
TECHNISCHE
UNIVERSITÄT
DARMSTADT
UKP Lab
Darmstadt, Germany

Graph Neural Networks for Knowledge Base Question Answering

Daniil Sorokin
@daniilmagpie 👋
https://daniilsorokin.github.io

UKP

South England NLP Meetup
January 21, 2019

Me

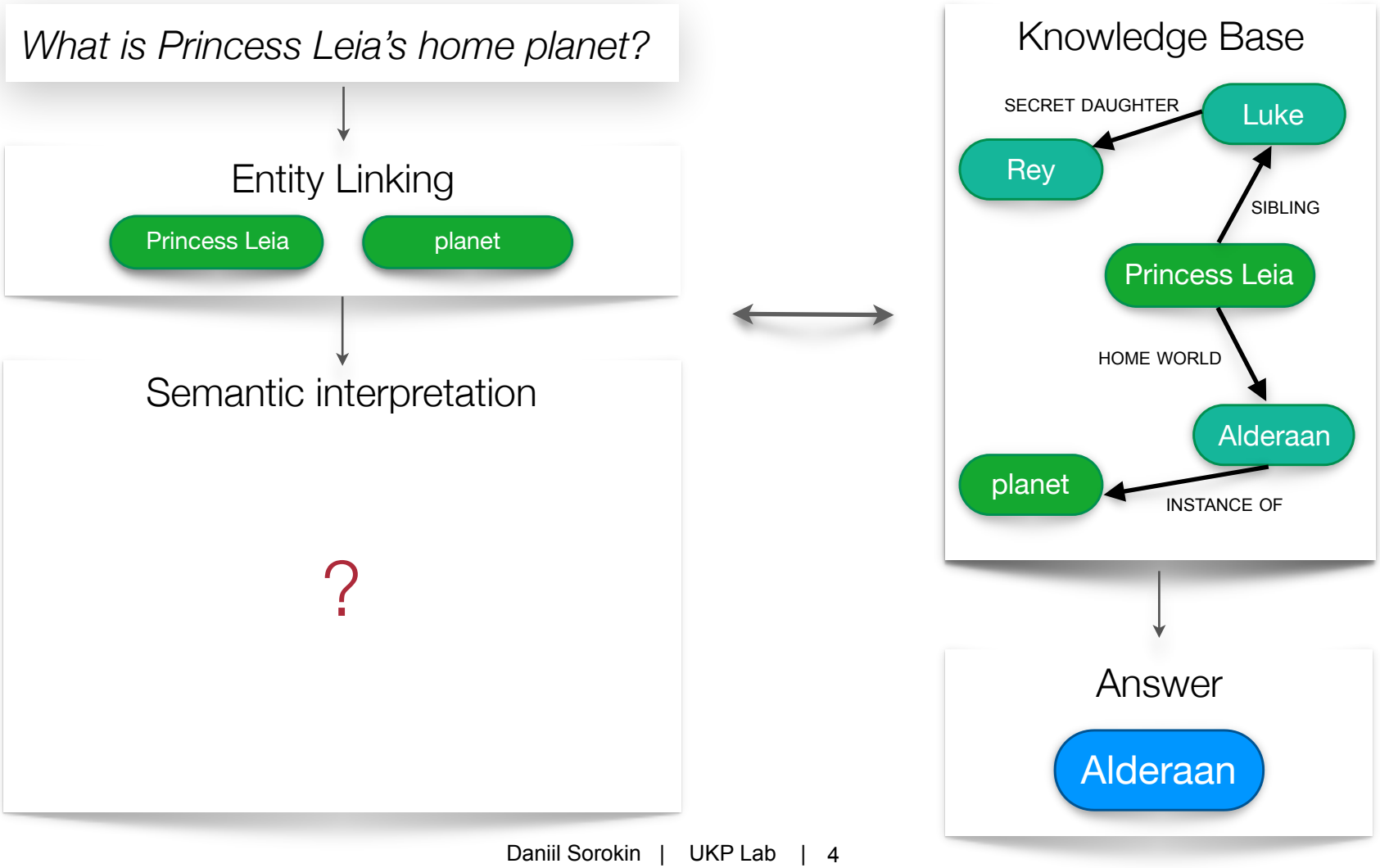
Bachelor in Applied Linguistics
at St. Petersburg State University, Russia
2011

Master in Computational Linguistics
at University of Tübingen, Germany, 2014

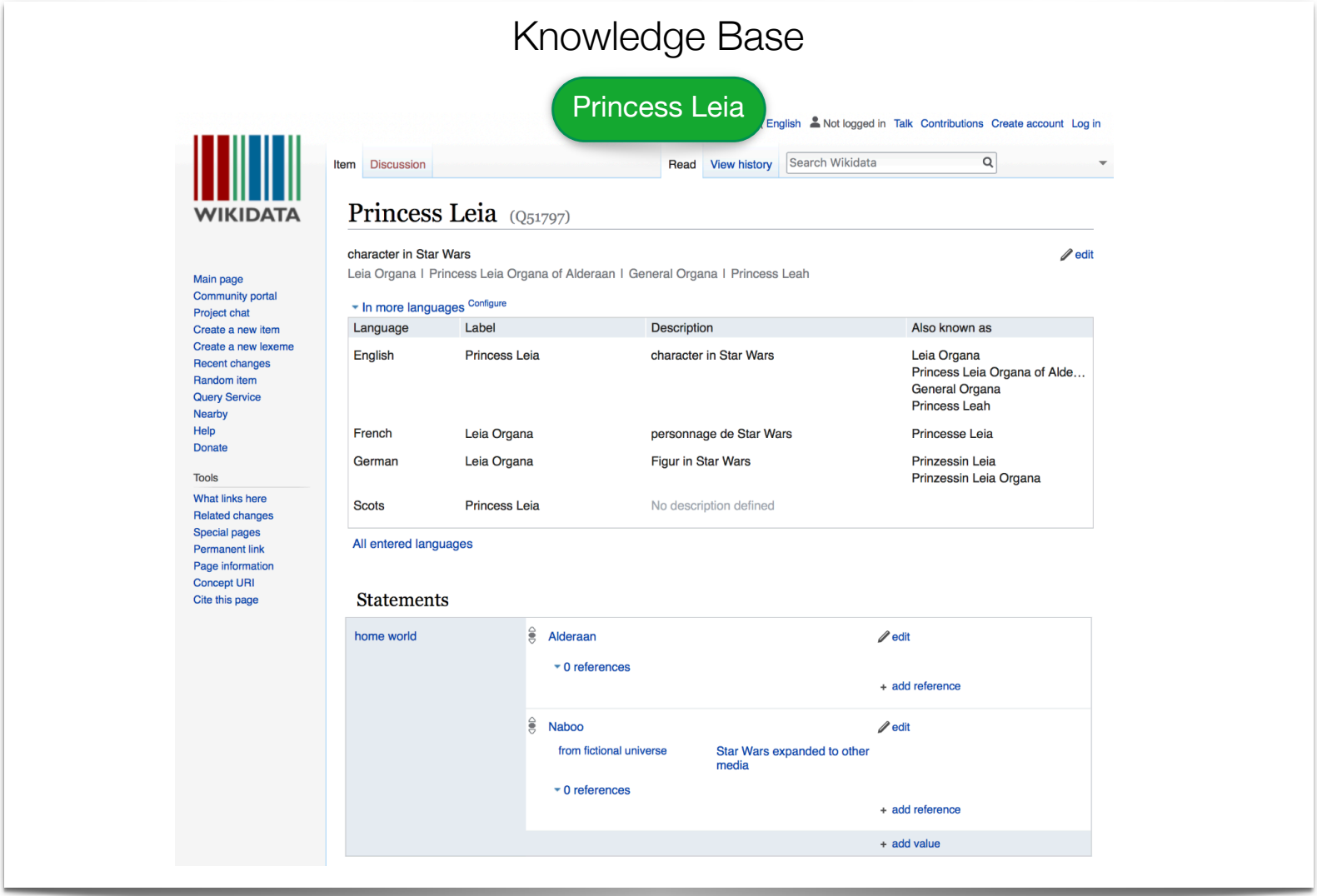
PhD in NLP
at the UKP Lab,
Darmstadt University of Technology, Germany,
since 2015,
http://ukp.tu-darmstadt.de

Research intern at
Factmata, London
November 2018 - January 2019,
https://factmata.com

