Carnegie Mellon University
Language Technologies Institute

Google

Text as a Virtual Knowledge Base

Bhuwan Dhingra
Language Technologies Institute
Carnegie Mellon University

Work done with: Haitian Sun, Manzil Zaheer, Vidhisha Balachandran,
Graham Neubig, Russ Salakhutdinov, William Cohen

Question Answering

When did Kendrick Lamar's first album come out?

A. July 2, 2011

Source: Natural Questions (https://ai.google.com/research/NaturalQuestions/dataset)

Sources of Information

Text

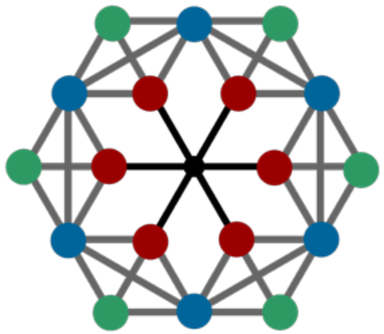
WIKIPEDIA
The Free Encyclopedia

REUTERS

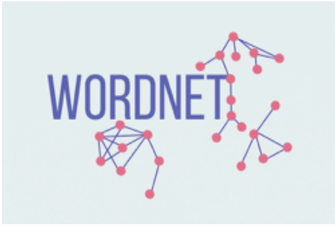
reddit

Knowledge Bases

Freebase



WIKIDATA



Sources of Information

Text

[Moldovan, 2002], [Voorhees, 1999], [Ferrucci, 2012], [Yih, 2013], [Hermann, 2016], [Chen, 2017], [Seo, 2017], [Peters, 2018], [Devlin, 2018] …

Knowledge Bases

[Kwiatkowski, 2013], [Berant, 2013], [Reddy, 2014], [Pasupat, 2015], [Bordes, 2015], [Yih, 2015], [Jain, 2016], [Liang, 2017], [Das, 2017] …

Sources of Information

Text

[Moldovan, 2002], [Voorhees, 1999], [Ferrucci, 2012], [Yih, 2013], [Hermann, 2016], [Chen, 2017], [Seo, 2017], [Peters, 2018], [Devlin, 2018] …

- Lexical pattern matching
- No grounding
- High recall

Knowledge Bases

[Kwiatkowski, 2013], [Berant, 2013], [Reddy, 2014], [Pasupat, 2015], [Bordes, 2015], [Yih, 2015], [Jain, 2016], [Liang, 2017], [Das, 2017] …

- Clear semantics via parsing
- Grounded
- High precision

Sources of Information

Text

Knowledge Bases

[Gardner & Krishnamurthy, 2017], [Ryu, 2017]
Universal Schema [Riedel, 2013], [Verga, 2016], [Das, 2017] …
Our Work

- Lexical pattern matching
- No grounding
- High recall

- Clear semantics via parsing
- Grounded
- High precision

Why Text + KBs?

Why Text + KBs?

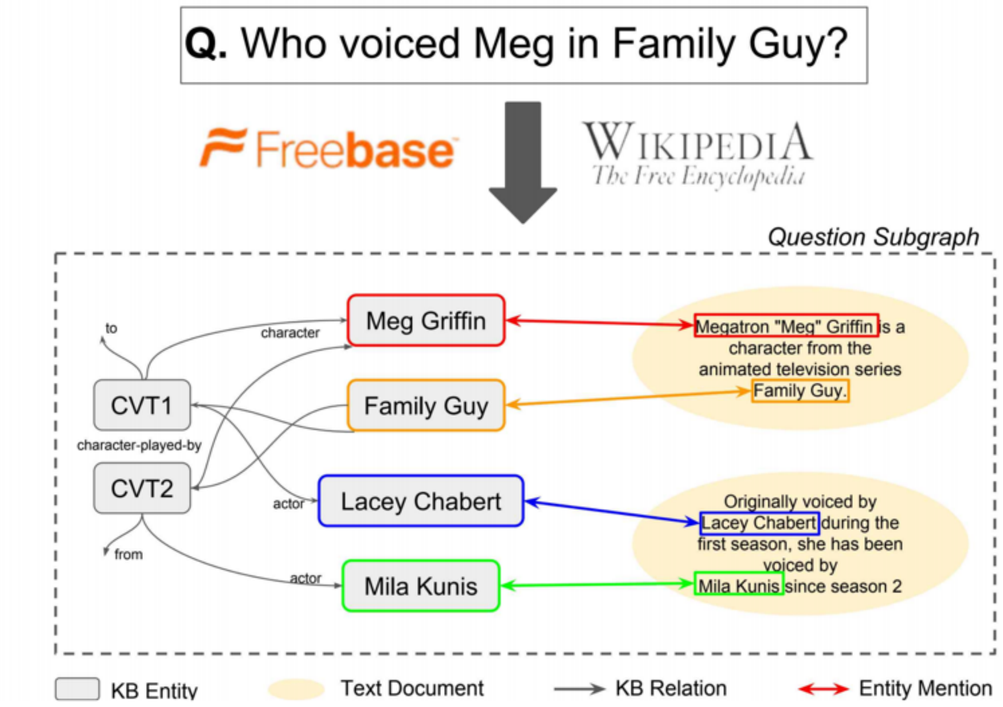
- Engineering motivation - QA performance
  - Text can complete missing information in KBs
  - KBs can provide background context for understanding text

Why Text + KBs?

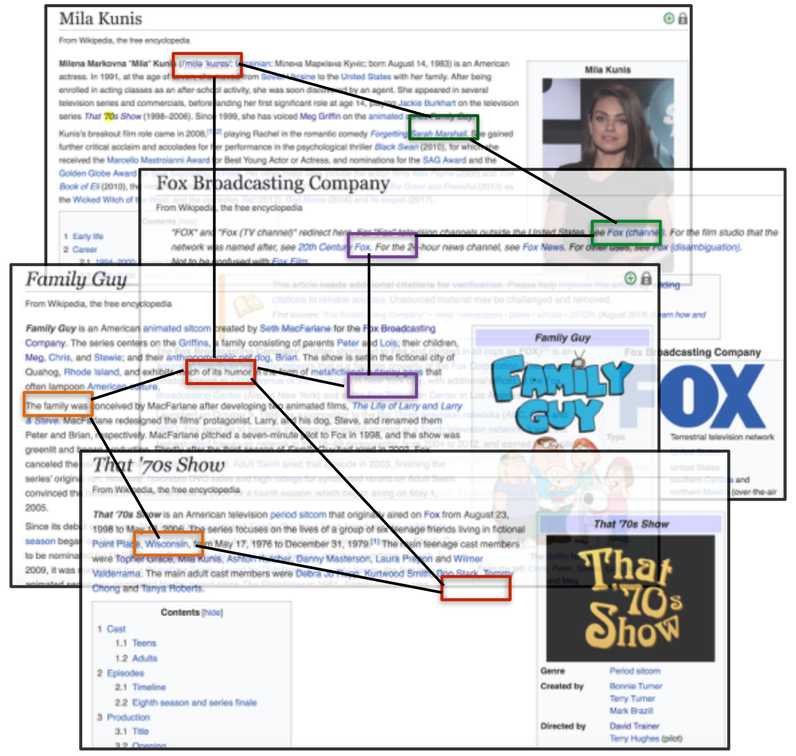
- Engineering motivation - QA performance
  - Text can complete missing information in KBs
  - KBs can provide background context for understanding text
- Scientific motivation - Knowledge Representation
  - Text is expressive
  - KBs support reasoning

This Talk

1. “Reading” heterogeneous graphs of facts and text [EMNLP’18]



2. Traversing text corpora like KBs [ongoing]



Open-domain QA using Early Fusion of KBs & Text

Haitian Sun*, Bhuwan Dhingra*, Manzil Zaheer,
Kathryn Mazaitis, Ruslan Salakhutdinov, William Cohen

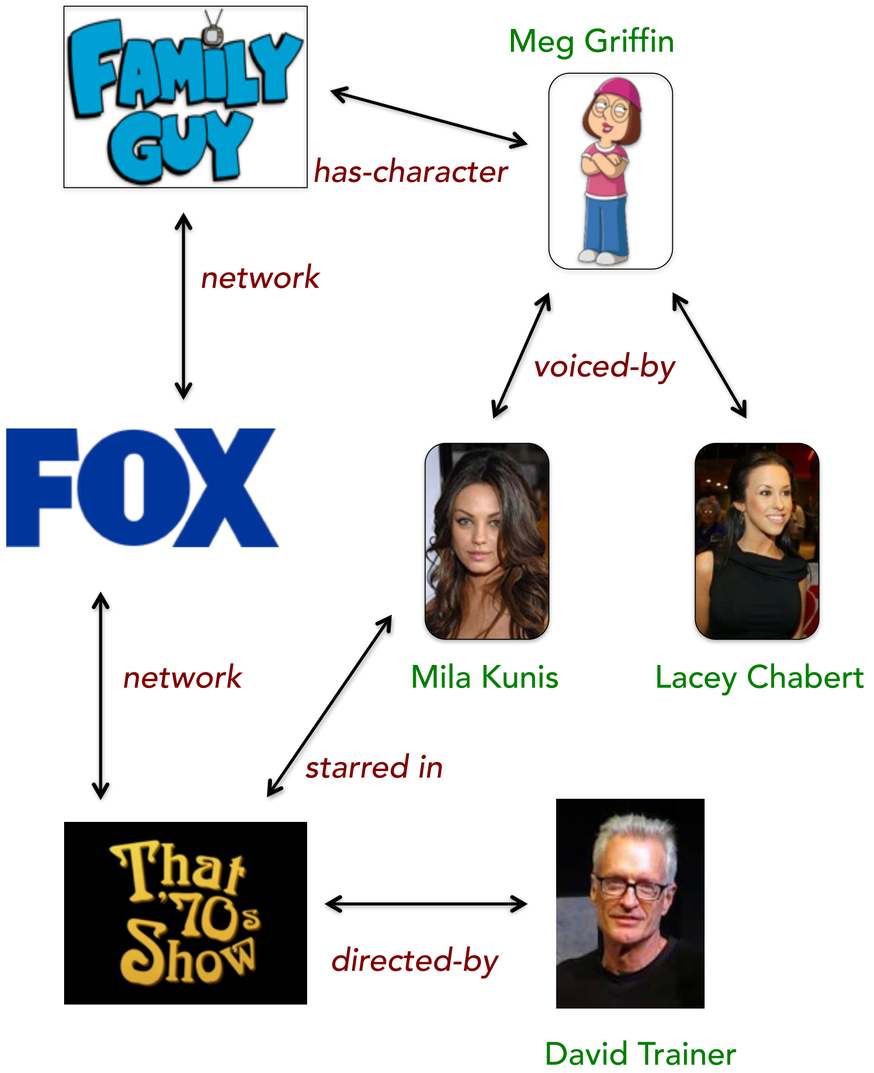
**equal contribution*

EMNLP 2018



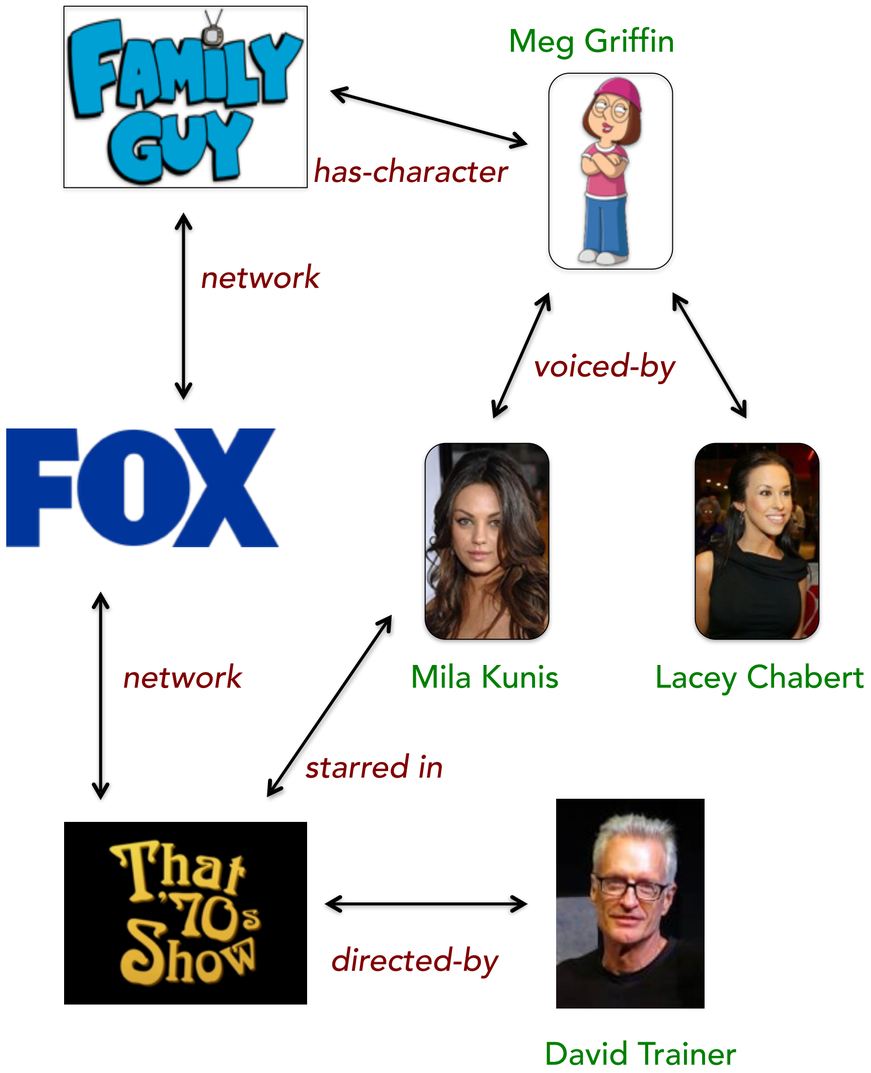


Entity-Relation Knowledge Bases



- Collection of *(subject, relation, object)* facts
- Organize information around *entity* nodes
- Freebase: 44M entities, >2B facts

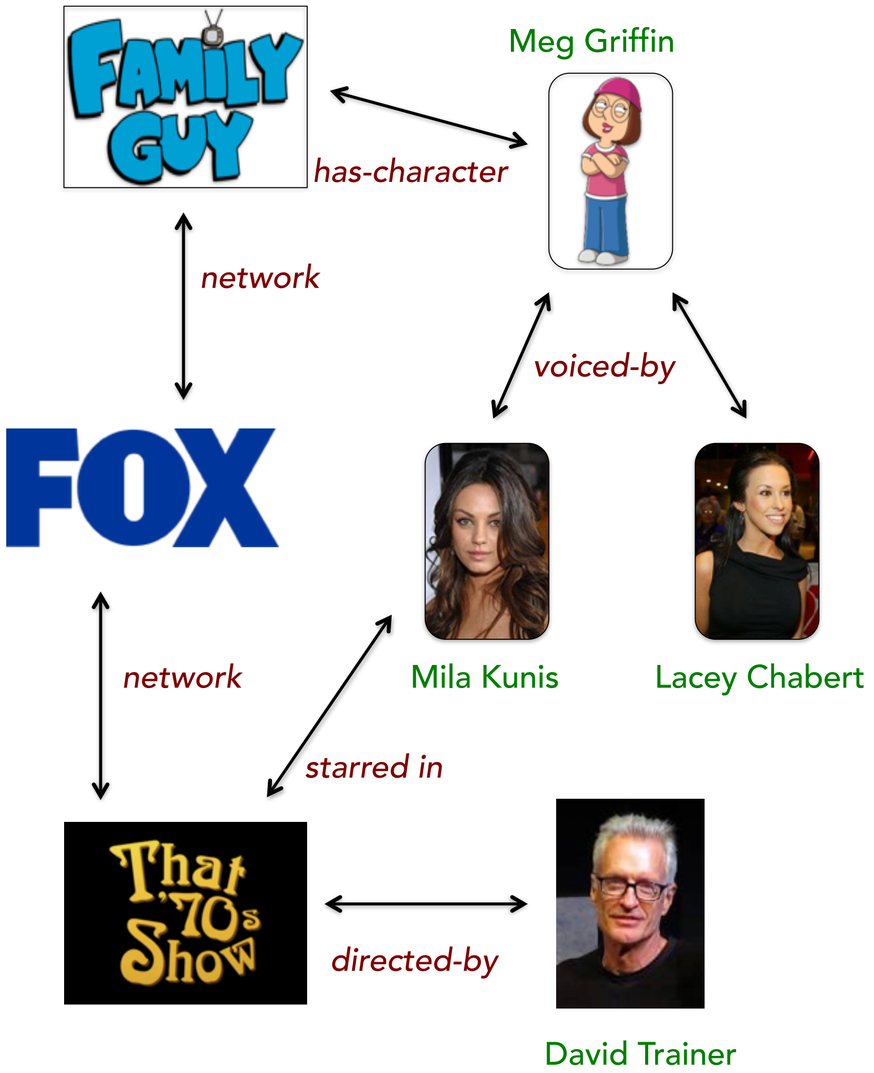
Reasoning in KBs



Relation following operation:

$$Y = X.\text{follow}(R) = \{x' : \exists x \in X \text{ s.t. } R(x, x') \text{ holds}\}$$

