Large Scale Learning for Information Extraction

Alan Ritter
Computer Science and Engineering
The Ohio State University

THE OHIO STATE UNIVERSITY

Q: Why are we so good at Speech, MT (but not NLU)?

People ***naturally*** translate and transcribe.

Q: Why are we so good at Speech, MT (but not NLU)?

People ***naturally*** translate and transcribe.

Large, End-to-End Datasets for NLU?

- Textual Entailment, Paraphrase
  - SNLI, PPDB, Twitter Paraphrases
- Question Answering
  - SQuaD, MS MARCO, TriviaQA

Q: Why are we so good at Speech, MT (but not NLU)?

People ***naturally*** translate and transcribe.

Large, End-to-End Datasets for NLU?

Learning from web-scale Conversations (e.g. conversational agents)

Data Driven Response Generation in Social Media
Alan Ritter, Colin Cherry, Bill Dolan
Proceedings of EMNLP 2011

Deep Reinforcement Learning for Dialogue Generation
Jiwei Li, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao, Dan Jurafsky.
Proceedings of EMNLP 2016

Adversarial Learning for Neural Dialogue Generation
Jiwei Li, Will Monroe, Tianlin Shi, Sébastien Jean, Alan Ritter, Dan Jurafsky
Proceedings of EMNLP 2017

Q: Why are we so good at Speech, MT (but not NLU)?

People ***naturally*** translate and transcribe.

Large, End-to-End Datasets for NLU?

Learning from web-scale Conversations (e.g. conversational agents)

Minimally Supervised Information Extraction (relations, events, etc…)

This Talk

Structured Learning for Neural Relation Extraction
Fan Bai and Alan Ritter
In Submission

Data Driven Response Generation in Social Media
Alan Ritter, Colin Cherry, Bill Dolan
Proceedings of EMNLP 2011

Deep Reinforcement Learning for Dialogue Generation
Jiwei Li, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao, Dan Jurafsky.
Proceedings of EMNLP 2016

Adversarial Learning for Neural Dialogue Generation
Jiwei Li, Will Monroe, Tianlin Shi, Sébastien Jean, Alan Ritter, Dan Jurafsky
Proceedings of EMNLP 2017

Q: Why are we so good at Speech, MT (but not NLU)?

People ***naturally*** translate and transcribe.

Large, End-to-End Datasets for NLU?

Learning from web-scale Conversations (e.g. conversational agents)

Minimally Supervised Information Extraction (relations, events, etc…)

This Talk

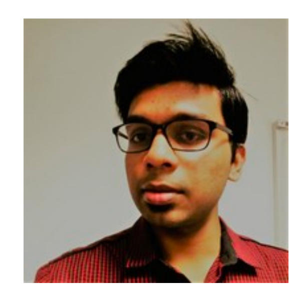
Data Driven Response Generation in Social Media
Alan Ritter, Colin Cherry, Bill Dolan
Proceedings of EMNLP 2011

Deep Reinforcement Learning for Dialogue Generation
Jiwei Li, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao, Dan Jurafsky.
Proceedings of EMNLP 2016

Adversarial Learning for Neural Dialogue Generation
Jiwei Li, Will Monroe, Tianlin Shi, Sébastien Jean, Alan Ritter, Dan Jurafsky
Proceedings of EMNLP 2017

Structured Learning for Neural Relation Extraction
Fan Bai and Alan Ritter
In Submission

"i have a feeling trump will win..................": Forecasting Winners and Losers from User Predictions on Twitter
Sandesh Swamy, Alan Ritter and Marie-Catherine de Marneffe
Proceedings of EMNLP 2017



Distant Supervision for Relation Extraction (A Timeline)

Time

Distant Supervision for Relation Extraction (A Timeline)

Time

(Mintz et. al. ACL 2009)

KB (Freebase) as Distant Supervision

Distant Supervision for Relation Extraction (A Timeline)

Time

- (Mintz et. al. ACL 2009) — **KB (Freebase) as Distant Supervision**
- (Riedel et. al. ECML 2010) — **NYT / Freebase Dataset + MIL**

Distant Supervision for Relation Extraction (A Timeline)

Time

- (Mintz et. al. ACL 2009) — **KB (Freebase) as Distant Supervision**
- (Riedel et. al. ECML 2010) — **NYT / Freebase Dataset + MIL**
- (Hoffmann et. al. ACL 2011) — **Structured Perceptron + Determ. OR**
  Mention-Level Evaluation

Distant Supervision for Relation Extraction (A Timeline)

Time

- (Mintz et. al. ACL 2009) — **KB (Freebase) as Distant Supervision**
- (Riedel et. al. ECML 2010) — **NYT / Freebase Dataset + MIL**
- (Hoffmann et. al. ACL 2011) — **Structured Perceptron + Determ. OR Mention-Level Evaluation**
- (Surdeanu et. al. EMNLP 2012) — **Multi-Instance, Multi-Label Learning**
- (Ritter et. al. TACL 2013) — **Missing Data (Mention-Level Evaluation)**

Distant Supervision for Relation Extraction (A Timeline)

Time

- (Mintz et. al. ACL 2009) — **KB (Freebase) as Distant Supervision**
- (Riedel et. al. ECML 2010) — **NYT / Freebase Dataset + MIL**
- (Hoffmann et. al. ACL 2011) — **Structured Perceptron + Determ. OR Mention-Level Evaluation**
- (Surdeanu et. al. EMNLP 2012) — **Multi-Instance, Multi-Label Learning**
- (Ritter et. al. TACL 2013) — **Missing Data (Mention-Level Evaluation)**
- (Zeng et. al. ACL 2015) — **Piecewise CNN**

Distant Supervision for Relation Extraction (A Timeline)

Time

- (Mintz et. al. ACL 2009) — **KB (Freebase) as Distant Supervision**
- (Riedel et. al. ECML 2010) — **NYT / Freebase Dataset + MIL**
- (Hoffmann et. al. ACL 2011) — **Structured Perceptron + Determ. OR Mention-Level Evaluation**
- (Surdeanu et. al. EMNLP 2012) — **Multi-Instance, Multi-Label Learning**
- (Ritter et. al. TACL 2013) — **Missing Data (Mention-Level Evaluation)**
- (Zeng et. al. ACL 2015) — **Piecewise CNN**
- (Lin et. al. ACL 2016) — **Piecewise CNN + Selective Attention**

Distant Supervision for Relation Extraction (A Timeline)

Time

- (Mintz et. al. ACL 2009) — **KB (Freebase) as Distant Supervision**

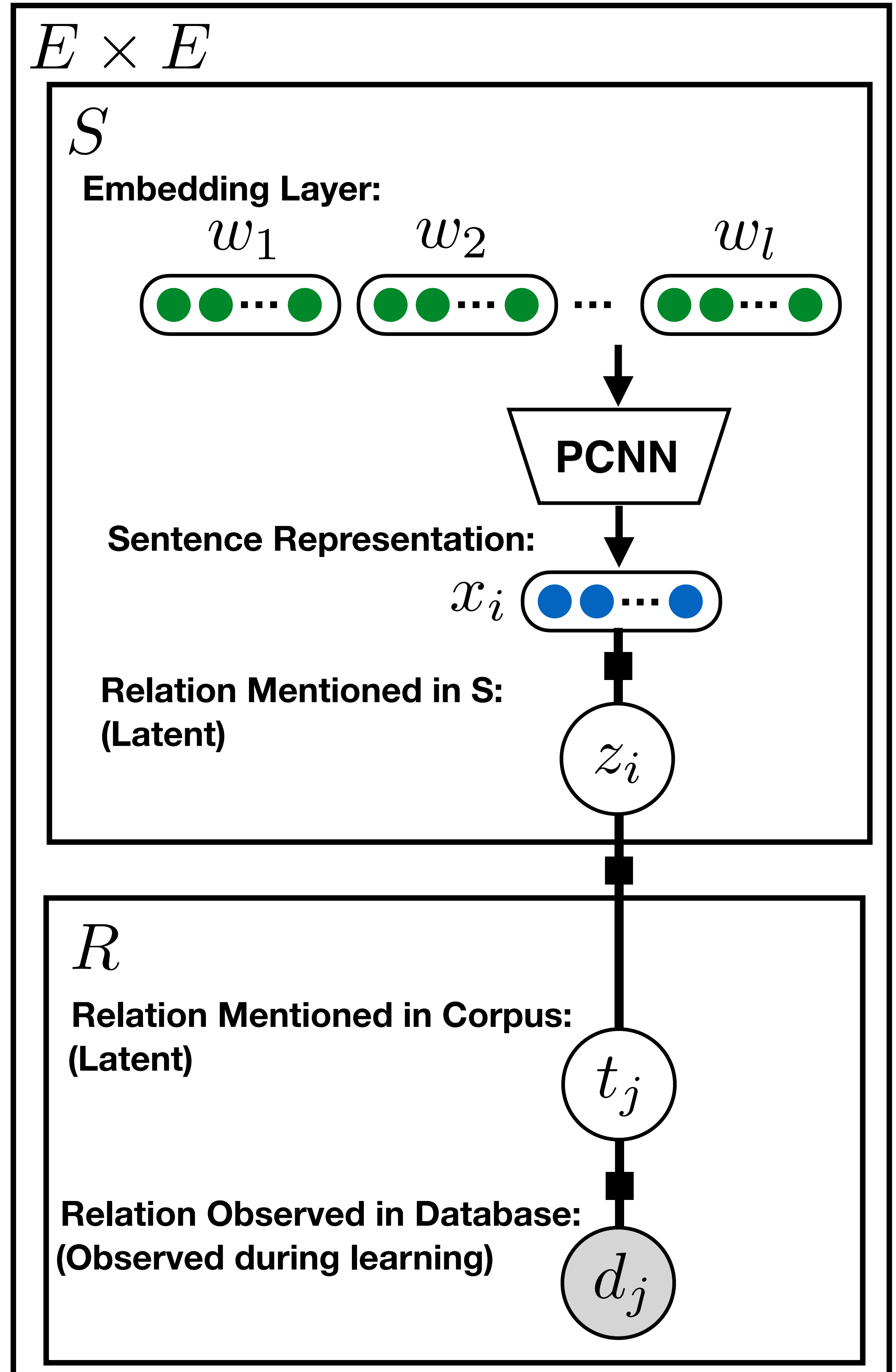
Structured

- (Riedel et. al. ECML 2010) — **NYT / Freebase Dataset + MIL**
- (Hoffmann et. al. ACL 2011) — **Structured Perceptron + Determ. OR Mention-Level Evaluation**
- (Surdeanu et. al. EMNLP 2012) — **Multi-Instance, Multi-Label Learning**
- (Ritter et. al. TACL 2013) — **Missing Data (Mention-Level Evaluation)**

Neural

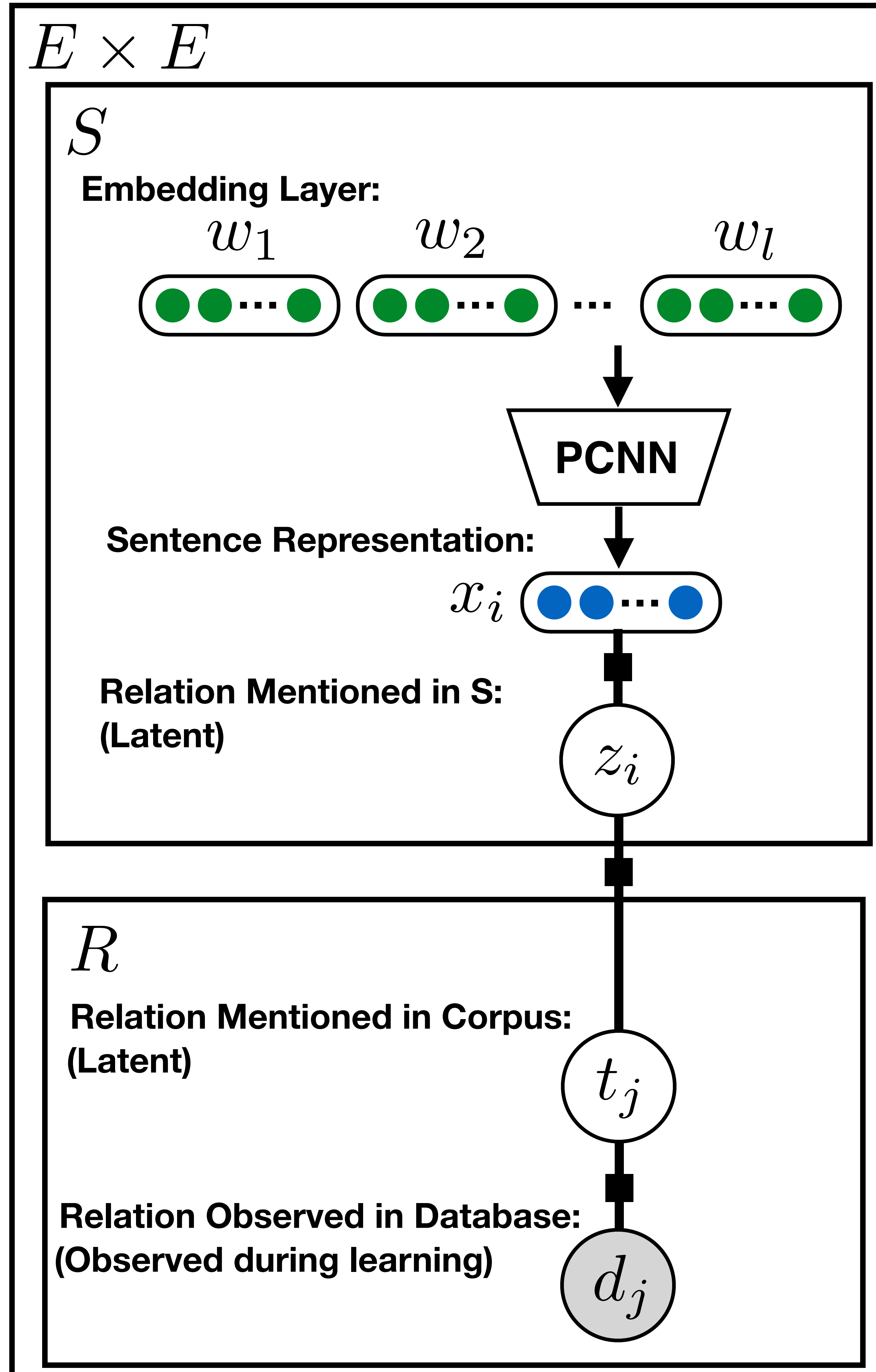
- (Zeng et. al. ACL 2015) — **Piecewise CNN**
- (Lin et. al. ACL 2016) — **Piecewise CNN + Selective Attention**

[Bai and Ritter, In Submission]



[Bai and Ritter, In Submission]

Idea: Combine the benefits of **structured** and **neural** approaches

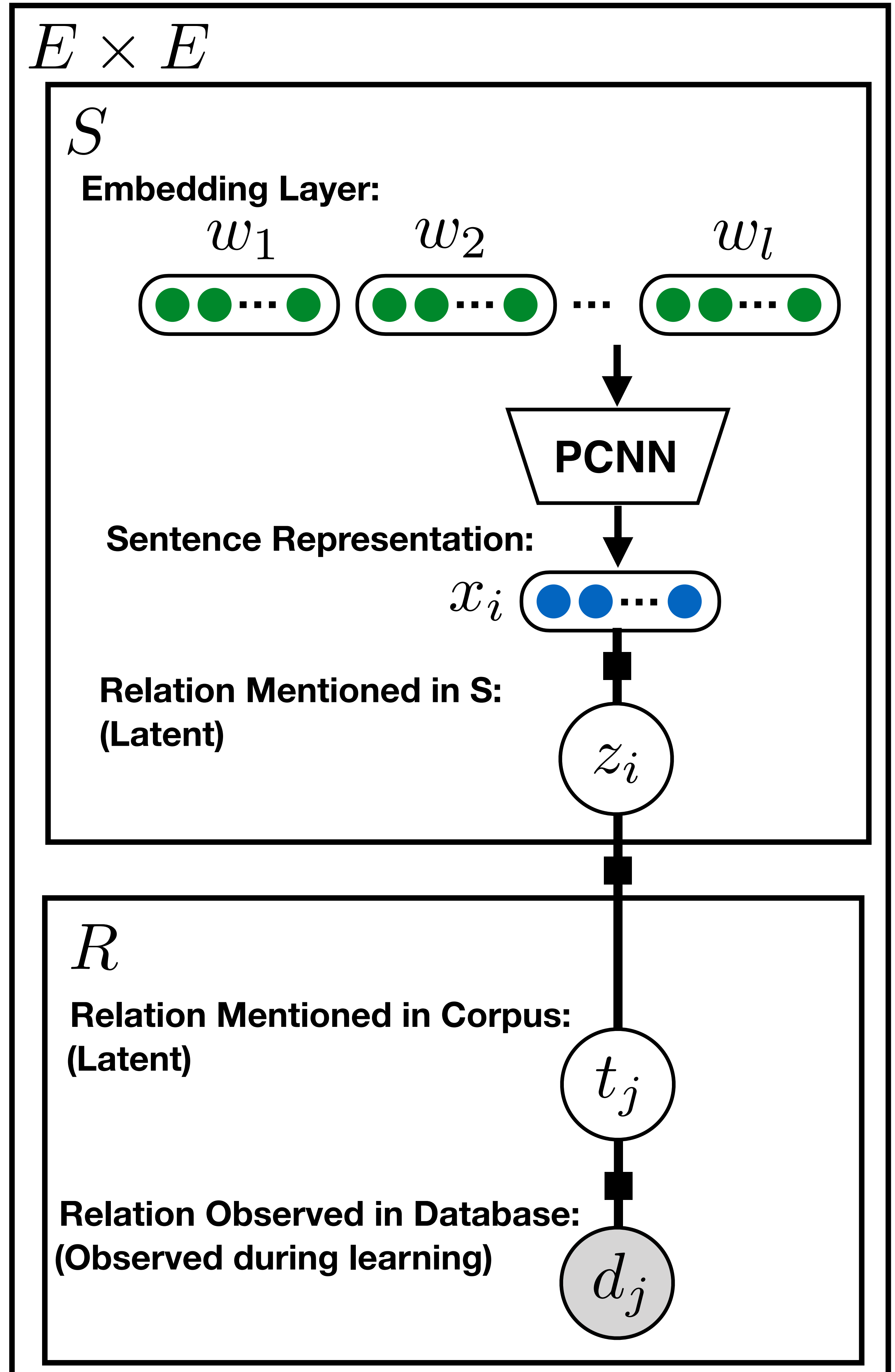


[Bai and Ritter, In Submission]

Idea: Combine the benefits of **structured** and **neural** approaches

★ Learned representations

★ Multi-task learning (e.g. word embeddings)



[Bai and Ritter, In Submission]

Idea: Combine the benefits of **structured** and **neural** approaches

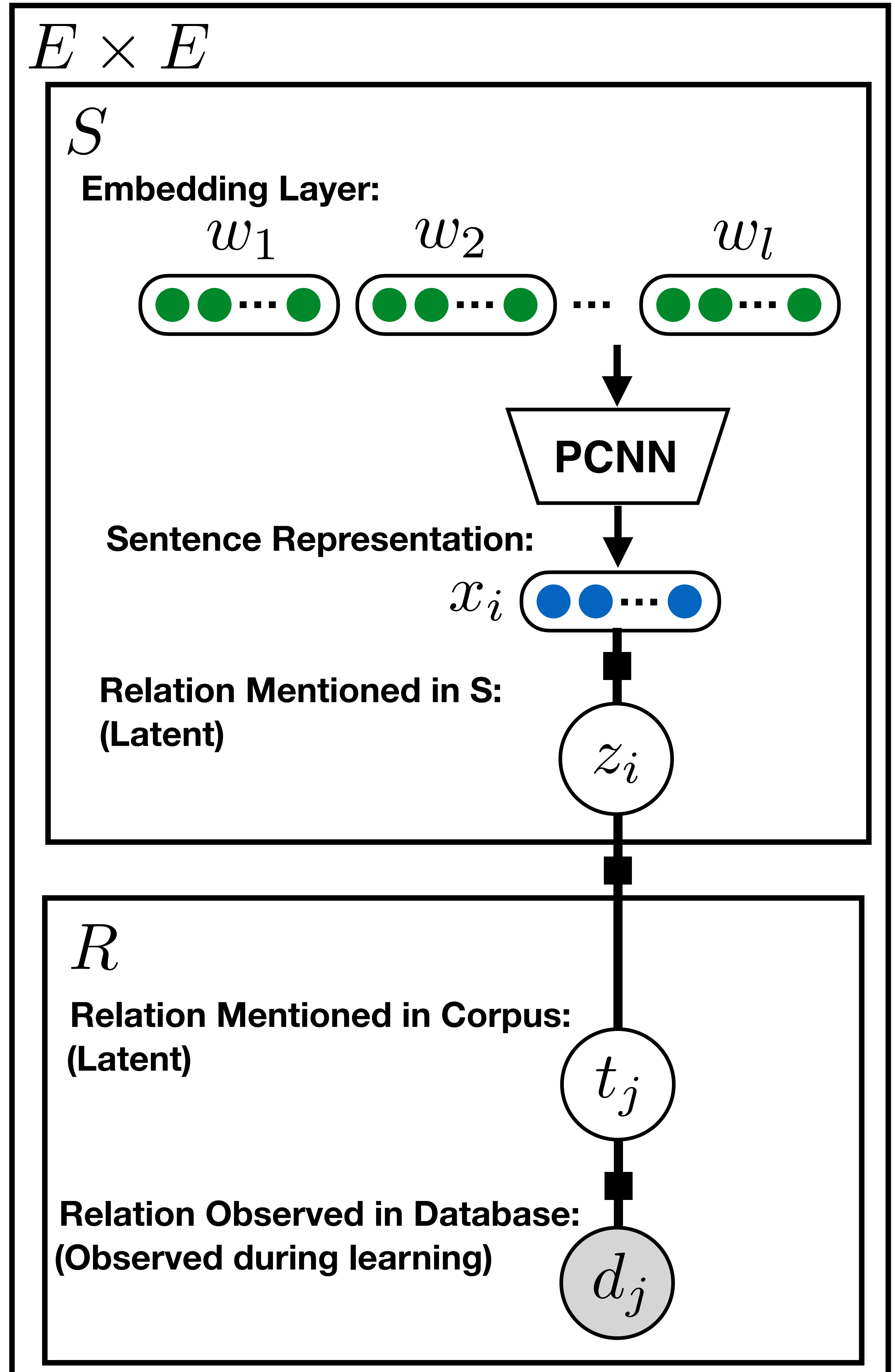
★ Learned representations

★ Multi-task learning (e.g. word embeddings)