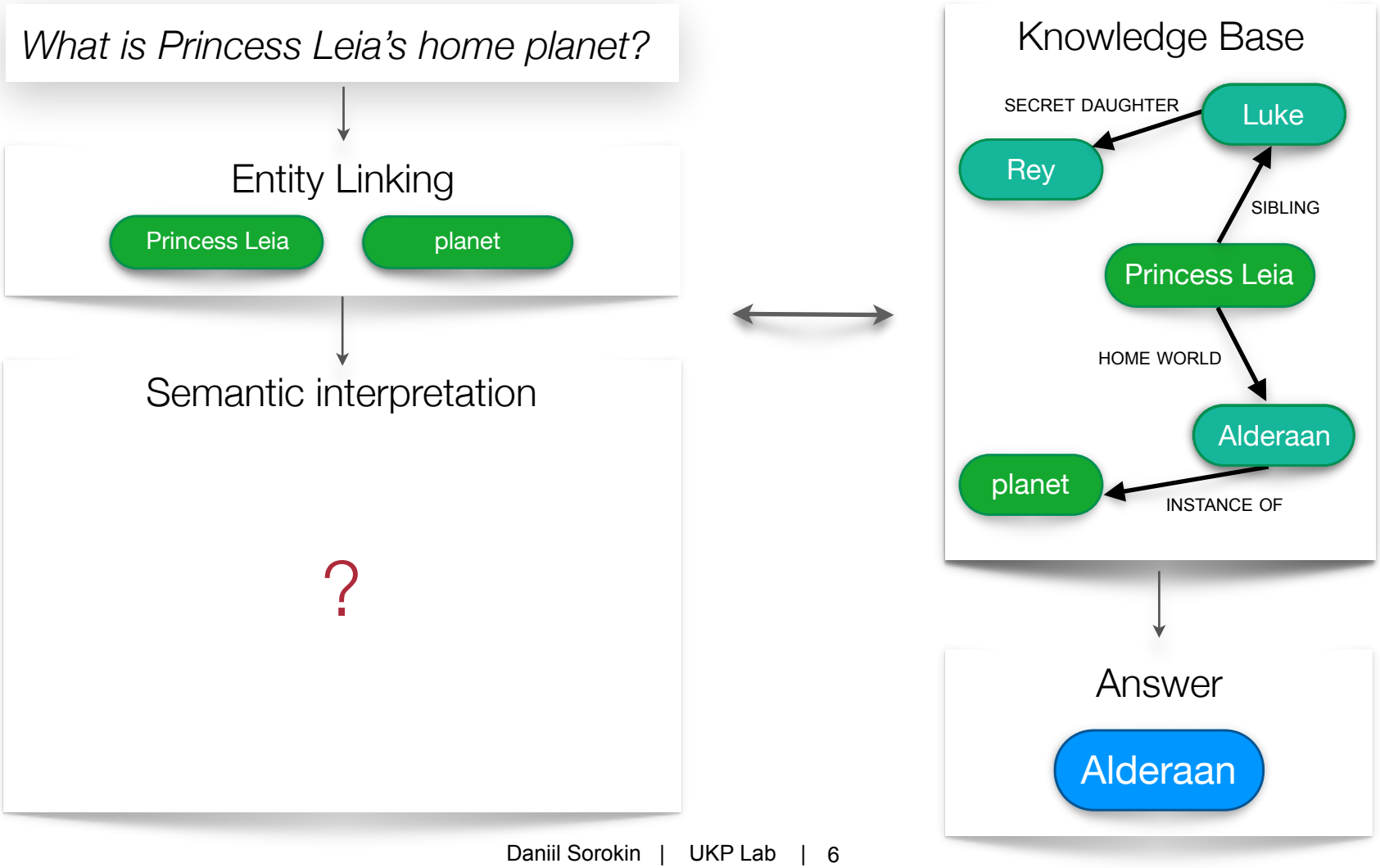
Knowledge Base Question Answering



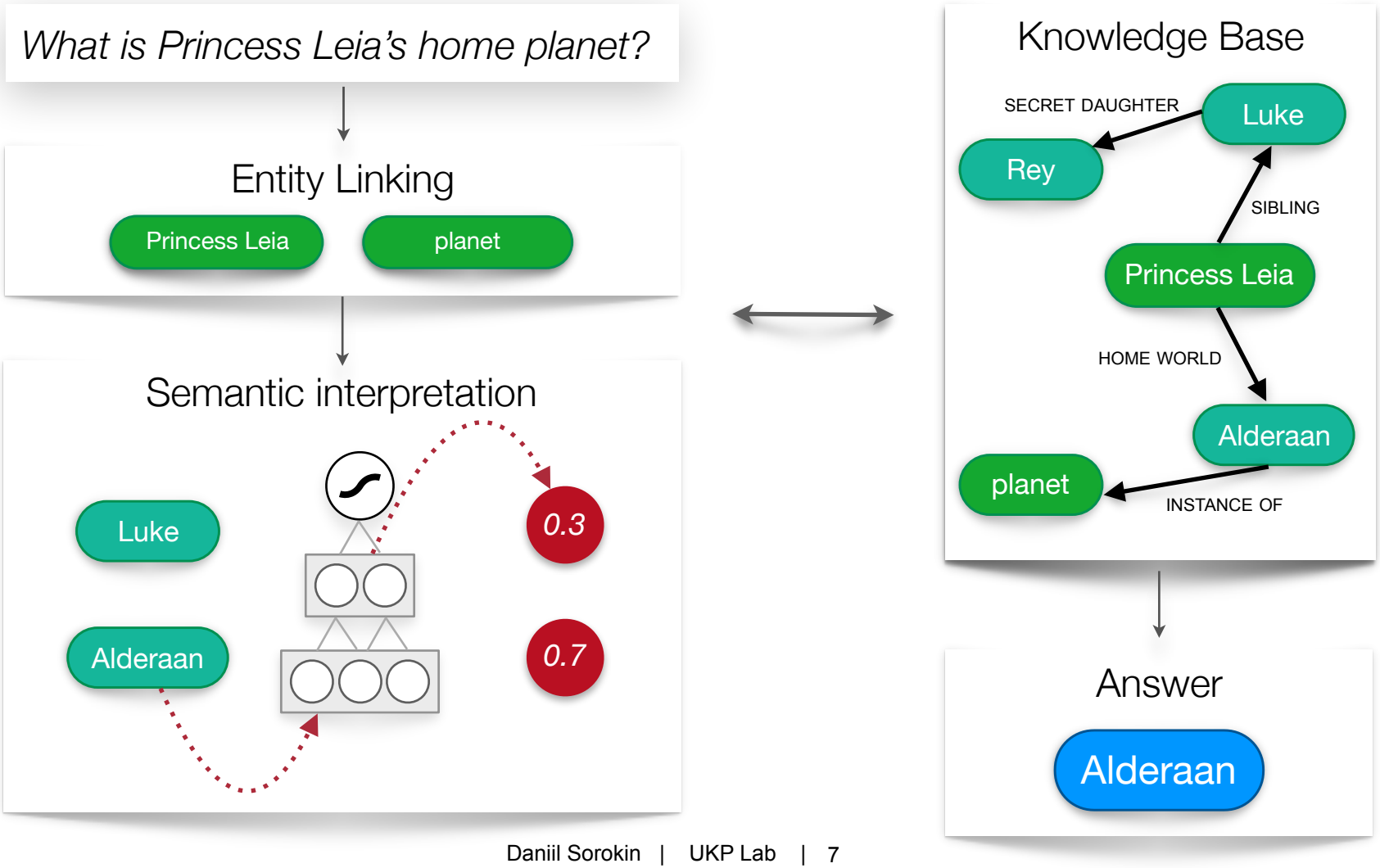
Knowledge Base



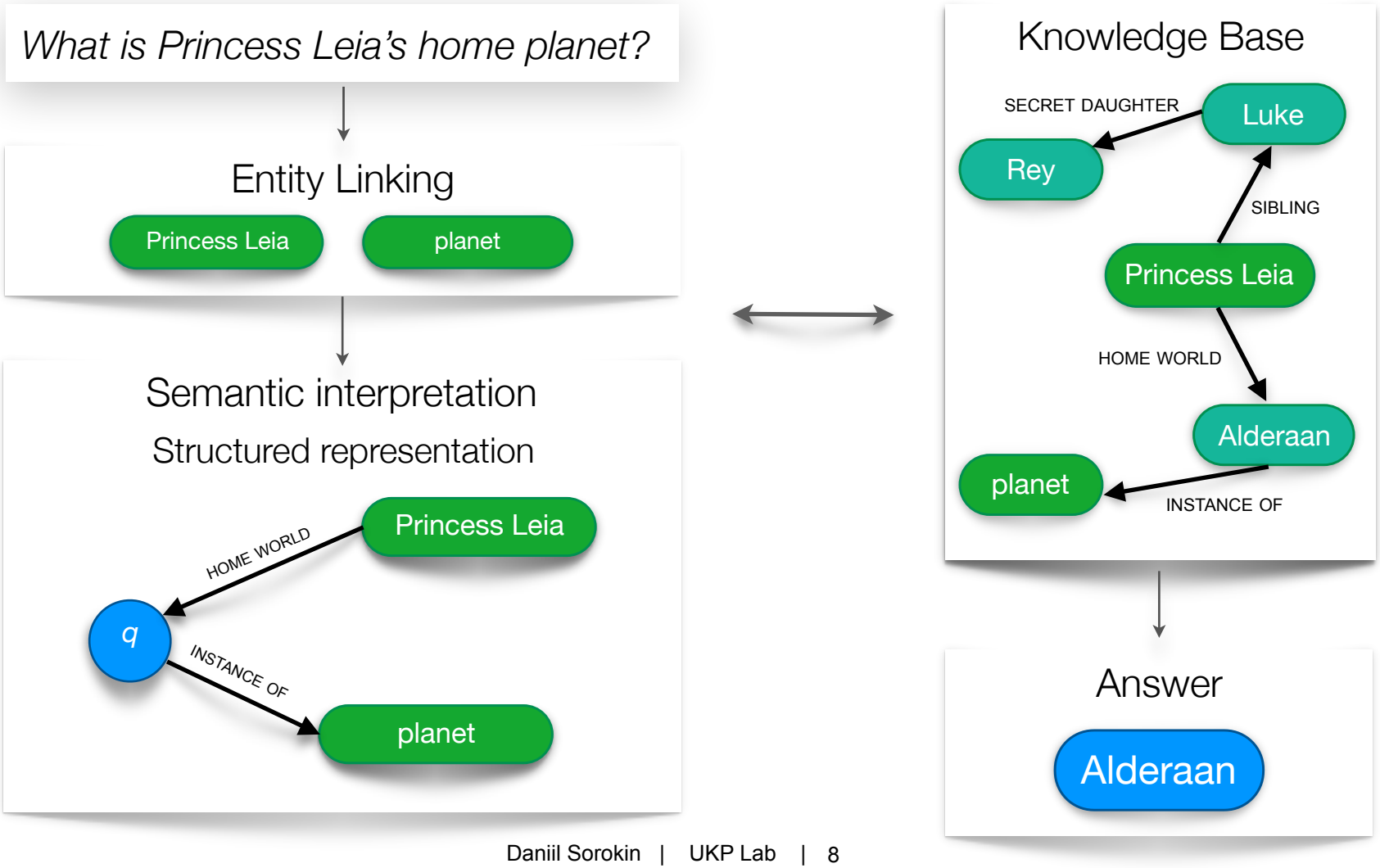
Knowledge Base Question Answering



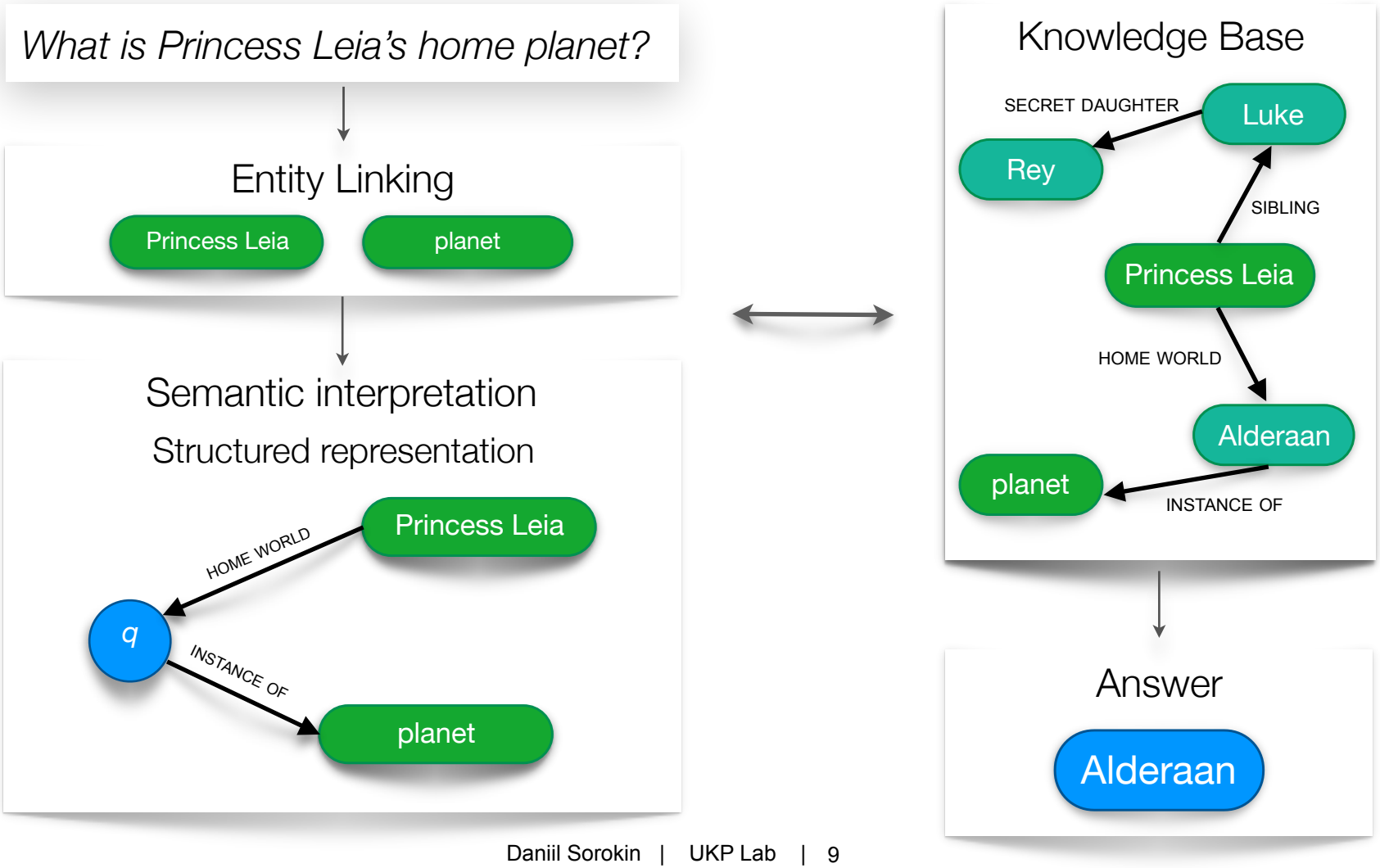
Knowledge Base Question Answering



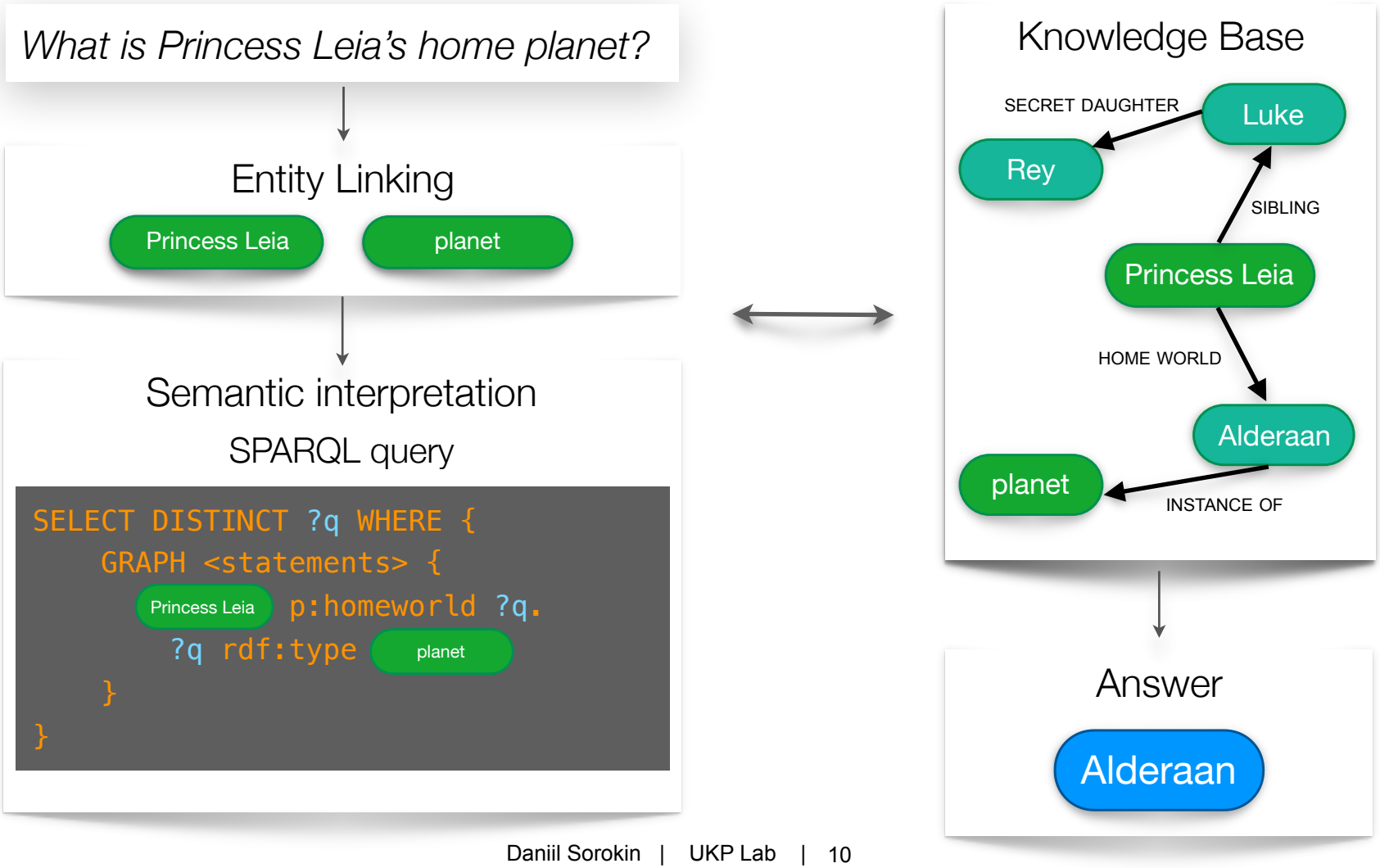
Knowledge Base Question Answering



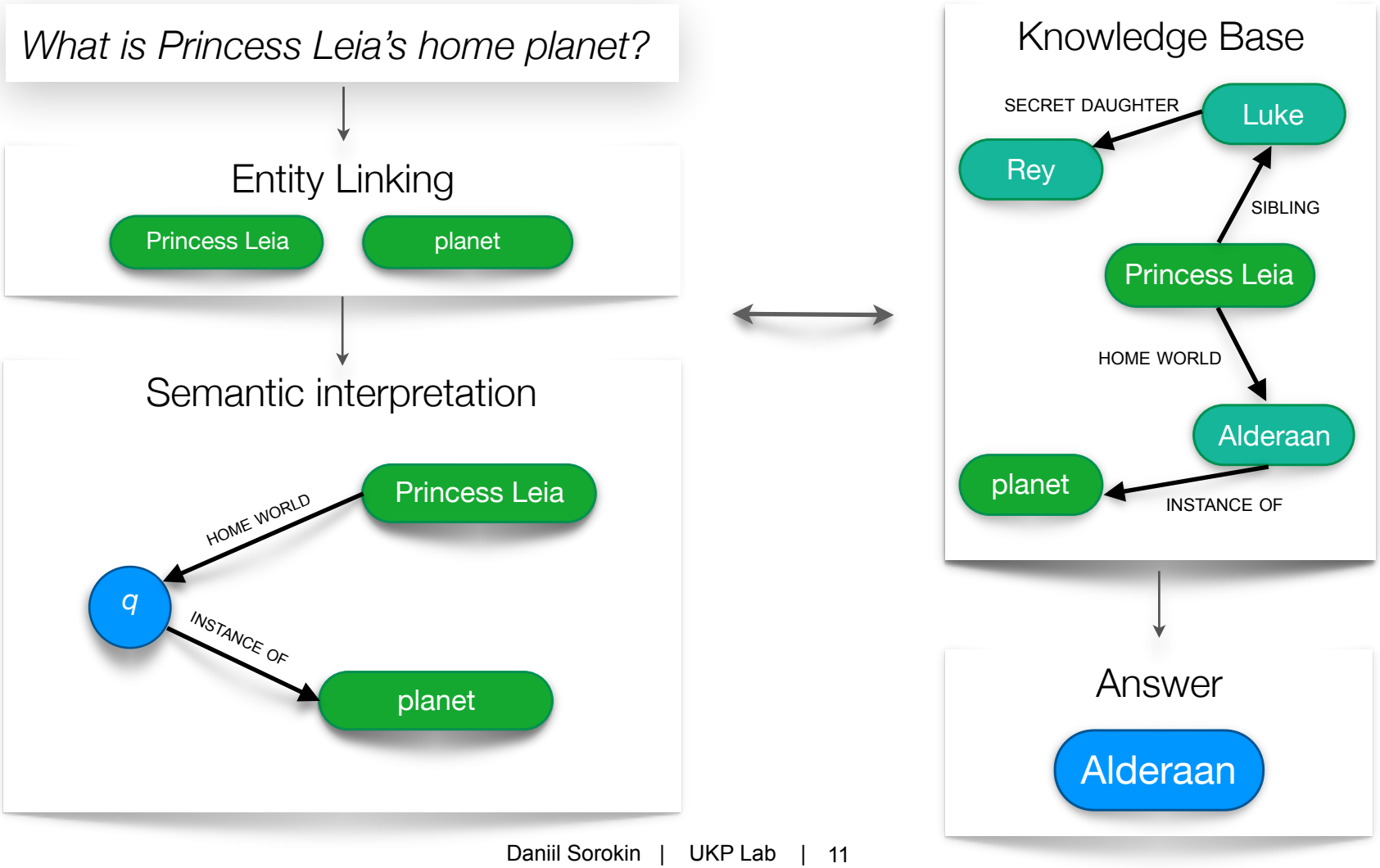
Knowledge Base Question Answering



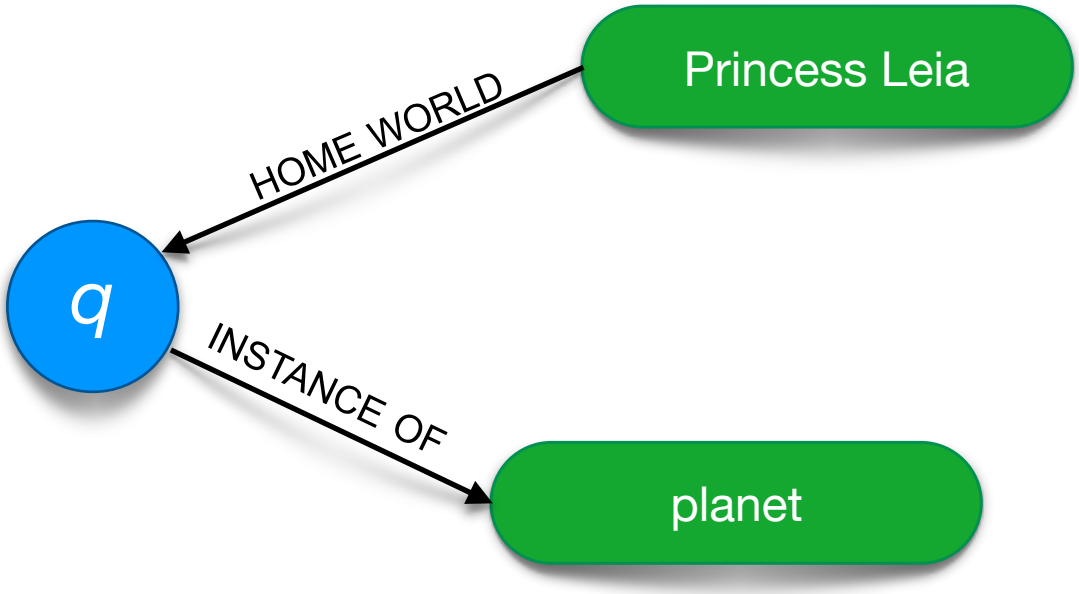
Knowledge Base Question Answering



Knowledge Base Question Answering



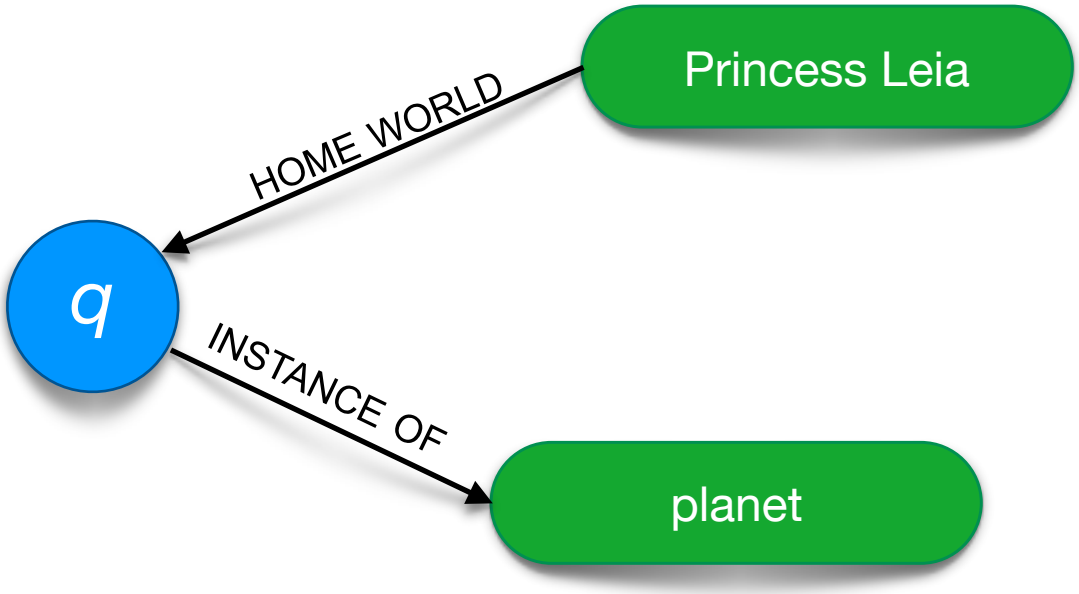
Structured Representation Approach



What is Princess Leia's home planet?

