A social-event based approach to sentiment analysis of identities and behaviors in text

Kenny Joseph

Kathleen M. Carley

Carnegie Mellon

Center for Computational Analysis of Social and Organizational Systems
http://www.casos.cs.cmu.edu/

The core research questions

How do "people" feel about different identities?

Can we learn this from text?

Data comes from the "Arab Spring"

- Newspaper data (700K articles)
  - LexisNexis, centered on 16 MENA countries
  - Major news outlets
  - 7/10 – 12/12

We use social events to infer affect

The New York Times Last week, Egyptian officials accused **demonstrators**

Traditionally:

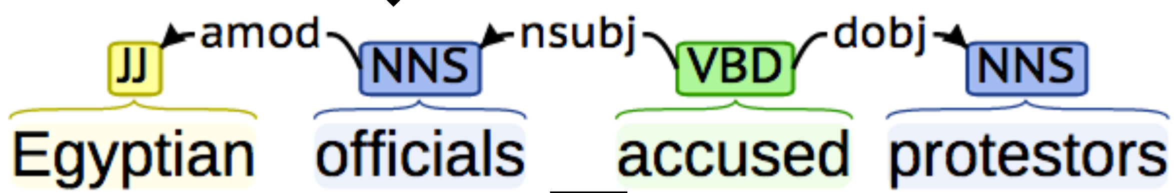
"good" demonstrators **"bad" demonstrators**

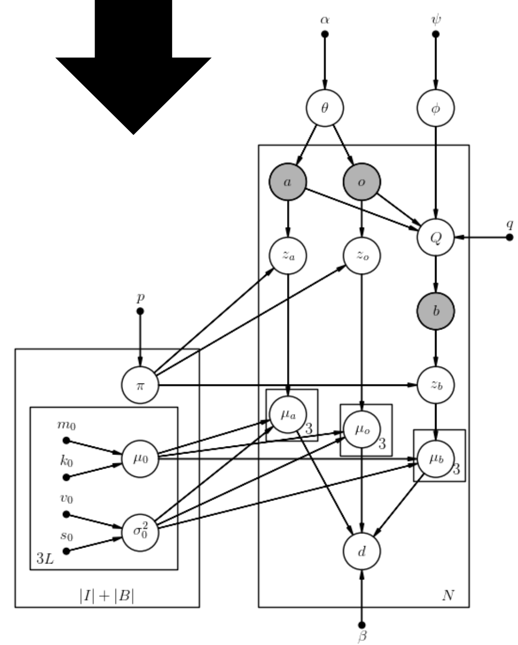
The present work:

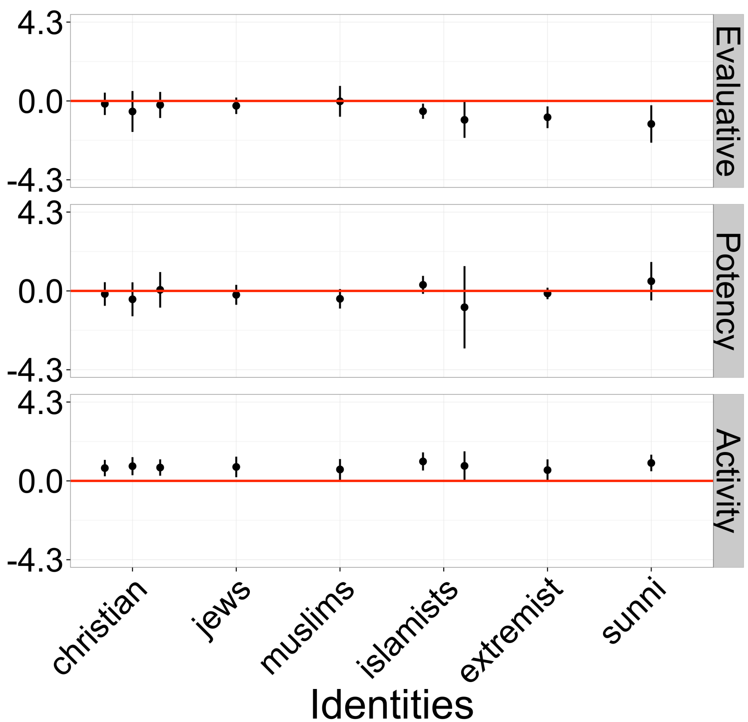
Egyptian official	**accuse**	**demonstrators**
Identity	**behavior**	**Other Identity**

An overview of the approach

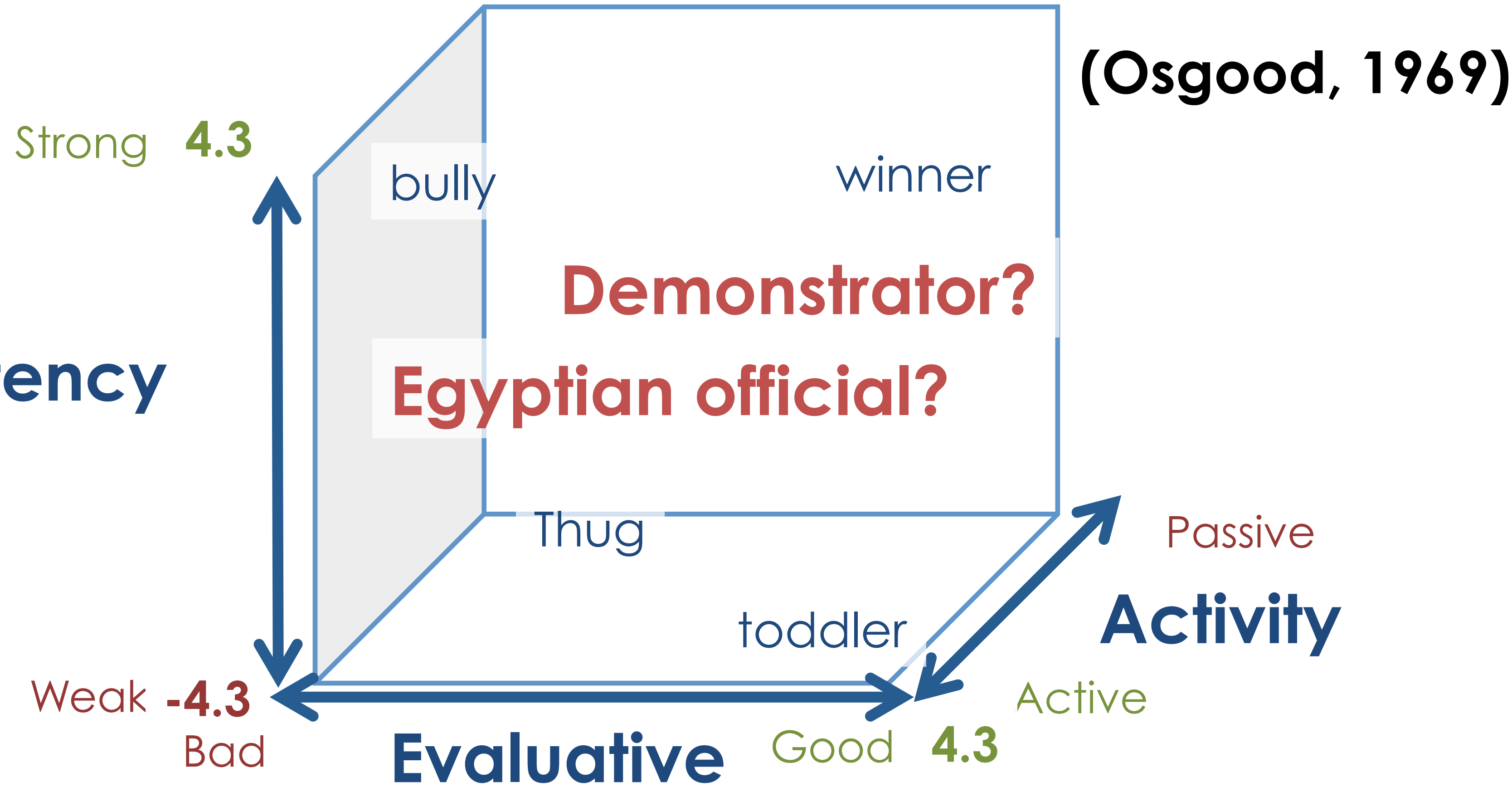
Last week, Egyptian officials accused **demonstrators**