★ Inference during learning

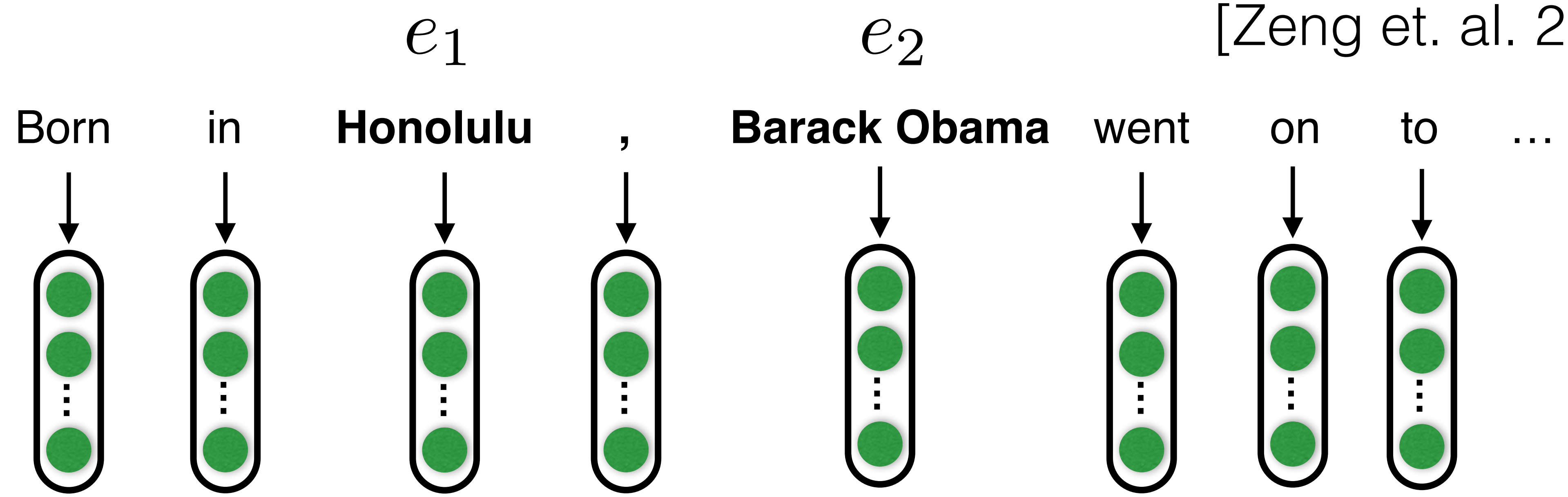
★ Modeling missing data

★ Overlapping relations



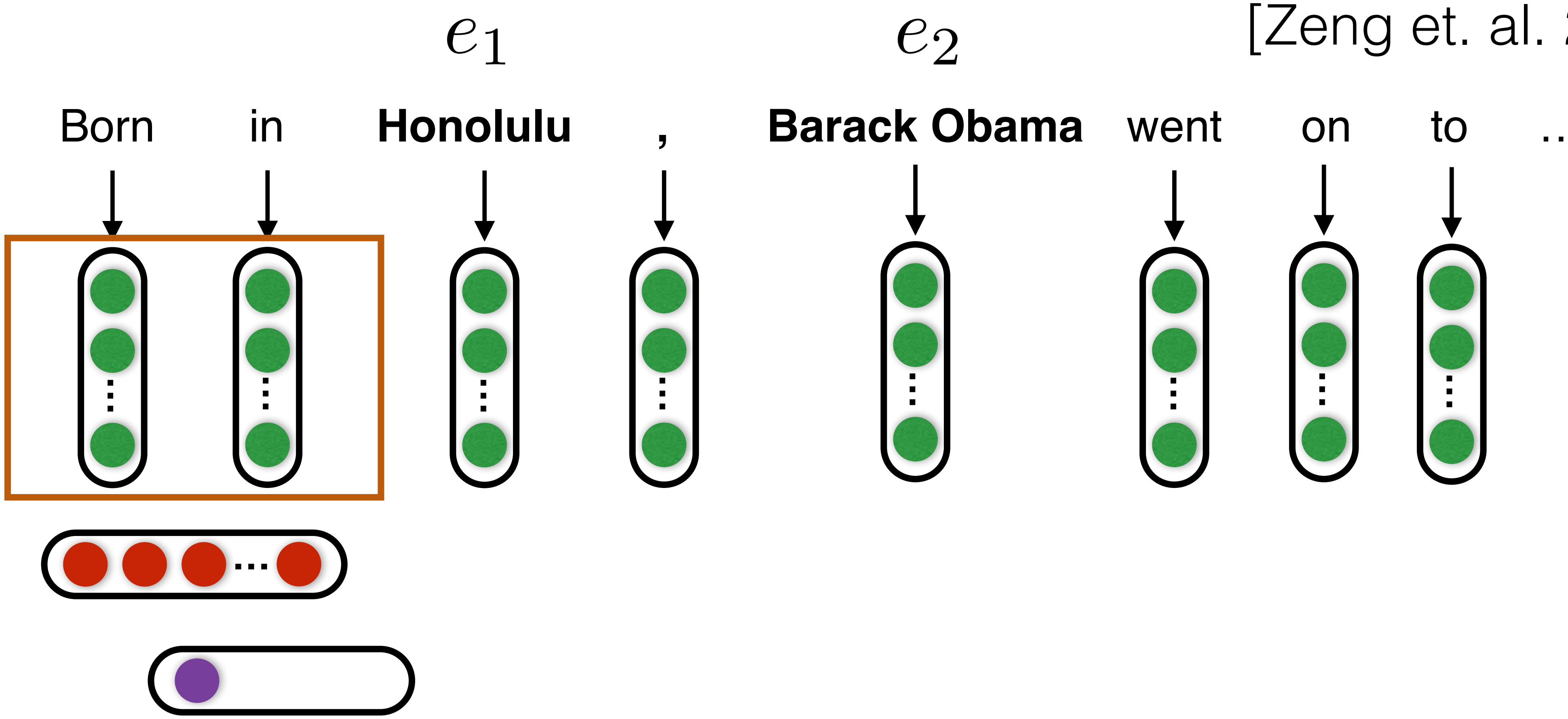
Piecewise Convolutional Neural Networks (PCNN)

[Zeng et. al. 2015]



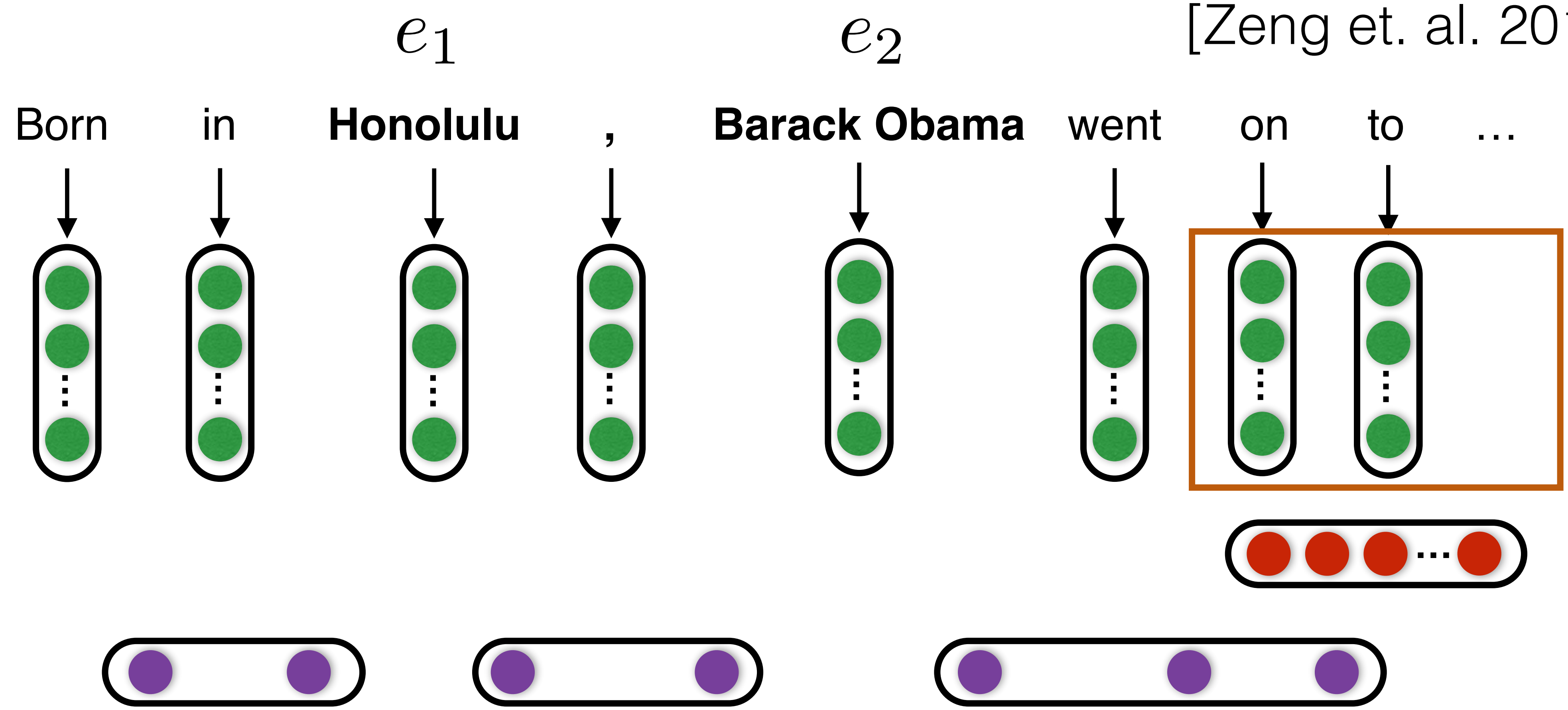
Piecewise Convolutional Neural Networks (PCNN)

[Zeng et. al. 2015]



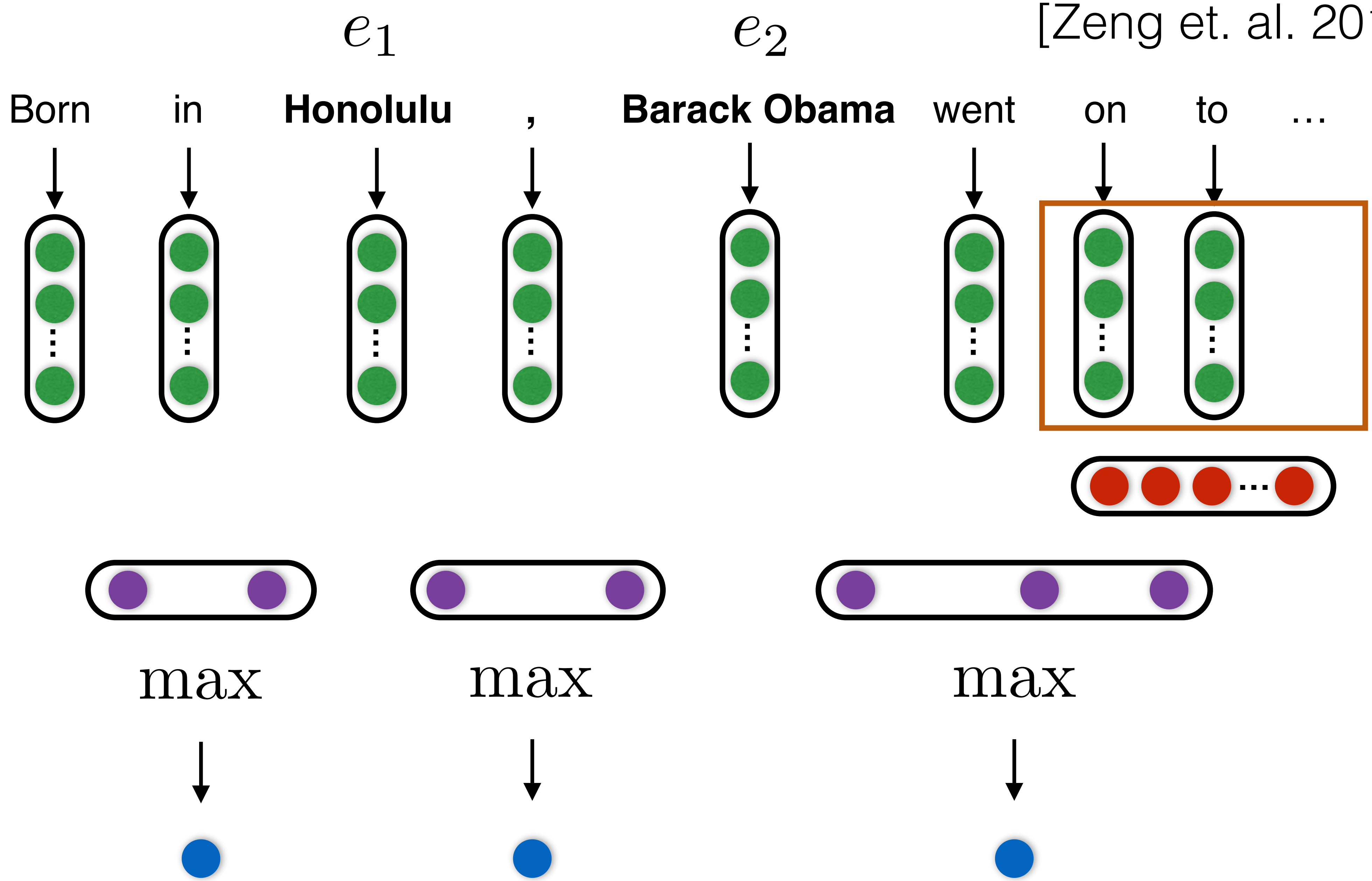
Piecewise Convolutional Neural Networks (PCNN)

[Zeng et. al. 2015]



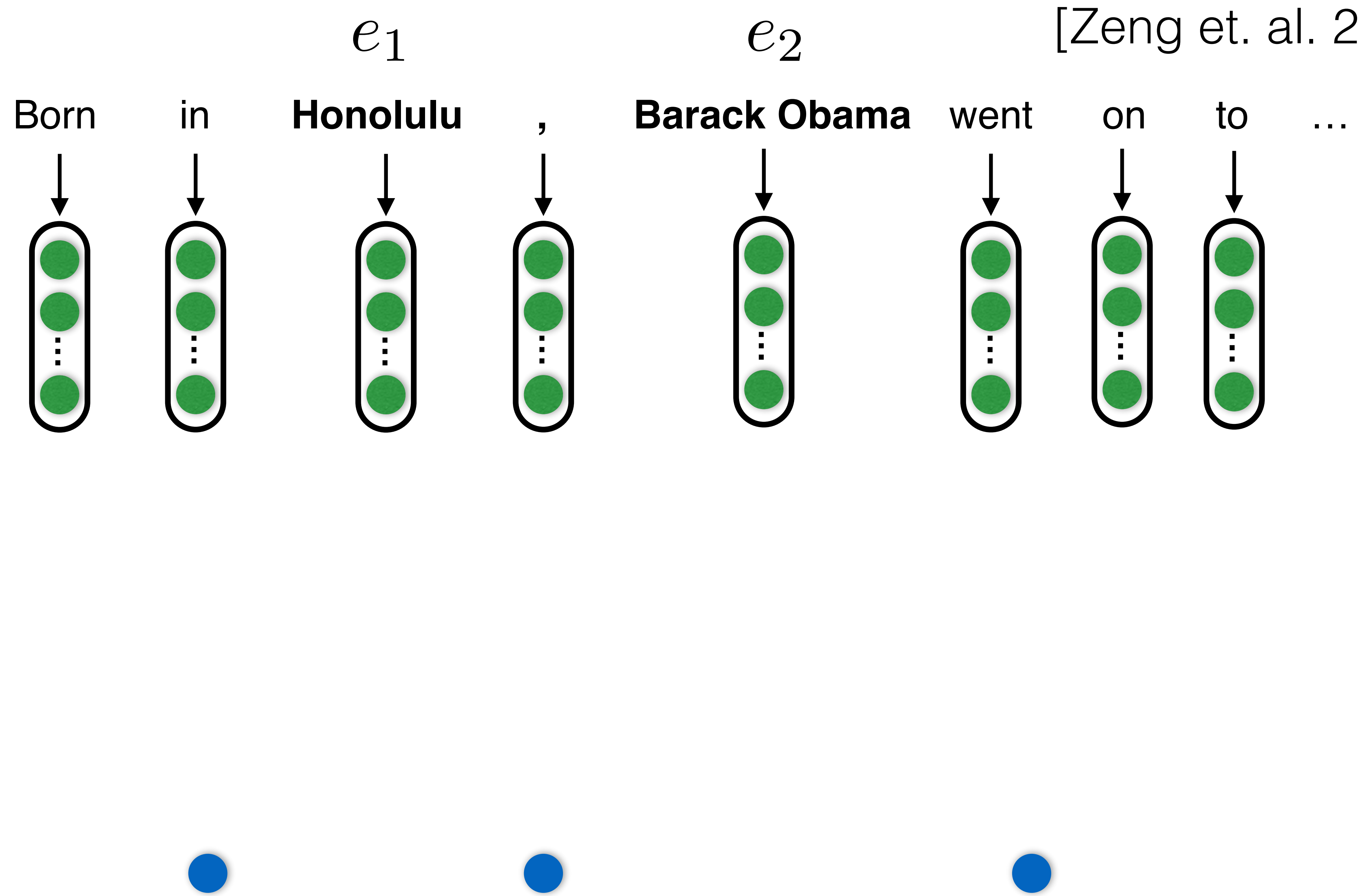
Piecewise Convolutional Neural Networks (PCNN)

[Zeng et. al. 2015]



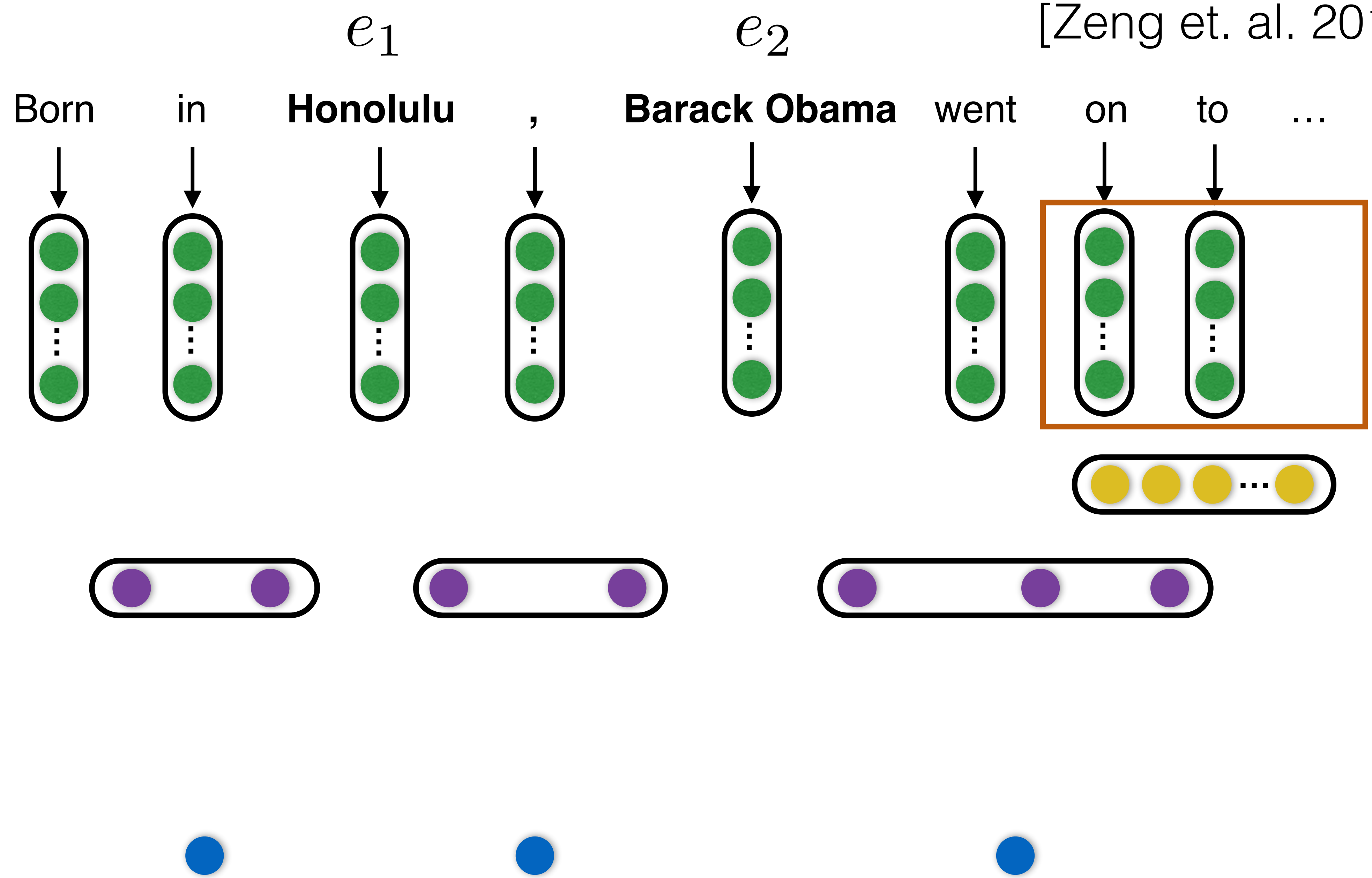
Piecewise Convolutional Neural Networks (PCNN)

[Zeng et. al. 2015]



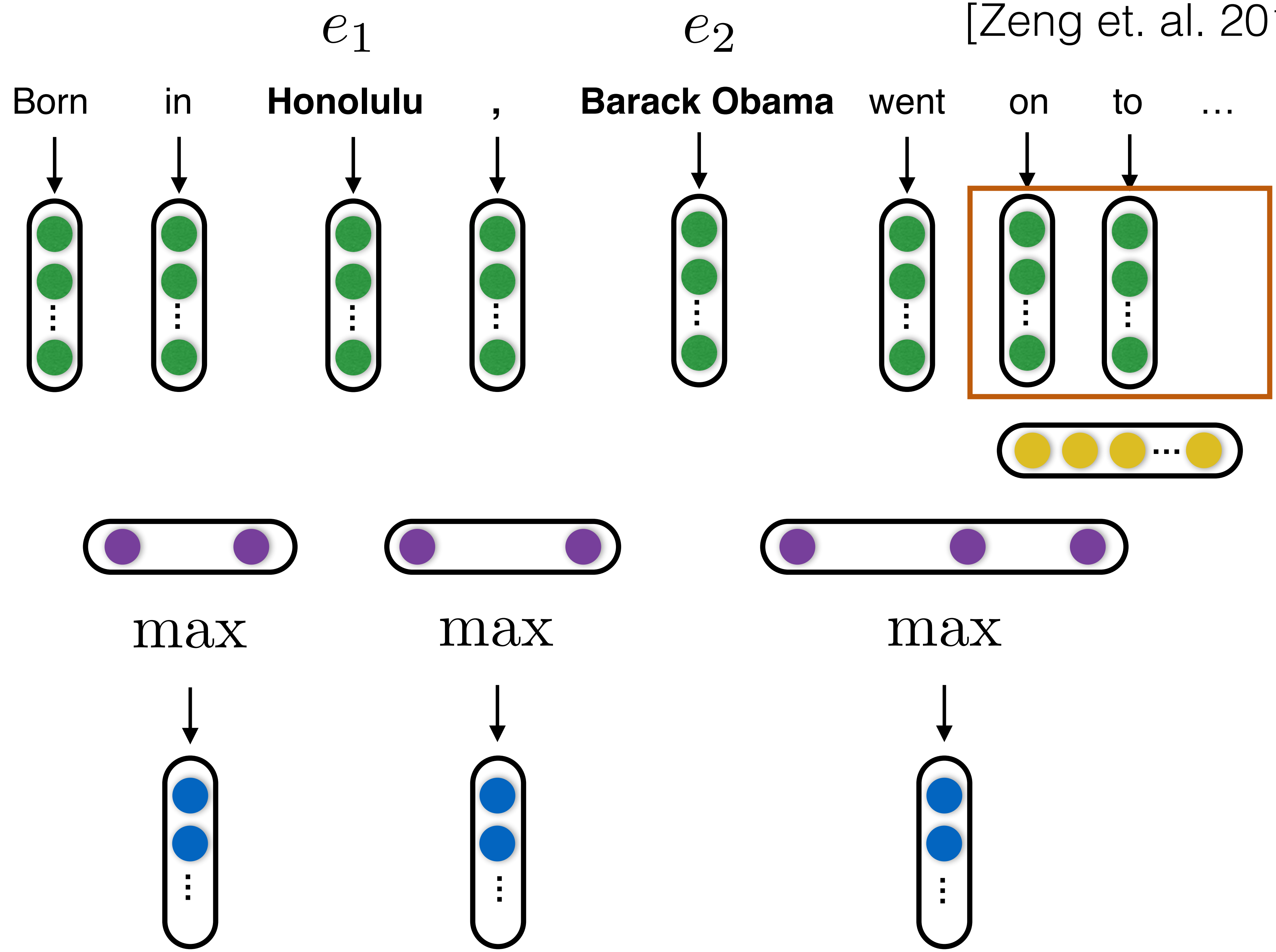
Piecewise Convolutional Neural Networks (PCNN)

[Zeng et. al. 2015]



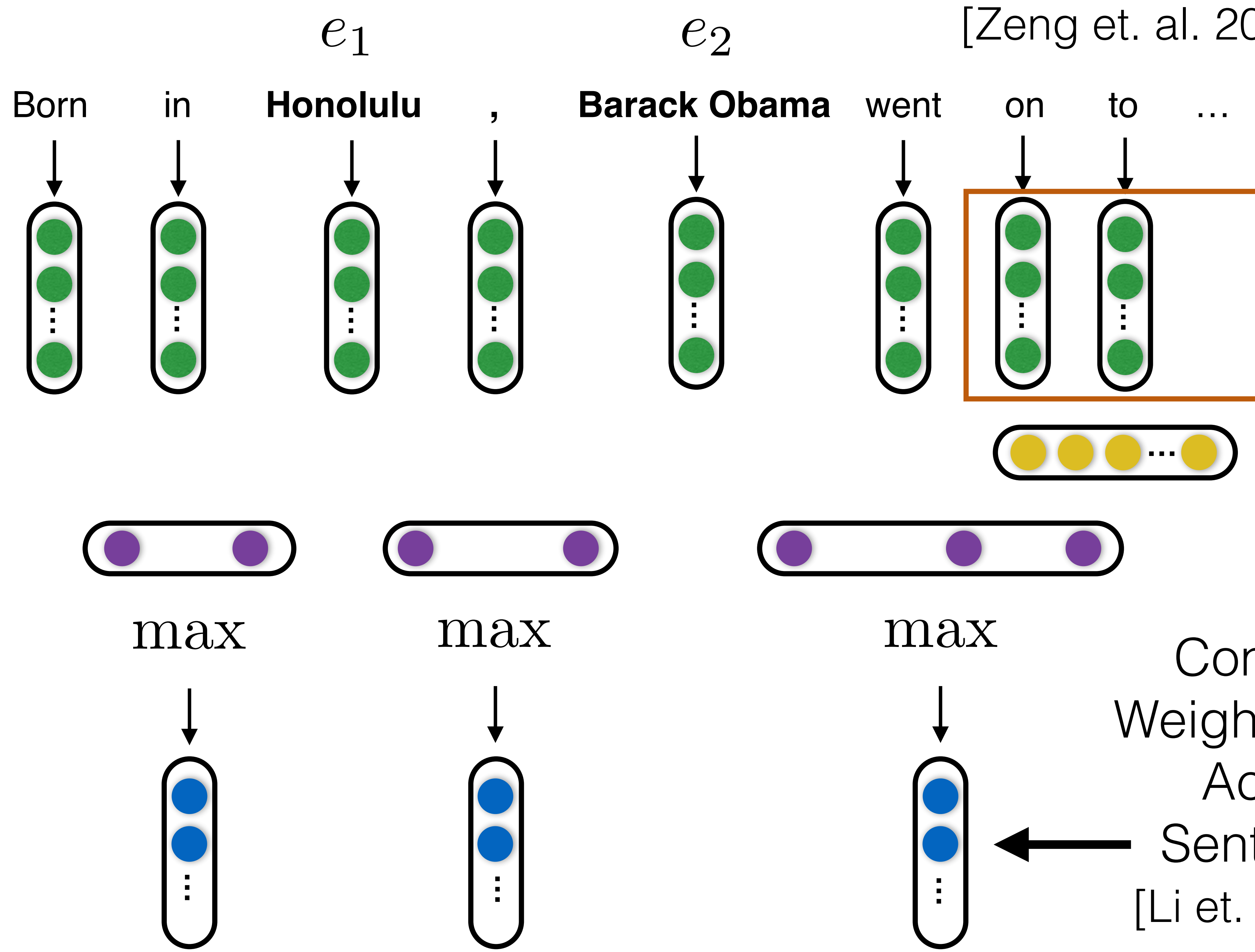
Piecewise Convolutional Neural Networks (PCNN)

