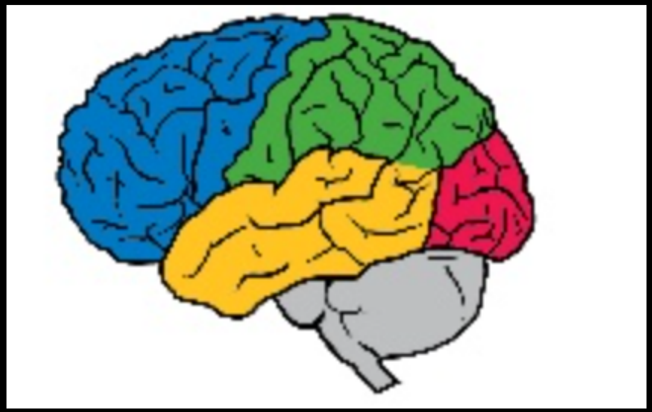
Knowledge Representation and Reasoning with Deep Neural Networks

Arvind Neelakantan

Advisor: Andrew McCallum

Committee: Rajesh Bhatt, Samuel Bowman, Subhransu Maji, Benjamin Marlin, Vijay Saraswat

UMASS AMHERST



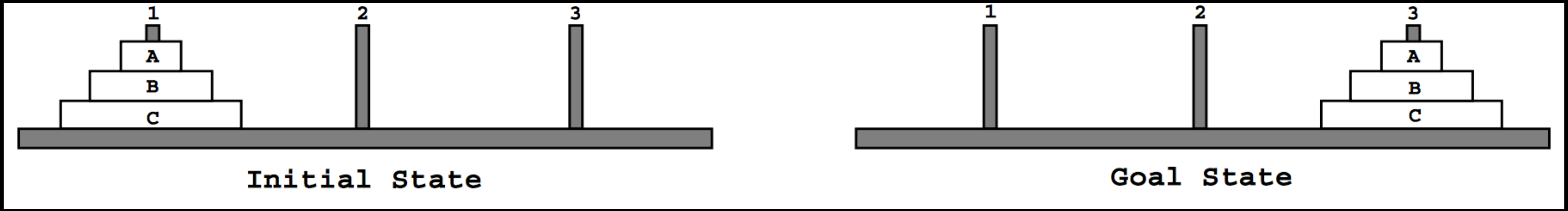
Knowledge Representation and Reasoning

- Represent world knowledge so that computers can use it
- Manipulating available knowledge to produce desired behavior
- Language understanding, robotics, ….

Early Systems

- Symbolic Representation
- Reasoning/Inference with search
- General Problem Solver (Simon et al., 1959), Cyc (Lenat et al., 1986), ...
- Precise

Early Systems



Knowledge: Permissible Transformations
Reasoning: Search Algorithm

Example

Date	Time	Opponent#	Rank#	Site	TV	Result	Attendance
September 3	6:00 PM	Missouri State*	#15	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	PPV	**W** 51–7	70,607
September 10	6:00 PM	New Mexico*	#14	War Memorial Stadium • Little Rock, AR	ESPNU	**W** 52–3	52,606
September 17	6:30 PM	Troy*	#14	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	CSS	**W** 38–28	69,861
September 24	2:30 PM	at #3 Alabama	#14	Bryant–Denny Stadium • Tuscaloosa, AL	CBS	**L** 14–38	101,821
October 1	11:00 AM	vs. #14 Texas A&M*	#18	Cowboys Stadium • Arlington, TX (Southwest Classic)	ESPN	**W** 42–38	69,838
October 8	6:00 PM	#15 Auburn	#10	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	ESPN	**W** 38–14	74,191

Example

Date	Time	Opponent#	Rank#	Site	TV	Result	Attendance
September 3	6:00 PM	Missouri State*	#15	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	PPV	**W** 51–7	70,607
September 10	6:00 PM	New Mexico*	#14	War Memorial Stadium • Little Rock, AR	ESPNU	**W** 52–3	52,606
September 17	6:30 PM	Troy*	#14	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	CSS	**W** 38–28	69,861
September 24	2:30 PM	at #3 Alabama	#14	Bryant–Denny Stadium • Tuscaloosa, AL	CBS	**L** 14–38	101,821
October 1	11:00 AM	vs. #14 Texas A&M*	#18	Cowboys Stadium • Arlington, TX (Southwest Classic)	ESPN	**W** 42–38	69,838
October 8	6:00 PM	#15 Auburn	#10	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	ESPN	**W** 38–14	74,191

Which venue has the biggest turnout ?

Example

Date	Time	Opponent#	Rank#	Site	TV	Result	Attendance
September 3	6:00 PM	Missouri State*	#15	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	PPV	**W** 51–7	70,607
September 10	6:00 PM	New Mexico*	#14	War Memorial Stadium • Little Rock, AR	ESPNU	**W** 52–3	52,606
September 17	6:30 PM	Troy*	#14	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	CSS	**W** 38–28	69,861
September 24	2:30 PM	at #3 Alabama	#14	Bryant–Denny Stadium • Tuscaloosa, AL	CBS	**L** 14–38	101,821
October 1	11:00 AM	vs. #14 Texas A&M*	#18	Cowboys Stadium • Arlington, TX (Southwest Classic)	ESPN	**W** 42–38	69,838
October 8	6:00 PM	#15 Auburn	#10	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	ESPN	**W** 38–14	74,191

Way more ambiguity than games

less direct

Which venue had the biggest turnout ?

1. Pick column Attendance
2. Get Position of Max entry
3. Print corresponding entry from column Site

Example

Date	Time	Opponent#	Rank#	Site	TV	Result	Attendance
September 3	6:00 PM	Missouri State*	#15	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	PPV	W 51–7	70,607
September 10	6:00 PM	New Mexico*	#14	War Memorial Stadium • Little Rock, AR	ESPNU	W 52–3	52,606
September 17	6:30 PM	Troy*	#14	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	CSS	W 38–28	69,861
September 24	2:30 PM	at #3 Alabama	#14	Bryant–Denny Stadium • Tuscaloosa, AL	CBS	L 14–38	101,821
October 1	11:00 AM	vs. #14 Texas A&M*	#18	Cowboys Stadium • Arlington, TX (Southwest Classic)	ESPN	W 42–38	69,838
October 8	6:00 PM	#15 Auburn	#10	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	ESPN	W 38–14	74,191

symbols

discrete processing

Which venue had the biggest turnout ?

select site (max attendance)

Manipulating Symbols and Discrete Processing!

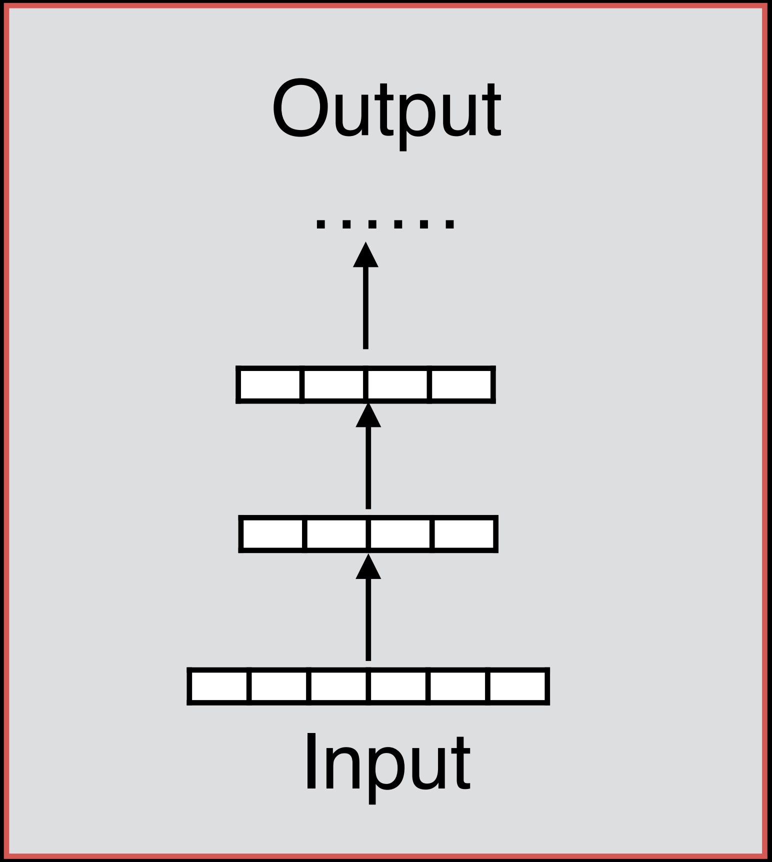
Early Systems: Issues

- Real-world data is challenging
- Lack of generalization to large number of symbols
- No learning

Recent Work

- Markov Logic Networks (Richardson & Domingos, 2006), Probabilistic Soft Logic (Kimmig et al., 2012), Semantic Parsers (Zelle & Mooney, 1996), ….
- Some components are learned
- Symbolic, most of the problems remain