Affect Control Theory and sentiment space



- ACT scholars obtain this data via survey
- Assumed to be culturally-shared

Social events provide info on sentiment

? officials

− criticize

\+ women

Affect Control Theory (ACT) gives a mathematical model for this

ACT defines a mathematical model of this

officials criticize women

$$f = \boxed{\begin{pmatrix} A_e & A_p & A_a \end{pmatrix}} \quad \boxed{\begin{matrix} B_e & B_p & B_a \end{matrix}} \quad \boxed{\begin{matrix} O_e & O_p & O_a \end{matrix}}$$

Unknown — Known — Known

Minimize $Deflection = \sum_{i}^{9} (f_i - M_{i*}^{T} G(f))^2$

$$\boxed{\begin{pmatrix} A_e & A_p & A_a \end{pmatrix}}$$

Known

Carnegie Mellon

institute for SOFTWARE RESEARCH

Inference is harder with incomplete info

? officials

– accused ?

? demonstrators

We can use many events to appease this

?	officials	?	officials
−	criticize	−	accused
+	women	?+	demonstrators

Meaning can/should diffuse through multiple events

Caveat – multiple meanings for identities

Sometimes,

officials $\neq$ officials

Solution:

Assume multiple latent "senses" of each identity/behavior

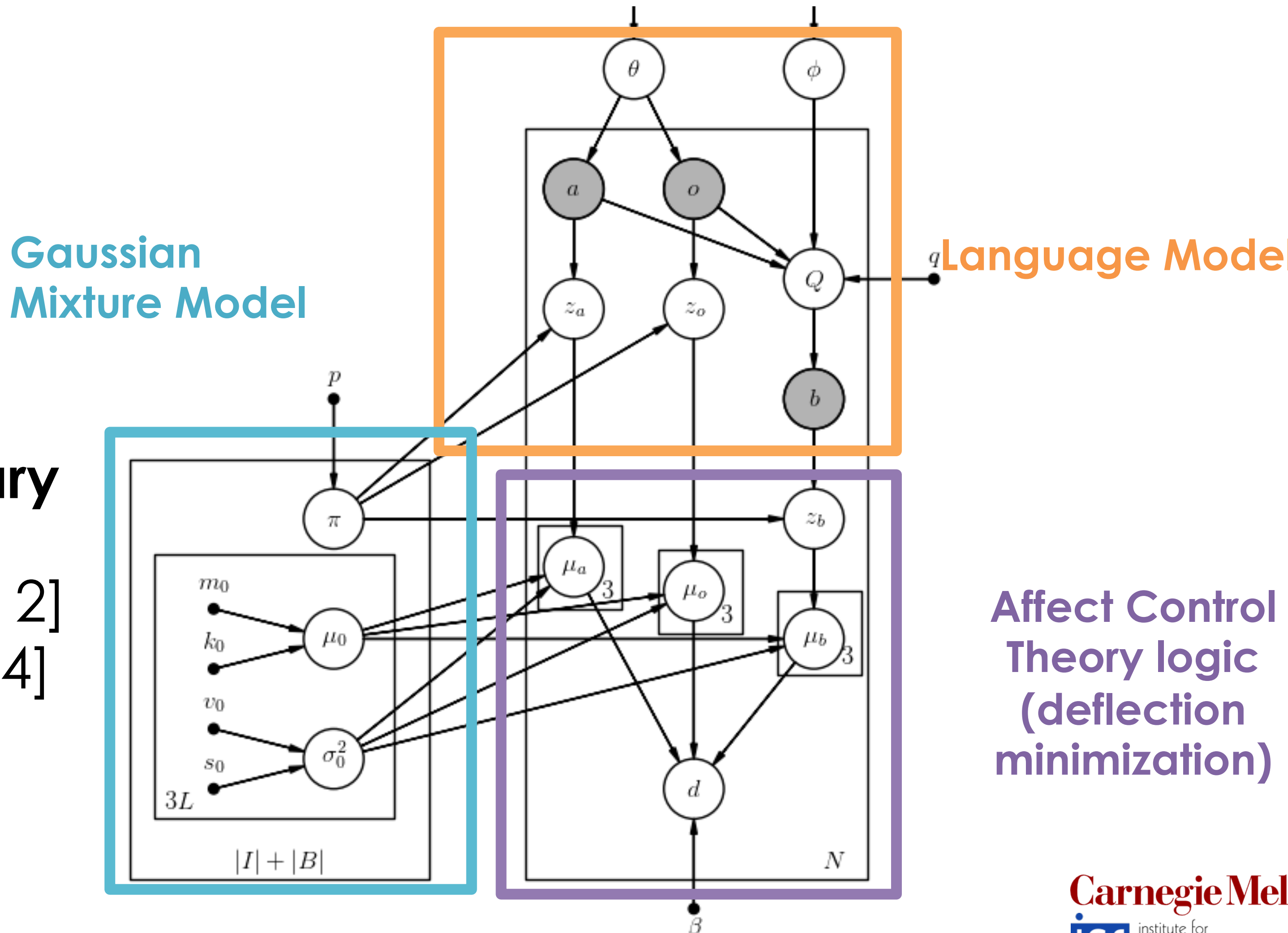
More on extracting social events, identities