[Zeng et. al. 2015]



Piecewise Convolutional Neural Networks (PCNN)

[Zeng et. al. 2015]

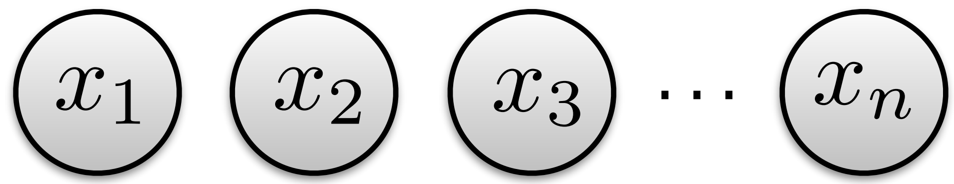


Concat + Weighted Sum Across Sentences
[Li et. al. 2016]

Mention-Level Model

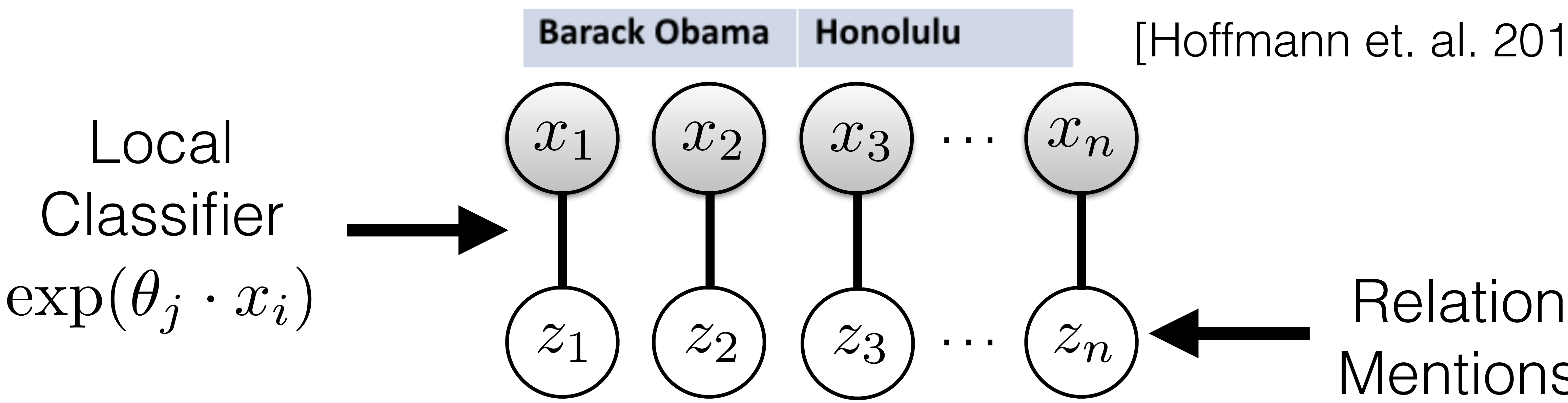
[Hoffmann et. al. 2011]

Barack Obama	**Honolulu**



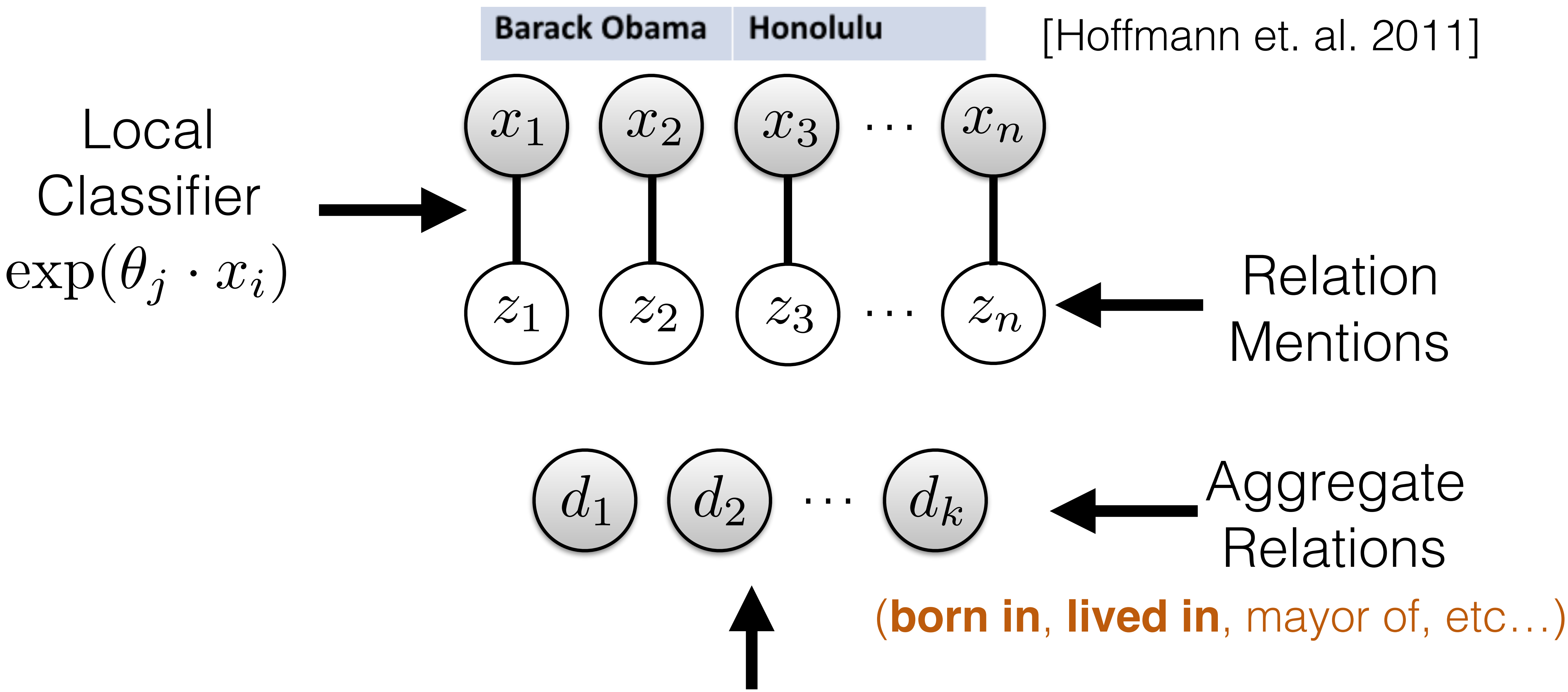
Mention-Level Model

[Hoffmann et. al. 2011]



Mention-Level Model

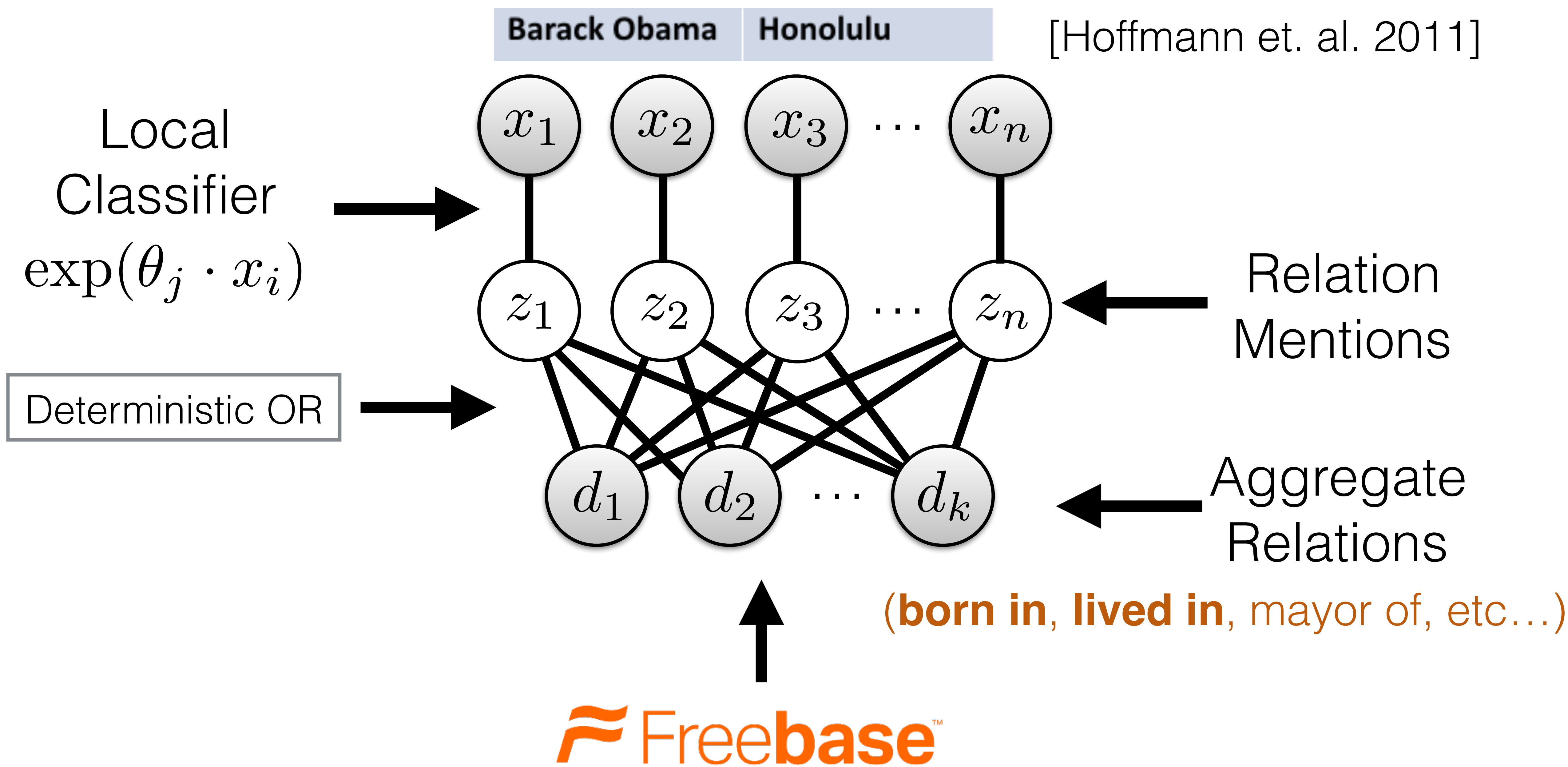
[Hoffmann et. al. 2011]



Freebase™

Mention-Level Model

[Hoffmann et. al. 2011]



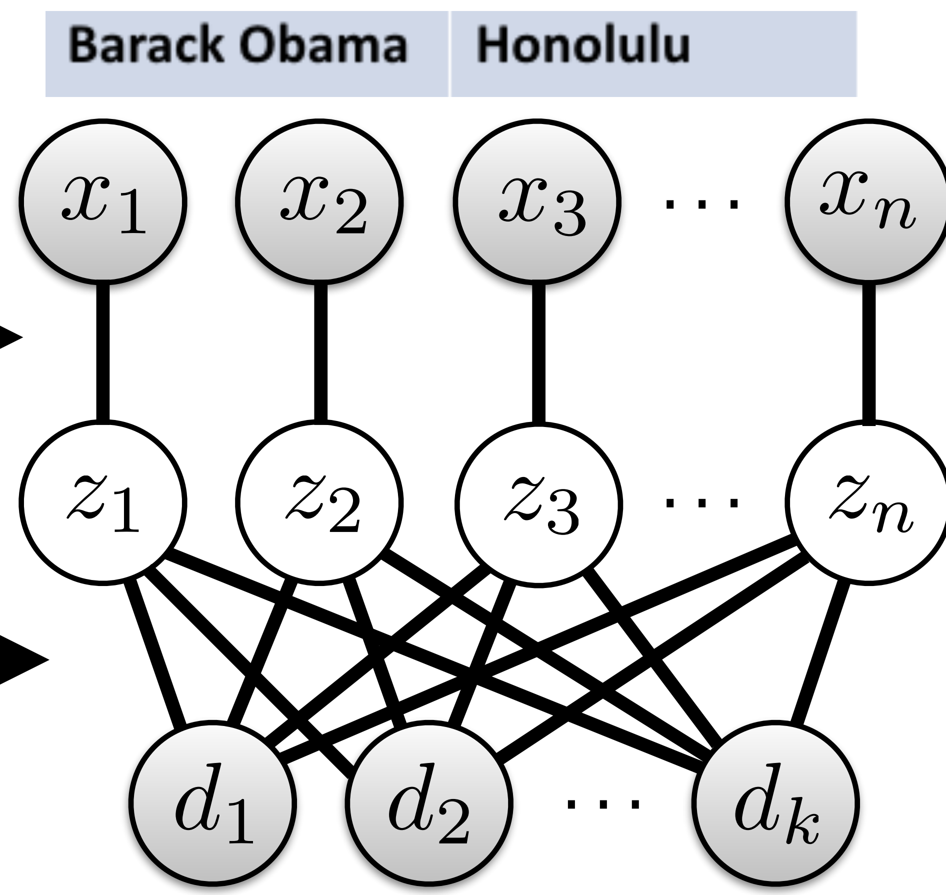
Mention-Level Model

[Hoffmann et. al. 2011]

Local Classifier

$\exp(\theta_j \cdot x_i)$

Deterministic OR



Relation Mentions

Aggregate Relations

(**born in**, **lived in**, mayor of, etc…)

Freebase

Maximize Conditional Likelihood:

$$\sum_z P(z, d|s; \theta)$$

Missing Data Problems

Two assumptions Drive Learning:

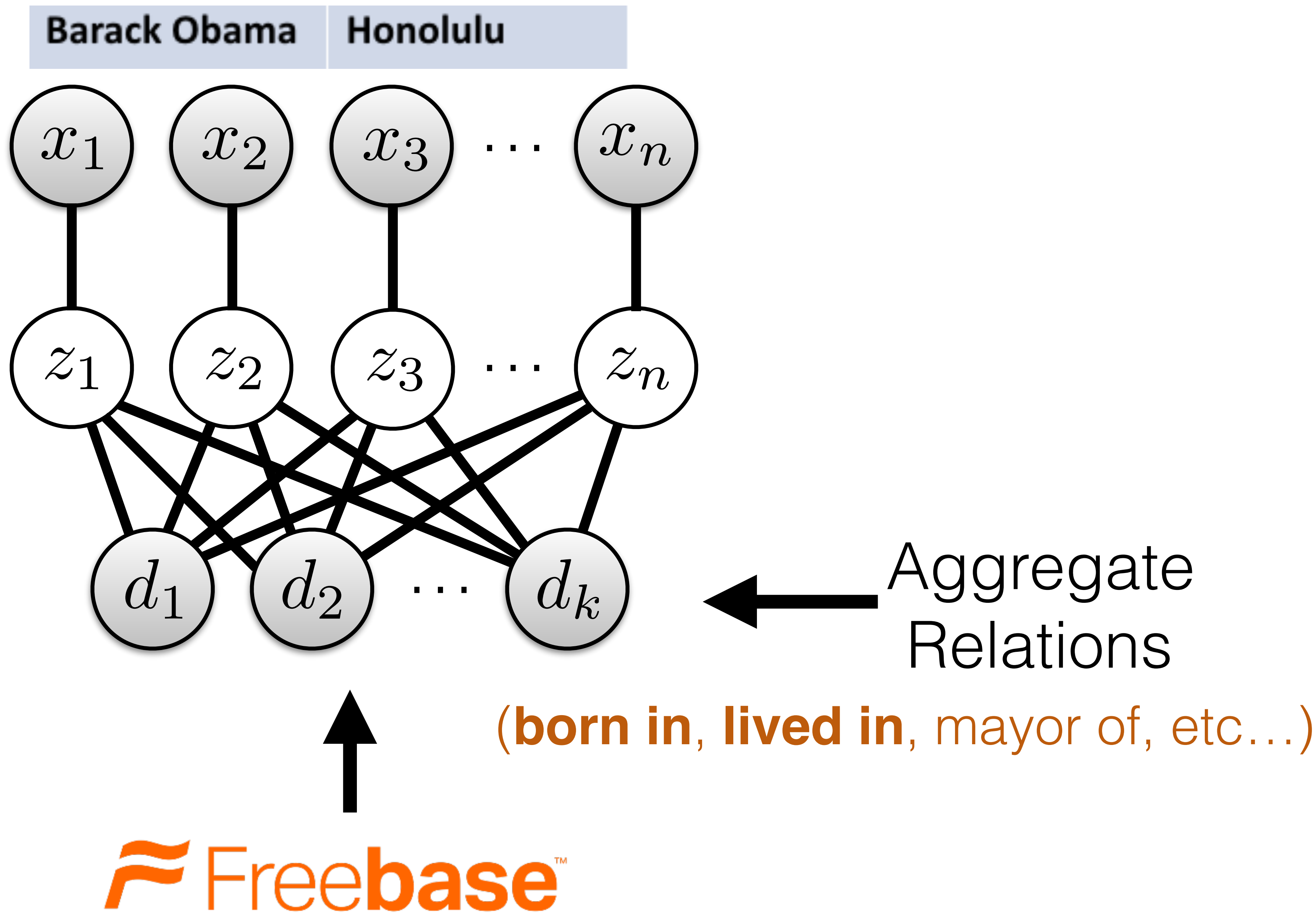
- Not in KB → Not in text
- In KB → Must be mentioned at least once

Leads to errors in training data:

- False negatives
- False positives

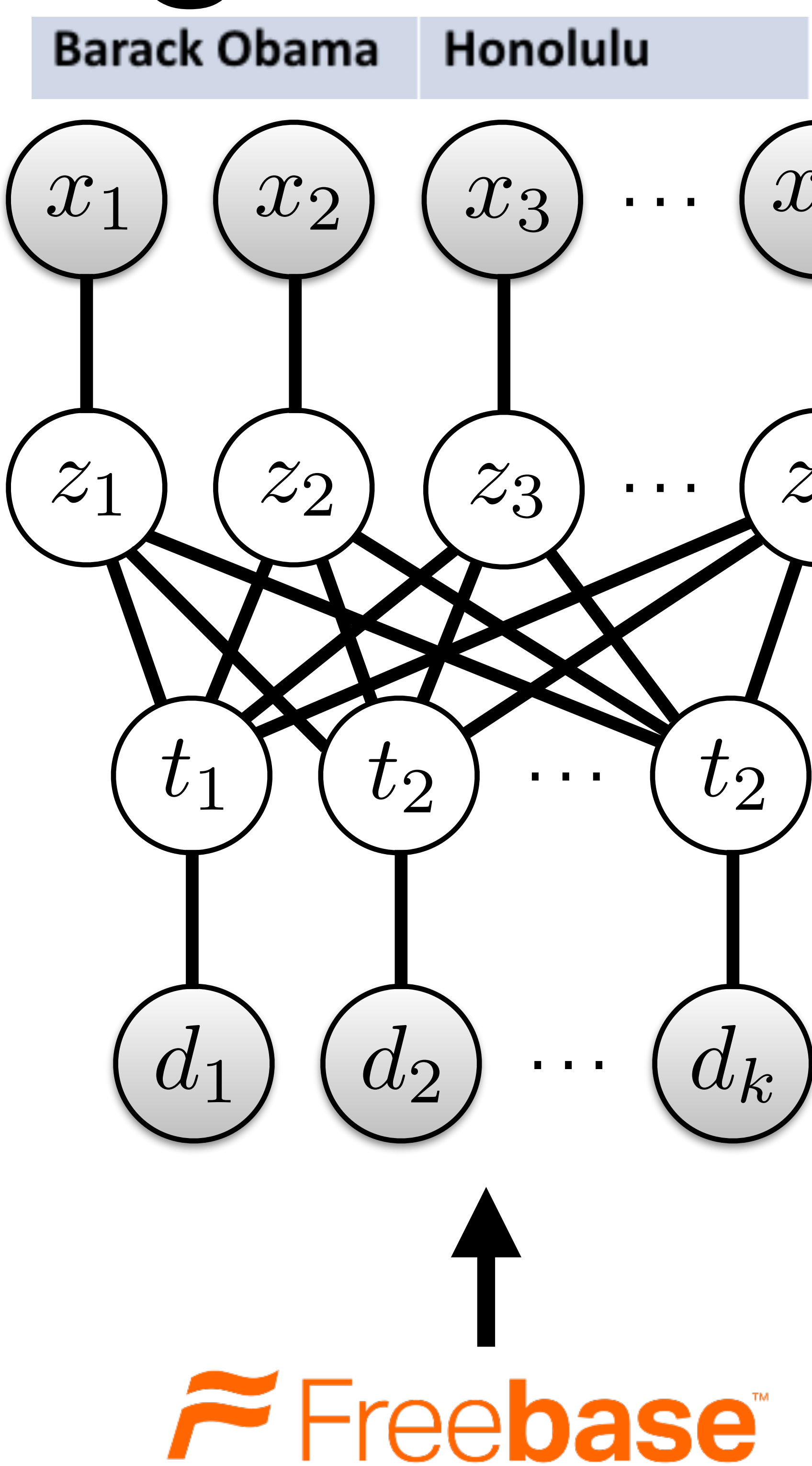
[Ritter, et. al. TACL 2013]

Modeling Missing Data



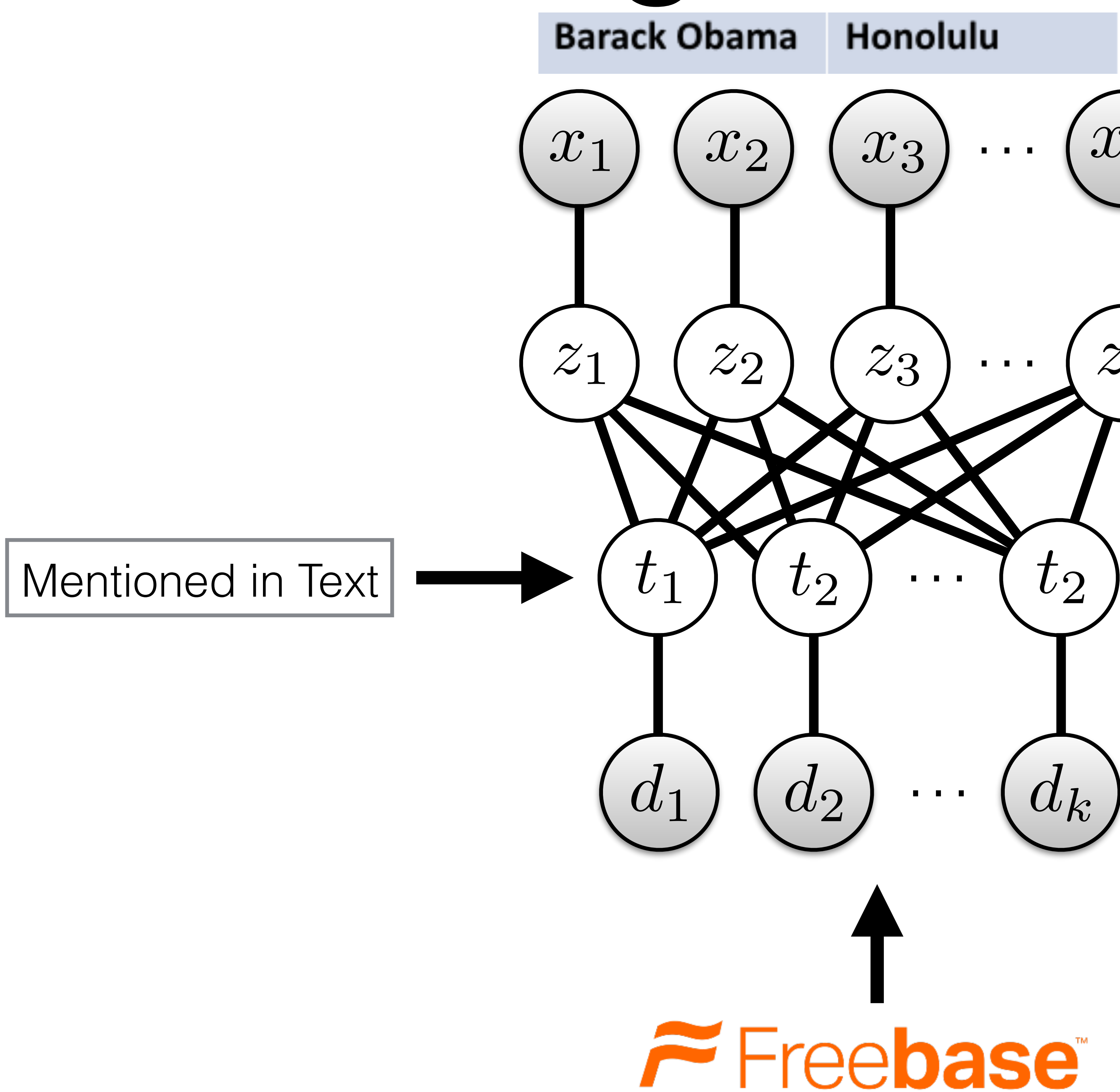
[Ritter, et. al. TACL 2013]