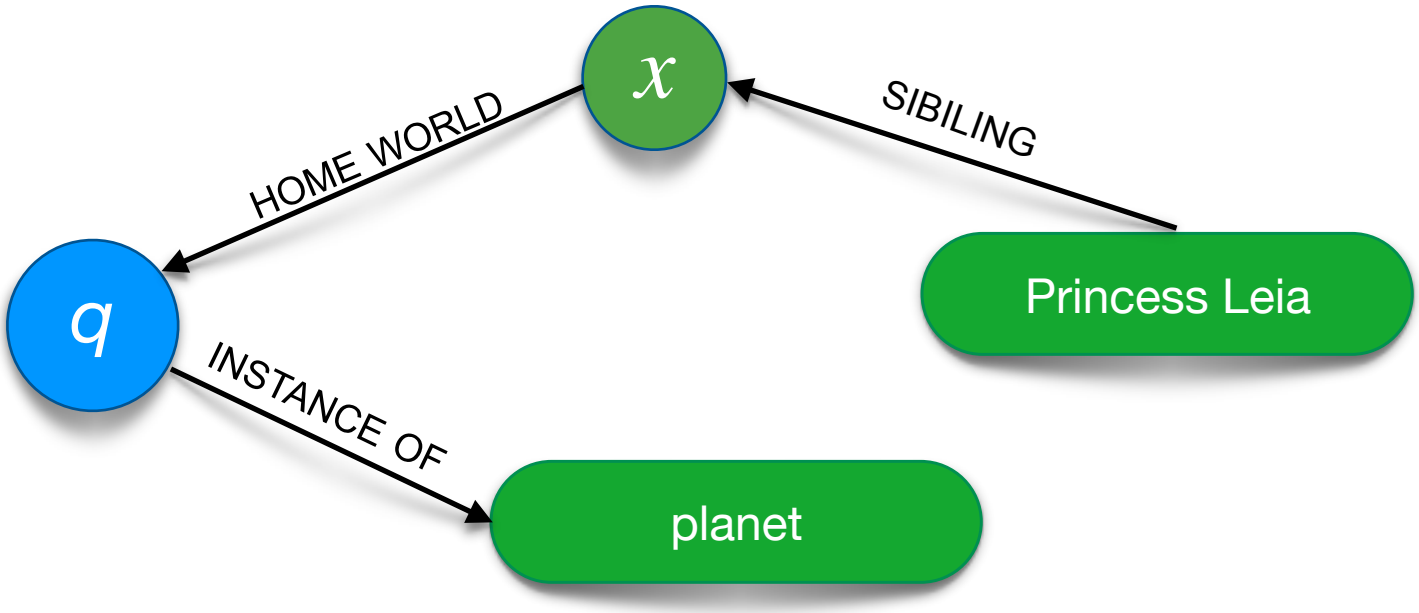
Possible Semantic Graphs

Yih et al. (2015)



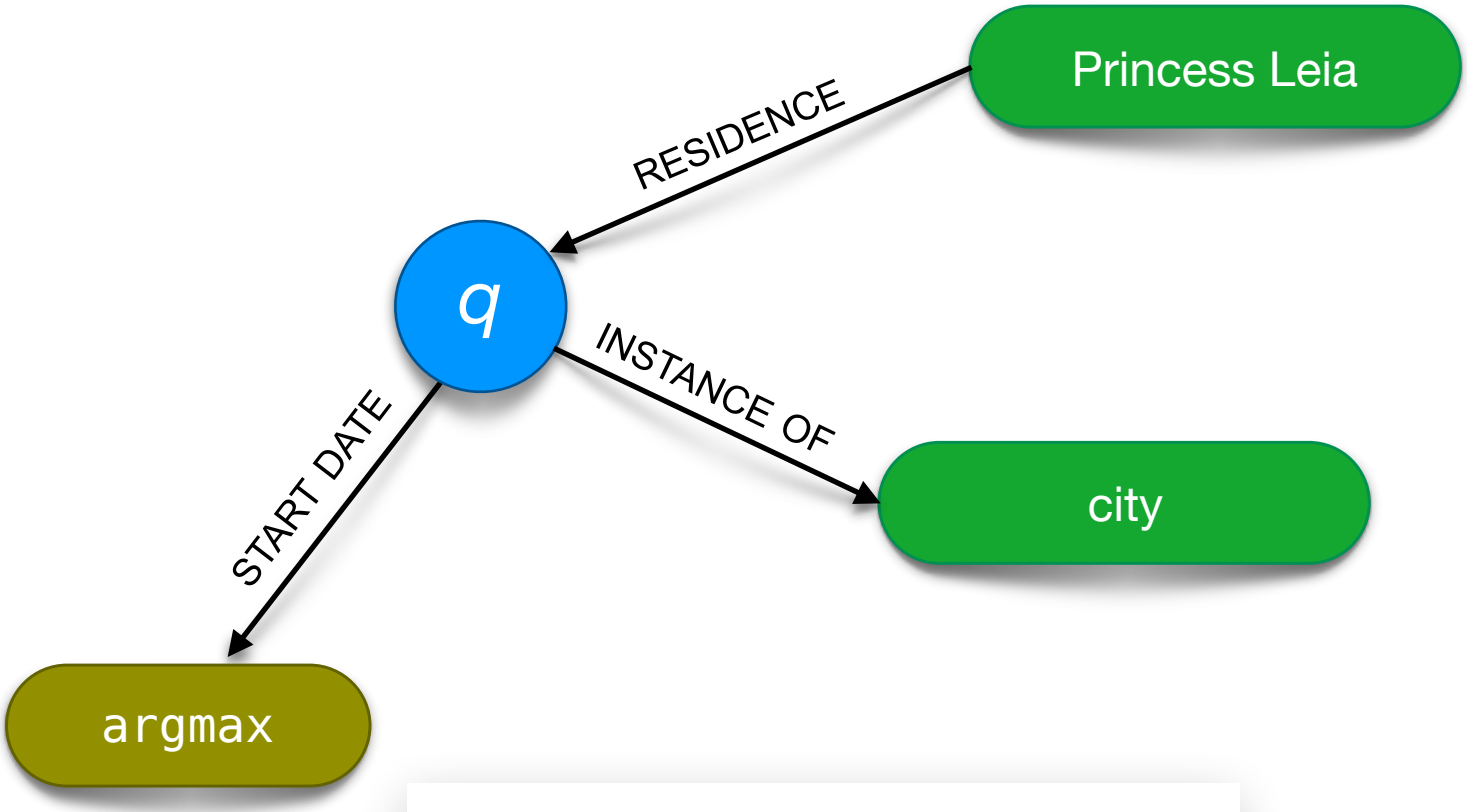
What is Princess Leia's home planet?

Possible Semantic Graphs



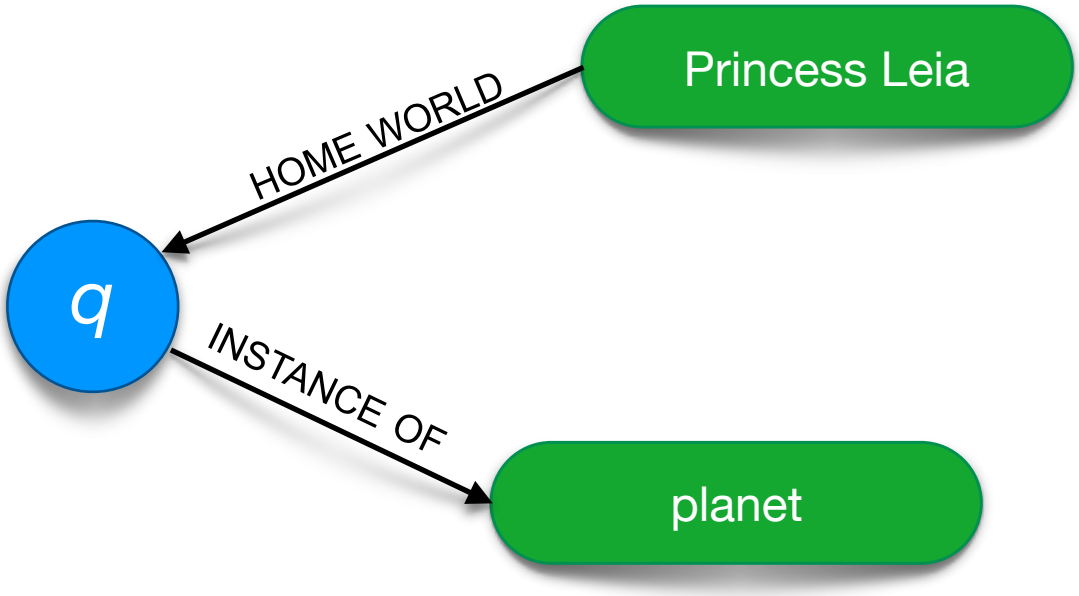
What is the home planet of Princess Leia's family?

Possible Semantic Graphs



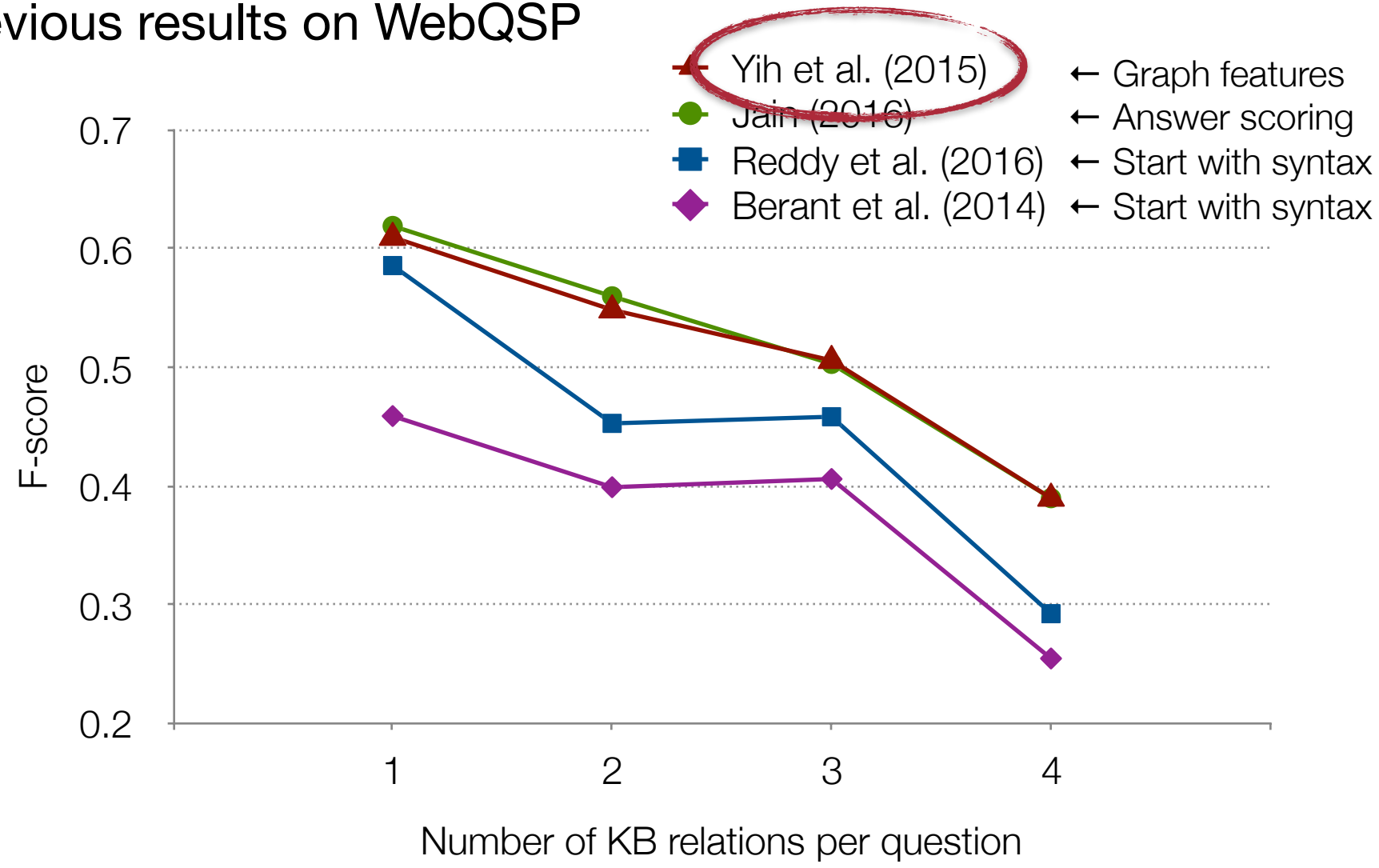
Where was Princess Leia's first home?

Possible Semantic Graphs



Representation structure is a graph and we model it as such

Previous results on WebQSP



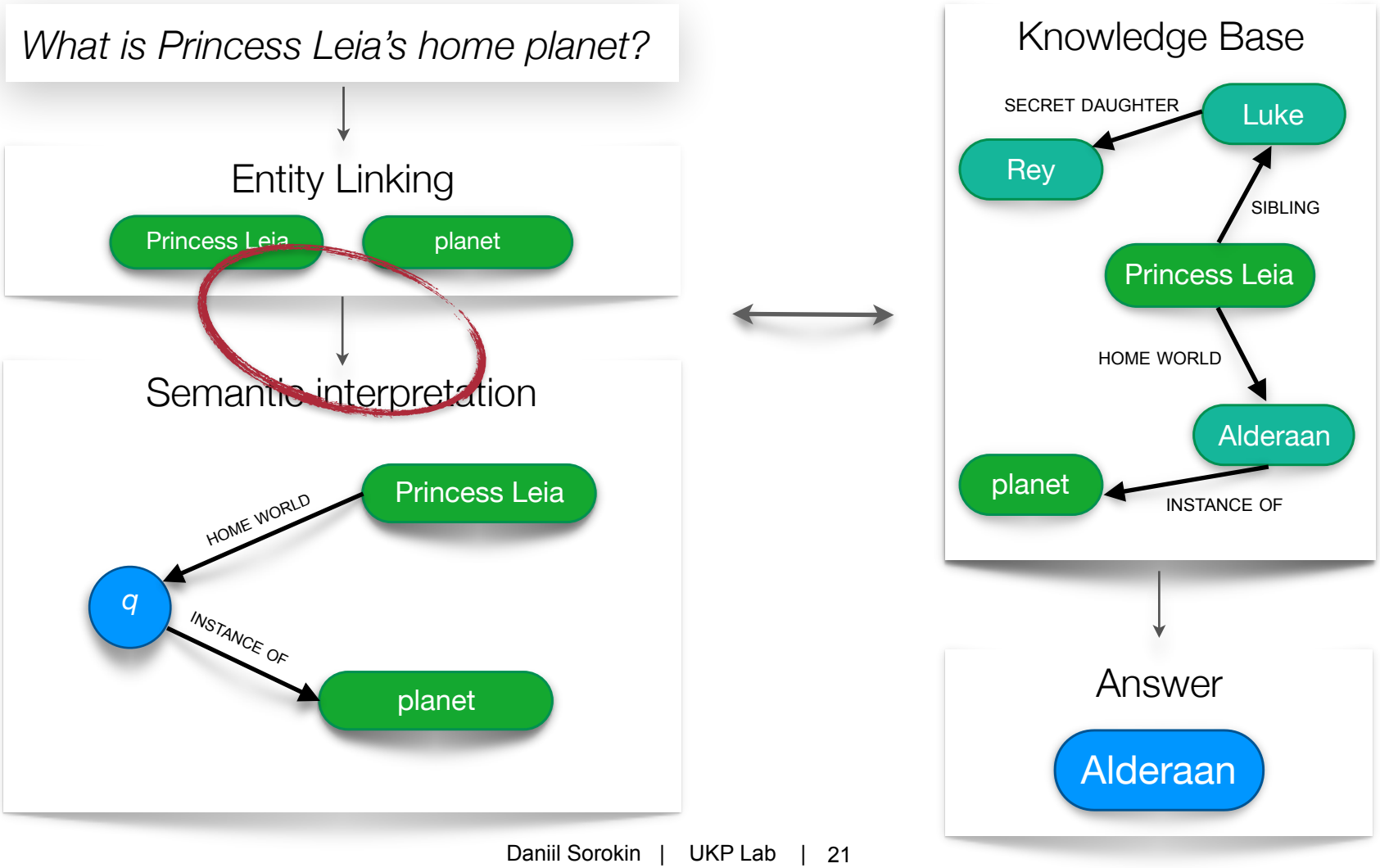
Freebase

Previous approaches do not integrate the structure of the representation into the model

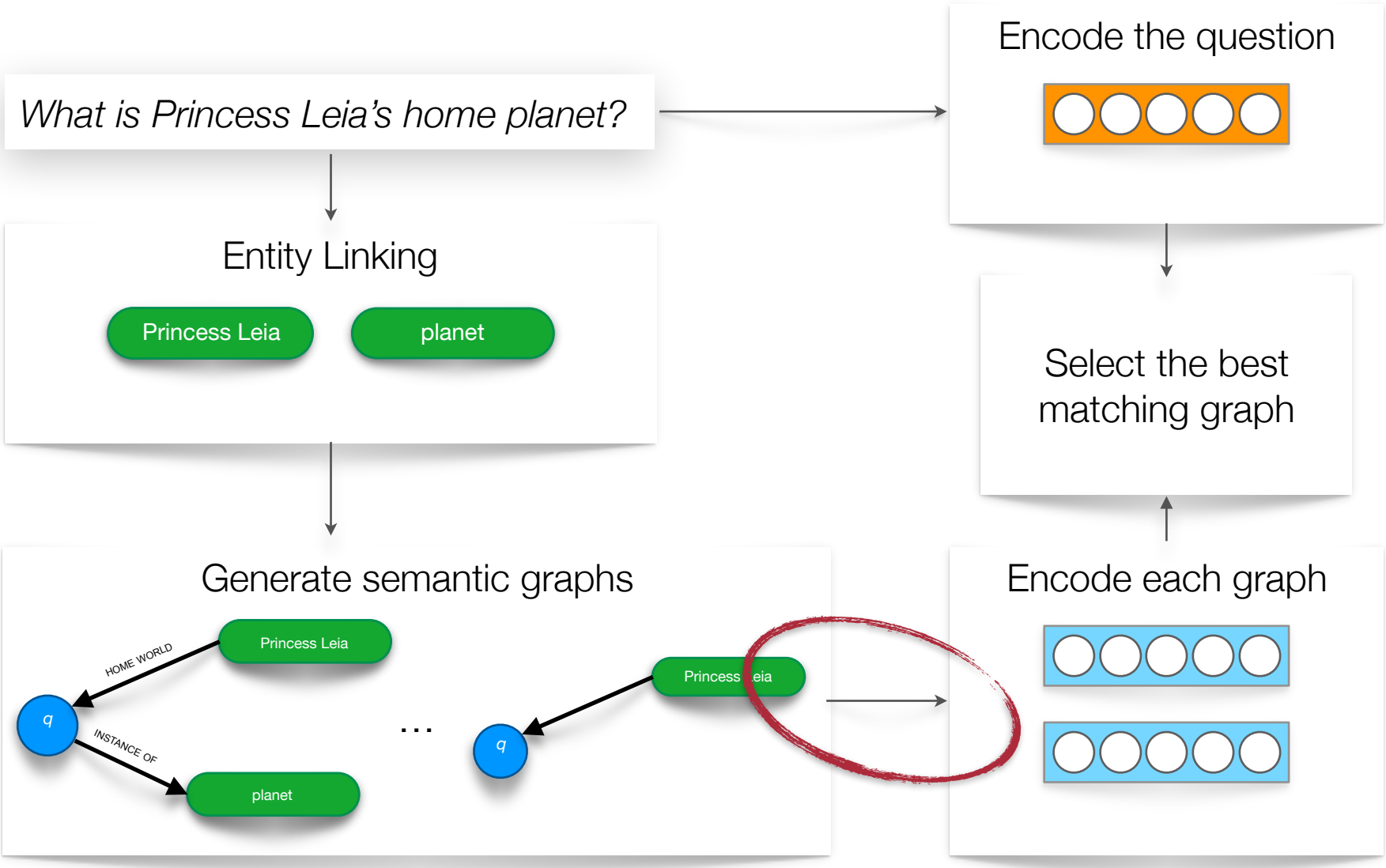
Modeling the knowledge base graph and other graphs