1. Ran dependency parser, extracted all N -> V -> N
2. Cleaned text using, e.g., stemming (accused -> accuse)
3. Hand-curated list of identities and behaviors

- 102 identities, 87 behaviors, 10K events
- Only 44% of identities in ACT dicts

A Bayesian Network to extract sentiment

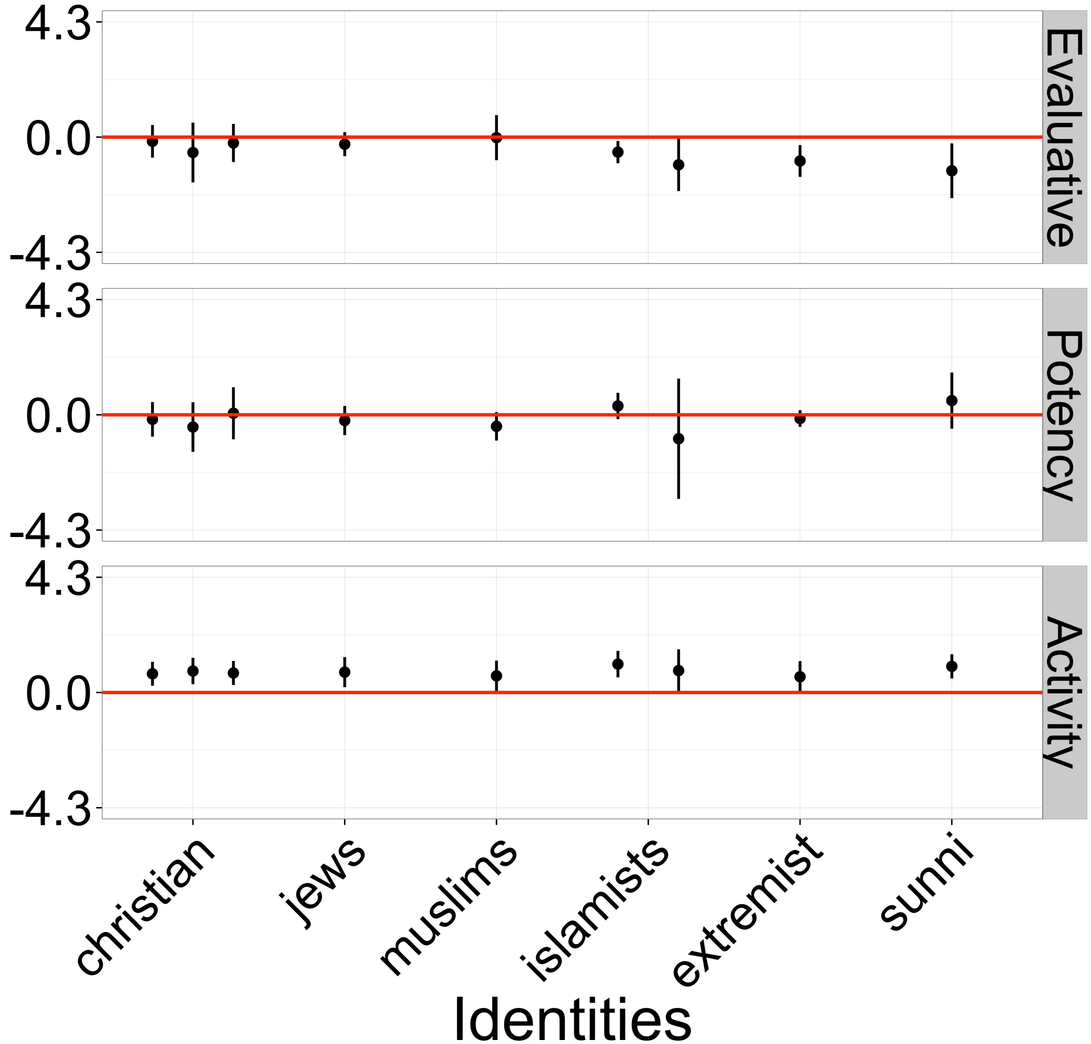
Officials criticize demonstrators



ACT Dictionary

criticize [-4 3 2]
women [4 1 4]
....

Results for a subset of (religious) identities)



Concluding

- Current work
  – New approach to sentiment mining for identities
  – Several parts of the process need work
- Future/Thesis work
  – Extract identities/behaviors from text automatically
  – Intermingling cognitive frameworks and Affect Control Theory

Thanks!!

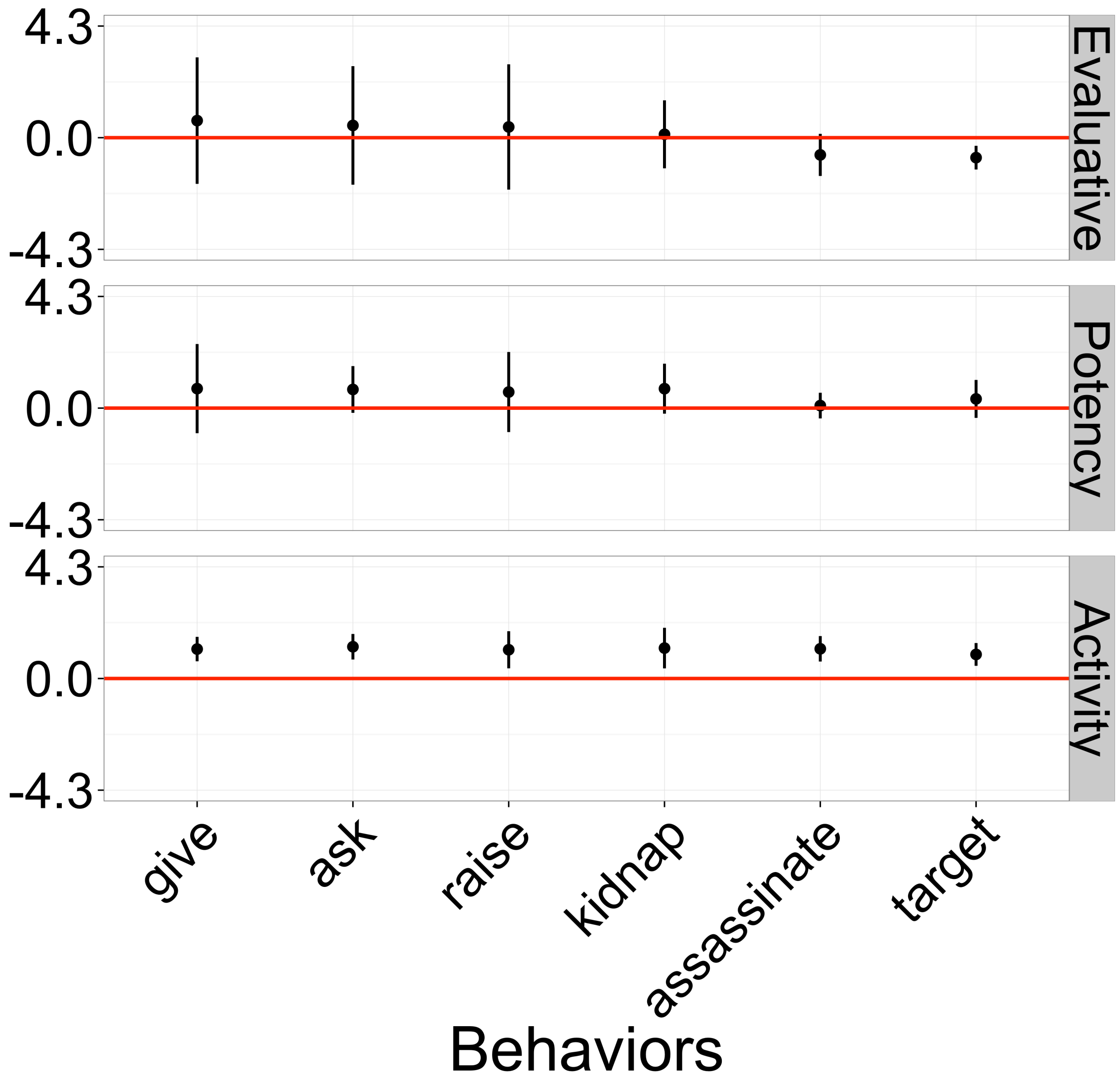
- In submission
- Replication data+code
  github.com/kennyjoseph/act_paper_public
- Me:
  Kenny Joseph
  cs.cmu.edu/~kjoseph
  @_kenny_joseph
  Looking for jobs…

