Automatic Fact Verification with Semi-Structured Knowledge

William Wang
UC **SANTA BARBARA**

Joint work w. Wenhu Chen, Jianshu Chen, Xin Wang etc.

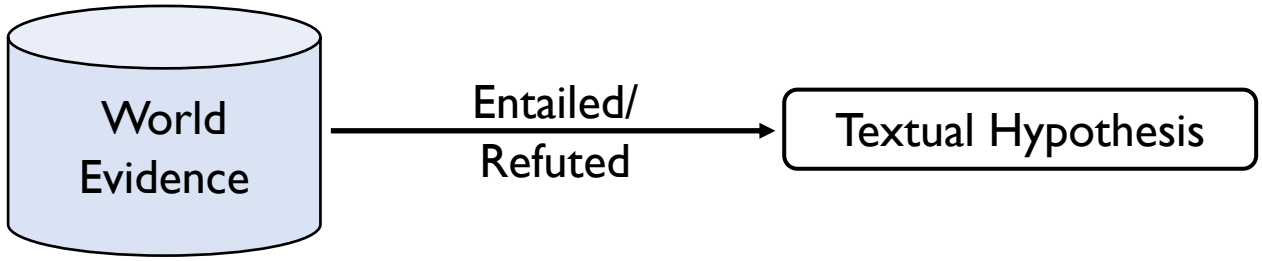
Outline

- Background
- Dataset
- Algorithm
- Conclusion
- Overview of UCSB NLP Group

Natural Language Understanding

- Fact Verification is a major problem in natural language understanding.

- It can benefit many downstream tasks like misinformation detection, fake news detection, etc.

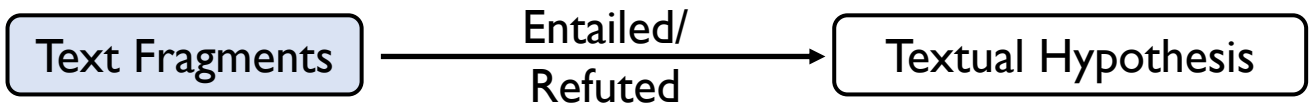


UCSB

Existing Fact Verification

- Recognizing Textual Entailment (Dagan et al. 2006)
  - Recognizing whether meaning of first sentence is contained in the second sentence.

- Natural Language Inference (Bowman et al. 2015)
  - Recognizing whether a premise text can entail another hypothesis text.



UCSB

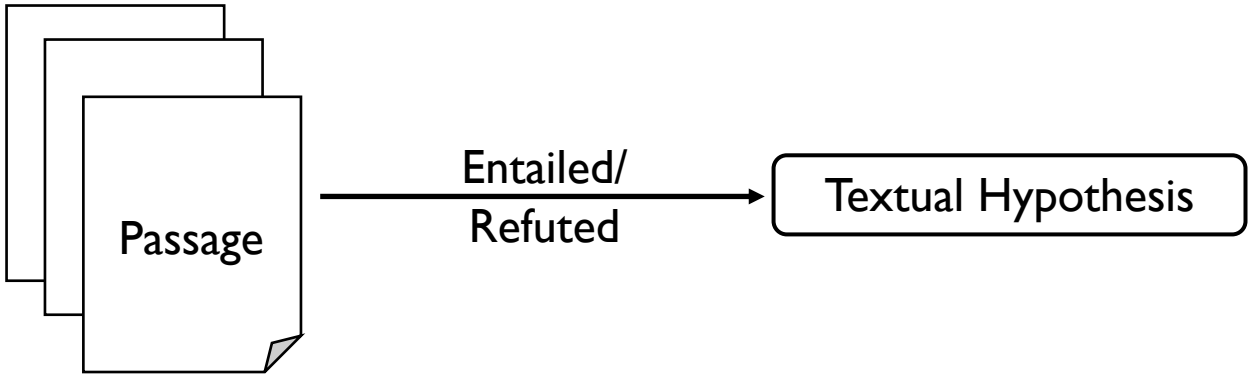
Automated Fake News Detection

- LIAR Benchmark Dataset (Wang, ACL 2017)
  - Short claims from PolitiFact.com.
  - Variety of sources.
  - Diverse selection of speakers.
  - Text only.
  - No fact-checking.
  - Relatively low accuracy in six-way classification.

UCSB

Existing Fact Verification

- Fact Checking (Thorne et al. 2018)
  - Verify claims modified from Wikipedia



UCSB

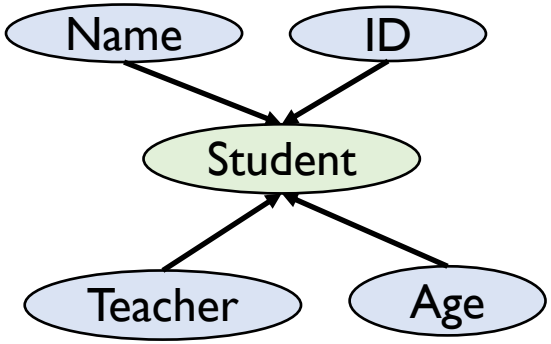
Structured Evidence

- Existing work mainly considers free-form language as evidence representation

- Structured data is also ubiquitous form of world knowledge like Graph, Table, Website.

ID	Name	Age
1	xxx	12
2	yyy	13
3	zzz	14

```
<University>
<Student Id="1">
<Age>18</Age>
<Name>Chen</Name>
</Student>
</University>
```



UCSB

Fact Verification on Semi-structured Tables

- We explore the fact verification problem under semi-structured Wikipedia tables.
  - Ubiquitous in real-world applications
  - Contain both structured and un-structured forms

- TabFact Dataset
  - 16K open-domain Wikipedia tables
  - 118K statements divided into entailed or refuted categories.

UCSB

TabFact Examples

United States House of Representatives Elections, 1972

District	Incumbent	Party	Result	Candidates
California 3	John E. Moss	democratic	re-elected	John E. Moss (d) 69.9% John Rakus (r) 30.1%
California 5	Phillip Burton	democratic	re-elected	Phillip Burton (d) 81.8% Edlo E. Powell (r) 18.2%
California 8	George Paul Miller	democratic	lost renomination democratic hold	Pete Stark (d) 52.9% Lew M. Warden , Jr. (r) 47.1%
California 14	Jerome R. Waldie	republican	re-elected	Jerome R. Waldie (d) 77.6% Floyd E. Sims (r) 22.4%
California 15	John J. Mcfall	republican	re-elected	John J. Mcfall (d) unopposed

1. John E. Moss and Phillip Burton are both re-elected in the house of representative election in 1972.

UCSB

TabFact Examples

United States House of Representatives Elections, 1972

District	Incumbent	Party	Result	Candidates
California 3	John E. Moss	democratic	re-elected	John E. Moss (d) 69.9% John Rakus (r) 30.1%
California 5	Phillip Burton	democratic	re-elected	Phillip Burton (d) 81.8% Edlo E. Powell (r) 18.2%
California 8	George Paul Miller	democratic	lost renomination democratic hold	Pete Stark (d) 52.9% Lew M. Warden , Jr. (r) 47.1%
California 14	Jerome R. Waldie	republican	re-elected	Jerome R. Waldie (d) 77.6% Floyd E. Sims (r) 22.4%
California 15	John J. Mcfall	republican	re-elected	John J. Mcfall (d) unopposed

2. In the election, four out of five incumbents are re-elected.

UCSB

TabFact Examples

United States House of Representatives Elections, 1972

District	Incumbent	Party	Result	Candidates
California 3	John E. Moss	democratic	re-elected	John E. Moss (d) 69.9% John Rakus (r) 30.1%
California 5	Phillip Burton	democratic	re-elected	Phillip Burton (d) 81.8% Edlo E. Powell (r) 18.2%
California 8	George Paul Miller	democratic	lost renomination democratic hold	Pete Stark (d) 52.9% Lew M. Warden , Jr. (r) 47.1%
California 14	Jerome R. Waldie	republican	re-elected	Jerome R. Waldie (d) 77.6% Floyd E. Sims (r) 22.4%
California 15	John J. Mcfall	republican	re-elected	John J. Mcfall (d) unopposed

3. John E. Moss and George Paul Miller are both re-elected in the house of representative election.

UCSB

TabFact Examples

United States House of Representatives Elections, 1972

District	Incumbent	Party	Result	Candidates
California 3	John E. Moss	democratic	re-elected	John E. Moss (d) 69.9% John Rakus (r) 30.1%
California 5	Phillip Burton	democratic	re-elected	Phillip Burton (d) 81.8% Edlo E. Powell (r) 18.2%
California 8	George Paul Miller	democratic	lost renomination democratic hold	Pete Stark (d) 52.9% Lew M. Warden , Jr. (r) 47.1%
California 14	Jerome R. Waldie	republican	re-elected	Jerome R. Waldie (d) 77.6% Floyd E. Sims (r) 22.4%
California 15	John J. Mcfall	republican	re-elected	John J. Mcfall (d) unopposed

4. There are five candidates in total, two of them are democrats and three of them are republicans.

UCSB

Challenges

- Mixed Reasoning in Semi-structured Input

UCSB

Challenges

- Mixed Reasoning in Semi-structured Input
  - Linguistic Reasoning on semantic-level

Result	Candidates
re-elected	John E. Moss (d) 69.9% John Rakus (r) 30.1%
re-elected	Phillip Burton (d) 81.8% Edlo E. Powell (r) 18.2%
lost renomination democratic hold	Pete Stark (d) 52.9% Lew M. Warden , Jr. (r) 47.1%
re-elected	Jerome R. Waldie (d) 77.6% Floyd E. Sims (r) 22.4%
re-elected	John J. Mcfall (d) unopposed