Deep Neural Networks



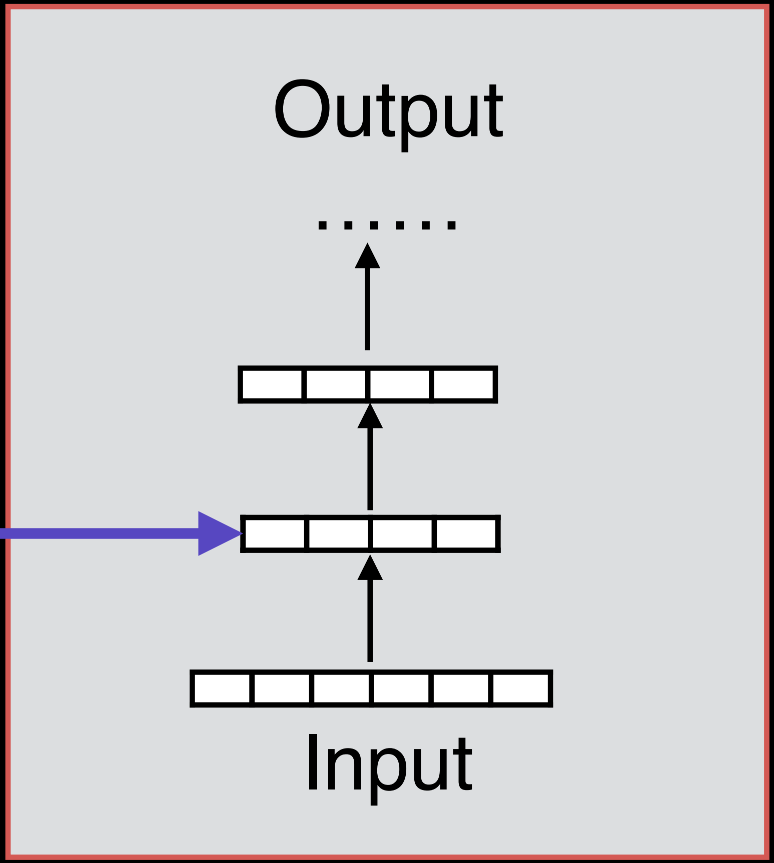
When I started Grad school, DNNs were creating lot of excitement

- Speech Recognition: ~5% absolute accuracy improvement
- Image Recognition: ~10% absolute accuracy improvement

(Dahl et al., 2012) and (Krizhevsky et al., 2012)

Deep Neural Networks

real-valued vector (distributed representations)



just a set of numbers, no symbols

all the processing is done in continuous space

Continuous data and processing through real-numbers

Transformation from input to output is learned from data using backpropagation algorithm

Perception vs Reasoning

- Input: Continuous Data vs Discrete Symbols
- Processing: Fuzzy vs Programs containing discrete operations, Rules, …

Parallel Research Areas

Deep Neural Networks for Knowledge Representation and Reasoning

Deep Neural Networks for Knowledge Representation and Reasoning

1. Can we represent symbols with distributed representations and learn them ?

2. Can we learn neural networks to perform reasoning with these representations ?

Deep Neural Networks for Knowledge Representation and Reasoning

1. Generalization via distributed representations

2. Powerful non-linear models

3. Learn end-to-end, handle *messy* real-world data

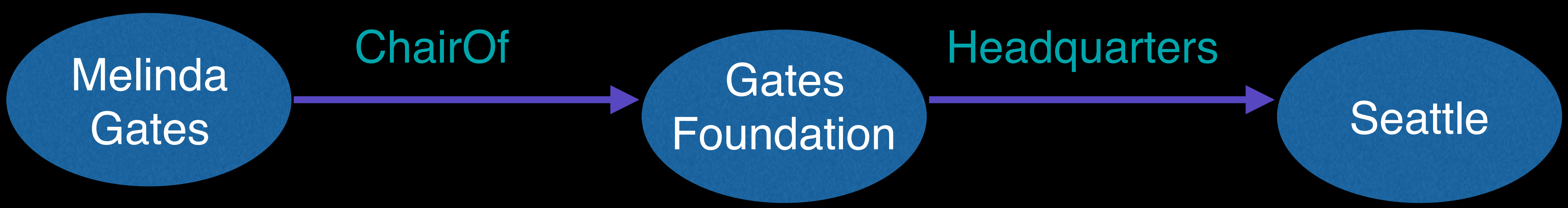
Deep Neural Networks for Knowledge Representation and Reasoning

1. Can we represent symbols with distributed representations and learn them ?

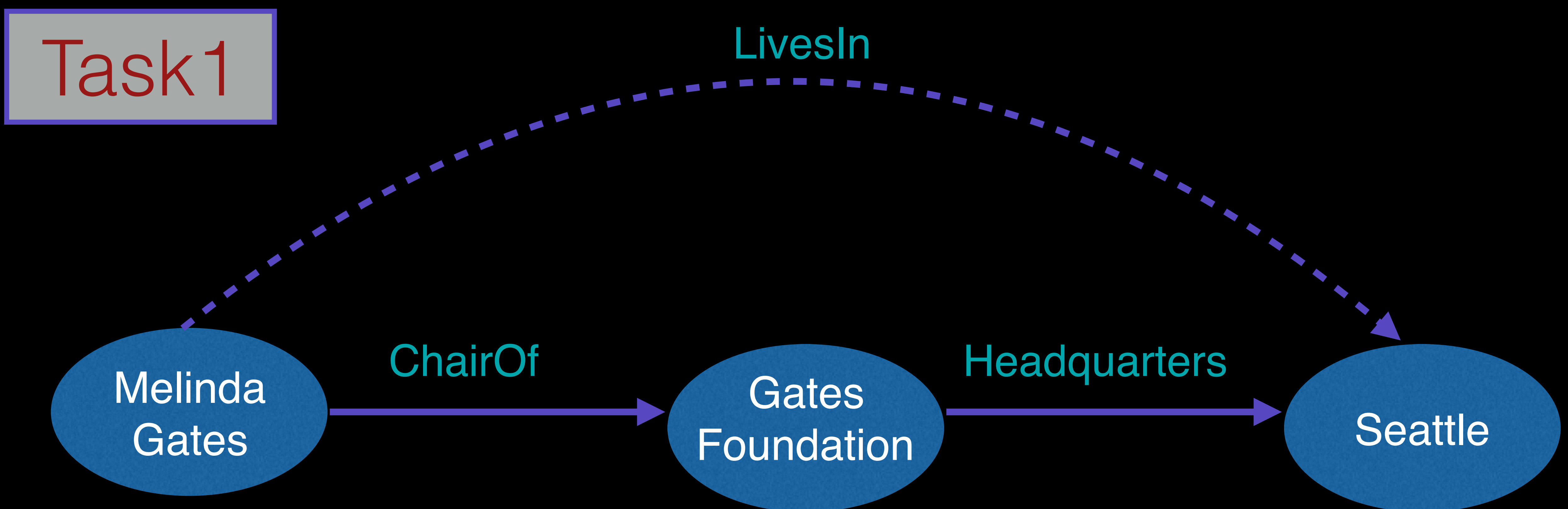
2. Can we train neural networks to perform reasoning with these representations ?

- Massive Structured Knowledge Base
- Semi-Structured Web Tables

Knowledge Graphs



Knowledge Graph Path Queries



ChairOf (A, X) ^ Headquarters (X, B) → LivesIn (A, B)

Program Induction/Semantic Parsing

Date	Time	Opponent#	Rank#	Site	TV	Result	Attendance
September 3	6:00 PM	Missouri State*	#15	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	PPV	**W** 51–7	70,607
September 10	6:00 PM	New Mexico*	#14	War Memorial Stadium • Little Rock, AR	ESPNU	**W** 52–3	52,606
September 17	6:30 PM	Troy*	#14	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	CSS	**W** 38–28	69,861
September 24	2:30 PM	at #3 Alabama	#14	Bryant–Denny Stadium • Tuscaloosa, AL	CBS	**L** 14–38	101,821
October 1	11:00 AM	vs. #14 Texas A&M*	#18	Cowboys Stadium • Arlington, TX (Southwest Classic)	ESPN	**W** 42–38	69,838
October 8	6:00 PM	#15 Auburn	#10	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	ESPN	**W** 38–14	74,191

Program Induction/Semantic Parsing

Date	Time	Opponent#	Rank#	Site	TV	Result	Attendance
September 3	6:00 PM	Missouri State*	#15	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	PPV	**W** 51–7	70,607
September 10	6:00 PM	New Mexico*	#14	War Memorial Stadium • Little Rock, AR	ESPNU	**W** 52–3	52,606
September 17	6:30 PM	Troy*	#14	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	CSS	**W** 38–28	69,861
September 24	2:30 PM	at #3 Alabama	#14	Bryant–Denny Stadium • Tuscaloosa, AL	CBS	**L** 14–38	101,821
October 1	11:00 AM	vs. #14 Texas A&M*	#18	Cowboys Stadium • Arlington, TX (Southwest Classic)	ESPN	**W** 42–38	69,838
October 8	6:00 PM	#15 Auburn	#10	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	ESPN	**W** 38–14	74,191

Which venue had the biggest turnout ? => select site (max attendance)

how many games were telecasted in CBS ? => count(location == CBS)