Extra Slides

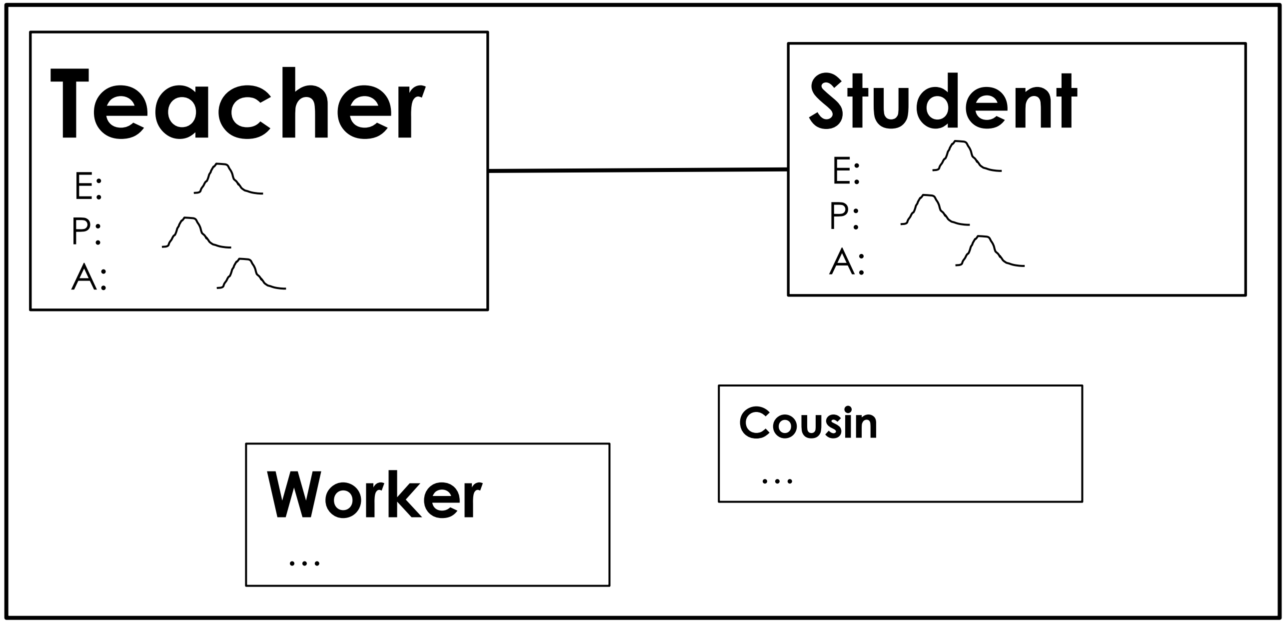
More details on the model

- Inference performed using Monte-Carlo EM
- Difficult parts of the inference
  - Inferring which latent topic/“sense” for a given event
  - Efficient usage of deflection principle
- Tested model using held-out events compared to various baselines

Results for behaviors



Can choose identities in context



f(Activation, Relationships, EPA)

student
worker?
classmate?
traveler?
cousin?

Teacher			advises					
2.45	1.75	0.29	2.12	1.59	0.98	1.49	0.31	0.75

More details on the model

1. Draw an actor and an object; $a \sim \mathrm{Cat}(\theta) \qquad o \sim \mathrm{Cat}(\theta)$

2. Draw a behavior; $b \sim \mathrm{Cat}(Q)$

3. Draw a latent sense for a, b and o; $z_a \sim \mathrm{Cat}(\pi_a) \qquad z_b \sim \mathrm{Cat}(\pi_b) \qquad z_o \sim \mathrm{Cat}(\pi_o)$

4. Draw EPA profiles for a, b, and o

   - $\mu_{a,e} \sim \mathrm{N}(\mu_{0,z_a,e}; \sigma^2_{0,z_a,e}) \quad \mu_{a,p} \sim \mathrm{N}(\mu_{0,z_a,p}; \sigma^2_{0,z_a,p}) \quad \mu_{a,a} \sim \mathrm{N}(\mu_{0,z_a,a}; \sigma^2_{0,z_a,a})$
   - $\mu_{b,e} \sim \mathrm{N}(\mu_{0,z_b,e}; \sigma^2_{0,z_b,e}) \quad \mu_{b,p} \sim \mathrm{N}(\mu_{0,z_b,p}; \sigma^2_{0,z_b,p}) \quad \mu_{b,a} \sim \mathrm{N}(\mu_{0,z_b,a}; \sigma^2_{0,z_b,a})$
   - $\mu_{o,e} \sim \mathrm{N}(\mu_{0,z_o,e}; \sigma^2_{0,z_o,e}) \quad \mu_{o,p} \sim \mathrm{N}(\mu_{0,z_o,p}; \sigma^2_{0,z_o,p}) \quad \mu_{o,a} \sim \mathrm{N}(\mu_{0,z_o,a}, \sigma^2_{0,z_o,a})$

5. Draw a deflection score for the event

   - $d \sim \mathrm{Laplace}(\sum_i^9 (f_i - M_{i*}^T G(f))^2, \beta)$ where $f = [\mu_{a_e}, \mu_{a_p}, \mu_{a_a}, \mu_{b_e}, \mu_{b_p}, \mu_{b_a}, \mu_{o_e}, \mu_{o_p}, \mu_{o_a}]$ lon