UCSB

Challenges

- Mixed Reasoning in Semi-structured Input
  - Linguistic Reasoning on semantic-level

Result	Candidates
re-elected	John E. Moss (d) 69.9% John Rakus (r) 30.1%
re-elected	Phillip Burton (d) 81.8% Edlo E. Powell (r) 18.2%
lost renomination democratic hold	Pete Stark (d) 52.9% Lew M. Warden , Jr. (r) 47.1%
re-elected	Jerome R. Waldie (d) 77.6% Floyd E. Sims (r) 22.4%
re-elected	John J. Mcfall (d) unopposed

  - Symbolic Reasoning on structure-level

District	Incumbent	Party
California 3	John E. Moss	democratic
California 5	Phillip Burton	democratic
California 8	George Paul Miller	democratic
California 14	Jerome R. Waldie	republican
California 15	John J. Mcfall	republican

UCSB

Models

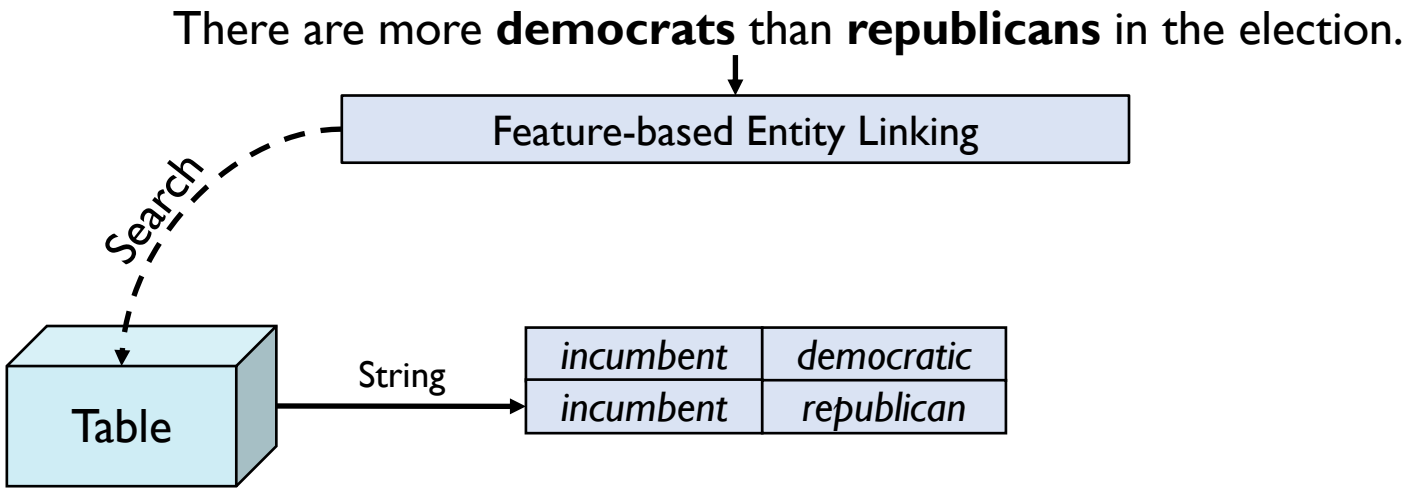
- Latent Program Analysis
  - Semantic-parsing baseline
  - Synthesize the latent logic forms for the statement and execute the logic form against table to verify.

- Table BERT
  - NLI baseline
  - Linearize the table as paragraph of sentences, then use large-scale pre-trained language model to verify.