Program Induction/Semantic Parsing

Task 2

Date	Time	Opponent#	Rank#	Site	TV	Result	Attendance
September 3	6:00 PM	Missouri State*	#15	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	PPV	**W** 51–7	70,607
September 10	6:00 PM	New Mexico*	#14	War Memorial Stadium • Little Rock, AR	ESPNU	**W** 52–3	52,606
September 17	6:30 PM	Troy*	#14	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	CSS	**W** 38–28	69,861
September 24	2:30 PM	at #3 Alabama	#14	Bryant–Denny Stadium • Tuscaloosa, AL	CBS	**L** 14–38	101,821
October 1	11:00 AM	vs. #14 Texas A&M*	#18	Cowboys Stadium • Arlington, TX (Southwest Classic)	ESPN	**W** 42–38	69,838
October 8	6:00 PM	#15 Auburn	#10	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	ESPN	**W** 38–14	74,191

Which venue had the biggest turnout ? => select site (max attendance)
how many games were telecasted in CBS ? => count(location == CBS)

Related Work in Reasoning

- Natural Language Inference/Textual Entailment
- Visual Question Answering
- Reading Comprehension
- ……

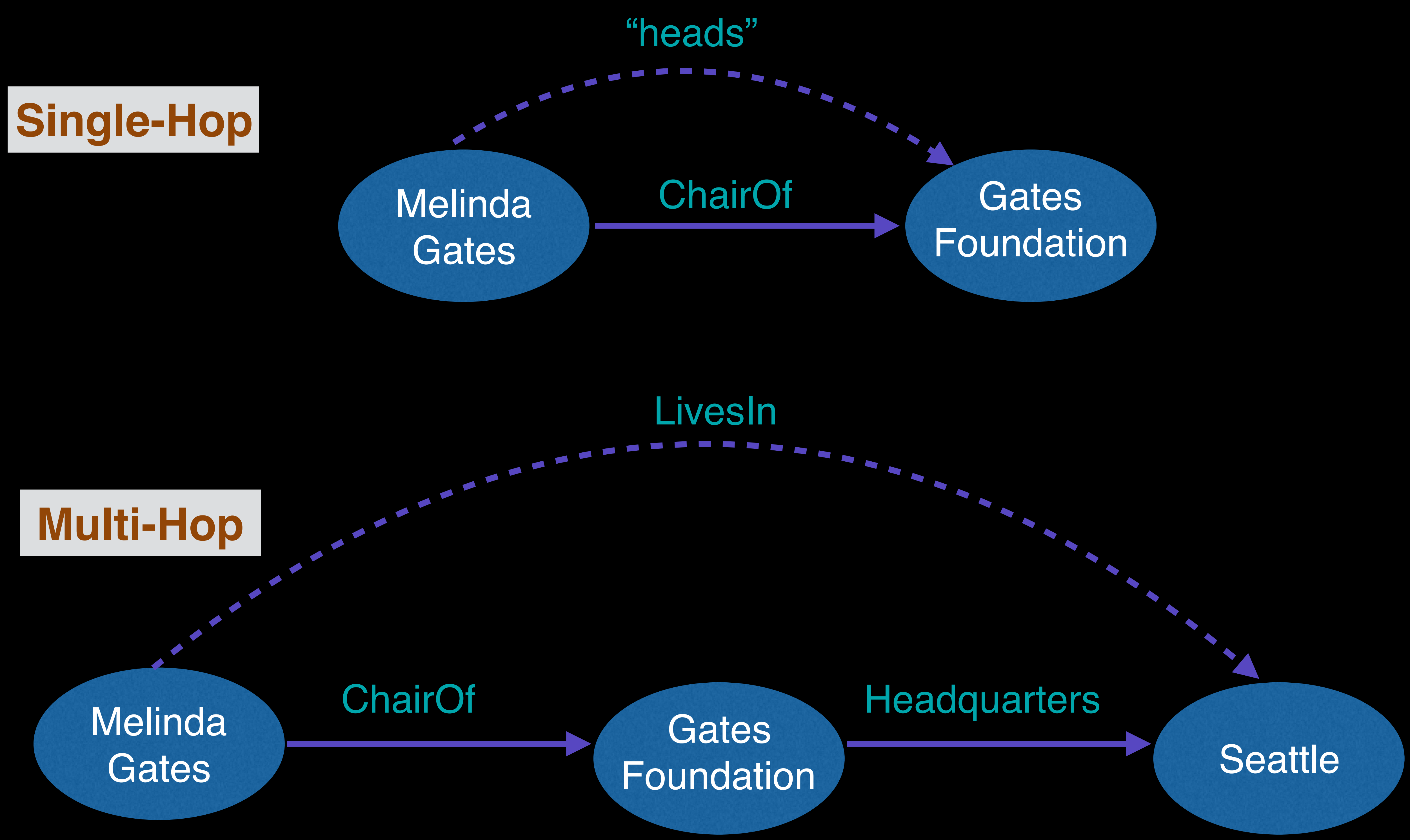
Task 1:
Knowledge Graph Path Queries

Arvind Neelakantan, Benjamin Roth, and Andrew McCallum. Knowledge base completion using compositional vector space models. Workshop on Automated Knowledge Base Construction at NIPS, 2014

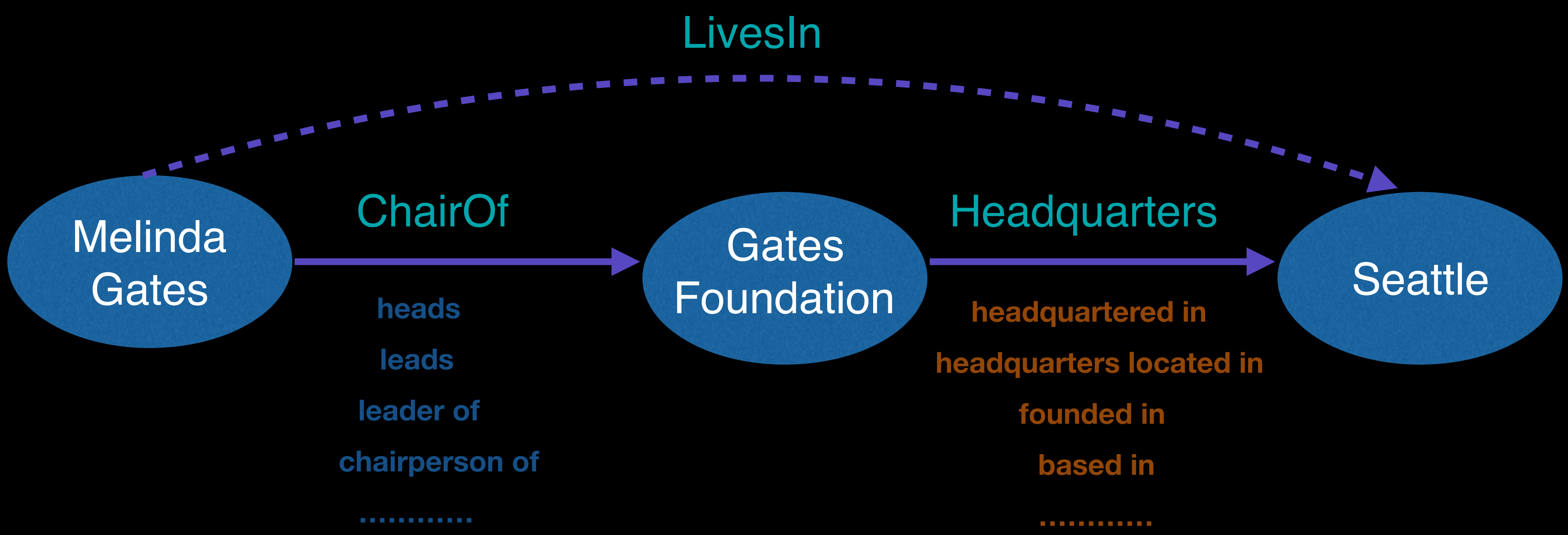
Arvind Neelakantan, Benjamin Roth, and Andrew McCallum. Compositional vector space models for knowledge base completion. ACL, 2015

Rajarshi Das, Arvind Neelakantan, David Belanger, and Andrew McCallum. Chains of reasoning over entities, relations, and text using recurrent neural networks. EACL, 2017

Path Queries



Motivation



Previous Work:

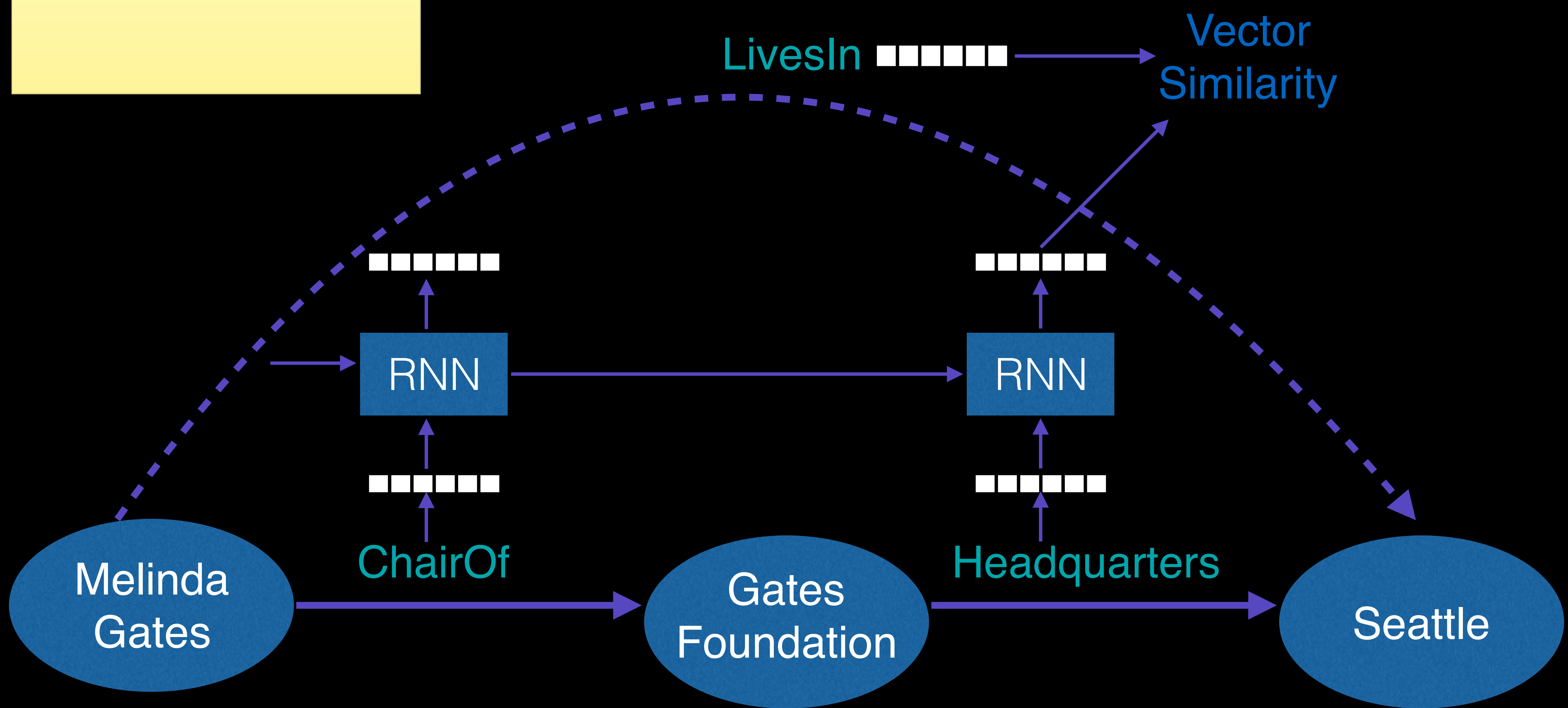
- Symbolic
  - Path Ranking Algorithm (Lao et al., 2011) & Sherlock (Schoenmackers et al. , 2010)
  - Combinatorial Explosion => Poor Generalization

Multi-hop Reasoning: Current Methods do not generalize to unseen paths

Model

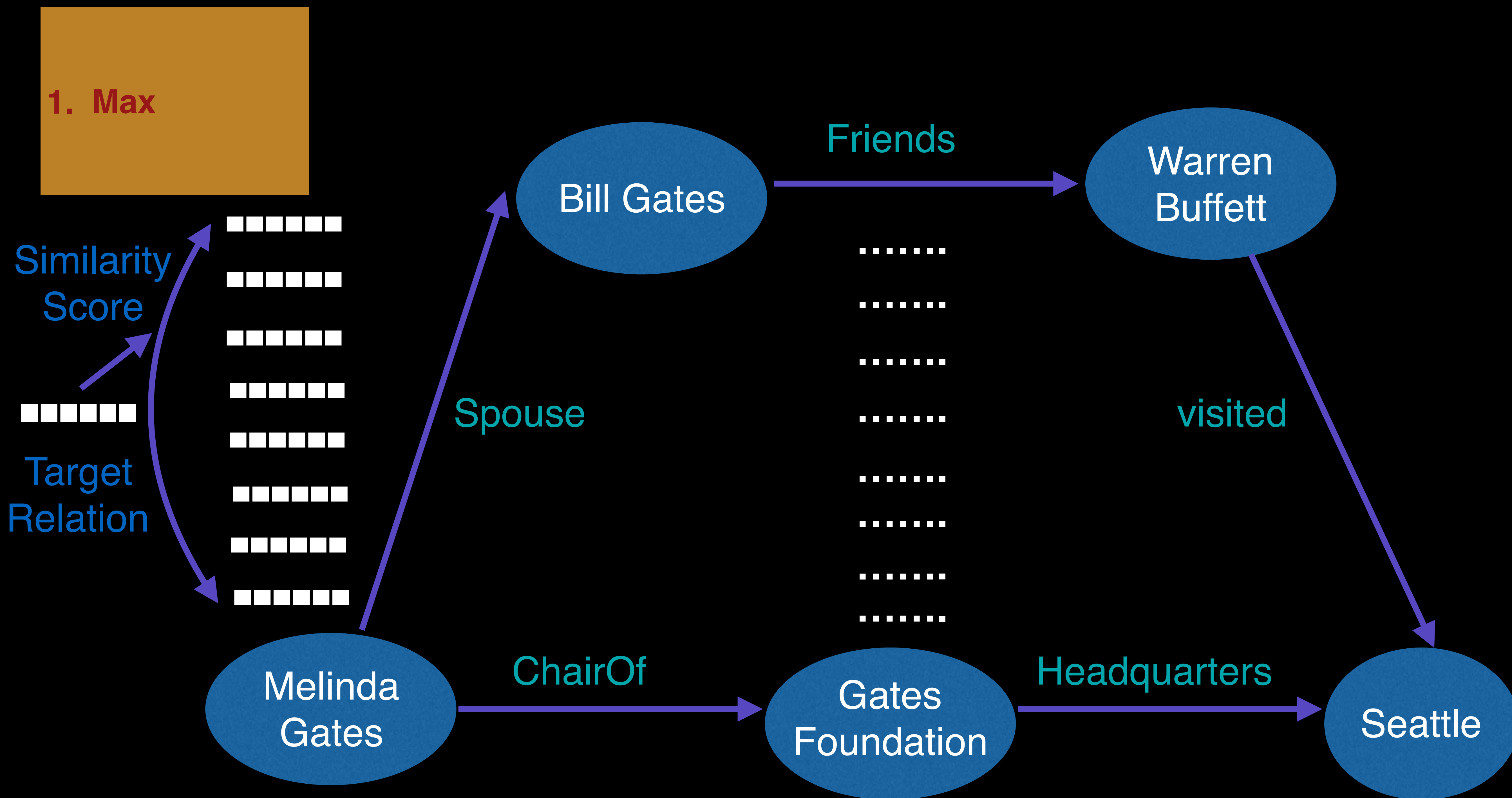
explain arrow

(Neelakantan, Roth, McCallum, 2014)



Generalize to Unseen Paths!

Selection/Attention



Train with backprop!

Data

Entity Pairs	**3.2M**
Facts	**52M**
Relations	**51K**
Relation types tested	**46**
Total # paths	**191M**
Average Path Length	**4.7**
Maximum Path Length	**7**

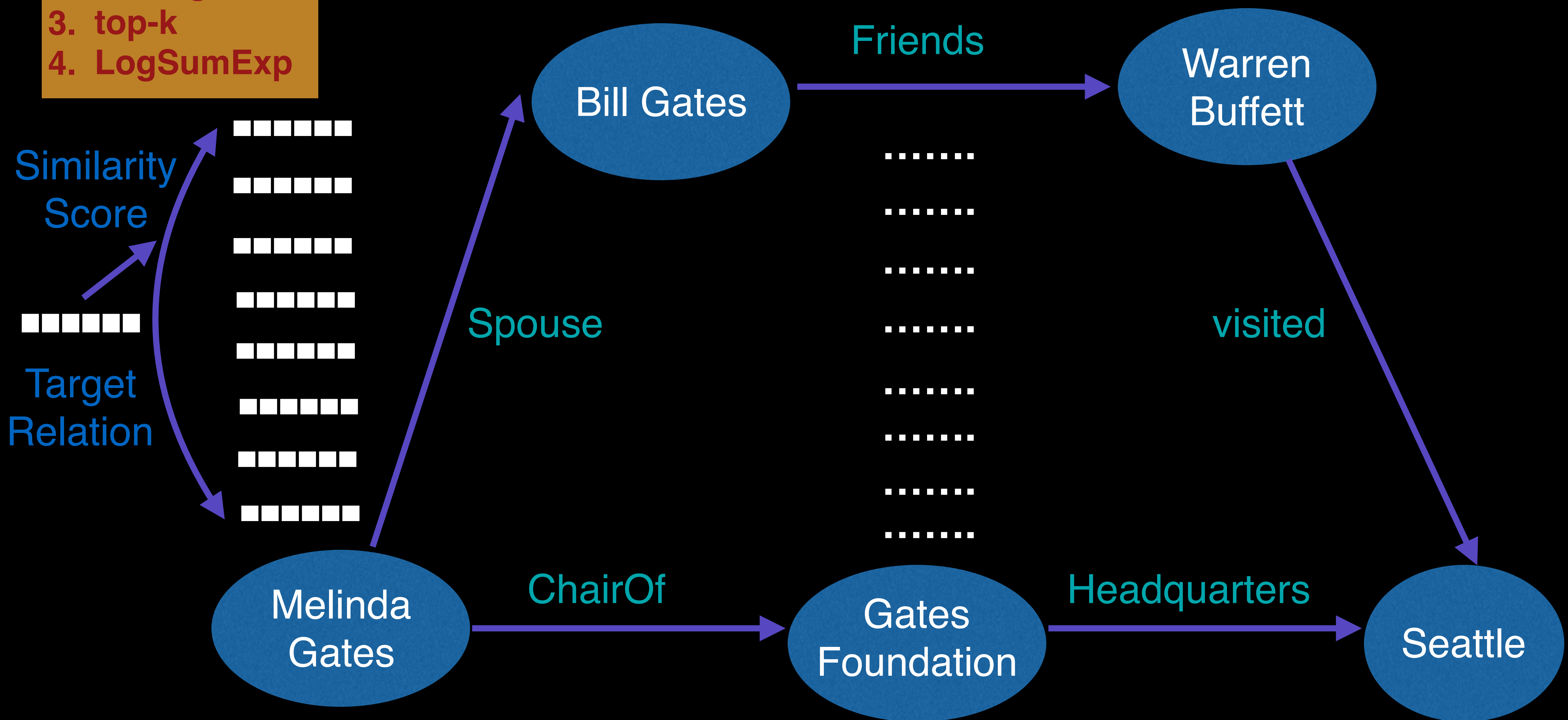
Results - Attention

Method	**Mean Average Precision**
Path Ranking Algorithm	64.4
Path Ranking Algorithm + bigram	64.9
RNN (max)	**65.2**

