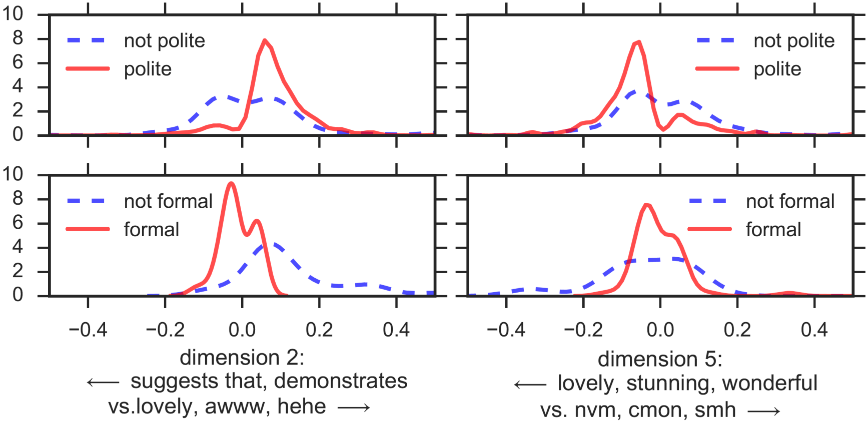
Word intrusion task

suggests that	demonstrates		conclude	demonstrated

suggests that	demonstrates	awww	conclude	demonstrated

intruder term

- Inter-rater reliability: Krippendorf's α = 0.73
- Model precision: 0.82 (chance precision = 0.20)


10
8
6
4
2
0
not polite
polite
not polite
polite
10
8
6
4
2
0
not formal
formal
not formal
formal
−0.4
−0.2
0.0
0.2
0.4
−0.4
−0.2
0.0
0.2
0.4
dimension 2:
⟵ suggests that, demonstrates
vs.lovely, awww, hehe ⟶
dimension 5:
⟵ lovely, stunning, wonderful
vs. nvm, cmon, smh ⟶

Predicting cross-posting in multiple subreddits

Predicting cross-posting in multiple subreddits

Classification task

Predict whether a pair of subreddits has higher or lower membership overlap

Dataset

Paired each of the top 100 popular subreddits with five other subreddits with highest and lowest membership overlap

High-Overlap Pairs	Low-Overlap Pairs
r/blog, r/announcements	r/soccer, r/nosleep
r/politics, r/technology	r/Minecraft, r/personalfinance
r/pokemon, r/wheredidthesodago	r/programming, r/gonewild

MODEL VALIDATION: EXTRINSIC EVALUATIONS

Predicting cross-posting in multiple subreddits

Models & Results

Model	Accuracy
BOW-COSINE	
STANCE-COSINE	
BOW-SVD	
STANCE-SVD	

MODEL VALIDATION: EXTRINSIC EVALUATIONS

Predicting cross-posting in multiple subreddits

Models & Results

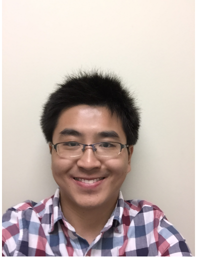
Model	Accuracy
BOW-COSINE	66.13%
STANCE-COSINE	64.31%
BOW-SVD	77.48%
STANCE-SVD	**84.93%**

Interactional style predicts cross-posting behavior

Computational models of social meaning

- What is the role of computation in understanding social meaning in language?
- Why do we need representation learning?
- **Is social meaning relevant to "core tasks" in NLP?**

Overcoming language variation in sentiment analysis with social attention (Yang & Eisenstein, 2017)



Yi Yang

- Can we exploit social network homophily to make document classification more robust to social variation?

Language variation: a challenge for NLP

"I would like to believe he's **sick** rather than just mean and evil."

Language variation: a challenge for NLP

"I would like to believe he's **sick** rather than just mean and evil."

"You could've been getting down to this **sick** beat."

(Yang & Eisenstein, 2017)

Personalization by ensemble

- Goal: personalized conditional likelihood, $P(y \mid \boldsymbol{x}, a)$, where a is the author.
- Problem: We have labeled examples for only a few authors.