UCSB

Latent Program Analysis

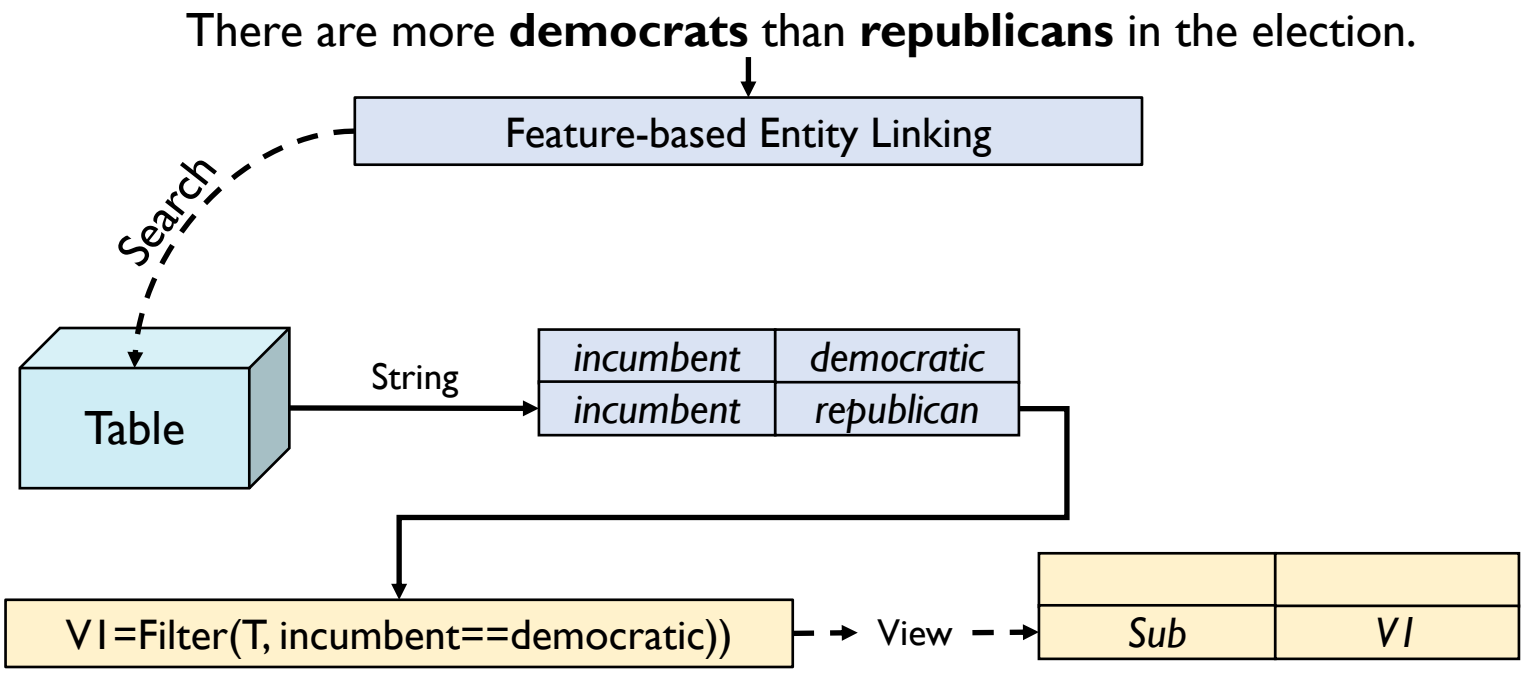
- Latent Program Analysis



UCSB

Latent Program Analysis

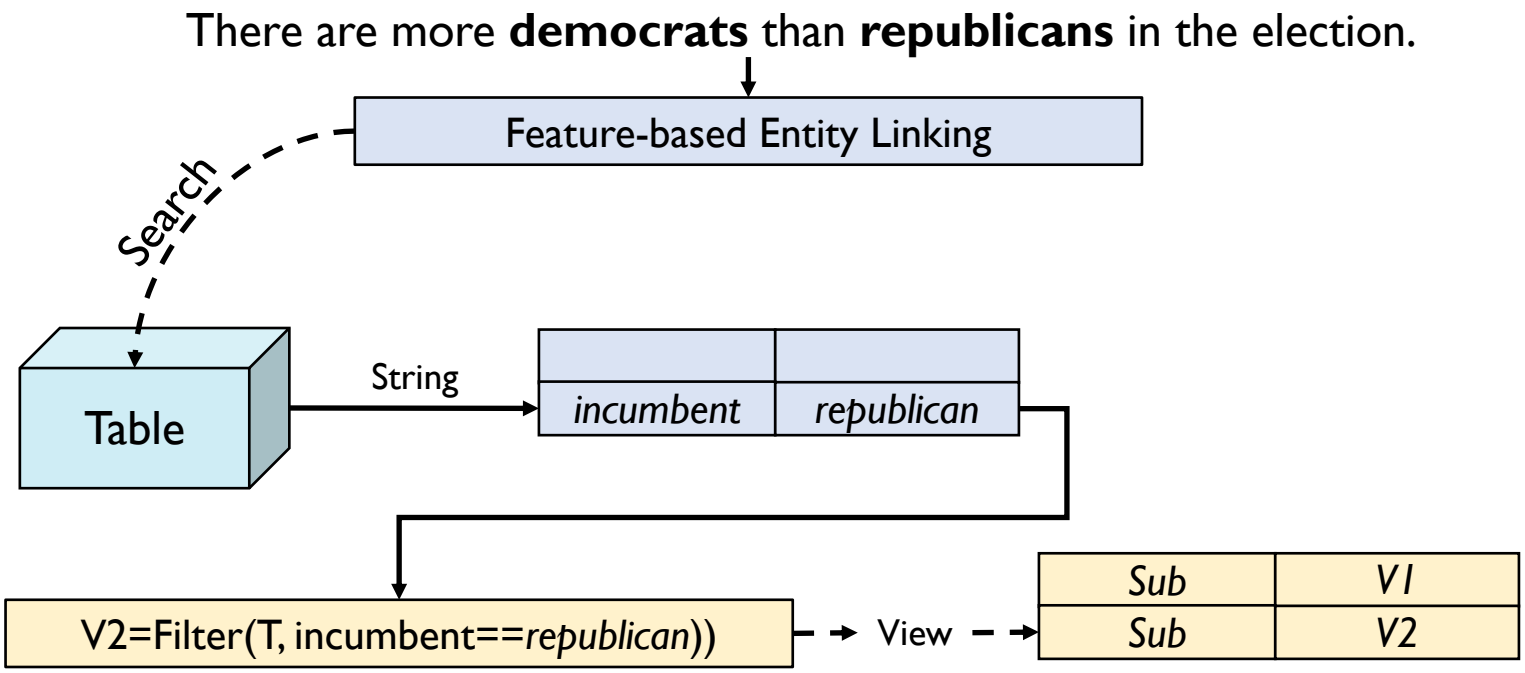
- Latent Program Analysis



UCSB

Latent Program Analysis

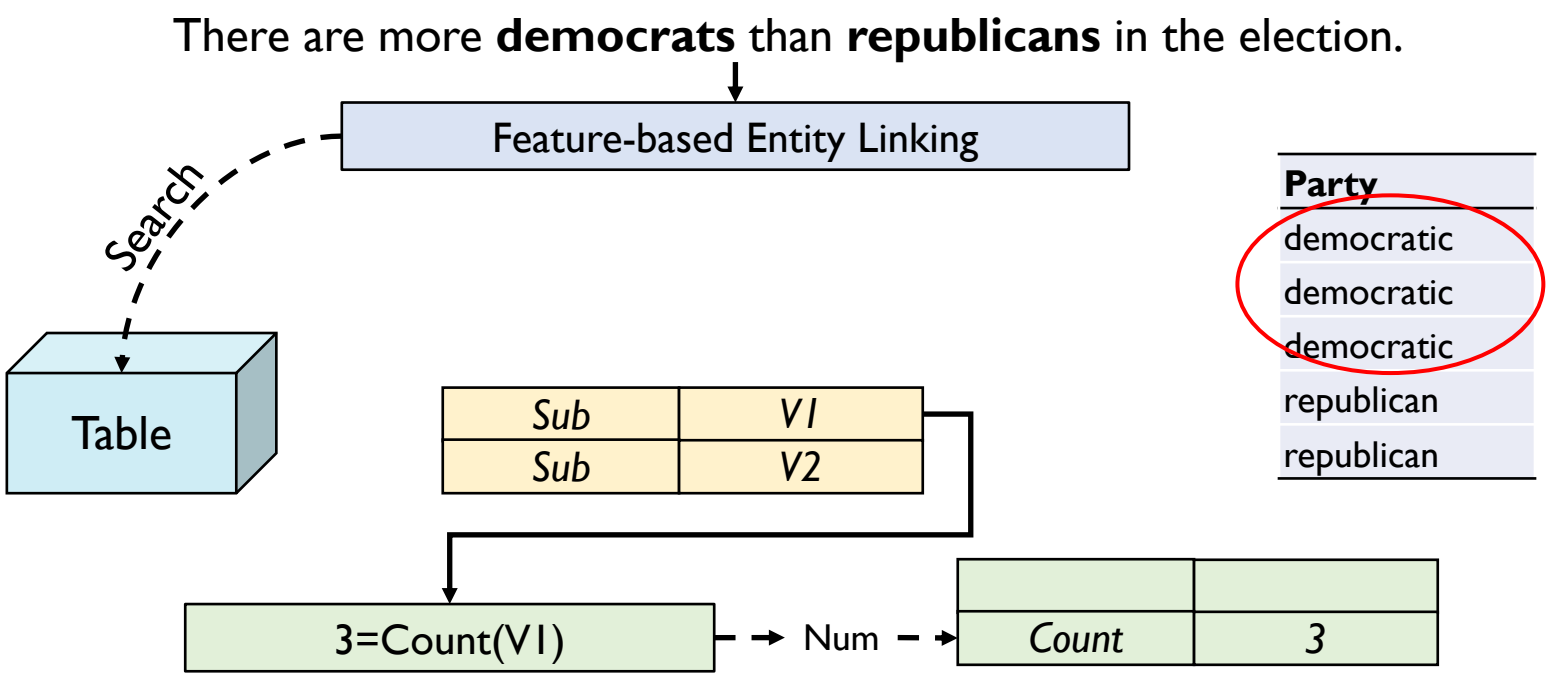
- Latent Program Analysis



UCSB

Latent Program Analysis

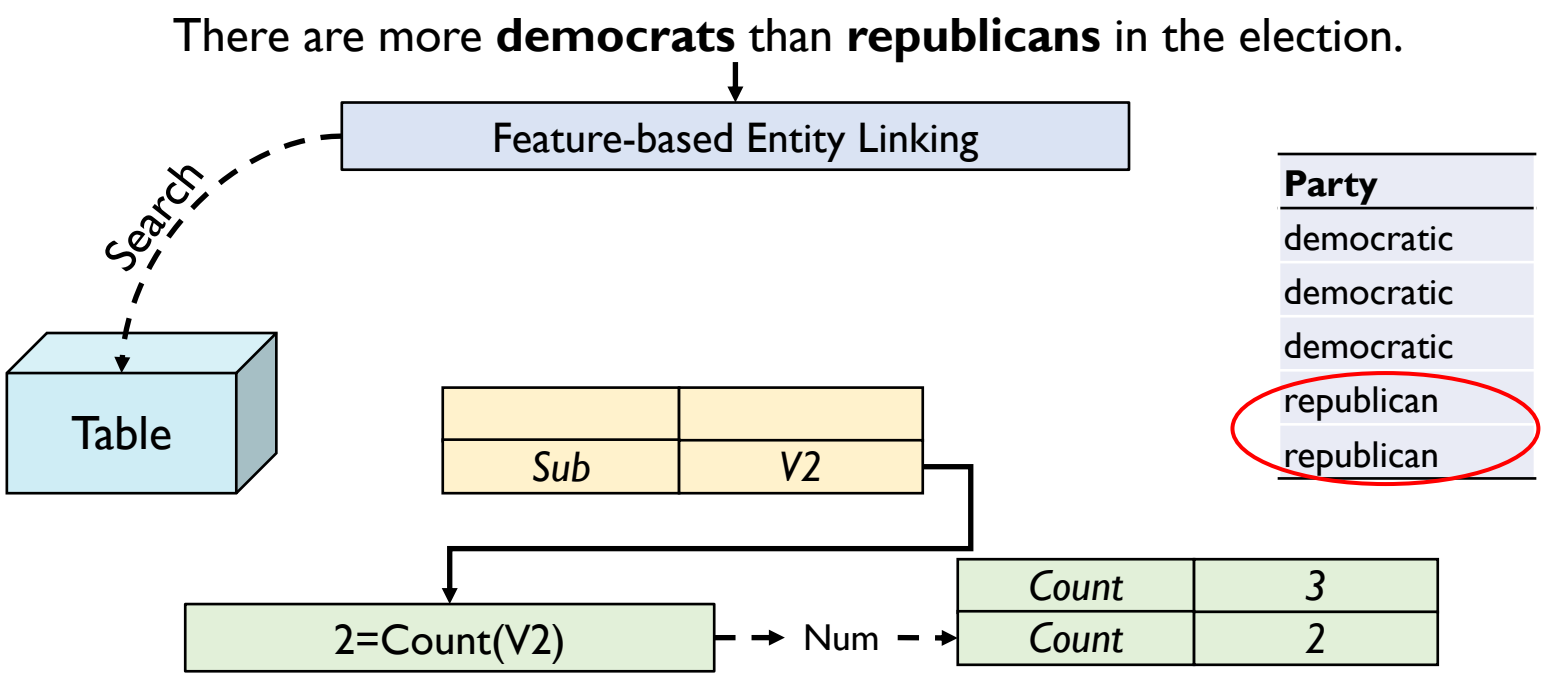
- Latent Program Analysis



UCSB

Latent Program Analysis

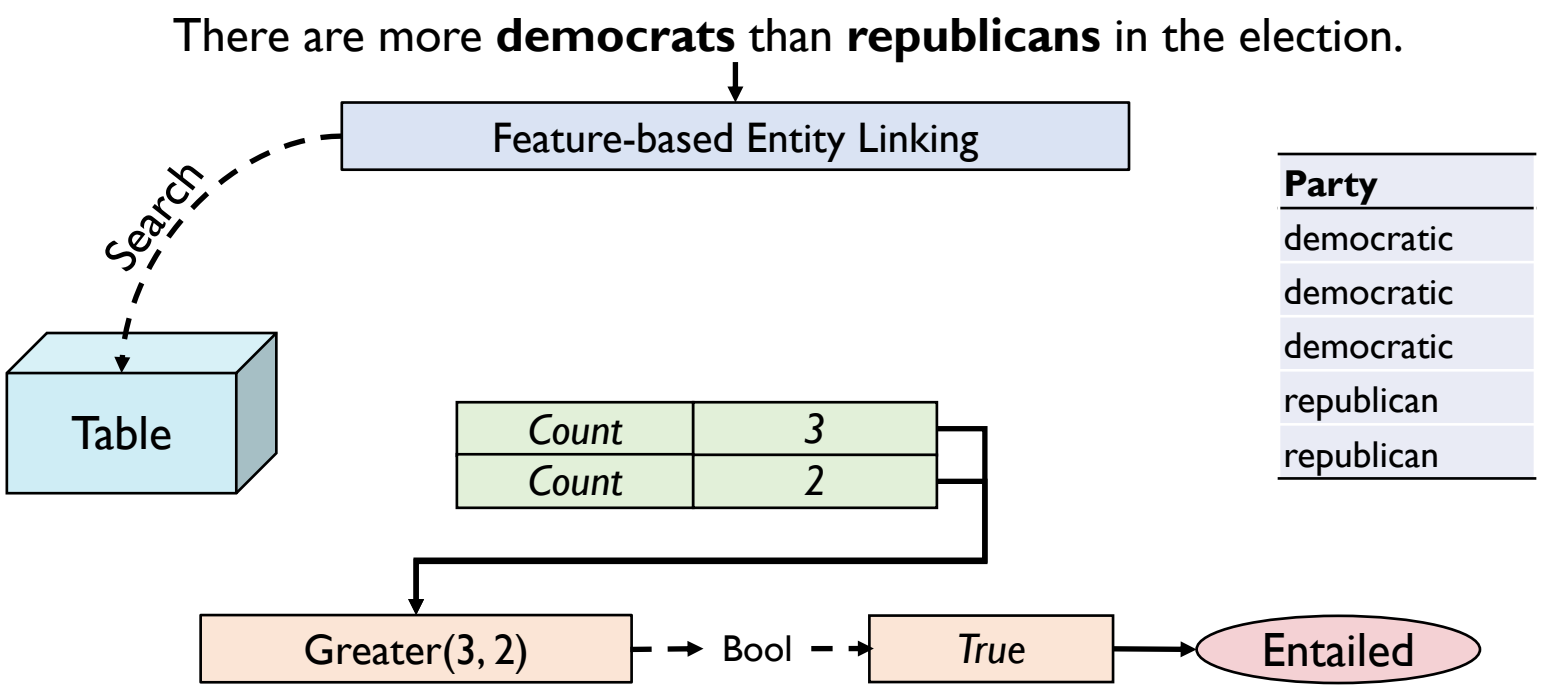
- Latent Program Analysis



UCSB

Latent Program Analysis

- Latent Program Analysis



UCSB

Table-BERT

- Scanning the Table

Game	Date	Opponent	Score
51	February 3 , 2009	Florida	3-4
52	February 4 , 2009	Buffalo	0-5
53	February 7 , 2010	Montreal	5-2

Horizontal Scan

UCSB

Table-BERT

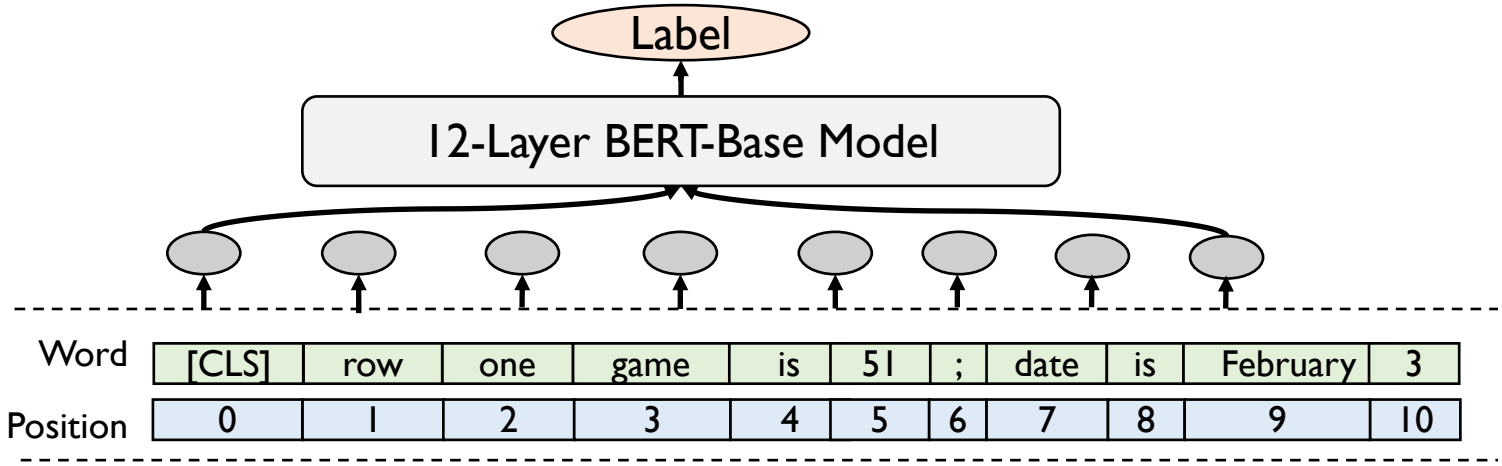
- Template the table with natural language

Word	[CLS]	row	one	game	is	51	;	date	is	February	3
Position	0	1	2	3	4	5	6	7	8	9	10

UCSB

Table-BERT

- Inference



Experimental Results

Model	Val	Test	Test (simple)	Test (complex)	Small Test
BERT classifier w/o Table	50.9	50.5	51.0	50.1	50.4
Table-BERT-Horizontal-F+T-Concatenate	50.7	50.4	50.8	50.0	50.3
Table-BERT-Vertical-F+T-Template	56.7	56.2	59.8	55.0	56.2
Table-BERT-Vertical-T+F-Template	56.7	57.0	60.6	54.3	55.5
Table-BERT-Horizontal-F+T-Template	66.0	65.1	79.0	58.1	67.9
Table-BERT-Horizontal-T+F-Template	**66.1**	**65.1**	**79.1**	**58.2**	**68.1**
LPA-Voting w/o Discriminator	57.7	58.2	68.5	53.2	61.5
LPA-Weighted-Voting	62.5	63.1	74.6	57.3	66.8
LPA-Ranking w/ Transformer	**65.2**	**65.0**	**78.4**	**58.5**	**68.6**
Human Performance	-	-	-	-	**92.1**

Experimental Results

- Latent Program Analysis
  - Pros: robust and explainable
  - Cons: Heavily rely on entity linking, the execution pipeline requires hand-written rules

