Structured Probabilistic Models for Online Dialogue and Text

Dhanya Sridhar

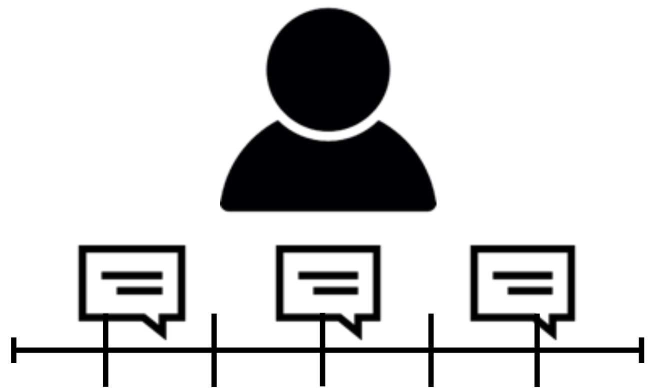
3.22.18

Stanford NLP Seminar

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

Socio-behavior modeling with text

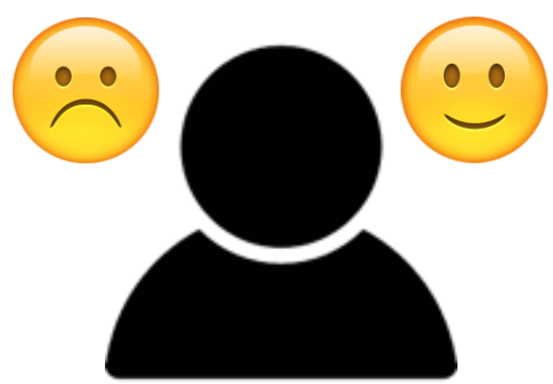
Data:



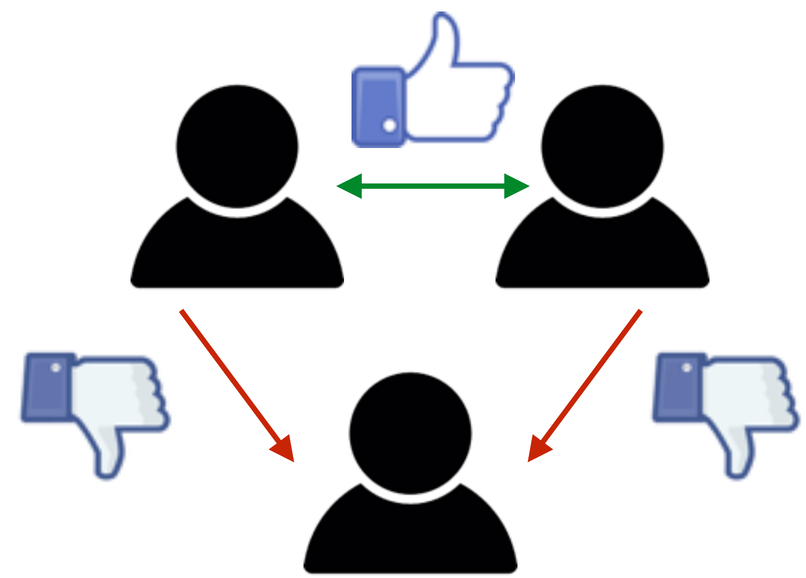
Longitudinal

Dialogue

Inferences:



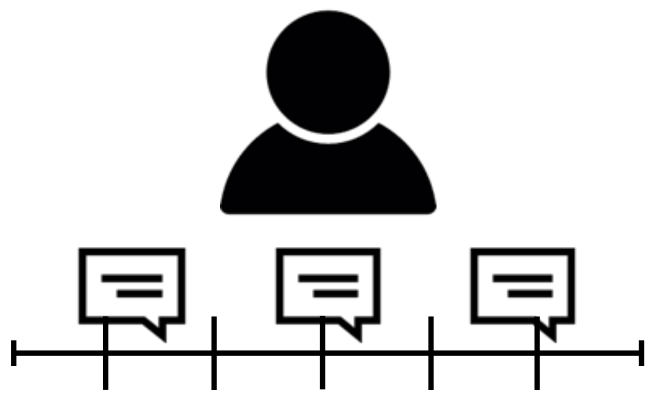
Mood modeling



Group attitudes

Socio-behavior modeling challenges

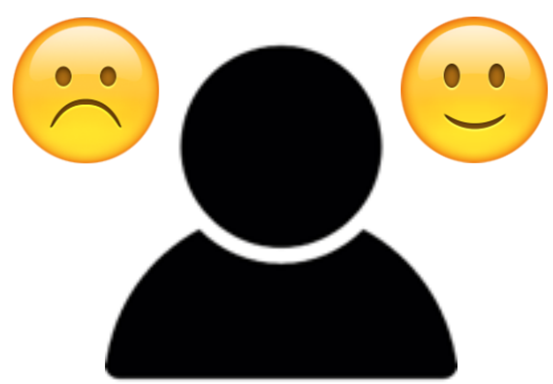
Data:



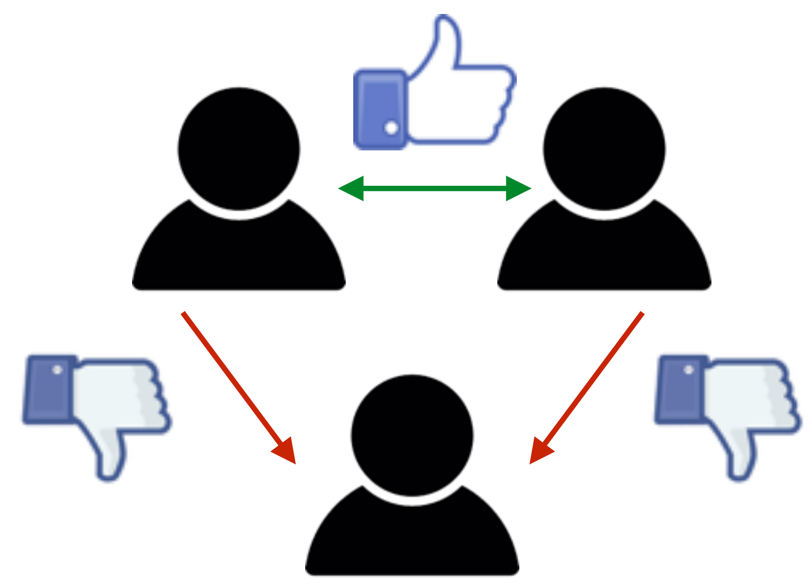
Longitudinal

Dialogue

Inferences:



Mood modeling



Group attitudes

Challenges: interrelated inferences, heterogenous data, knowledge discovery

Key contributions

Probabilistic models that

Exploit structure	**Fuse signals**
Support **causal** inference	**Discover** knowledge

for modeling online dialogue and text

Understanding stances on issues

ars TECHNICA
BIZ & IT TECH SCIENCE POLICY CARS GAMING & CULTURE

PRE-DETERMINED OUTCOME —

Here's Ajit Pai's "proof" that killing net neutrality created more broadband

CNBC
HOME U.S. NEWS MARKETS INVESTING TECH MAKE IT VIDEO

New York attorney general will sue to stop the FCC's 'illegal rollback of net neutrality'

Social media indicates stances

Today, I'm proposing to repeal the heavy-handed Internet regulations imposed by the Obama

Go Mr. Pai. I watched as you fought Obama's costly regulatory abuse for years.

THANK YOU for having the bravery to stand for giant corporations and ignore the good of the people.

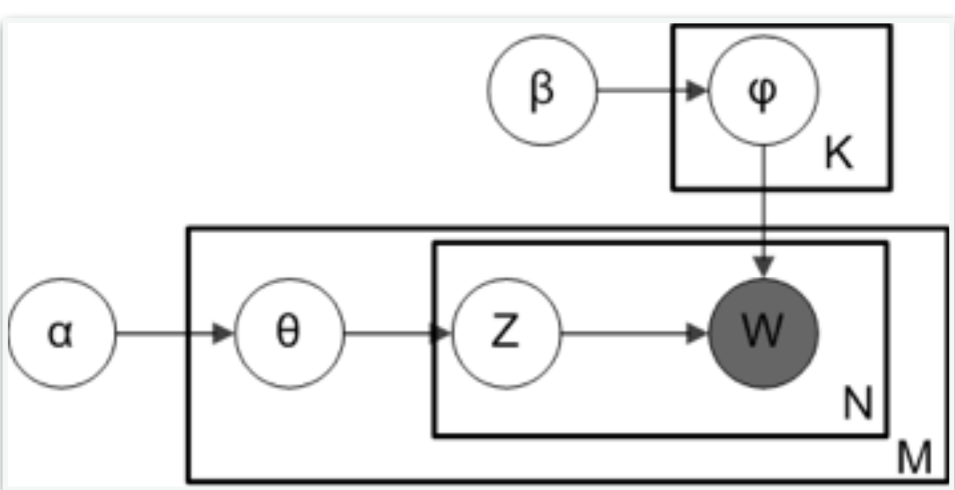
I will sue to stop FCC's illegal rollback of net neutrality.

What was illegal was the govt's seizure of the internet. Only a matter of time before THEY restricted access.

False. The government is required to protect the rights of all Americans as per the Constitution.

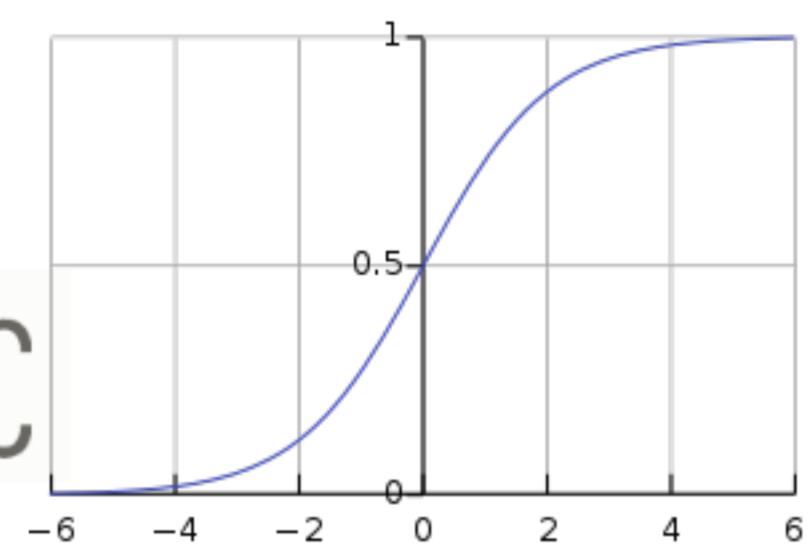
Modeling text documents

Today, I'm proposing to repeal the heavy-handed Internet regulations imposed by the Obama	Go Mr. Pai. I watched as you fought Obama's costly regulatory abuse for years.
THANK YOU for having the bravery to stand for giant corporations and ignore the good of the people.	I will sue to stop FCC's illegal rollback of net neutrality.
What was illegal was the govt's seizure of the internet. Only a matter of time before THEY restricted access.	False. The government is required to protect the rights of all Americans as per the Constitution.



word2vec

LIWC

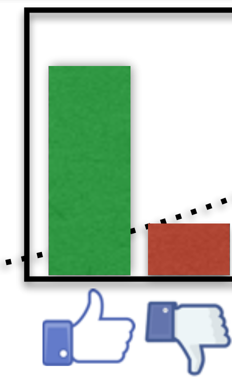


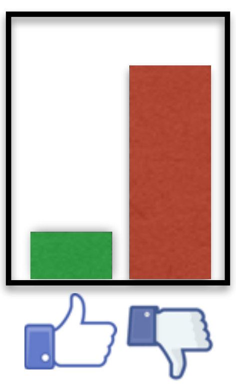
Improving upon existing methods

Today, I’m proposing to repeal the heavy-handed Internet regulations imposed by the Obama Administration …

Go Mr. Pai. I watched as you fought Obama’s costly regulatory abuse for years.

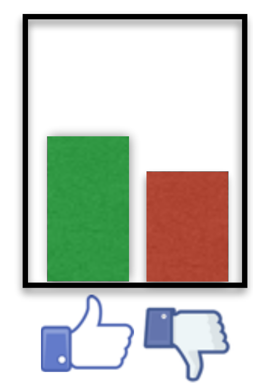
THANK YOU for having the bravery to stand for giant corporations and ignore the good of the people.

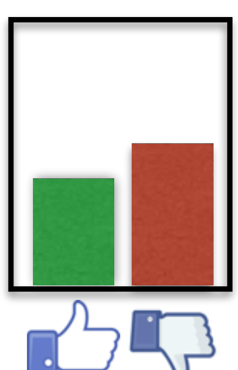




I will sue to stop FCC’s illegal rollback of net neutrality.

What was illegal was the govt’s seizure of the internet. Only a matter of time before THEY restricted access.





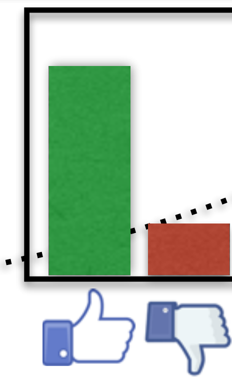
False. The government is required to protect the rights of all Americans as per the Constitution.

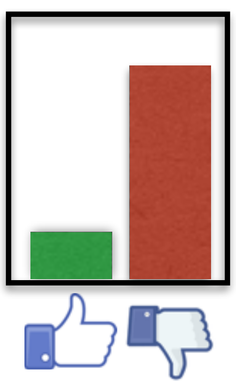
Improving upon existing methods

Today, I'm proposing to repeal the heavy-handed Internet regulations imposed by the Obama Administration …

Go Mr. Pai. I watched as you fought Obama's costly regulatory abuse for years.

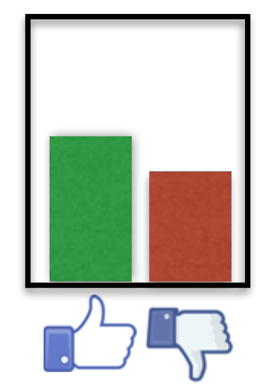
THANK YOU for having the bravery to stand for giant corporations and ignore the good of the people.

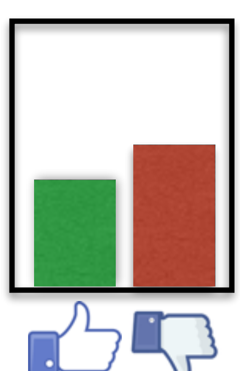




I will sue to stop FCC's illegal rollback of net neutrality.

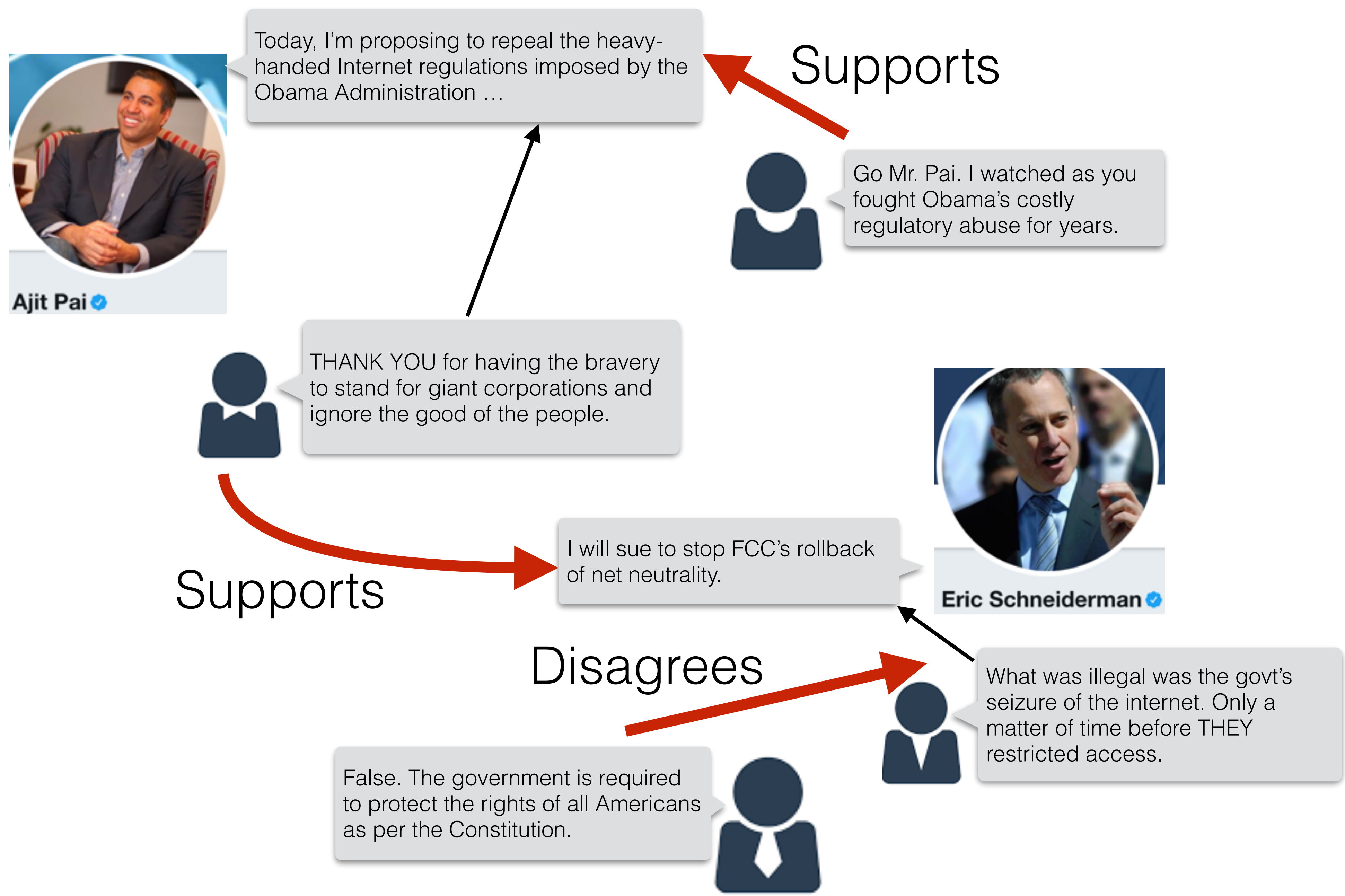
What was illegal was the govt's seizure of the internet. Only a matter of time before THEY restricted access.





False. The government is required to protect the rights of all Americans as per the Constitution.