Reasoning in KBs

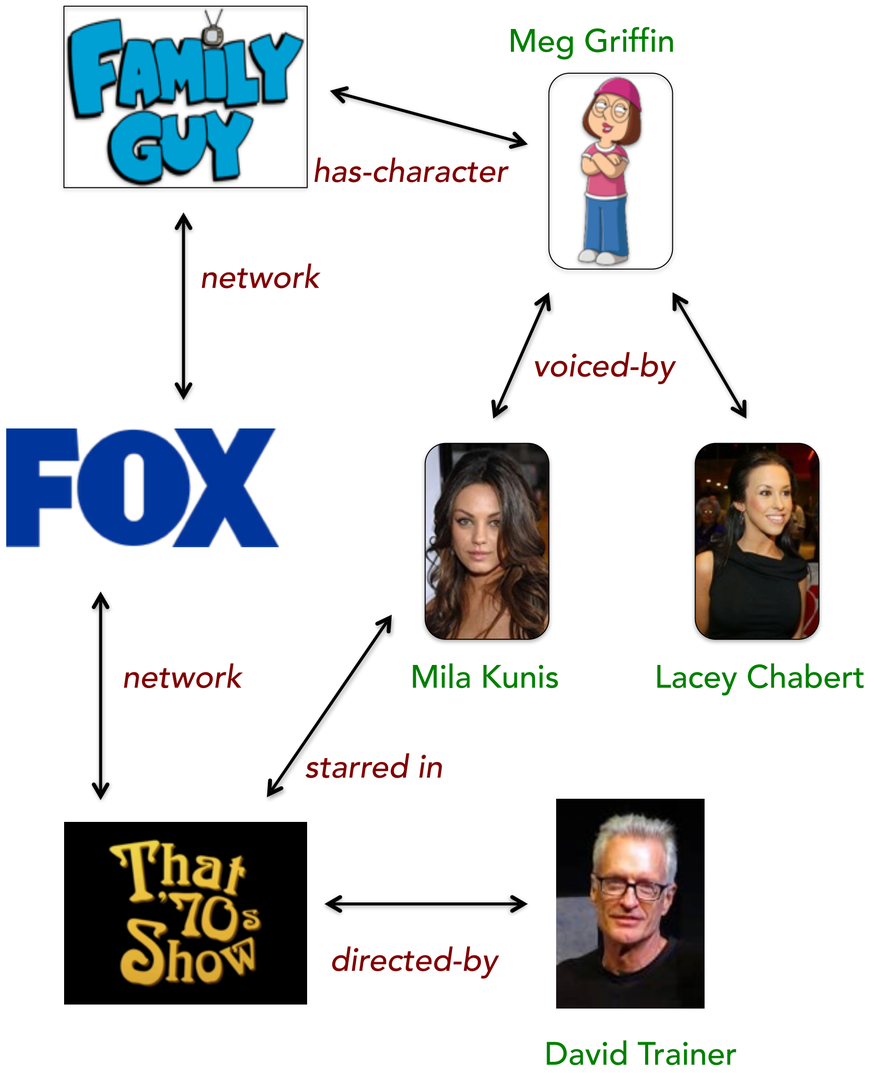


Relation following operation:

$$Y = X.\text{follow}(R) = \{x' : \exists x \in X \text{ s.t. } R(x, x') \text{ holds}\}$$

- Given a set of entities *X*

Reasoning in KBs

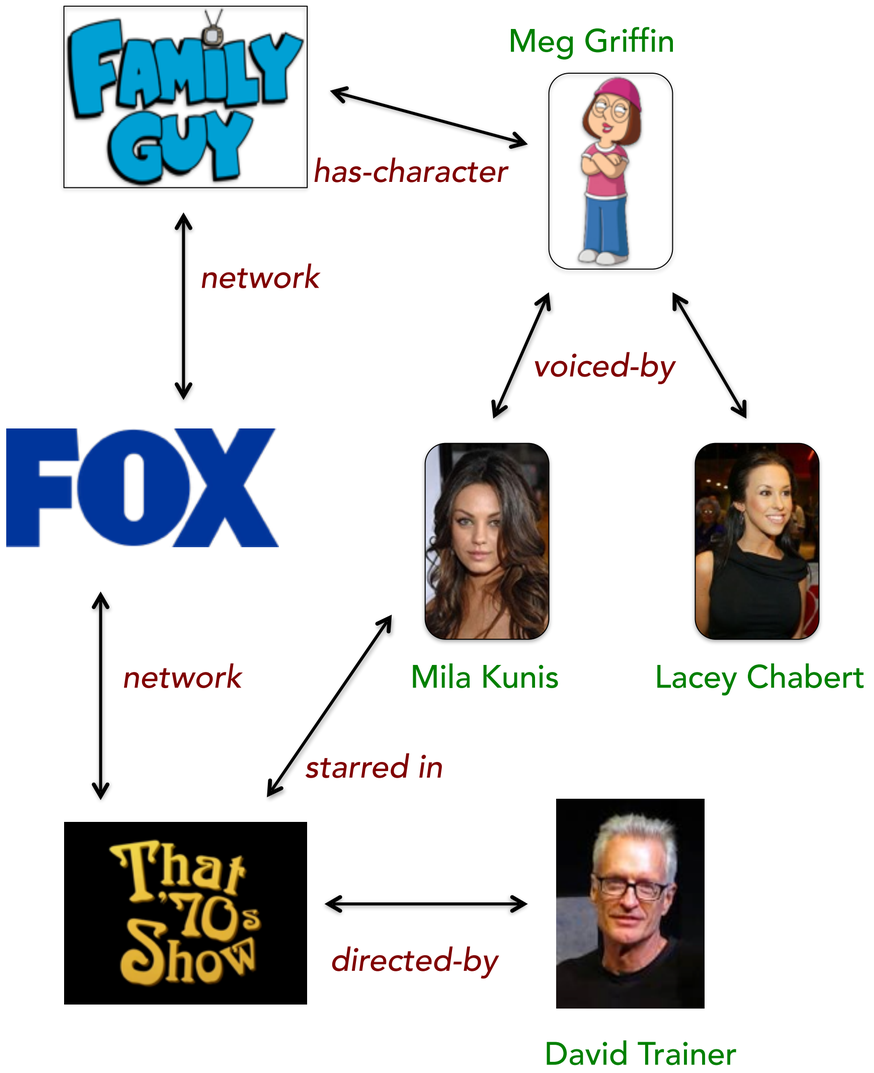


Relation following operation:

$$Y = X.\text{follow}(R) = \{x' : \exists x \in X \text{ s.t. } R(x, x') \text{ holds}\}$$

- Given a set of entities *X*
- Follow edges labeled with the relation *R*

Reasoning in KBs

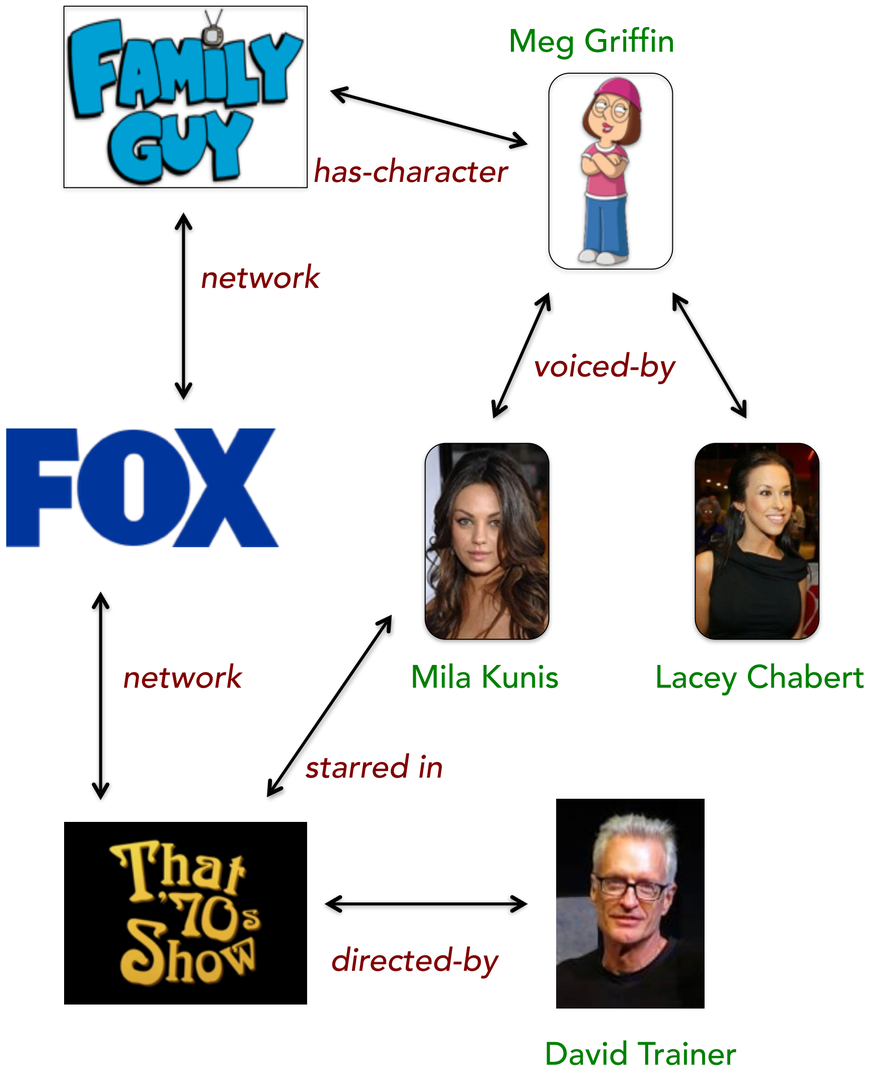


Relation following operation:

$$Y = X.\text{follow}(R) = \{x' : \exists x \in X \text{ s.t. } R(x, x') \text{ holds}\}$$

- Given a set of entities *X*
- Follow edges labeled with the relation *R*
- To arrive at a set of entities *Y*

Reasoning in KBs



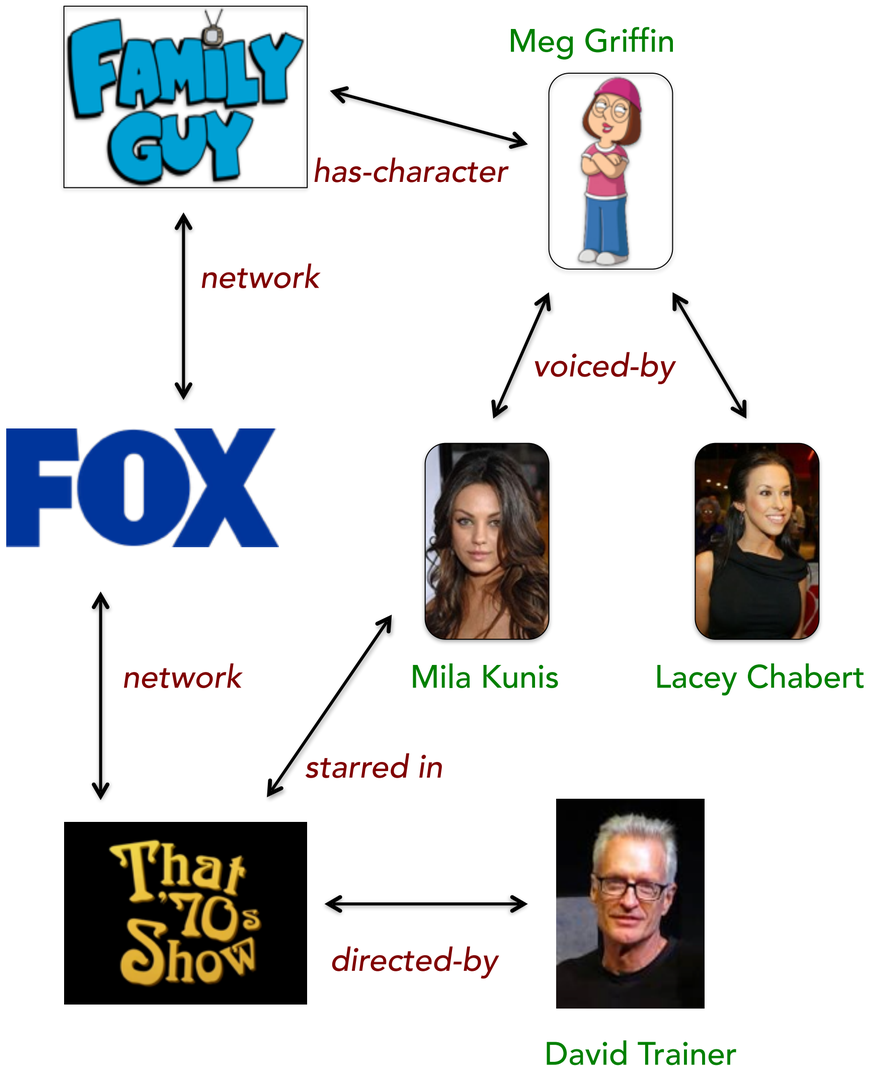
Relation following operation:

$$Y = X.\text{follow}(R) = \{x' : \exists x \in X \text{ s.t. } R(x, x') \text{ holds}\}$$

- Given a set of entities *X*
- Follow edges labeled with the relation *R*
- To arrive at a set of entities *Y*

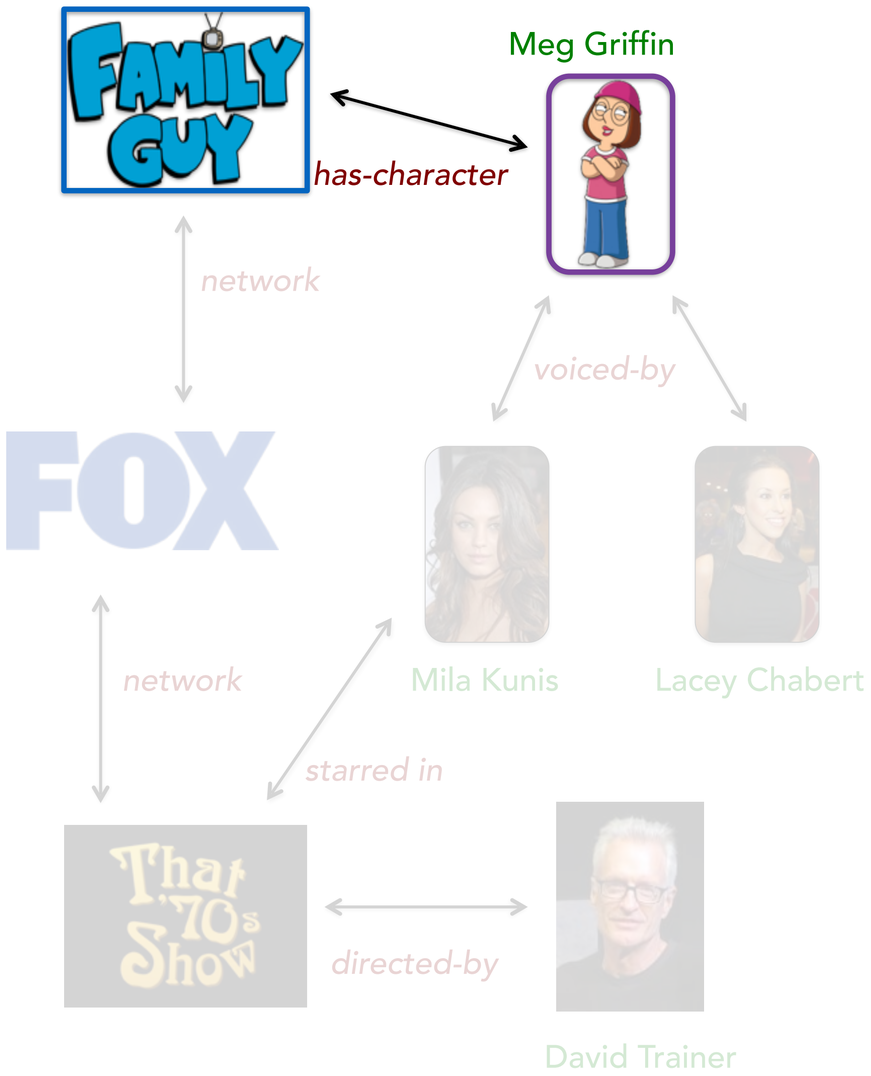
E.g. {Family_Guy, That_70s_Show}.follow(network) = {Fox}