Personalization by ensemble

- Goal: personalized conditional likelihood, $P(y \mid \boldsymbol{x}, a)$, where a is the author.
- Problem: We have labeled examples for only a few authors.
- **Personalization ensemble**

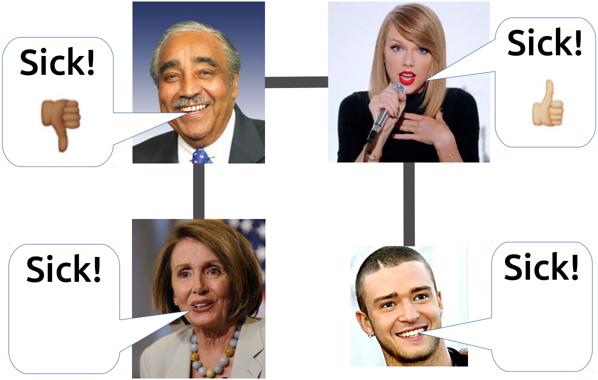
$$P(y \mid \boldsymbol{x}, a) = \sum_k P_k(y \mid \boldsymbol{x})\pi_a(k)$$

  - $P_k(y \mid \boldsymbol{x})$ is a basis model
  - $\boldsymbol{\pi}_a(\cdot)$ are the ensemble weights for author a

Homophily to the rescue?

Labeled data

Unlabeled data



Are language styles **assortative** on the social network?

Evidence for linguistic homophily

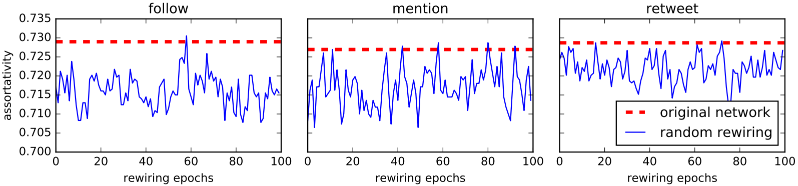
Pilot study: is classifier accuracy **assortative** on the Twitter social network?

$$\text{assort}(G) = \frac{1}{\#|G|} \sum_{(i,j) \in G} \delta(y_i = \hat{y}_i)\delta(y_j = \hat{y}_j) \\ + \delta(y_i \neq \hat{y}_i)\delta(y_j \neq \hat{y}_j)$$

Evidence for linguistic homophily

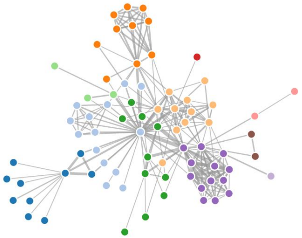
Pilot study: is classifier accuracy **assortative** on the Twitter social network?

$$\text{assort}(G) = \frac{1}{\#|G|} \sum_{(i,j) \in G} \delta(y_i = \hat{y}_i)\delta(y_j = \hat{y}_j) + \delta(y_i \neq \hat{y}_i)\delta(y_j \neq \hat{y}_j)$$



Network-driven personalization

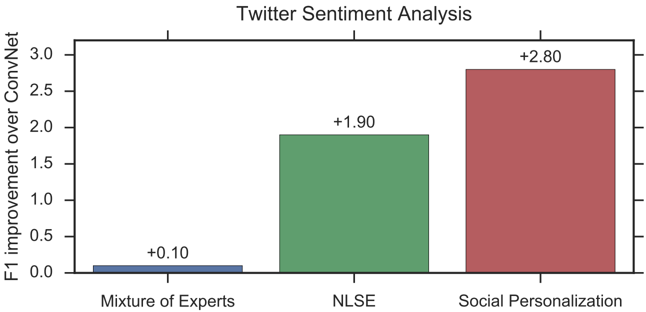
- For each author, estimate a **node embedding** $\boldsymbol{e}_a$ (Tang et al., 2015).
- Nodes who share neighbors get similar embeddings.



$$\boldsymbol{\pi}_a = \text{SoftMax}(f(\boldsymbol{e}_a))$$

$$P(y \mid \boldsymbol{x}, a) = \sum_{k=1}^{K} P_k(y \mid \boldsymbol{x}) \pi_a(k)$$

Results



Improvements over ConvNet baseline:

- +2.8% on Twitter Sentiment Analysis
- +2.7% on Ciao Product Reviews

NLSE is prior state-of-the-art (Astudillo et al., 2015).

Variable sentiment words

	More positive	More negative
1	banging loss fever broken fucking	dear like god yeah wow
2	chilling cold ill sick suck	satisfy trust wealth strong lmao
3	ass damn piss bitch shit	talent honestly voting win clever
4	insane bawling fever weird cry	lmao super lol haha hahaha
5	ruin silly bad boring dreadful	*lovatics* wish *beliebers arianators kendall*

What's the right representation for social meaning?

Word senses are only ever defined relative to a set of interests. The set of senses defined by a dictionary may or may not match the set that is relevant for an NLP application (Kilgarriff, 1997).

What's the right representation for social meaning?

Word senses are only ever defined relative to a set of interests. The set of senses defined by a dictionary may or may not match the set that is relevant for an NLP application (Kilgarriff, 1997).