Exploit structural dependencies



Interaction network between users induces useful dependencies for consistent predictions

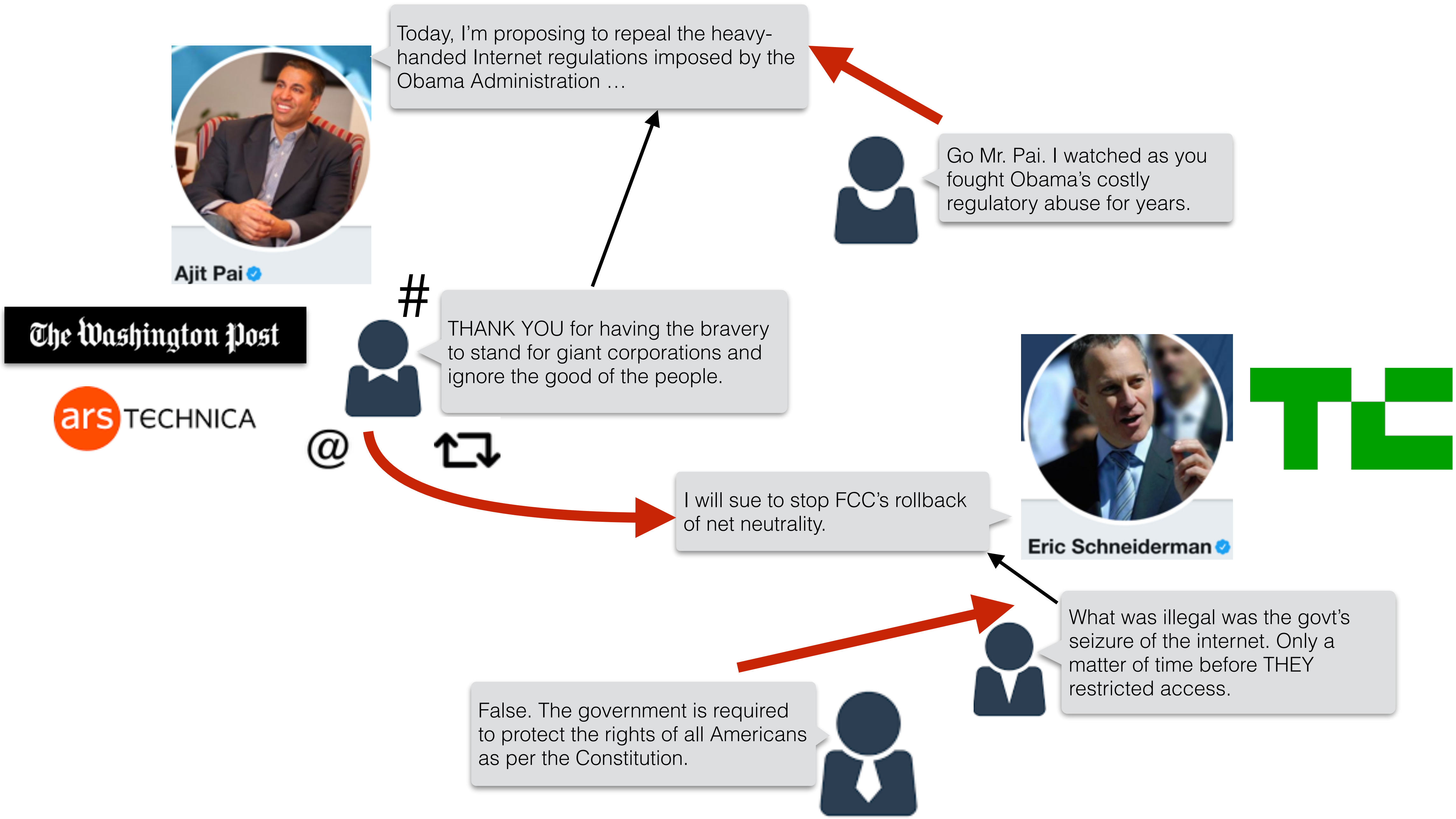
Fuse heterogenous signals

Today, I'm proposing to repeal the heavy-handed Internet regulations imposed by the Obama Administration …

I will sue to stop FCC's rollback of net neutrality.

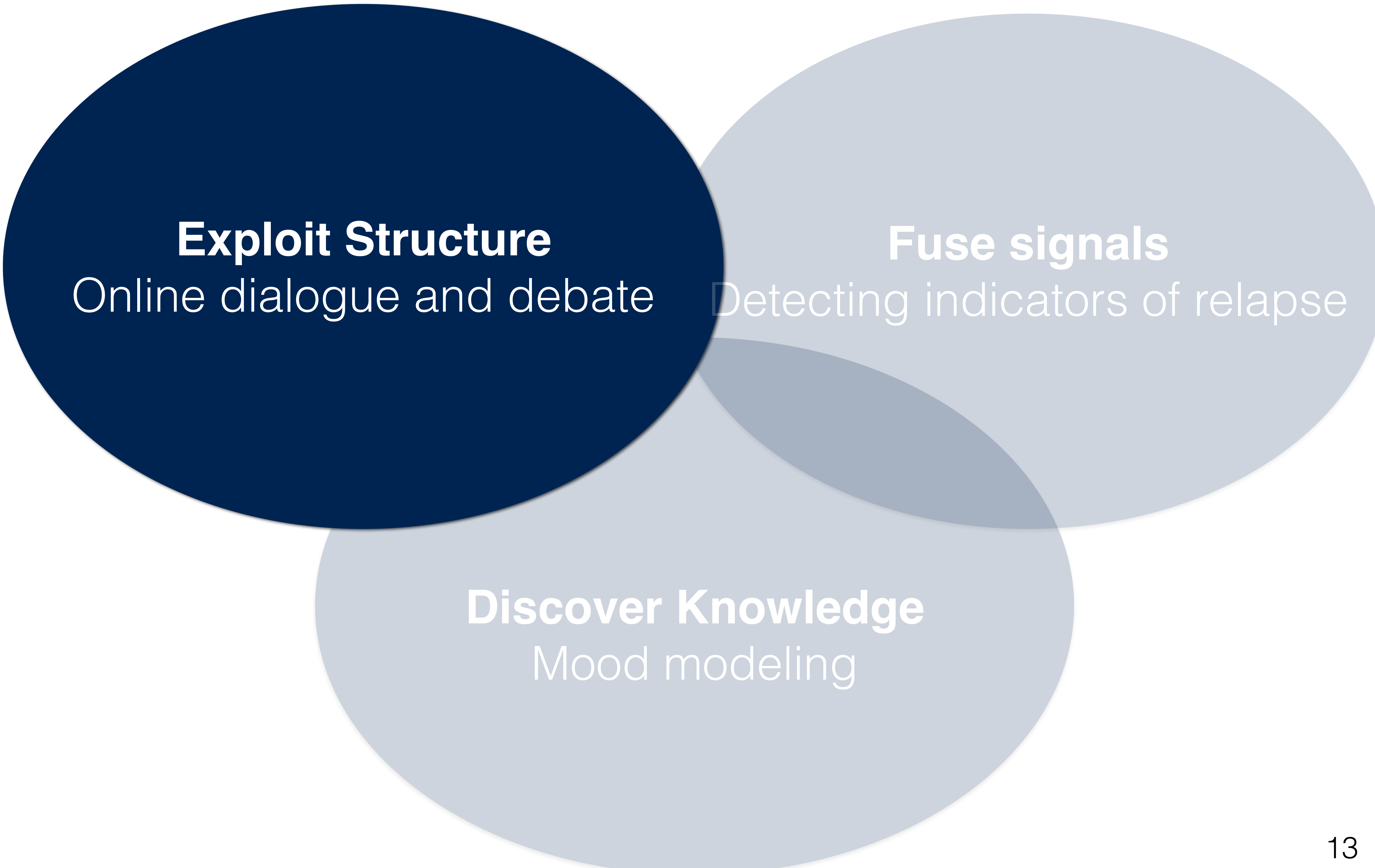
Combine additional information sources of varying reliability

Discover knowledge

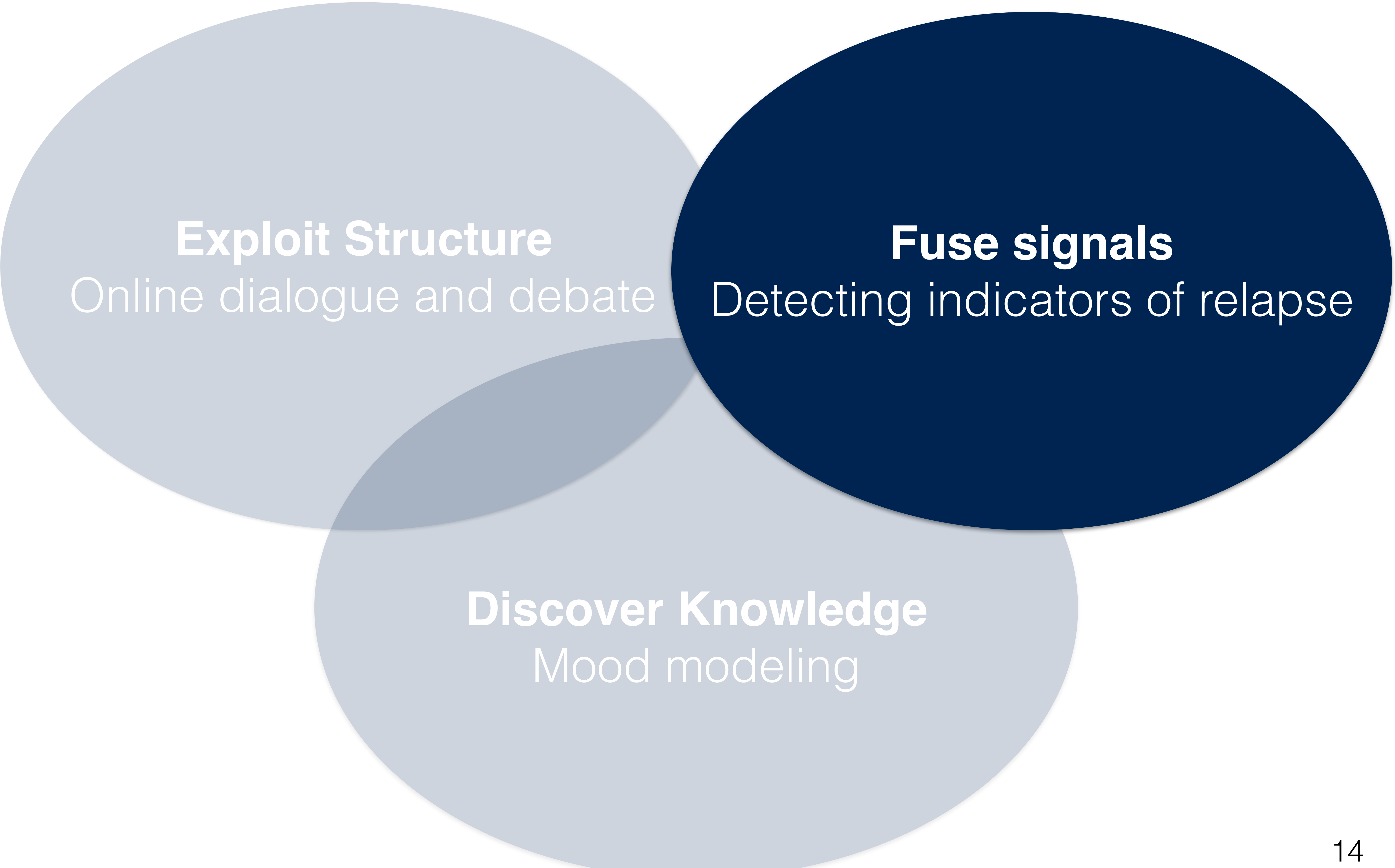


Users who retweet those followed by top users share stance

Roadmap of my talk



Roadmap of my talk



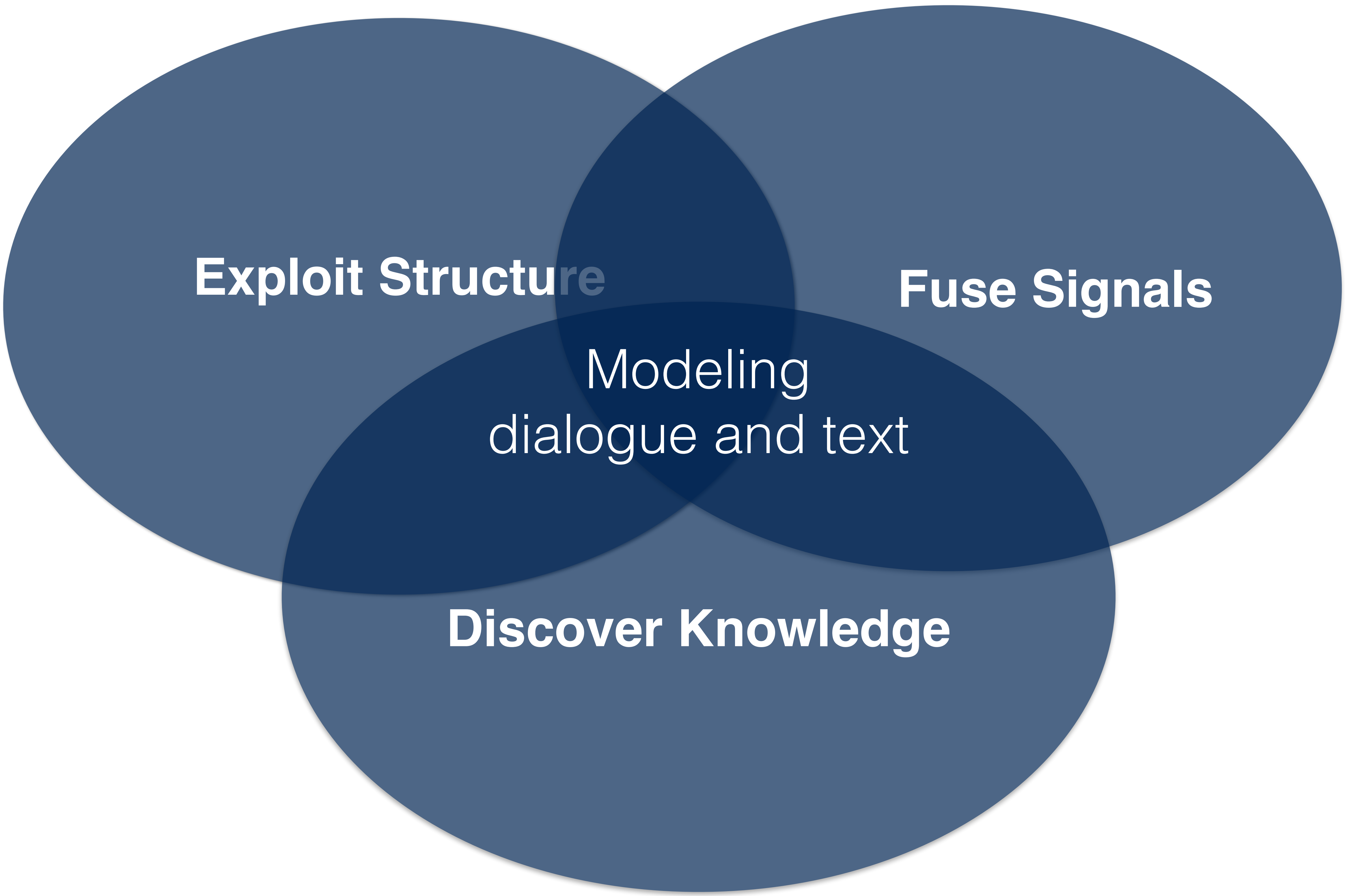
Roadmap of my talk

Exploit Structure
Online dialogue and debate

Fuse signals
Detecting indicators of relapse

Discover Knowledge
Mood modeling

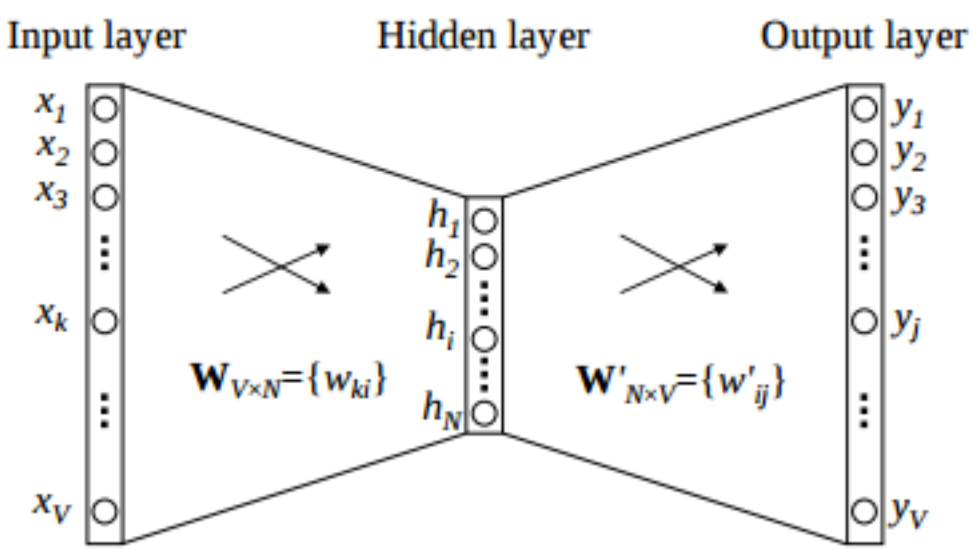
Roadmap of my talk

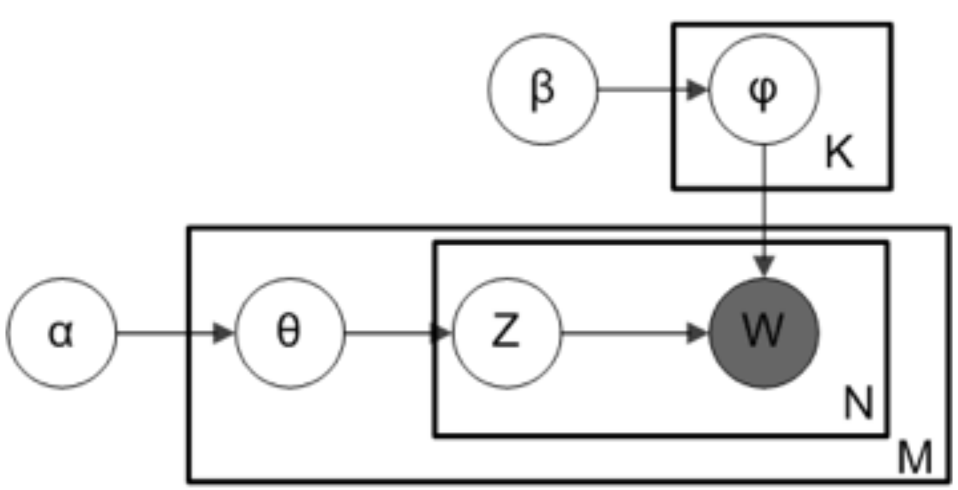


Background

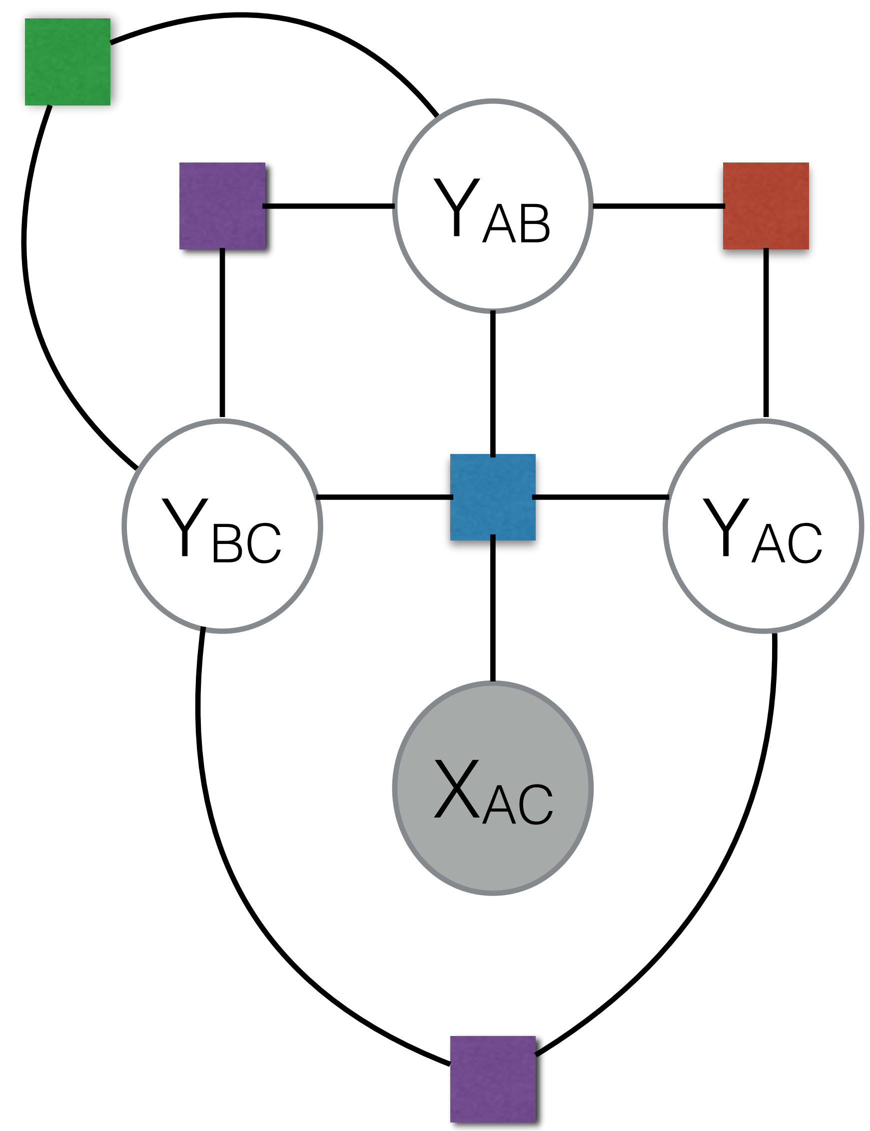
Factor graphs for structured prediction

Structure in input space:

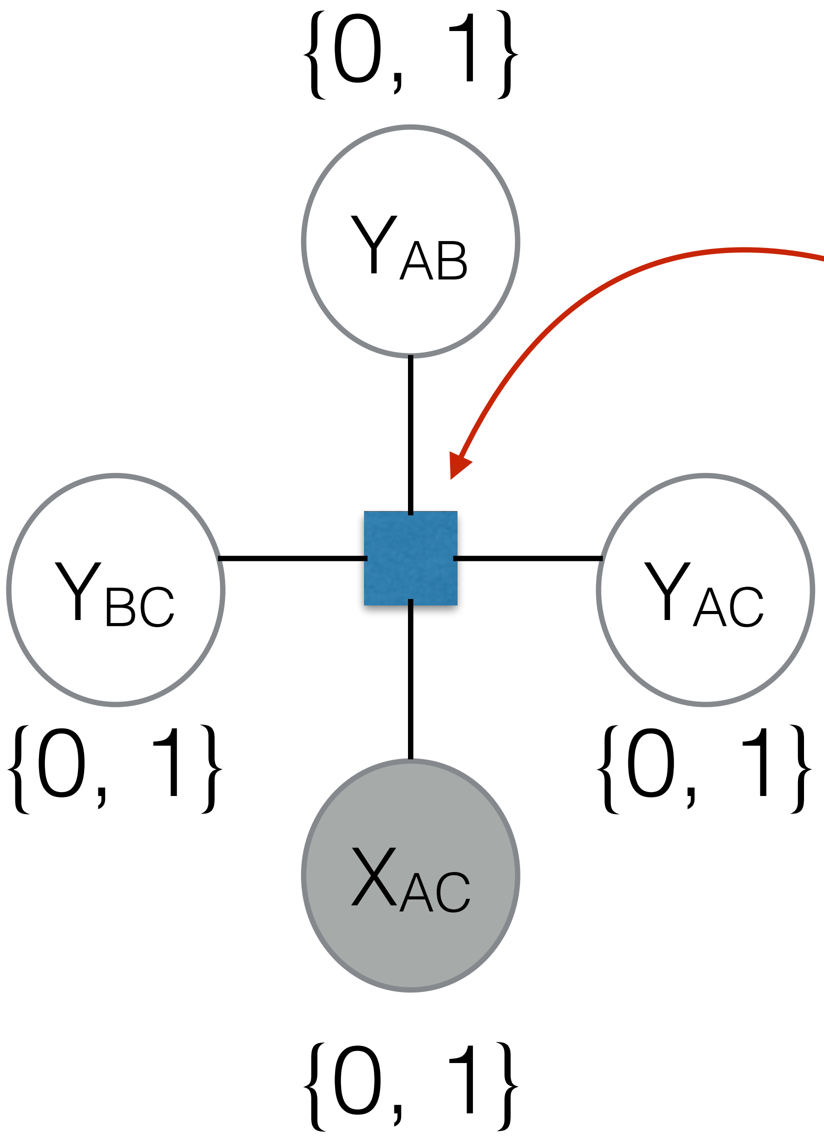




Structure in output space:



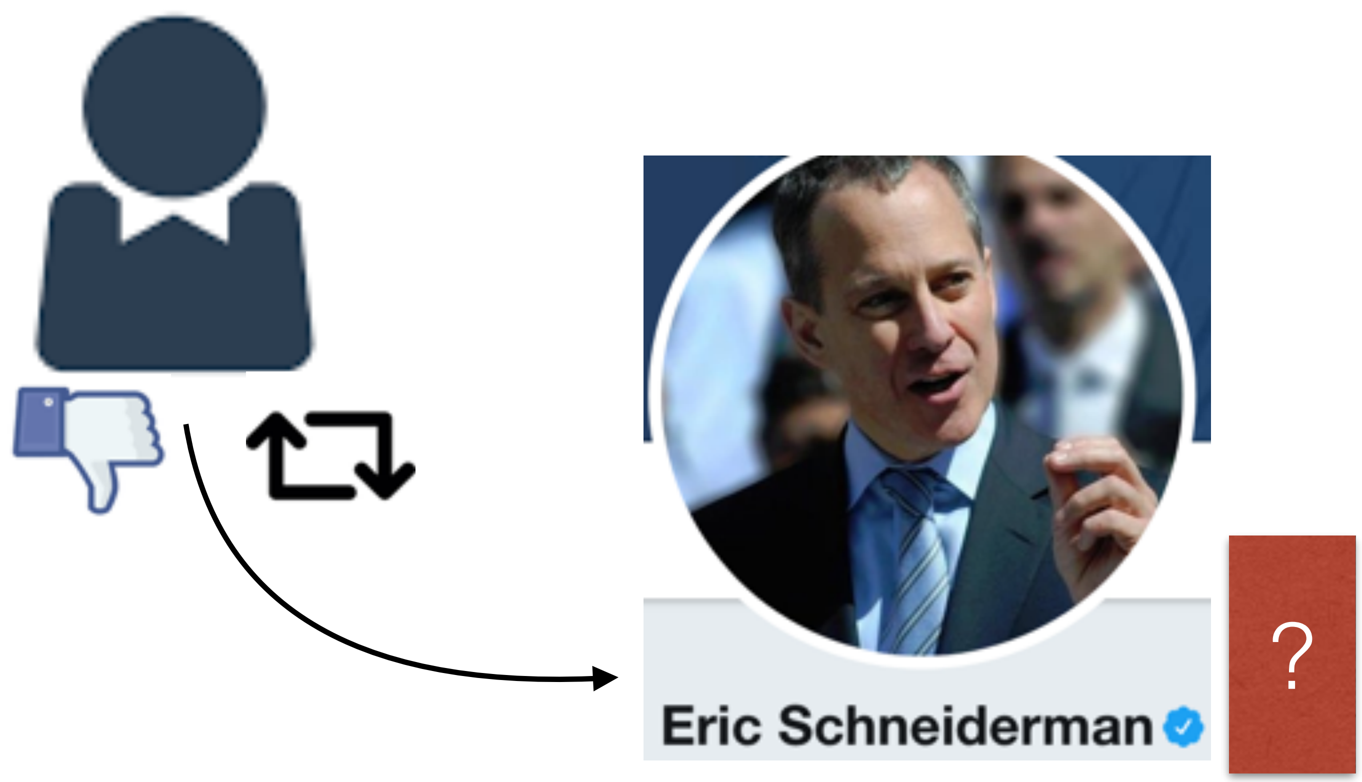
Defining feature functions with logic



$$\mathrm{Y}(A,B) \wedge \mathrm{X}(A,C) \wedge \mathrm{Y}(B,C) \rightarrow \mathrm{Y}(A,C)$$

Logic represents rich relationships

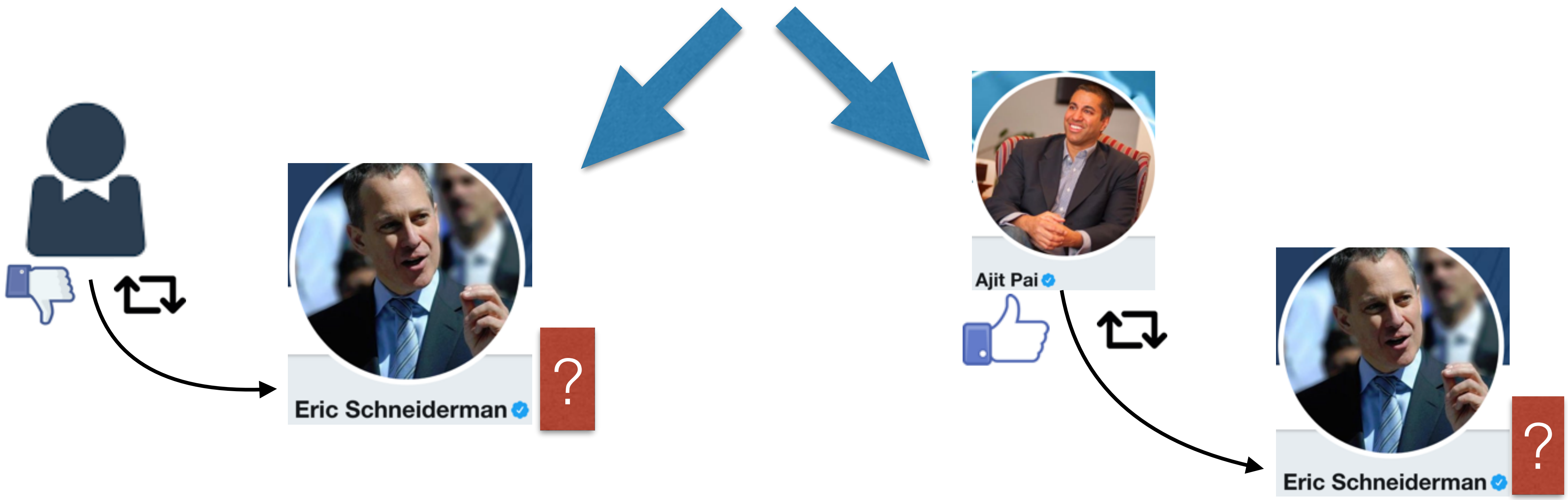
```
Retweets(U1, U2) & Side(U1, P) → Side(U2, P)
```



Logic is a powerful representation for capturing relationships and constraints

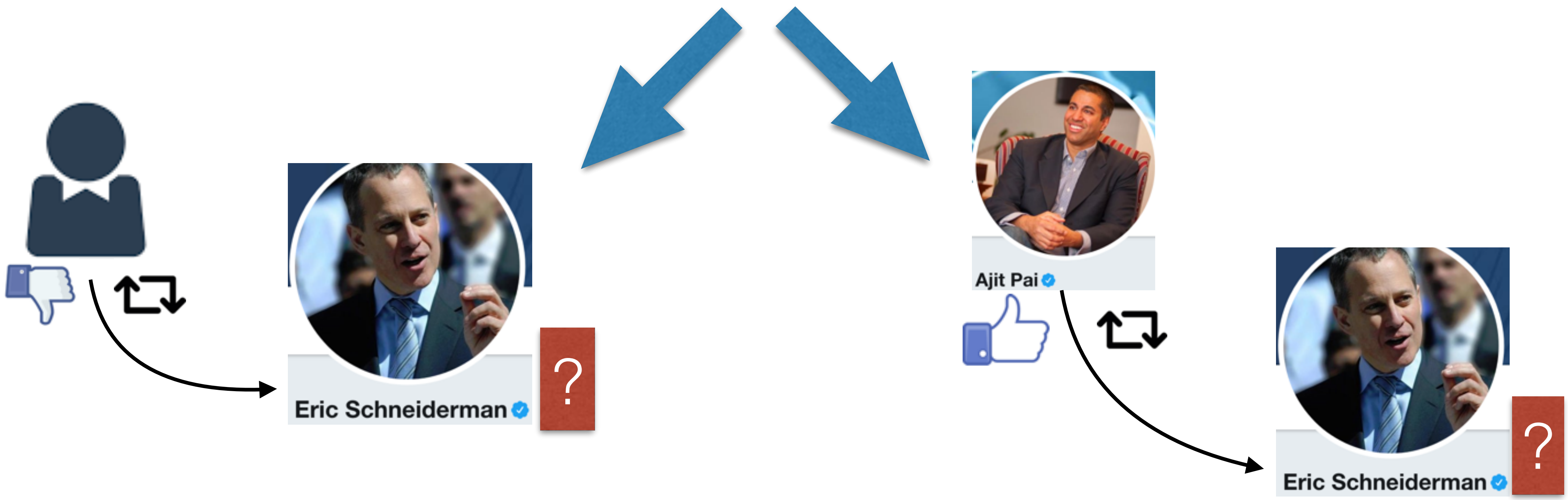
Logical satisfaction intractable

```
Retweets(U1, U2) & Side(U1, P) → Side(U2, P)
```



Logical satisfaction intractable

`Retweets(U1, U2) & Side(U1, P) → Side(U2, P)`



Problems with logic:

- Conflicts lead to infeasible states
- Combinatorial optimization doesn't scale

Probabilistic soft logic

$$\underbrace{\texttt{Retweets(U1, U2)}}_{\substack{[0,1] \\ R_{12}}} \texttt{ \& } \underbrace{\texttt{Side(U1, P)}}_{\substack{[0,1] \\ S_1}} \rightarrow \underbrace{\texttt{Side(U2, P)}}_{\substack{[0,1] \\ S_2}}$$

Bach et al. 2017, *Hinge-loss Markov random fields and Probabilistic Soft Logic*. In JMLR.

Probabilistic soft logic

$$\underbrace{\texttt{Retweets(U1, U2)}}_{\substack{[0,1] \\ R_{12}}} \texttt{ \& } \underbrace{\texttt{Side(U1, P)}}_{\substack{[0,1] \\ S_1}} \rightarrow \underbrace{\texttt{Side(U2, P)}}_{\substack{[0,1] \\ S_2}}$$

Given values, apply
relaxation of rule satisfaction

If satisfied, return 0
else penalty

Probabilistic soft logic

$$\underbrace{\texttt{Retweets(U1, U2)}}_{\substack{[0,1] \\ R_{12}}} \ \texttt{\&} \ \underbrace{\texttt{Side(U1, P)}}_{\substack{[0,1] \\ S_1}} \rightarrow \underbrace{\texttt{Side(U2, P)}}_{\substack{[0,1] \\ S_2}}$$

Given values, apply relaxation of rule satisfaction

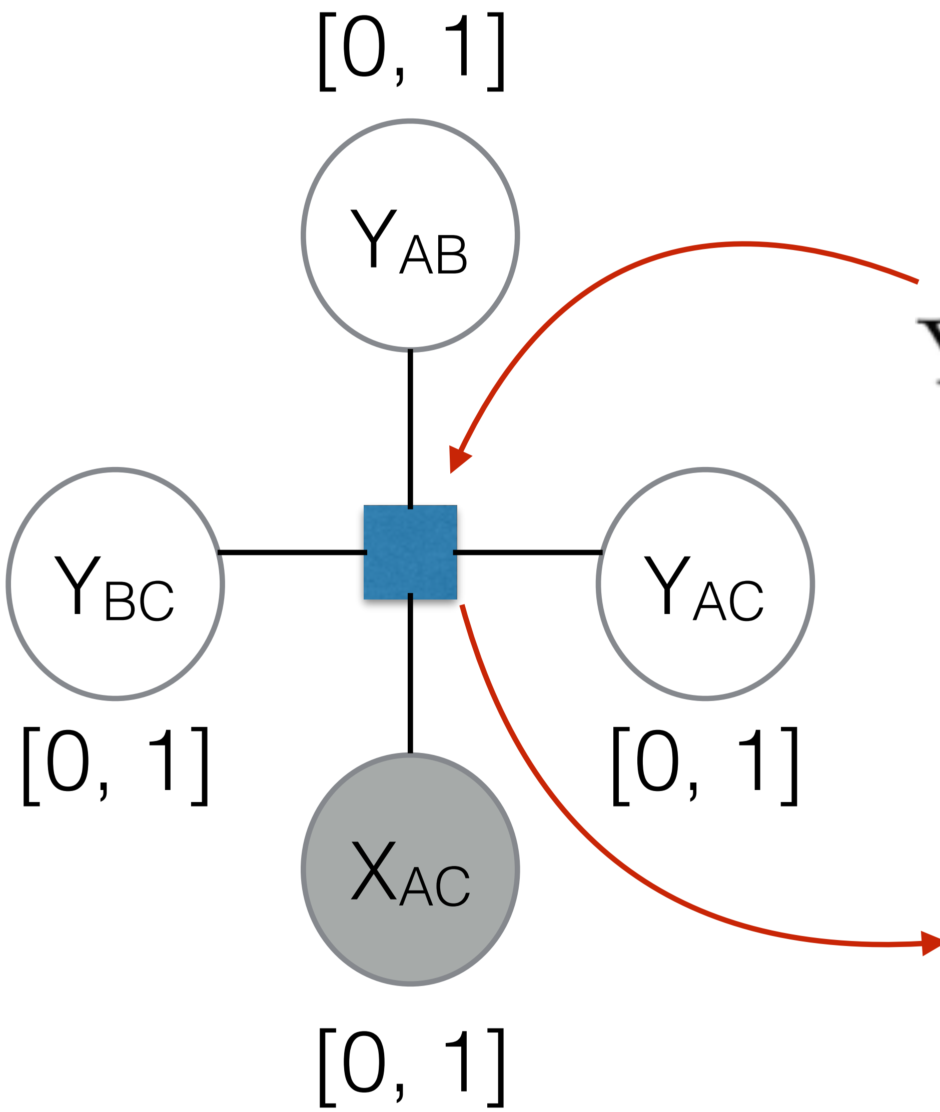
If satisfied, return 0 else penalty

$$\max\{f(R_{12}, S_1, S_2), 0\}$$

Linear function

Bach et al. 2017, *Hinge-loss Markov random fields and Probabilistic Soft Logic*. In JMLR.

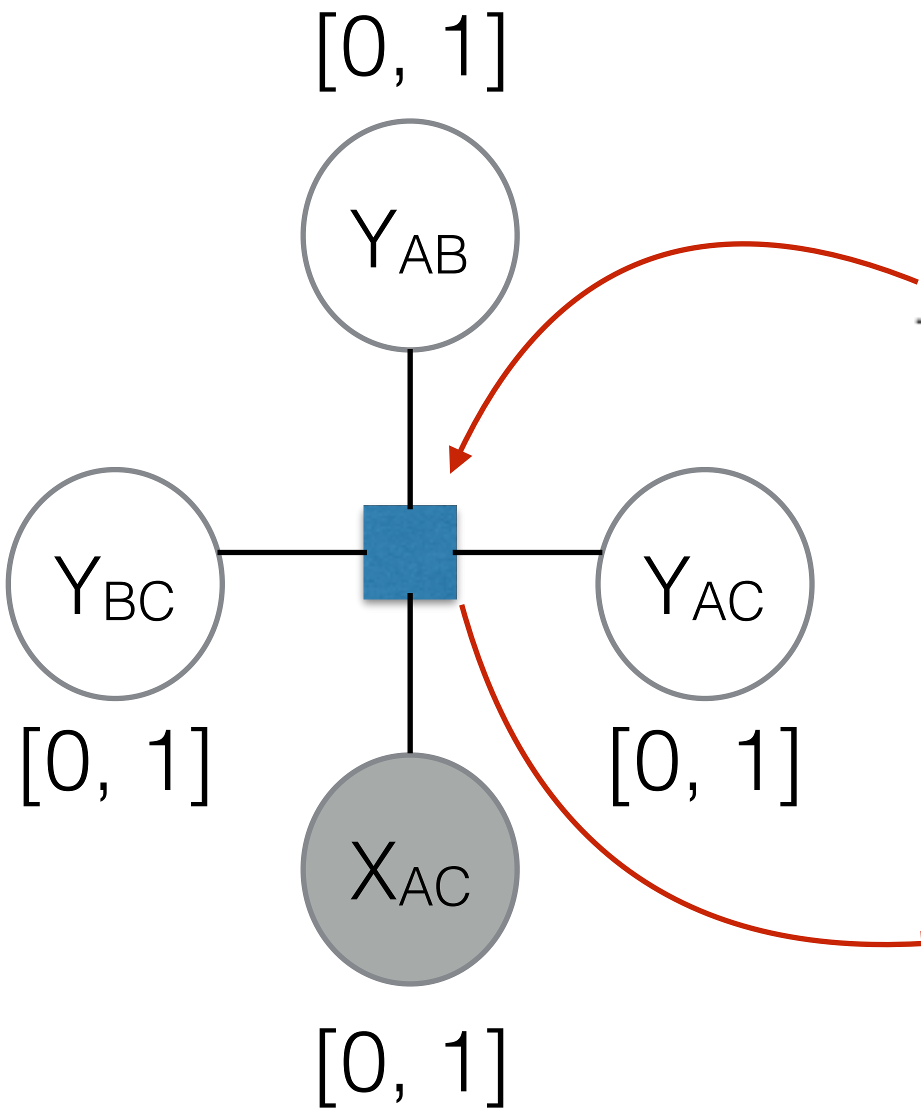
Defining feature functions with logic



$$\mathrm{Y}(A,B) \,\tilde{\wedge}\, \mathrm{X}(A,C) \,\tilde{\wedge}\, \mathrm{Y}(B,C) \,\tilde{\rightarrow}\, \mathrm{Y}(A,C)$$