Question Answering



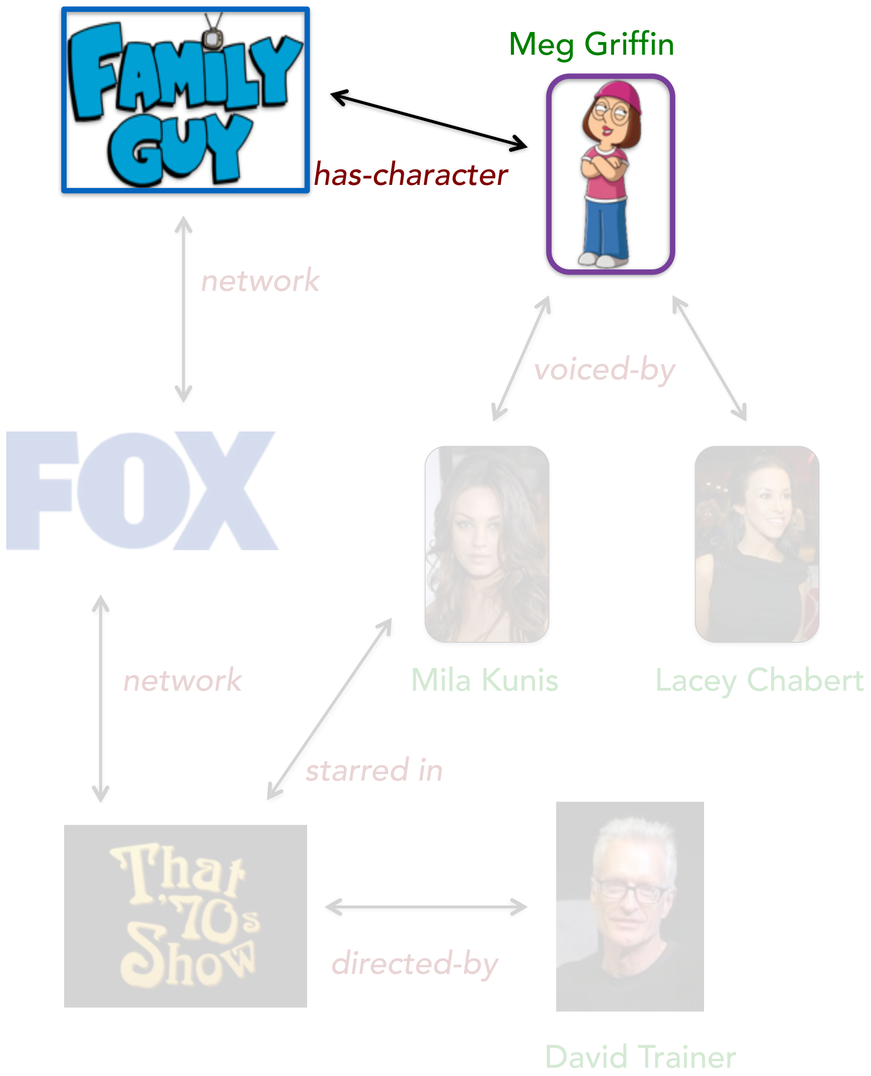
Who voiced Meg in Family Guy?

Question Answering



Who voiced Meg in Family Guy?

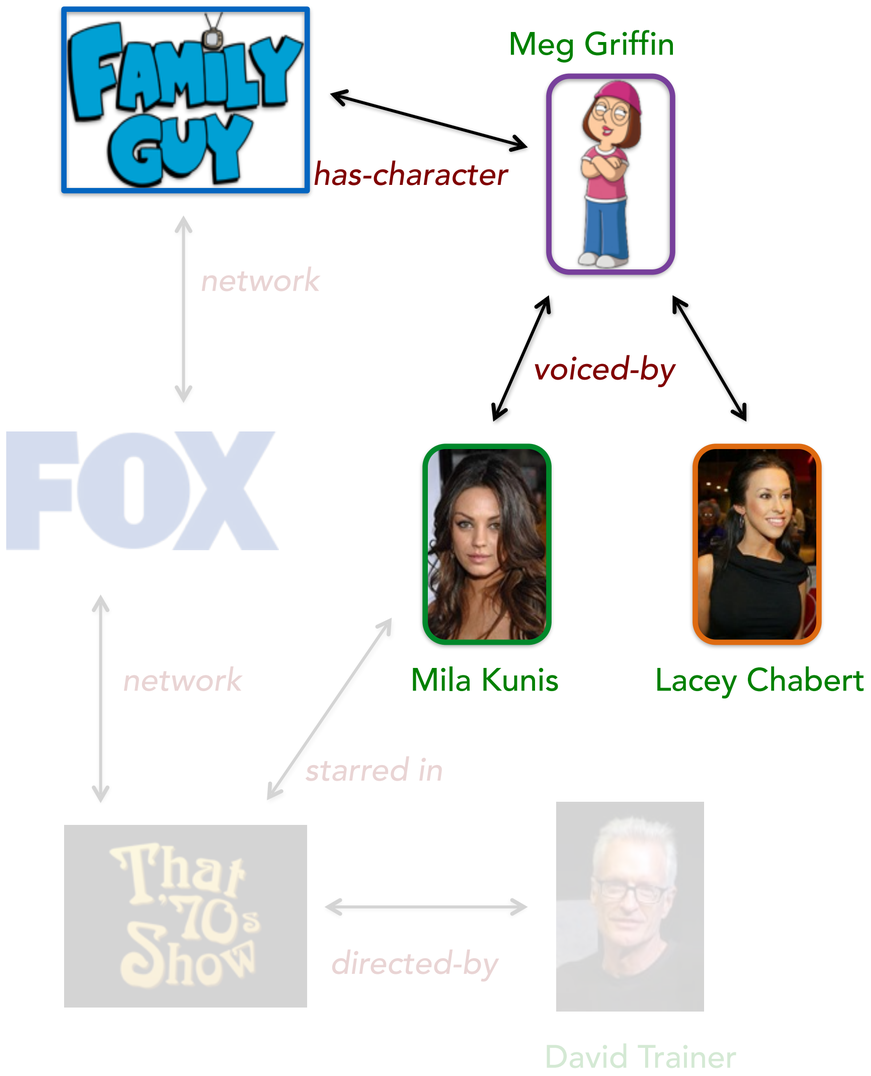
Question Answering



Who voiced Meg in Family Guy?

```
{Meg_Griffin}.follow(voiced-by)
```

Question Answering

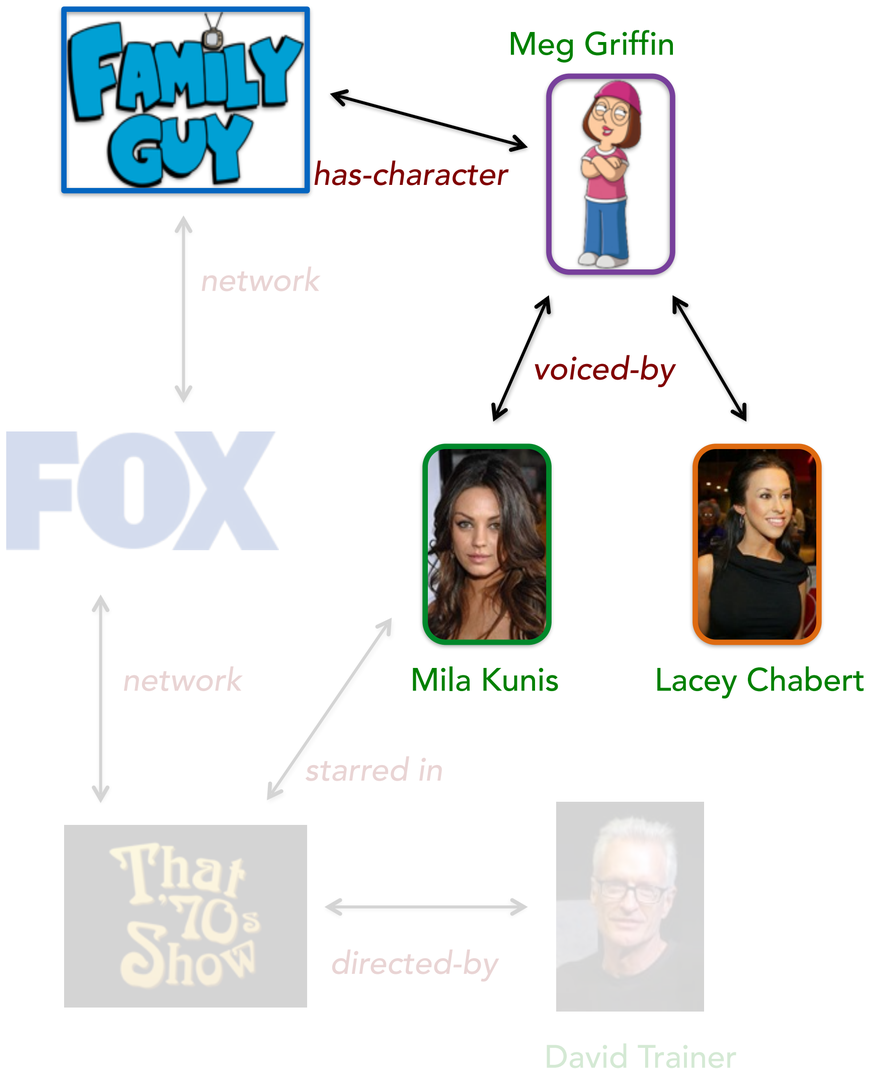


Who voiced Meg in Family Guy?

{Meg_Griffin}.follow(voiced-by)

{Mila_Kunis, Lacey_Chabert}

Question Answering



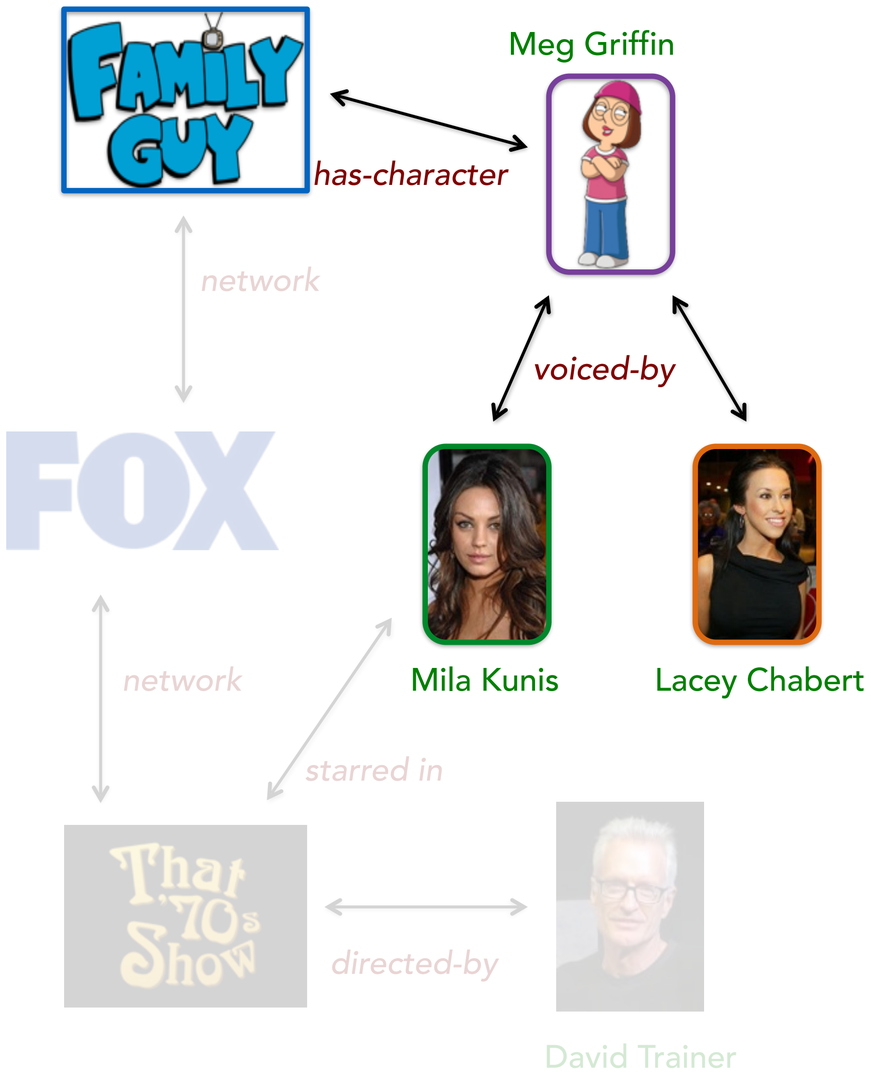
Who voiced Meg in Family Guy?

Annotating semantic parses of questions is expensive

```
{Meg_Griffin}.follow(voiced-by)
```

```
{Mila_Kunis, Lacey_Chabert}
```

Question Answering



Who voiced Meg in Family Guy?

Search for parses which lead to the correct answer in the KB

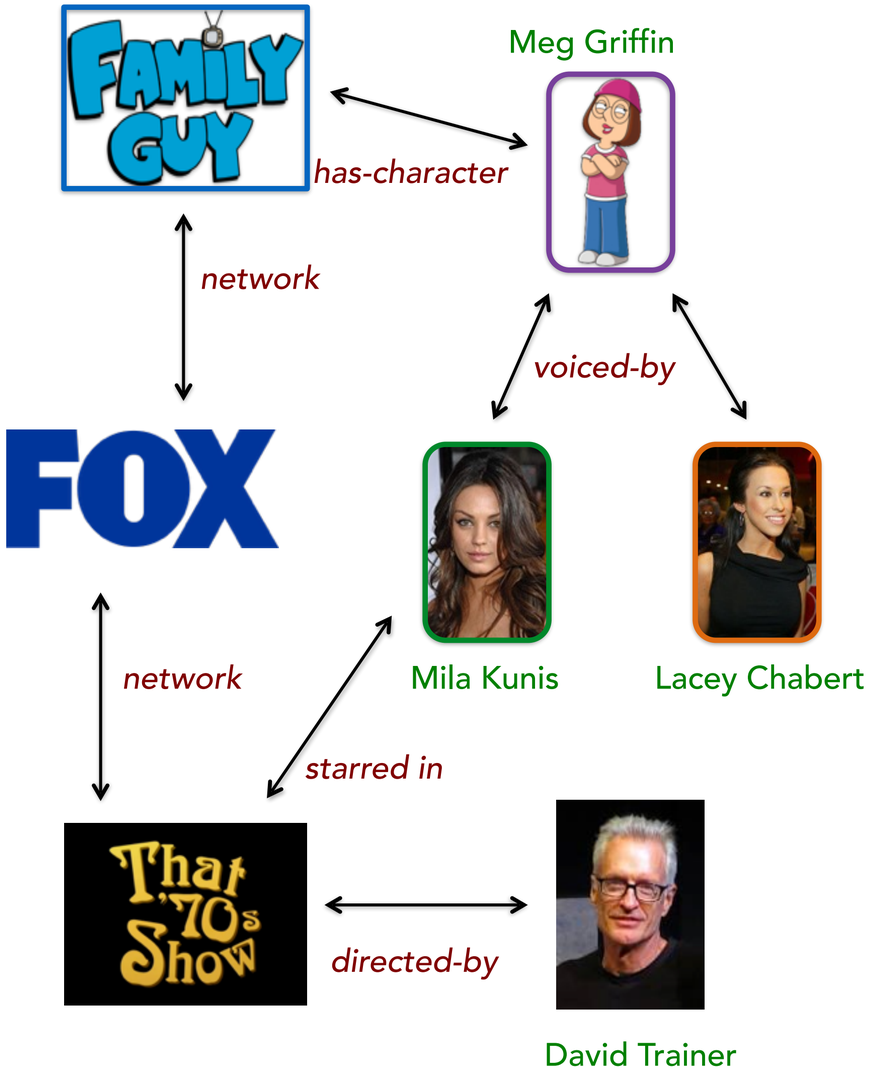
???