Experimental Results

- Latent Program Analysis
  - Pros: robust and explainable
  - Cons: Heavily rely on entity linking, the execution pipeline requires hand-written rules

- Table BERT
  - Pros: simple and general
  - Cons: training unstable, black-box without explainability

Conclusion https://tabfact.github.io

- TabFact investigates an underexplored fact verification problem under structured data and release a dataset to promote research in this area.

- Both Proposed Symbolic and Neural models have pros and cons.

- How to design a more powerful model to combine linguistic and symbolic reasoning remains an open problem.

UCSB NLP Group

- Information Extraction and Reasoning
- Computational Social Science
- Language and Vision
- Natural Language Generation

We are a Top-3 most productive group in Natural Language Processing (CSRankings.org by NLP publications in 2018 and 2019)

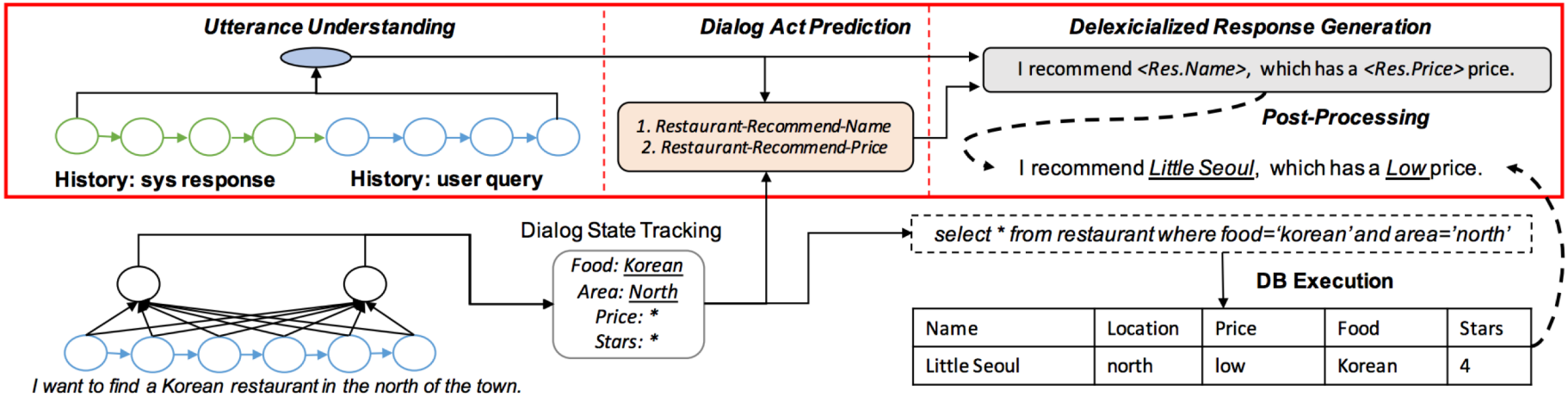
#	Institution	Count	Faculty
1	Carnegie Mellon University	21.0	16
2	Cornell University	11.5	9
3	University of California - Santa Barbara	8.8	5
4	Johns Hopkins University	8.0	6
5	University of Massachusetts Amherst	7.6	4
6	University of Pennsylvania	7.4	5
7	Stanford University	7.2	7
8	University of North Carolina	6.6	2
9	University of California - Berkeley	6.3	5
10	New York University	6.1	3

Research Activities at UCSB's NLP Lab

Natural Language Generation

(Dialog Systems and Summarization)

Semantically Conditioned Transformer for Dialogue Generation (ACL 2019)



Distributional Semantic Rewards (Li et al., EMNLP 2019)

- Prior work on Reinforcement Learning for NLP uses only BLEU score and discrete lexical rewards.
- We propose the use of distributional semantic rewards based on BERT to improve learning.
- Strong performance on abstractive summarization.