1. Graph Convolutional Networks
   - Kipf and Welling (2017) – document classification
   - Marcheggiani and Titov (2017) – encode syntax dependencies for SRL
   - Schlichtkrull et al. (2018) – KB embeddings
2. Gated Graph Neural Networks
   - Scarselli et al. (2009) – Graph Neural Networks
   - Li et al. (2016) – Gated variant
   - Chen et al. (2018) – Structured Dialogue Policy
   - Sun et al. (2018) – QA with Fusion of Knowledge Bases and Text

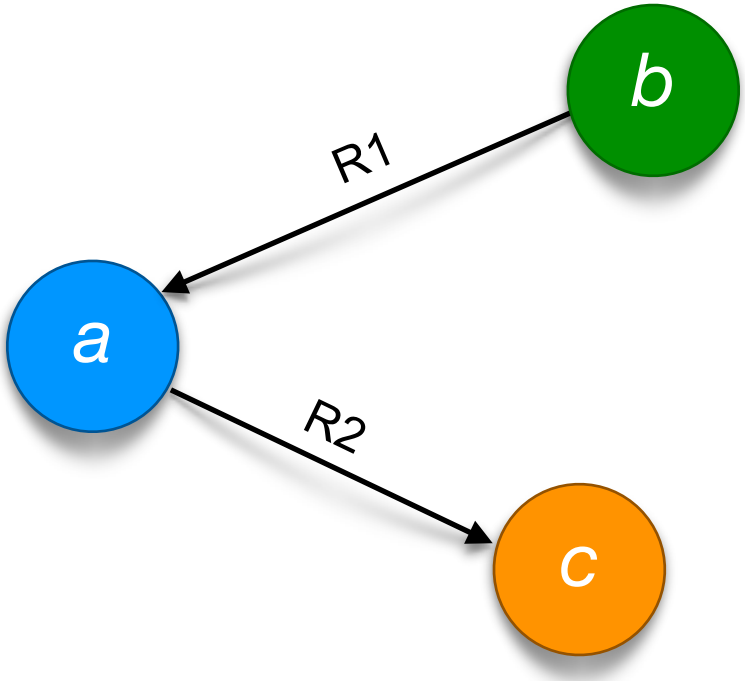
Knowledge Base Question Answering



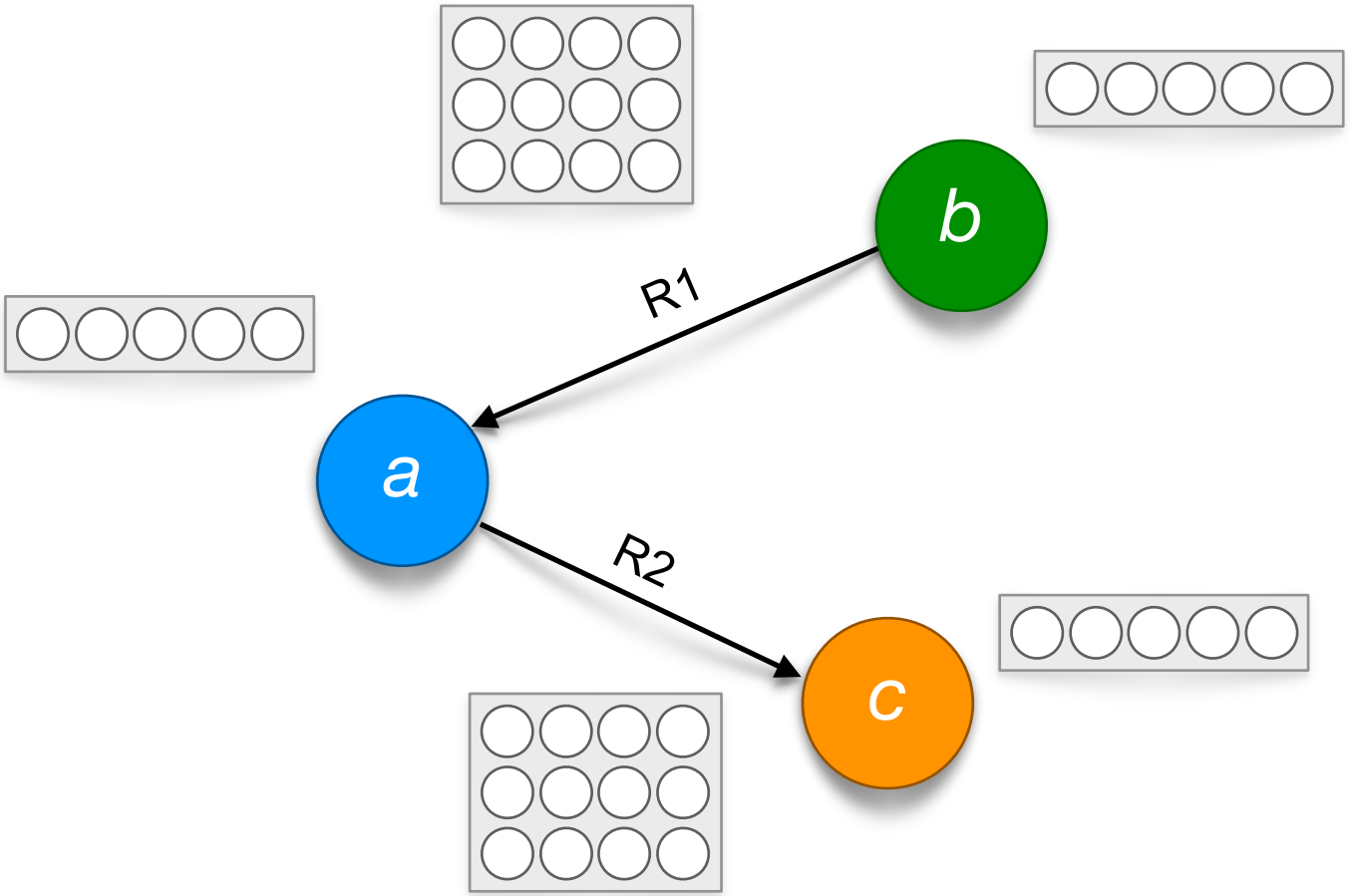
Representation Learning



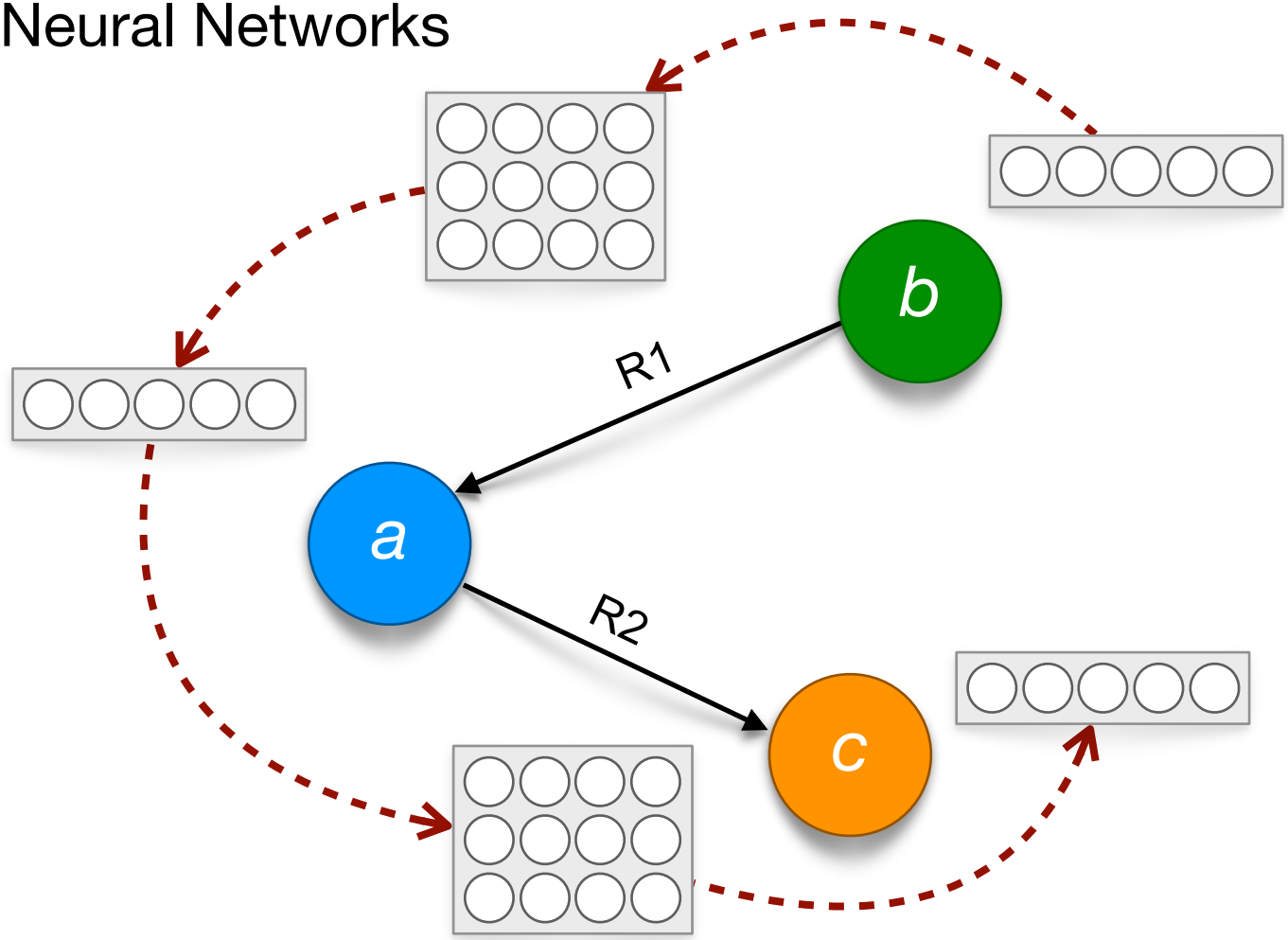
Graph Neural Networks – Scarselli et al. (2009)