Continuous penalty from
relaxed rule satisfaction

Defining feature functions with logic

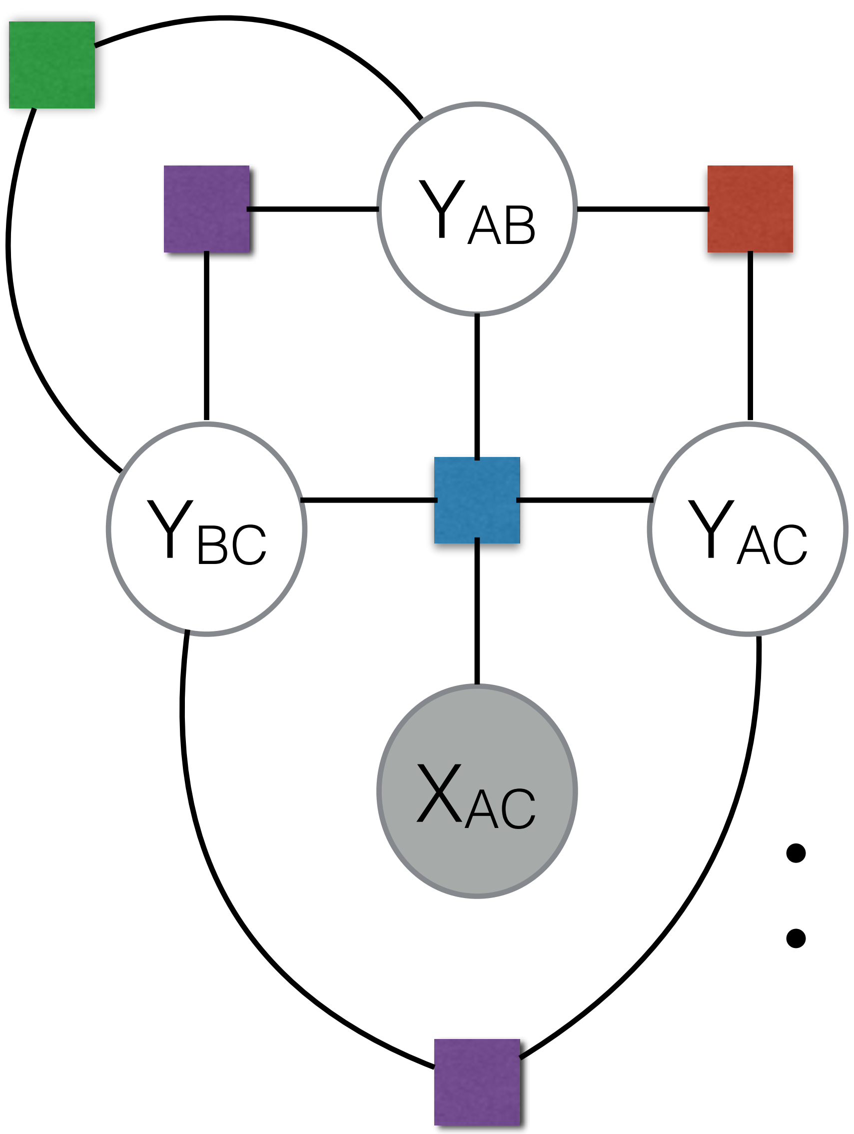


$$\mathrm{Y}(A,B)\ \tilde{\wedge}\ \mathrm{X}(A,C)\ \tilde{\wedge}\ \mathrm{Y}(B,C) \ \tilde{\rightarrow}\ \mathrm{Y}(A,C)$$

$$\max\{l(Y_{AB}, Y_{AC}, Y_{BC}, X_{AC}), 0\}$$

Hinge-loss Markov random field

$$P(\mathbf{Y}=\mathbf{y}|\mathbf{X}=\mathbf{x}) = \frac{1}{\mathcal{Z}} \exp\Big(-\sum_{r=1}^{M} w_r \left(\max\{l_r(\mathbf{y},\mathbf{x}),0\}\right)\Big)$$



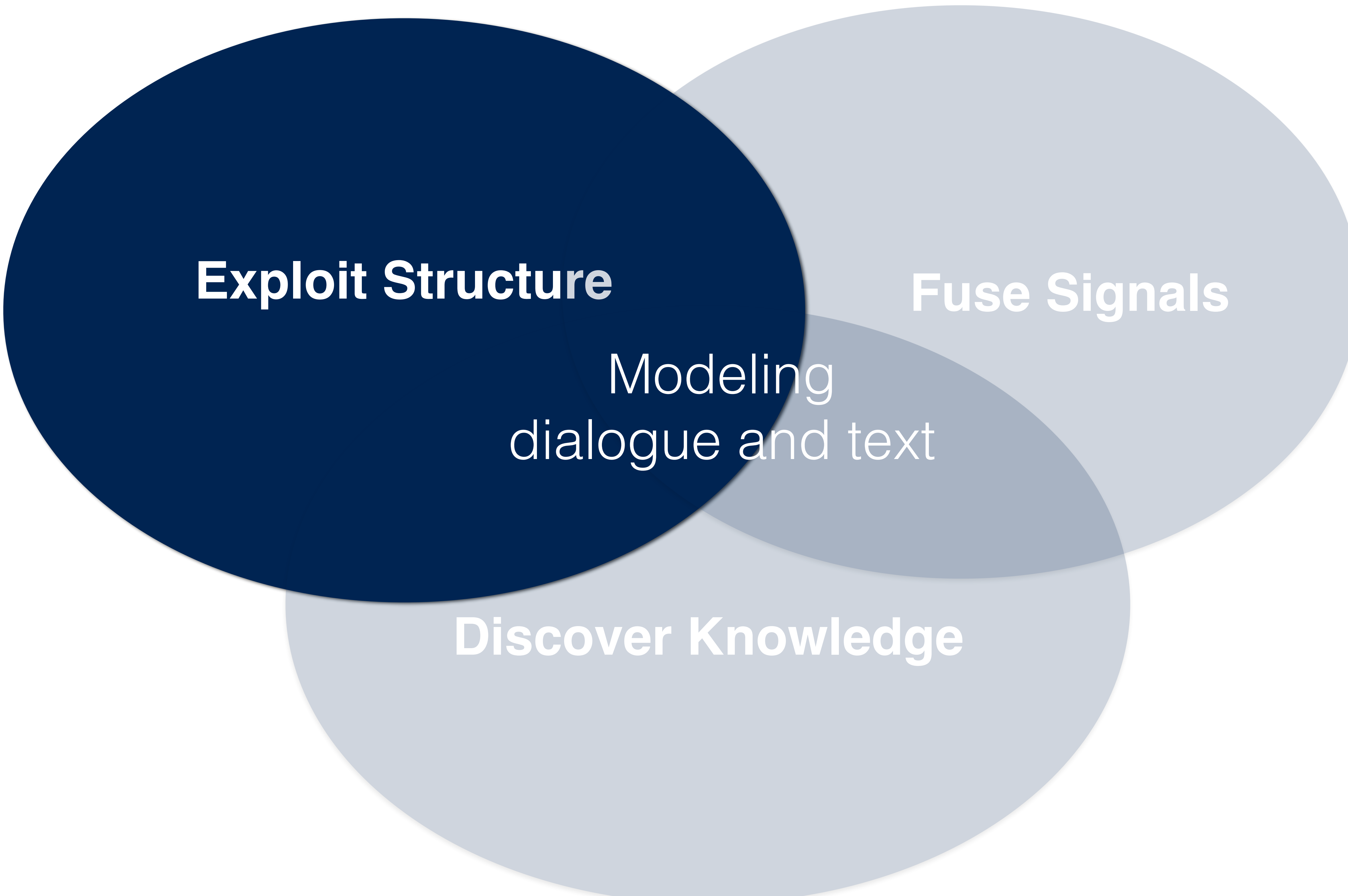
Convex MAP inference:

$$\arg\min_{\mathbf{y}\in[0,1]^n} \sum_{r=1}^{m} w_r \max\{l_r(\mathbf{y},\mathbf{x}),0\}$$

- Solved with ADMM message passing algorithm
- MLE and EM for learning weights from training data

Open-source software: psl.linqs.org

Roadmap of my talk



Understanding stances on issues

NATIONAL GEOGRAPHIC

Polar Bears Really Are Starving Because of Global Warming, Study Shows

[–] **filmfiend999** 38 points 2 hours ago
Why was this even up for debate? It sounds perfectly plausible if not totally obvious.
permalink embed save report reply

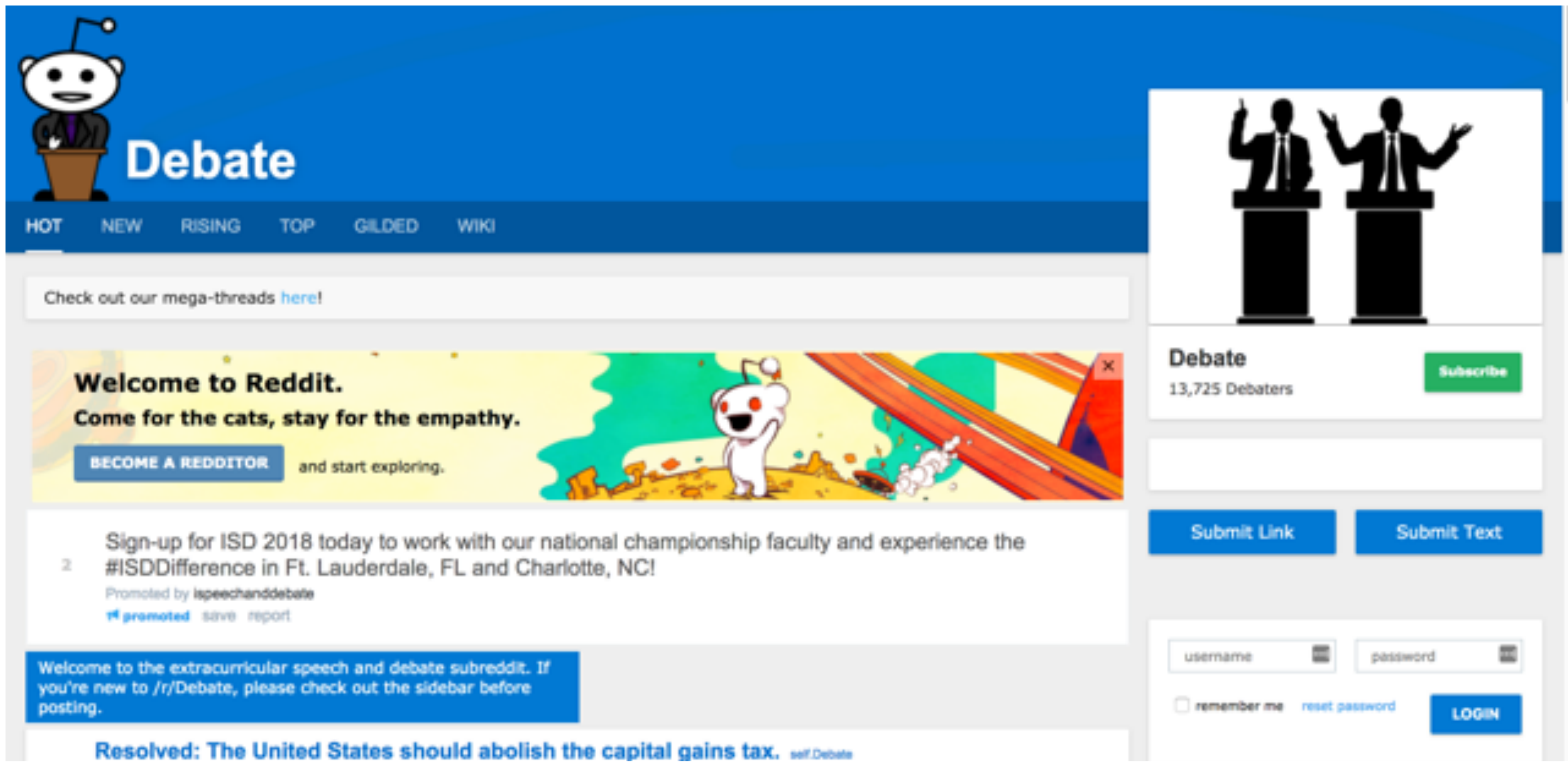
[–] **ContraPositive** 43 points an hour ago
Because US politics is full of climate change deniers.
permalink embed save parent report reply

[–] **yeetboy** 22 points an hour ago
Unfortunately, it's spreading. Ignorance breeds ignorance.
permalink embed save parent report reply

[–] **swiddie** [score hidden] an hour ago
Good ideas spread. Bad ideas die. Perhaps the idea that climate change isn't caused by burning oil is spreading because it is a good idea. Oil and gas and exhaust is natural, renewable, organic and healthy for the environment and economy.

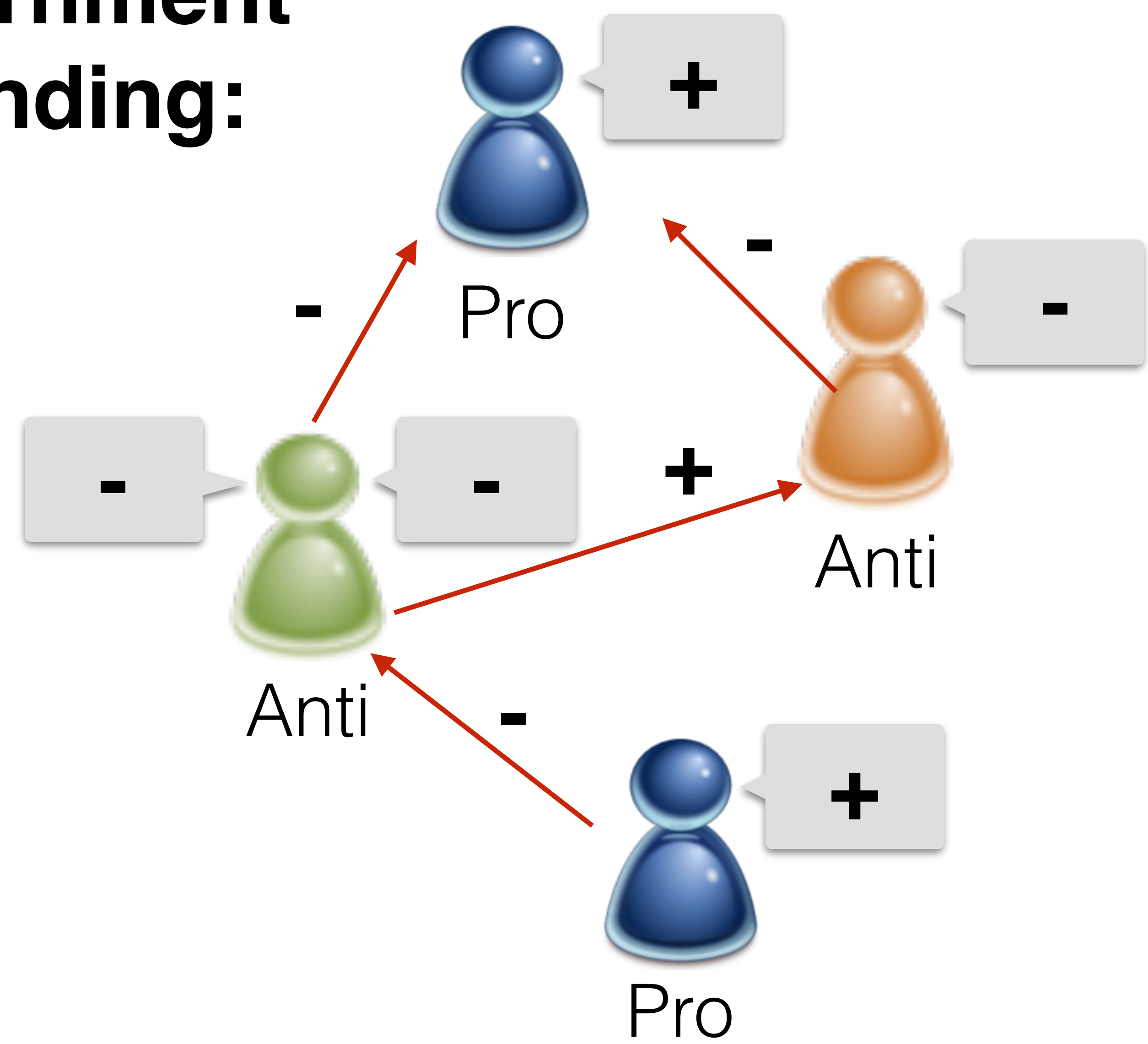
Identifying stances on issues is key to studying evolving ideologies, and biases.

Online debate forums provide rich dataset



Discussions in online debates

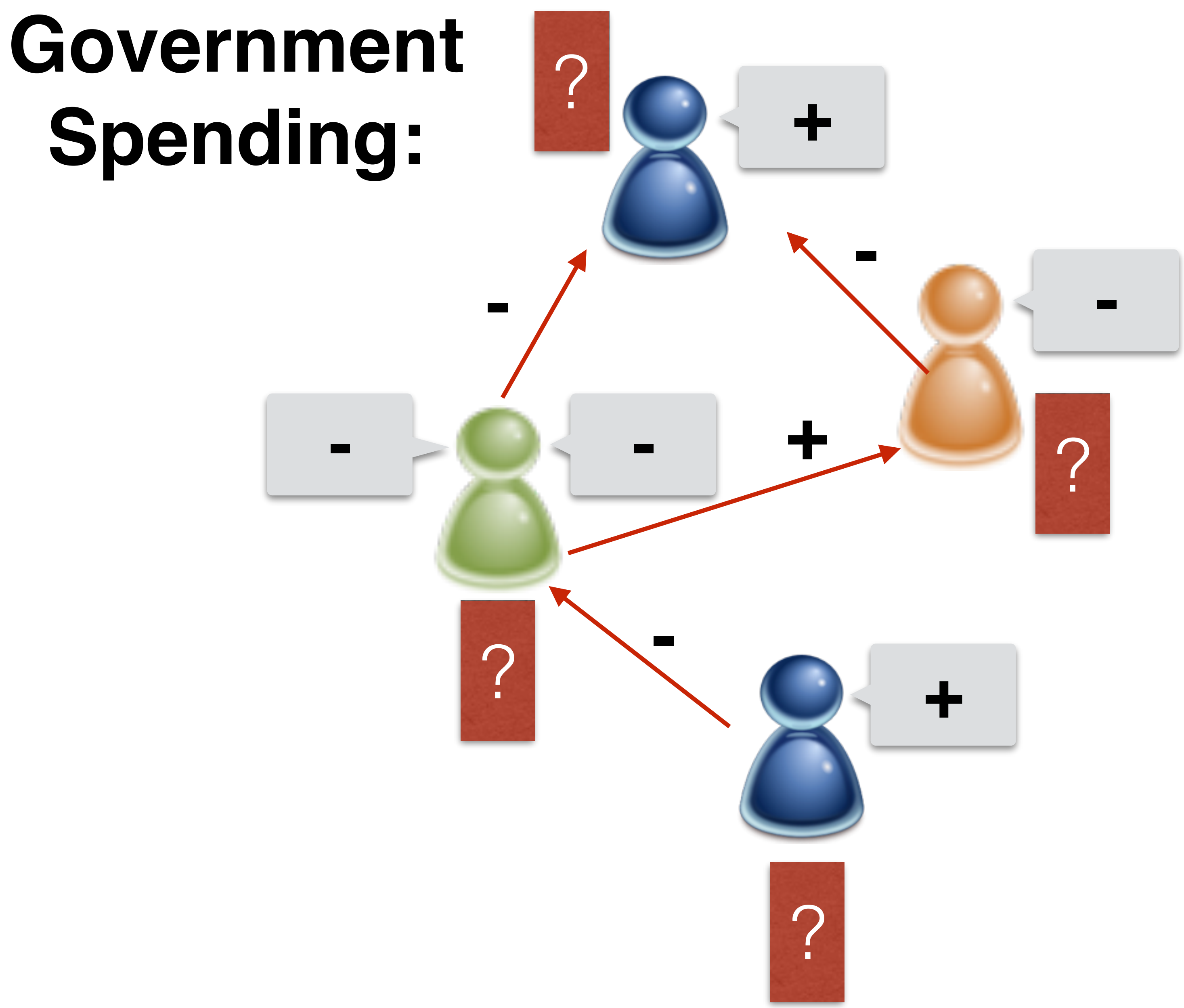
Government Spending:



Users initiate threads and reply to posts, signaling their position on issues and towards each other

Sridhar, Foulds, Huang, Getoor, Walker. 2015, Joint Models of Stance and Disagreement in Online Debates. ACL.