Language and Vision

(Video Captioning, Multimodal Captioning, Visual Storytelling, Vision-Language Navigation)

Hierarchical Deep Reinforcement Learning for Video Captioning (CVPR 2018)

http://arxiv.org/abs/1711.11135

Caption #1: A woman offers her dog some food.

Caption #2: A woman is eating and sharing food with her dog.

Caption #3: A woman is sharing a snack with a dog.

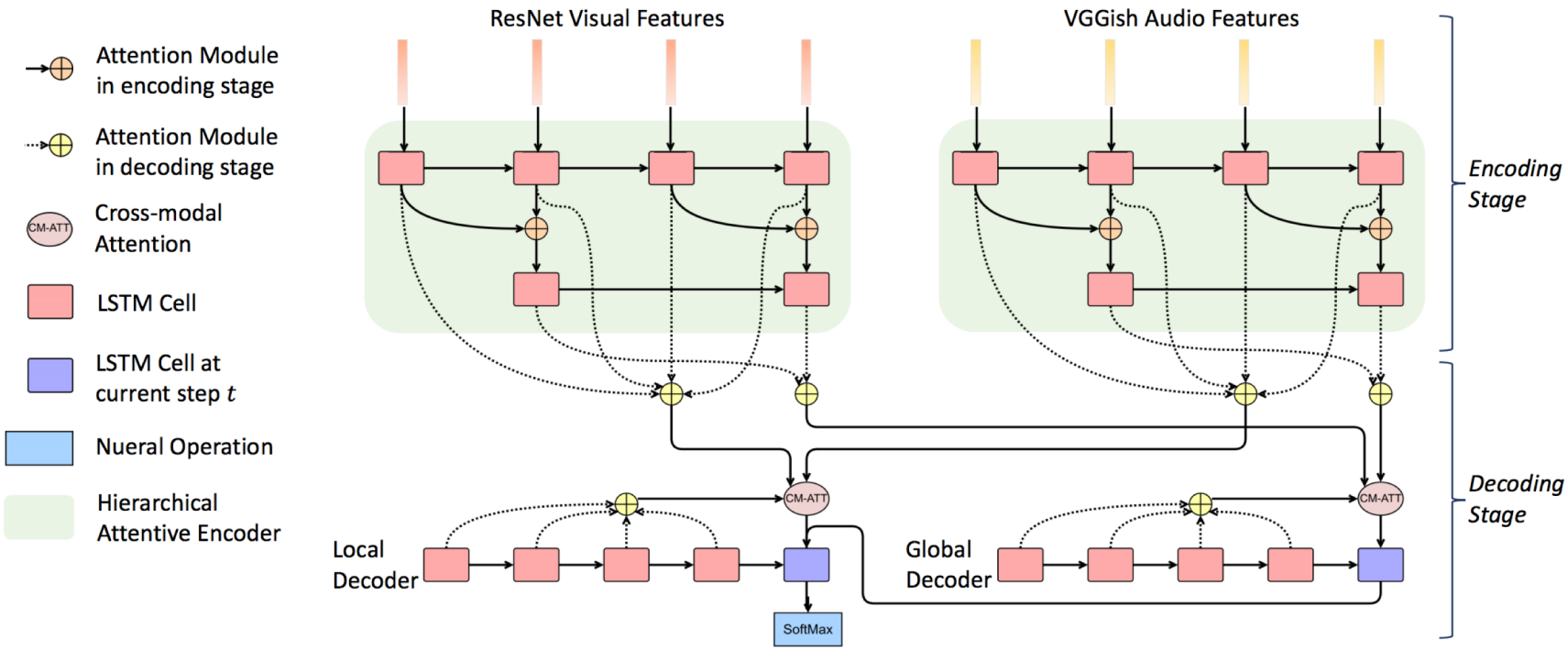
Caption: A person sits on a bed and puts a laptop into a bag. The person stands up, puts the bag on one shoulder, and walks out of the room.

Watch, Listen, and Describe: Globally and Locally Aligned Cross-Modal Attentions for Video Captioning (NAACL 2018)

https://arxiv.org/pdf/1804.05448.pdf

Ground Truth: A girl is singing.
A girl sings to a song.
Video Only: A woman is talking in a room.
Video + Audio: A girl is singing a song.

Figure 1: A video captioning example.



No Metrics Are Perfect: Adversarial Reward Learning for Visual Storytelling (ACL 2018)

(a) (b) (c) (d) (e)

Captions:

(a) A small boy and a girl are sitting together.
(b) Two kids sitting on a porch with their backpacks on.
(c) Two young kids with backpacks sitting on the porch.
(d) Two young children that are very close to one another.
(e) A boy and a girl smiling at the camera together.

Story #1: The brother and sister were ready for the first day of school. They were excited to go to their first day and meet new friends. They told their mom how happy they were. They said they were going to make a lot of new friends . Then they got up and got ready to get in the car .

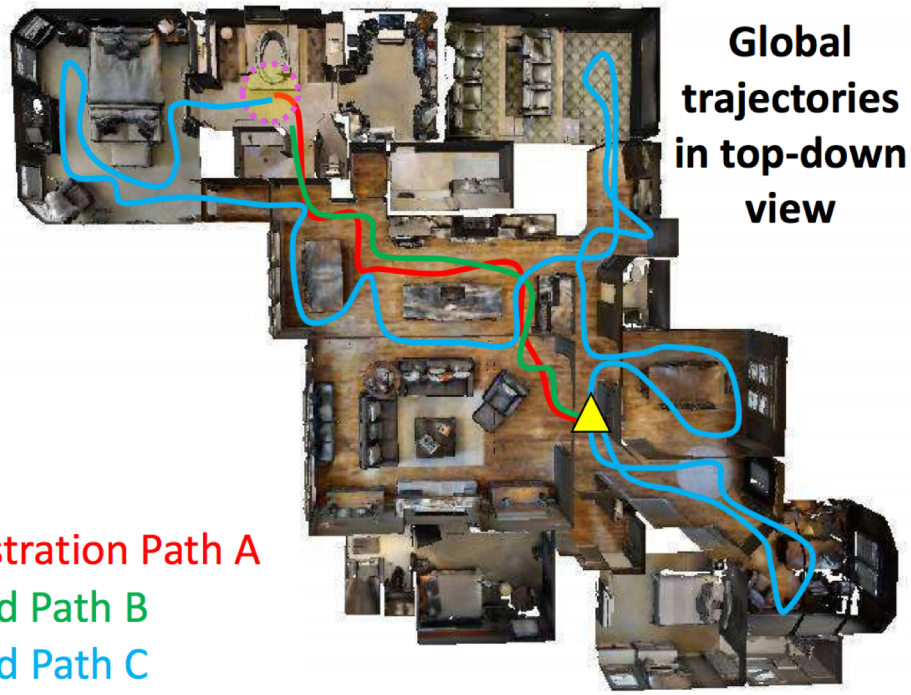
Story #2: The brother did not want to talk to his sister. The siblings made up. They started to talk and smile. Their parents showed up. They were happy to see them.

Exploration of Unseen Environment (CVPR 2019 Best Student Paper Award)

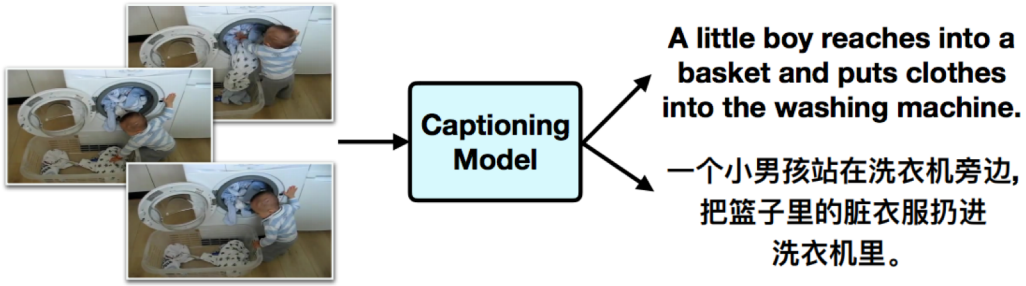
Instruction

Turn right and head towards the *kitchen*. Then turn left, pass a *table* and enter the *hallway*. Walk down the hallway and turn into the *entry way* to your right *without doors*. Stop in front of the *toilet*.