{Mila_Kunis, Lacey_Chabert}

Learning from Denotations

[Liang, 2011], [Berant, 2013], [Reddy, 2014], [Krishnamurthy, 2017]

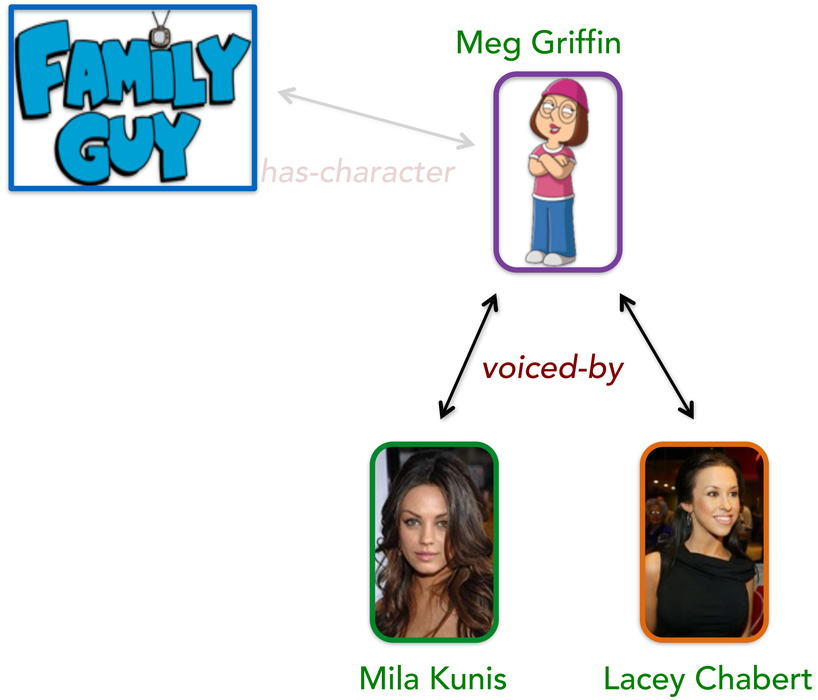
But KBs are often incomplete



Freebase relation types	Incompleteness
/people/person/education	0.792
/people/person/employment_history	0.923
/people/person/nationality*	0.785
/people/person/parents*	0.988
/people/person/place_of_birth*	0.938
/people/person/places_lived*	0.966

Min et al, NAACL 2013

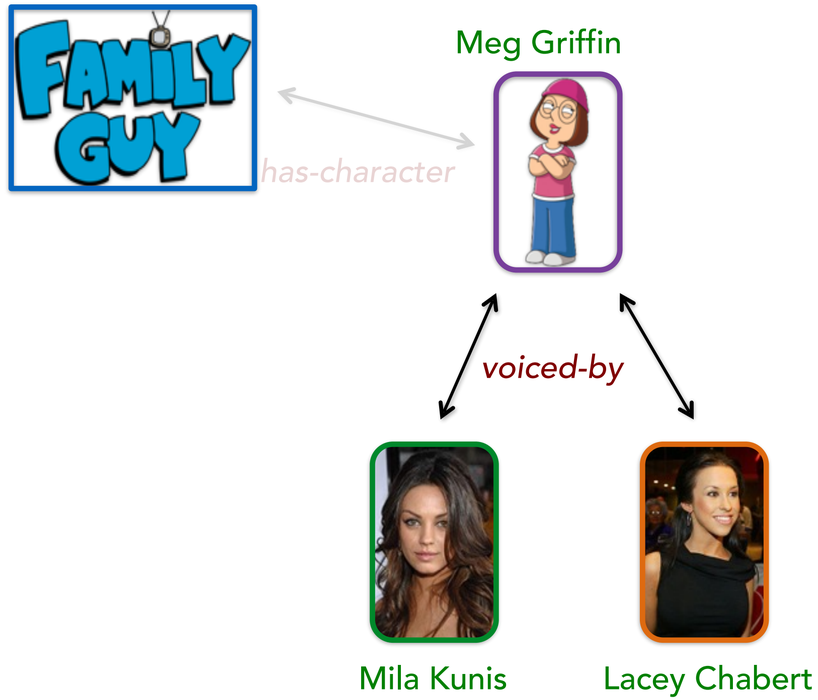
Graphs of Facts and Text



Who voiced Meg in Family Guy?

WIKIPEDIA
The Free Encyclopedia

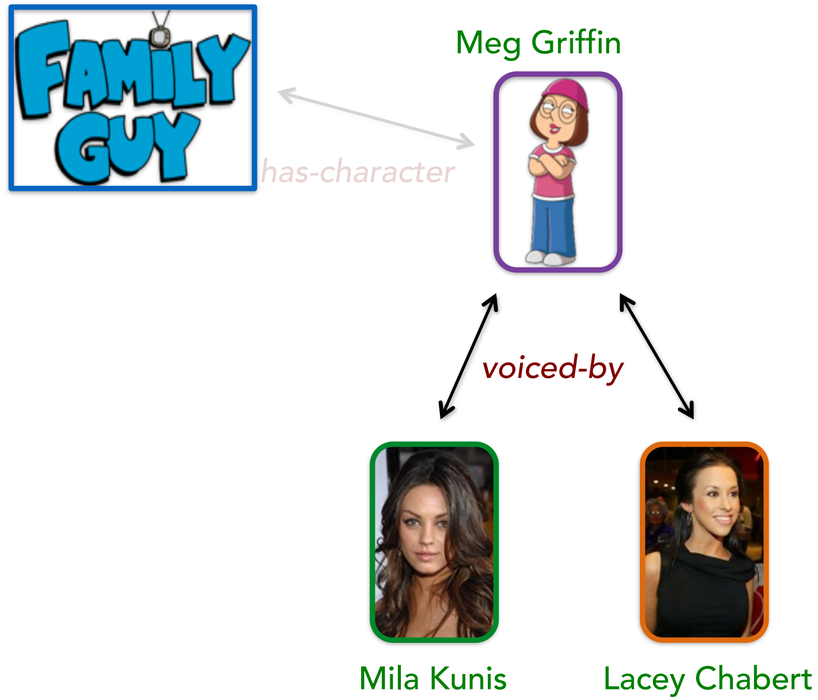
Graphs of Facts and Text



Inject relevant text into the graph

Who voiced Meg in Family Guy?

Graphs of Facts and Text



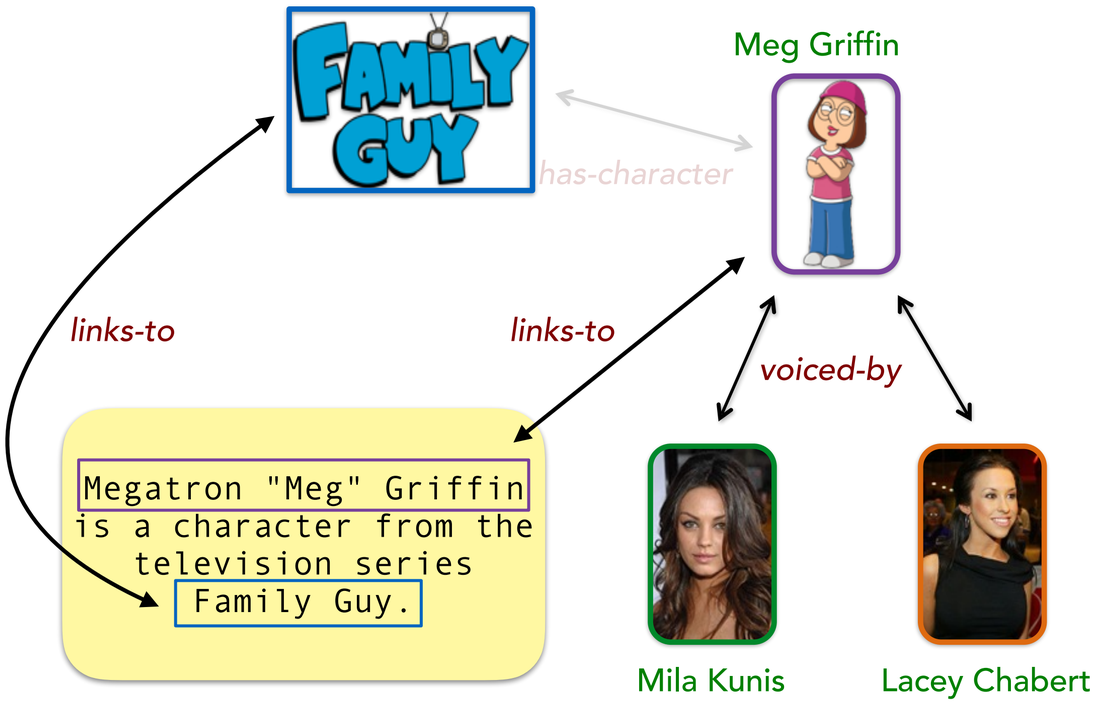
Inject relevant text into the graph

Who voiced Meg in Family Guy?

TF-IDF Retrieval

WIKIPEDIA
The Free Encyclopedia

Graphs of Facts and Text



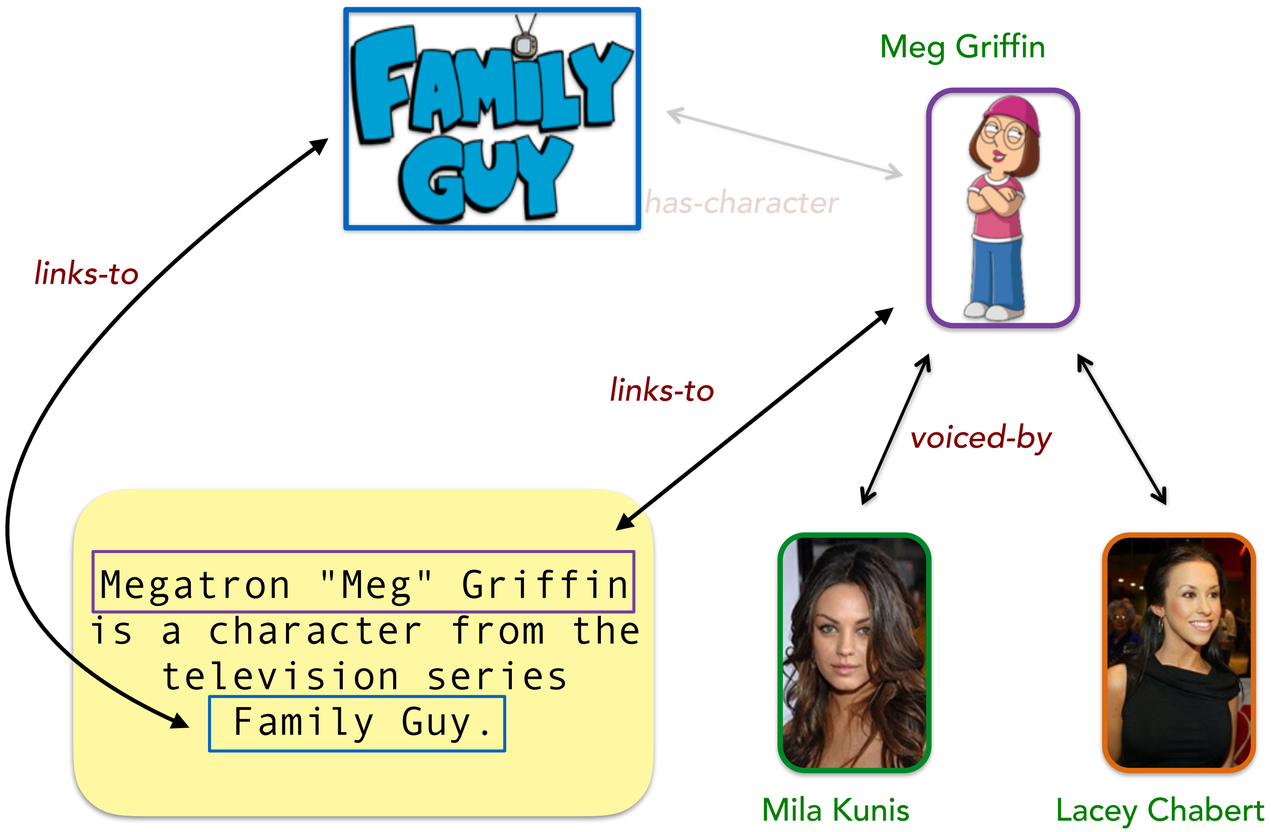
Inject relevant text into the graph

Who voiced Meg in Family Guy?