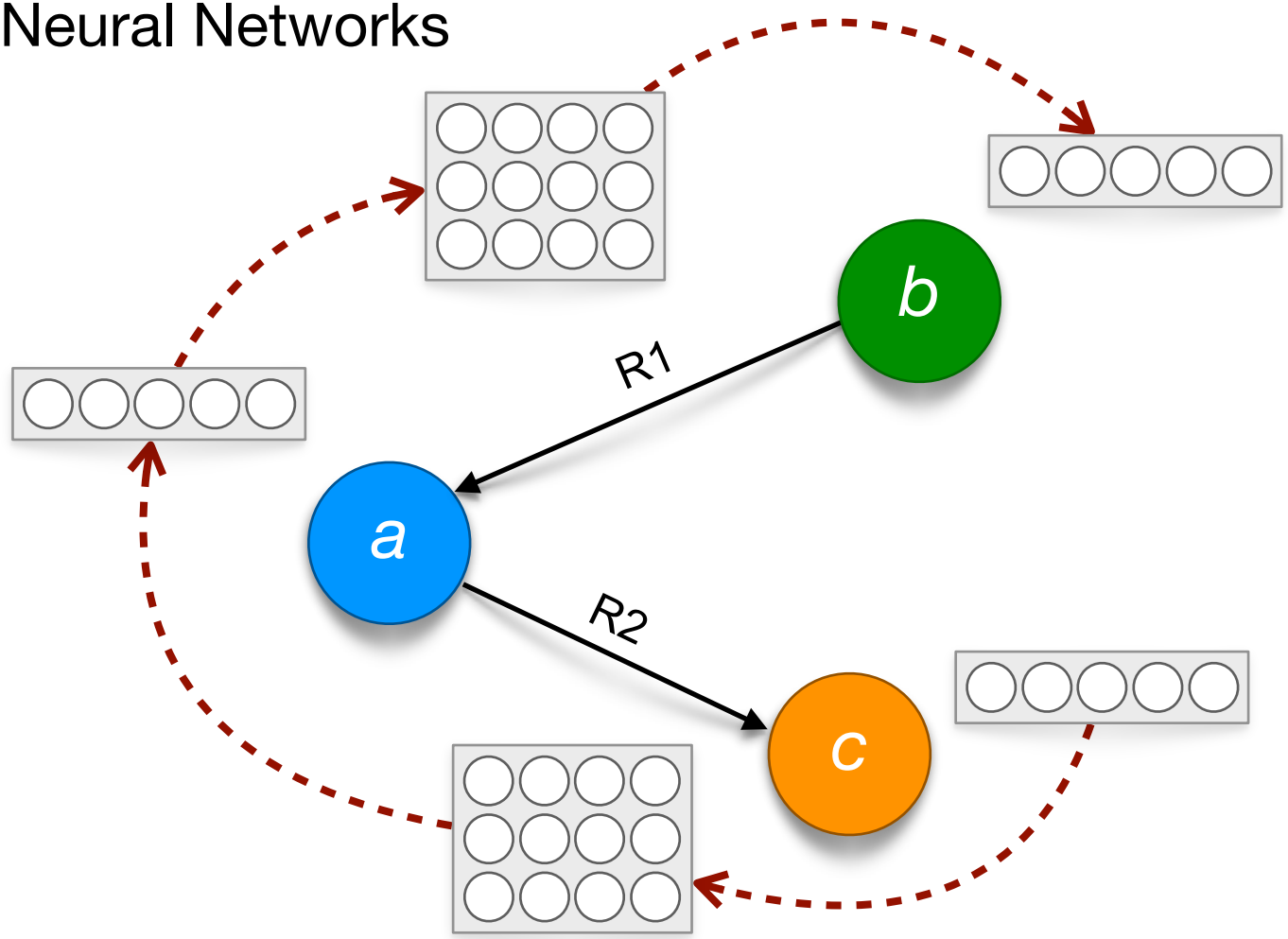
Graph Neural Networks



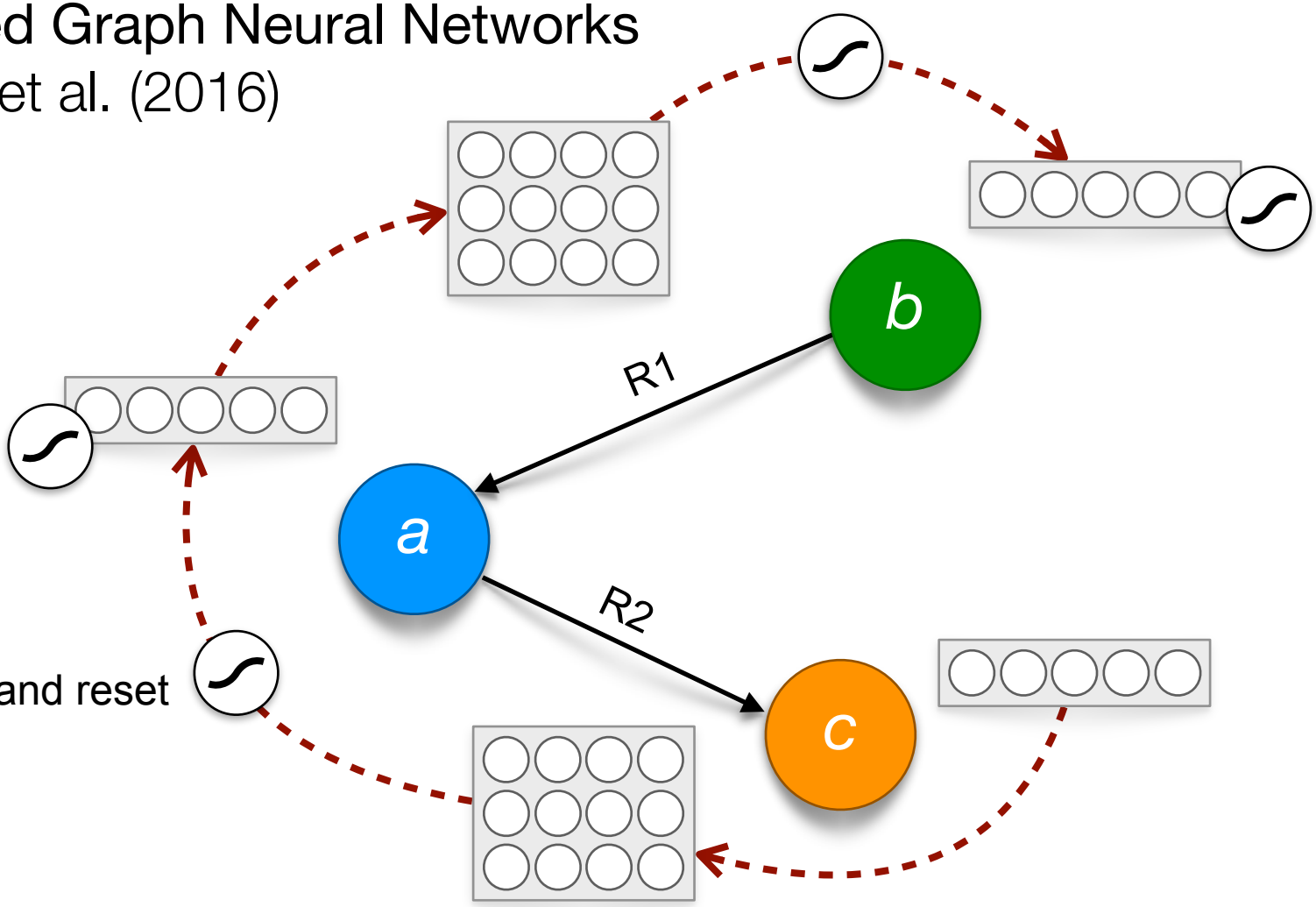
Graph Neural Networks



Graph Neural Networks

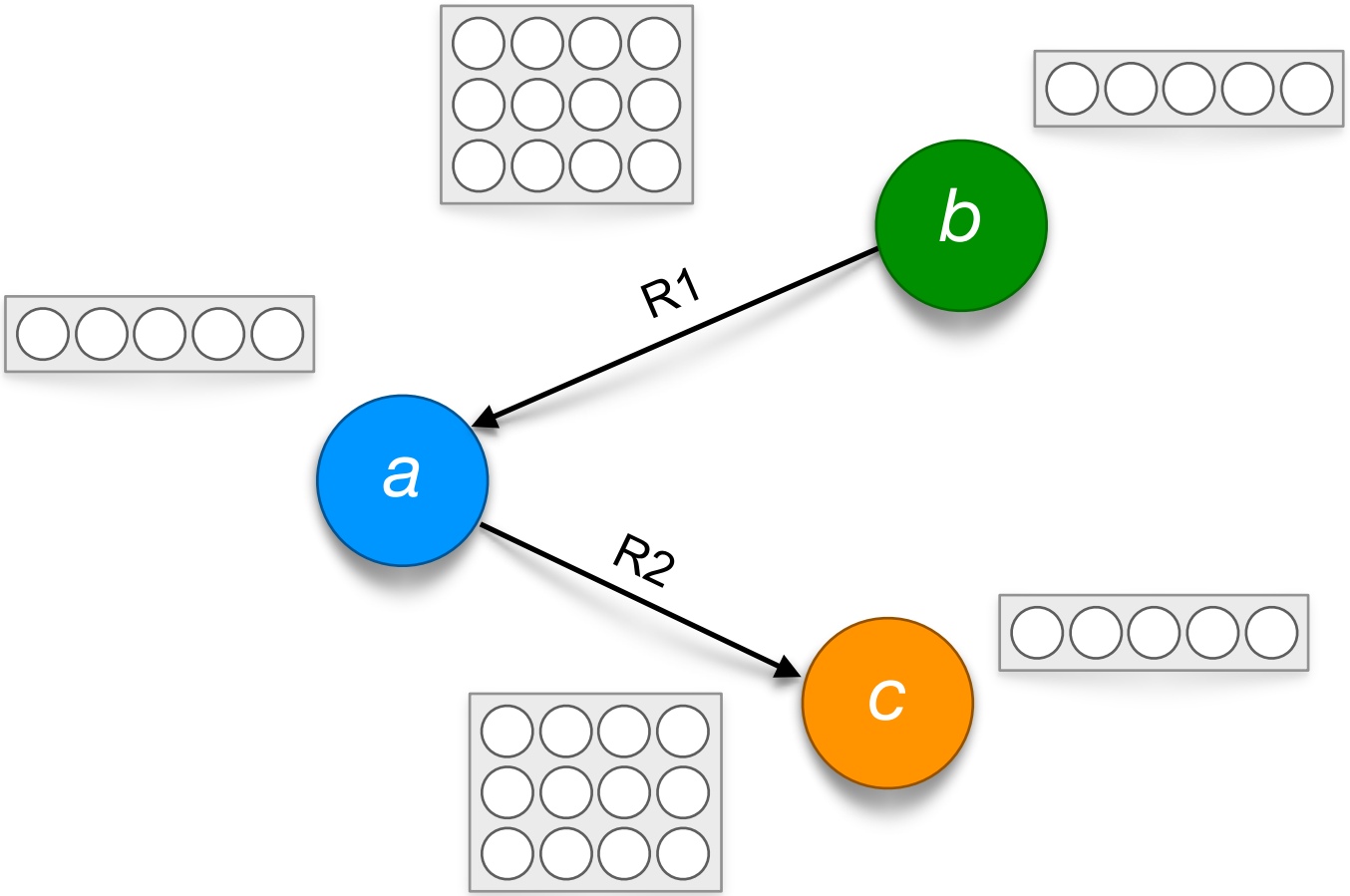


Gated Graph Neural Networks
– Li et al. (2016)

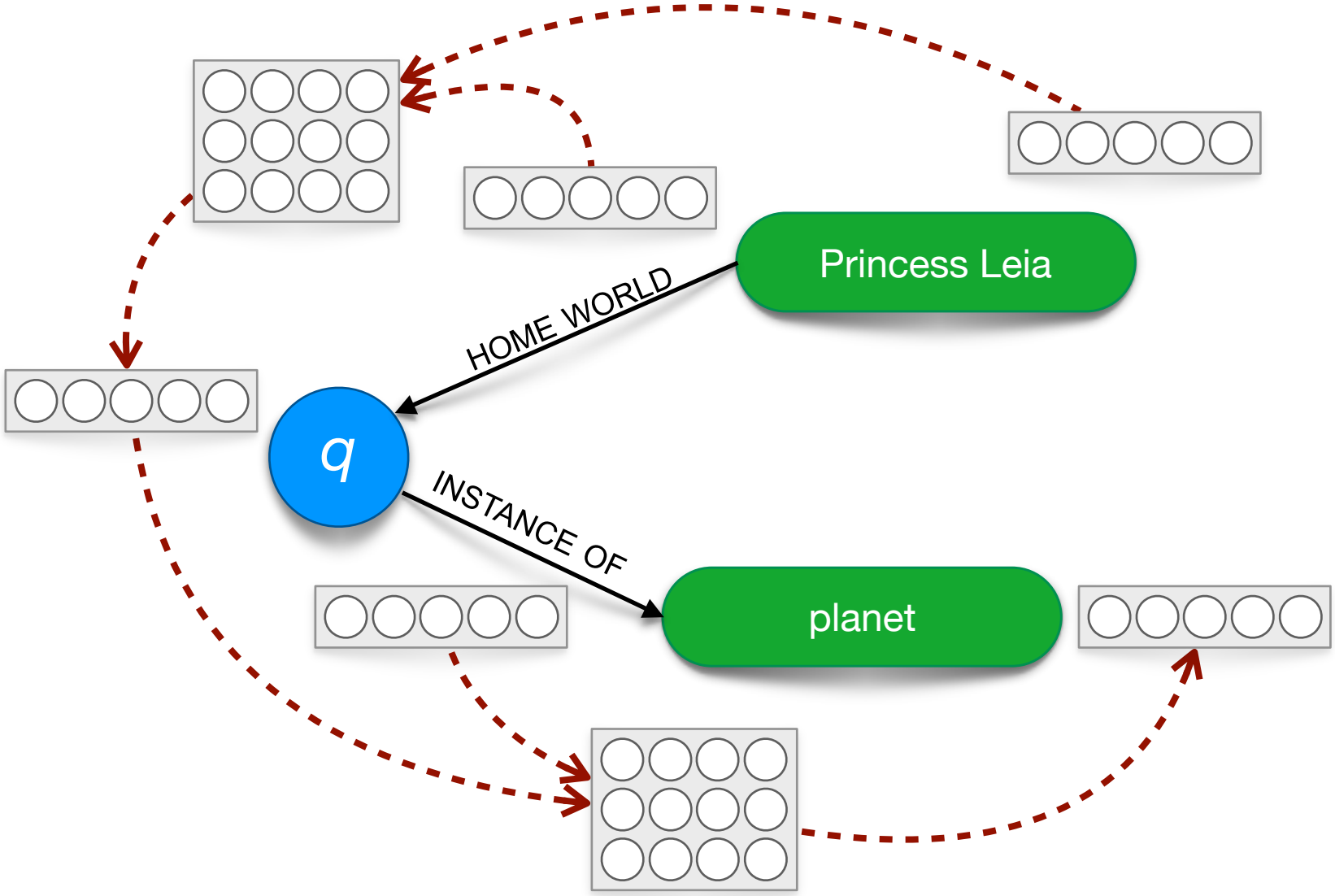


Update and reset gates

Gated Graph Neural Networks



Gated Graph Neural Networks with Labels



Propagation Model

Previous state

$$\left[\mathbf{h}_1^{(t-1)\top} \ldots \mathbf{h}_{|\mathscr{V}|}^{(t-1)\top}\right]$$

Relation encodings

$$\mathbf{h}'_r = \mathbf{W}_{\rightarrow}\mathbf{h}_{l_r} \quad \mathbf{h}''_r = \mathbf{W}_{\leftarrow}\mathbf{h}_{l_r}$$

Update vector

$$\mathbf{a}_v^{(t)} = \mathbf{A}_{v:}^{\top}\left[\mathbf{h}_1^{(t-1)\top} \ldots \mathbf{h}_{|\mathscr{V}|}^{(t-1)\top}\right] + \mathbf{A}'^{\top}_{r:}\left[\mathbf{h}_1'^{\top} \ldots \mathbf{h}_{|\mathscr{R}|}'^{\top}, \mathbf{h}_1''^{\top} \ldots \mathbf{h}_{|\mathscr{R}|}''^{\top}\right]$$

Update and reset gates

$$\mathbf{z}_v^t = \sigma\left(\mathbf{W}^z\mathbf{a}_v^{(t)} + \mathbf{U}^z\mathbf{h}_v^{(t-1)} + \mathbf{b}^z\right)$$

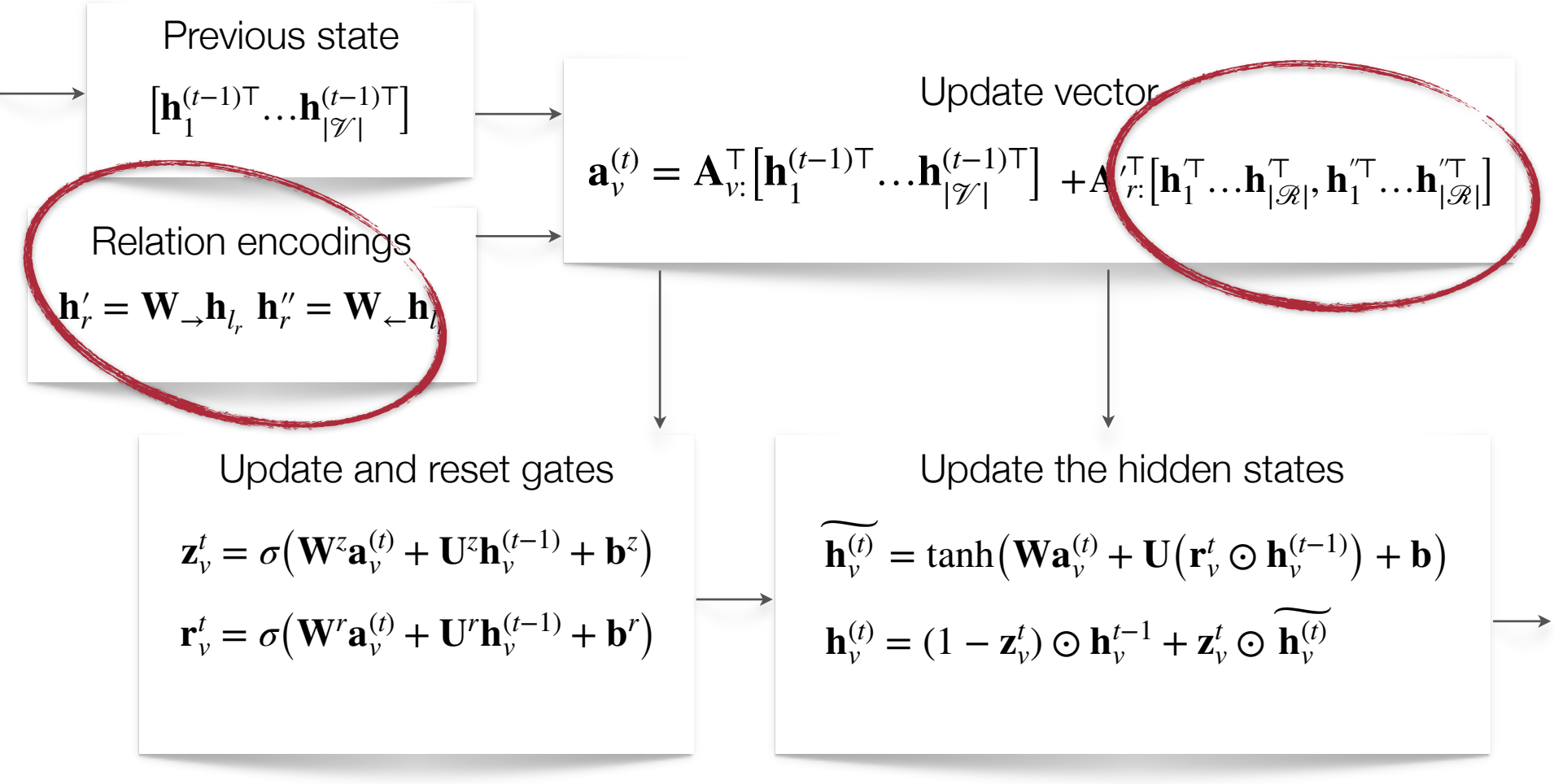
$$\mathbf{r}_v^t = \sigma\left(\mathbf{W}^r\mathbf{a}_v^{(t)} + \mathbf{U}^r\mathbf{h}_v^{(t-1)} + \mathbf{b}^r\right)$$

Update the hidden states

$$\widetilde{\mathbf{h}_v^{(t)}} = \tanh\left(\mathbf{W}\mathbf{a}_v^{(t)} + \mathbf{U}\left(\mathbf{r}_v^t \odot \mathbf{h}_v^{(t-1)}\right) + \mathbf{b}\right)$$

$$\mathbf{h}_v^{(t)} = (1 - \mathbf{z}_v^t) \odot \mathbf{h}_v^{t-1} + \mathbf{z}_v^t \odot \widetilde{\mathbf{h}_v^{(t)}}$$

Propagation Model



Propagation Model

Previous state

$$\left[\mathbf{h}_1^{(t-1)\top} \ldots \mathbf{h}_{|\mathscr{V}|}^{(t-1)\top}\right]$$

Relation encodings

$$\mathbf{h}'_r = \mathbf{W}_{\rightarrow}\mathbf{h}_{l_r} \quad \mathbf{h}''_r = \mathbf{W}_{\leftarrow}\mathbf{h}_{l_r}$$

Update vector

$$\mathbf{a}_v^{(t)} = \mathbf{A}_{v:}^{\top}\left[\mathbf{h}_1^{(t-1)\top} \ldots \mathbf{h}_{|\mathscr{V}|}^{(t-1)\top}\right] + \mathbf{A}'^{\top}_{r:}\left[\mathbf{h}_1'^{\top} \ldots \mathbf{h}_{|\mathscr{R}|}'^{\top}, \mathbf{h}_1''^{\top} \ldots \mathbf{h}_{|\mathscr{R}|}''^{\top}\right]$$

Update and reset gates

$$\mathbf{z}_v^t = \sigma\left(\mathbf{W}^z\mathbf{a}_v^{(t)} + \mathbf{U}^z\mathbf{h}_v^{(t-1)} + \mathbf{b}^z\right)$$

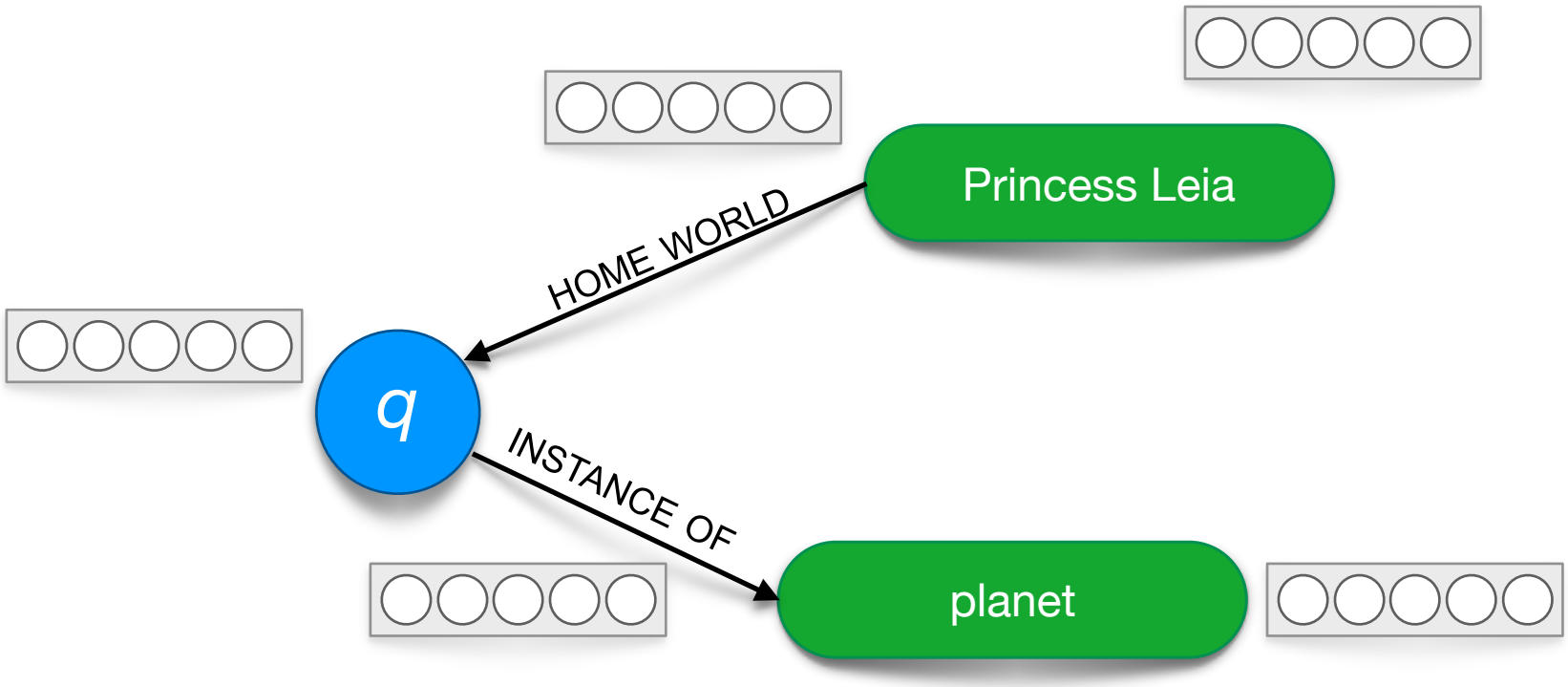
$$\mathbf{r}_v^t = \sigma\left(\mathbf{W}^r\mathbf{a}_v^{(t)} + \mathbf{U}^r\mathbf{h}_v^{(t-1)} + \mathbf{b}^r\right)$$