Modeling Missing Data



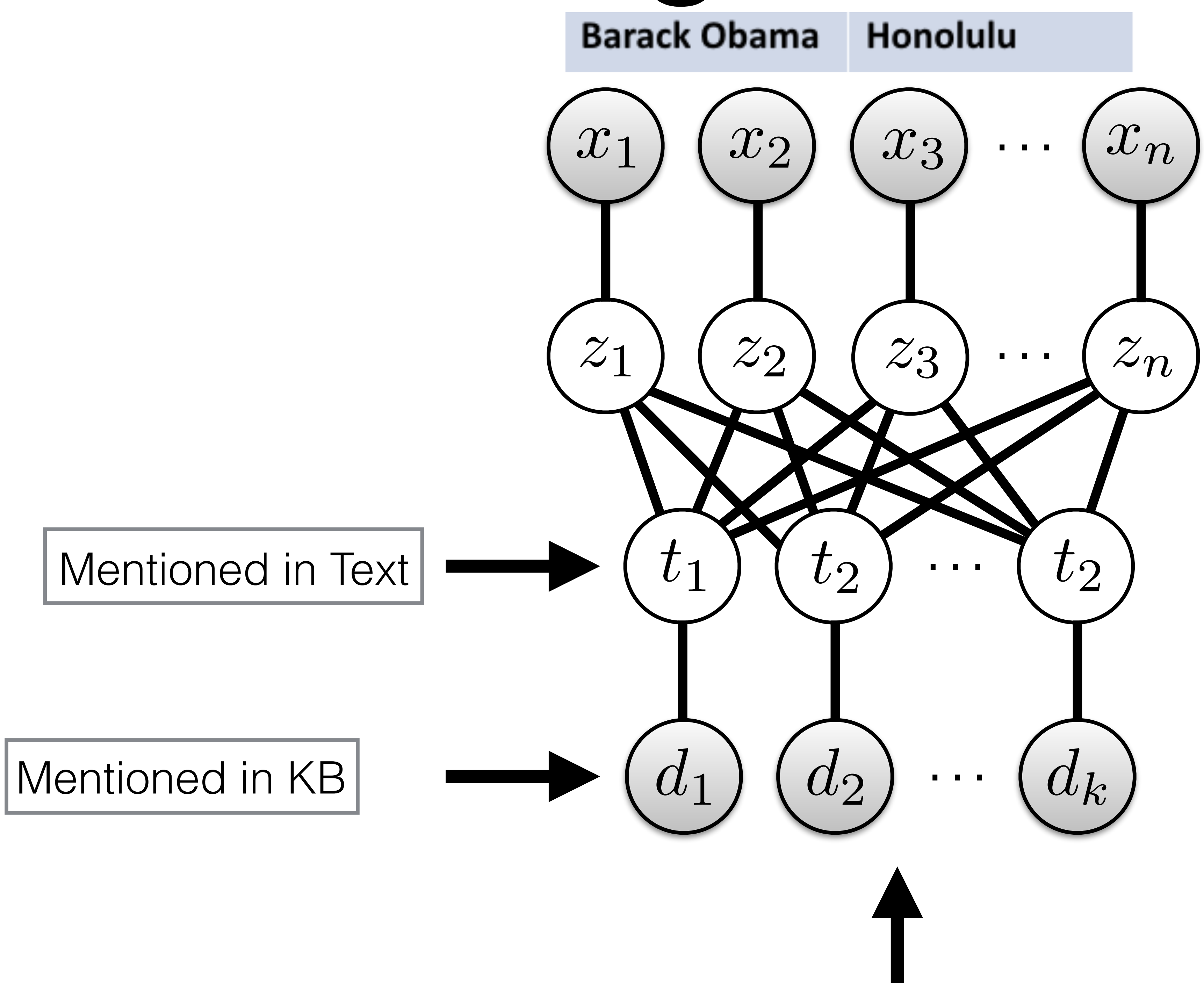
[Ritter, et. al. TACL 2013]

Modeling Missing Data



[Ritter, et. al. TACL 2013]

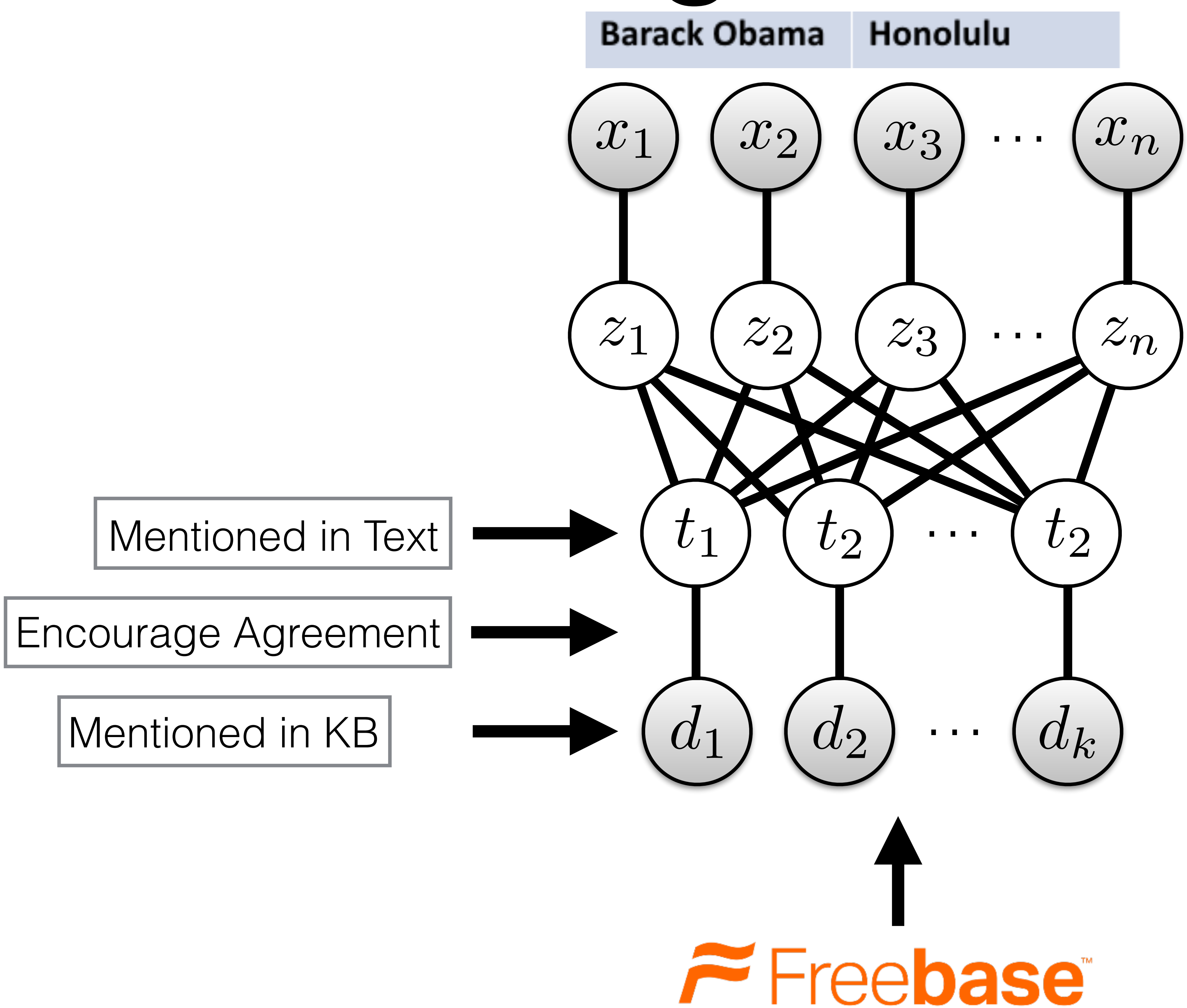
Modeling Missing Data



Freebase™

[Ritter, et. al. TACL 2013]

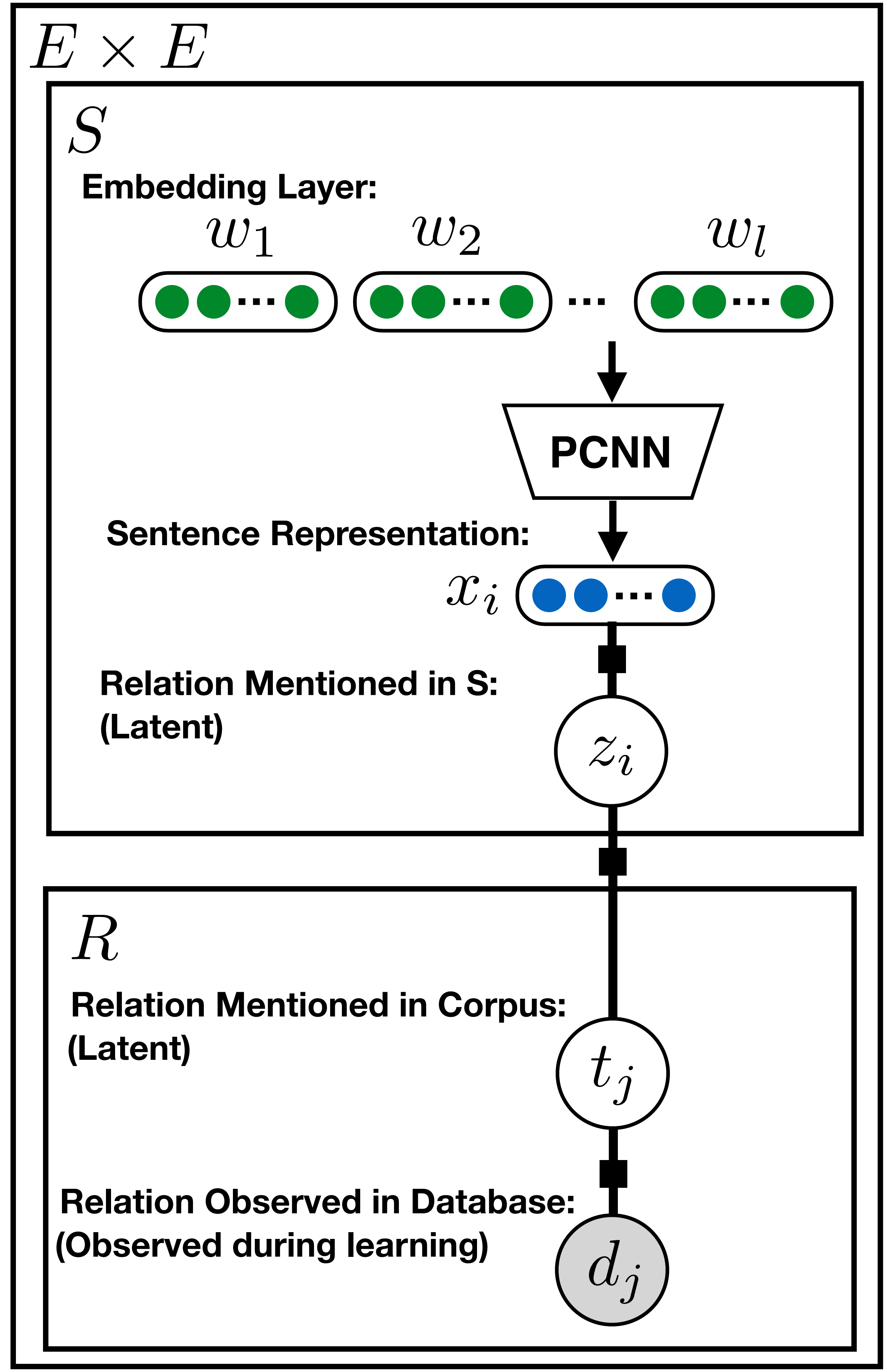
Modeling Missing Data



Implementation Details

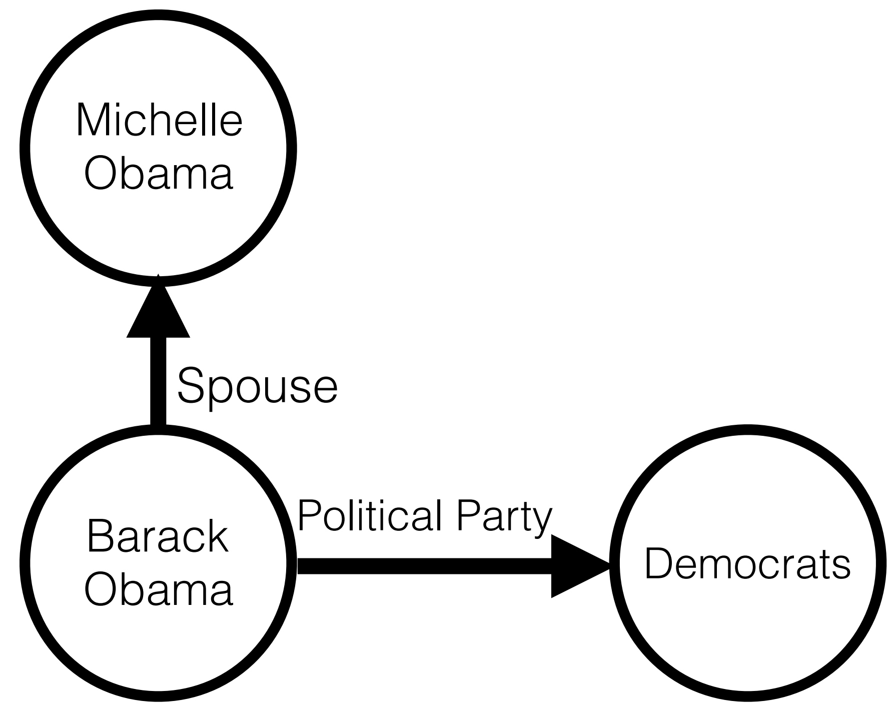
[Bai and Ritter, In Submission]

- Pre-trained sentence embeddings - PCNN+ATT (Lin et. al. 2016)
- MIRA (Cramer and Singer, 2003)
  - Crucial for continuous representations
- Hyperparameter search on dev set



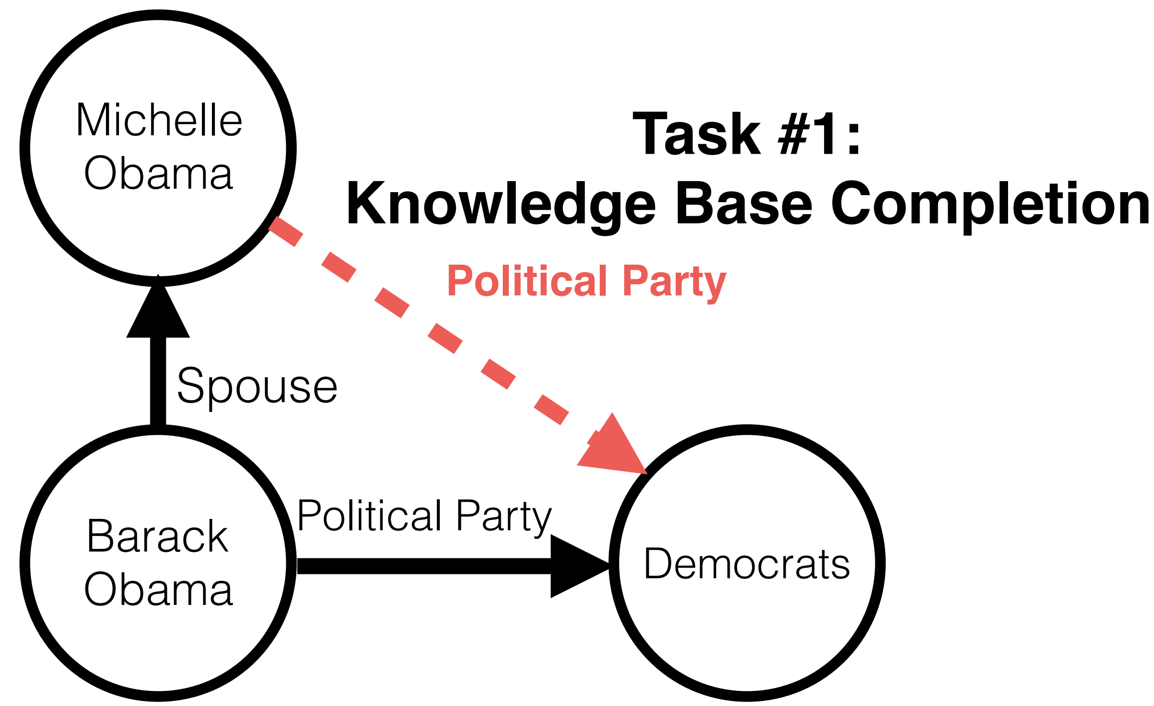
Knowledge Base Completion vs. Relation Extraction

A (very small) knowledge base:



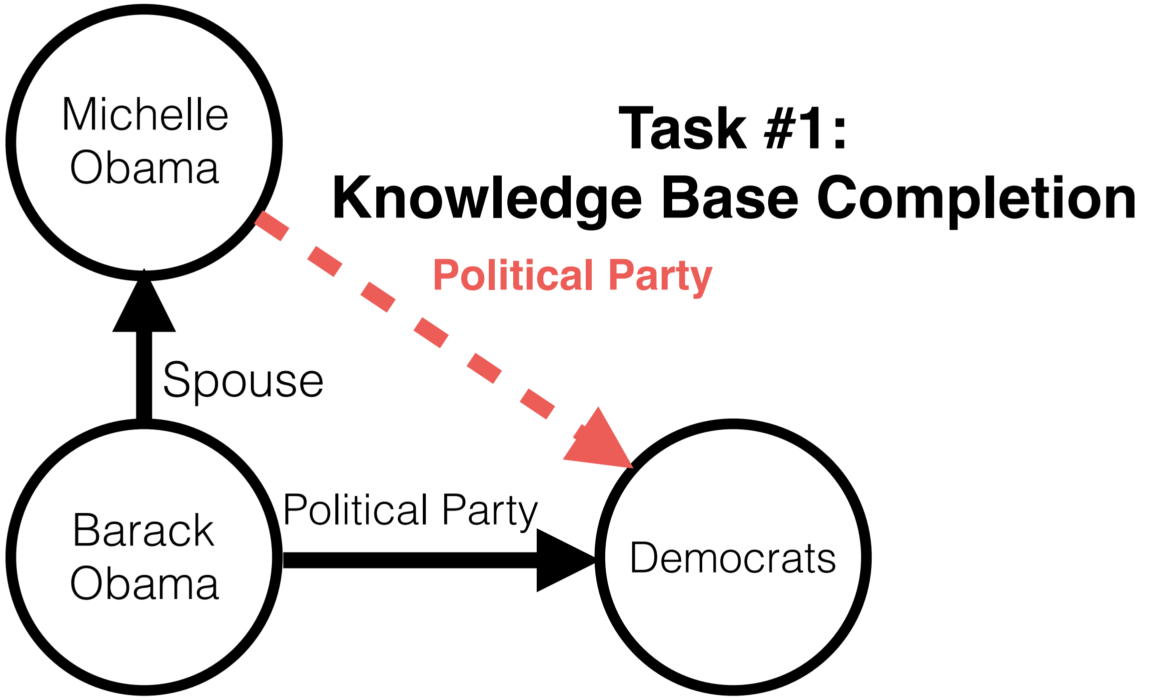
Knowledge Base Completion vs. Relation Extraction

A (very small) knowledge base:



Knowledge Base Completion vs. Relation Extraction

A (very small) knowledge base:



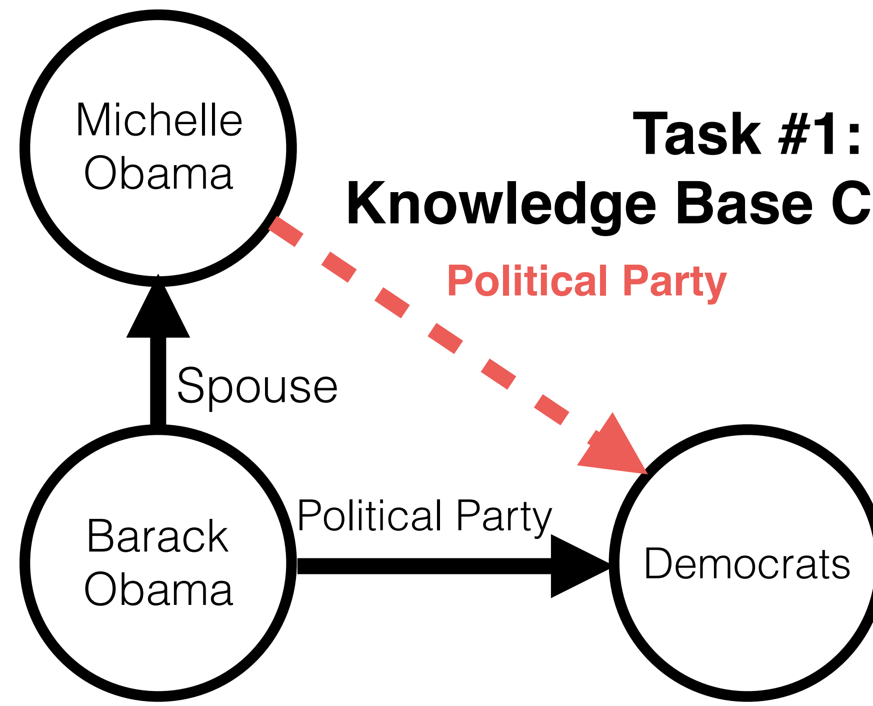
"For the last eight years, **Michelle Obama** has been a rock star in the **Democratic Party**."