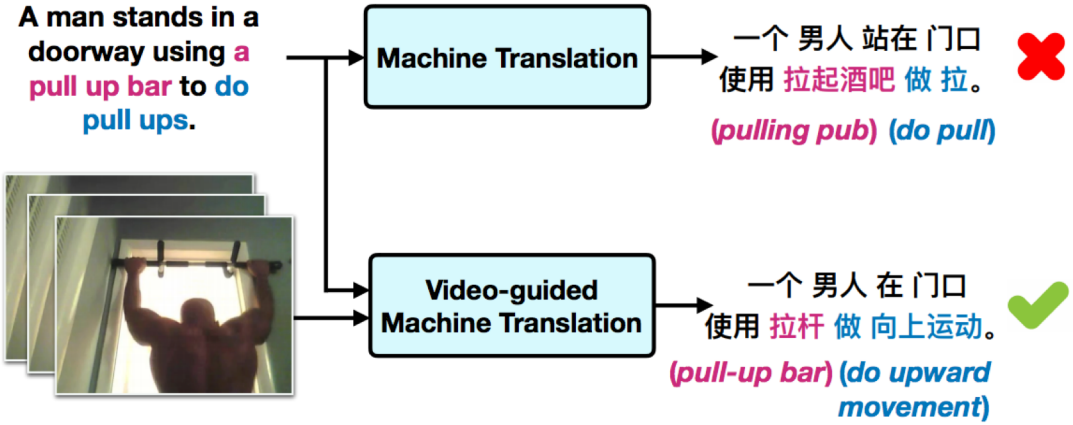
Local visual scene



VATEX:Video-Guided MT (ACL 2020 ALVR Challenge)



(a) Multilingual Video Captioning



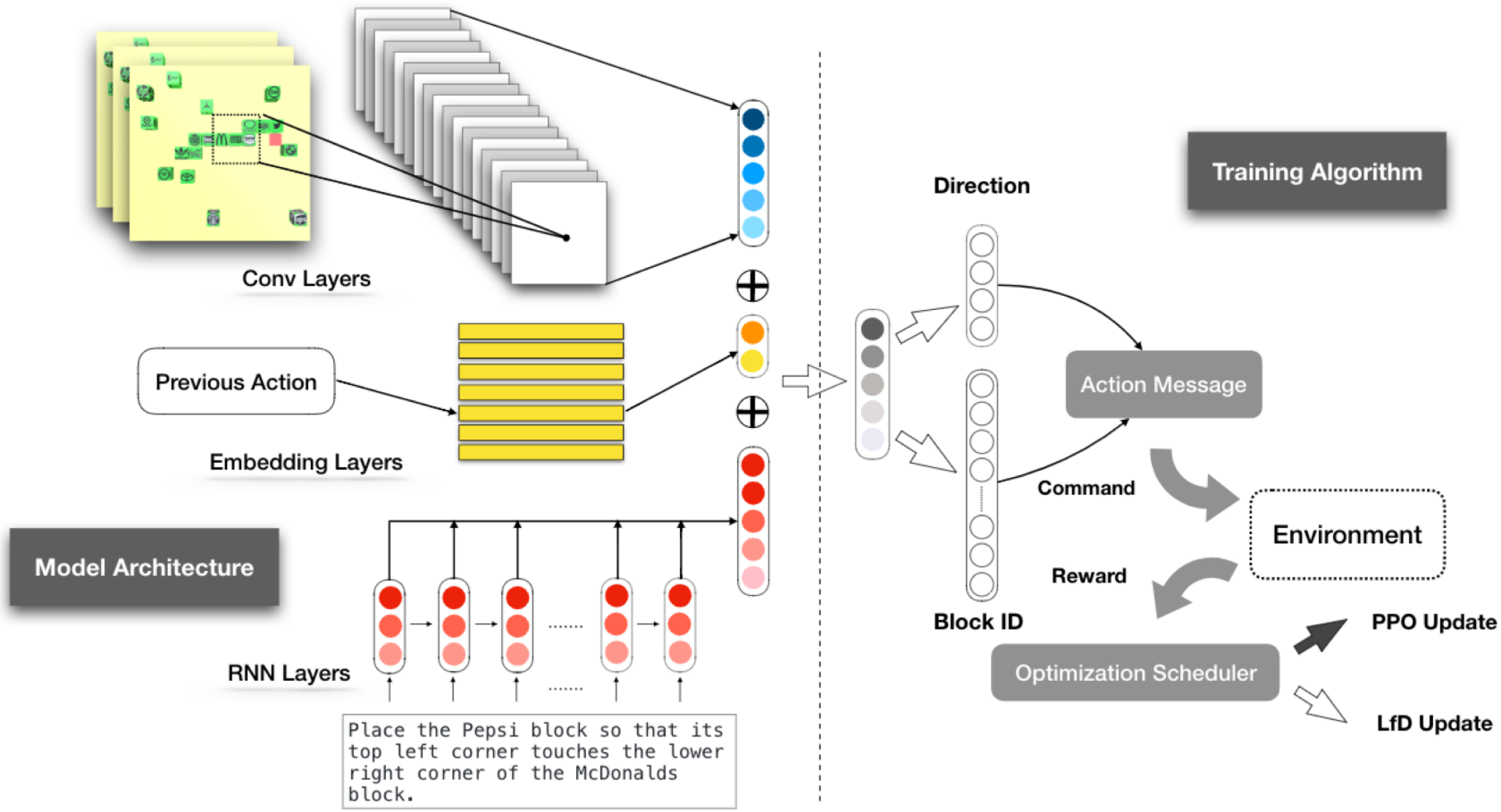
(b) Video-guided Machine Translation

Computational Social Science

A New Dataset for Hate Speech Intervention (Qian et al., EMNLP 2019)

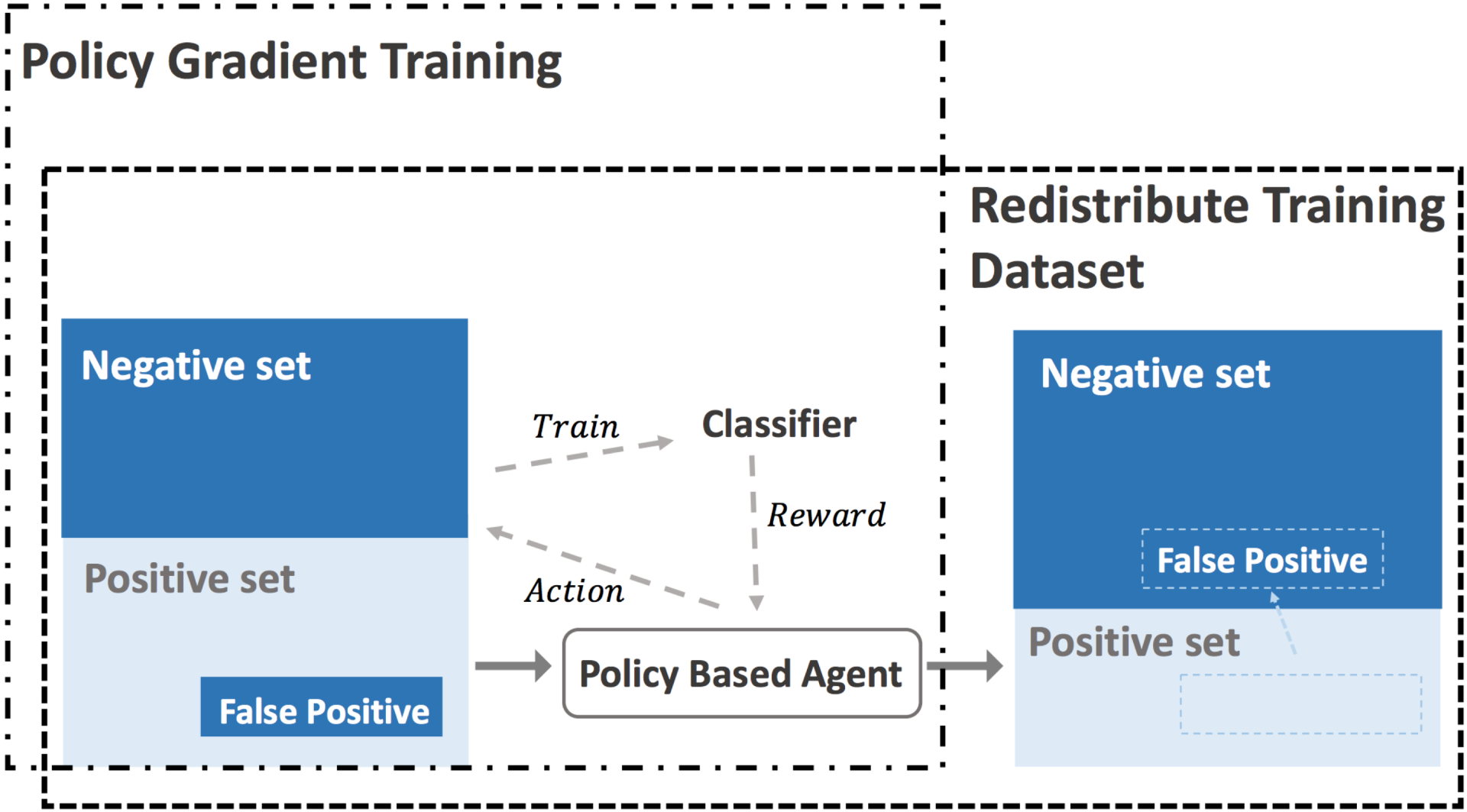
Deep Reinforcement Learning

Scheduled Policy Optimization (IJCAI-ECAI 2018)