ACT is a “first order social effects” model

Teacher beats up student

School official child

- What we learn should “filter” to encompassing concepts
- Also makes sense that related concepts somehow “influence” perceptions

Future work: incorporate cognitive relations

Egyptian officials

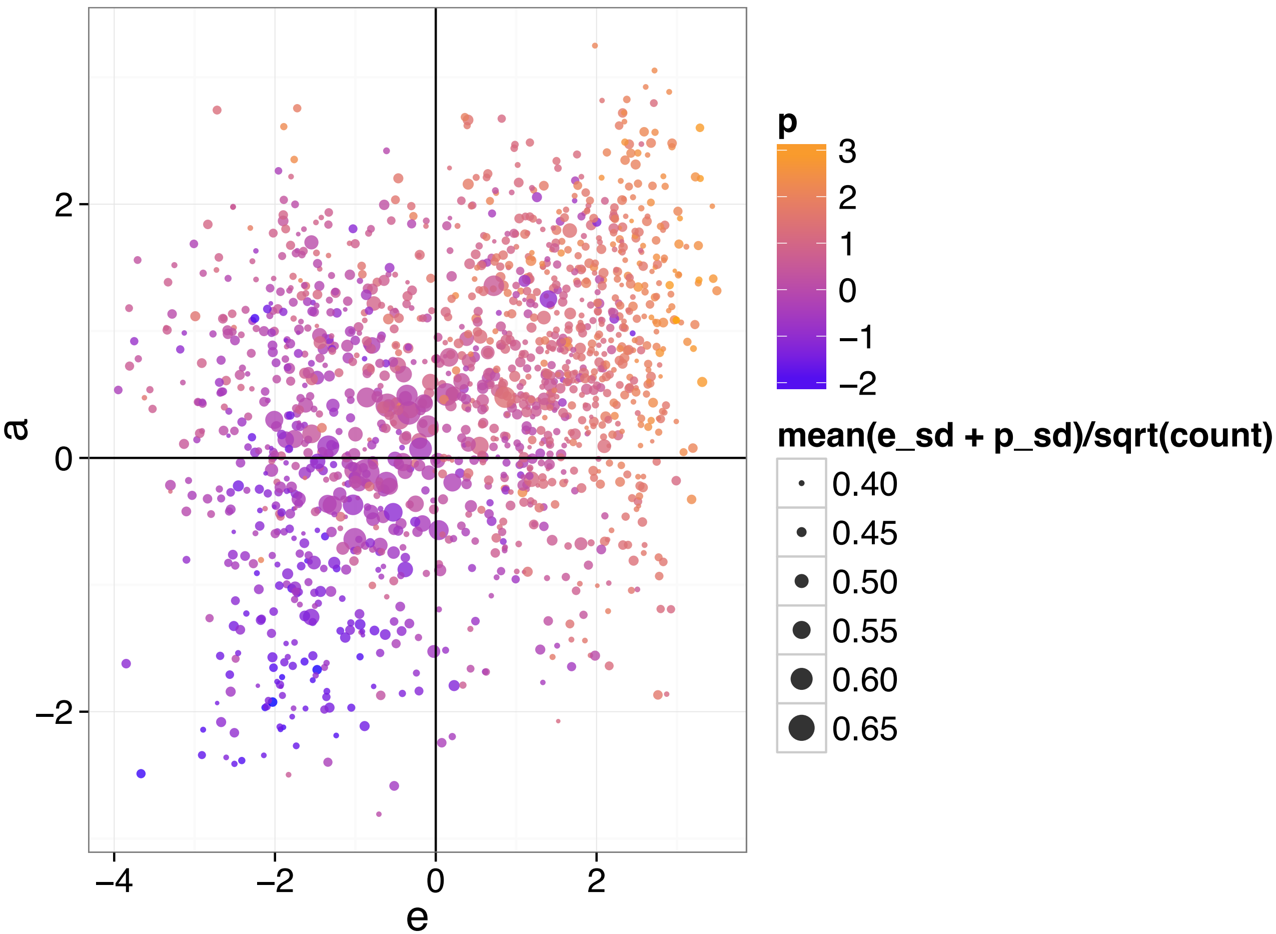
Egyptian officials

Egyptians

Can also consider other relations between identities

EPA Correlations in existing ACT dictionaries

- From the Indiana survey



ACT provides a mathematical model of how these meanings change and are maintained during *social events*

Teacher			advises			student		
2.45	1.75	0.29	2.12	1.59	0.98	1.49	0.31	0.75
Good/ Bad	Strong/ Weak	Active/ Passive						

$$f = \begin{pmatrix} A_e & A_p & A_a & B_e & B_p & B_a & O_e & O_p & O_a \end{pmatrix}$$

$$Deflection = \sum_i^9 (f_i - M_{i*}^T G(f))^2 \quad = .08$$

(This) Teacher			beats up			(this) student		
2.73	1.34	0.95	2.12	1.59	0.98	1.95	-0.58	0.51

Deflection = 15.4

(This) Teacher			(This) student		
-2.15	1.04	1.61	-2.23	0.65	-0.02

ACT defines what social events change

You see:

You think:

Then you see:

Teacher			beats up			student		
2.45	1.75	0.29	-1.92	1.00	1.62	1.49	0.31	0.75
Good/ Bad	Strong/ Weak	Active/ Passive						

Now you think: $M_{i*}^{T}G(f)$

(This) Teacher			(This) student		
-2.15	1.04	1.61	-2.23	0.65	-0.02

ACT provides a mathematical model of how the meanings change and are maintained during *social events*

$$\begin{array}{c|c|c} \textbf{Teacher} & \textbf{instructs} & \textbf{student} \\ f = (A_e \quad A_p \quad A_a & B_e \quad B_p \quad B_a & O_e \quad O_p \quad O_a) \end{array}$$