$\textsc{Party}(\text{Michelle Obama, Democratic Party})$

Task #2
(Mention-Level) Relation Extraction

Knowledge Base Completion vs. Relation Extraction

A (very small) knowledge base:



Task #1: Knowledge Base Completion

"For the last eight years, **Michelle Obama** has been a rock star in the **Democratic Party**."

$$\textsc{Party}(\text{Michelle Obama}, \text{Democratic Party})$$

I am interested in this → **Task #2 (Mention-Level) Relation Extraction**

Issues with Evaluation…

[Bai and Ritter, In Submission]

- **Freebase / NYT Dataset**

Freebase™

Linguistic Data Consortium

The New York Times Annotated Corpus

Issues with Evaluation…

[Bai and Ritter, In Submission]

- **Freebase / NYT Dataset**
  - Originally created by Riedel, Yao and McCallum 2010

Freebase™

LDC

Linguistic Data Consortium

The New York Times Annotated Corpus

Issues with Evaluation…

[Bai and Ritter, In Submission]

- **Freebase / NYT Dataset**
    - Originally created by Riedel, Yao and McCallum 2010
    - Several Versions of this Dataset in Use

Freebase™

LDC Linguistic Data Consortium

The New York Times Annotated Corpus

Issues with Evaluation…

[Bai and Ritter, In Submission]

- **Freebase / NYT Dataset**
  - Originally created by Riedel, Yao and McCallum 2010
  - Several Versions of this Dataset in Use
  - A Recent Version Contains Overlapping Entity Pairs in Train / Test Split (we removed these)

Freebase™

LDC Linguistic Data Consortium

The New York Times Annotated Corpus

Mention-Level Prediction in PCNN+ATT Baseline

- Use attention parameters to score mentions
  - Pre-trained using PCNN+ATT (Lin et. al. 2016)
- Softmax over relations

$$P(z|x_i) = \frac{\exp(r_z \cdot x_i)}{\sum_{k=1}^{n_r} \exp(r_k \cdot x_i)}$$

Mention-Level Prediction in PCNN+ATT Baseline

- Use attention parameters to score mentions
  - Pre-trained using PCNN+ATT (Lin et. al. 2016)
- Softmax over relations

$$P(z|x_i) = \frac{\exp(r_z \cdot x_i)}{\sum_{k=1}^{n_r} \exp(r_k \cdot x_i)}$$

Note: Attention Layer of PCNN+ATT uses softmax over sentences

Mention-Level Results (Riedel et. al. dataset)

Model	**DEV**	**TEST**
MultiR_sparse	66.2	63.2
MultiR_sparse_MIRA	76.1	75.4
MultiR_continuous	74.3	66.6
MultiR_continuous_MIRA	80.9	74.7
DNMAR_sparse	77.8	74.4
DNMAR_sparse_MIRA	79.4	76.2
DNMAR_continuous	78.7	73.1
DNMAR_continuous_MIRA	**83.7**	**78.3**
PCNN+ATT	81.4	76.4

Mention-Level Results (Riedel et. al. dataset)

Model	DEV	TEST
MultiR_sparse	66.2	$\boxed{63.2}$
MultiR_sparse_MIRA	76.1	$\boxed{75.4}$
MultiR_continuous	74.3	66.6
MultiR_continuous_MIRA	80.9	74.7
DNMAR_sparse	77.8	$\boxed{74.4}$
DNMAR_sparse_MIRA	79.4	$\boxed{76.2}$
DNMAR_continuous	78.7	73.1
DNMAR_continuous_MIRA	**83.7**	**78.3**
PCNN+ATT	81.4	$\boxed{76.4}$

$$P(z|x_i) = \frac{\exp(r_z \cdot x_i)}{\sum_{k=1}^{n_r} \exp(r_k \cdot x_i)}$$

Mention-Level Results (Riedel et. al. dataset)

Model	**DEV**	**TEST**
MultiR_sparse	66.2	63.2
MultiR_sparse_MIRA	76.1	75.4
MultiR_continuous	74.3	66.6
MultiR_continuous_MIRA	80.9	74.7
DNMAR_sparse	77.8	74.4
DNMAR_sparse_MIRA	79.4	76.2
DNMAR_continuous	78.7	73.1
DNMAR_continuous_MIRA	**83.7**	**78.3**
PCNN+ATT	81.4	76.4

MIRA helps in all cases and is crucial for continuous representations

Mention-Level Results (Lin et. al. dataset)

Model	**DEV**	**TEST**
MultiR_continuous	66.1	59.9
MultiR_continuous_MIRA	69.2	67.8
DNMAR_continuous	67.4	62.9
DNMAR_continuous_MIRA	**74.8**	**73.4**
PCNN+ATT	67.9	72.1

Mention-Level Results (Lin et. al. dataset)

Model	**DEV**	**TEST**
MultiR_continuous	66.1	59.9
MultiR_continuous_MIRA	69.2	67.8
DNMAR_continuous	67.4	62.9
DNMAR_continuous_MIRA	**74.8**	**73.4**
PCNN+ATT	67.9	72.1

$$P(z|x_i) = \frac{\exp(r_z \cdot x_i)}{\sum_{k=1}^{n_r} \exp(r_k \cdot x_i)}$$

Mention-Level Results (Lin et. al. dataset)

Model	DEV	TEST
MultiR_continuous	66.1	59.9
MultiR_continuous_MIRA	69.2	$\boxed{67.8}$
DNMAR_continuous	67.4	62.9
DNMAR_continuous_MIRA	**74.8**	$\boxed{\mathbf{73.4}}$
PCNN+ATT	67.9	72.1

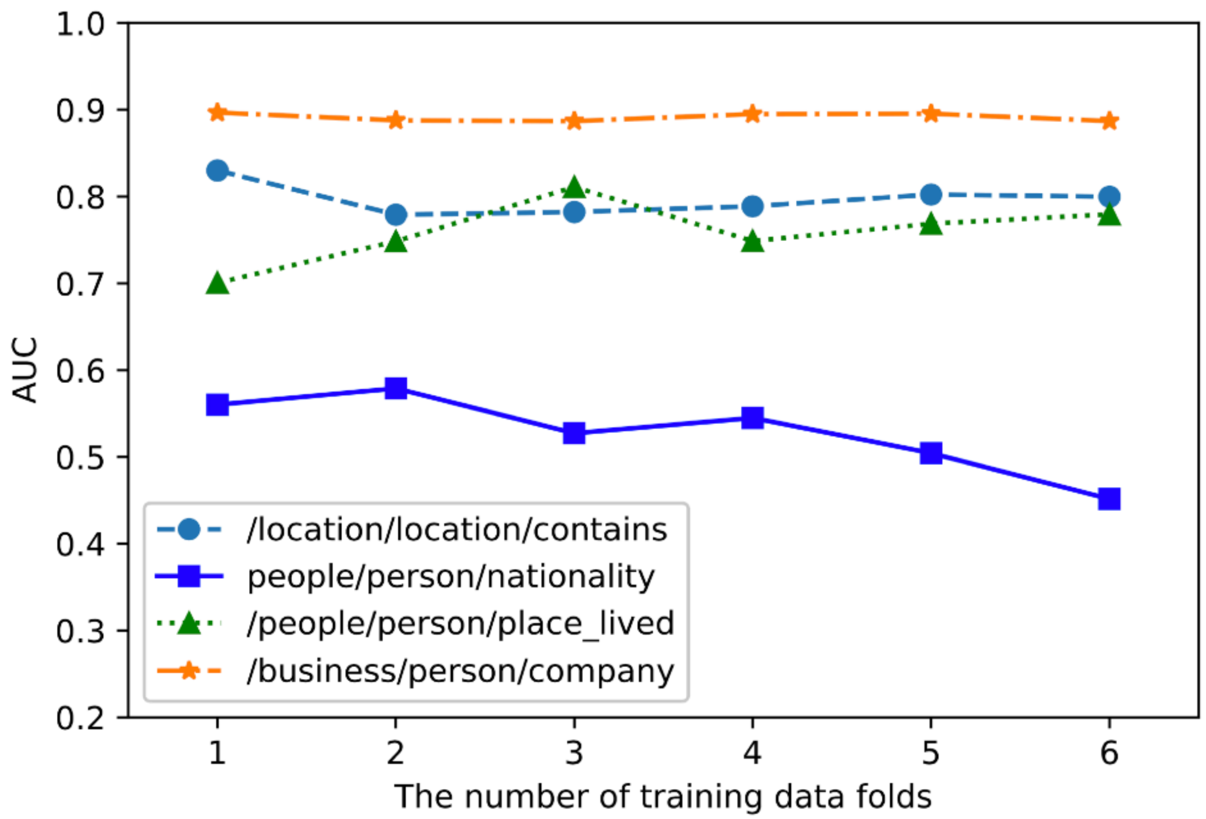
Mention-Level Results (Lin et. al. dataset)

Model	DEV	TEST	
MultiR_continuous	66.1	59.9	
MultiR_continuous_MIRA	69.2	67.8	
DNMAR_continuous	67.4	62.9	
DNMAR_continuous_MIRA	**74.8**	**73.4**	**78.3**
PCNN+ATT	67.9	72.1	

More data hurts performance?

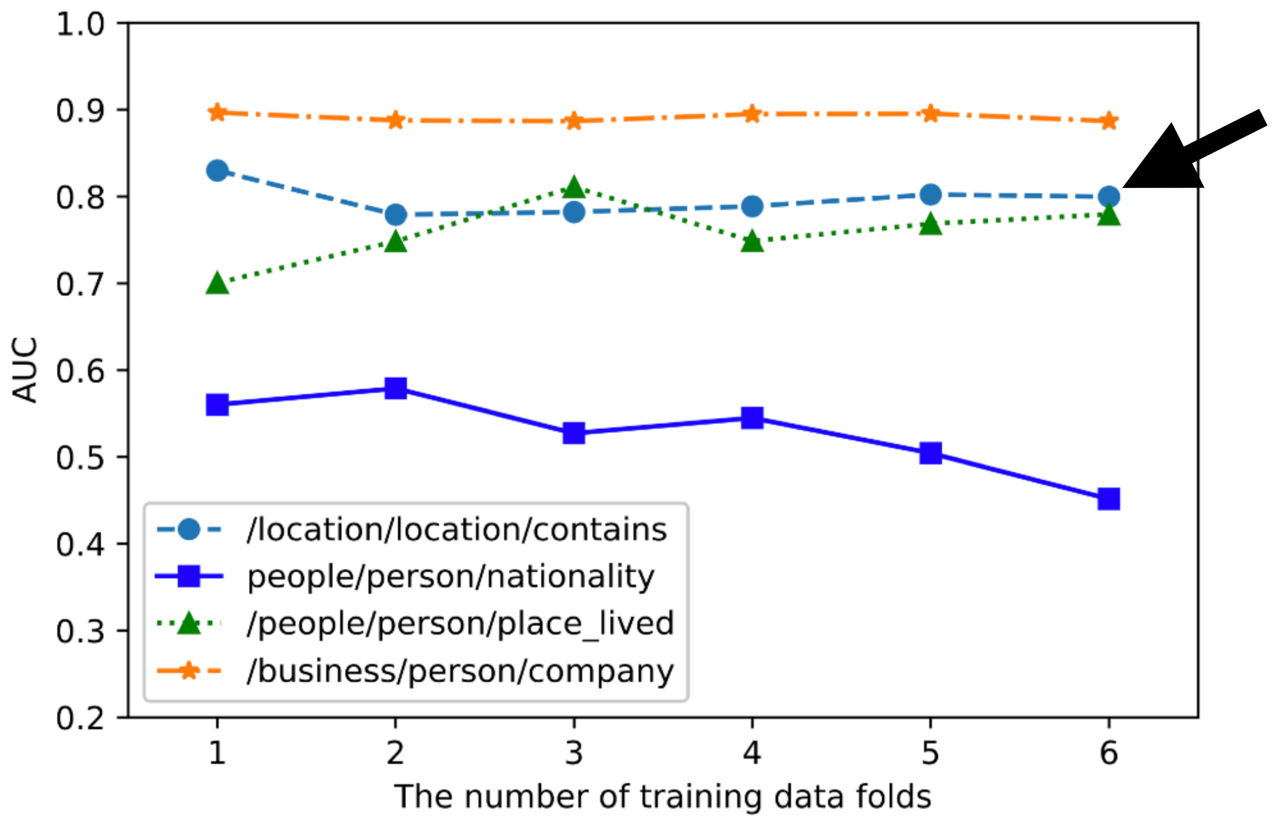
More data hurts performance?

PCNN+ATT



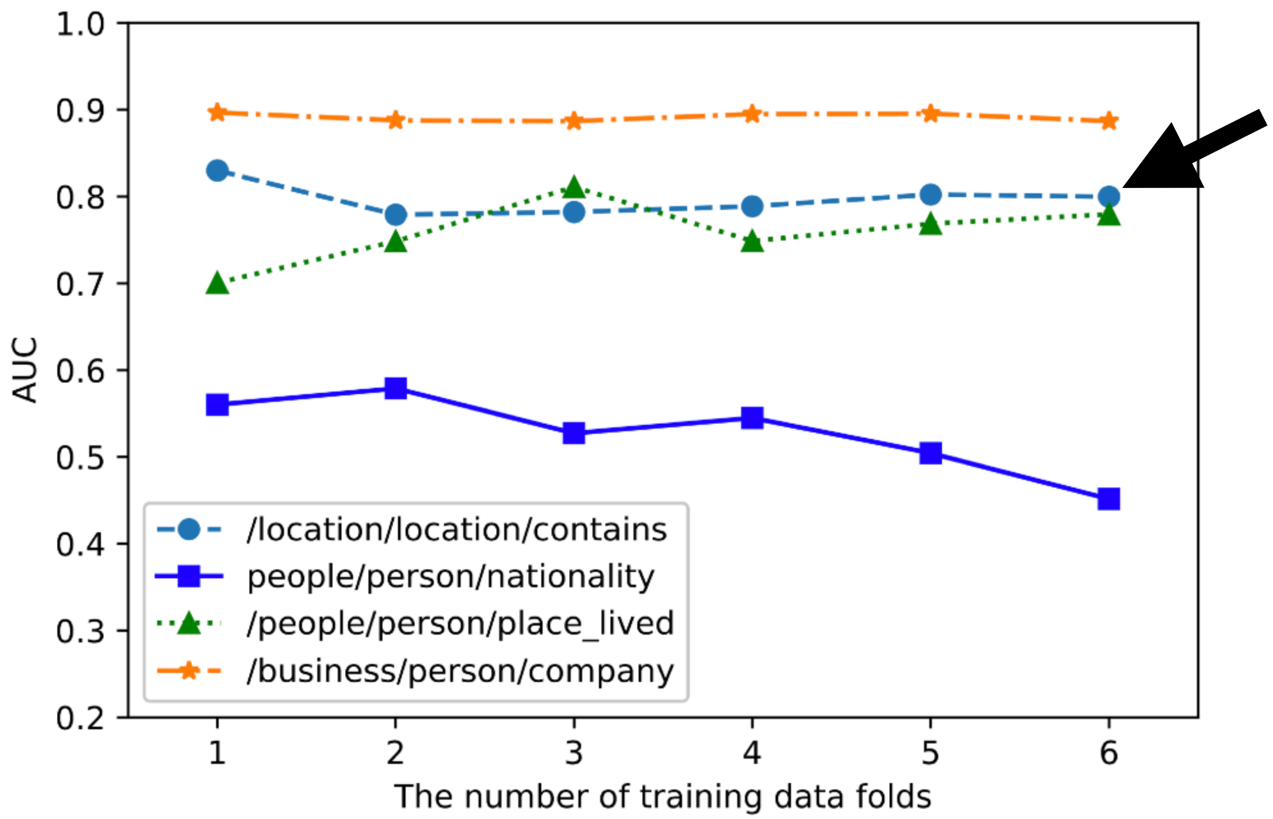
More data hurts performance?

PCNN+ATT

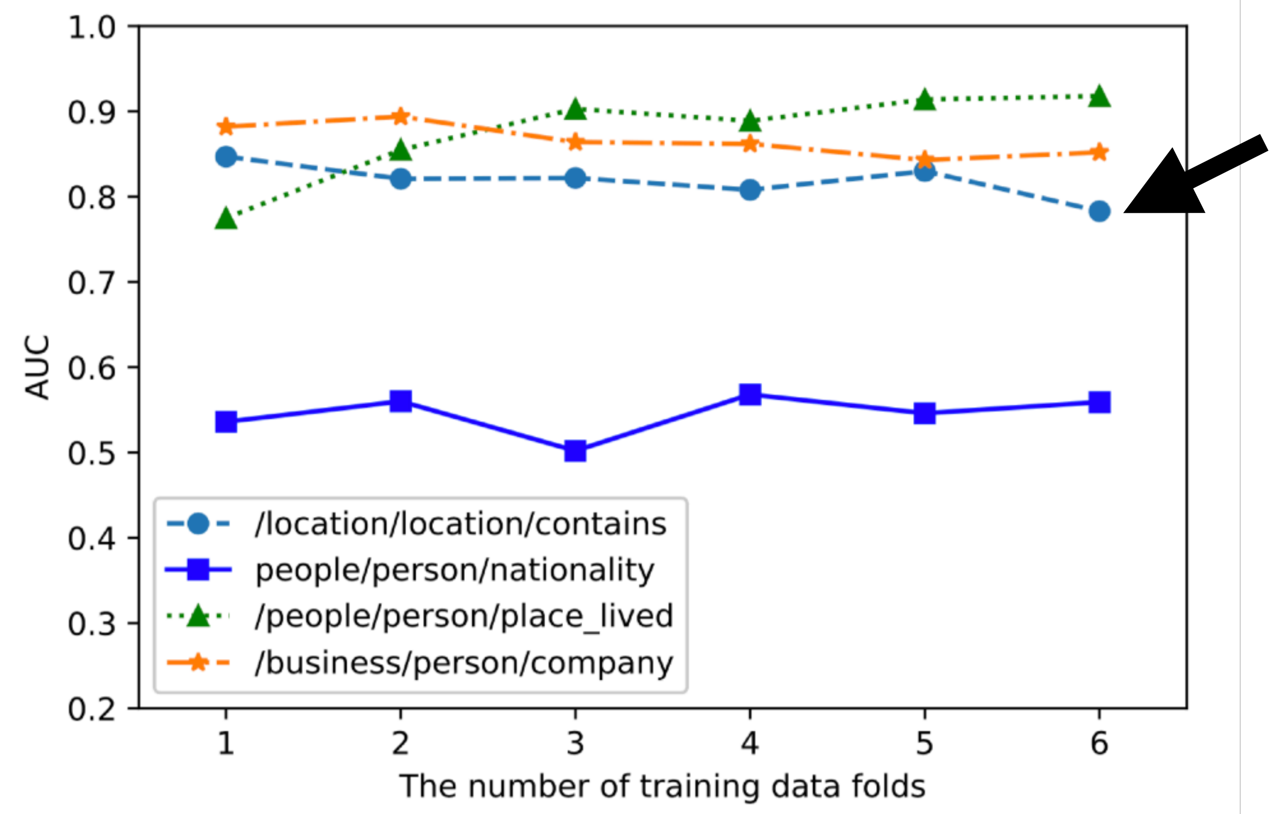


More data hurts performance?

PCNN+ATT

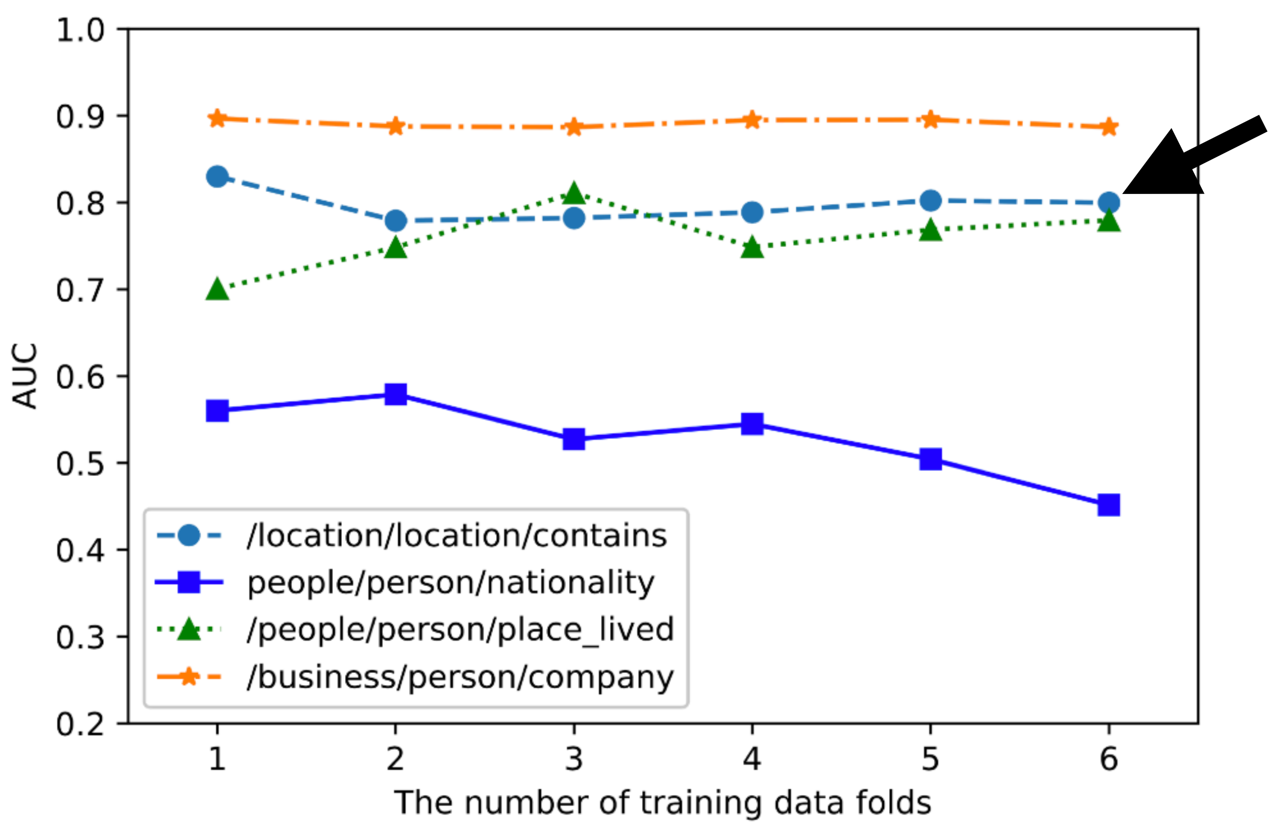


PCNN+Structured

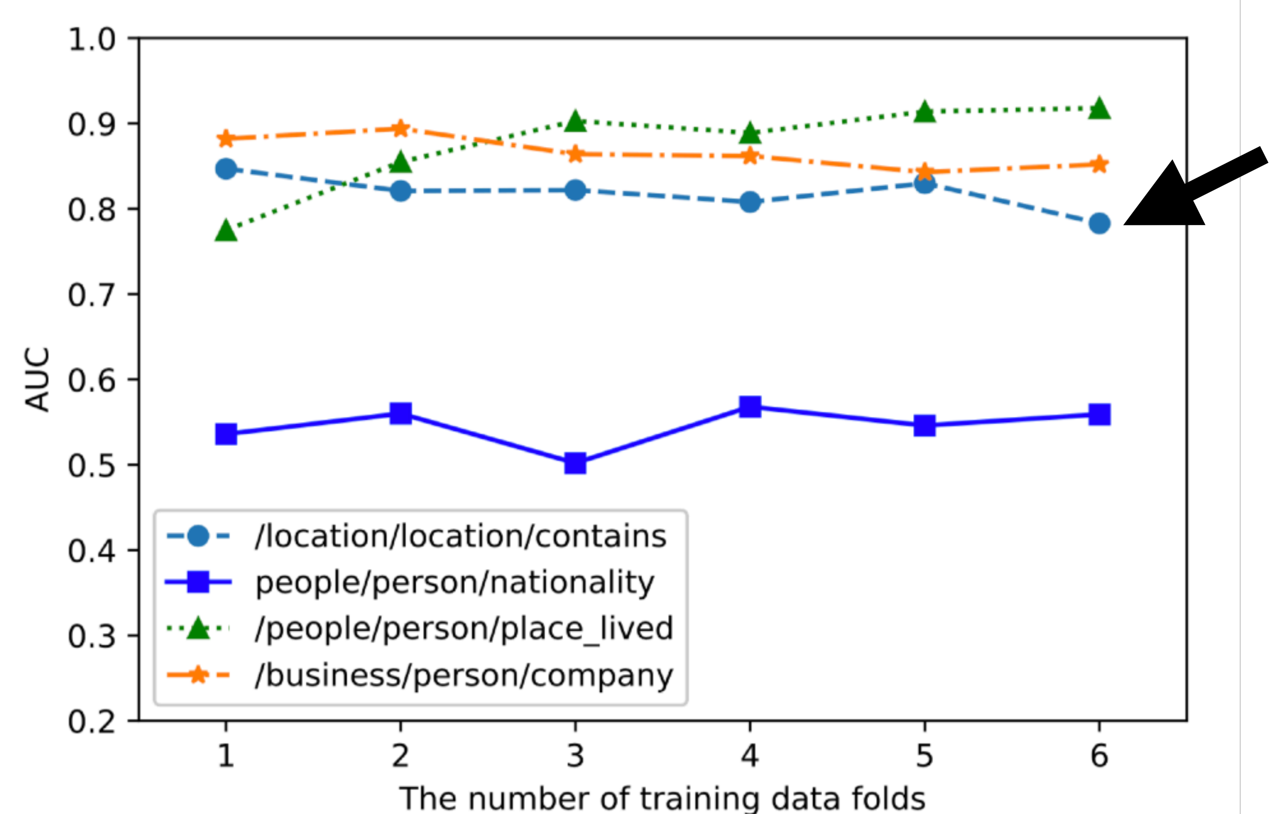


More data hurts performance?