Inferring stances of users

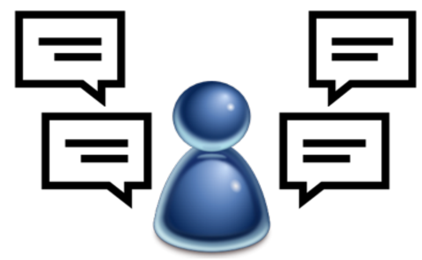


Supervised classification problem where labels are self-reported or annotated

Sridhar, Foulds, Huang, Getoor, Walker. 2015, Joint Models of Stance and Disagreement in Online Debates. ACL.

Important modeling questions

What is the right granularity to aggregate?



vs.

Author level | **Post level**

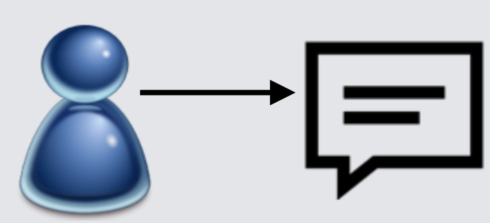
Features

w_1	w_2 ...	w_n
0	1	0

w_1	w_2 ...	w_n
1	1	1

Labels

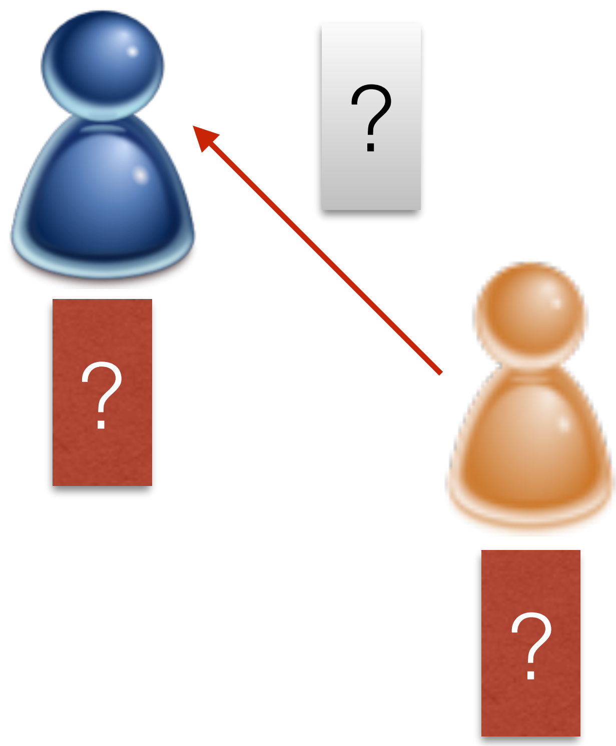
Majority()



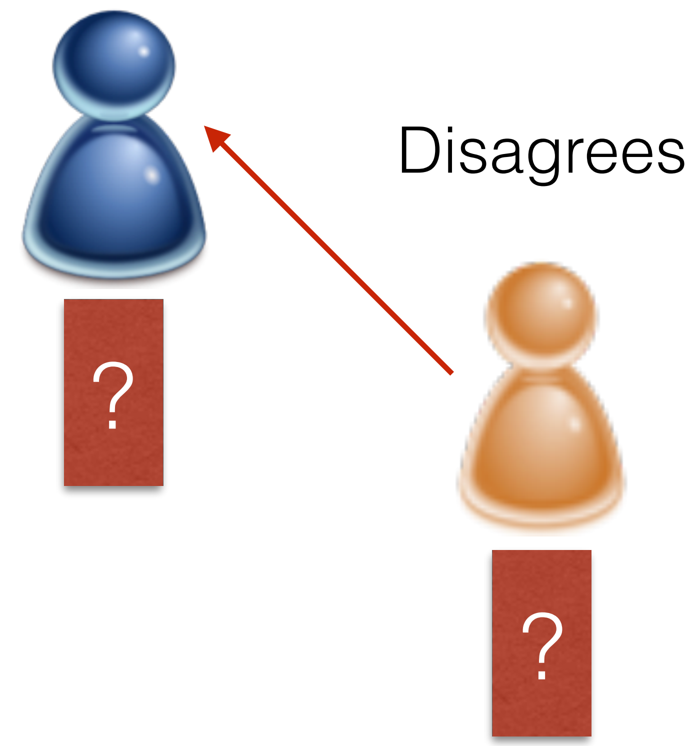
Hasan and Ng 2013, Extra-linguistic constraints on stance recognition in ideological debates. ACL.

Important modeling questions

How can we use context most effectively?



Model polarity of replies with stance



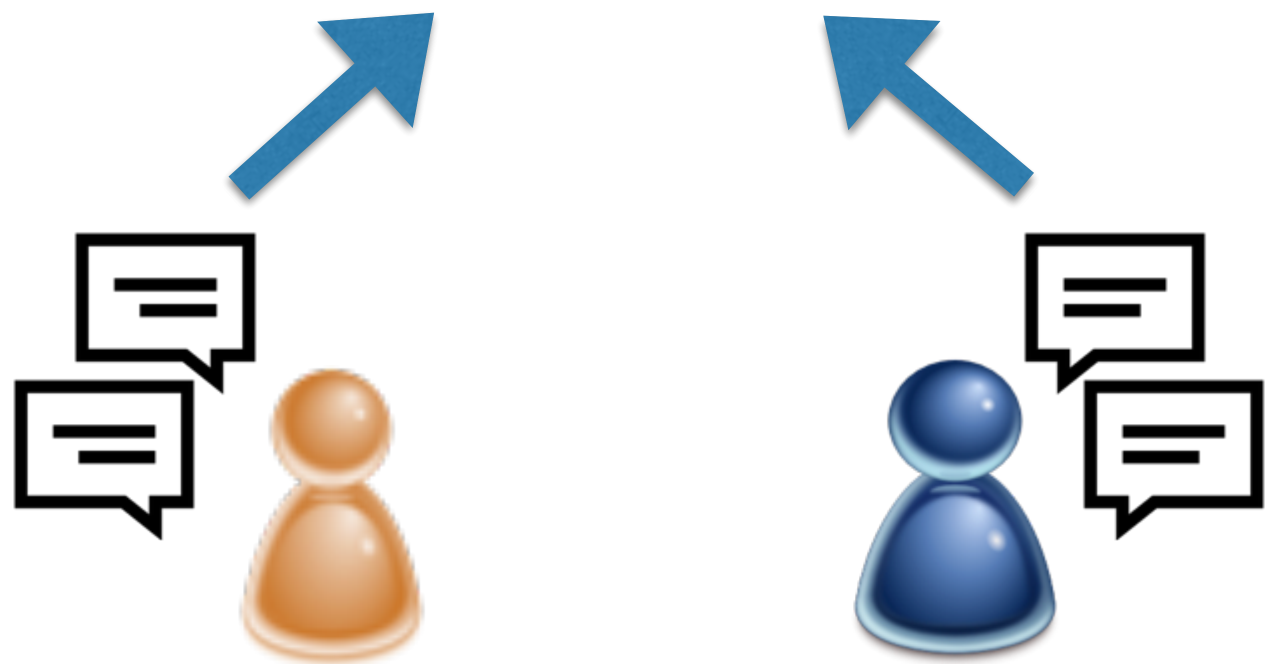
Treat replies as disagreement indicator

Hasan and Ng 2013, Extra-linguistic constraints on stance recognition in ideological debates. ACL.

Text classifiers as noisy local signals

$$\mathrm{LocalPro}(U) \rightarrow \mathrm{Pro}(U)$$

Logistic Regression



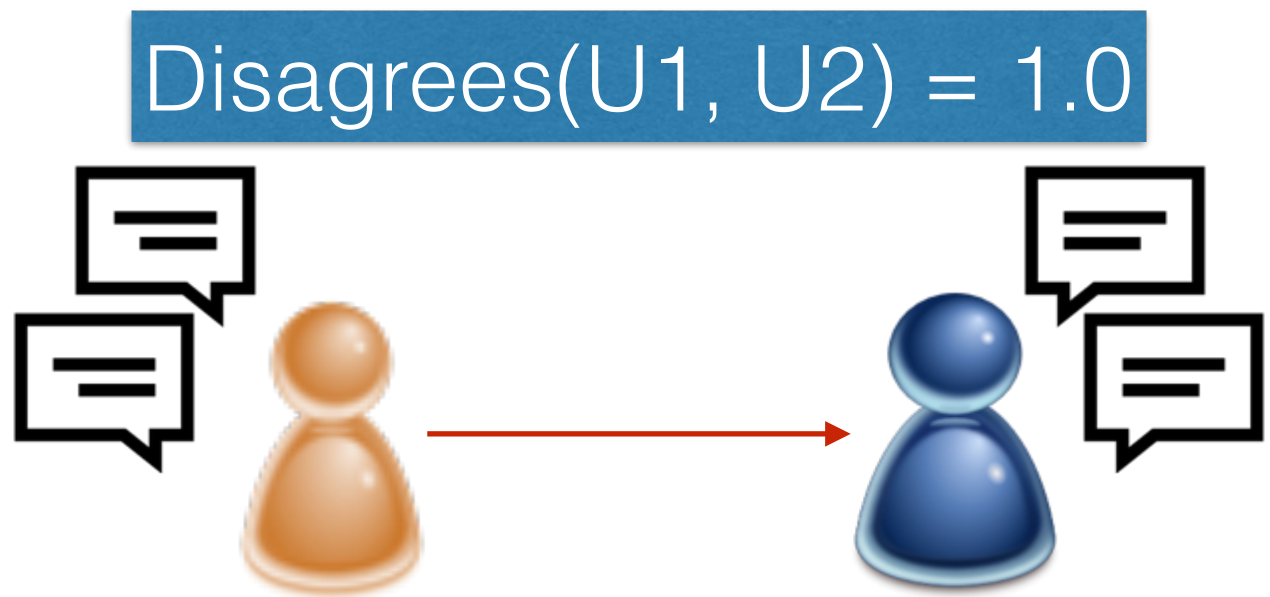
$Pr_{Local}(U_1=Pro) = 0.4$ $Pr_{Local}(U_2=Pro) = 0.6$

Sridhar, Foulds, Huang, Getoor, Walker. 2015, Joint Models of Stance and Disagreement in Online Debates. ACL.

Naive collective classification

$$\text{LocalPro}(U) \rightarrow \text{Pro}(U)$$

$$\text{Disagrees}(U_1, U_2) \wedge \text{Pro}(U_1) \rightarrow \neg\text{Pro}(U_2)$$



$Pr_{Local}(U_1\text{=Pro}) = 0.4$ $Pr_{Local}(U_2\text{=Pro}) = 0.6$

Sridhar, Foulds, Huang, Getoor, Walker. 2015, Joint Models of Stance and Disagreement in Online Debates. ACL.

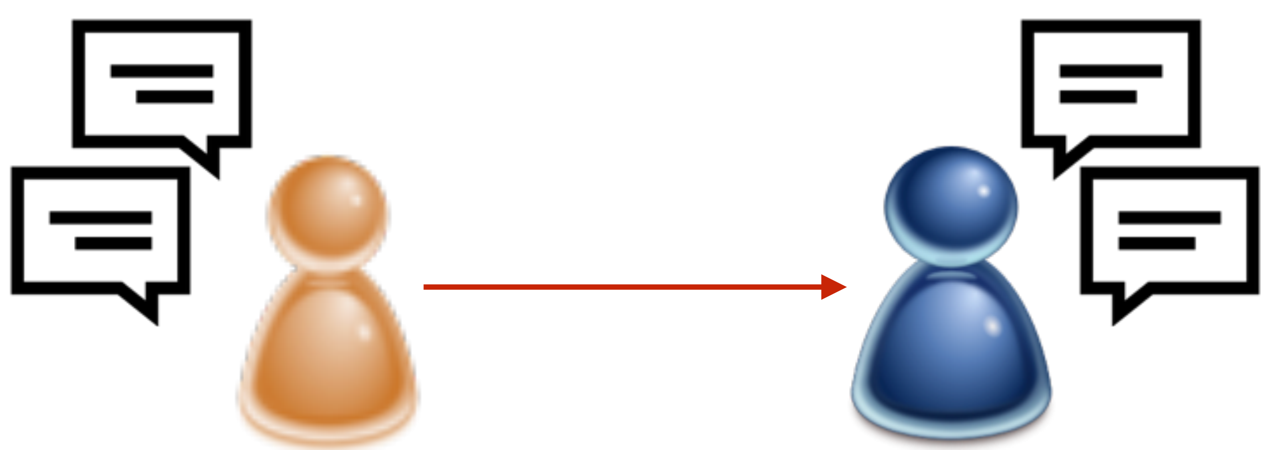
Jointly modeling stance and disagreement

$$\neg\text{Disagrees}(U_1, U_2) \wedge \text{Pro}(U_1) \rightarrow \text{Pro}(U_2)$$

$$\neg\text{Pro}(U_1) \wedge \text{Pro}(U_2) \rightarrow \text{Disagrees}(U_1, U_2)$$

$$\text{Pro}(U_1) \wedge \text{Pro}(U_2) \rightarrow \neg\text{Disagrees}(U_1, U_2)$$

$\text{Pr}_{\text{Local}}(U_1,U_2{=}\text{Dis}) = 0.3$



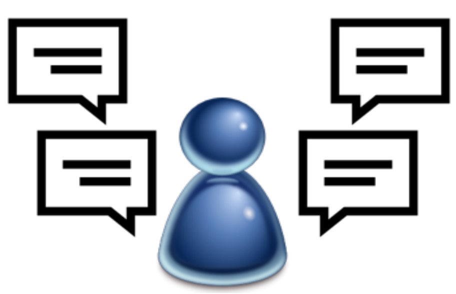
$\text{Pr}_{\text{Local}}(U_1{=}\text{Pro}) = 0.4$ $\text{Pr}_{\text{Local}}(U_2{=}\text{Pro}) = 0.6$

Evaluating on two debate forums

4Forums CreateDebate

- 4 topics
- 300 users
- 4 to 19 posts per user

Compare models **local, collective, joint** along with:



vs.

Author Level

- Aggregate features
- Majority post label

Post Level

- Separate features
- Apply author's label

Experimental highlights

Accuracy on Evolution topic in 4Forums

	Post Stance	Author Stance
Logistic Regression	73.0	77.3
Simple Collective	68.3	74.4
Joint-Author	**80.3**	**78.7**
Joint-Post	73.9	76.7

Granularity of aggregating information has ramifications

Experimental highlights

Accuracy on Evolution topic in 4Forums

	Post Stance	Author Stance
Logistic Regression	73.0	77.3
Simple Collective	**68.3**	**74.4**
Joint-Author	**80.3**	**78.7**
Joint-Post	73.9	76.7

Naive assumptions are harmful — simple collective model worse than baseline in nuanced topic

Benefits of joint modeling

Text	Stance
Post: I agree with everything except the last part. Safe gun storage is very important, and sensible storage requirements have two important factors.	ANTI
Reply: I can agree with this. And in case it seemed otherwise, I know full well how to store guns safely, and why it's necessary. My point was that I don't like the idea of such a law, especially when you consider the problem of enforcement.	ANTI

Post-reply pair whose stance is correctly predicted by joint but not collective

Sridhar, Foulds, Huang, Getoor, Walker. 2015, Joint Models of Stance and Disagreement in Online Debates. ACL.

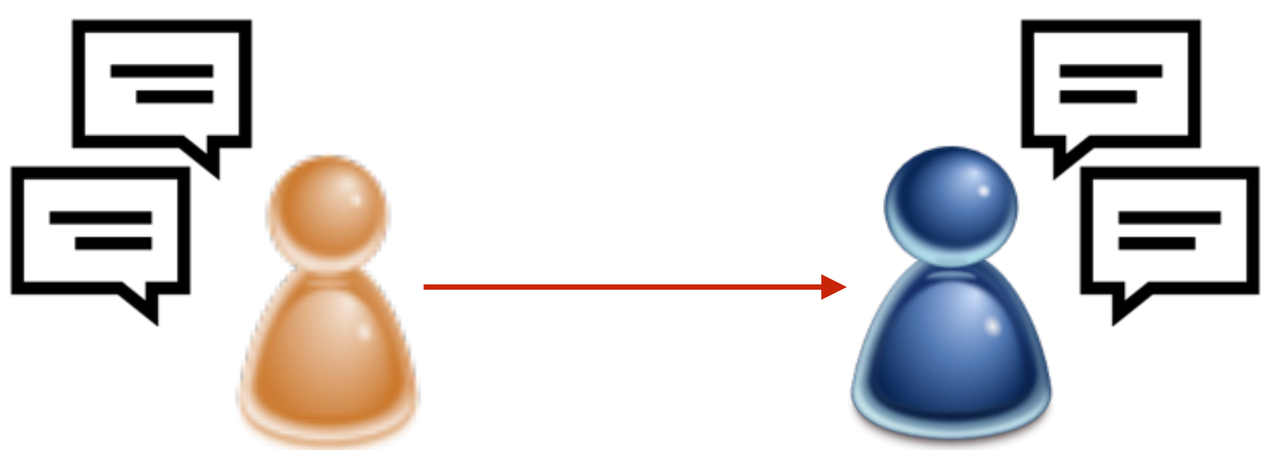
Takeaway

$$\neg\mathrm{Disagrees}(U_1, U_2) \wedge \mathrm{Pro}(U_1) \rightarrow \mathrm{Pro}(U_2)$$

$$\neg\mathrm{Pro}(U_1) \wedge \mathrm{Pro}(U_2) \rightarrow \mathrm{Disagrees}(U_1, U_2)$$

$$\mathrm{Pro}(U_1) \wedge \mathrm{Pro}(U_2) \rightarrow \neg\mathrm{Disagrees}(U_1, U_2)$$

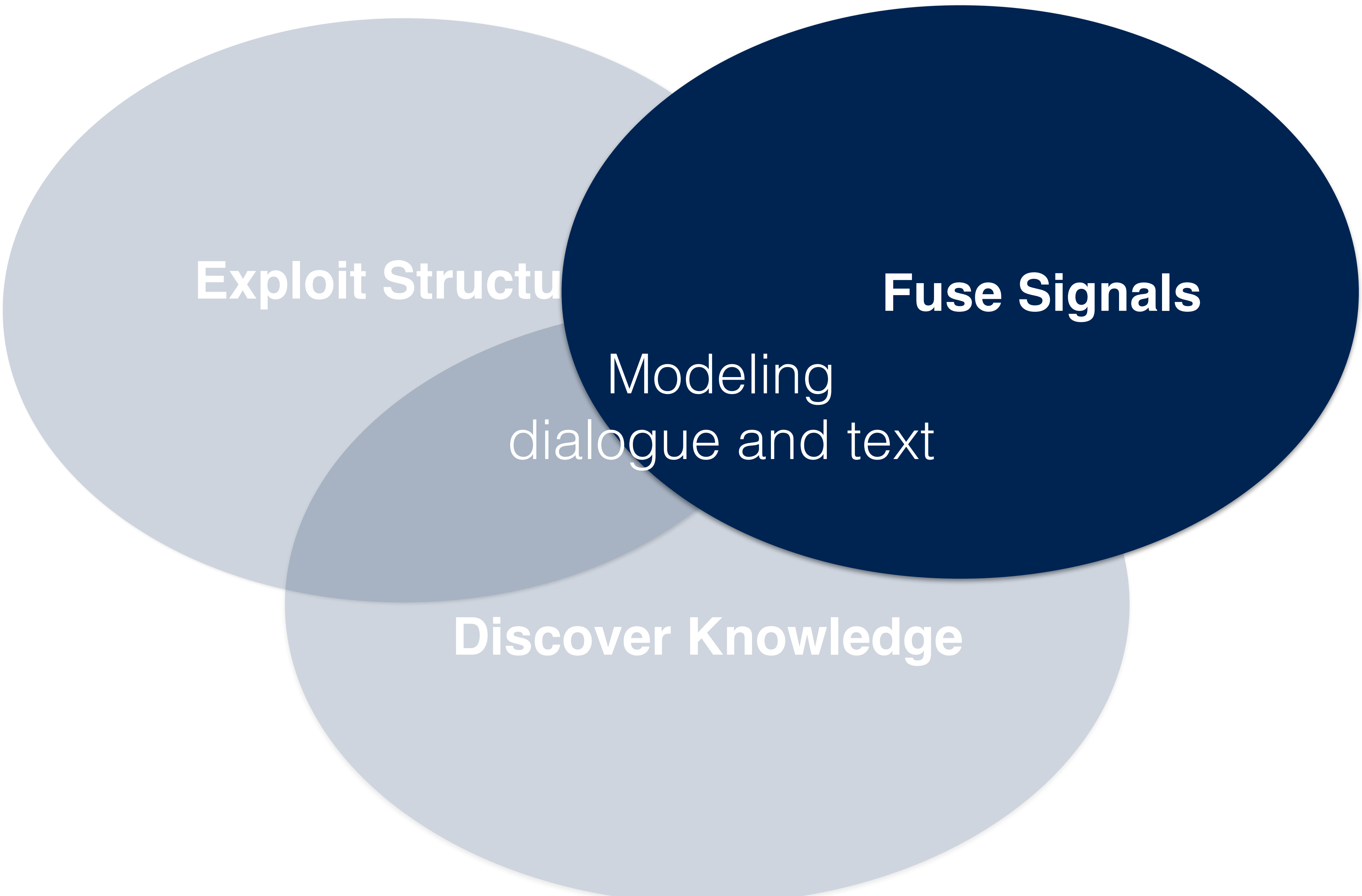
$\mathrm{Pr}_{\mathrm{Local}}(\mathrm{U}_1,\mathrm{U}_2\mathrm{=Dis}) = 0.3$