TF-IDF Retrieval

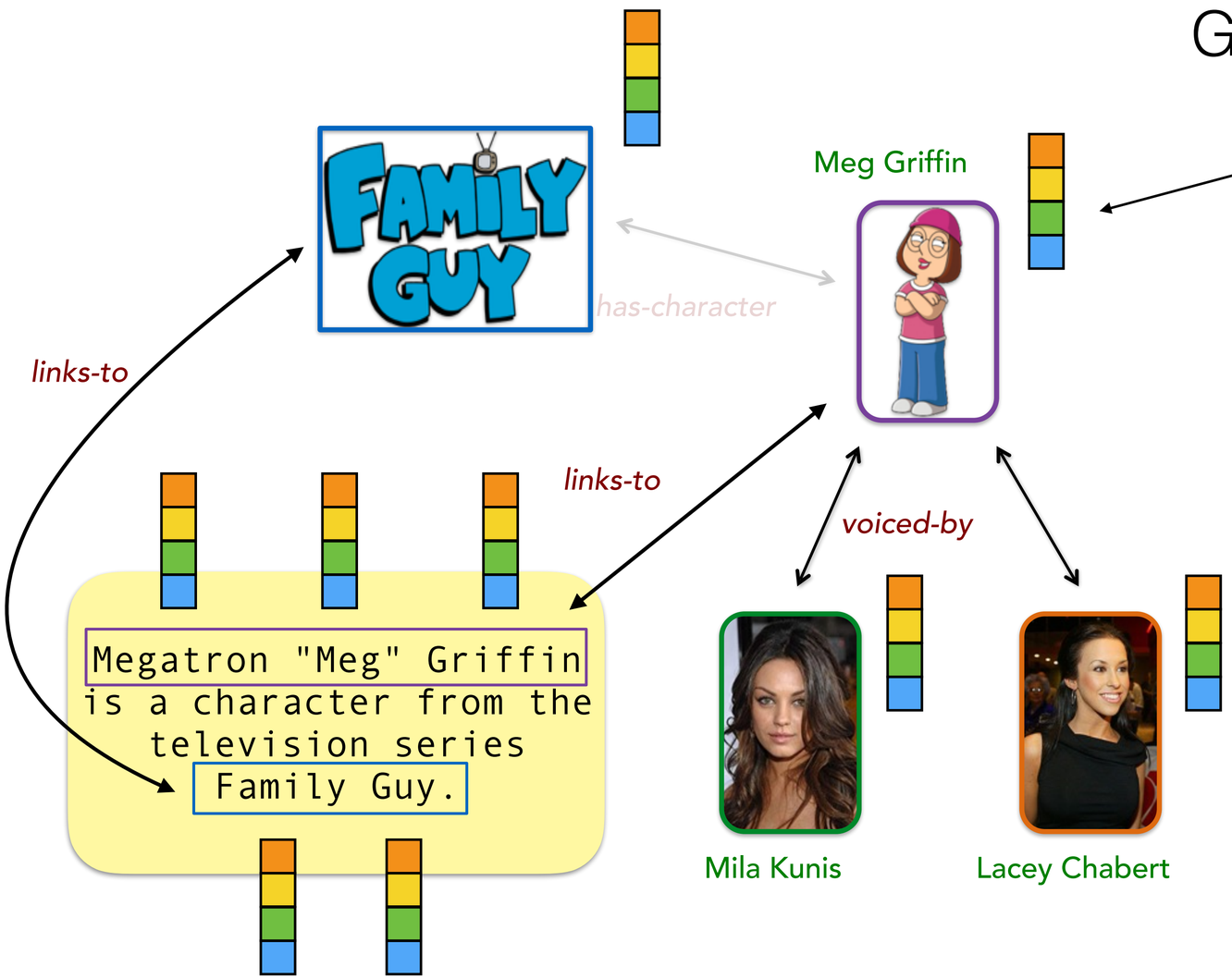
Entity Linking

Search via Representation Learning



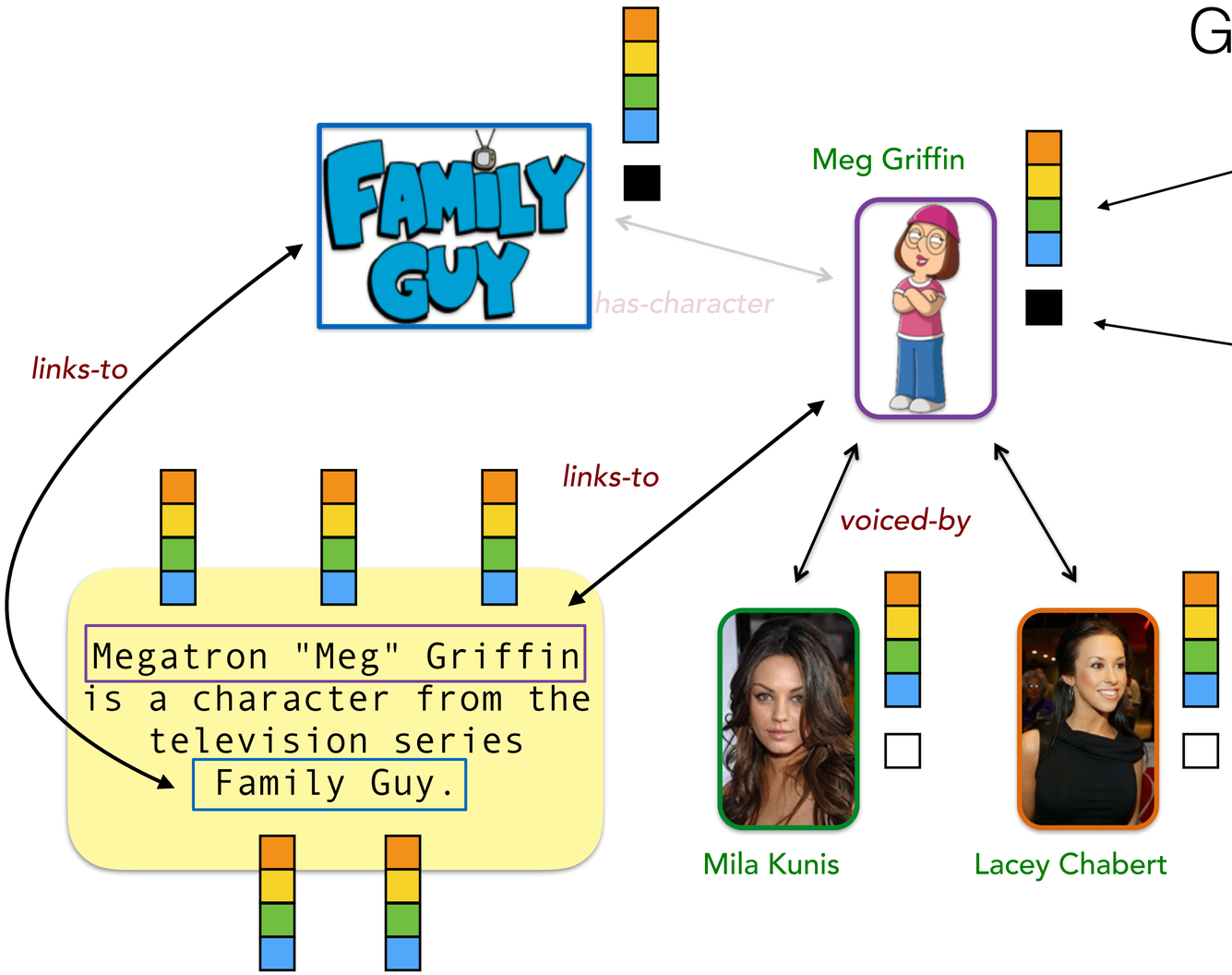
Who voiced Meg in Family Guy?

Search via Representation Learning



Who voiced Meg in Family Guy?

Search via Representation Learning



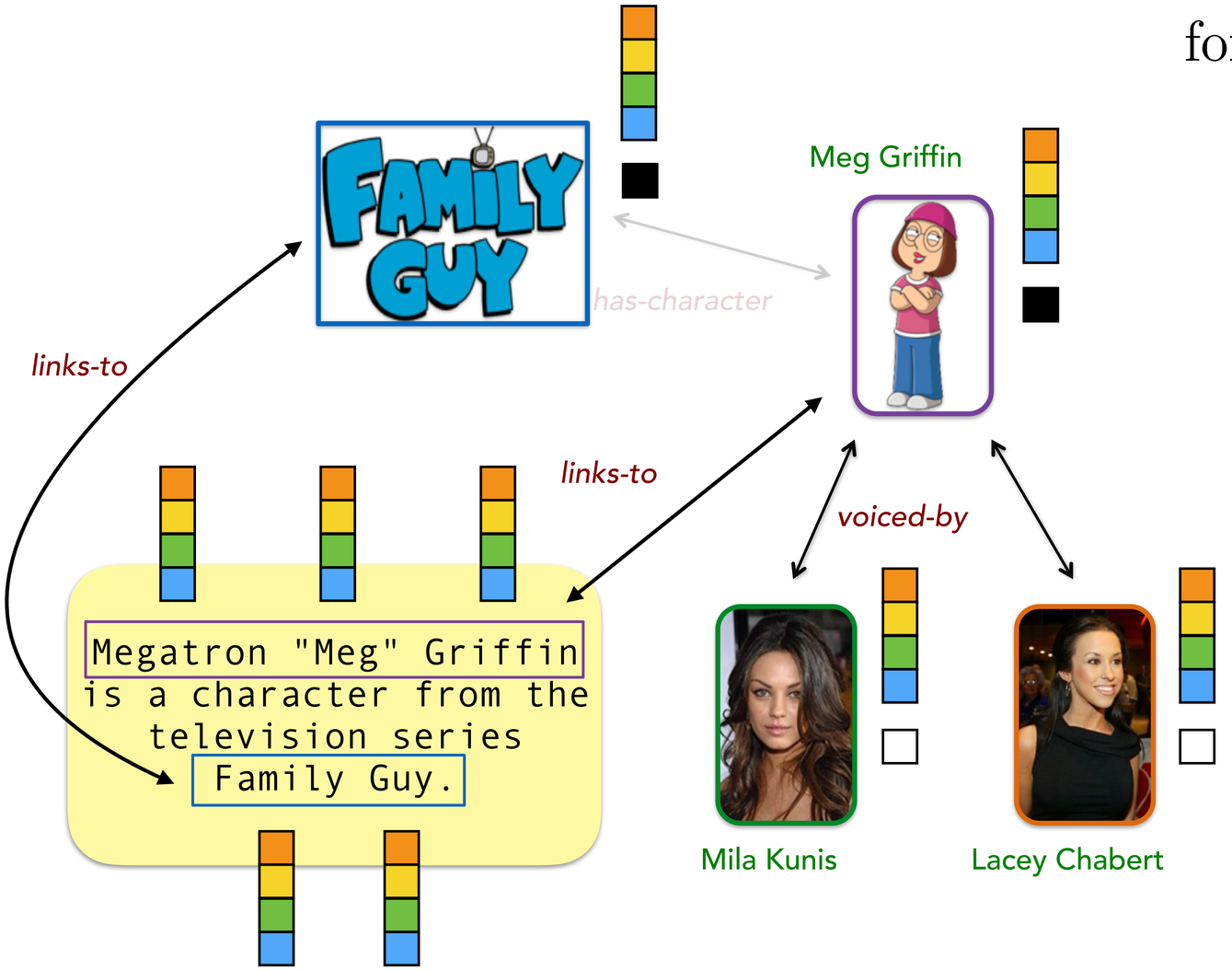
Graph Neural Networks

Embedding Vector

"Pagerank" score
- Initially uniform over entities in the question

Who voiced Meg in Family Guy?

Search via Representation Learning

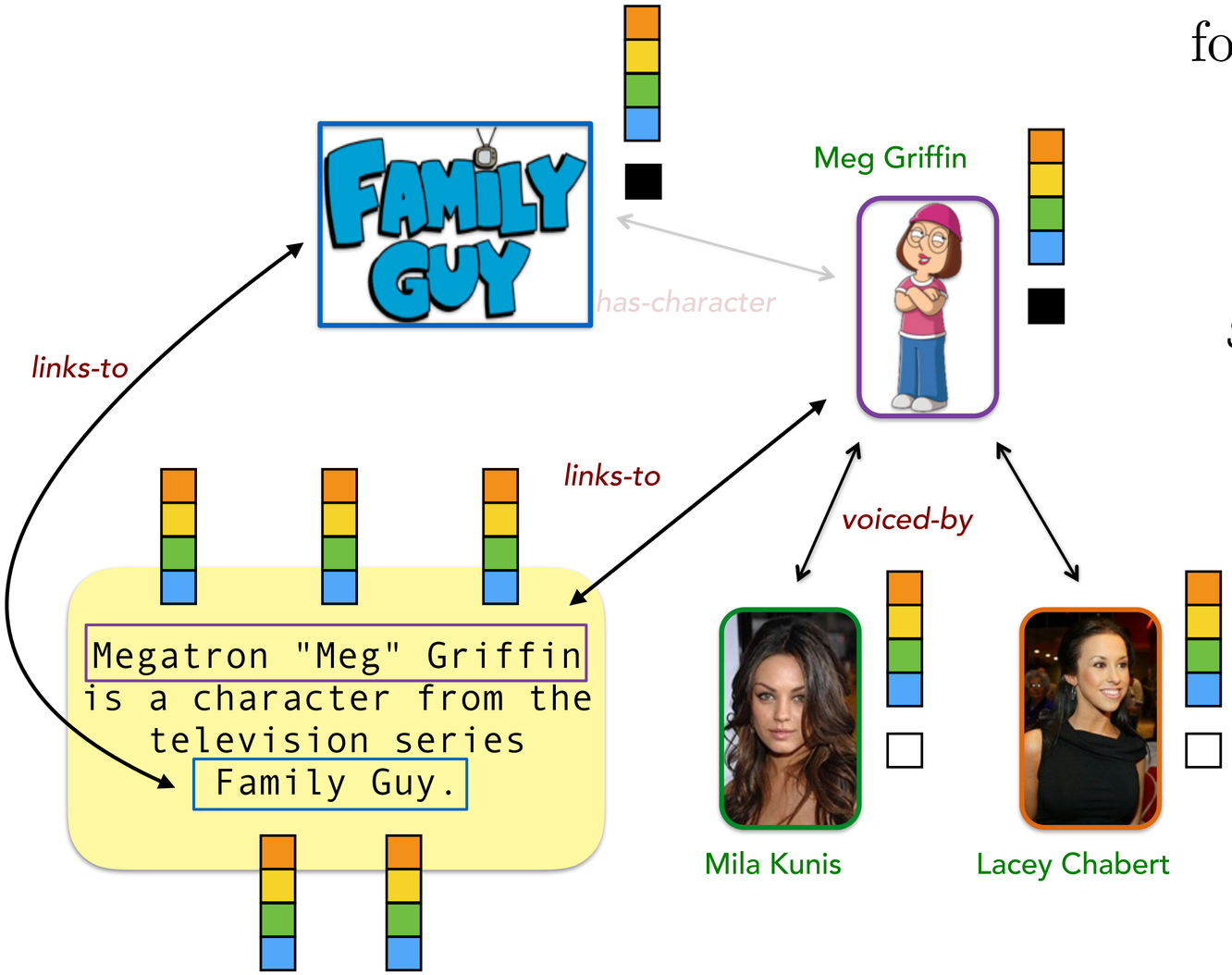


$$\text{for } t = 1 \dots T:$$

$$h_v^{(t)} = f(h_v^{(t-1)}, \sum_{v' \in N(v)} \alpha_{v'} h_{v'}^{(t-1)}, \sum_{d: v \in d} h_{d_w}^{(t-1)})$$

Who voiced Meg in Family Guy?

Search via Representation Learning



$$\text{for } t = 1 \dots T:$$

$$h_v^{(t)} = f(h_v^{(t-1)}, \sum_{v' \in N(v)} \alpha_{v'} h_{v'}^{(t-1)}, \sum_{d: v \in d} h_{d_w}^{(t-1)})$$

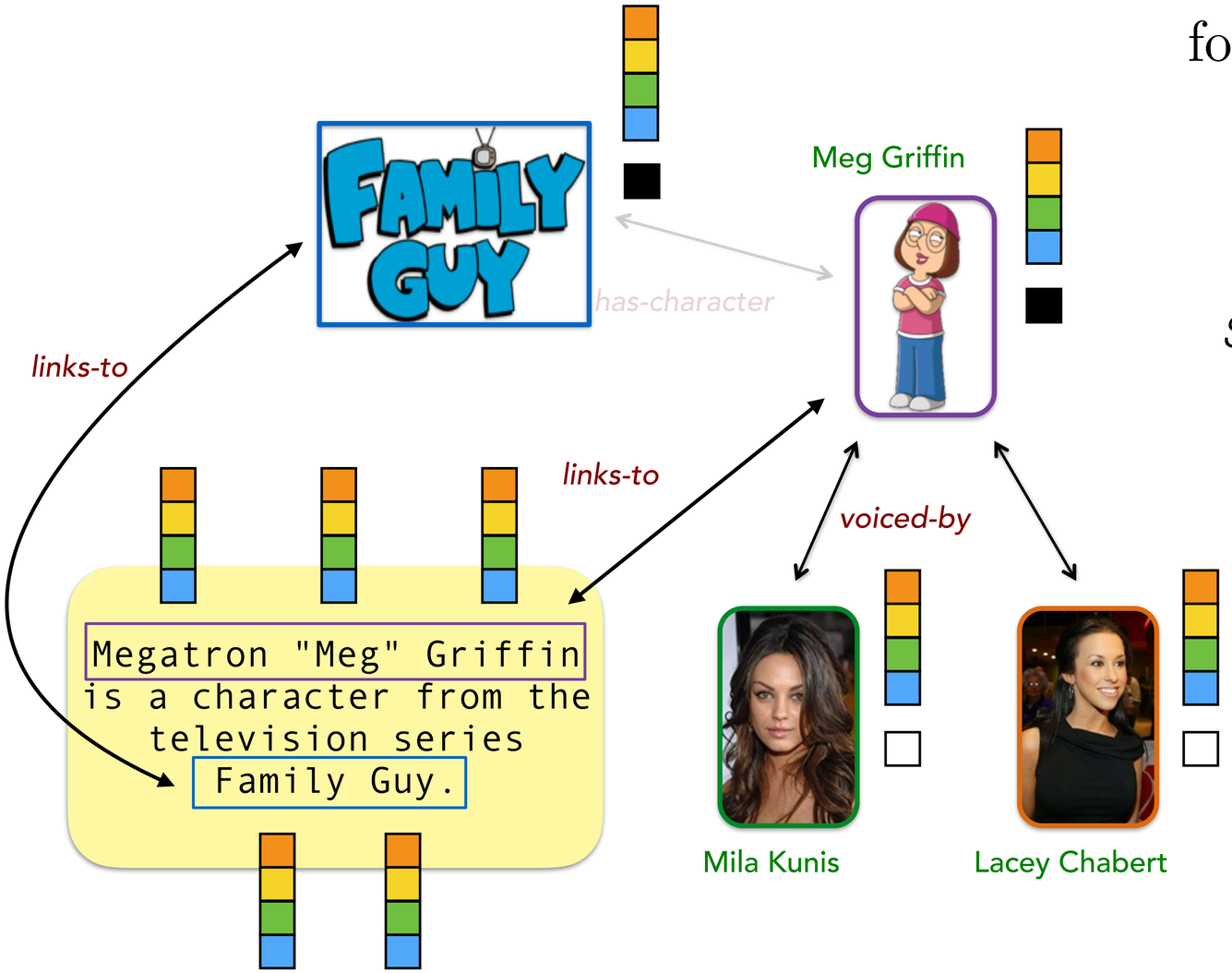
Self connection

Entity neighbors

Text mentions

Who voiced Meg in Family Guy?