PCNN+ATT



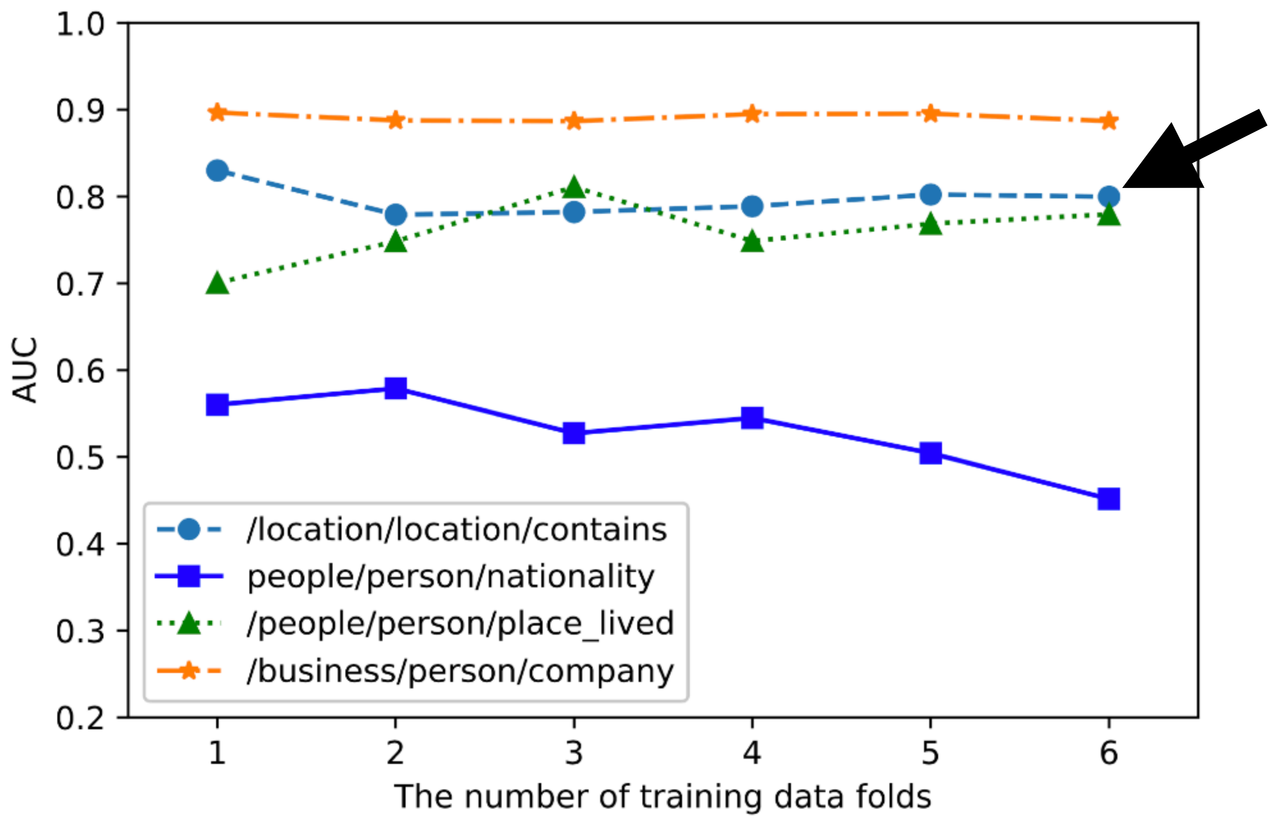
PCNN+Structured



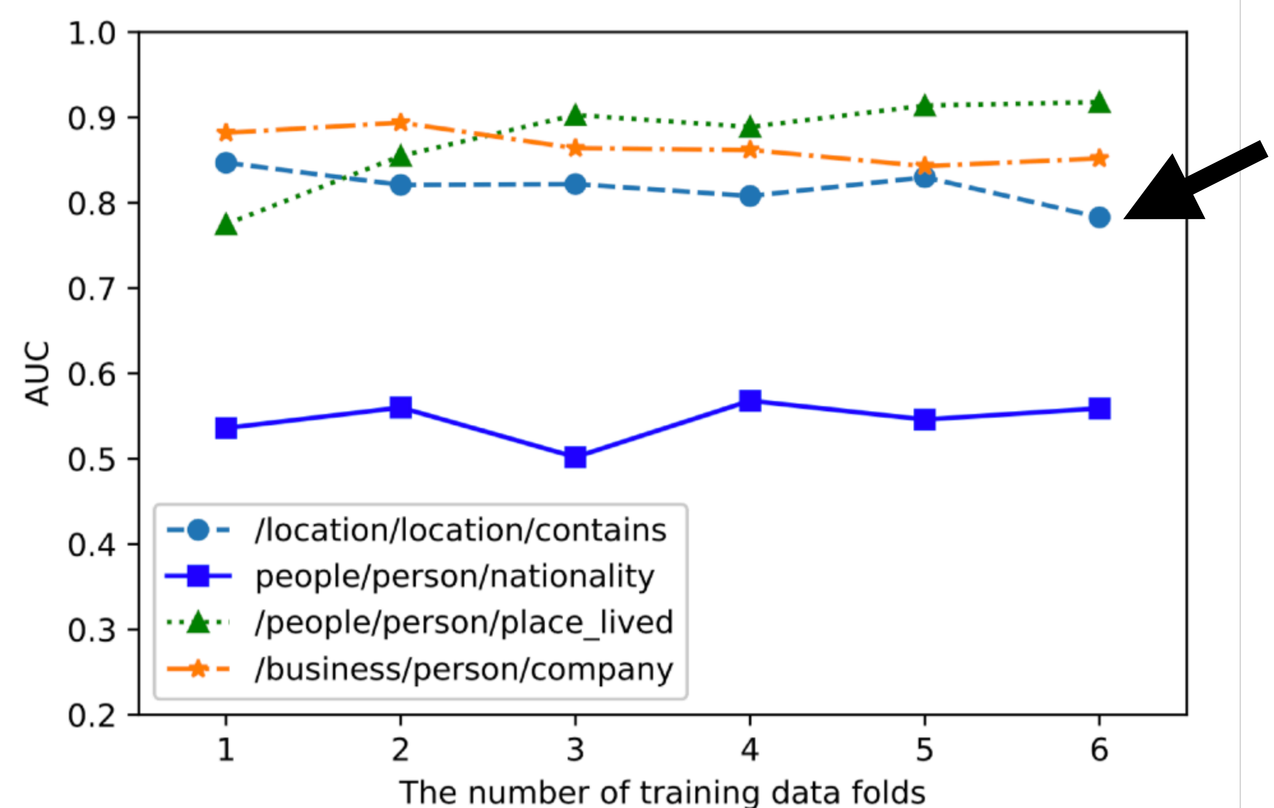
Hypothesis: better learning to predict what relations are likely to appear in Freebase

More data hurts performance?

PCNN+ATT



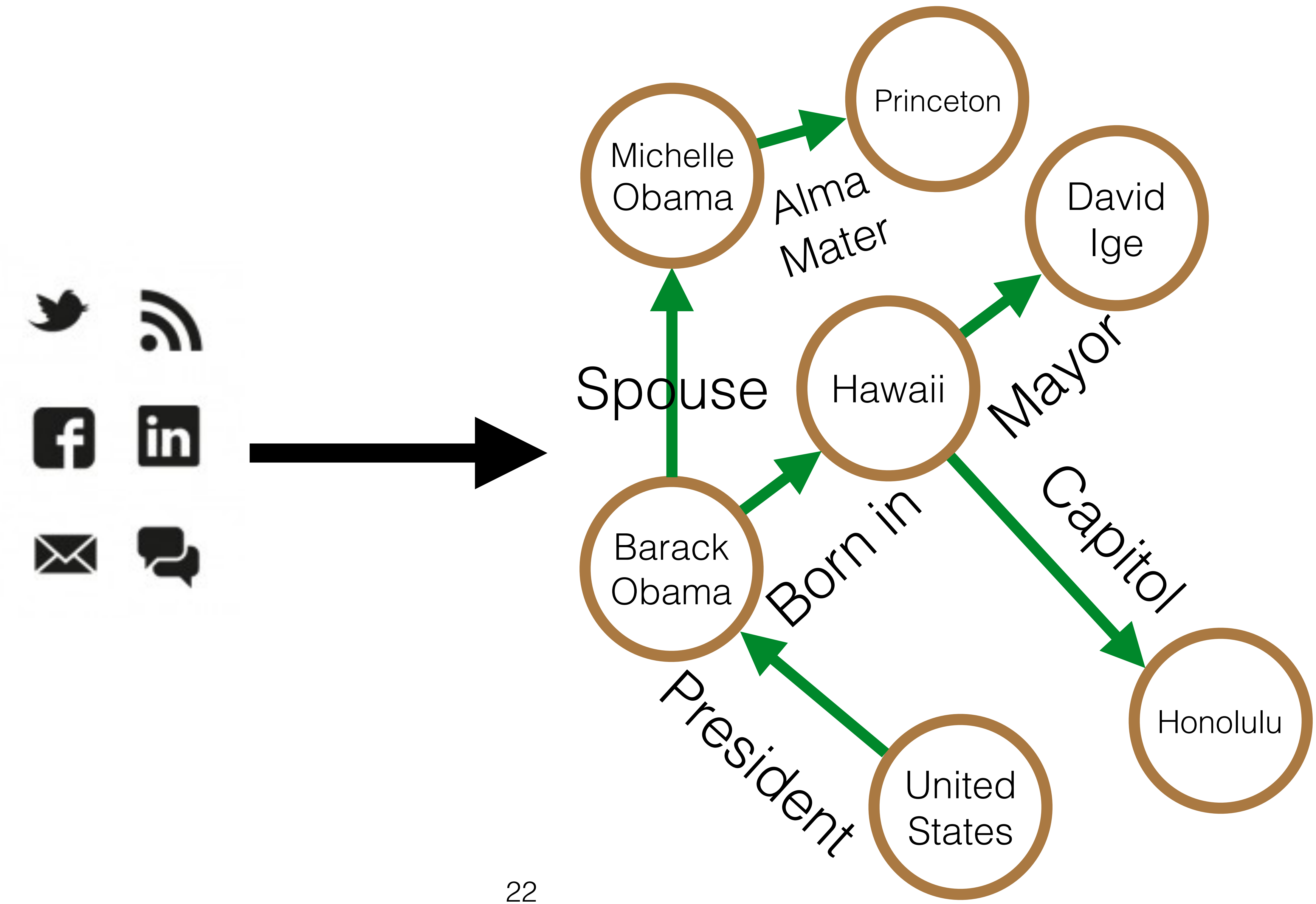
PCNN+Structured



Hypothesis: better learning to predict what relations are likely to appear in Freebase

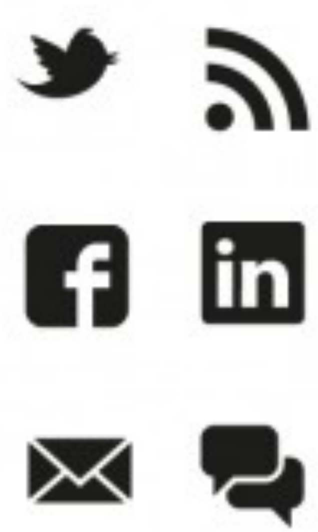
Structured + PCNN model is able to better take advantage of larger training datasets

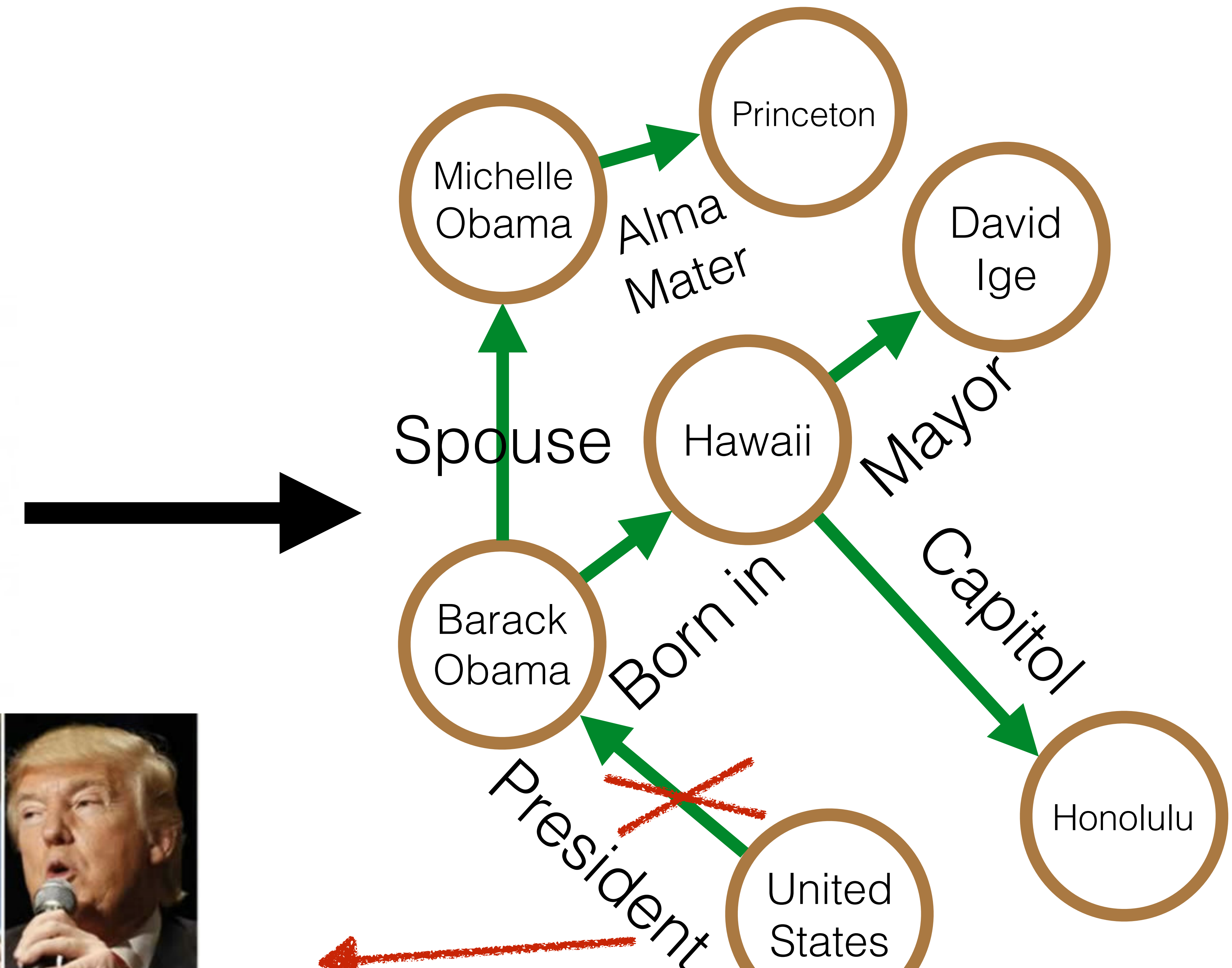
Extracting Knowledge Graphs



Extracting Knowledge Graphs

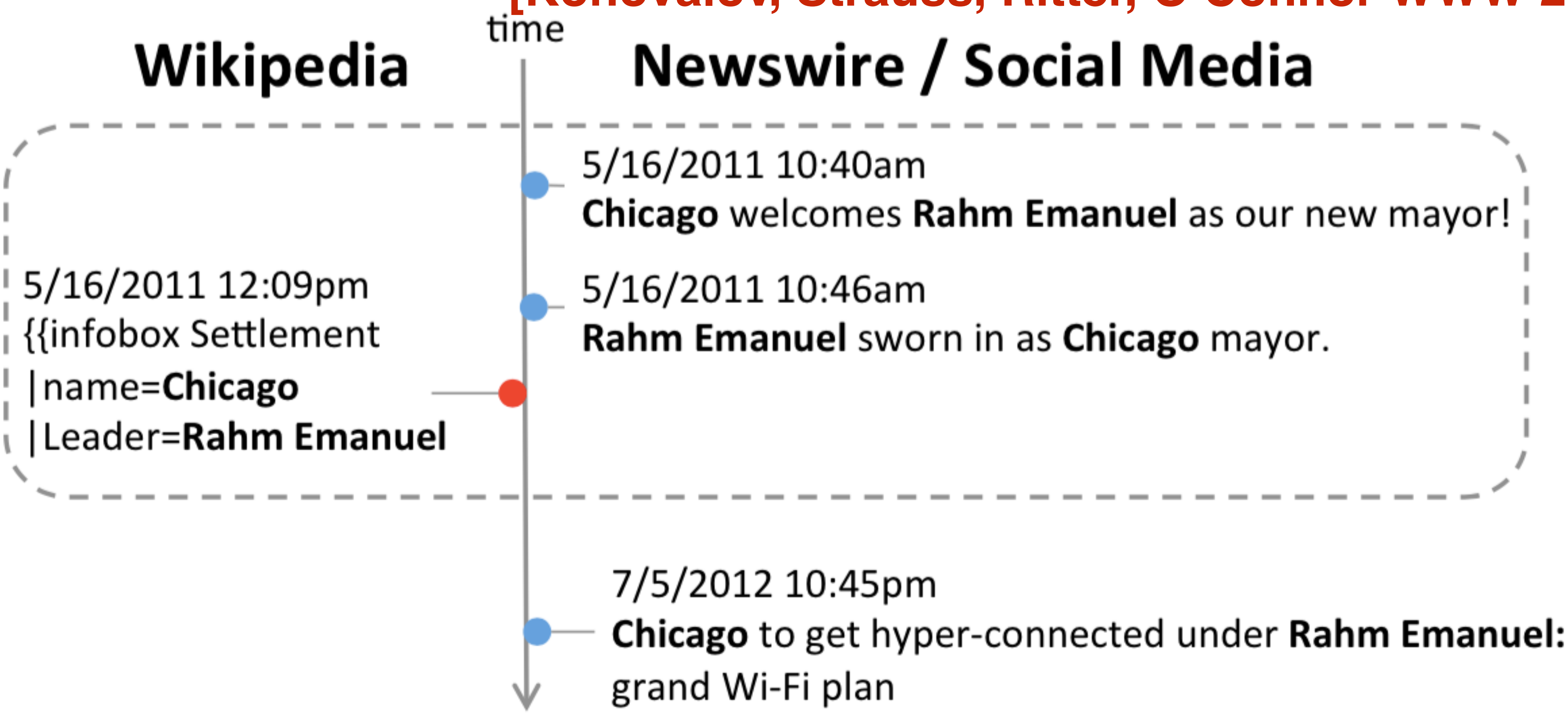
[Konovalov, Strauss, Ritter, O'Connor WWW 2017]





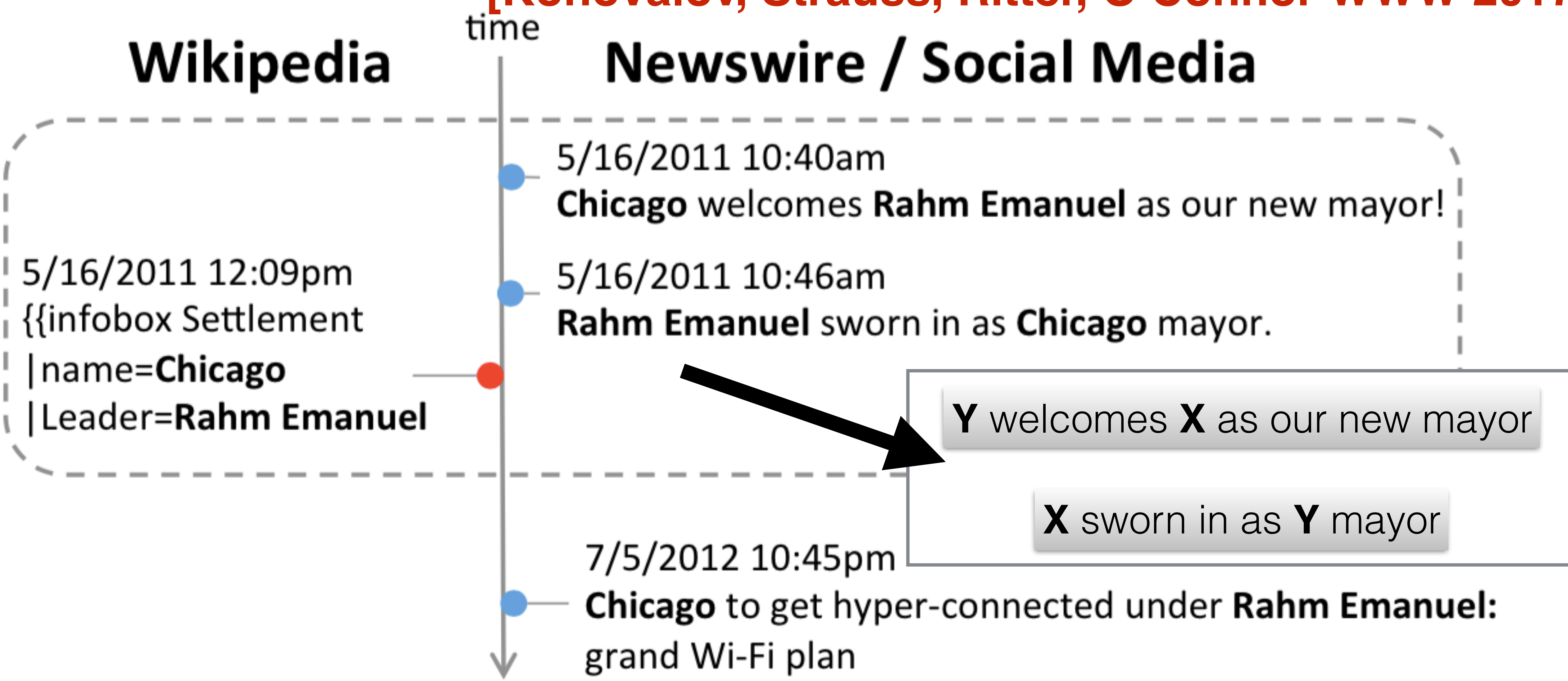
Learning to Extract Events from Wikipedia Edits

[Konovalov, Strauss, Ritter, O'Connor WWW 2017]



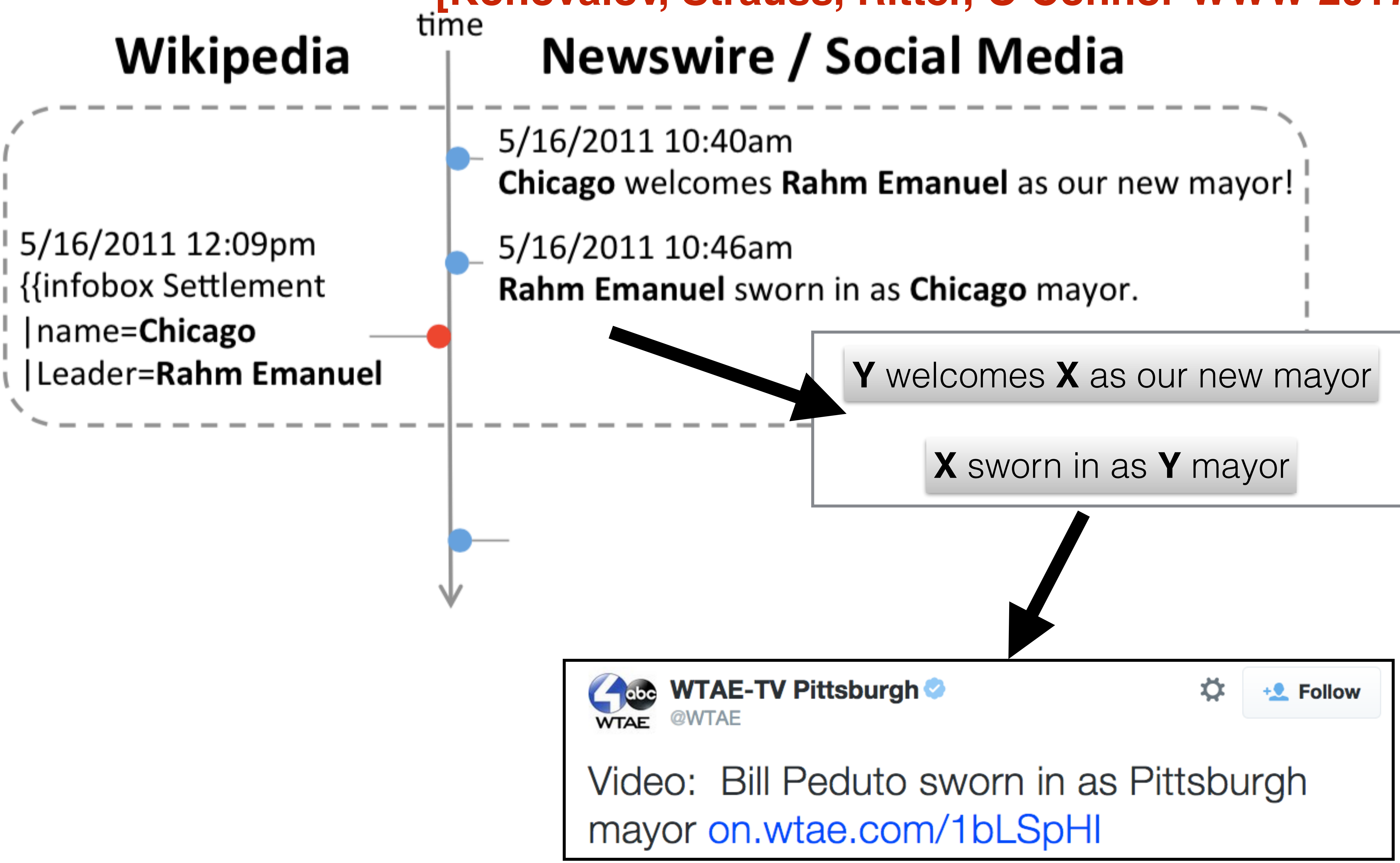
Learning to Extract Events from Wikipedia Edits

[Konovalov, Strauss, Ritter, O'Connor WWW 2017]



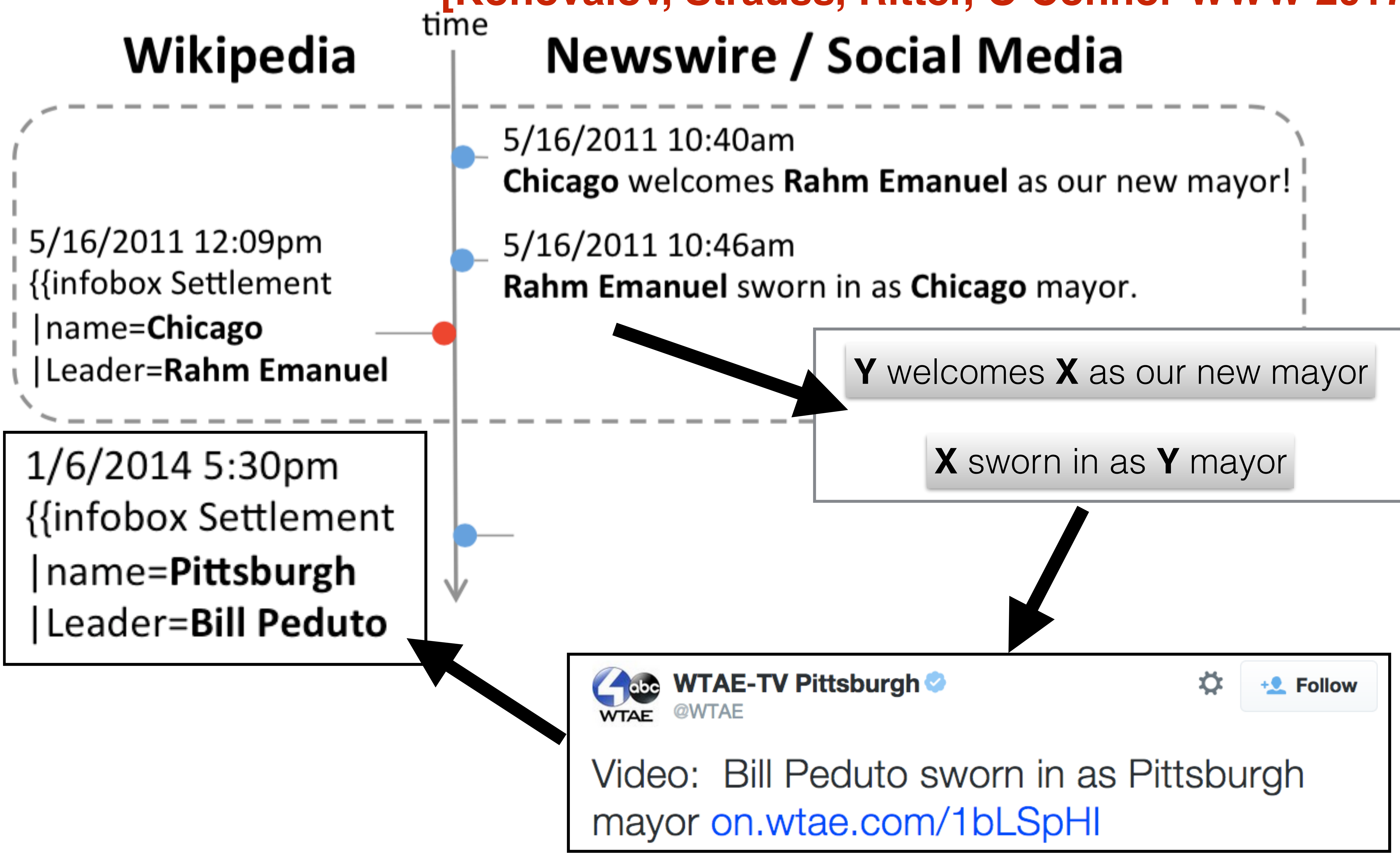
Learning to Extract Events from Wikipedia Edits