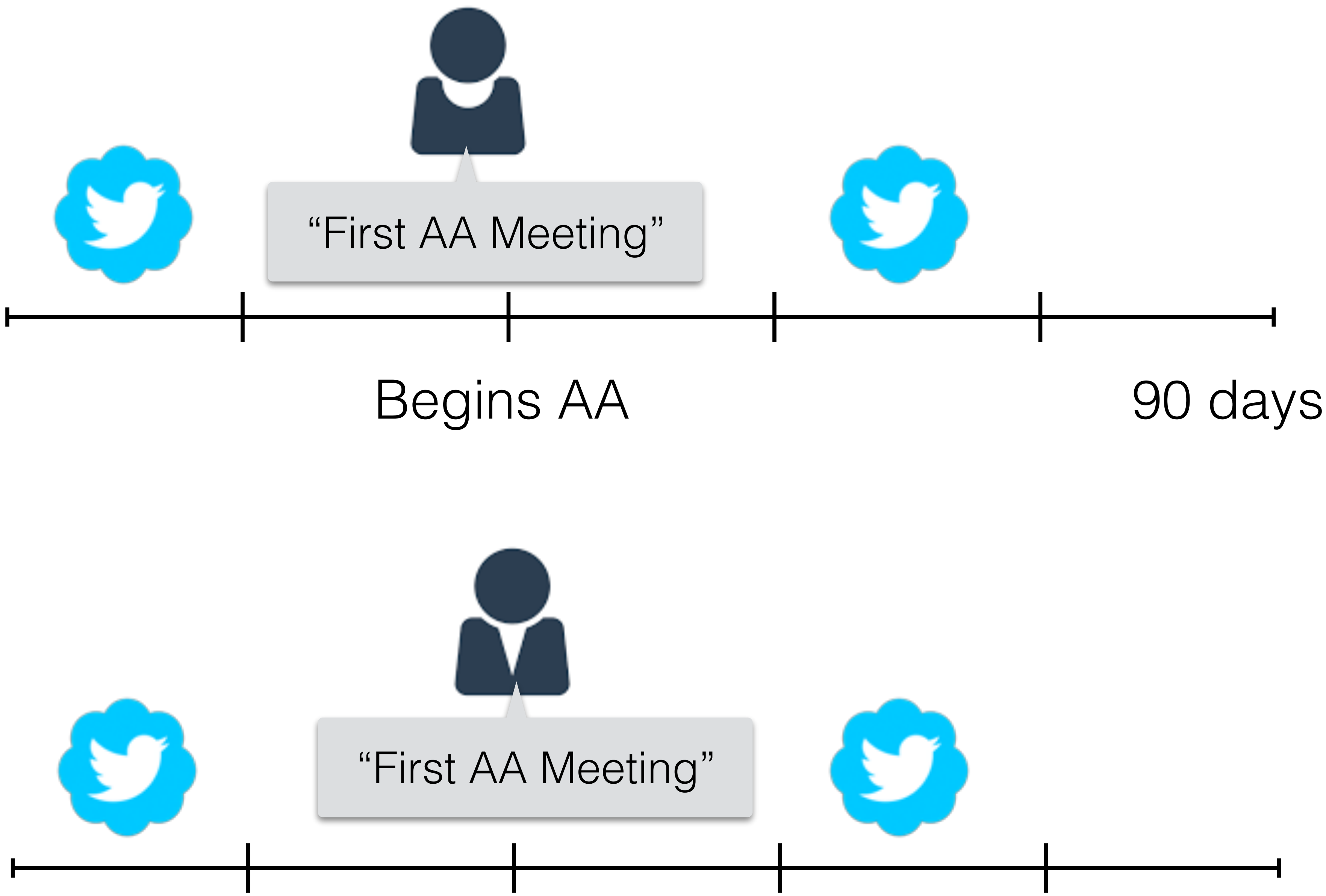
Using (dis)similarity to inform predictions is powerful, general template

Sridhar, Foulds, Huang, Getoor, Walker. 2015, Joint Models of Stance and Disagreement in Online Debates. ACL.

Roadmap of my talk



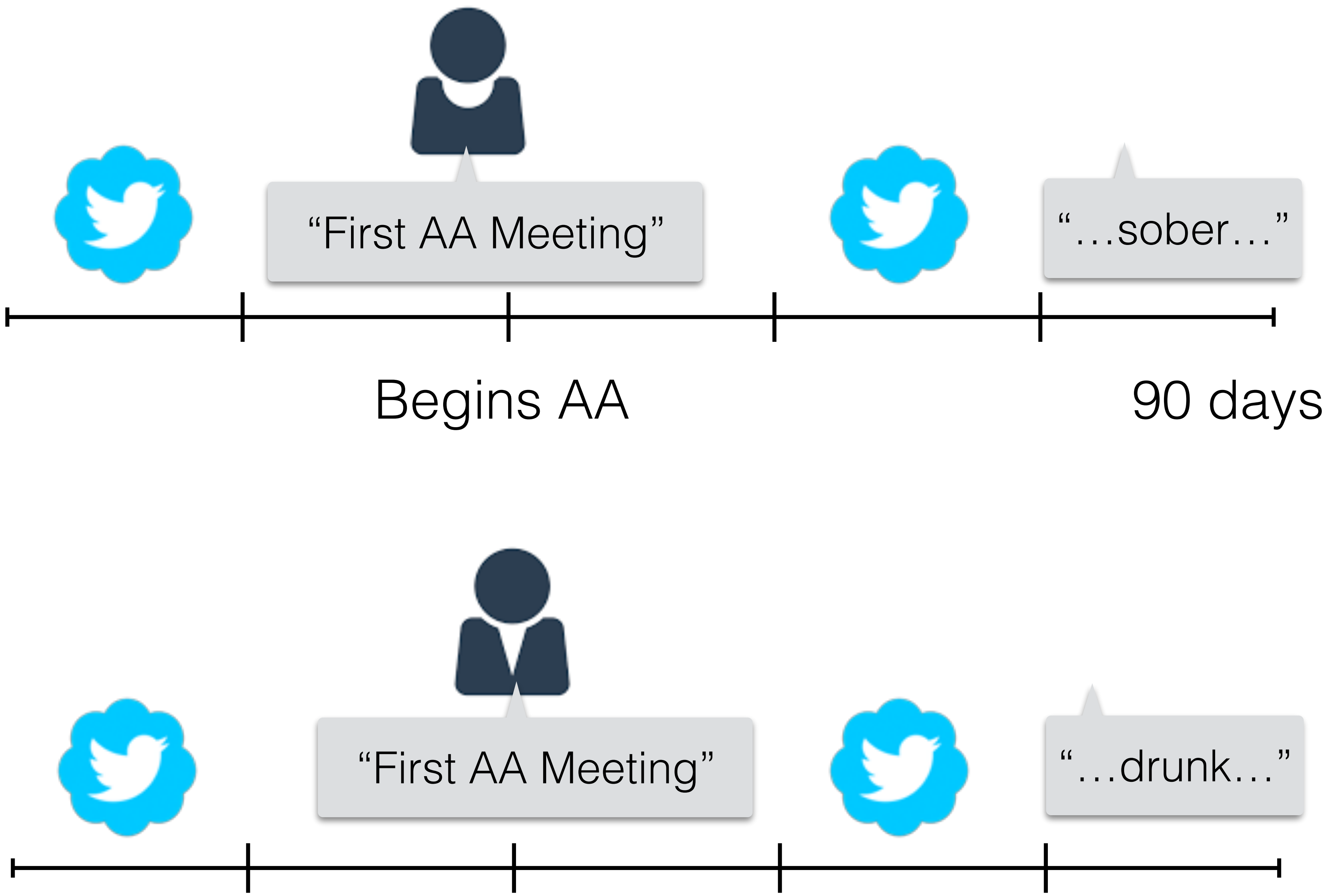
Detecting alcoholism relapse from Twitter



Identify tweets mentioning AA meeting, collecting tweets before and after

Zhang, Ramesh, Golbeck, Sridhar, Getoor, 2018, A Structured Approach to Understanding Recovery and Relapse in AA. WWW.

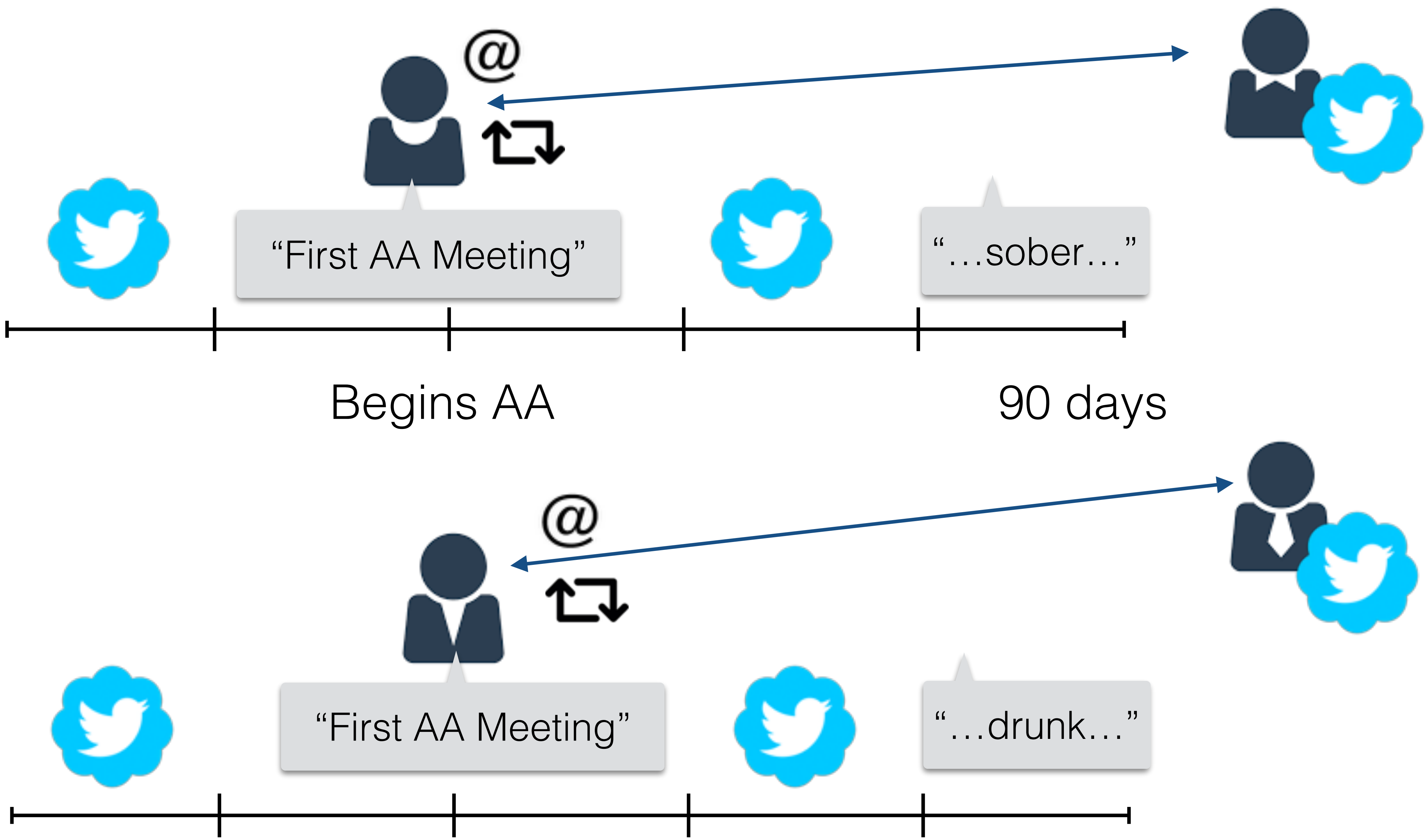
Detecting alcoholism relapse from Twitter



Label relapse using clear indicators after 90 days

Zhang, Ramesh, Golbeck, Sridhar, Getoor, 2018, A Structured Approach to Understanding Recovery and Relapse in AA. WWW.

Detecting alcoholism relapse from Twitter

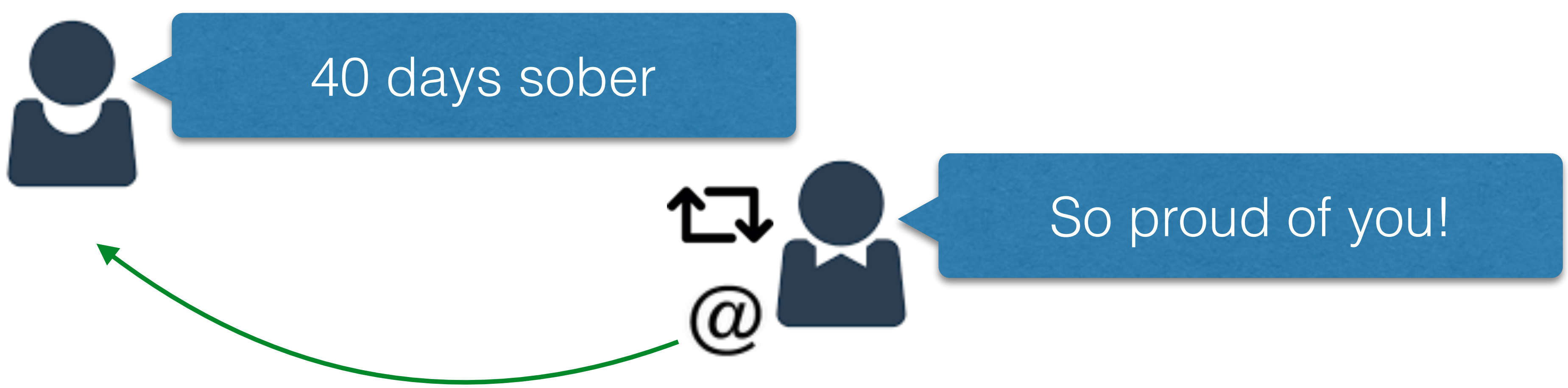


Tweets from friends that users mention and retweet

Zhang, Ramesh, Golbeck, Sridhar, Getoor, 2018, A Structured Approach to Understanding Recovery and Relapse in AA. WWW.

Capturing supportive friend behavior

Negative interaction



Positive interaction

Combine multiple language signals

I feel like getting drunk

Sobriety sucks

40 days sober

Glad I’m not drinking

First dry week successful

Combine multiple language signals

I feel like getting drunk

Sobriety sucks

40 days sober

Glad I'm not drinking

First dry week successful

Alcohol/sober word dictionary

- UsesAlcoholWord(User)
- UsesSoberWord(User)

Combine multiple language signals

LIWC and Sentiwordnet for affect

- PosAffect(User)
- PosSentiment(Tweet)

Combine multiple language signals

I feel like getting drunk

Sobriety sucks

40 days sober

Glad I'm not drinking

First dry week successful

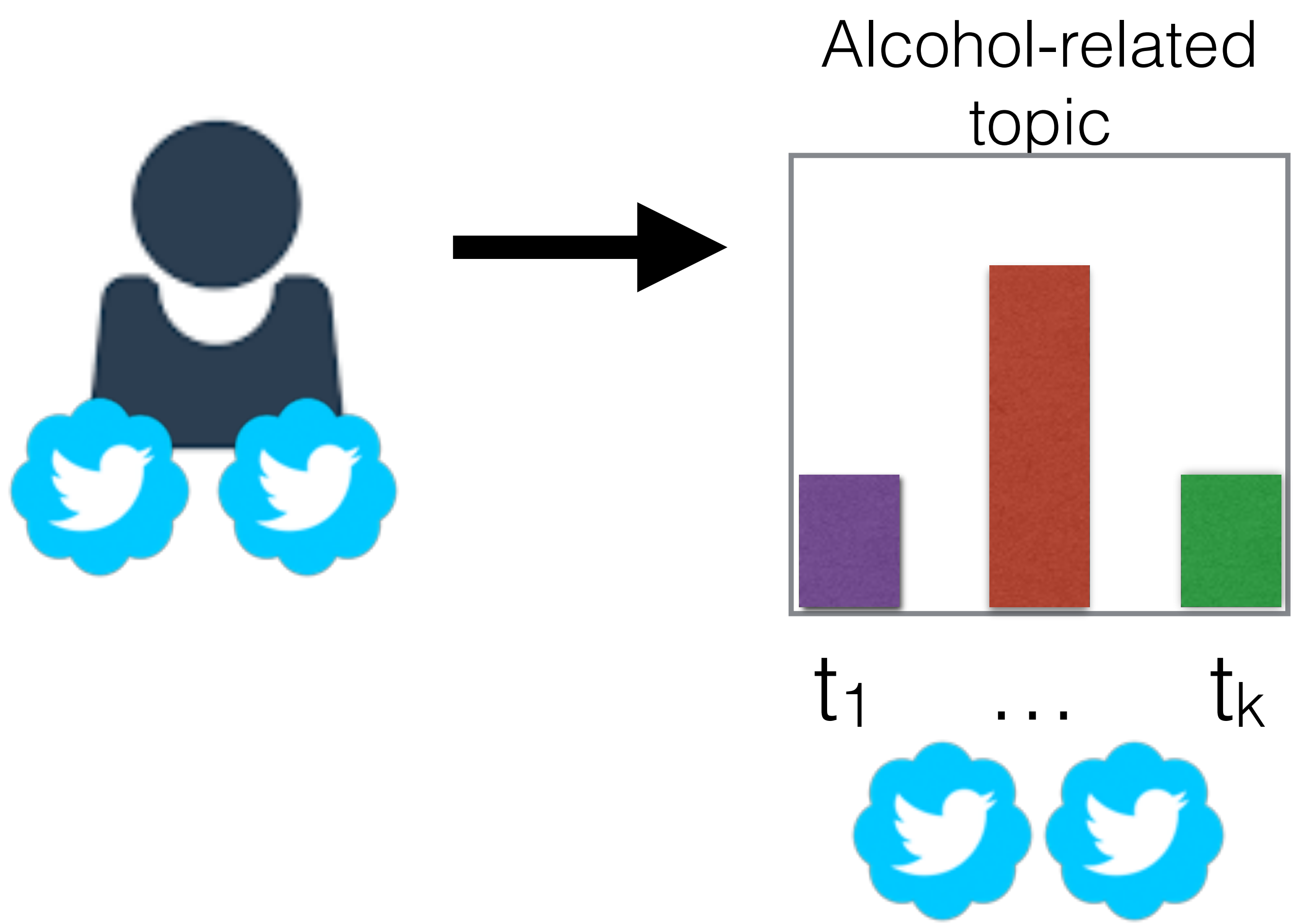
Seeded LDA with alcohol/sober words

- TweetTopic(Tweet, LatentTopic)
- UserTopic(User, LatentTopic)

Jagarlamudi et. 2012, Incorporating lexical priors into topic models. EACL.