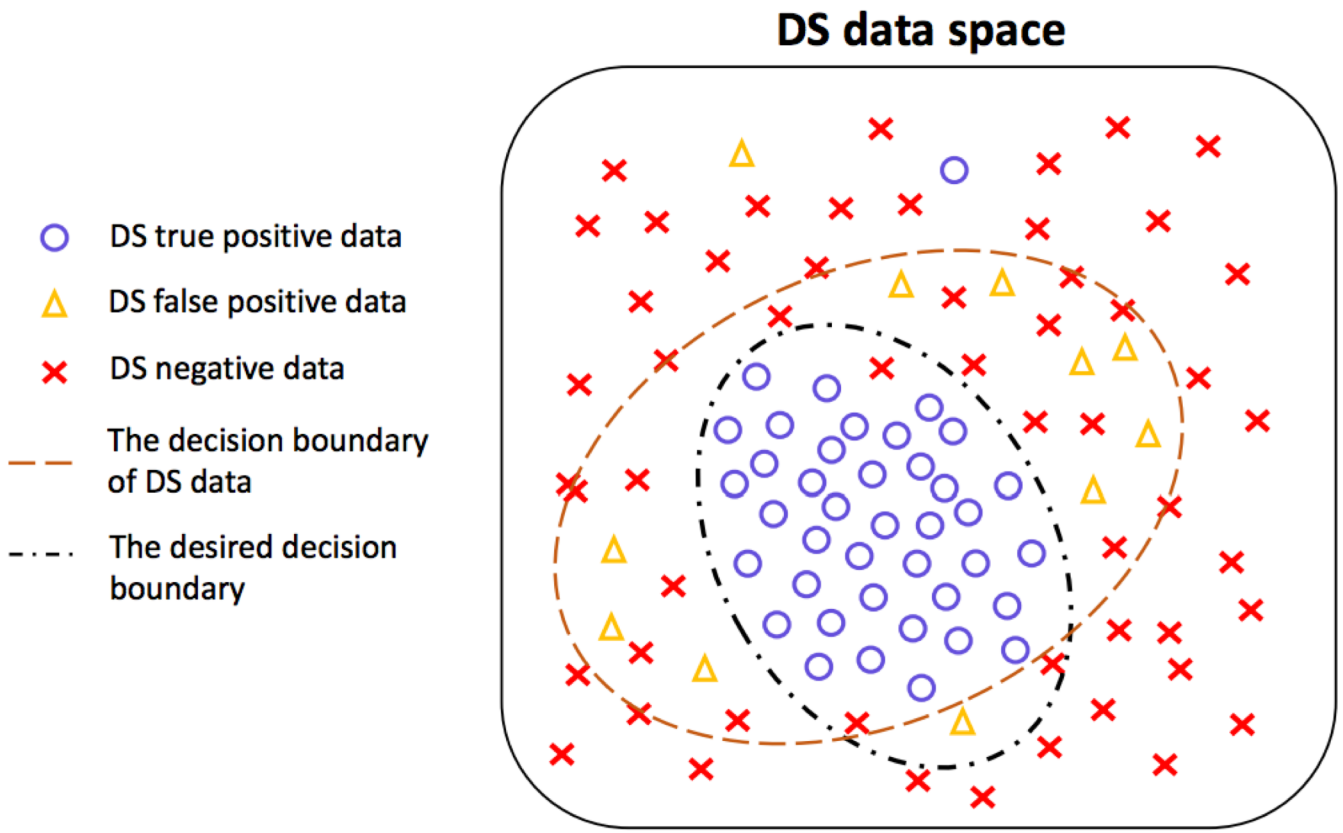


Information Extraction and Knowledge Graph

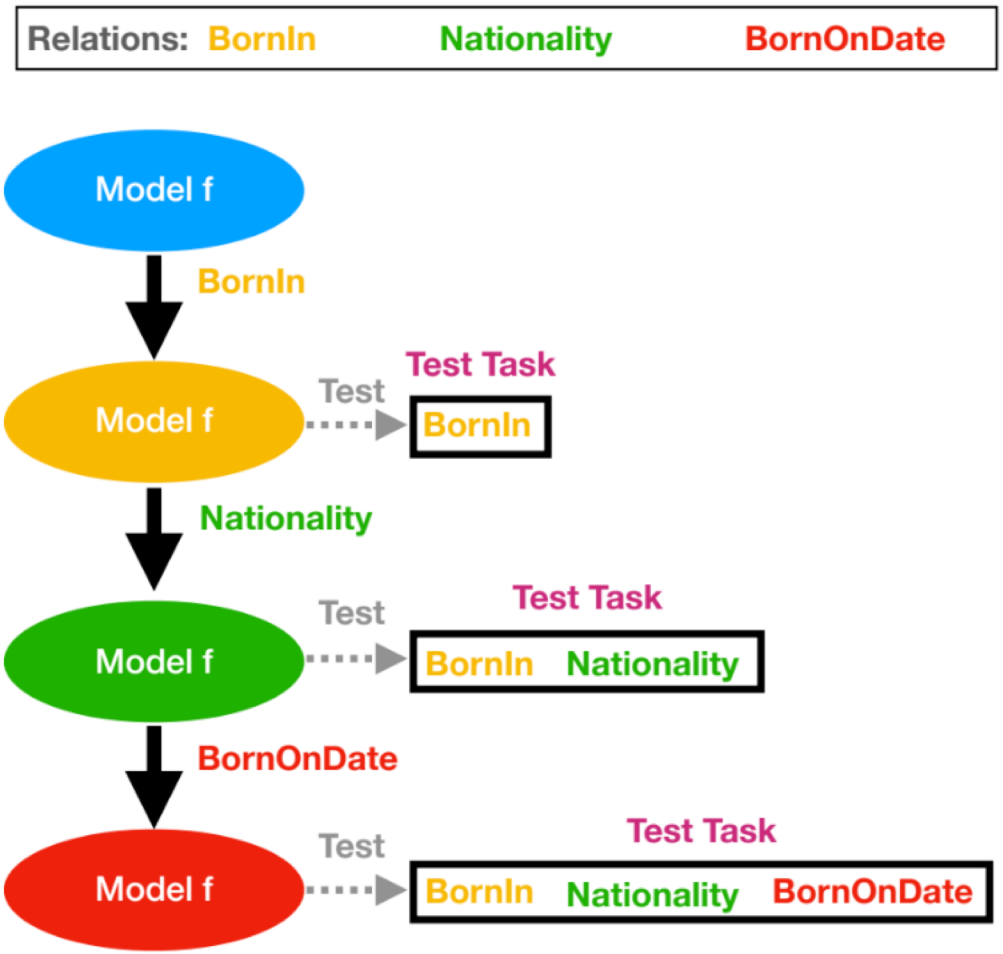
Robust Distant Supervision for Relation Extraction (ACL 2018)



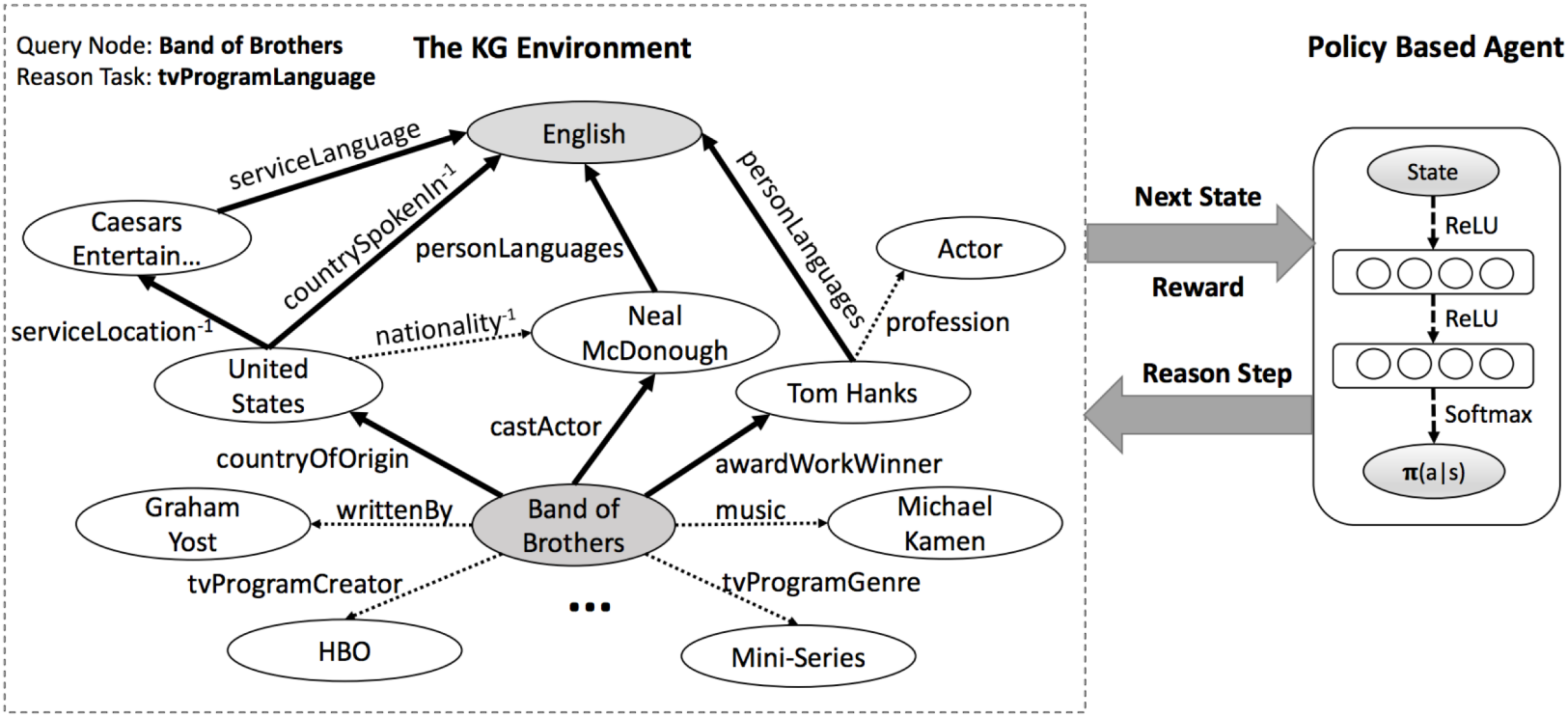
DSGAN: Adversarial Learning for Distant Supervision IE (ACL 2018)