Search via Representation Learning



$$\text{for } t = 1 \dots T:$$

$$h_v^{(t)} = f(h_v^{(t-1)}, \sum_{v' \in N(v)} \alpha_{v'} h_{v'}^{(t-1)}, \sum_{d:v \in d} h_{d_w}^{(t-1)})$$

Self connection

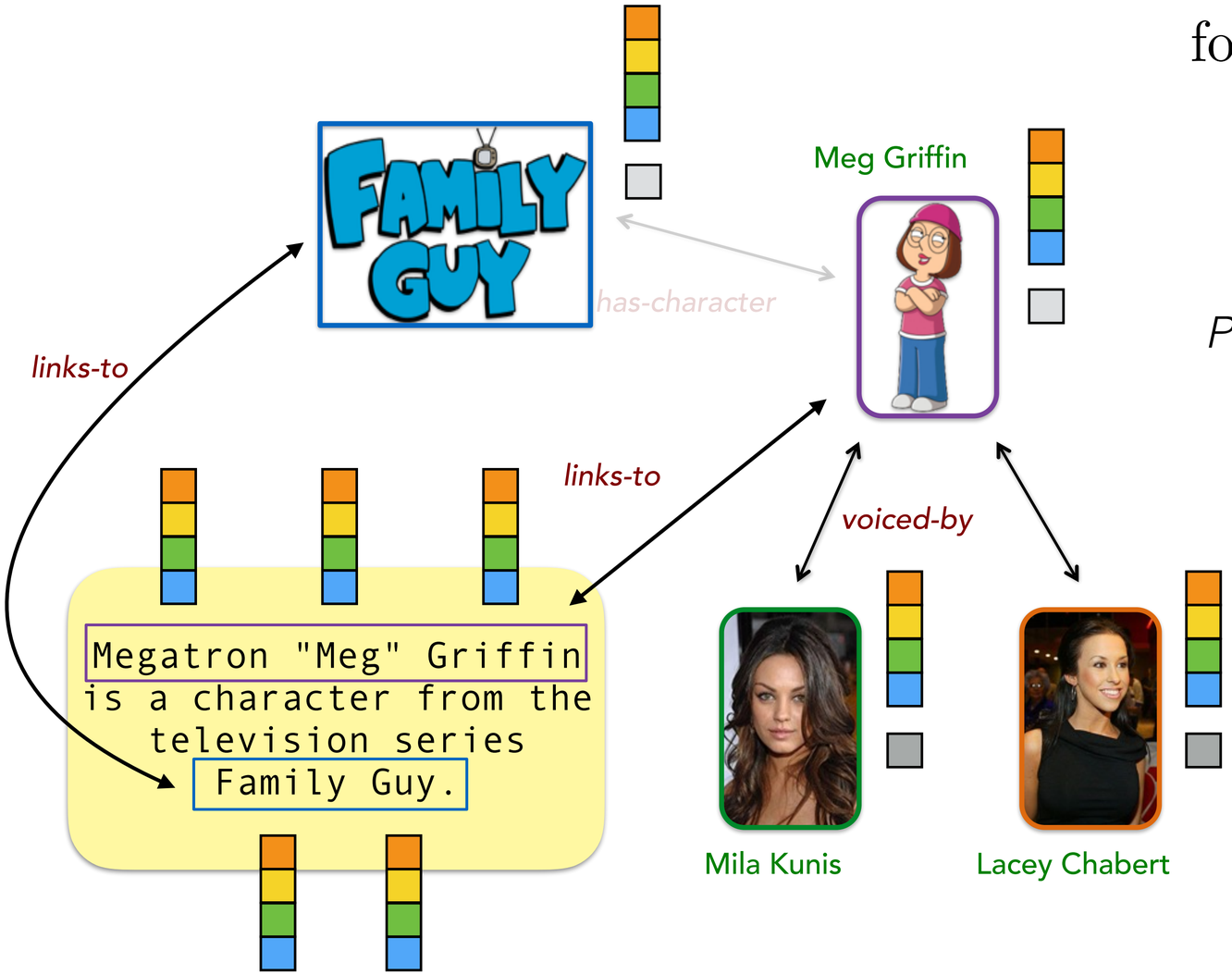
Entity neighbors

Text mentions

- Only propagate embeddings from nodes with **non-zero pagerank score**
- This constrains learning along **valid paths** in the graph

Who voiced Meg in Family Guy?

Search via Representation Learning



$$\text{for } t = 1 \dots T:$$

$$pr_v^{(t)} = (1-\lambda) pr_v^{(t-1)} + \lambda \sum_{v' \in N_r(v)} \alpha_{v'} pr_{v'}^{(t-1)}$$

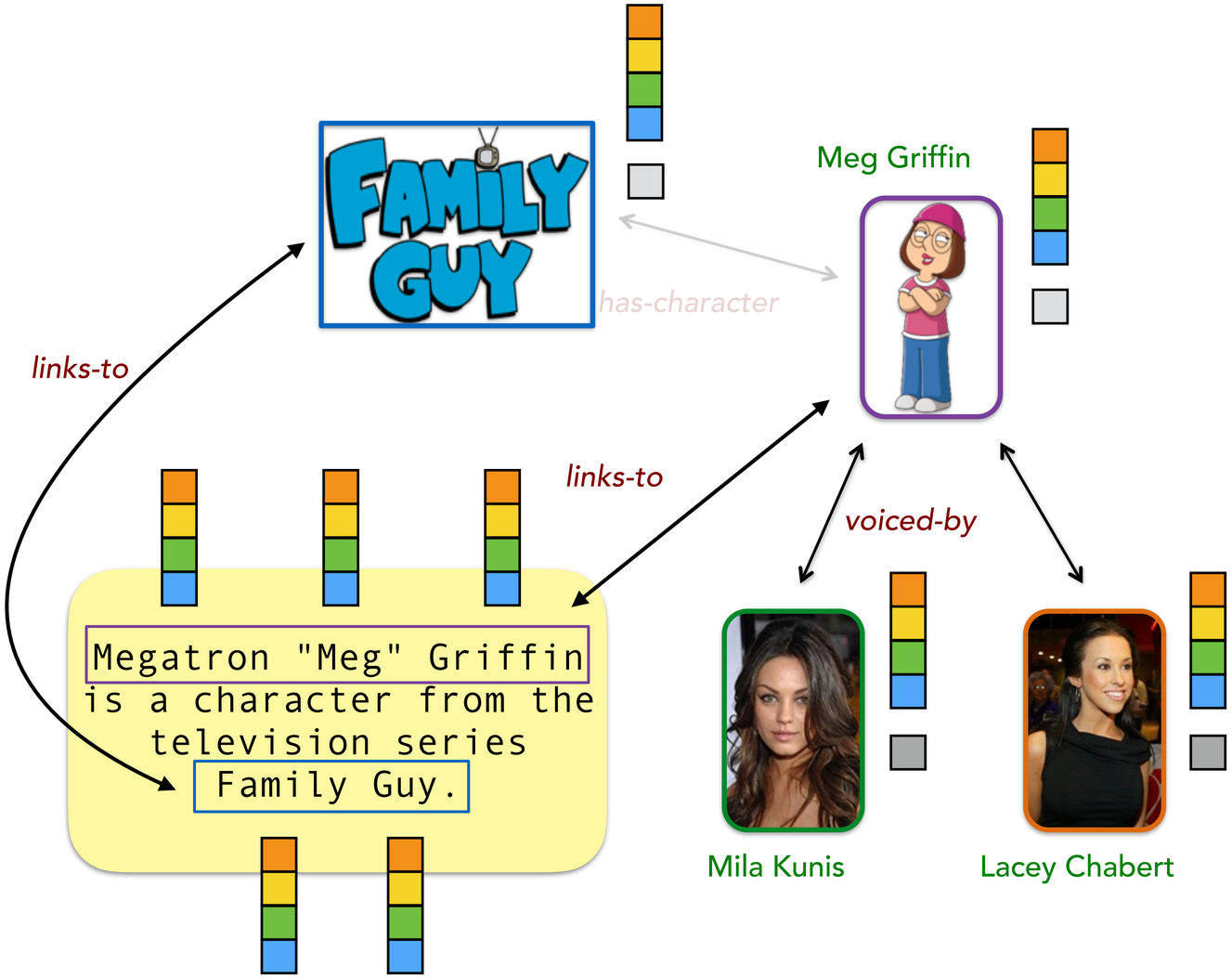
Pagerank update

Learned weights

- Only propagate embeddings from nodes with **non-zero pagerank score**
- This constrains learning along **valid paths** in the graph

Who voiced Meg in Family Guy?

Search via Representation Learning

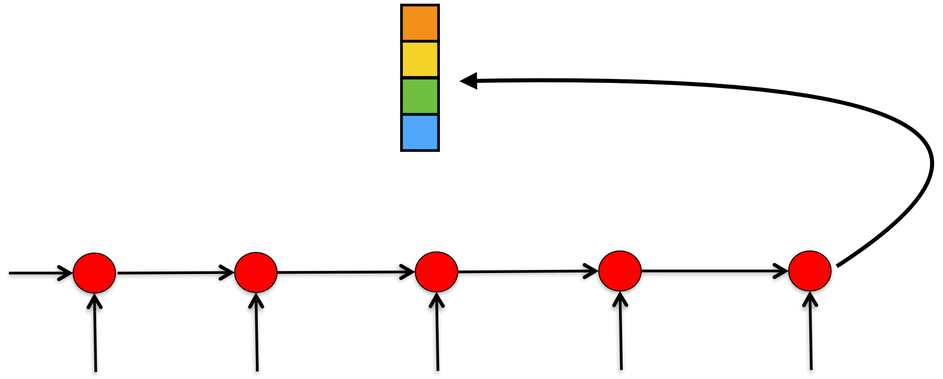


Classify each entity as Answer / Not-Answer:

$$\log Pr(v|q) \propto (h_q \cdot h_v^T)$$

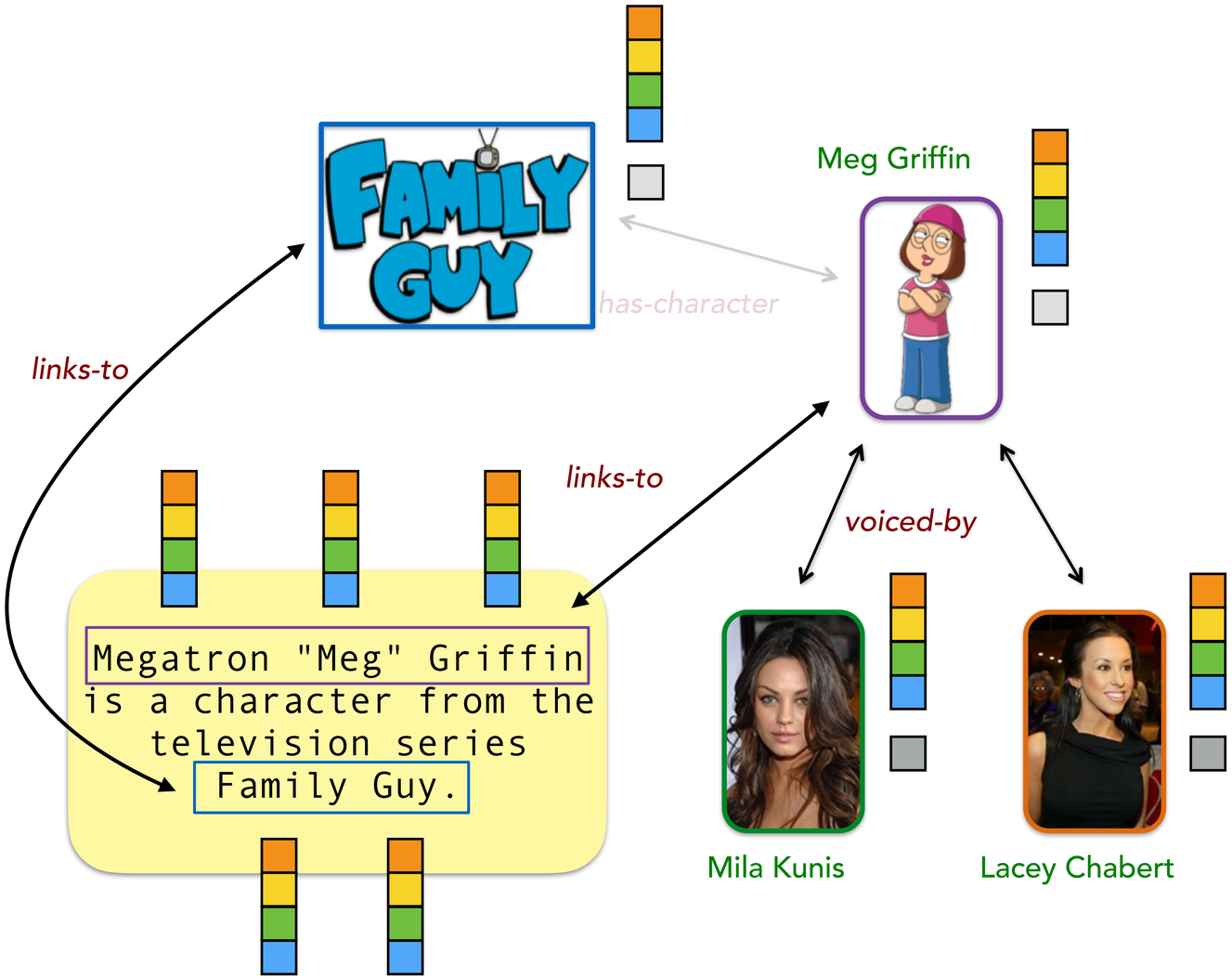
h_q : Question Representation

h_v^T : Entity Representation



Who voiced Meg in Family Guy?

Search via Representation Learning



Summary:

1. Inject text into KB using retrieval
2. Propagate entity embeddings along paths starting from question entities
3. Classify each entity as answer / no-answer

Who voiced Meg in Family Guy?

Evaluation — WebQuestionsSP & WikiMovies

Evaluation — WebQuestionsSP & WikiMovies

- WebQuestionsSP [Berant, 2013; Yih, 2016]:
  E.g. *"What language do they speak in Afghanistan?" – Pashto language*

  - KB - Freebase, Text - Wikipedia

  - 5K QA pairs for training

Evaluation — WebQuestionsSP & WikiMovies

- WebQuestionsSP [Berant, 2013; Yih, 2016]:
  E.g. *"What language do they speak in Afghanistan?" – Pashto language*
  - KB - Freebase, Text - Wikipedia
  - 5K QA pairs for training
- Wikimovies [Miller, 2016]:
  E.g. *"What movies did Quentin Tarantino direct?" – Reservoir dogs, Pulp fiction, …*
  - KB - Subset of OMDB, Text - Subset of Wikipedia
  - 10K QA pairs for training

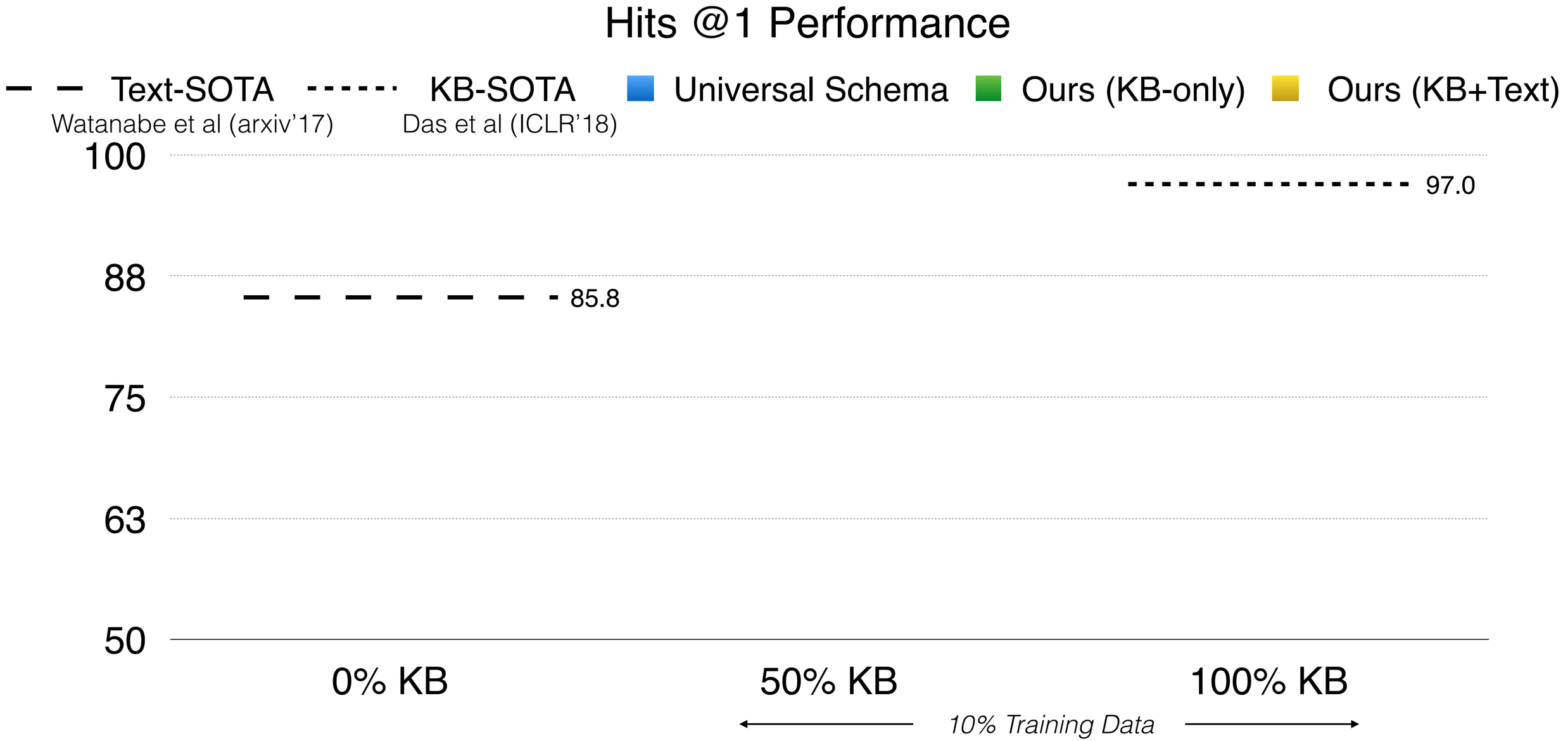
Evaluation — WebQuestionsSP & WikiMovies