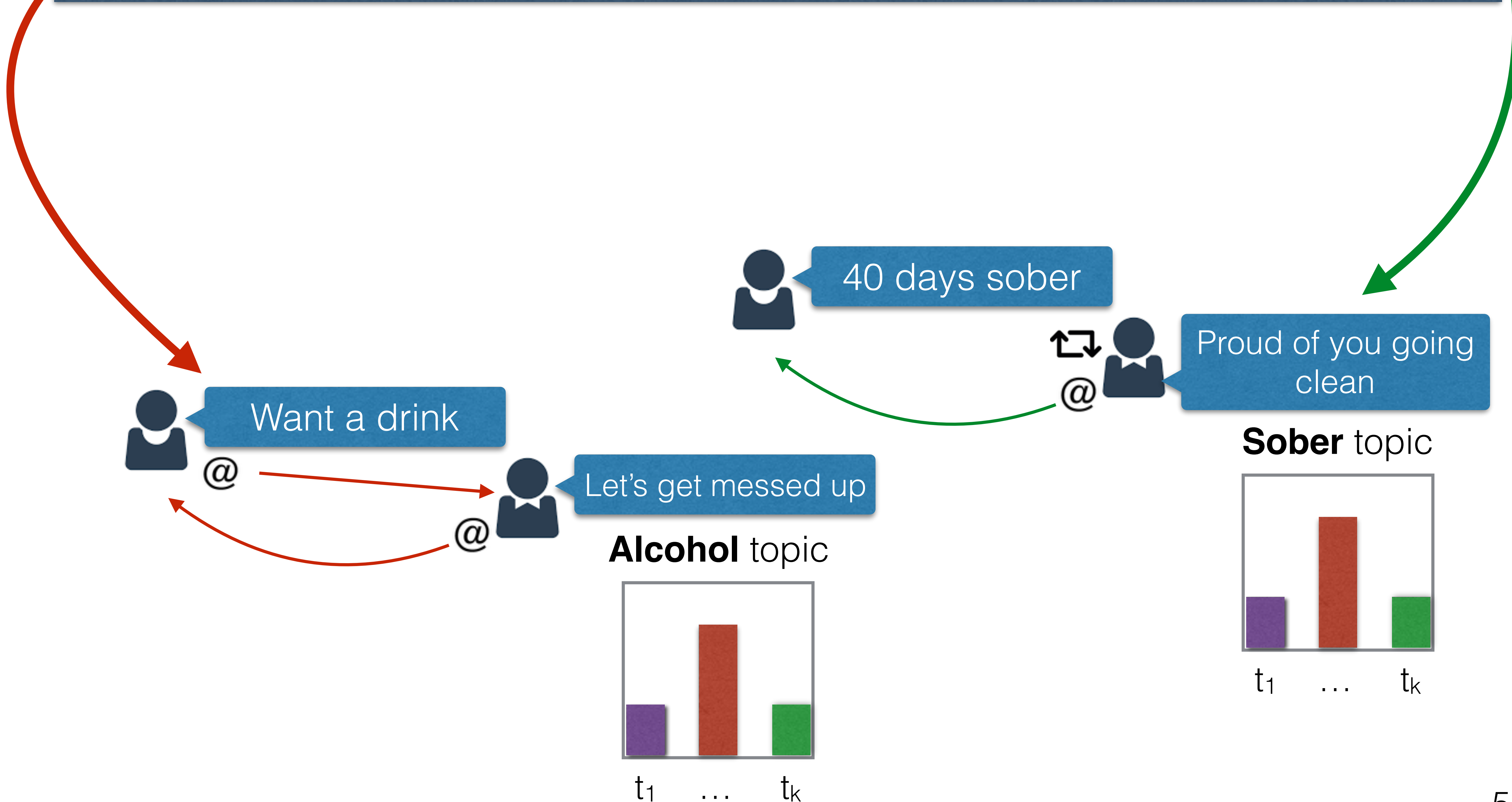
Local text signals from tweets

```
UserTopic(U, alcohol) → !Recovers(U)
UserTopic(U, sober) → Recovers(U)
```



Modeling interactions with friends

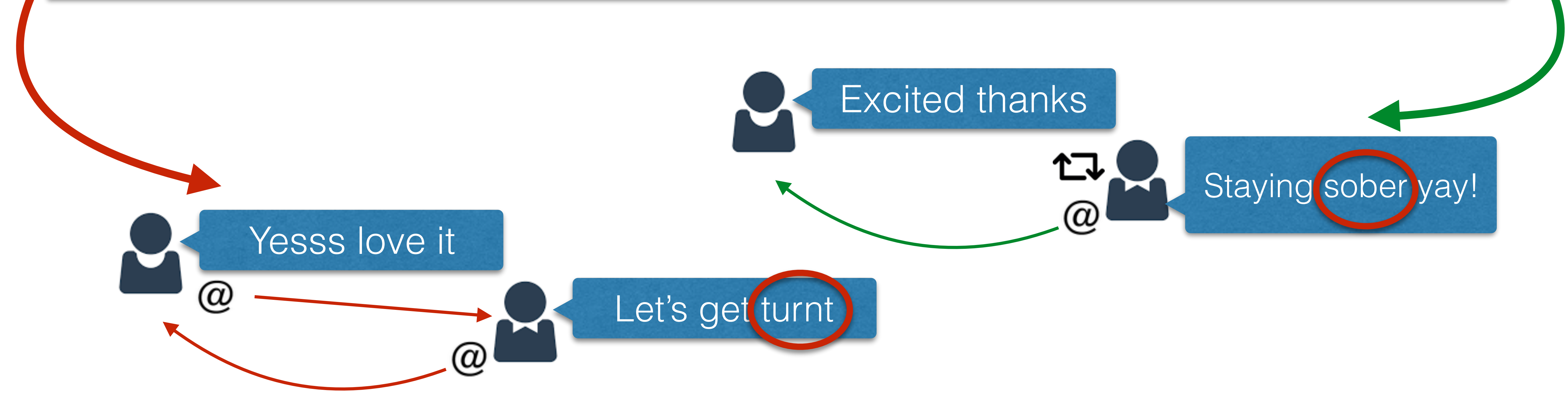
```
Retweets(U2, U1, T) & TweetTopic(T, sober) & PosSent(T) → Recovers(U1)
Replies(U2, U1, T) & TweetTopic(T, alcohol) & PosSent(T) → !Recovers(U1)
```



Modeling interactions with friends

```
Retweets(U2, U1, T) & TweetTopic(T, sober) & PosSent(T) → Recovers(U1)
Replies(U2, U1, T) & TweetTopic(T, alcohol) & PosSent(T) → !Recovers(U1)
```

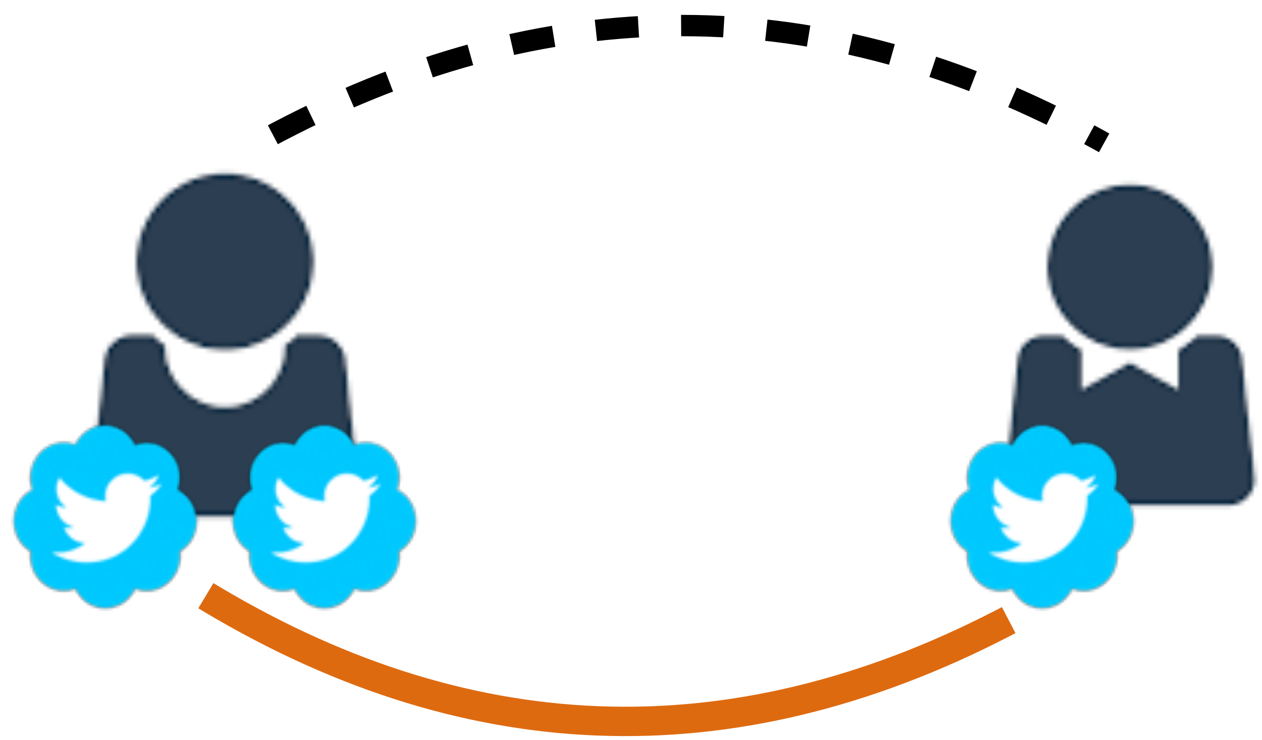
```
Friends(U2, U1) & UsesSoberWord(U2) & PosAffect(U1) → Recovers(U1)
Friends(U2, U1) & UsesAlcoholWord(U2) & PosAffect(U1) → !Recovers(U1)
```



Collective inference with similarities

```
TweetSimilarity(U1, U2) & Recovers(U2) → Recovers(U1)
TweetSimilarity(U1, U2) & !Recovers(U2) → !Recovers(U1)
```

Similar behavior



Cosine similarity of tweets

Advantages of combined approach

Method	AUC-PR	AUC-ROC
Logistic Regression	0.58	0.71
PSL approach	**0.75**	**0.90**

Outperforms text baseline for predicting relapse after 90 days

Advantages of combined approach

@... drink your **beer** snort your gear.

RT @...: I need **vodka**.

@... it's okay cause we were **drunk** everyday. What're you plans for the day?!

Captures real examples of **enabling** behavior

Advantages of combined approach

@... I struggled with holidays in early **sobriety** I had a plan. Go to a meeting, call my sponsor or have coffee with a **sober** friend

@... Do you need a **sober** companion? We're here for you.

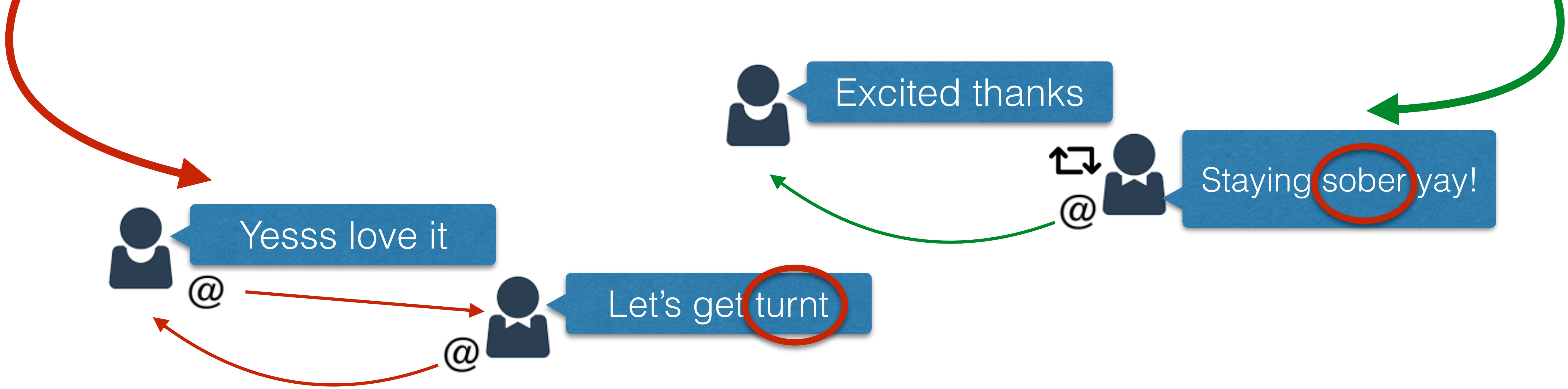
RT @...: Tips for the **sober** beginner! I contributed to @XXX's blog, which is run by the UK nonprofit

Captures real examples of **supportive** behavior

Takeaway

```
Retweets(U2, U1, T) & TweetTopic(T, sober) & PosSent(T) → Recovers(U1)
Replies(U2, U1, T) & TweetTopic(T, alcohol) & PosSent(T) → !Recovers(U1)
```

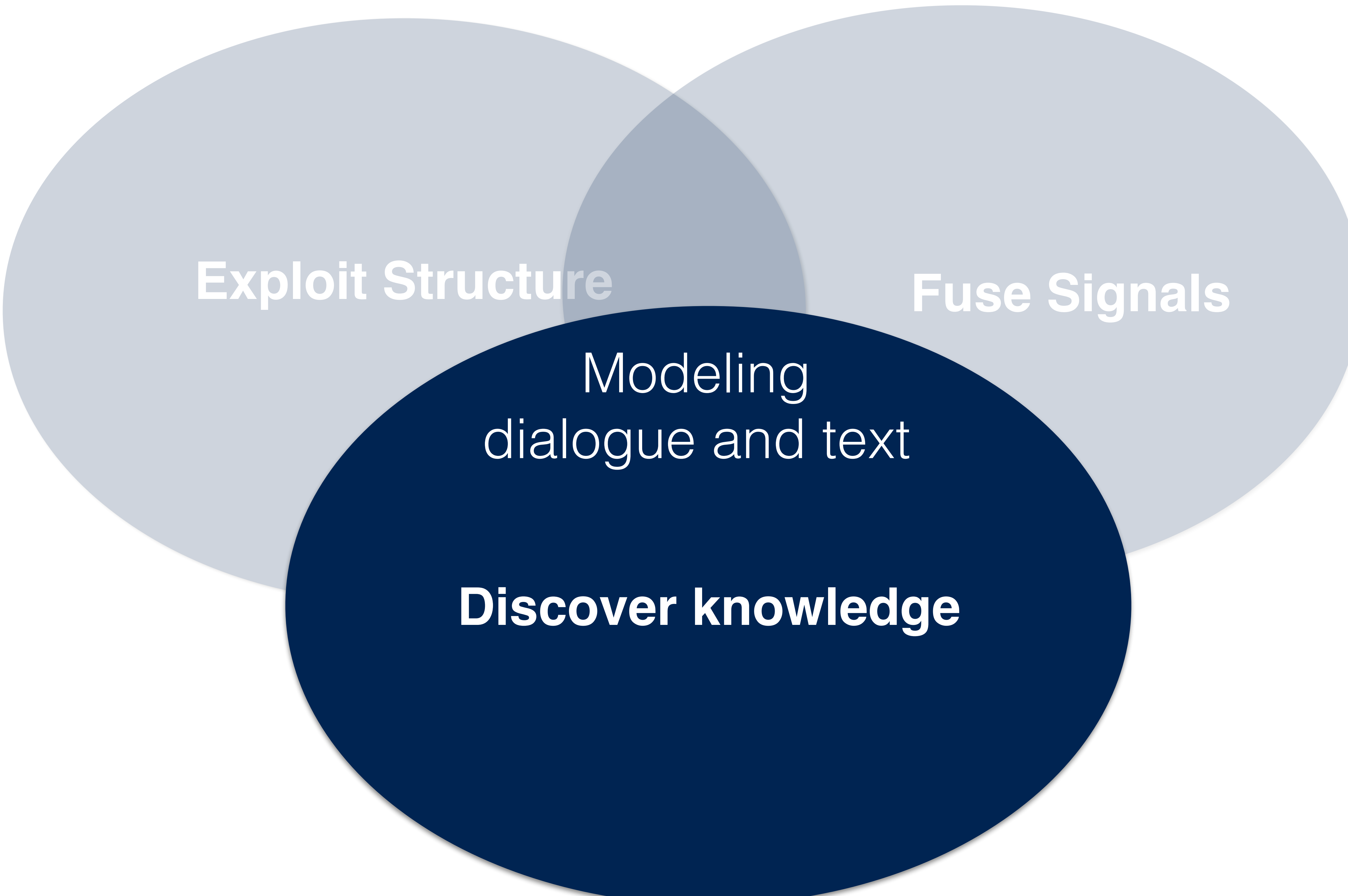
```
Friends(U2, U1) & UsesSoberWord(U2) & PosAffect(U1) → Recovers(U1)
Friends(U2, U1) & UsesAlcoholWord(U2) & PosAffect(U1) → !Recovers(U1)
```



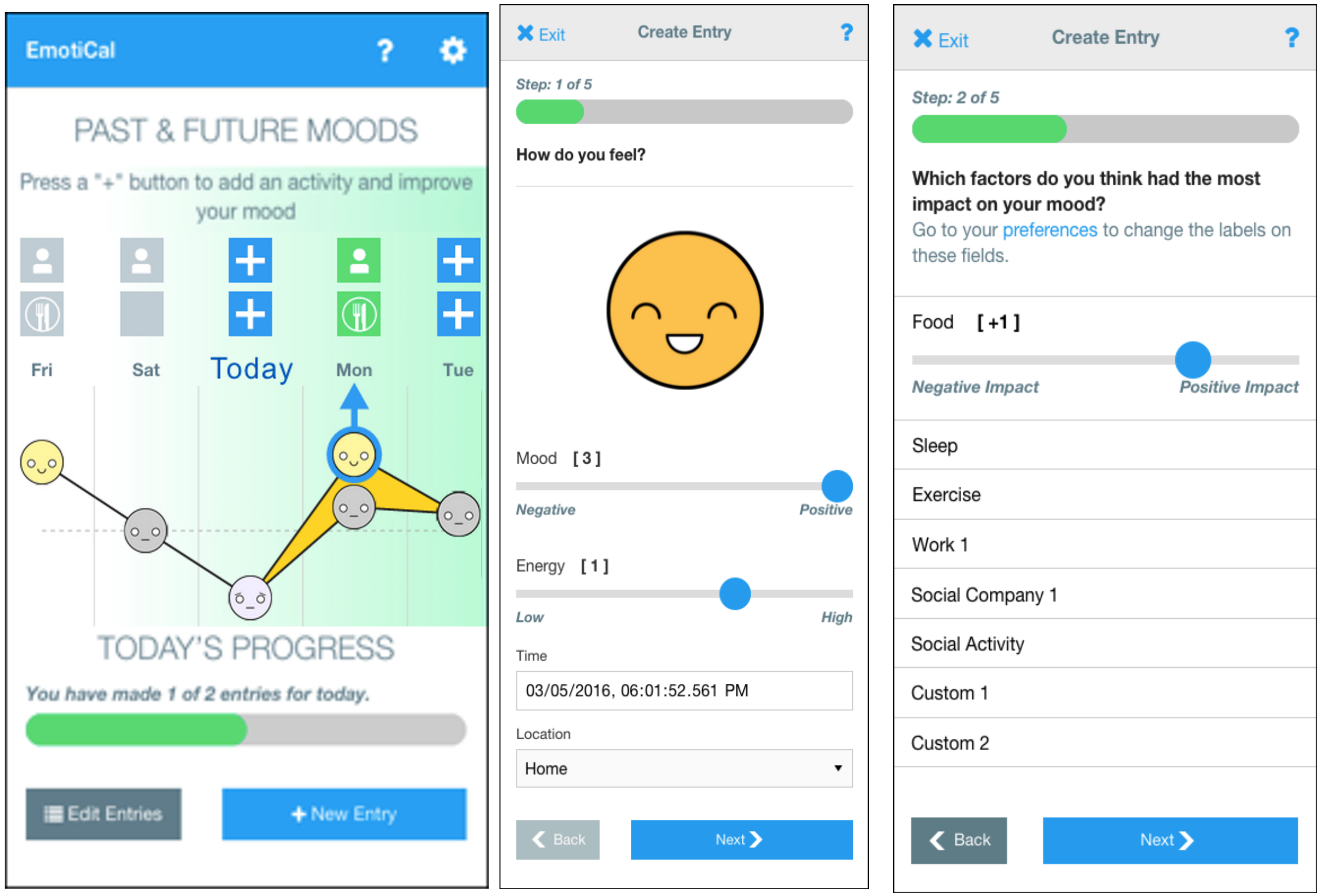
Capture nuanced dependencies and multiple models

Zhang, Ramesh, Golbeck, Sridhar, Getoor, 2018, A Structured Approach to Understanding Recovery and Relapse in AA. WWW.