Selection/Attention

(Das, Neelakantan, Belanger, McCallum, 2016)

1. Max
2. Average
3. top-k
4. LogSumExp



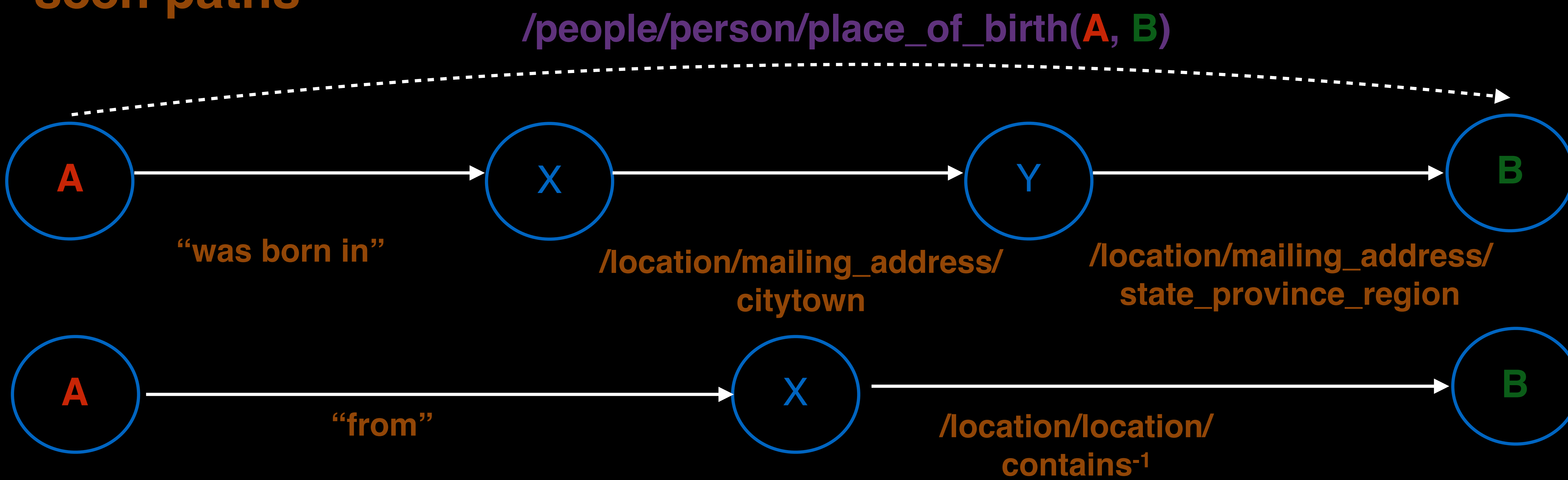
Train with backprop!

Results - Attention

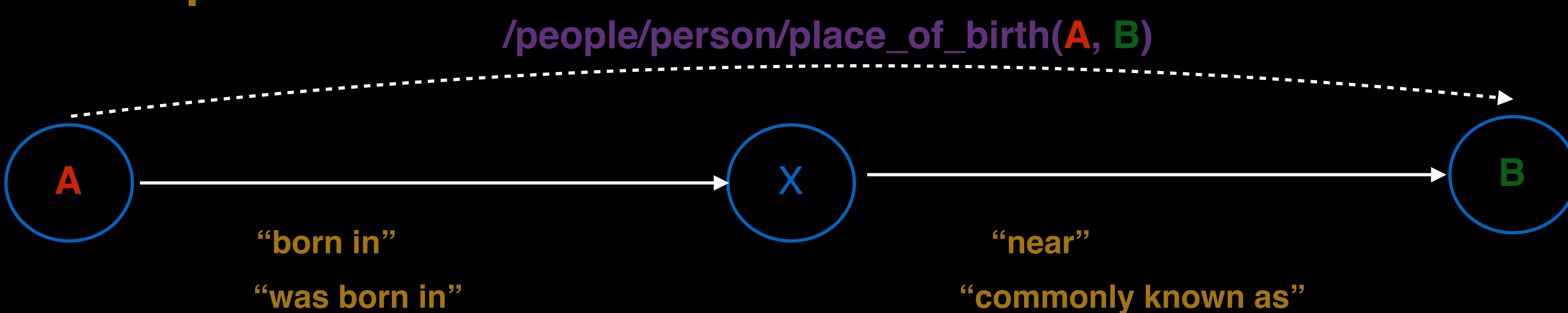
Method	Mean Average Precision
Path Ranking Algorithm	64.4
Path Ranking Algorithm + bigram	64.9
RNN (max)	**65.2**
RNN (avg)	**55.0**
RNN (top-k)	**68.2**
RNN (logsumexp)	**70.1**

Predictive Paths

seen paths



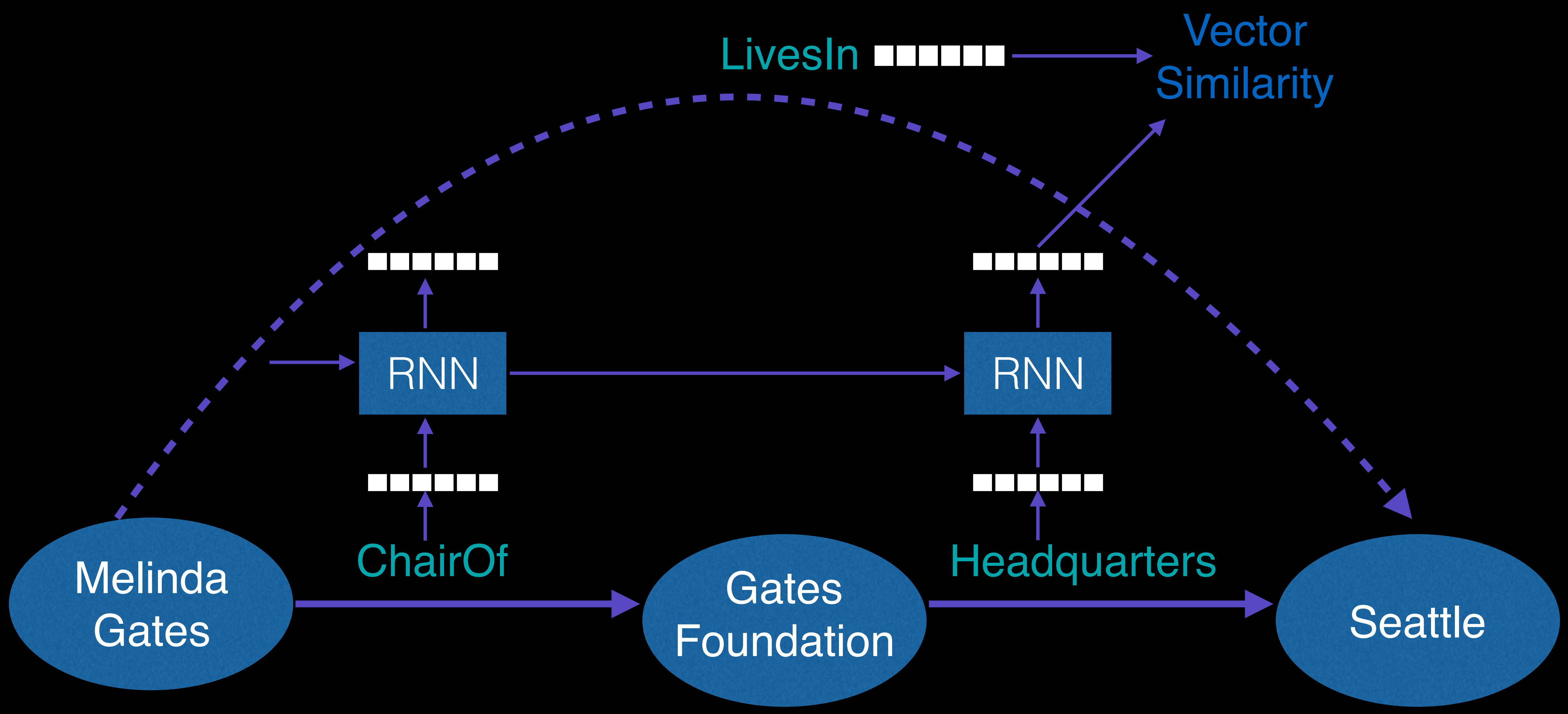
unseen paths



Multi-hop Reasoning: Current Methods do not generalize to unseen paths

Recurrent Neural Networks achieve state-of-the-art results on answering path queries

Zero-Shot



Predict relations not explicitly trained on!