- Annotating the full social meaning of an utterance is not possible.
- The way forward: representation learning **plus linguistics** for model design, lexicon and metadata selection.
- Single-membership models are only the first step!

Acknowledgments

- **Collaborators**: David Bamman, Ming-Wei Chang, Jim Fitzpatrick, Scott Kiesling, Vinodh Krishnan, Umashanthi Pavalanathan, Tyler Schnoebelen, Yi Yang
- **Sponsors**: National Science Foundation, Air Force Office of Scientific Research, National Institutes for Health.

References I

Argamon, S., Koppel, M., Pennebaker, J. W., & Schler, J. (2007). Mining the blogosphere: Age, gender and the varieties of self-expression. *First Monday*, *12*(9).

Astudillo, R. F., Amir, S., Lin, W., Silva, M., & Trancoso, I. (2015). Learning word representations from scarce and noisy data with embedding sub-spaces. In *Proceedings of the Association for Computational Linguistics (ACL)*, Beijing.

Brown, P. (1987). *Politeness: Some universals in language usage*, volume 4. Cambridge University Press.

Danescu-Niculescu-Mizil, C., Sudhof, M., Jurafsky, D., Leskovec, J., & Potts, C. (2013). A computational approach to politeness with application to social factors. In *Proceedings of the Association for Computational Linguistics (ACL)*, (pp. 250–259)., Sophia, Bulgaria.

Du Bois, J. W. (2007). The stance triangle. In R. Engelbretson (Ed.), *Stancetaking in discourse* (pp. 139–182). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Eckert, P. (2000). *Linguistic variation as social practice*. Blackwell.

Eisenstein, J., O'Connor, B., Smith, N. A., & Xing, E. P. (2010). A latent variable model for geographic lexical variation. In *Proceedings of Empirical Methods for Natural Language Processing (EMNLP)*, (pp. 1277–1287).

Eisenstein, J., O'Connor, B., Smith, N. A., & Xing, E. P. (2014). Diffusion of lexical change in social media. *PLoS ONE*, *9*.

Eisenstein, J., Smith, N. A., & Xing, E. P. (2011). Discovering sociolinguistic associations with structured sparsity. In *Proceedings of the Association for Computational Linguistics (ACL)*, (pp. 1365–1374)., Portland, OR.

Grieve, J., Nini, A., & Guo, D. (2017). Analyzing lexical emergence in modern american english online 1. *English Language & Linguistics*, *21*(1), 99–127.

Heeringa, W. & Nerbonne, J. (2001). Dialect areas and dialect continua. *Language Variation and Change*, *13*(3), 375–400.

Heylighen, F. & Dewaele, J.-M. (1999). Formality of language: definition, measurement and behavioral determinants. *Interner Bericht, Center Leo Apostel, Vrije Universiteit Brüssel*.

Kilgarriff, A. (1997). I dont believe in word senses. *Computers and the Humanities*, *31*(2), 91–113.

References II

Krishnan, V. & Eisenstein, J. (2015). "You're Mr. Lebowski, I'm The Dude": Inducing address term formality in signed social networks. In *NAACL*.

Labov, W. (1972). The social stratification of (r) in new york city department stores. In *Sociolinguistic Patterns* (pp. 43–54). University of Pennsylvania of Press.

Nguyen, D., Gravel, R., Trieschnigg, D., & Meder, T. (2013). "how old do you think i am?" a study of language and age in twitter. In *Proceedings of the International Conference on Web and Social Media (ICWSM)*.

Pavlick, E. & Tetreault, J. (2016). An empirical analysis of formality in online communication. *Transactions of the Association for Computational Linguistics*, *4*, 61–74.

Rahimi, A., Cohn, T., & Baldwin, T. (2017). A neural model for user geolocation and lexical dialectology. In *Proceedings of Empirical Methods for Natural Language Processing (EMNLP)*.

Roller, S., Speriosu, M., Rallapalli, S., Wing, B., & Baldridge, J. (2012). Supervised text-based geolocation using language models on an adaptive grid. In *Proceedings of Empirical Methods for Natural Language Processing (EMNLP)*, (pp. 1500–1510).

Sankoff, D. & Labov, W. (1979). On the uses of variable rules. *Language in society*, *8*(2-3), 189–222.

Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J., & Mei, Q. (2015). Line: Large-scale information network embedding. In *Proceedings of the Conference on World-Wide Web (WWW)*, (pp. 1067–1077).

Yang, Y. & Eisenstein, J. (2017). Overcoming language variation in sentiment analysis with social attention. *Transactions of the Association for Computational Linguistics (TACL), in press*.