Roadmap of my talk



Mood modeling dataset

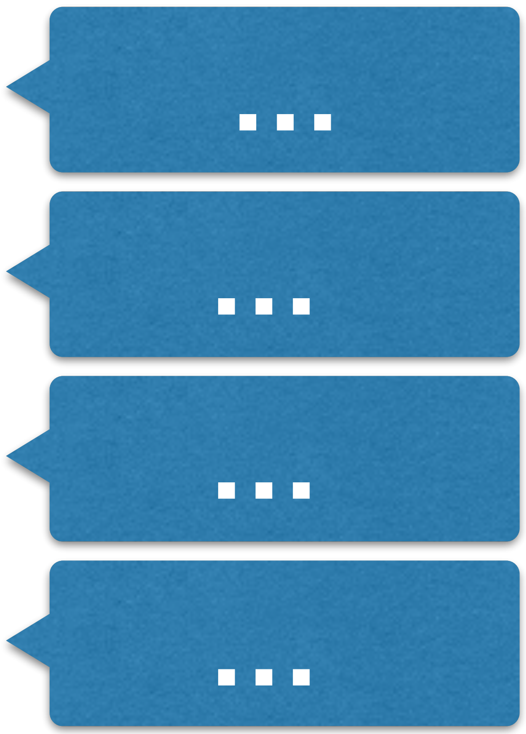


Users log mood and factors over range of time and describe their days

Hollis et al. 2017. *What Does All This Data Mean for My Future Mood? Actionable Analytics and Targeted Reflection for Emotional Well-Being*. Human–Computer Interaction.

Observational data for causal inference

Mood	Exercise	Sleep	...
-3	-2	-3	
2	3	1	
1	0	0	
-2	0	-3	



Unique opportunity to combine observational data with text

Observational data for causal inference

Mood	Exercise	Sleep	...
-3	-2	-3	
2	3	1	
1	0	0	
-2	0	-3	

Estimate causal effect of exercise on mood to validate against literature

Matching units for causal analysis

Treatment unit

Mood	Exercise	Sleep	...
-3	3	-3	

Control unit

Mood	Exercise	Sleep	...
-3	0	-3	

Perform matching to select most similar control for treatments

Estimation of causal effect

Treatment unit

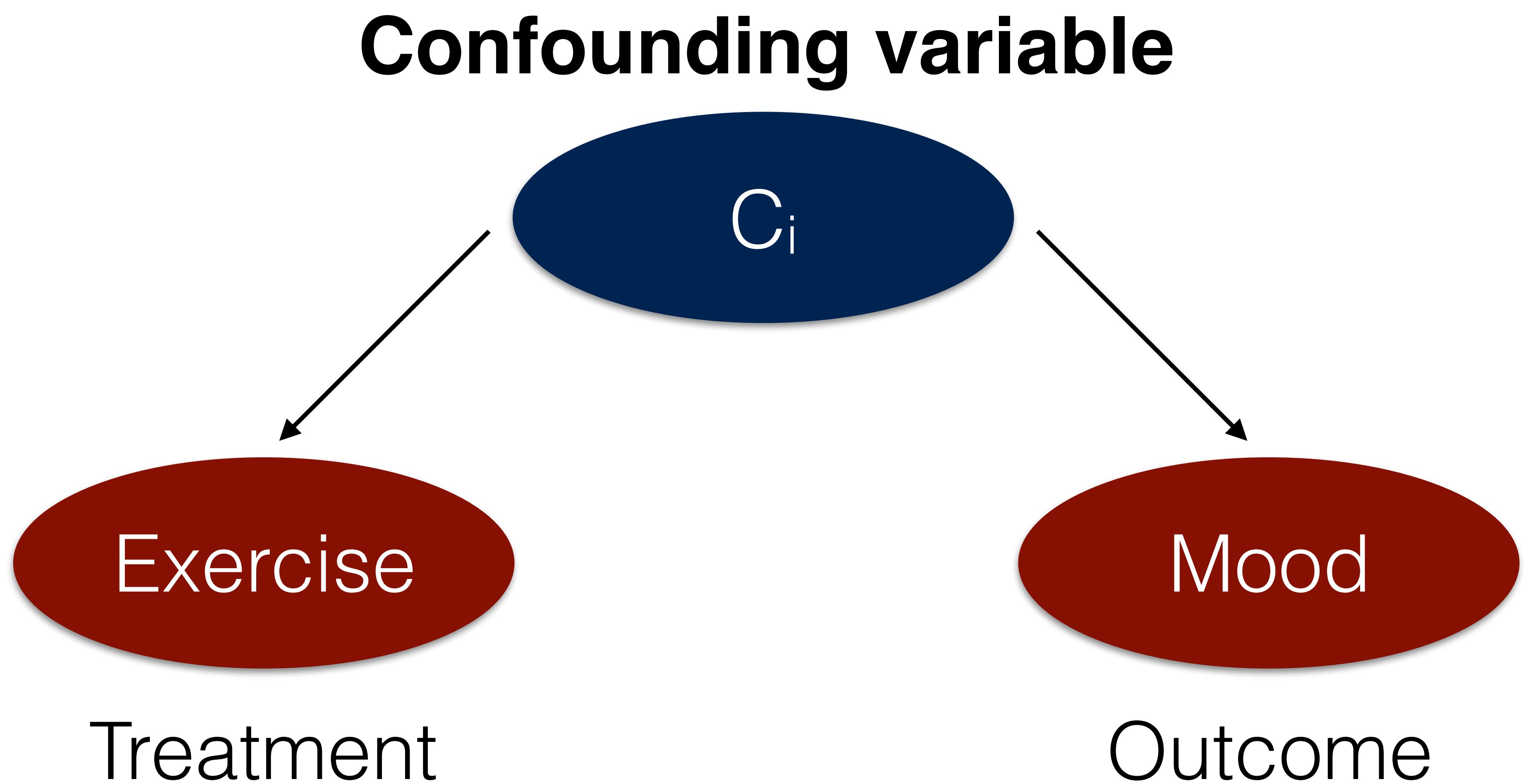
Mood	Exercise	Sleep	...
-3	3	-3	

Control unit

Mood	Exercise	Sleep	...
-3	0	-3	

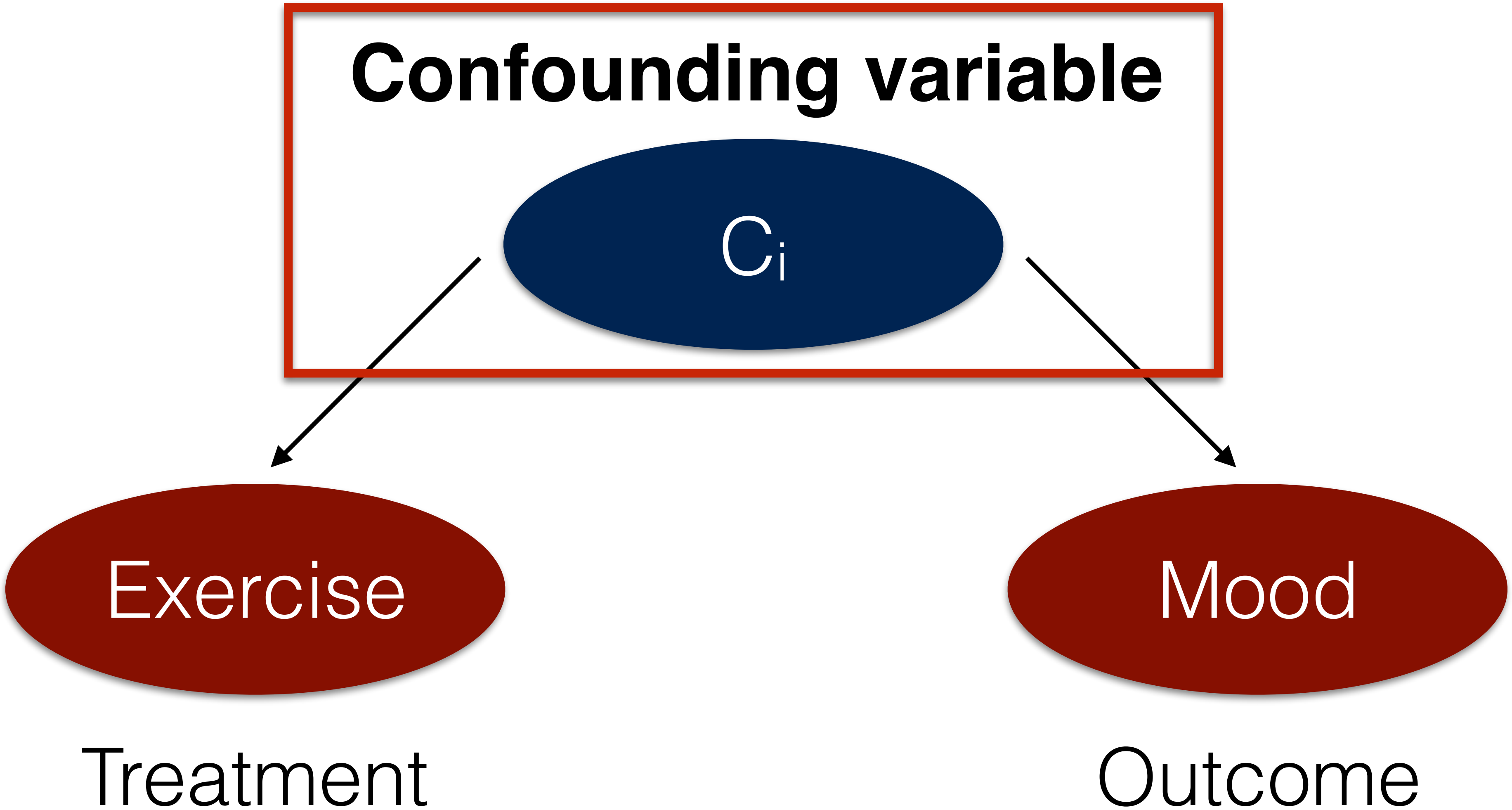
Many techniques support estimation including regression

Requirements for causal inference



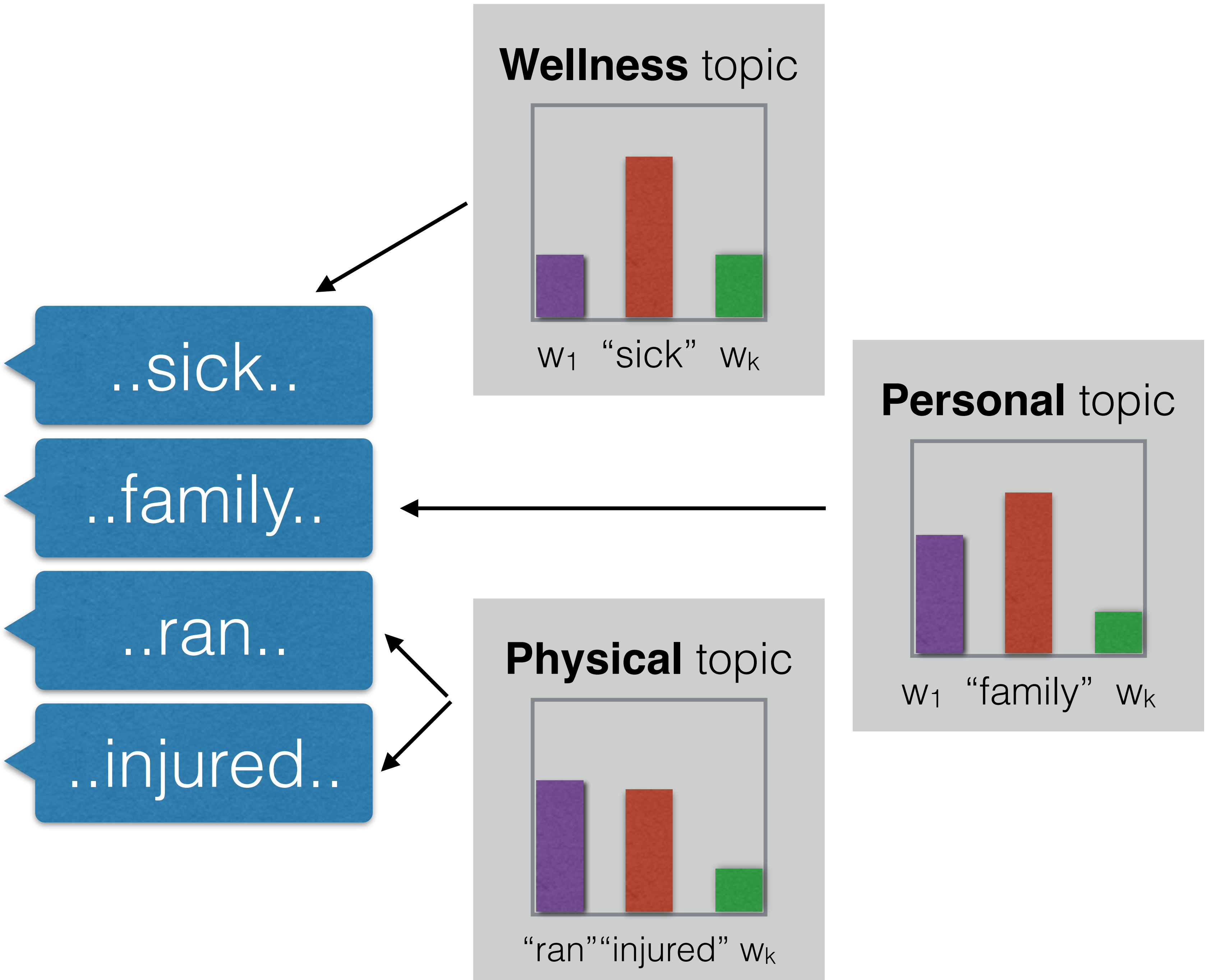
Need to include all common causes of treatment and outcome in matching and regression

Requirements for causal inference



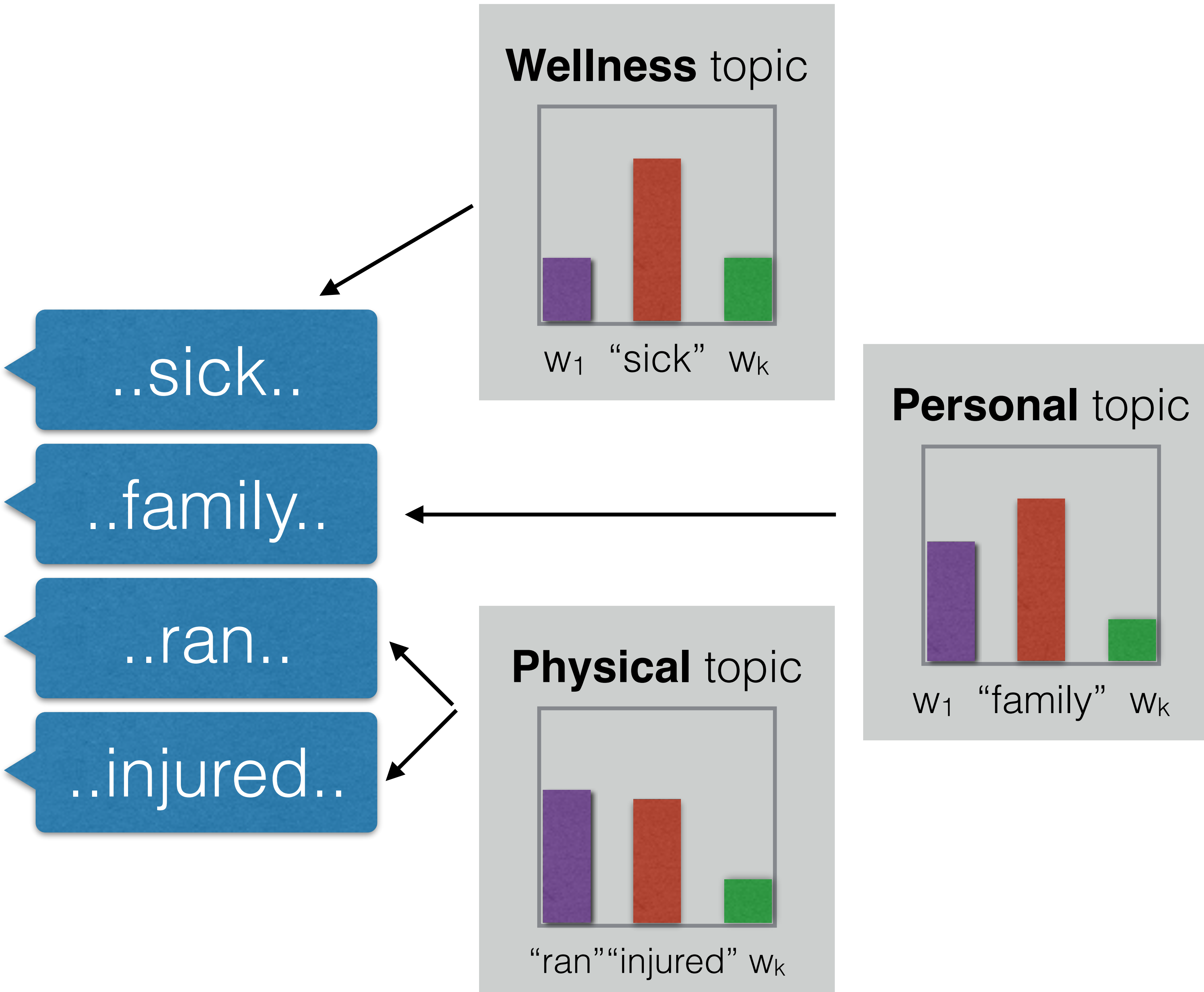
Many unmeasured, **latent** confounders

Approximate confounders from text



Discover latent confounding topics and words

Approximate confounders from text



Include as confounding variables in causal inference

Aggregating observations for users

Mood	Exercise	Sleep	…
-3	-2	-3	
2	3	1	
1	0	0	

“sick”

“ran”

Treatment = ?

Modeling structure across users’ observations

PSL for computational social science

- Discovering PSL rules from relational data
- Detecting cyber-bullying on social media
- Fusing data to characterize user personality types
- Relational bootstrapping for weakly supervised stance detection
- Identifying latent group attitudes on social media
- Learning patterns of engagement in MOOCs

Contributions in a nutshell

Broadly applicable templates for social science

Exploit structure

Fuse signals

Capture nuanced dependencies

Leverage new modes of evidence for causal inference

Causal

Discover

Infer models from observations for better reasoning

Contributions in a nutshell

Fair algorithms
Bias in datasets
More socio-behavioral domains
Causal inference
…

We're eager to collaborate!

References

[1] Yue Zhang, Arti Ramesh, Jennifer Golbeck, Dhanya Sridhar, and Lise Getoor. *A Structured Approach to Understanding Recovery and Relapse in AA*. In WWW, 2018.

[2] Golnoosh Farnadi, Jie Tang, Martine De Cock, Marie-Francine Moens. U*ser Profiling through Deep Multimodal Fusion*. In WSDM, 2018.

[3] Sabina Tomkins, Lise Getoor, Yunfei Chen, Yi Zhang. *Detecting Cyber-bullying from Sparse Data and Inconsistent Labels*. In Learning from Limited Data Workshop, 2017.

[4] Shachi H Kumar, Jay Pujara, Lise Getoor, David Mares, Dipak Gupta, Ellen Riloff. *Unsupervised Models to Predict Strategic Relations between Organizations.* In ASONAM, 2016.

[5] Dhanya Sridhar, James Foulds, Bert Huang, Lise Getoor, and Marilyn Walker. *Joint models of disagreement and stance in online debate*. In ACL, 2015.

[6] Arti Ramesh, Dan Goldwasser, Bert Huang, Hal Daume III, Lise Getoor. *Learning latent engagement patterns of students in online courses*. In AAAI, 2014.