Results

Method	Mean Average Precision
Random	7.6
RNN (zero-shot)	20.6
RNN (supervised)	50.1

Multi-hop Reasoning: Current Methods do not generalize to unseen paths

Recurrent Neural Networks achieve state-of-the-art results on answering path queries

RNNs can perform zero-shot learning!

Deep Neural Networks for Knowledge Representation and Reasoning

Recurrent Neural Networks achieve state-of-the-art results on answering knowledge graph path queries

Task 2:
Program Induction/Semantic Parsing

Arvind Neelakantan, Quoc V. Le, and Ilya Sutskever. Neural Programmer: Inducing latent programs with gradient descent. ICLR, 2016

Arvind Neelakantan, Quoc V Le, Martin Abadi, Andrew McCallum, and Dario Amodei. Learning a natural language interface with neural programmer. ICLR, 2017

Program Induction/Semantic Parsing

Date	Time	Opponent#	Rank#	Site	TV	Result	Attendance
September 3	6:00 PM	Missouri State*	#15	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	PPV	**W** 51–7	70,607
September 10	6:00 PM	New Mexico*	#14	War Memorial Stadium • Little Rock, AR	ESPNU	**W** 52–3	52,606
September 17	6:30 PM	Troy*	#14	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	CSS	**W** 38–28	69,861
September 24	2:30 PM	at #3 Alabama	#14	Bryant–Denny Stadium • Tuscaloosa, AL	CBS	**L** 14–38	101,821
October 1	11:00 AM	vs. #14 Texas A&M*	#18	Cowboys Stadium • Arlington, TX (Southwest Classic)	ESPN	**W** 42–38	69,838
October 8	6:00 PM	#15 Auburn	#10	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	ESPN	**W** 38–14	74,191

Program Induction/Semantic Parsing

Date	Time	Opponent#	Rank#	Site	TV	Result	Attendance
September 3	6:00 PM	Missouri State*	#15	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	PPV	**W** 51–7	70,607
September 10	6:00 PM	New Mexico*	#14	War Memorial Stadium • Little Rock, AR	ESPNU	**W** 52–3	52,606
September 17	6:30 PM	Troy*	#14	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	CSS	**W** 38–28	69,861
September 24	2:30 PM	at #3 Alabama	#14	Bryant–Denny Stadium • Tuscaloosa, AL	CBS	**L** 14–38	101,821
October 1	11:00 AM	vs. #14 Texas A&M*	#18	Cowboys Stadium • Arlington, TX (Southwest Classic)	ESPN	**W** 42–38	69,838
October 8	6:00 PM	#15 Auburn	#10	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	ESPN	**W** 38–14	74,191

Lookup Question: Which venue had the biggest turnout ?
Number Question: how many games were telecasted in CBS ?

Program Induction/Semantic Parsing

Date	Time	Opponent#	Rank#	Site	TV	Result	Attendance
September 3	6:00 PM	Missouri State*	#15	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	PPV	**W** 51–7	70,607
September 10	6:00 PM	New Mexico*	#14	War Memorial Stadium • Little Rock, AR	ESPNU	**W** 52–3	52,606
September 17	6:30 PM	Troy*	#14	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	CSS	**W** 38–28	69,861
September 24	2:30 PM	at #3 Alabama	#14	Bryant–Denny Stadium • Tuscaloosa, AL	CBS	**L** 14–38	101,821
October 1	11:00 AM	vs. #14 Texas A&M*	#18	Cowboys Stadium • Arlington, TX (Southwest Classic)	ESPN	**W** 42–38	69,838
October 8	6:00 PM	#15 Auburn	#10	Donald W. Reynolds Razorback Stadium • Fayetteville, AR	ESPN	**W** 38–14	74,191

Which venue had the biggest turnout ? => select site (max attendance)
how many games were telecasted in CBS ? => count(location == CBS)

Challenges

- Multi-step Reasoning

  Which section is the longest ? => select name (max kilometers)

- Weak Supervision

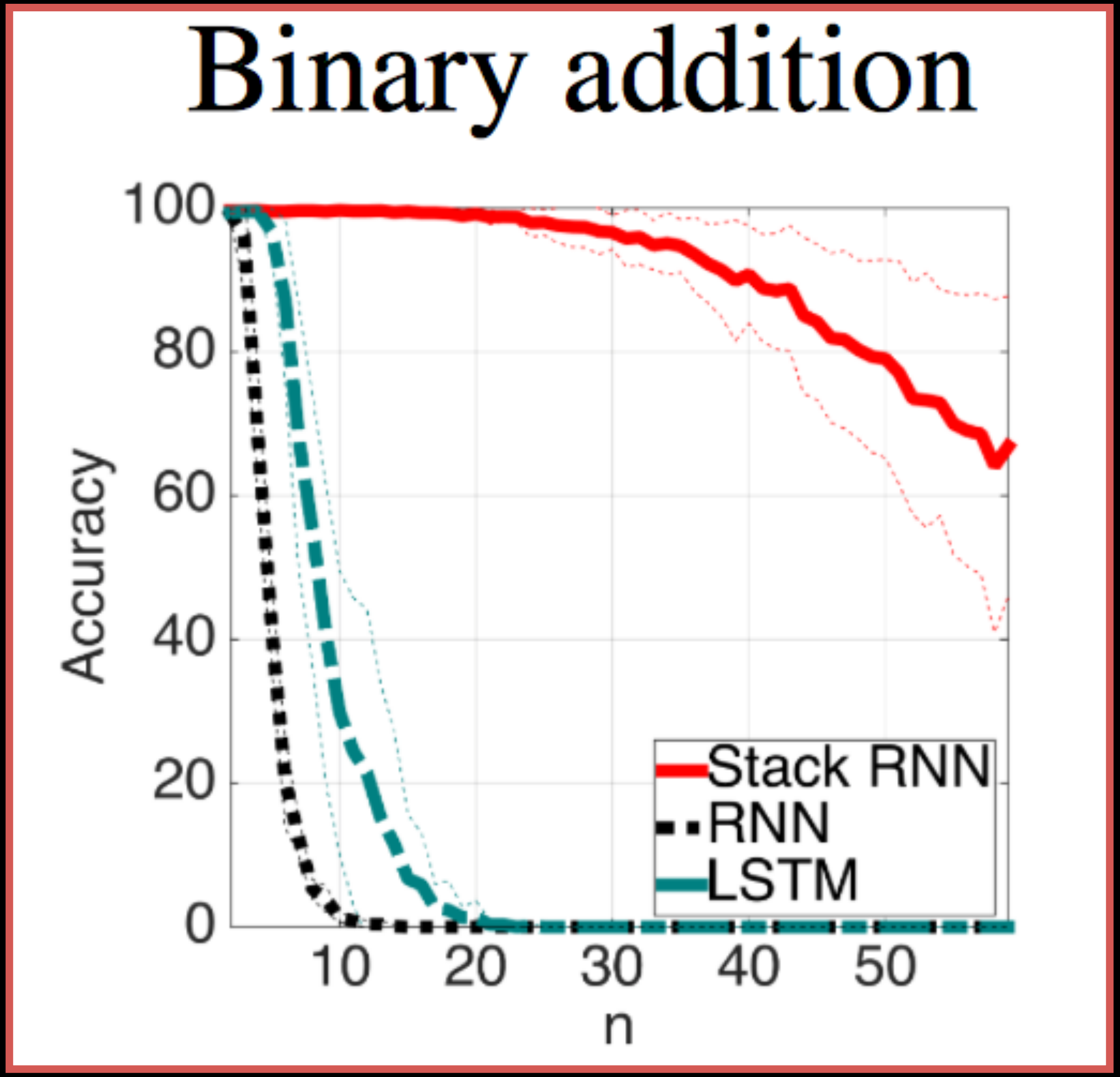
  Which section is the longest ? => ~~select name (max kilometers)~~
  => **IDF Checkpoint**

Motivation

	Strong Supervision	Weak Supervision (dataset specific rules to guide program search)
Non-Neural Network	Zelle & Mooney, (1996); Zettlemoyer & Collins, (2005)	Liang et al., (2011); Kwiatkowski et al., (2013); Pasupat & Liang, (2015)

End-to-End Neural Networks

Learning Discrete Functions is notoriously challenging!