Update the hidden states

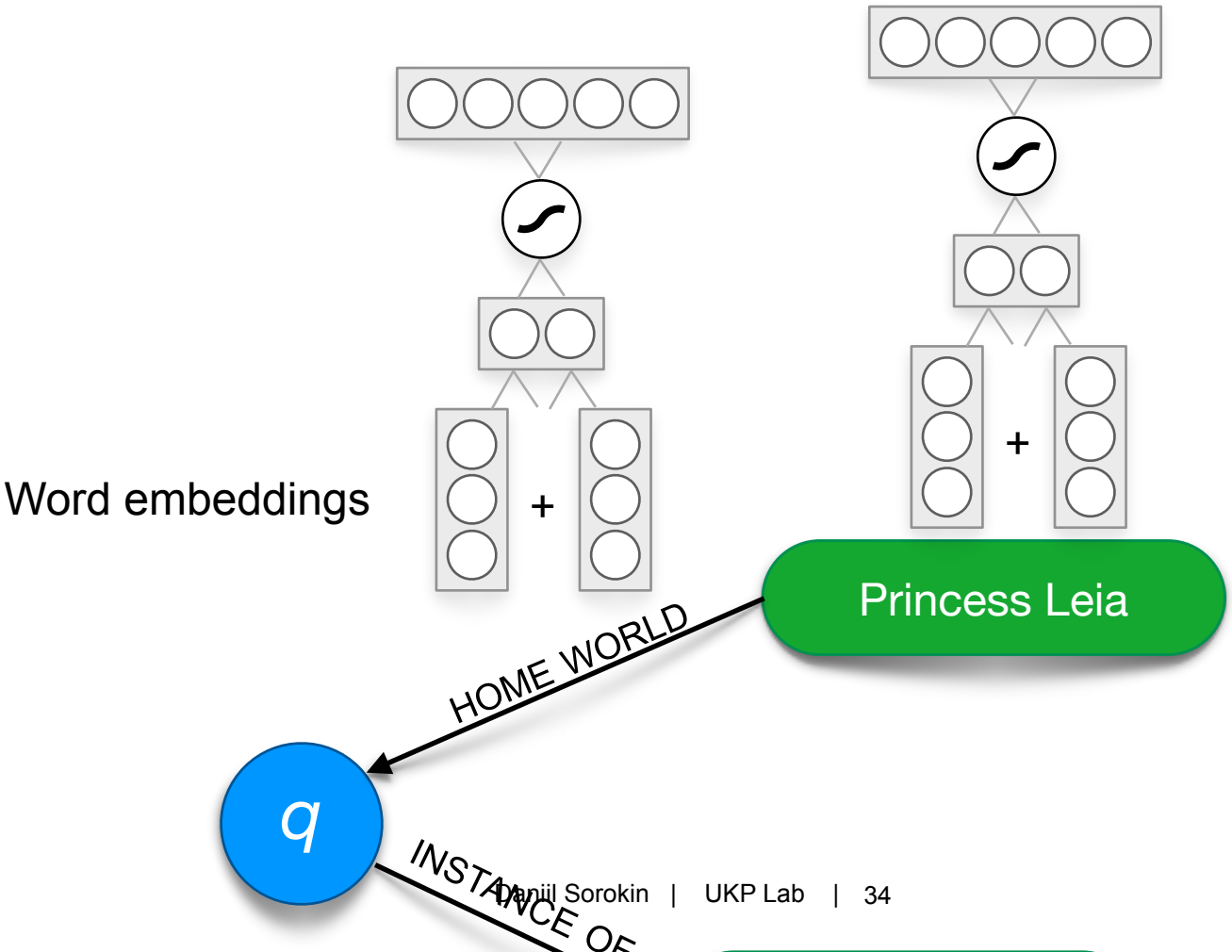
$$\widetilde{\mathbf{h}_v^{(t)}} = \tanh\left(\mathbf{W}\mathbf{a}_v^{(t)} + \mathbf{U}\left(\mathbf{r}_v^t \odot \mathbf{h}_v^{(t-1)}\right) + \mathbf{b}\right)$$

$$\mathbf{h}_v^{(t)} = (1 - \mathbf{z}_v^t) \odot \mathbf{h}_v^{t-1} + \mathbf{z}_v^t \odot \widetilde{\mathbf{h}_v^{(t)}}$$