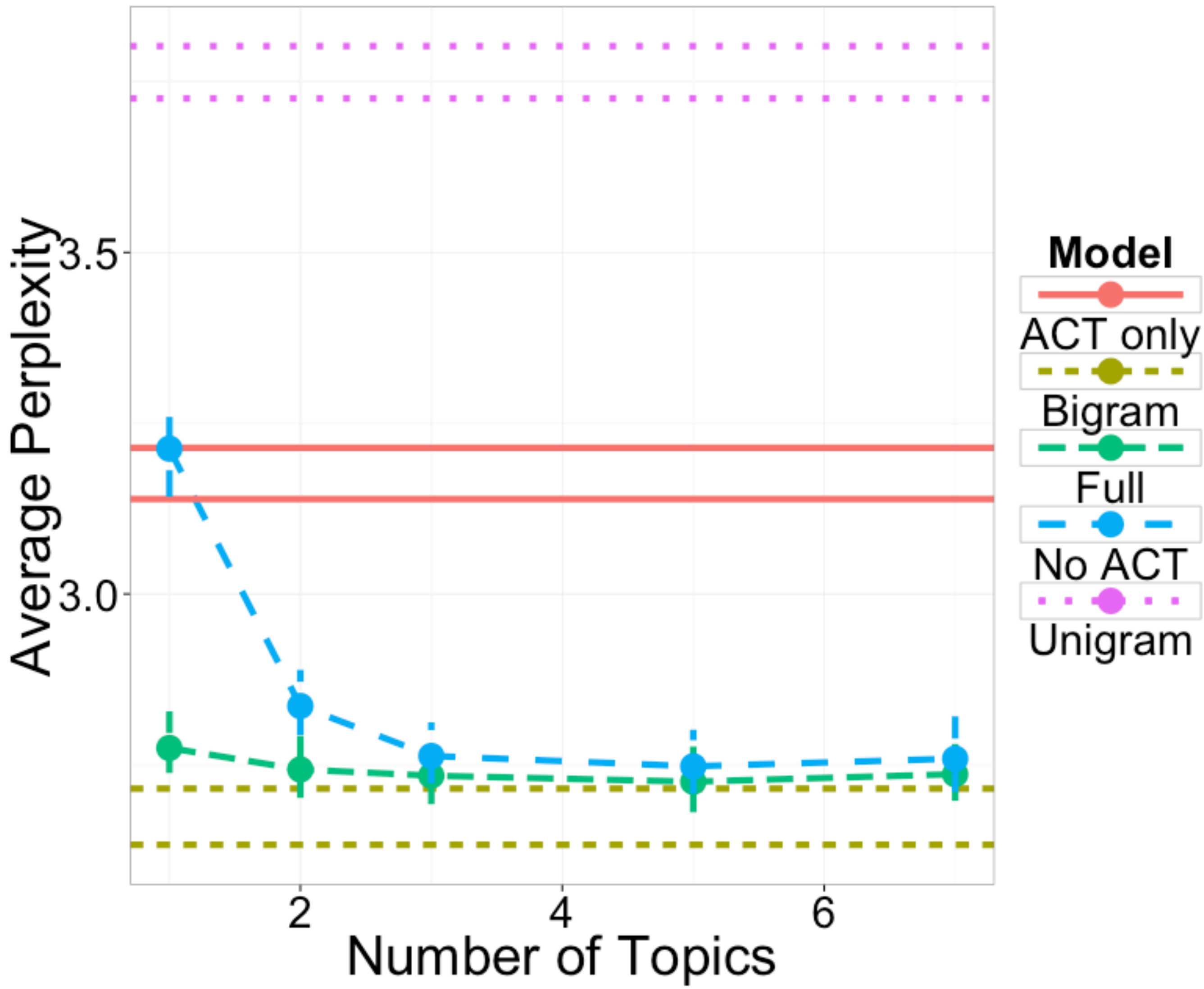
$$Deflection = \sum_{i}^{9}(f_i - M_{i*}^T G(f))^2$$

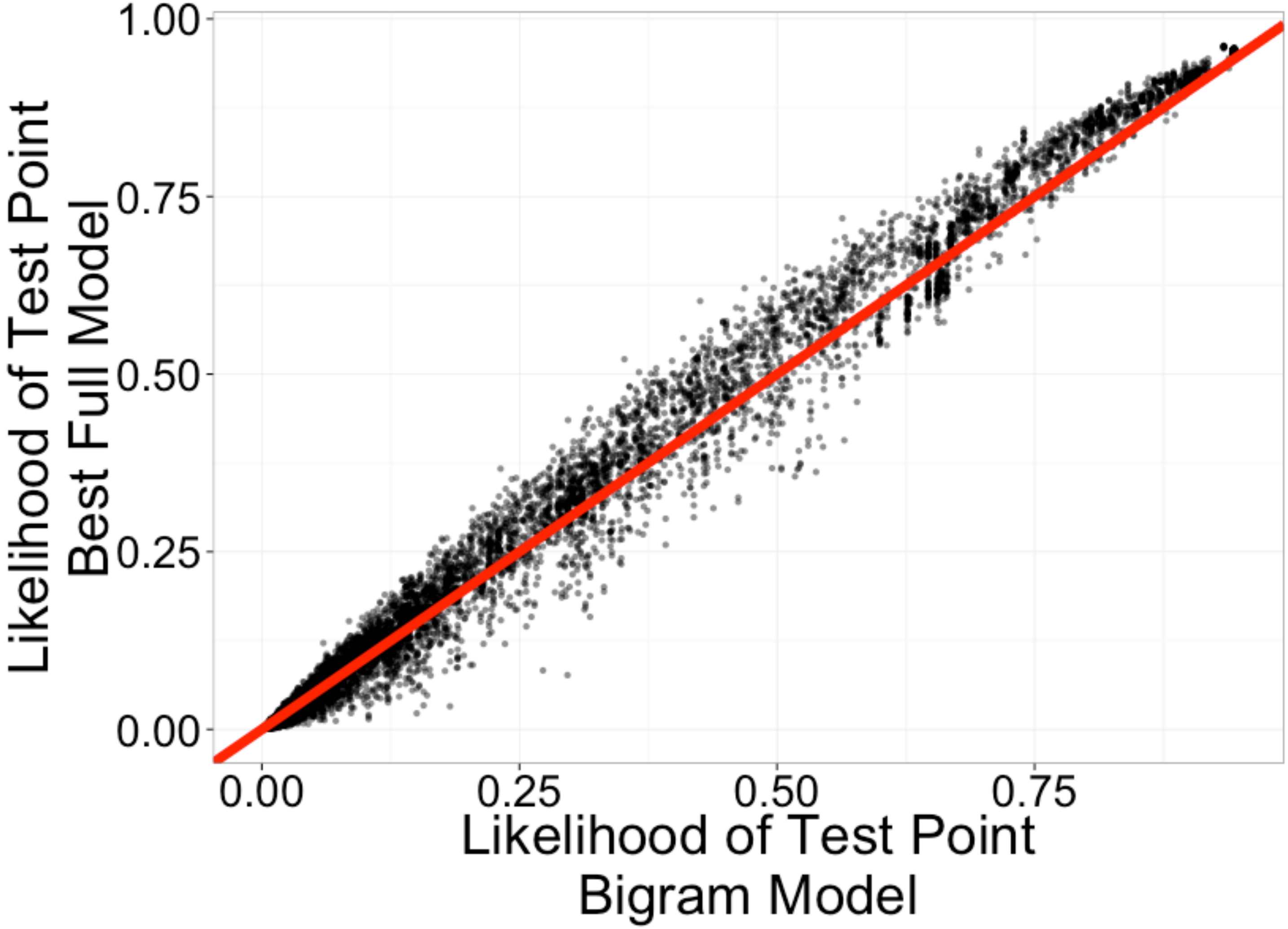
Teacher instructs student (Deflection: 0.8)