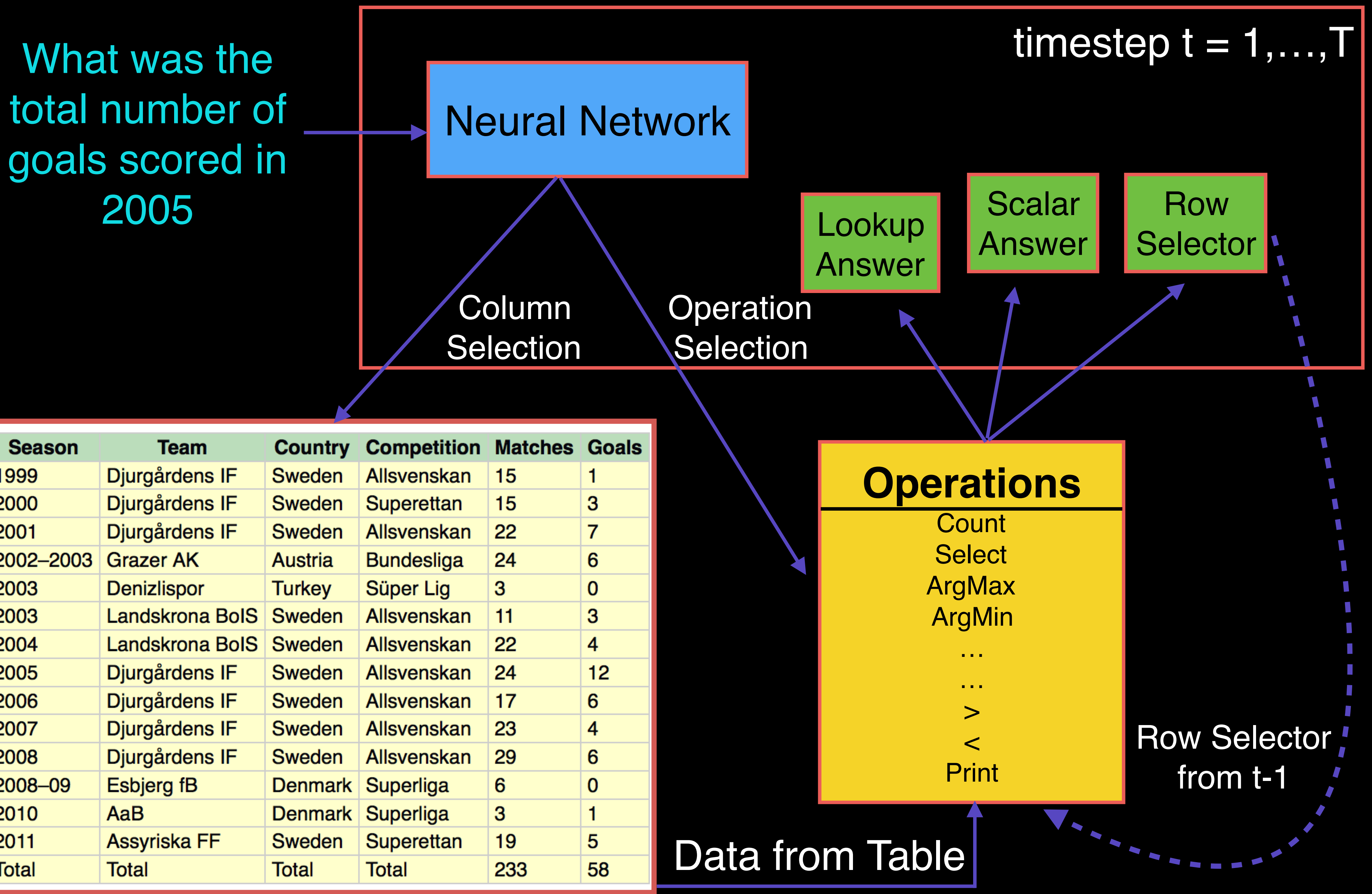


(Joulin & Mikolov, 2015)

Semantic Parsing: multi-step reasoning with discrete functions; weak supervision

Neural Programmer

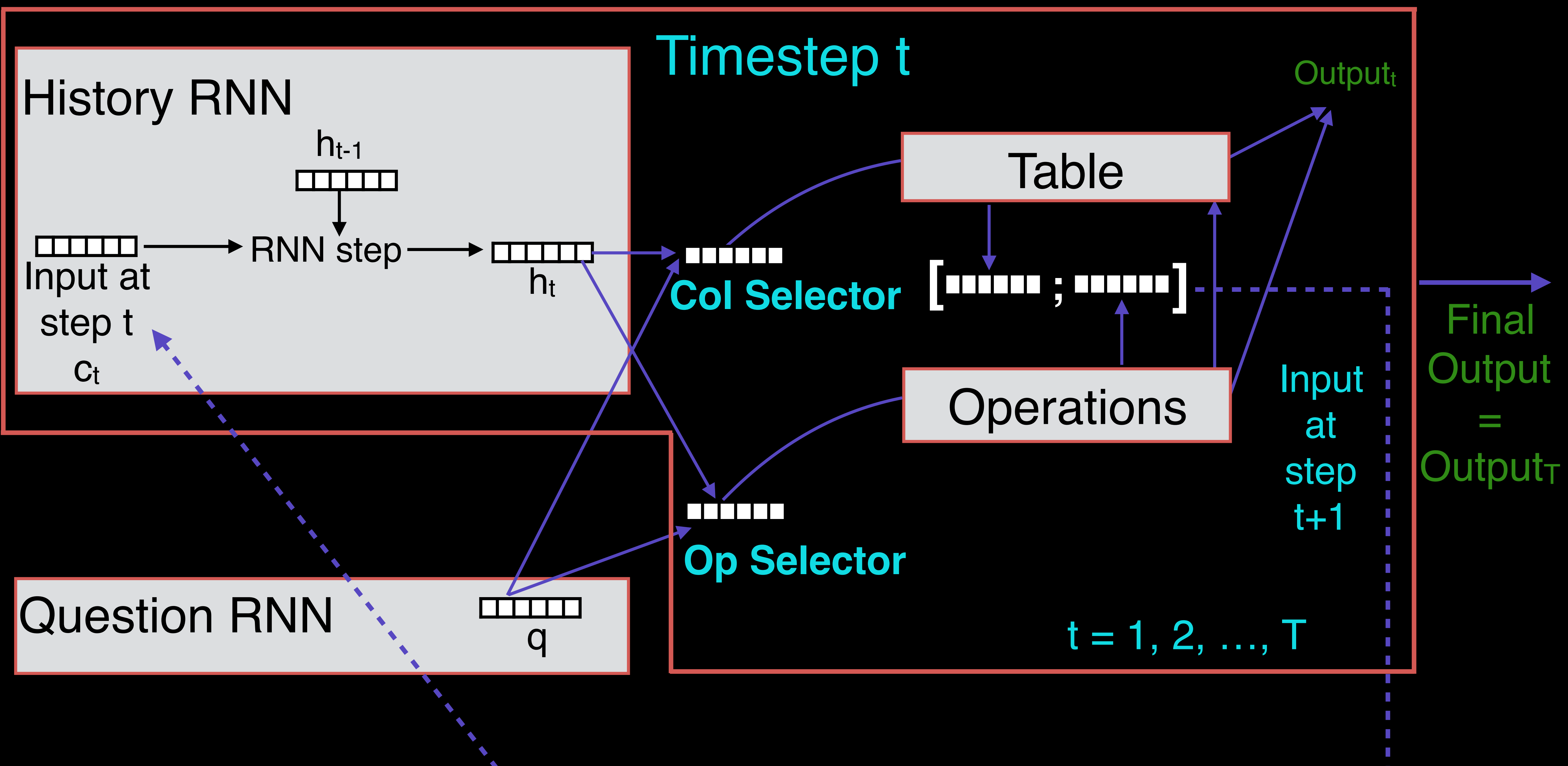
(Neelakantan, Le, Sustskever, ICLR 2016)



Season	Team	Country	Competition	Matches	Goals
1999	Djurgårdens IF	Sweden	Allsvenskan	15	1
2000	Djurgårdens IF	Sweden	Superettan	15	3
2001	Djurgårdens IF	Sweden	Allsvenskan	22	7
2002–2003	Grazer AK	Austria	Bundesliga	24	6
2003	Denizlispor	Turkey	Süper Lig	3	0
2003	Landskrona BoIS	Sweden	Allsvenskan	11	3
2004	Landskrona BoIS	Sweden	Allsvenskan	22	4
2005	Djurgårdens IF	Sweden	Allsvenskan	24	12
2006	Djurgårdens IF	Sweden	Allsvenskan	17	6
2007	Djurgårdens IF	Sweden	Allsvenskan	23	4
2008	Djurgårdens IF	Sweden	Allsvenskan	29	6
2008–09	Esbjerg fB	Denmark	Superliga	6	0
2010	AaB	Denmark	Superliga	3	1
2011	Assyriska FF	Sweden	Superettan	19	5
Total	Total	Total	Total	233	58

Table

Neural Programmer



Output: Scalar Answer, Lookup Answer, Row Selector

Operations

- Row Selector: vector with size equal to number of rows
    - Comparison: >, <, >=, <=
    - Superlative: argmax, argmin
    - Table Ops: select, first, last, prev, next, group_by_max
    - Reset/No-Op
- Scalar Answer: real number
    - Aggregation: count
- Lookup Answer: matrix with same dimension as table
    - Print

Example

Question		Step 1	Step 2	Step 3	Step 4
What was the total number of goals scored in 2005	Operation	No-Op	No-Op	select	print
	Column	-	-	season	goals

Weak Supervision

Question		Step 1	Step 2	Step 3	Step 4
What was the total number of goals scored in 2005	Operation Column	No-Op -	No-Op -	select season	print goals

Final Answer: 12

Soft Selection/Attention

(Bahdanau, Cho, Bengio, 2014)

- Average outputs of the different operations weighted by the probabilities from the model
- Train with backprop!