- WebQuestionsSP [Berant, 2013; Yih, 2016]:
  E.g. *"What language do they speak in Afghanistan?" – Pashto language*
  - KB - Freebase, Text - Wikipedia
  - 5K QA pairs for training
- Wikimovies [Miller, 2016]:
  E.g. *"What movies did Quentin Tarantino direct?" – Reservoir dogs, Pulp fiction, …*
  - KB - Subset of OMDB, Text - Subset of Wikipedia
  - 10K QA pairs for training
- Both datasets are answerable using KBs
  - But we simulate an incomplete KB setting

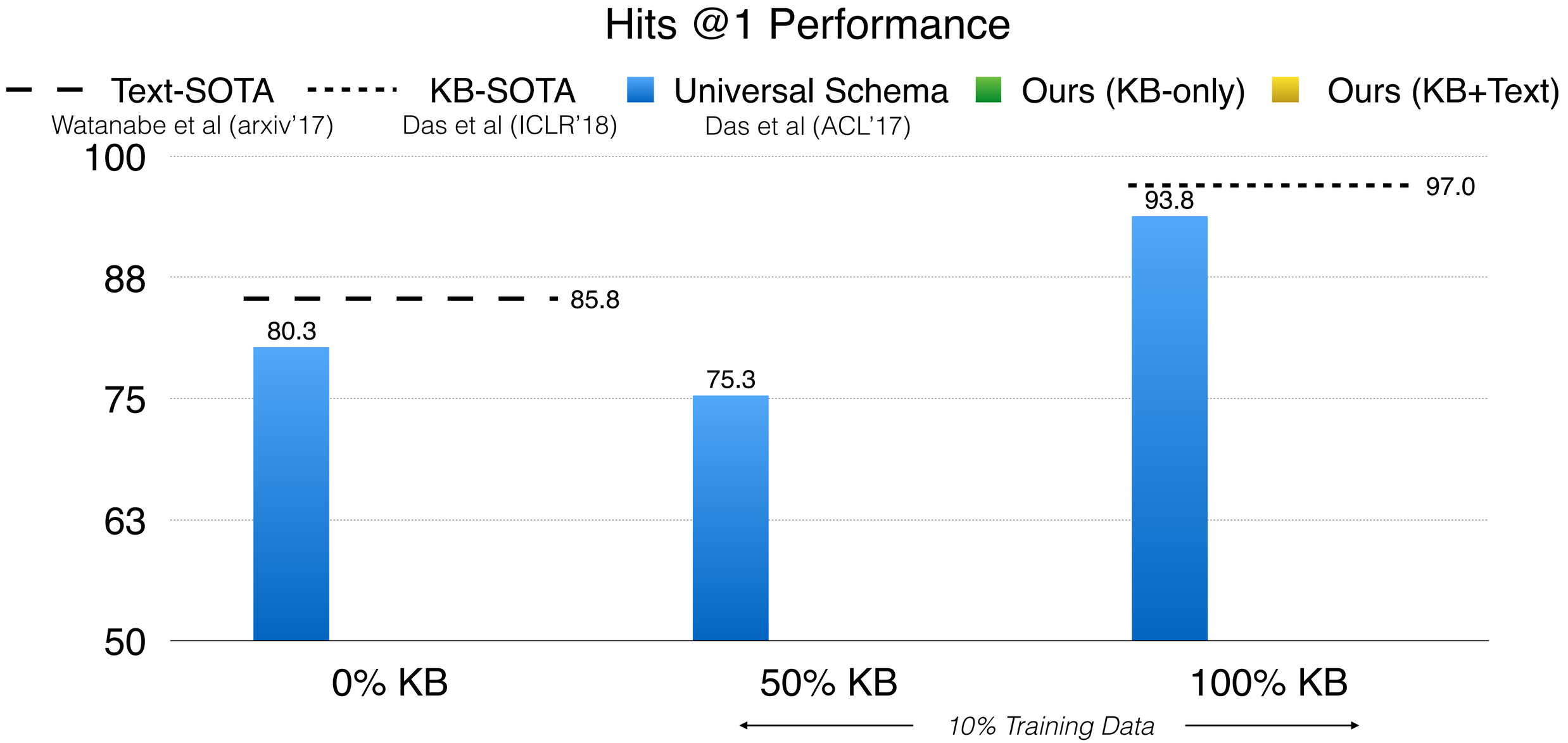
WikiMovies Results

Hits @1 Performance

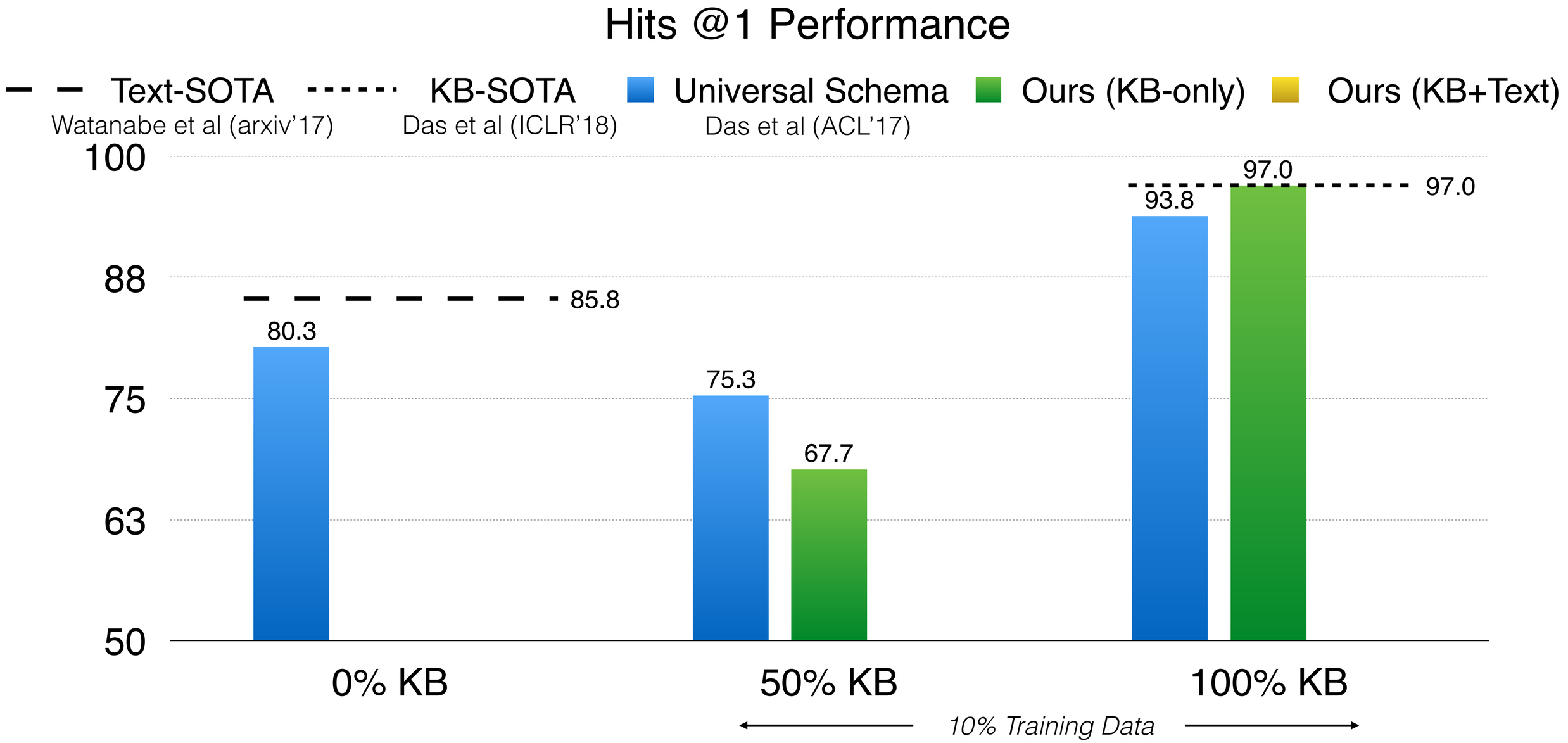
WikiMovies Results



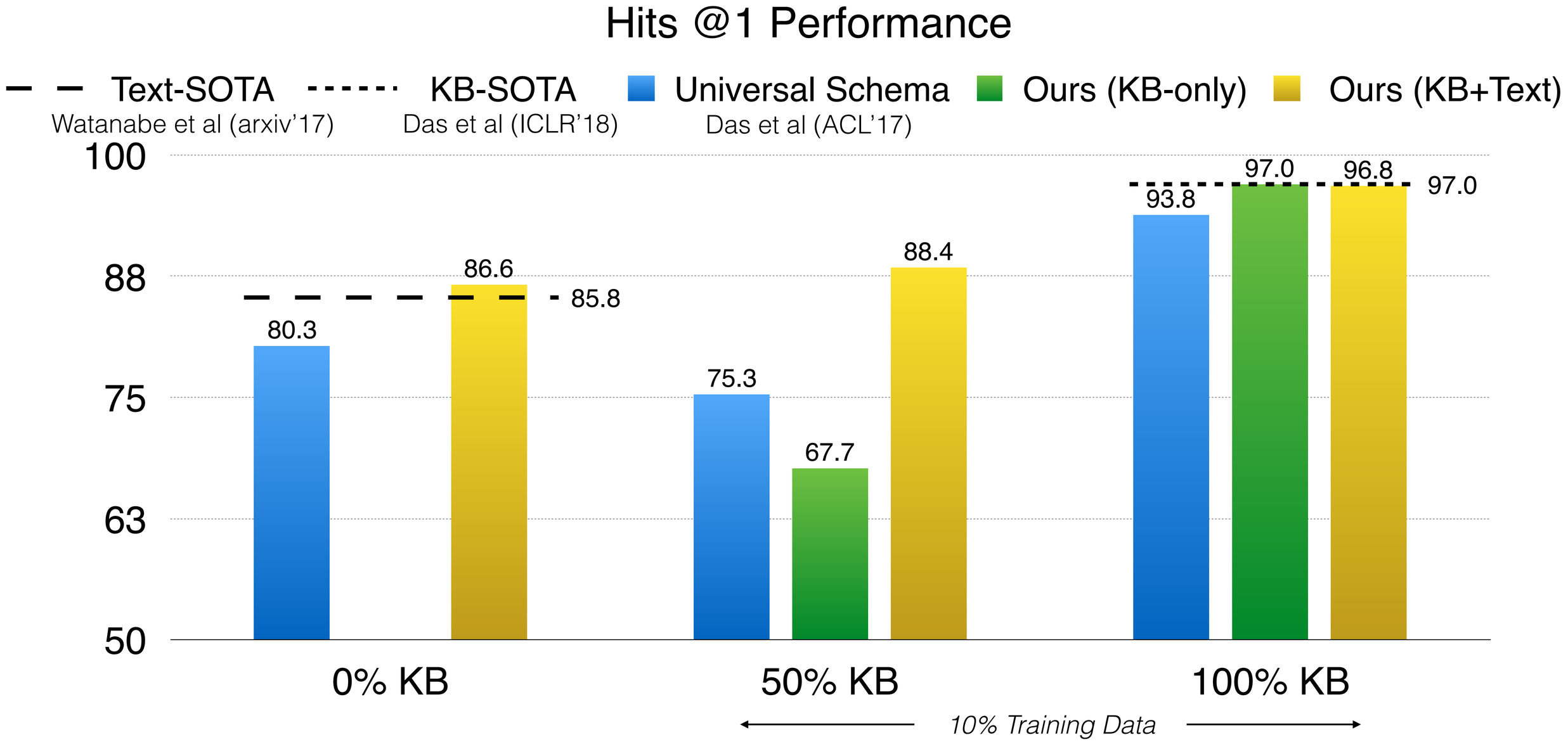
WikiMovies Results