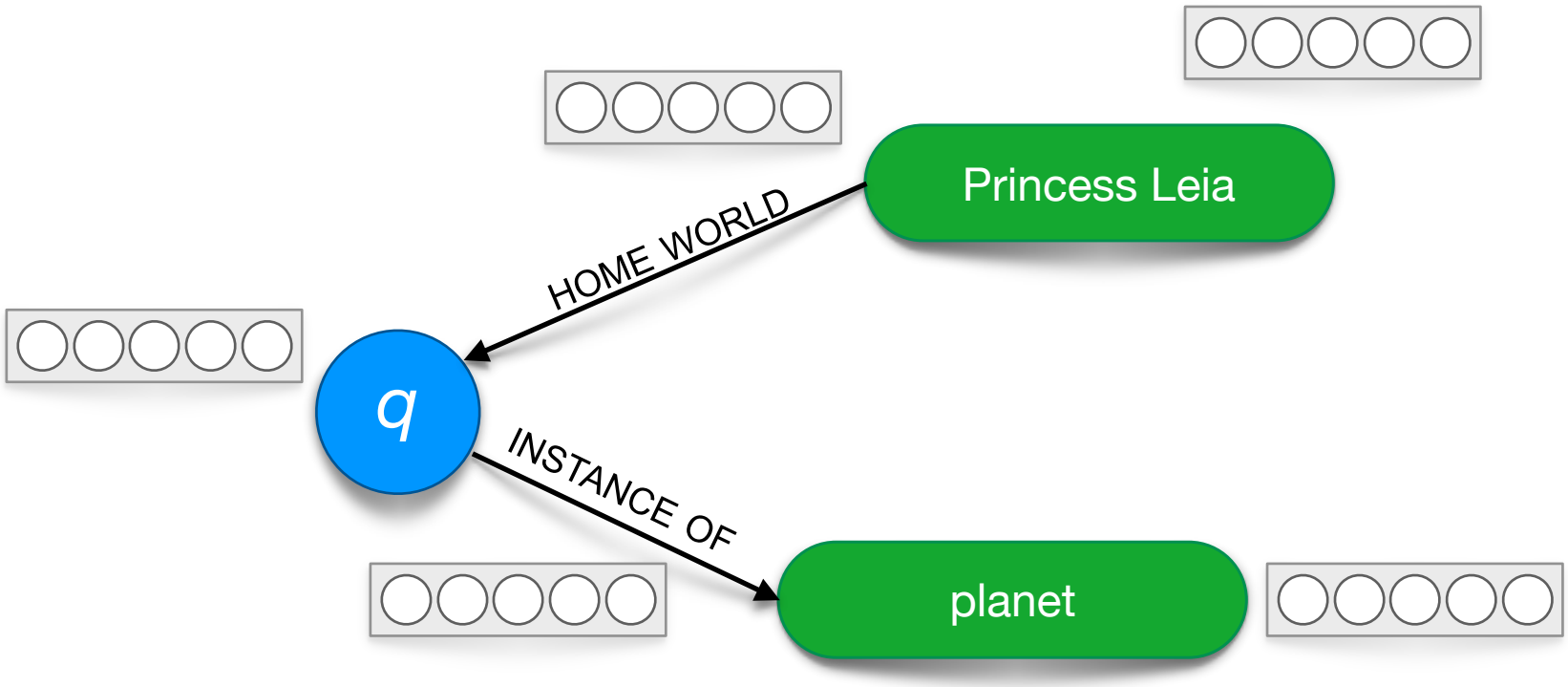
Entity and Relation Embeddings



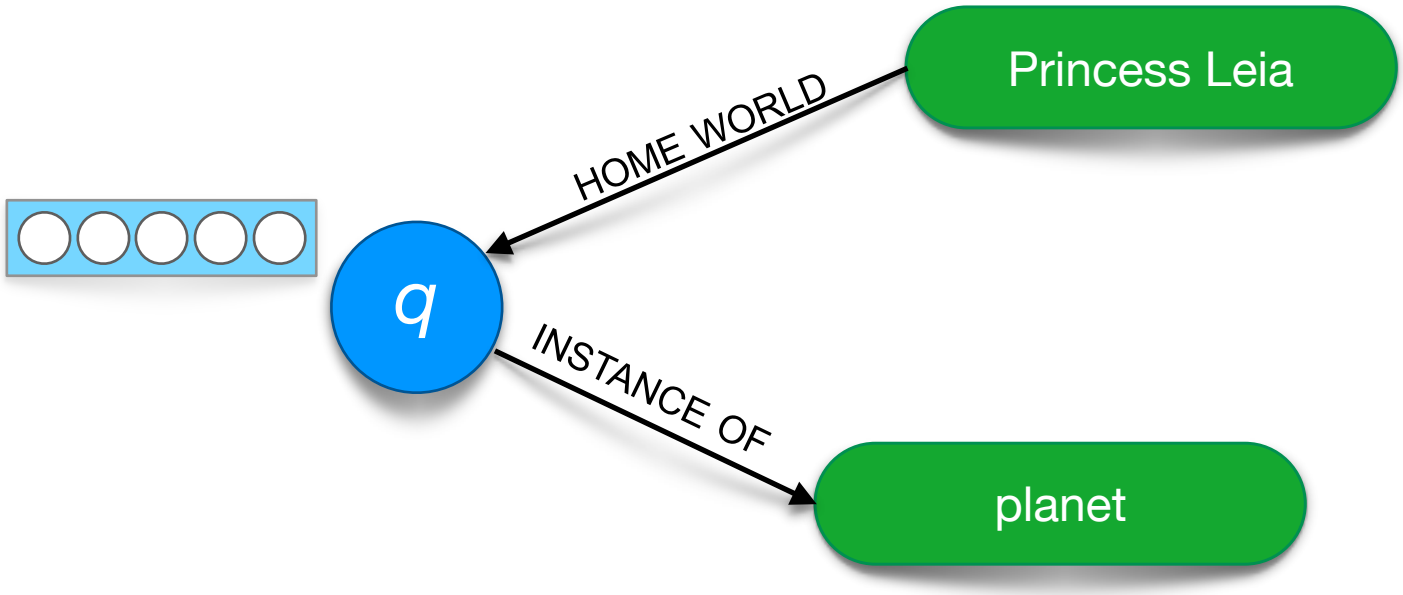
Entity and Relation Encodings



Graph Encoding

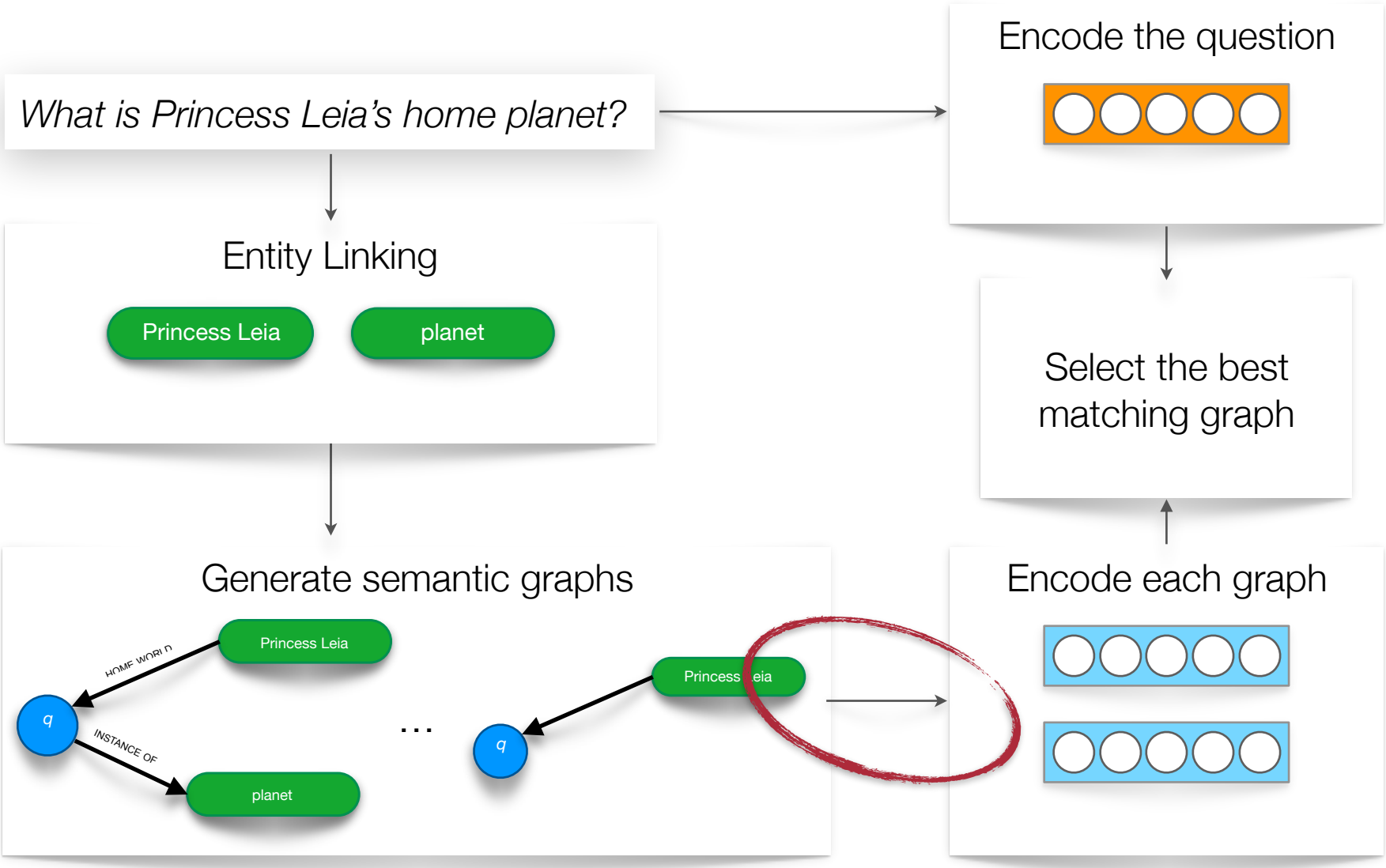


Graph Encoding

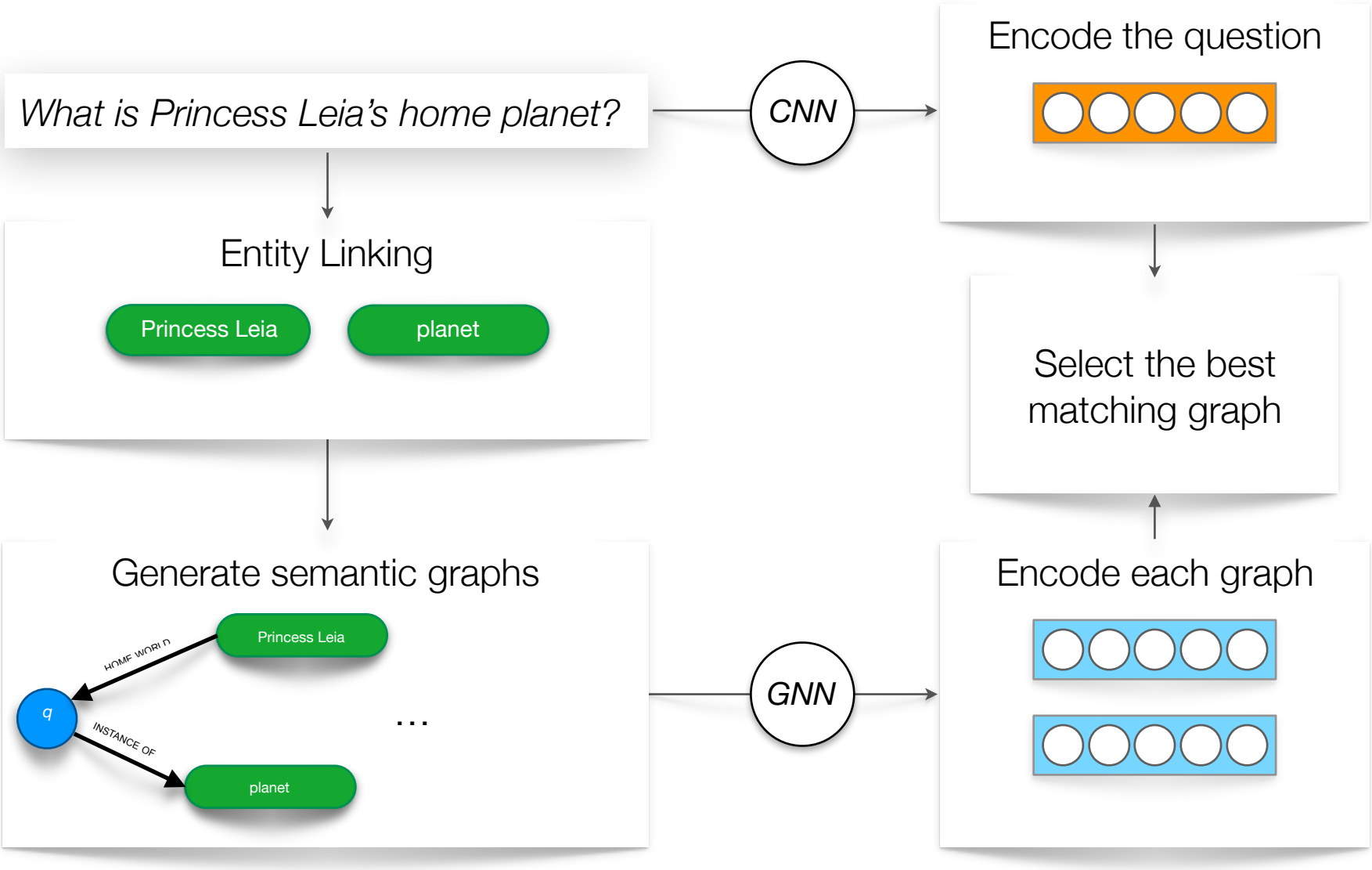


The Graph Neural Network models all aspects of the semantic graph representation

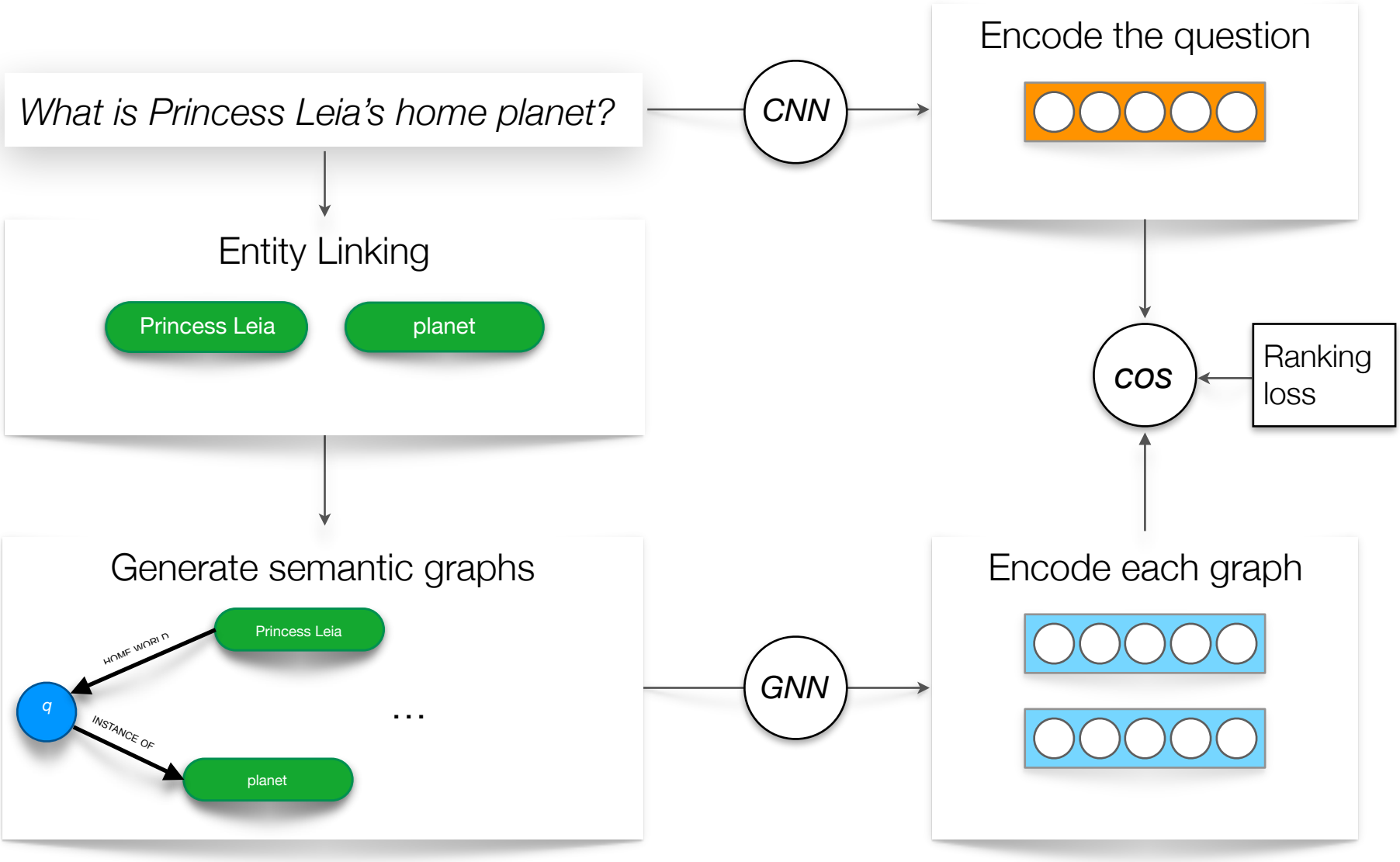
Question Answering Pipeline – Representation Learning



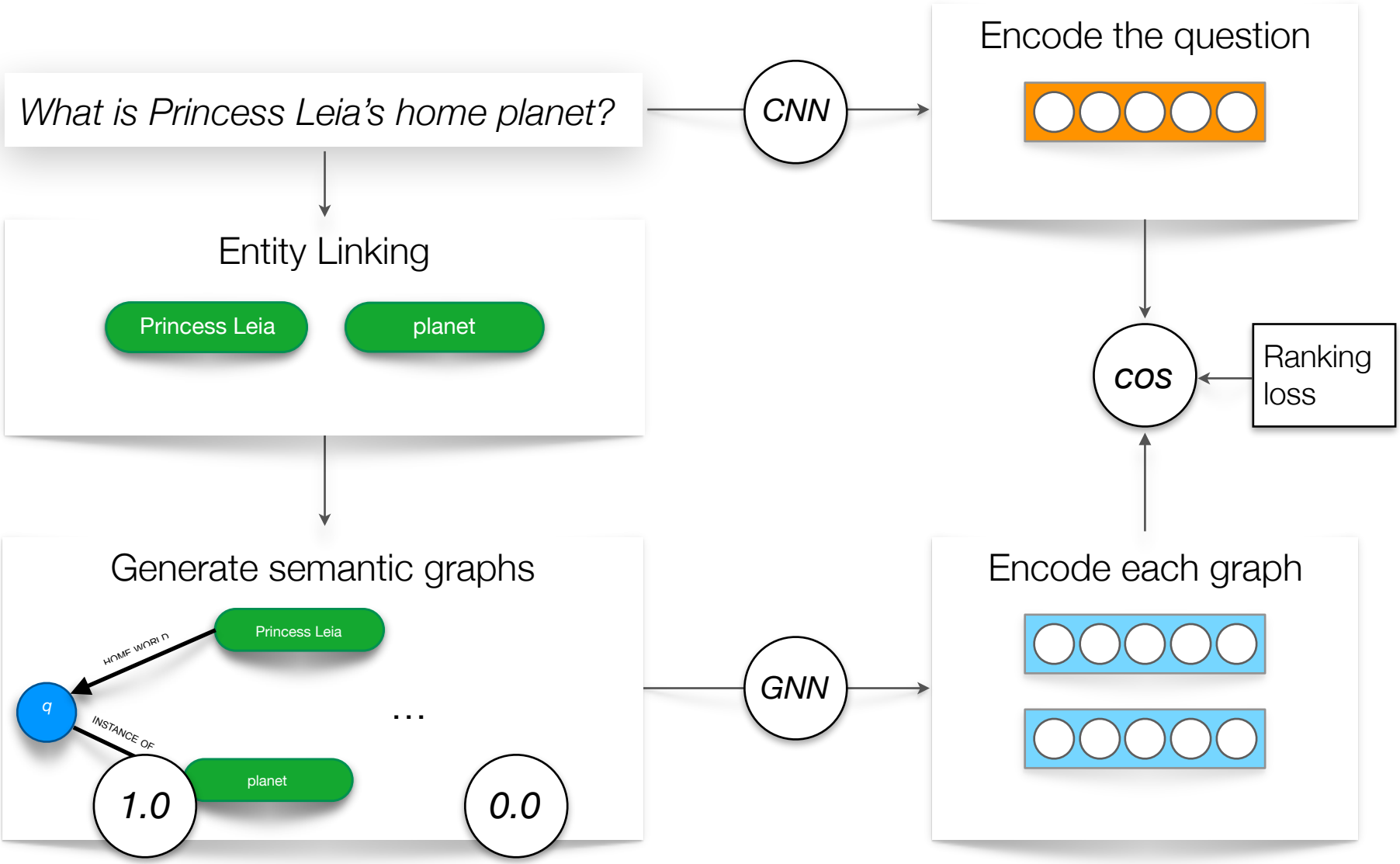
Question Answering Pipeline – Representation Learning



Question Answering Pipeline – Representation Learning



Question Answering Pipeline – Representation Learning



Models

- STAGG (Yih et al. 2015)
- Single Edge
- Pooled Edges
- GNN
- Gated GNN

Data set

WebQSP-WD

Based on WebQSP/WebQuestions

Mapped to Wikidata

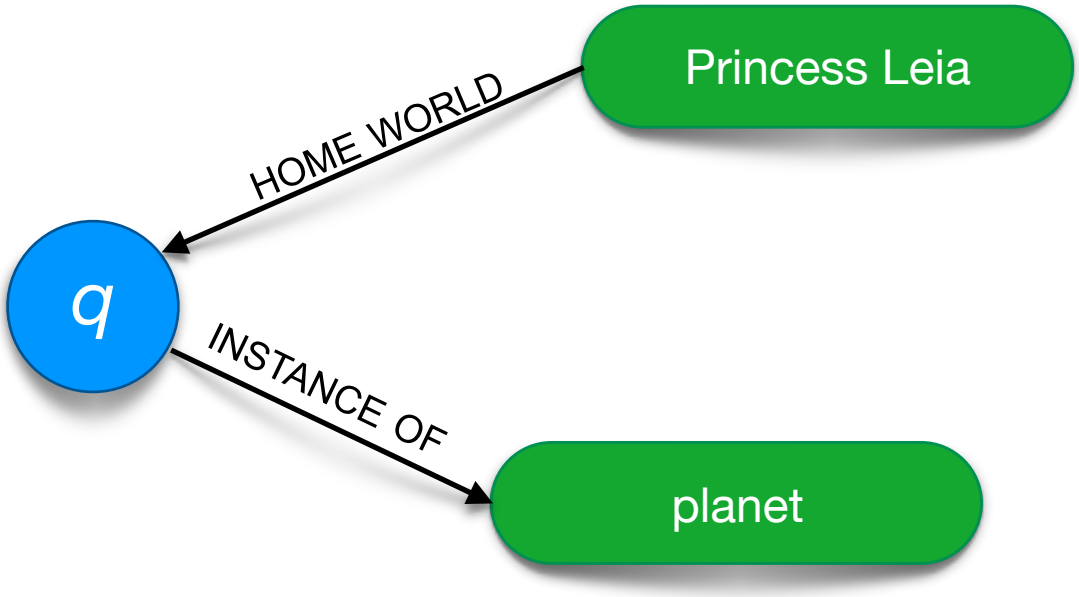
1033 test questions (293 complex)

2880 train questions (419 complex)

Evaluation

What is Princess Leia's home planet?

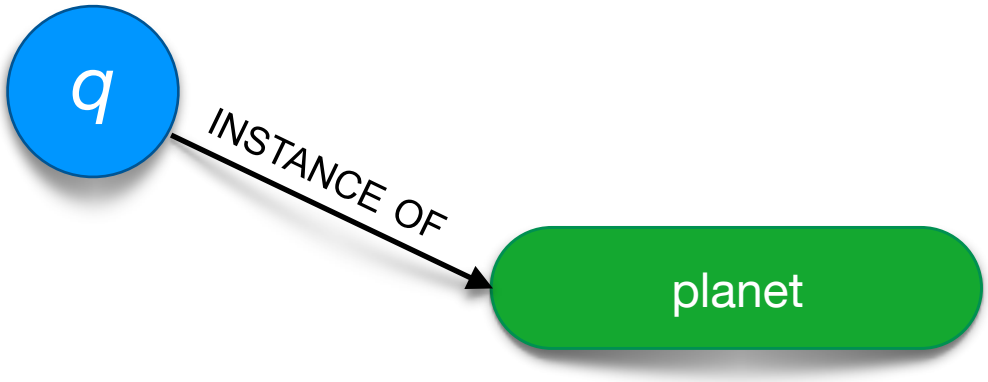
Precision	Recall	F-score
1.000	1.000	1.000



Evaluation

What is Princess Leia's home planet?

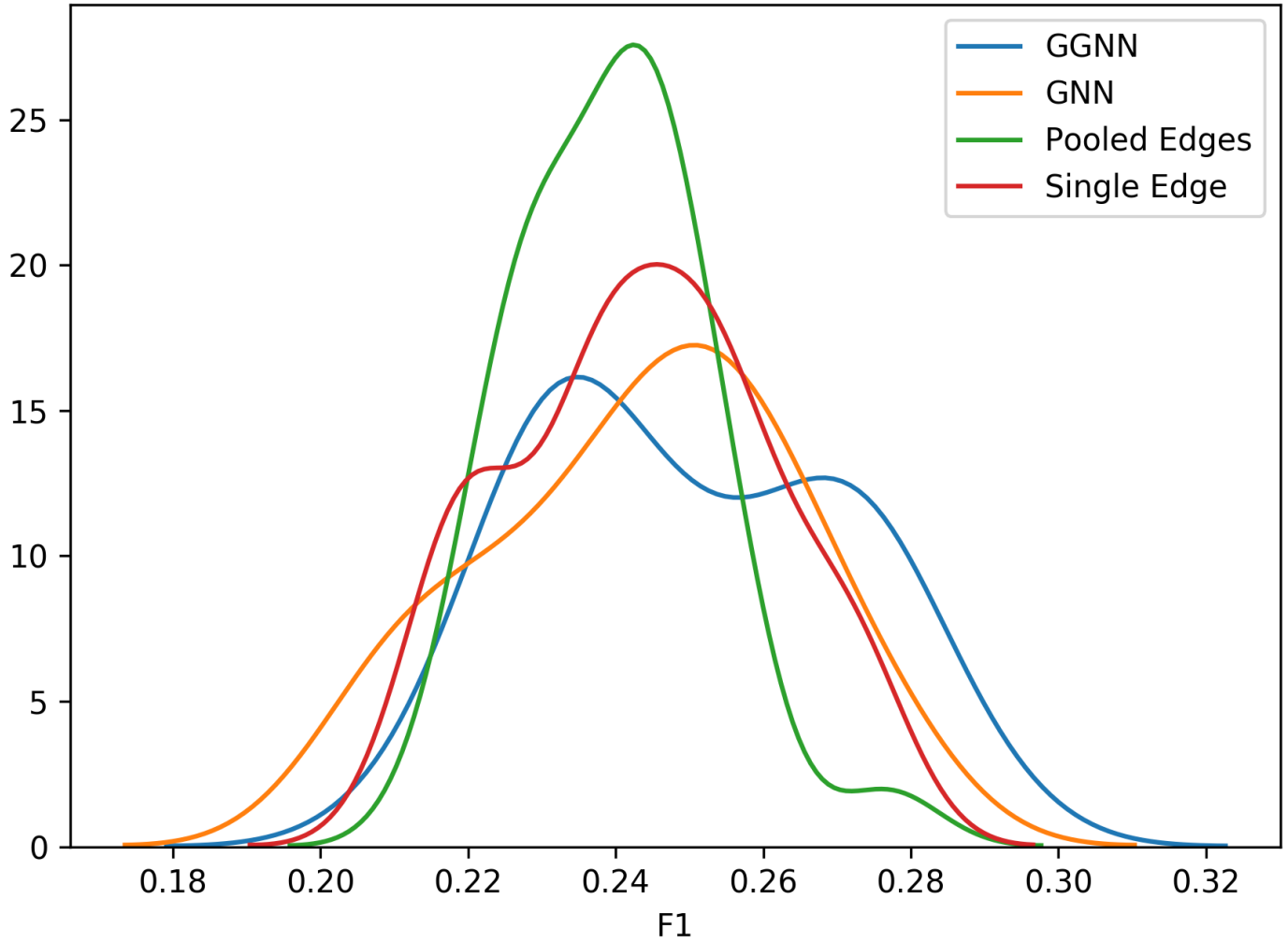
Precision	Recall	F-score
0.050	1.000	0.095