Soft Selection/Attention

		0.7	0.3
		Column A	**Column B**
0.6	**Operation A**	10	-5
0.4	**Operation B**	100	50

Output

0.6 x 0.7 x 10 +

0.6 x 0.3 x -5 +

0.4 x 0.7 x 100 +

0.4 x 0.3 x 50 +

Training Objective

- Final Answer
  - Number Answer: Square Loss
  - Lookup Answer: Average of loss on each entry
- Answer simply written down introduces ambiguity
  - Number could be generated or a table entry
  - Multiple table entries match the answer
  - Minimum of individual losses

Semantic Parsing: multi-step reasoning with discrete functions; weak supervision

Neural Programmer can be trained end-to-end with backprogragation using weak supervision

Previous Work

	Strong Supervision	Weak Supervision
Non-Neural Network	Zelle & Mooney, (1996); Zettlemoyer & Collins, (2005)	Liang et al., (2011); Kwiatkowski et al., (2013); Pasupat & Liang, (2015)
Neural Network	Jia & Liang, (2016); Neural Programmer Interpreter (Reed & De Freitas, 2015); Neural Enquirer (Yin et al., 2016)	Dynamic Neural Module Network (Andreas et al., 2016) not end-to-end

Experiments

- WikiTablesQuestions dataset (Pasupat & Liang, 2015)
- Database at test time are ***unseen*** during training
- 10k training examples with weak supervision
- Hard selection at test time
- 4 timesteps and 15 operations

Neural Networks

- Seq2Seq (Sutskever, Vinyals & Le, 2014)

  8.9% accuracy

- Pointer Networks (Vinyals, Fortunato & Jaitly, 2015)

  4.0% accuracy on lookup questions

Results

(Neelakantan, Le, Abadi, McCallum, Amodei, 2017)

Method	Dev Accuracy	Test Accuracy
Information Retrieval System	13.4	12.7
Simple Semantic Parser	23.6	24.3
Semantic Parser (Pasupat & Liang, 2015)	**37.0**	**37.1**
Neural Programmer - {dropout, weight decay }	30.3	-
Neural Programmer	34.2	34.2
Ensemble of 15 Neural Programmers	**37.5**	**37.7**

Training Data Size

	Textual Entailment	Textual Entailment	Reading Comprehension
# Training Examples	4.5k	550k	86k
Non-Neural Network	77.8	78.2	51.0
Neural Network	71.3	88.3	82.9

Semantic Parsing: multi-step reasoning with discrete functions; weak supervision

Neural Programmer can be trained end-to-end with backprogragation using weak supervision

Neural Programmer works surprisingly well on a small real-world dataset

Example Programs (1)

Question		Step 1	Step 2	Step 3	Step 4
What is the total number of teams ?	Operation	-	-	-	count
	Column	-	-	-	-
how many games had greater than 1500 in attendance ?	Operation	-	-	>=	count
	Column	-	-	attendance	-
what is the total number of runner-ups listed on the chart?	Operation	-	-	select	count
	Column	-	-	outcome	-

Example Programs (2)

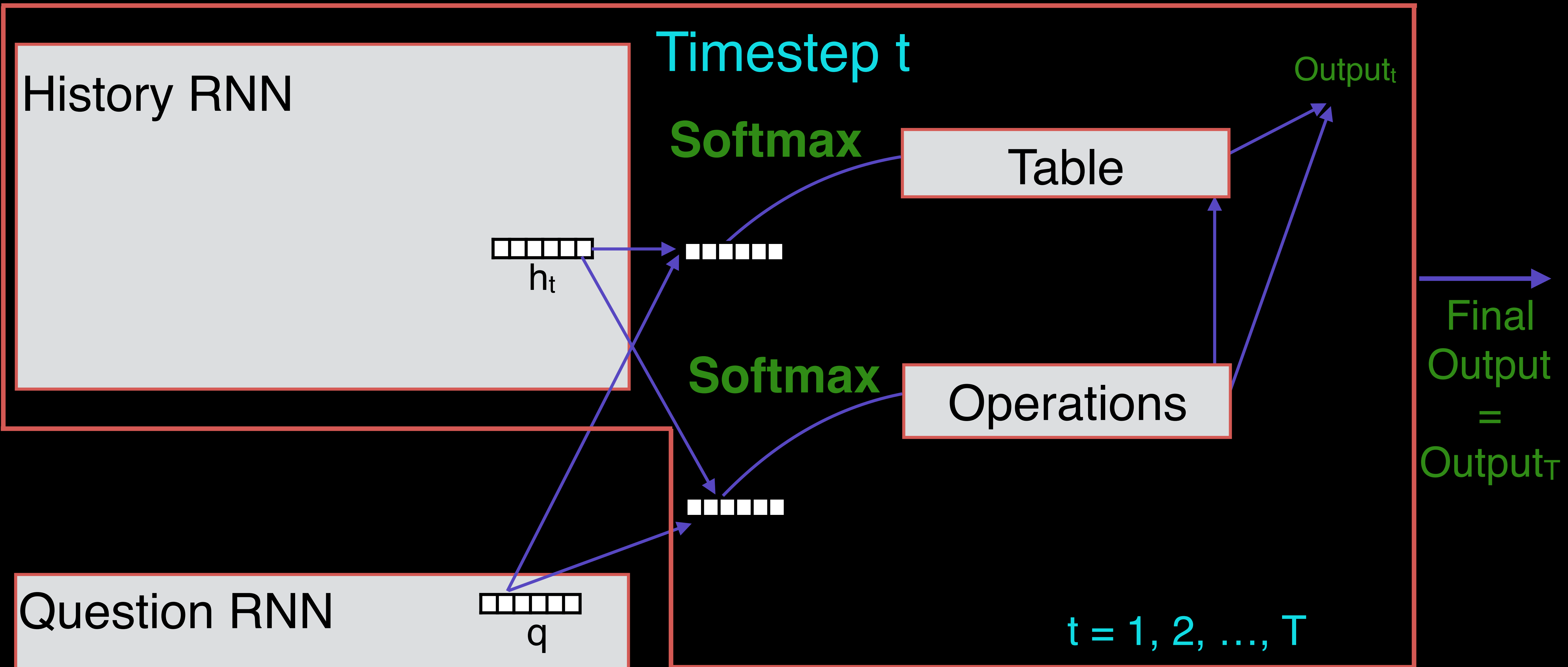
Question		Step 1	Step 2	Step 3	Step 4
which section is longest??	Operation	-	-	argmax	print
	Column	-	-	kilometers	name
Which engine(s) has the least amount of power ?	Operation	-	-	argmin	print
	Column	-	-	power	engine
Who had more silver medals, cuba or brazil ?	Operation	argmax	select	argmax	print
	Column	nation	nation	silver	nation

Example Programs (3)

Question		Step 1	Step 2	Step 3	Step 4
who was the next appointed director after lee p. brown?	Operation	select	next	last	print
	Column	name	-	-	name
what team is listed previous to belgium?	Operation	select	previous	first	print
	Column	team	-	-	team

Future Work

Neural Programmer



Scaling Up
Neural Programmer Training

- Soft selection requires every operation to be executed at every step

For every example: $O((MN)^T)$

M: # operations, N: # columns, T: # timesteps

- Fixed number of timesteps

Sparse Selection: top-k

- Top-k elements in the softmax (Rae et al., 2016)
- k operations and columns are selected
- k << M, N. Fast training!

For every example: $O((k^2)^T)$

Sparse Selection: Discrete

- Discrete selection
- One operation and one column are selected
- Train with Policy Gradient (Williams, 1992)

Reinforcement Learning

- Agent selects an action at every timestep
- Agent receives reward signals
- Lot of excitement recently
- Very high sample complexity- Shown to work well mostly in games and simulations

Learning To Halt

- Currently predefined maximum number of steps for every input
- Learning to halt to speed up the model
- Discrete Neuron to decide whether to continue (Graves, 2016)

Example

Question		Step 1	Step 2	Step 3	Step 4
What was the total number of goals scored in 2005	Operation Column	No-Op -	No-Op -	select season	print goals

Example

Question		Step 1	Step 2
What was the total number of goals scored in 2005	Operation Column	select season	print goals

Continue

Stop

Challenges and Experiments

- High variance in gradient estimation
- Experiments with Neural Programmer:
    1. Sparse Selection
    2. Learning to Halt
- Accuracy Vs Training Time

Summary

Deep Neural Networks for Knowledge Representation and Reasoning

Recurrent Neural Networks achieve state-of-the-art results on answering knowledge graph path queries

Neural Programmer achieves competitive results on a small real-world question answering dataset

Key Components

- Recurrent Neural Networks
- Attention/Selection Mechanism
- Backpropagation

Deep Neural Networks for Knowle ge
Representation and Reaso

Recurrent Neu tworks
achieve state e-art results on
answering kn ge graph path queries

Neural Programmer
ves competitive results on a small
eal-world question answering dataset

Code and Data are publicly available!

Keep Calm
and
Believe in Backprop!

Acknowledgements: IESL & Google Brain Team!

Thank You!