[Konovalov, Strauss, Ritter, O'Connor WWW 2017]



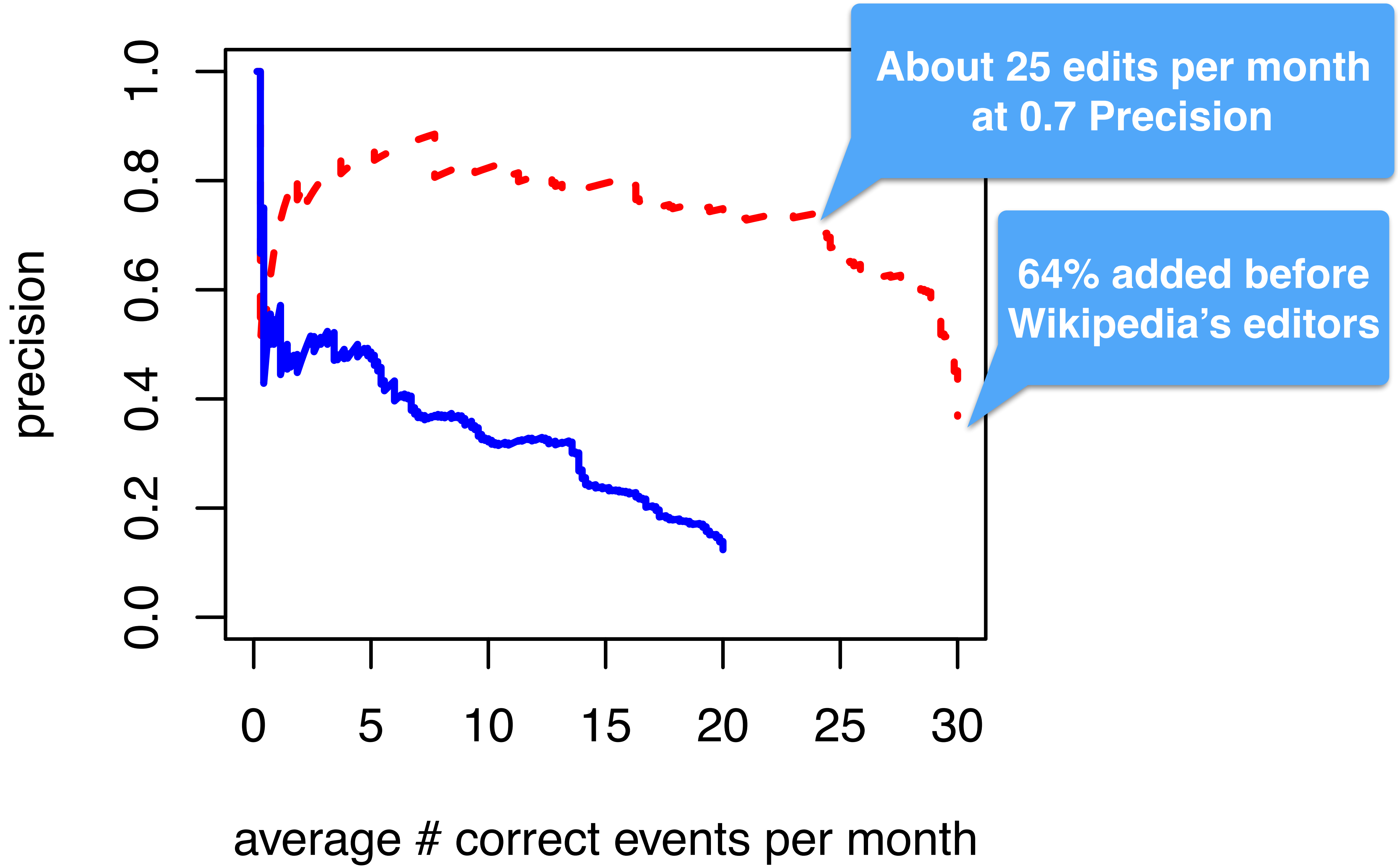
Learning to Extract Events from Wikipedia Edits

[Konovalov, Strauss, Ritter, O'Connor WWW 2017]



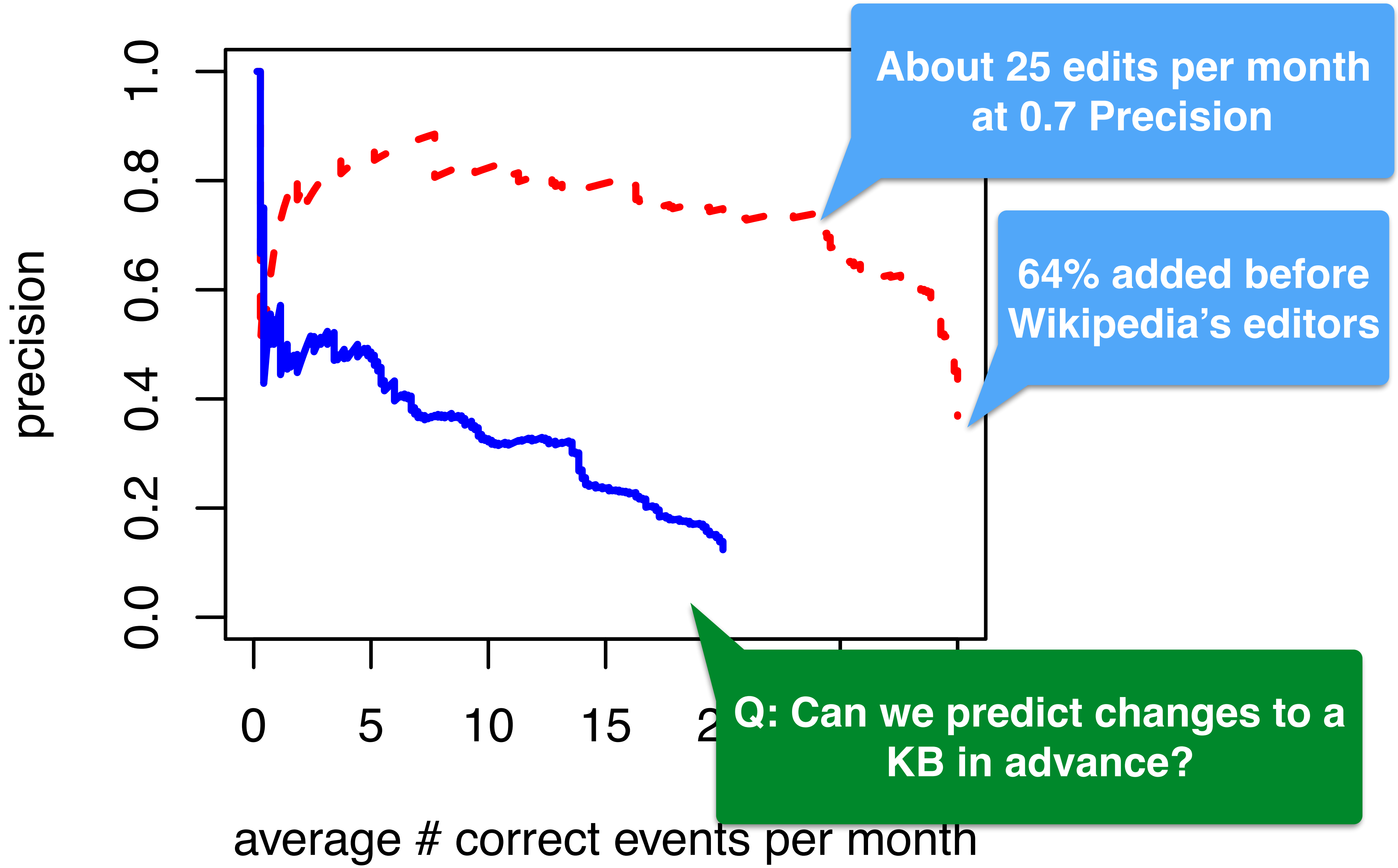
Extracting KB Edits from Text

[Konovalov, Strauss, Ritter, O'Connor WWW 2017]



Extracting KB Edits from Text

[Konovalov, Strauss, Ritter, O'Connor WWW 2017]



[Swamy, Ritter, de Marneffe EMNLP 2017]

Predicting Changes to a KB

[Swamy, Ritter, de Marneffe EMNLP 2017]

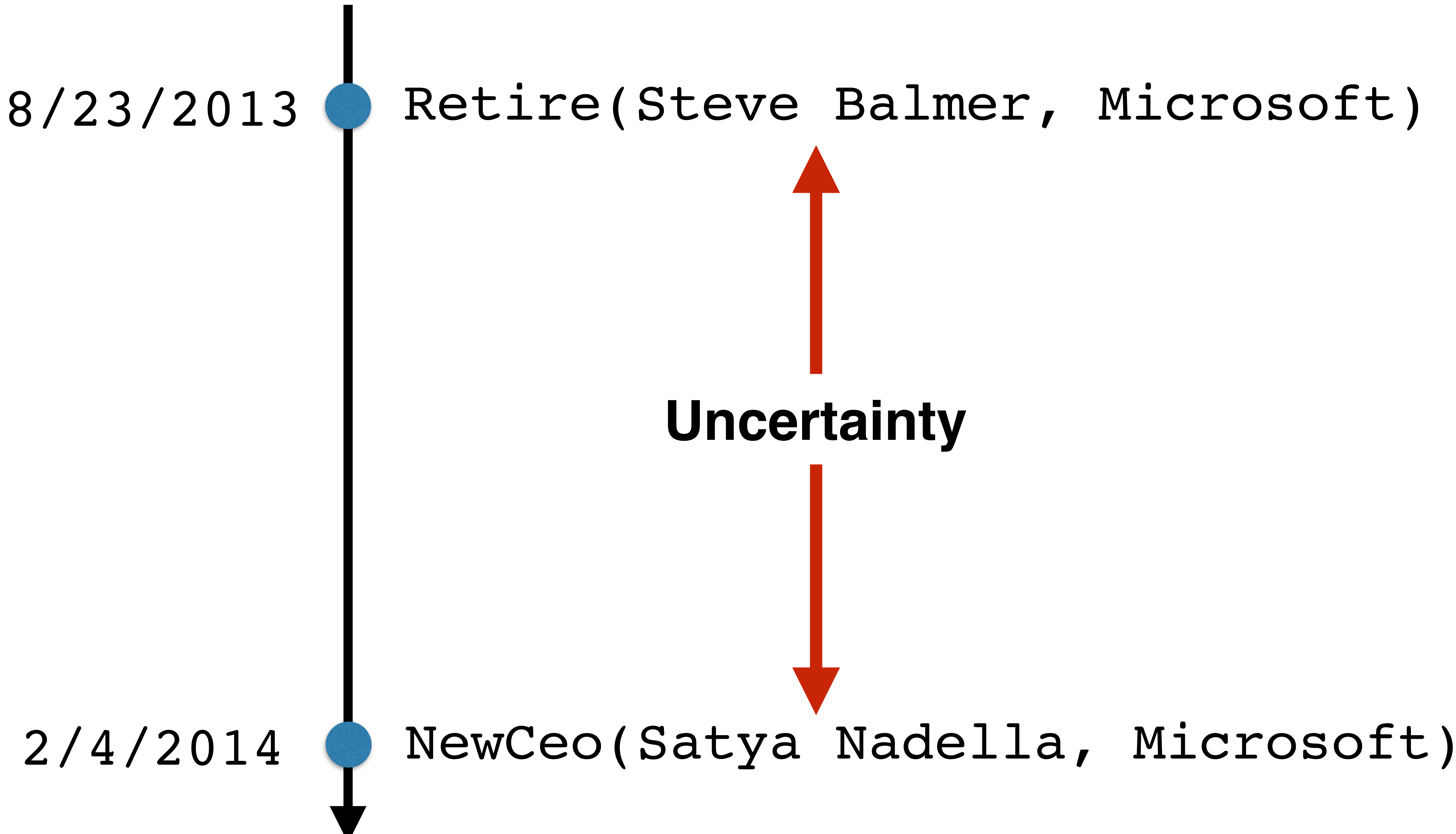
Predicting Changes to a KB

8/23/2013 Retire(Steve Balmer, Microsoft)

2/4/2014 NewCeo(Satya Nadella, Microsoft)

[Swamy, Ritter, de Marneffe EMNLP 2017]

Predicting Changes to a KB



[Swamy, Ritter, de Marneffe EMNLP 2017]

Predicting Changes to a KB

8/23/2013 Retire(Steve Balmer, Microsoft)

Follow

My bet is that Satya Nadella is named new $MSFT CEO on Ballmer's retirement.

9:32 AM - 23 Aug 2013

2/4/2014 NewCeo(Satya Nadella, Microsoft)

[Swamy, Ritter, de Marneffe EMNLP 2017]

Predicting Changes to a KB

8/23/2013 Retire(Steve Balmer, Microsoft)

Follow

My bet is that Satya Nadella is named new $MSFT CEO on Ballmer's retirement.

9:32 AM - 23 Aug 2013

Follow

I hope Ben Affleck is the new Microsoft CEO

7:59 PM - 23 Aug 2013

2/4/2014 NewCeo(Satya Nadella, Microsoft)

[Swamy, Ritter, de Marneffe EMNLP 2017]

Predicting Changes to a KB

8/23/2013 Retire(Steve Balmer, Microsoft)

Follow

My bet is that Satya Nadella is named new $MSFT CEO on Ballmer's retirement. ✓

9:32 AM - 23 Aug 2013

Follow

I hope Ben Affleck is the new Microsoft CEO ✗

7:59 PM - 23 Aug 2013

2/4/2014 NewCeo(Satya Nadella, Microsoft)

Data Collection (Example: Oscars)

WIKIPEDIA
The Free Encyclopedia

scrape

Leonardo DiCaprio
Bryan Cranston
Matt Daemon
Michael Fassbender
Eddie Redmayne

(Best actor, 2/28/2016)

Data Collection (Example: Oscars)

WIKIPEDIA
The Free Encyclopedia

scrape

Leonardo DiCaprio
Bryan Cranston
Matt Daemon
Michael Fassbender
Eddie Redmayne

(Best actor, 2/28/2016)

query

Oscars **Leonardo DiCaprio** win since:2016-2-27 until:2016-2-28

Oscars **Bryan Cranston** win since:2016-2-27 until:2016-2-28

⋮

Data Collection (Example: Oscars)

WIKIPEDIA
The Free Encyclopedia

scrape

Leonardo DiCaprio
Bryan Cranston
Matt Daemon
Michael Fassbender
Eddie Redmayne

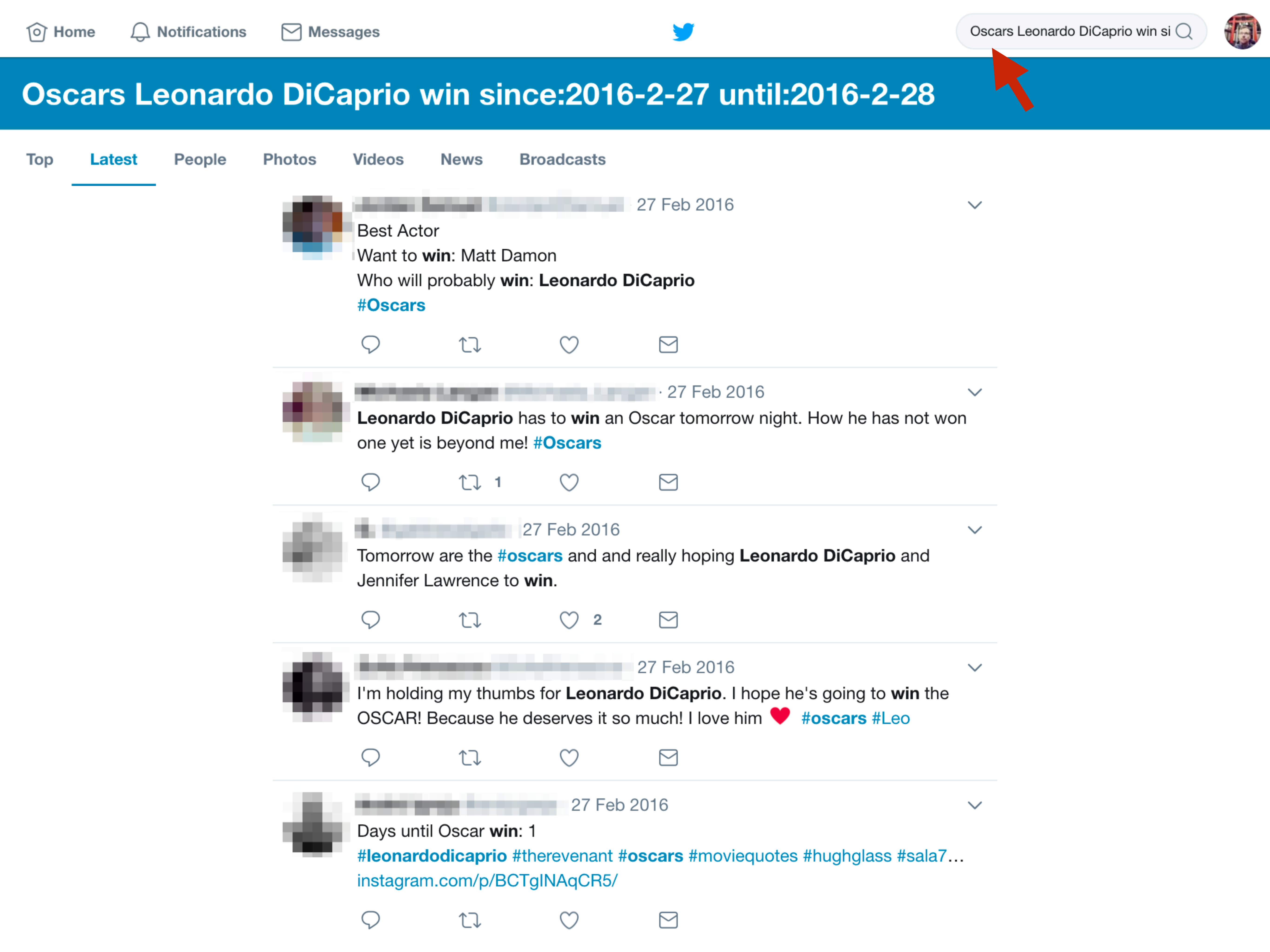
(Best actor, 2/28/2016)

query

Oscars **Leonardo DiCaprio** win since:2016-2-27 until:2016-2-28

Oscars **Bryan Cranston** win since:2016-2-27 until:2016-2-28

⋮


Home
Notifications
Messages
Oscars Leonardo DiCaprio win si
Oscars Leonardo DiCaprio win since:2016-2-27 until:2016-2-28
Top
Latest
People
Photos
Videos
News
Broadcasts
27 Feb 2016
Best Actor
Want to win: Matt Damon
Who will probably win: Leonardo DiCaprio
#Oscars
27 Feb 2016
Leonardo DiCaprio has to win an Oscar tomorrow night. How he has not won one yet is beyond me! #Oscars
1
27 Feb 2016
Tomorrow are the #oscars and and really hoping Leonardo DiCaprio and Jennifer Lawrence to win.
2
27 Feb 2016
I'm holding my thumbs for Leonardo DiCaprio. I hope he's going to win the OSCAR! Because he deserves it so much! I love him ❤️ #oscars #Leo
27 Feb 2016
Days until Oscar win: 1
#leonardodicaprio #therevenant #oscars #moviequotes #hughglass #sala7…
instagram.com/p/BCTgINAqCR5/

Data Collection (Summary)

Contest	#Events	#Tweets
2016 US Presidential Primaries	483	38,220
Tennis Grand Slams	52	30,530
2016 US Presidential Elections	76	13,645
Football World Cup (2010-2016)	12	13,618
Balon d'Or Award (2010-2016)	18	6,777
Cricket World Cup (2010-2016)	18	4,120
Oscars (2009-2016)	176	2,370
Eurovision (2010-2016)	162	1,682
2014 Indian General Elections	68	1,656
Rugby World Cup (2010-2016)	6	651
Total	**1071**	**113269**

Data Collection (Summary)

Contest	#Events	#Tweets
2016 US Presidential Primaries	483	38,220
Tennis Grand Slams	52	30,530
2016 US Presidential Elections	76	13,645
Football World Cup (2010-2016)	12	13,618
Balon d'Or Award (2010-2016)	18	6,777
Cricket World Cup (2010-2016)	18	4,120
Oscars (2009-2016)	176	2,370
Eurovision (2010-2016)	162	1,682
2014 Indian General Elections	68	1,656
Rugby World Cup (2010-2016)	6	651
Total	**1071**	**113269**

- **Not all are Positive Veridicality Predictions**
- **We Need a Classifier to Measure Veridicality**

Data Collection (Summary)

Contest	#Events	#Tweets
2016 US Presidential Primaries	483	38,220
Tennis Grand Slams	52	30,530
2016 US Presidential Elections	76	13,645
Football World Cup (2010-2016)	12	13,618
Balon d'Or Award (2010-2016)	18	6,777
Cricket World Cup (2010-2016)	18	4,120
Oscars (2009-2016)	176	2,370
Eurovision (2010-2016)	162	1,682
2014 Indian General Elections	68	1,656
Rugby World Cup (2010-2016)	6	651
Total	**1071**	**113269**

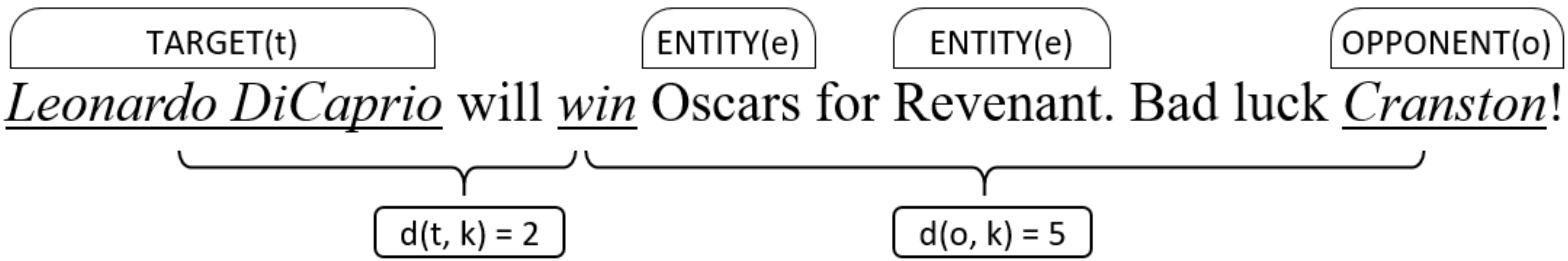
MTurk Annotation: 3,543 Tweets

Features

$$\log P(y = v | c, O, \text{tweet}) \propto \theta_v \cdot f(c, O, \text{tweet})$$

Features

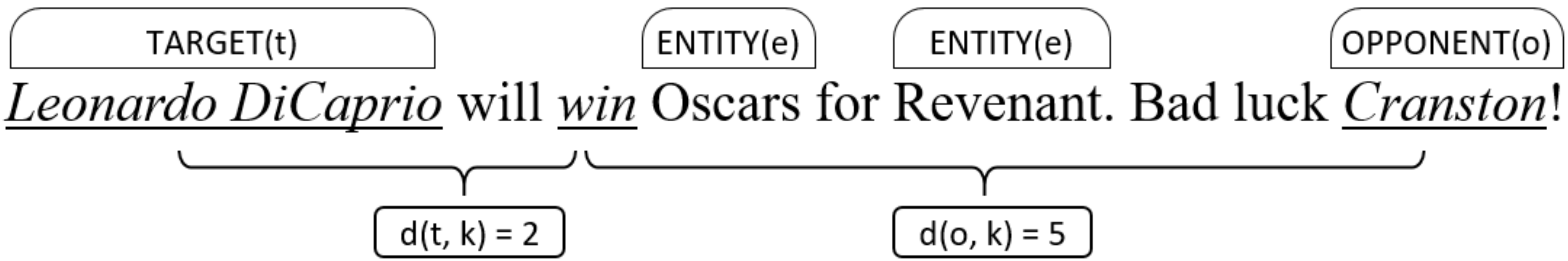
$$\log P(y = v | c, O, \text{tweet}) \propto \theta_v \cdot f(c, O, \text{tweet})$$