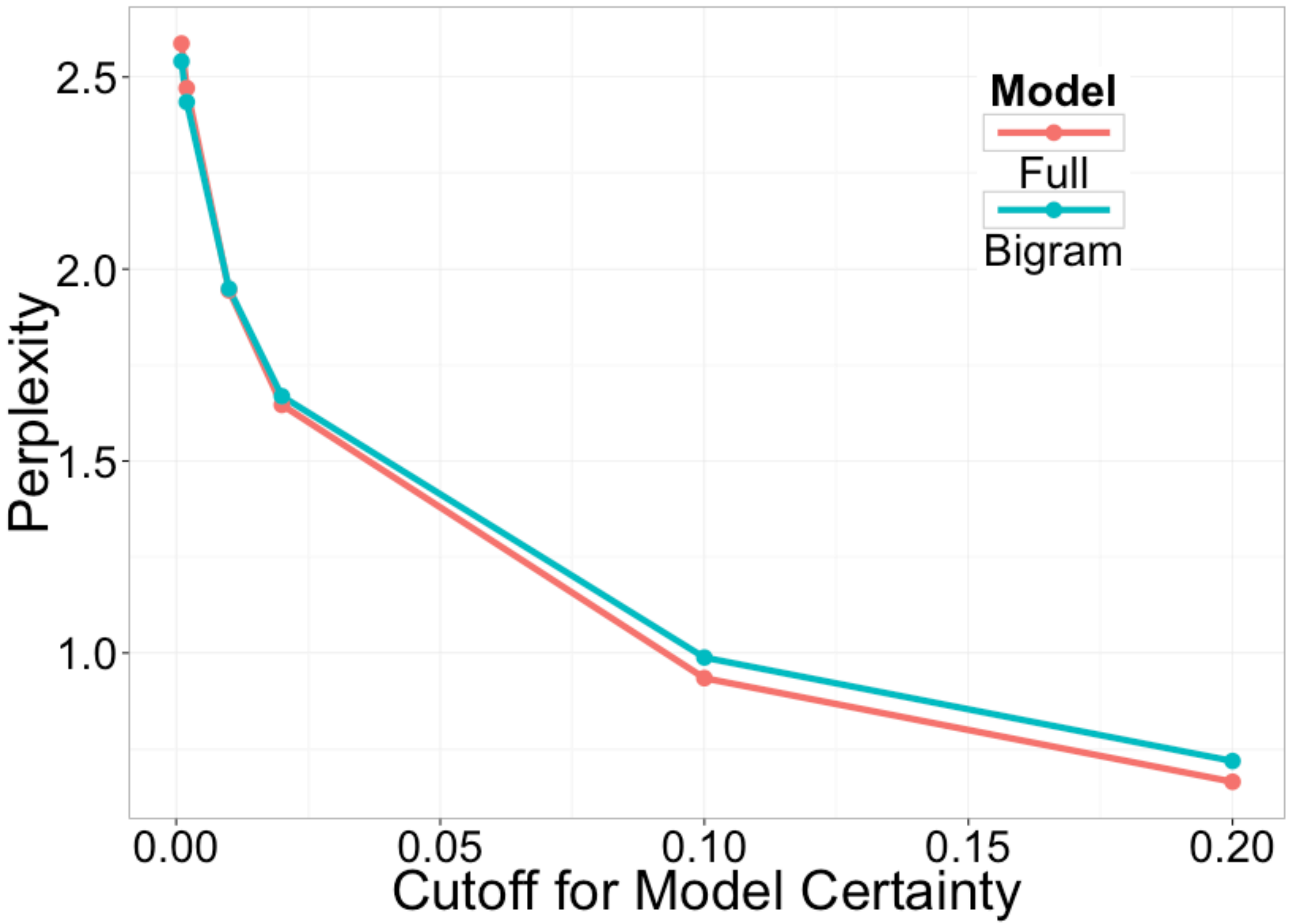
$$\begin{aligned} G(f) = [1 \quad & A_e \quad A_p \quad A_a \quad B_e \quad B_p \quad B_a \quad O_e \quad O_p \quad O_a \\ & B_eA_e \quad B_pA_e \quad O_eA_e \quad A_eO_p \quad A_pB_e \quad A_pB_p \\ & A_pO_p \quad A_pO_a \quad A_aB_a \quad B_eO_e \quad B_eO_p \quad O_eB_p \quad B_pO_p \\ & B_pO_a \quad B_aO_p \quad B_eO_eA_e \quad B_pA_eO_p \quad A_pB_pO_p \quad A_pB_pO_a] \end{aligned}$$