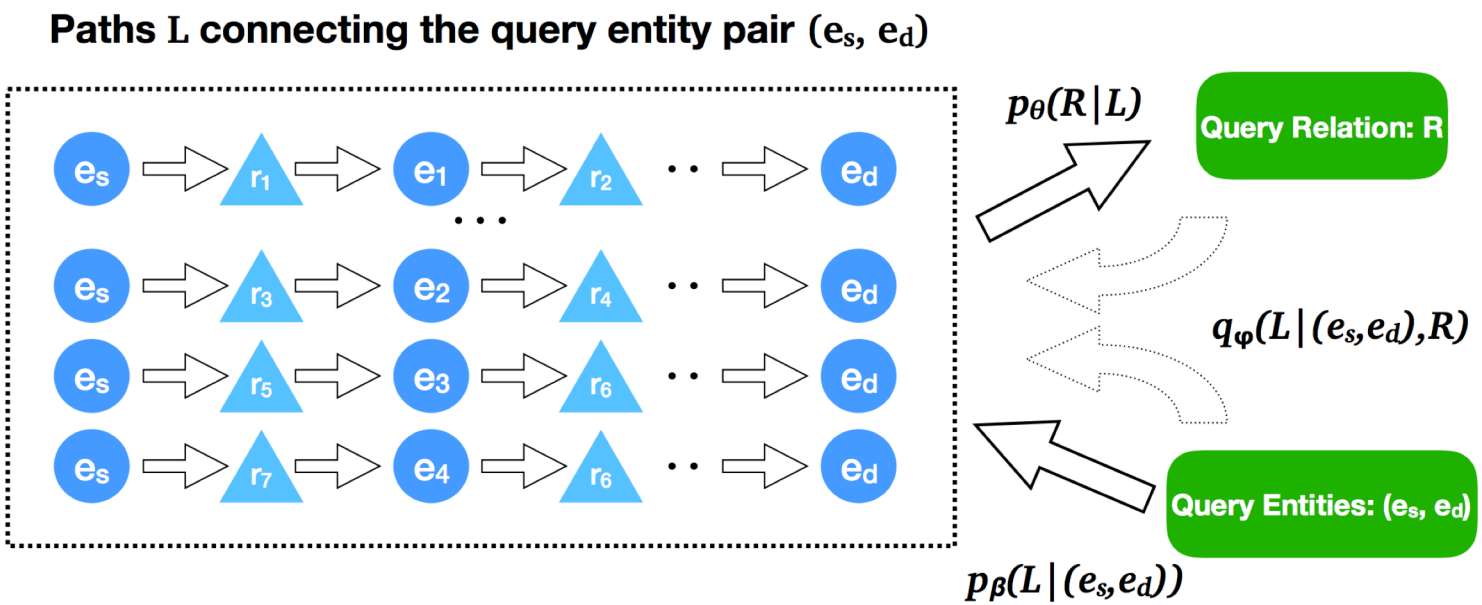
Lifelong Relation Extraction (NAACL 2019)



DeepPath: RL for KG Reasoning (EMNLP 2017)



DIVA: Variational Knowledge Graph Reasoning (NAACL 2018)

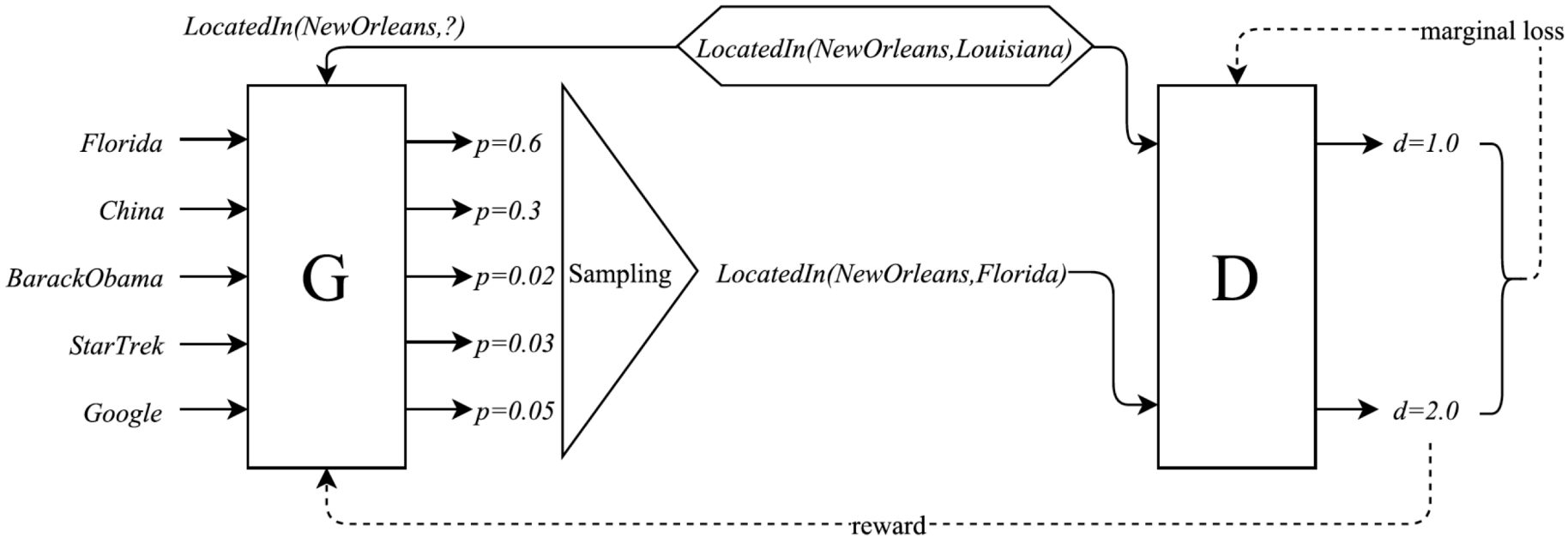


$$Obj = \sum_{(e_s,r,e_d)\in D} \log p(r|(e_s,e_d))$$

$$= \sum_{(e_s,r,e_d)\in D} \log \sum_{L} p_\theta(L|(e_s,e_d))p(r|L)$$

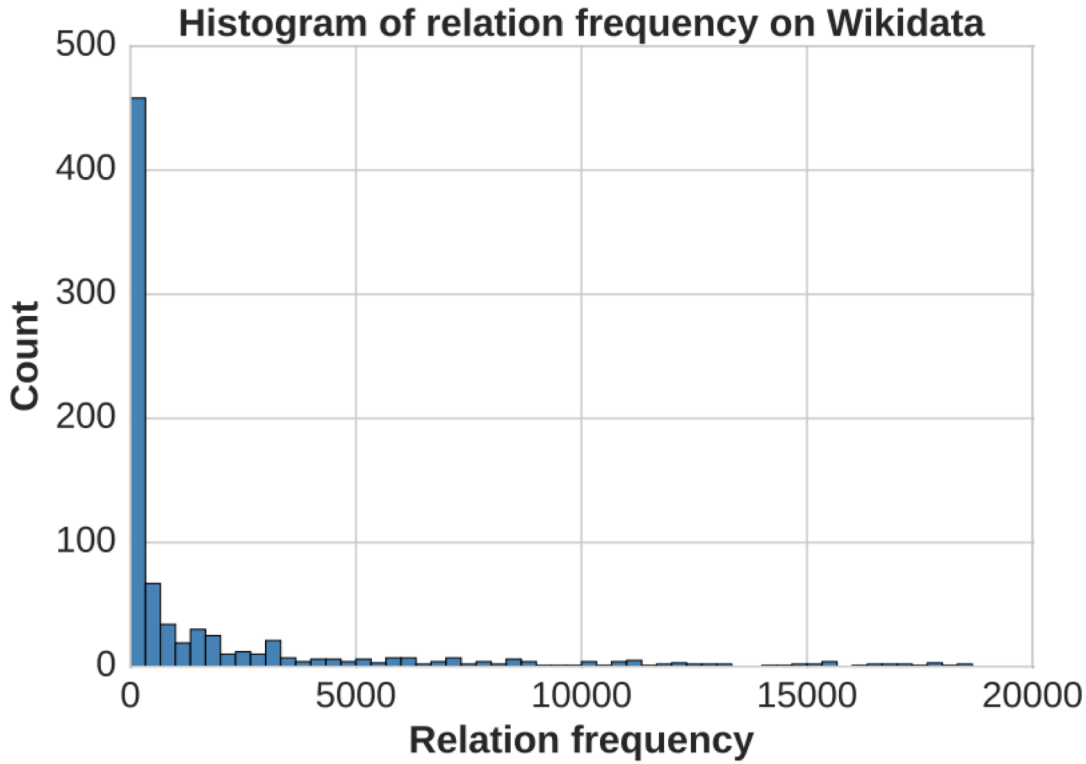
KBGAN: Adversarial Learning for Knowledge Graph Completion (NAACL 2018)

https://arxiv.org/abs/1711.04071



Idea: use adversarial learning to replace random sampling (from a uniform distribution).

One-Shot Relational Learning (EMNLP 2018)



Other related recent IE papers that I don't have time to cover

- Question Answering from both Knowledge Graph and Text Documents (ACL 2019)
- TweetQA: a new dataset for question answering on Twitter data (ACL 2019)
- Imposing Label-Relational Inductive Bias for Extremely Fine-Grained Entity Typing (NAACL 2019)
- A Variational Approach on Vocabulary Selection (NAACL 2019)

Acknowledgment

Sponsors: Adobe, Amazon, ByteDance, DARPA, Facebook, Google, IBM, Intel, LogMeIn, NVIDIA, and Tencent.

Thank you!

- UCSB NLP Group: nlp.cs.ucsb.edu

- AREL: https://github.com/eric-xw/AREL
- DeepPath: https://github.com/xwhan/DeepPath/
- Walk the Block: https://github.com/xwhan/walk_the_blocks
- Cross-Lingual Dialog State Tracking: https://github.com/wenhuchen/Cross-Lingual-NBT
- MojiTalk: https://github.com/claude-zhou/MojiTalk

UCSB