Positive Veridicality Feature Type	Feature	Weight
Keyword context	TARGET *will* KEYWORD	0.41
Keyword dep. path	TARGET $\rightarrow$ *to* $\rightarrow$ KEYWORD	0.38
Keyword dep. path	TARGET $\leftarrow$ *is* $\rightarrow$ *going* $\rightarrow$ *to* $\rightarrow$ KEYWORD	0.29
Target context	TARGET *is favored to win*	0.19
Keyword context	TARGET *are going to* KEYWORD	0.15
Target context	TARGET *predicted to win*	0.13
Pair context	TARGET1 *could win* TARGET2	0.13
Distance to keyword	TARGET *closer to* KEYWORD	0.11

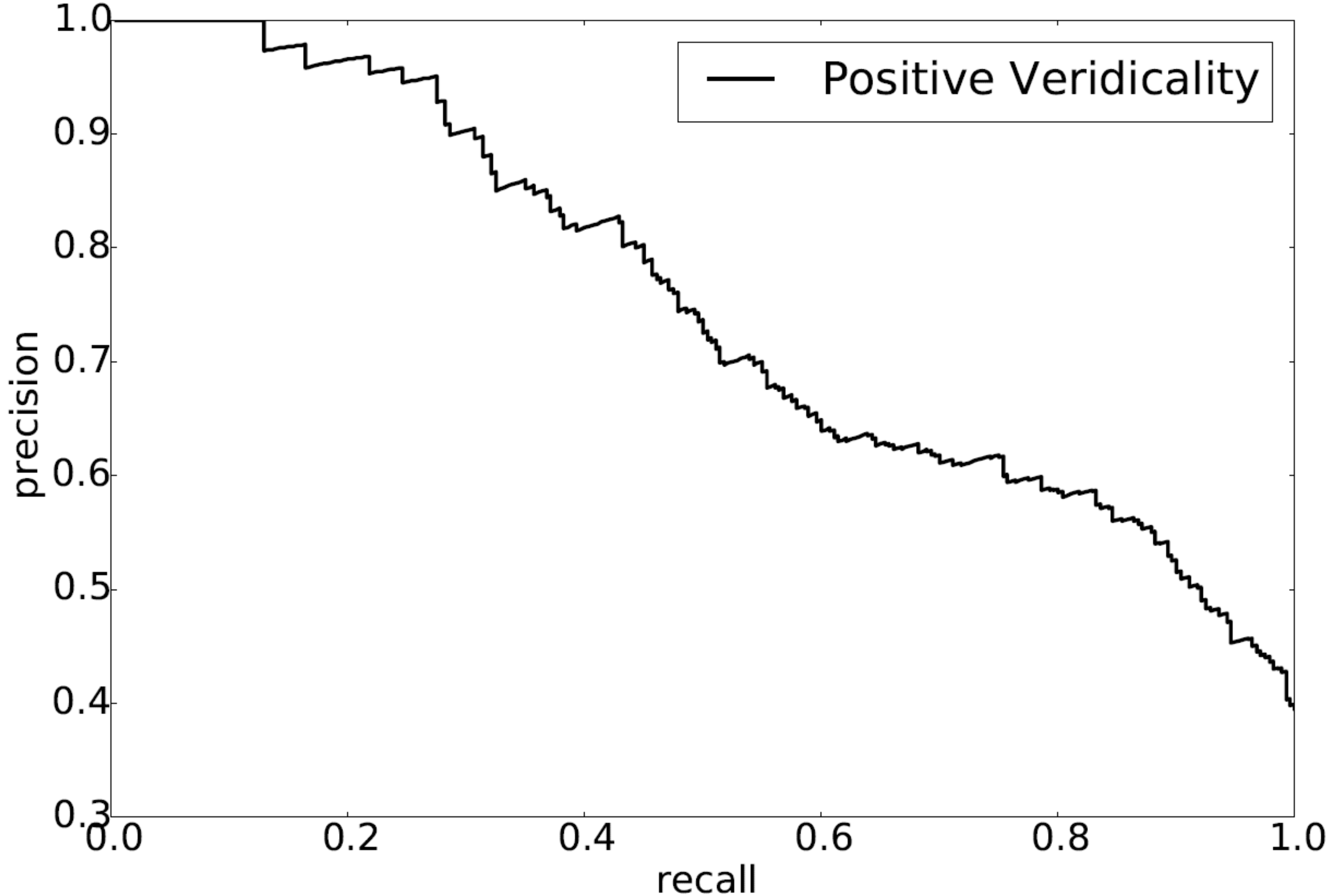
Features

$$\log P(y = v | c, O, \text{tweet}) \propto \theta_v \cdot f(c, O, \text{tweet})$$



Negative Veridicality Feature Type	Feature	Weight
Negated keyword	keyword is negated	0.47
Keyword context	TARGET *won't* KEYWORD	0.41
Opponent context	OPPONENT *will win*	0.37
Keyword dep. path	TARGET $\leftarrow$ *will* $\rightarrow$ *not* $\rightarrow$ KEYWORD	0.31
Distance to keyword	OPPONENT *closer to* KEYWORD	0.28
Target context	TARGET *may not win*	0.27
Keyword dep. path	OPPONENT $\leftarrow$ *will* $\rightarrow$ KEYWORD	0.23
Target context	TARGET *can't win*	0.18

Precision / Recall



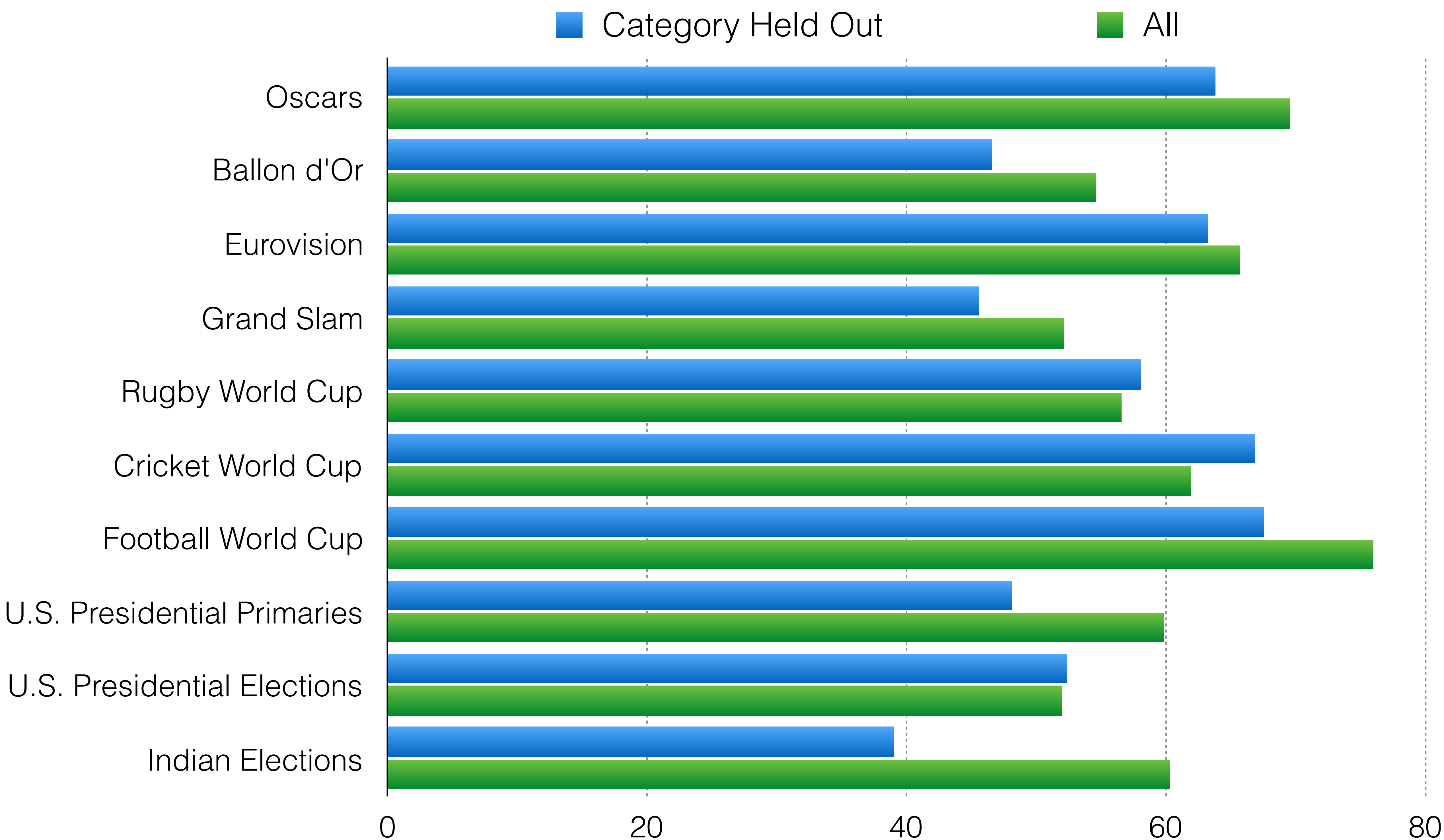
Error Analysis

Tweet	Gold	Predicted
The heart wants **Nadal** to win tomorrow but the mind points to a Djokovic win over 4 sets. Djokovic 7-5 4-6 7-5 6-4 **Nadal** for me.	?	

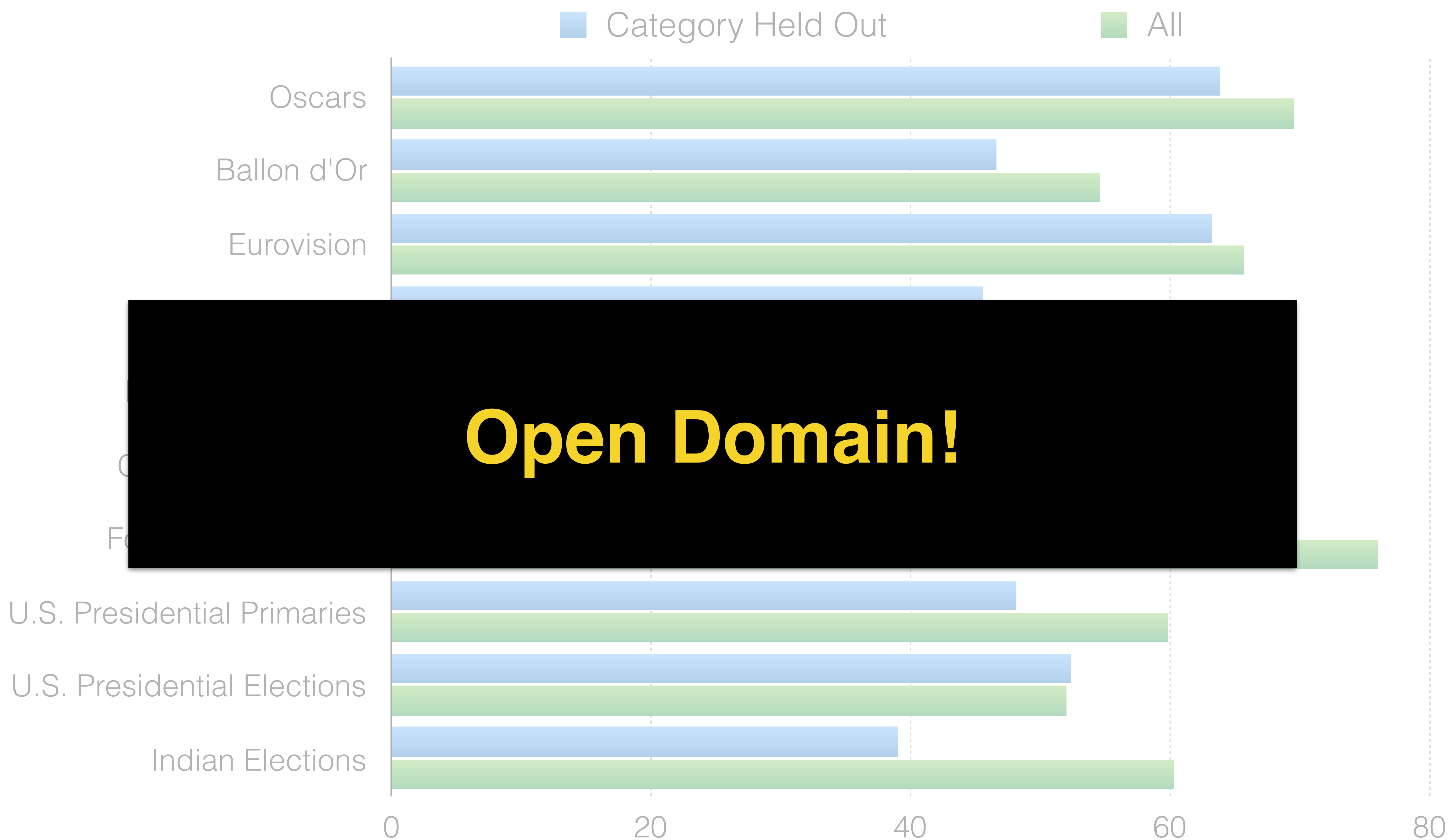
Error Analysis

Tweet	Gold	Predicted
The heart wants **Nadal** to win tomorrow but the mind points to a Djokovic win over 4 sets. Djokovic 7-5 4-6 7-5 6-4 **Nadal** for me.	negative	positive

Cross-Domain Experiments

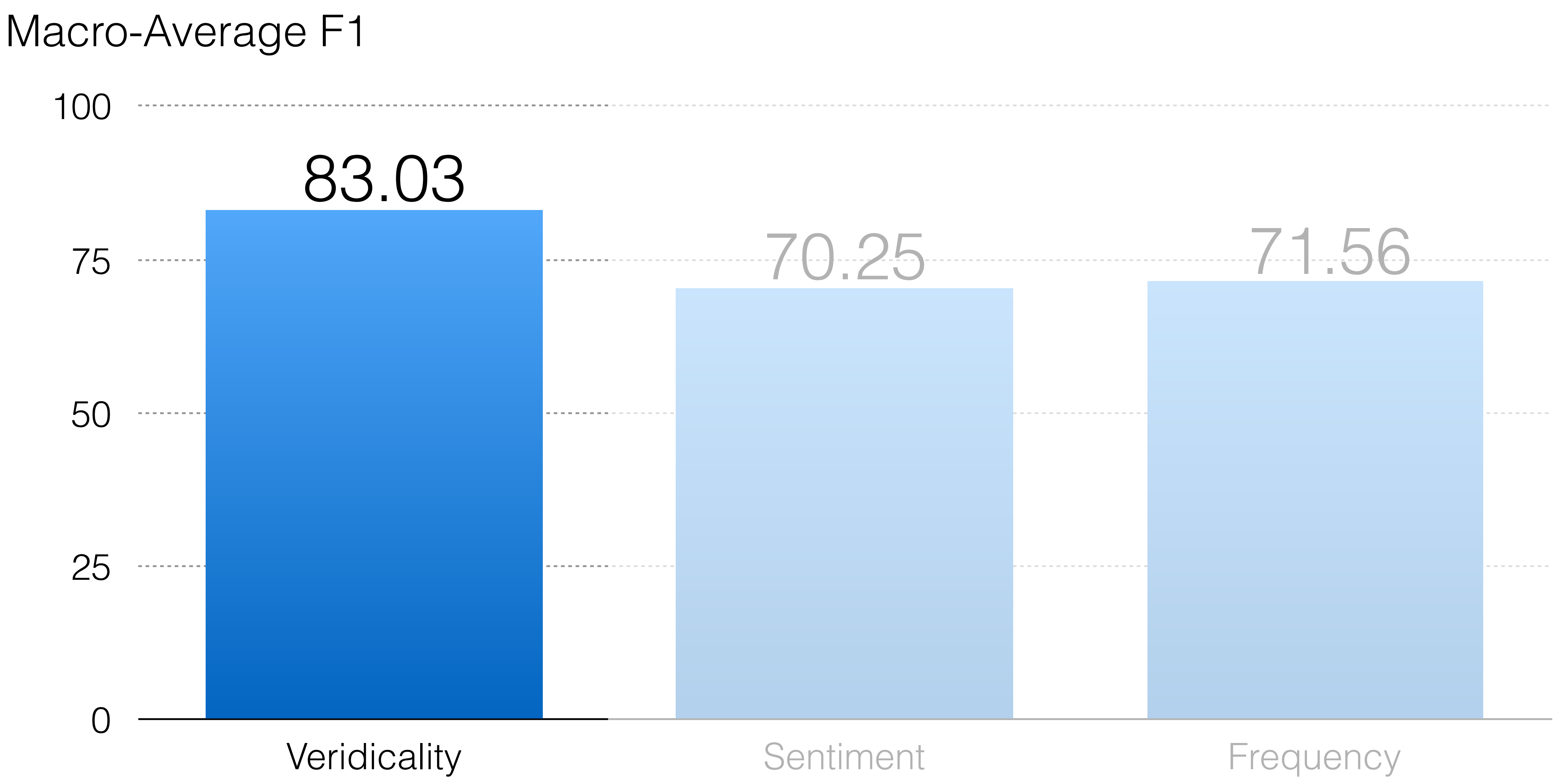


Cross-Domain Experiments

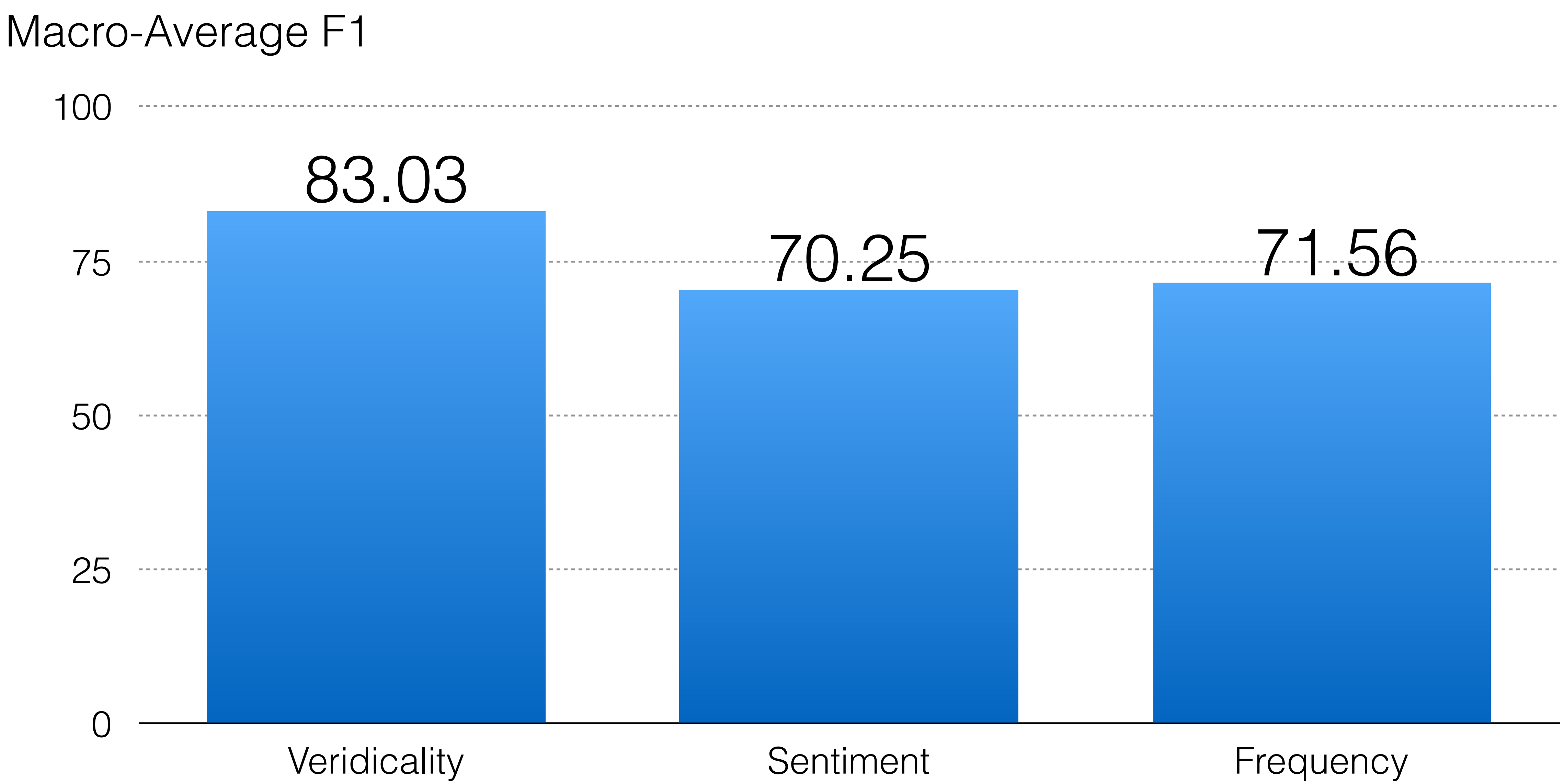


Open Domain!

Forecasting Performance



Baselines (Sentiment + Volume)



Surprise Outcomes

		Contender		Veridicality Score
OSCARS		**Leonardo DiCaprio**		0.97
		Natalie Portman		0.92
		Julianne Moore		0.91
		Daniel Day-Lewis		0.90
		Slumdog Millionaire		0.75
		Matthew McConaughey		0.74
	!	The Revenant		0.73
		Argo		0.71
		Brie Larson		0.70
		The Artist		0.67
PRIMARIES		**Trump**	South Carolina	0.96
		Clinton	Iowa	0.90
		Trump	Massachusetts	0.88
		Trump	Tennessee	0.88
		Sanders	Maine	0.87
		Sanders	Alaska	0.87
	!	Trump	Maine	0.87
		Sanders	Wyoming	0.86
		Trump	Louisiana	0.86
	!	Clinton	Indiana	0.85

Surprise Outcomes

		Contender		Veridicality Score
OSCARS		**Leonardo DiCaprio**		0.97
		Natalie Portman		0.92
		Julianne Moore		0.91
		Daniel Day-Lewis		0.90
		Slumdog Millionaire		0.75
		Matthew McConaughey		0.74
	!	The Revenant		0.73
		Argo		0.71
		Brie Larson		0.70
		The Artist		0.67
PRIMARIES		**Trump**	South Carolina	0.96
		Clinton	Iowa	0.90
		Trump	Massachusetts	0.88
		Trump	Tennessee	0.88
		Sanders	Maine	0.87
		Sanders	Alaska	0.87
	!	Trump	Maine	0.87
		Sanders	Wyoming	0.86
		Trump	Louisiana	0.86
	!	Clinton	Indiana	0.85


The Washington Post
Arts and Entertainment
How 'Spotlight's' surprise Best Picture win made Oscar history
By Amy Argetsinger and Geoff Edgers February 29, 2016

Surprise Outcomes

		Contender		Veridicality Score
OSCARS		**Leonardo DiCaprio**		0.97
		Natalie Portman		0.92
		Julianne Moore		0.91
		Daniel Day-Lewis		0.90
		Slumdog Millionaire		0.75
		Matthew McConaughey		0.74
	!	The Revenant		0.73
		Argo		0.71
		Brie Larson		0.70
		The Artist		0.67
PRIMARIES		**Trump**	South Carolina	0.96
		Clinton	Iowa	0.90
		Trump	Massachusetts	0.88
		Trump	Tennessee	0.88
		Sanders	Maine	0.87
		Sanders	Alaska	0.87
	!	Trump	Maine	0.87
		Sanders	Wyoming	0.86
		Trump	Louisiana	0.86
	!	Clinton	Indiana	0.85


The Washington Post
Arts and Entertainment
How 'Spotlight's' surprise Best Picture win made Oscar history
By Amy Argetsinger and Geoff Edgers February 29, 2016


HUFFPOST
Ted Cruz Upsets Donald Trump In Maine Republican Caucus
Cruz won despite Gov. Paul LePage's endorsement of Trump.
By Daniel Marans

Surprise Outcomes

		Contender		Veridicality Score
OSCARS		**Leonardo DiCaprio**		0.97
		Natalie Portman		0.92
		Julianne Moore		0.91
		Daniel Day-Lewis		0.90
		Slumdog Millionaire		0.75
		Matthew McConaughey		0.74
	!	The Revenant		0.73
		Argo		0.71
		Brie Larson		0.70
		The Artist		0.67
PRIMARIES		**Trump**	South Carolina	0.96
		Clinton	Iowa	0.90
		Trump	Massachusetts	0.88
		Trump	Tennessee	0.88
		Sanders	Maine	0.87
		Sanders	Alaska	0.87
	!	Trump	Maine	0.87
		Sanders	Wyoming	0.86
		Trump	Louisiana	0.86
	!	Clinton	Indiana	0.85

The Washington Post

Arts and Entertainment

How 'Spotlight's' surprise Best Picture win made Oscar history

By Amy Argetsinger and Geoff Edgers February 29, 2016

HUFFPOST

Ted Cruz Upsets Donald Trump In Maine Republican Caucus

Cruz won despite Gov. Paul LePage's endorsement of Trump.

By Daniel Marans

POLITICO

Sanders surprises Clinton in Indiana

Thank you!