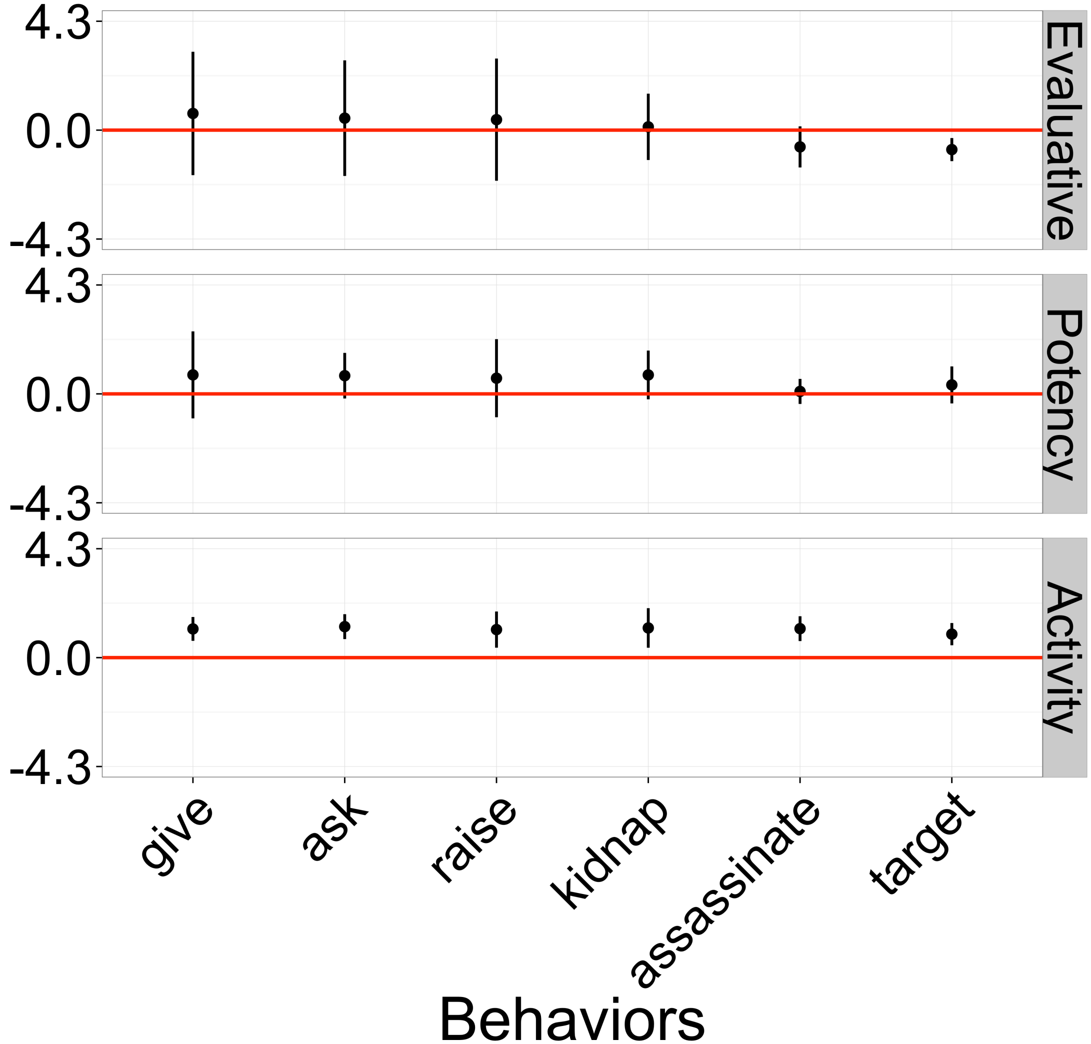
$$M = \begin{pmatrix} c_{1,1} & -.09 & \dots & c_{1,29} \\ .445 & c_{22} & \dots & c_{2,29} \\ \vdots & \vdots & \ddots & \vdots \\ .025 & 0 & \dots & c_{9,29} \end{pmatrix}$$

Predicts pretty well

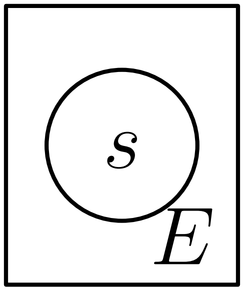


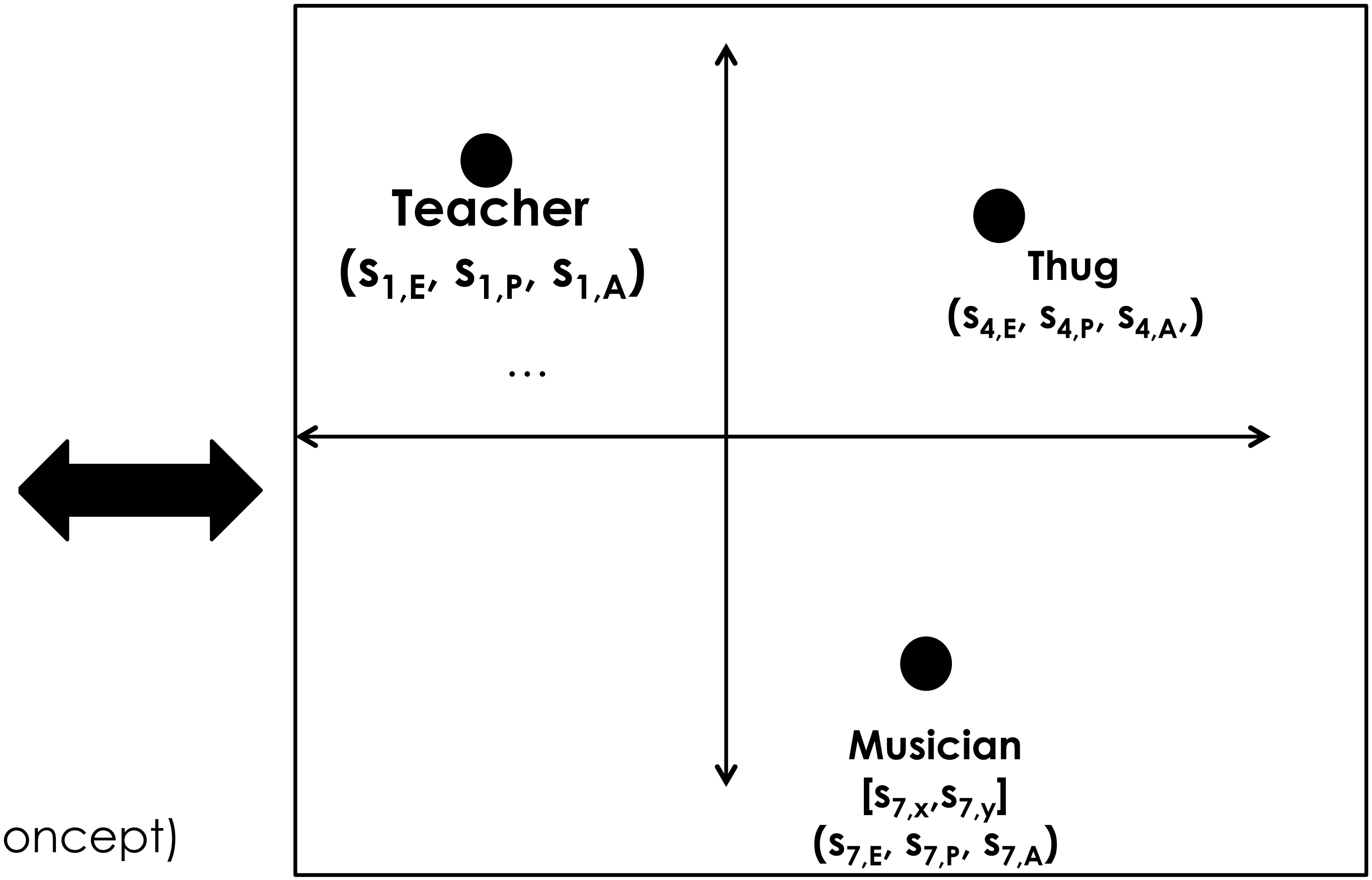

1.00
0.75
0.50
0.25
0.00
Likelihood of Test Point
Best Full Model
0.00
0.25
0.50
0.75
Likelihood of Test Point
Bigram Model


Model
Full
Bigram
Perplexity
2.5
2.0
1.5
1.0
0.00
0.05
0.10
0.15
0.20
Cutoff for Model Certainty


4.3
0.0
-4.3
Evaluative
Potency
Activity
give
ask
raise
kidnap
assassinate
target
Behaviors

Probabilistic graphical model of LCSS





s = schema
E = entity (concept)

Each variable *s* is a 6-dim. representation of a schema