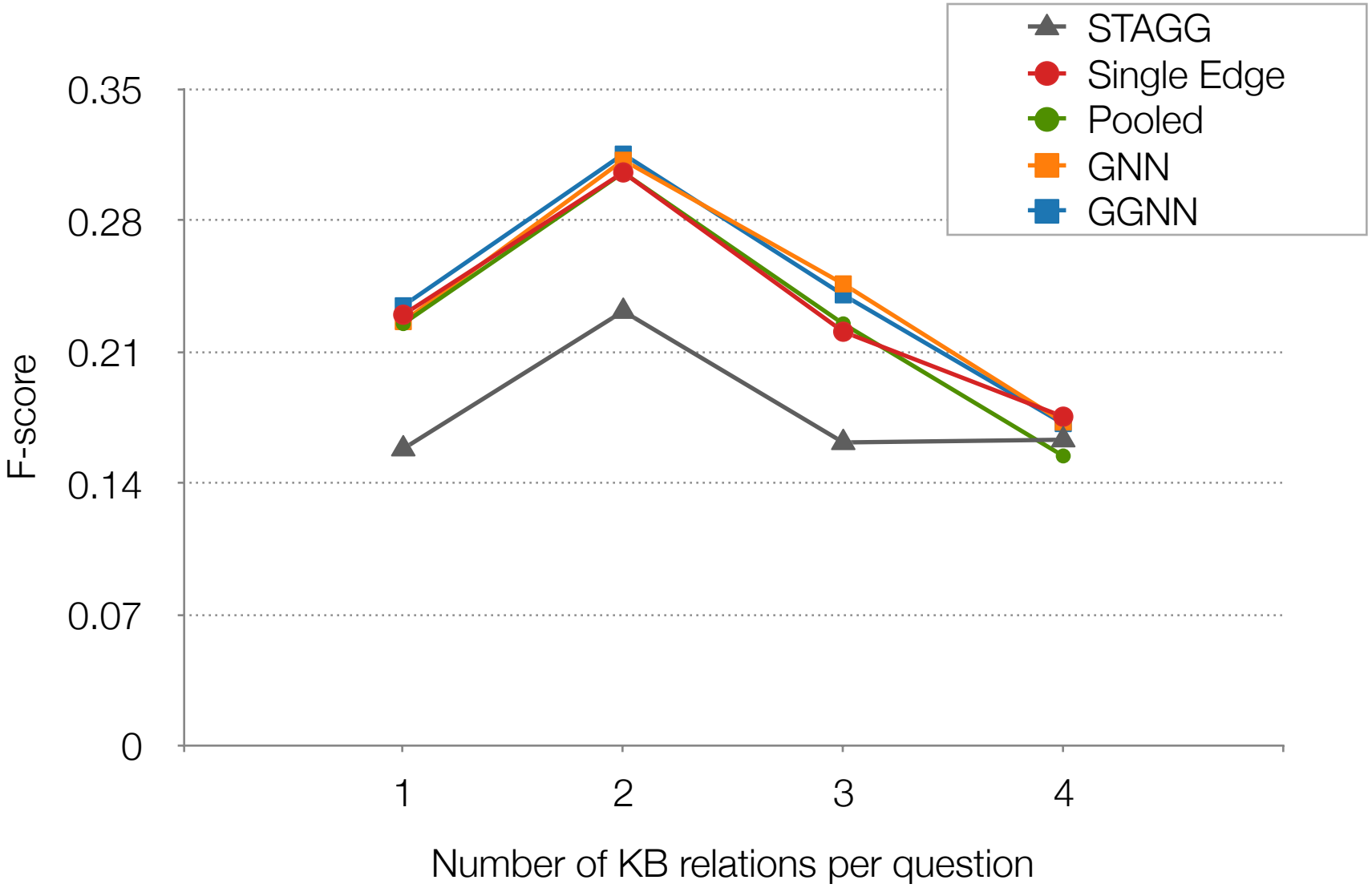


WebQSP-WD

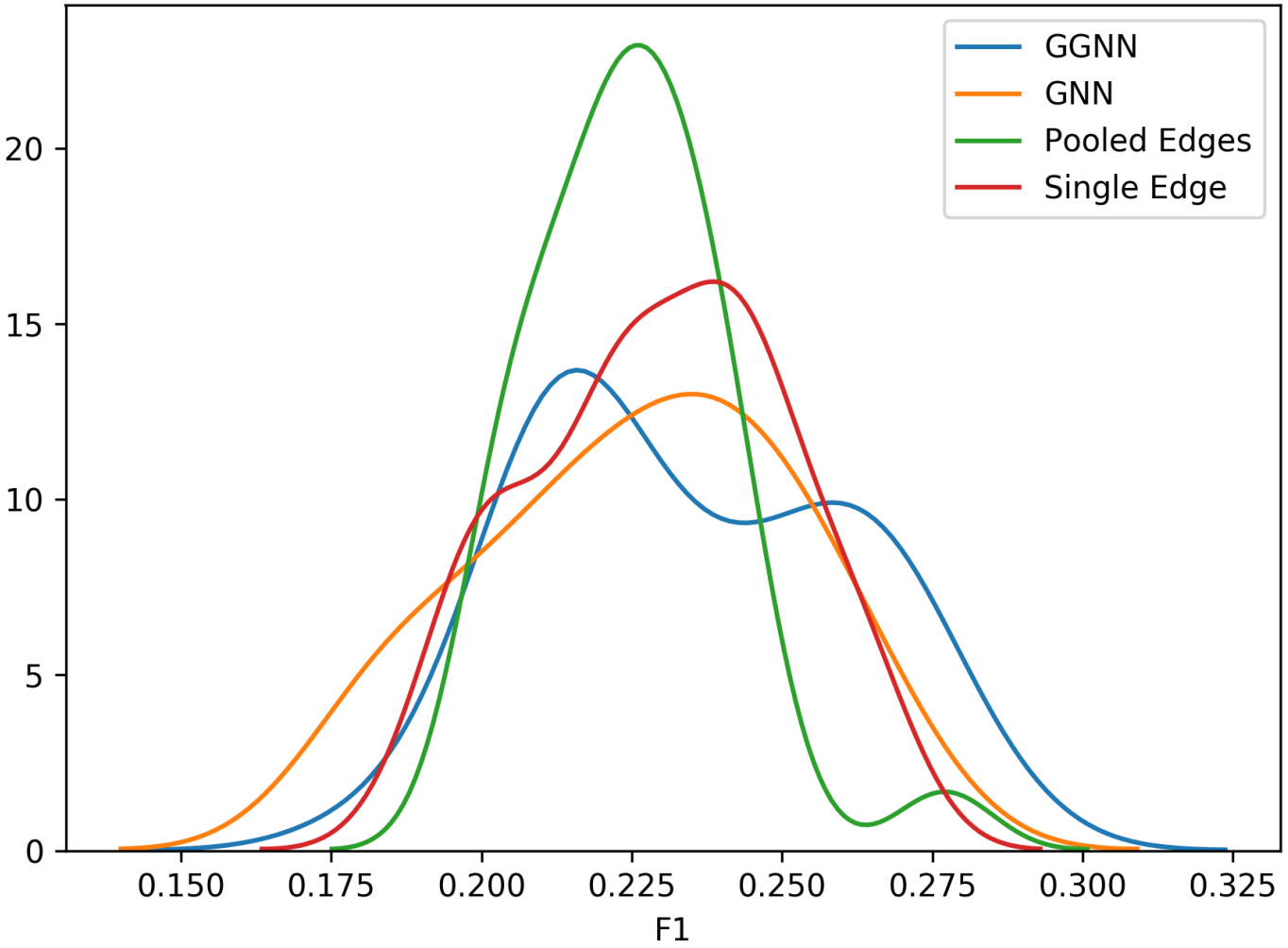
	Precision	Recall	F-score
STAGG	0.173	0.214	0.167
Single Edge	0.250	0.307	0.243
Pooled Edges	0.245	0.302	0.240
GNN	0.253	0.293	0.243
GGNN	0.259	0.305	**0.250**

WebQSP-WD

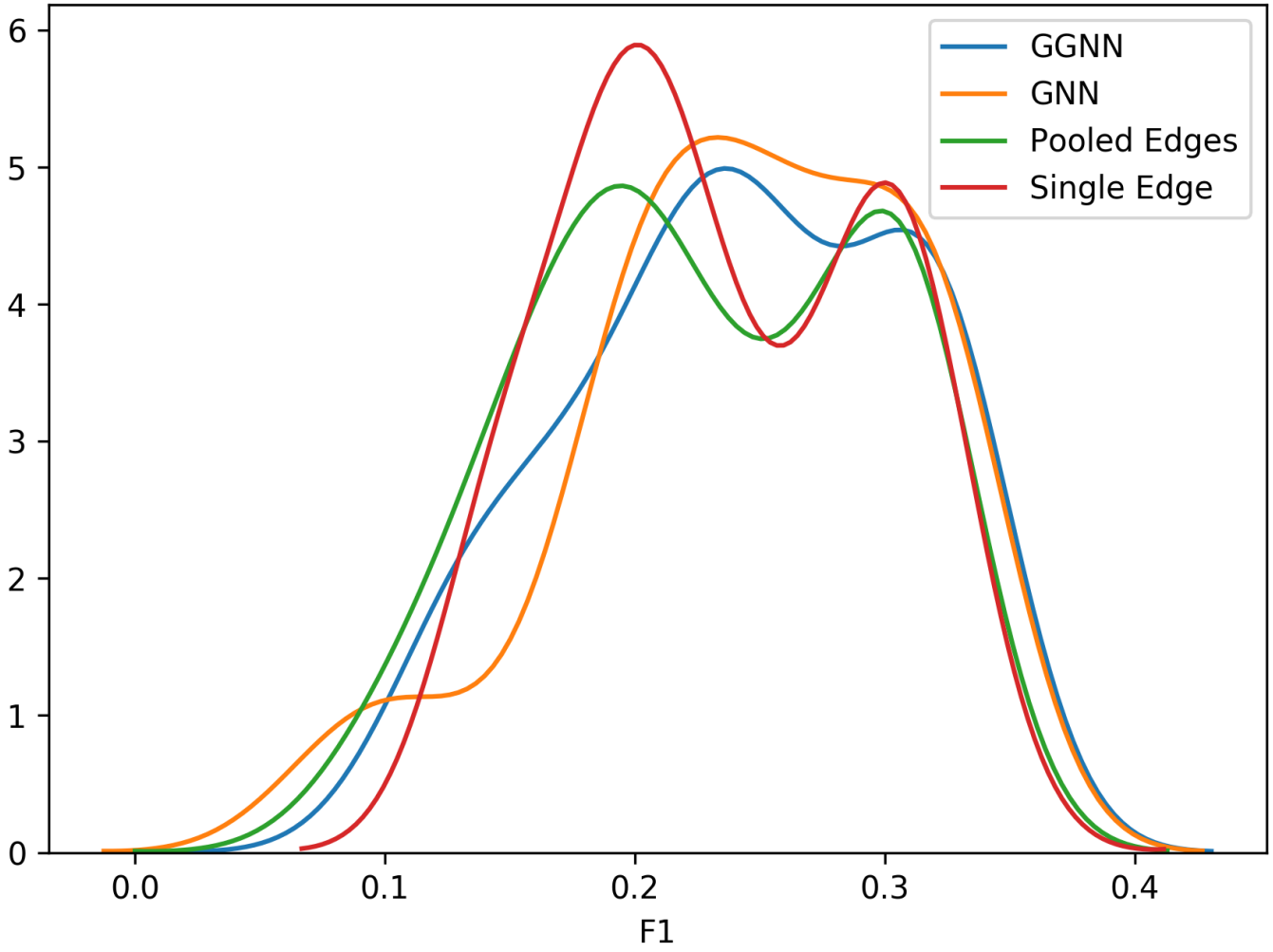

GGNN
GNN
Pooled Edges
Single Edge
25
20
15
10
5
0
0.18
0.20
0.22
0.24
0.26
0.28
0.30
0.32
F1


STAGG
Single Edge
Pooled
GNN
GGNN
0.35
0.28
0.21
0.14
0.07
0
F-score
1
2
3
4
Number of KB relations per question

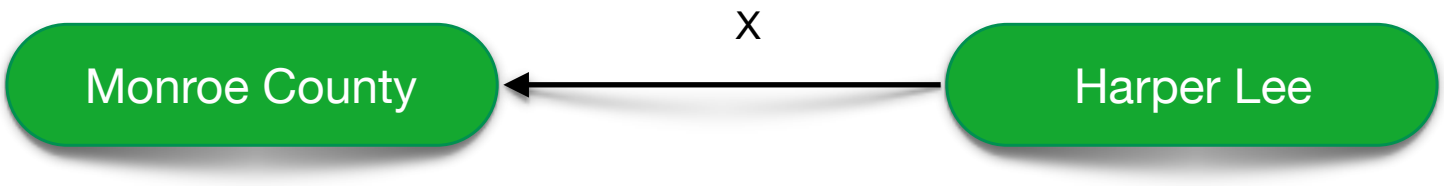
WebQSP-WD simple questions



WebQSP-WD complex questions

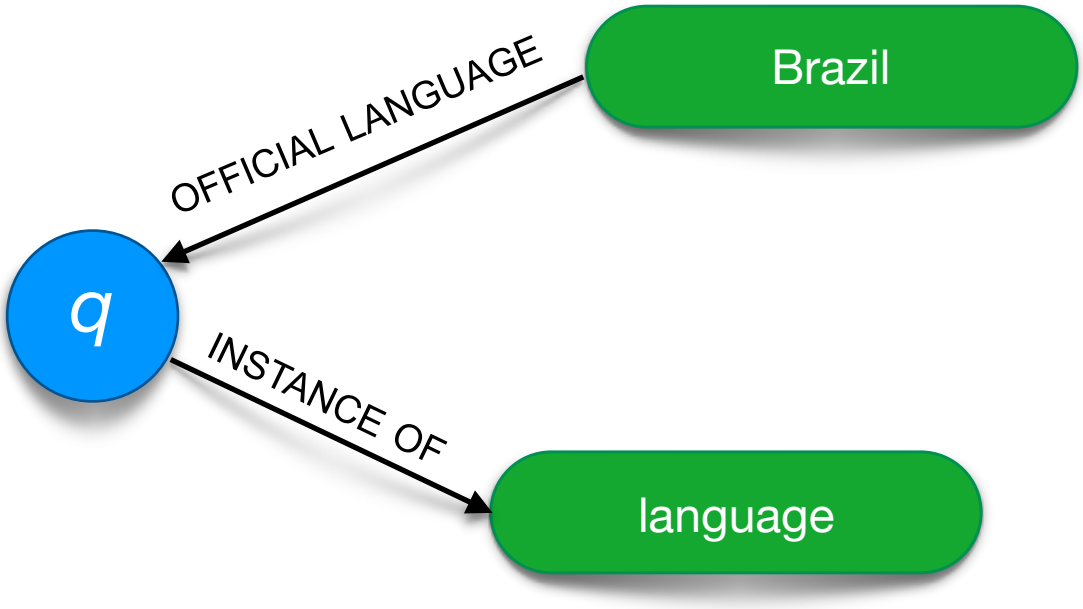


Example output



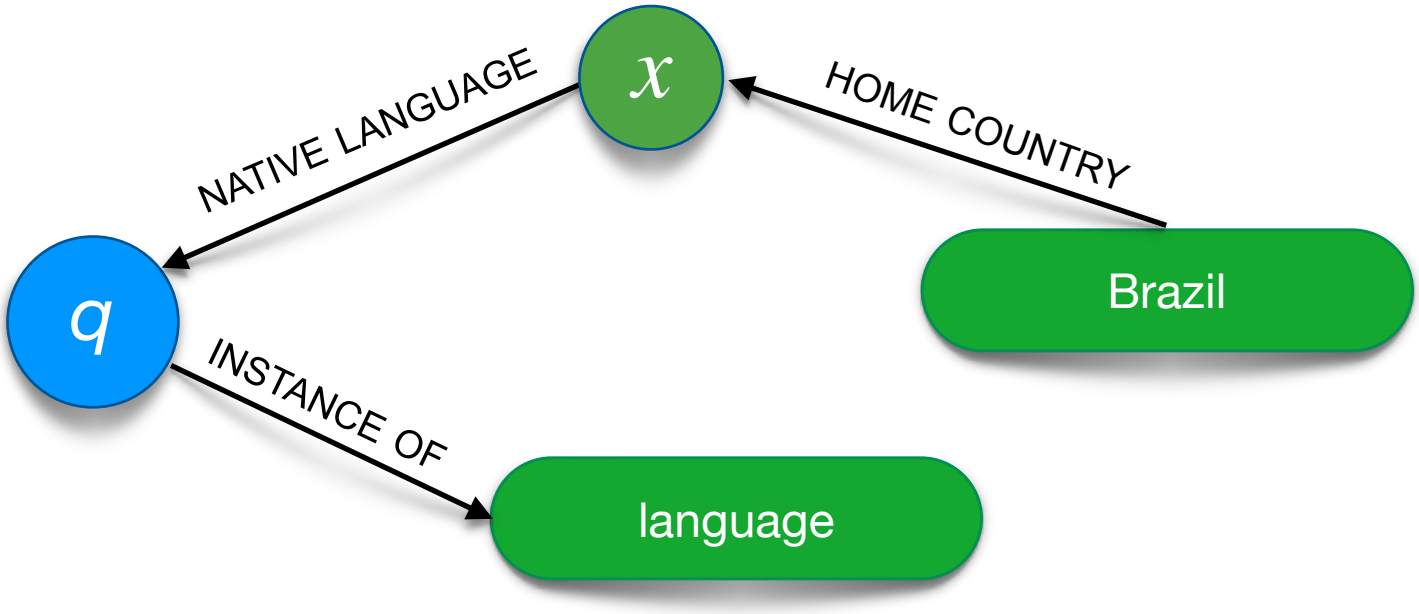
Where did Harper Lee attend high school?

Example output Pooled



What language do people speak in Brazil?

Example output GGNN



What language do people speak in Brazil?

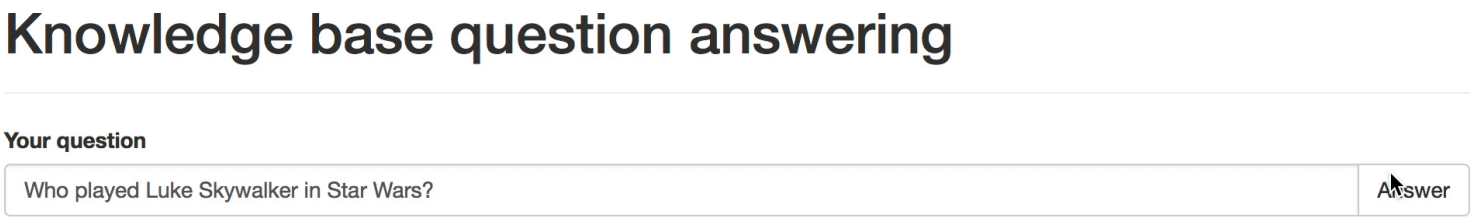
Limitations

1. The small training set limits the more complex GraphNNs architecture
2. Small test set
3. WebQSP contains a lot of data set errors and is not entirely covered by Wikidata (around 30% of errors)
4. Not many Wikidata data sets for a meaningful comparison
5. Entity linking is a bottle neck (10% of errors)

Modelling the semantic structure with **GraphNNs** improves the representation learning for knowledge base graphs

Demo and code

https://github.com/UKPLab/coling2018-graph-neural-networks-question-answering



Follow up?

- https://daniilsorokin.github.io ⬅️ slides
- sorokin@ukp.informatik.tu-darmstadt.de
- @daniilmagpie
- COLING'18 Paper: http://aclweb.org/anthology/C18-1280