WikiMovies Results



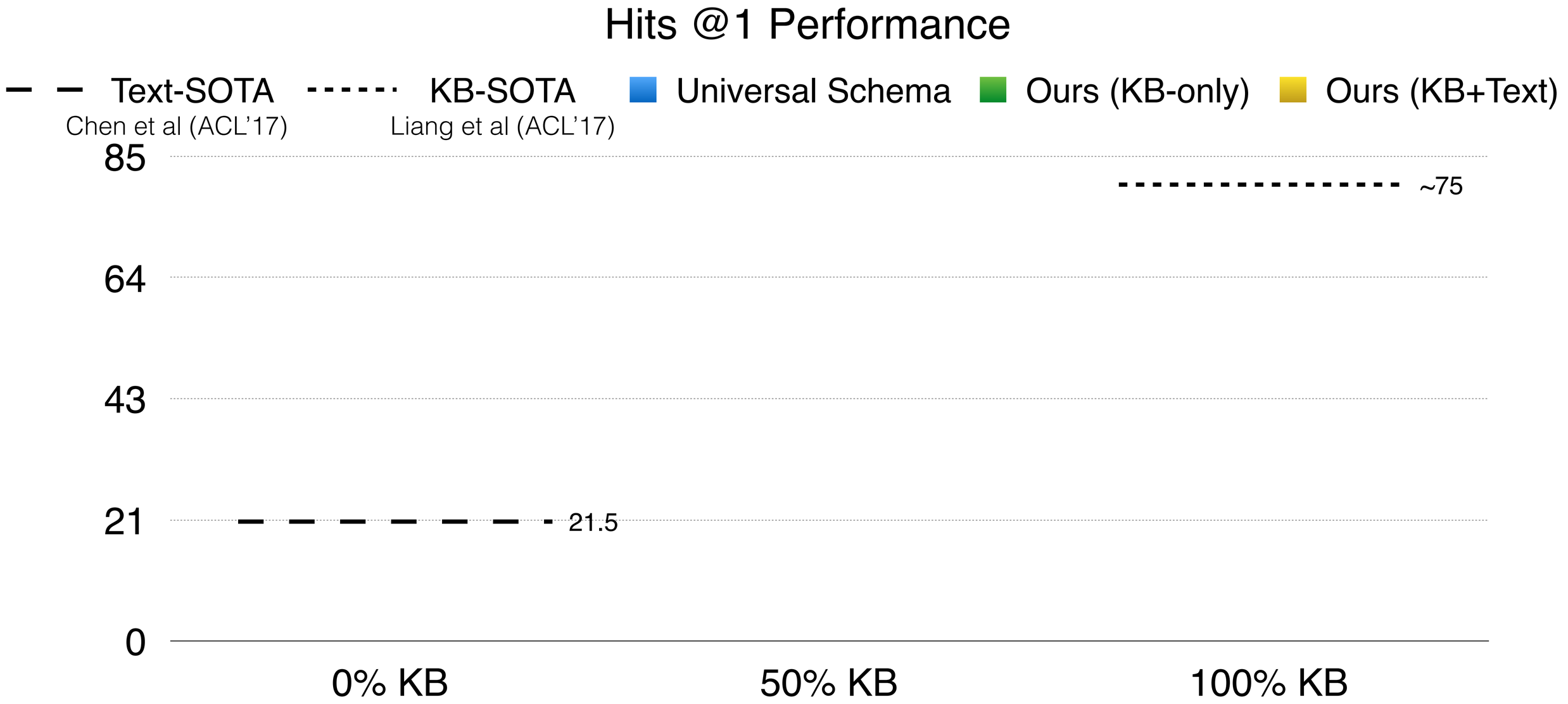
WikiMovies Results



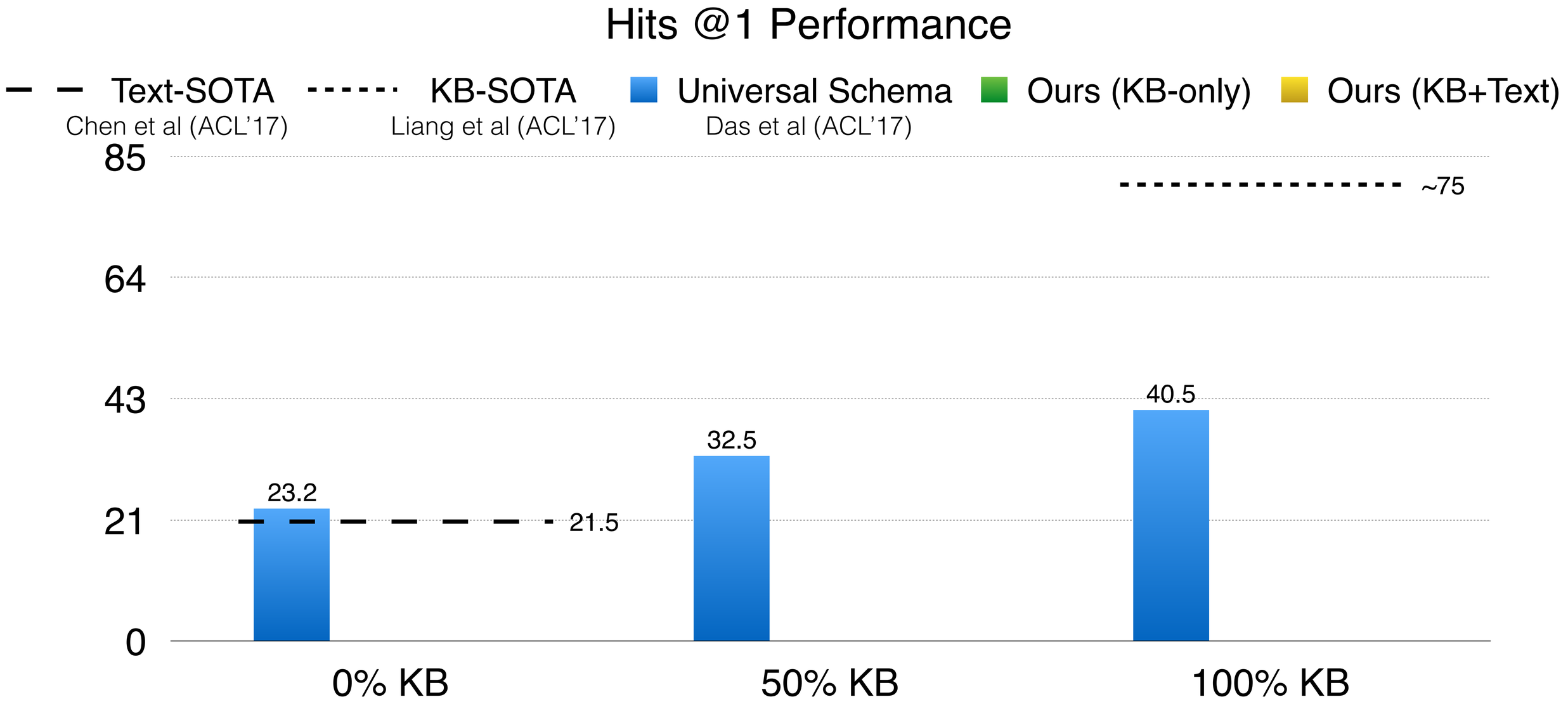
WebQuestionsSP Results

Hits @1 Performance

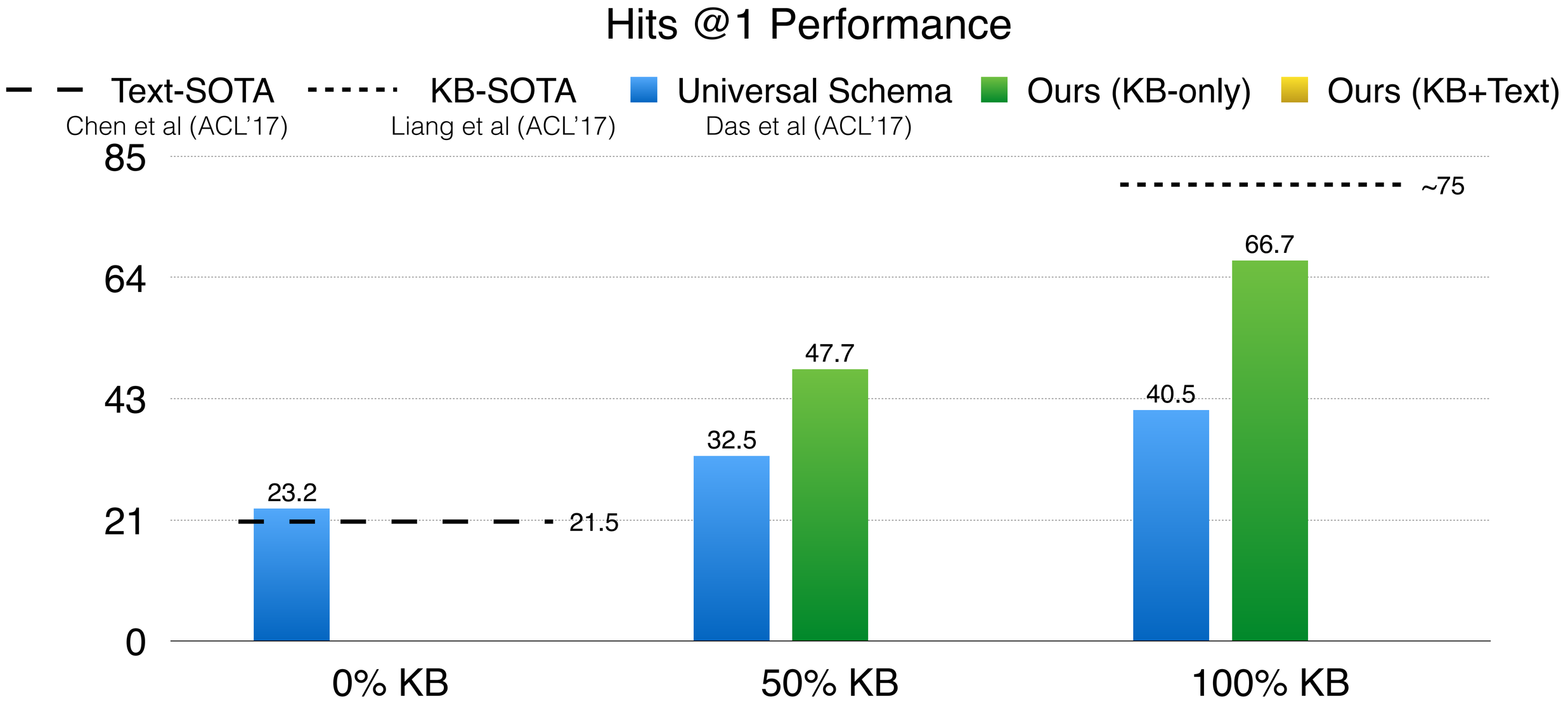
WebQuestionsSP Results



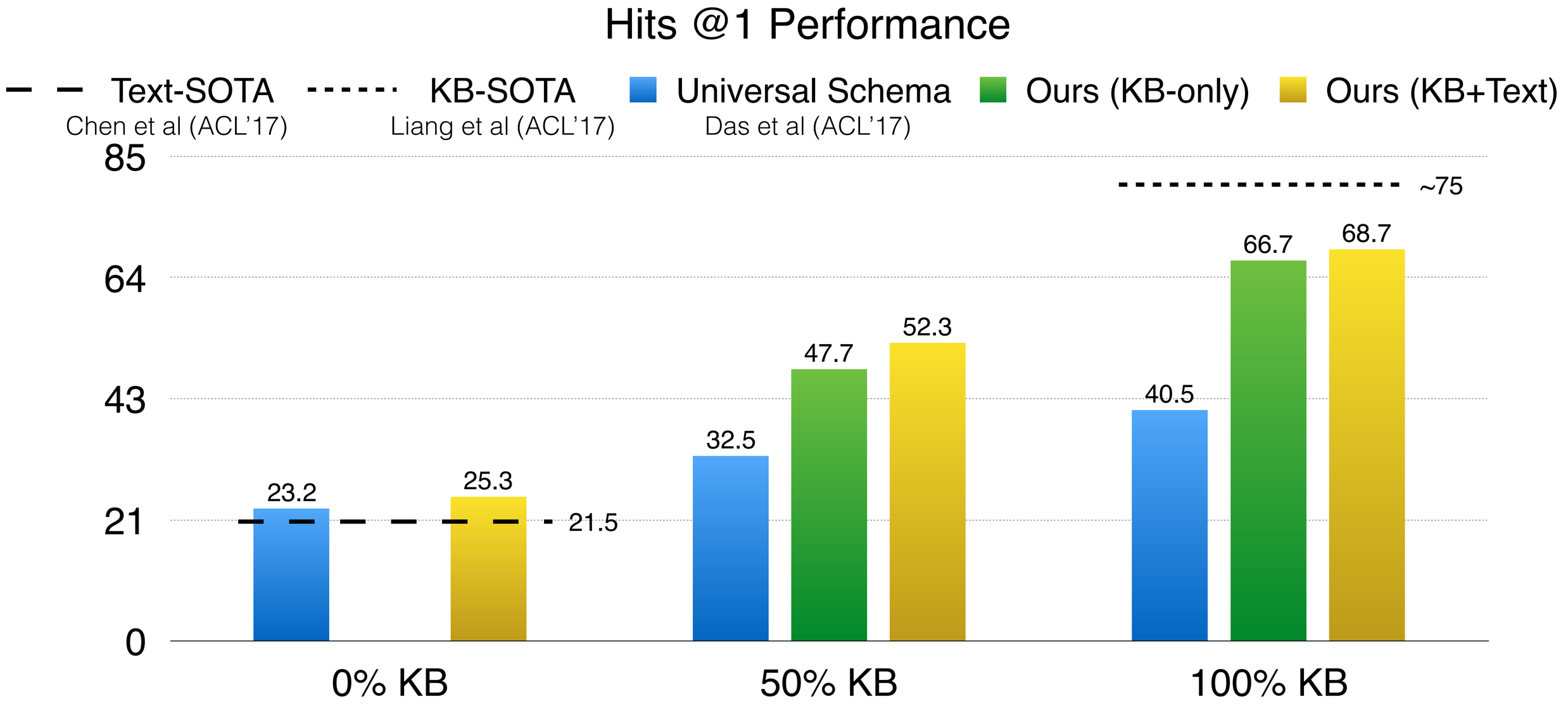
WebQuestionsSP Results



WebQuestionsSP Results



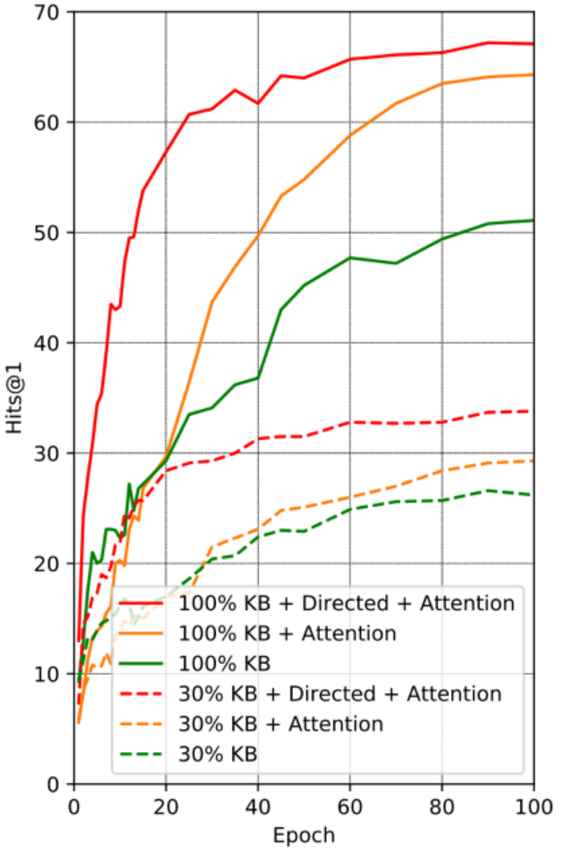
WebQuestionsSP Results



Analysis

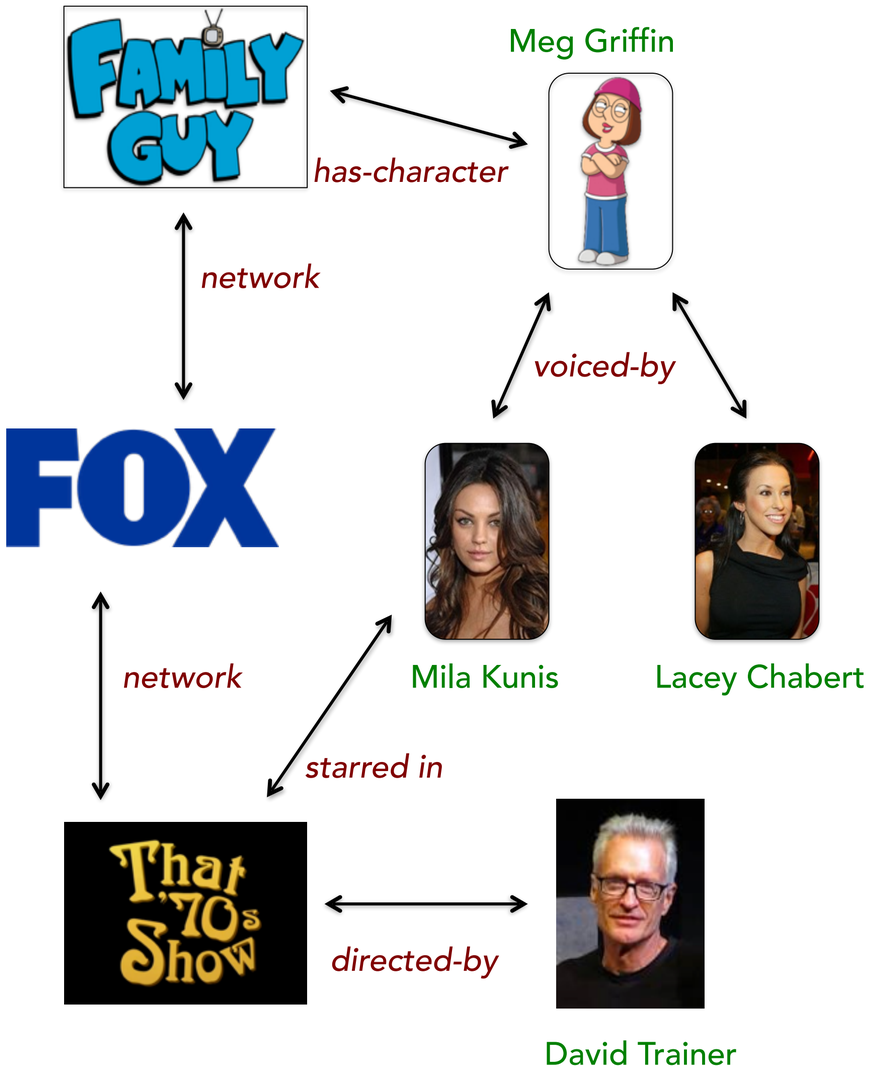
- *Pagerank* propagation leads to ~10% improvement
- Errors:
  - Retrieval (both text + facts) has only 90% recall
  - Complex questions:

    “Who *first* voiced Meg in Family Guy?”
    “Which club did Cristiano Ronaldo play for *in 2007*?”



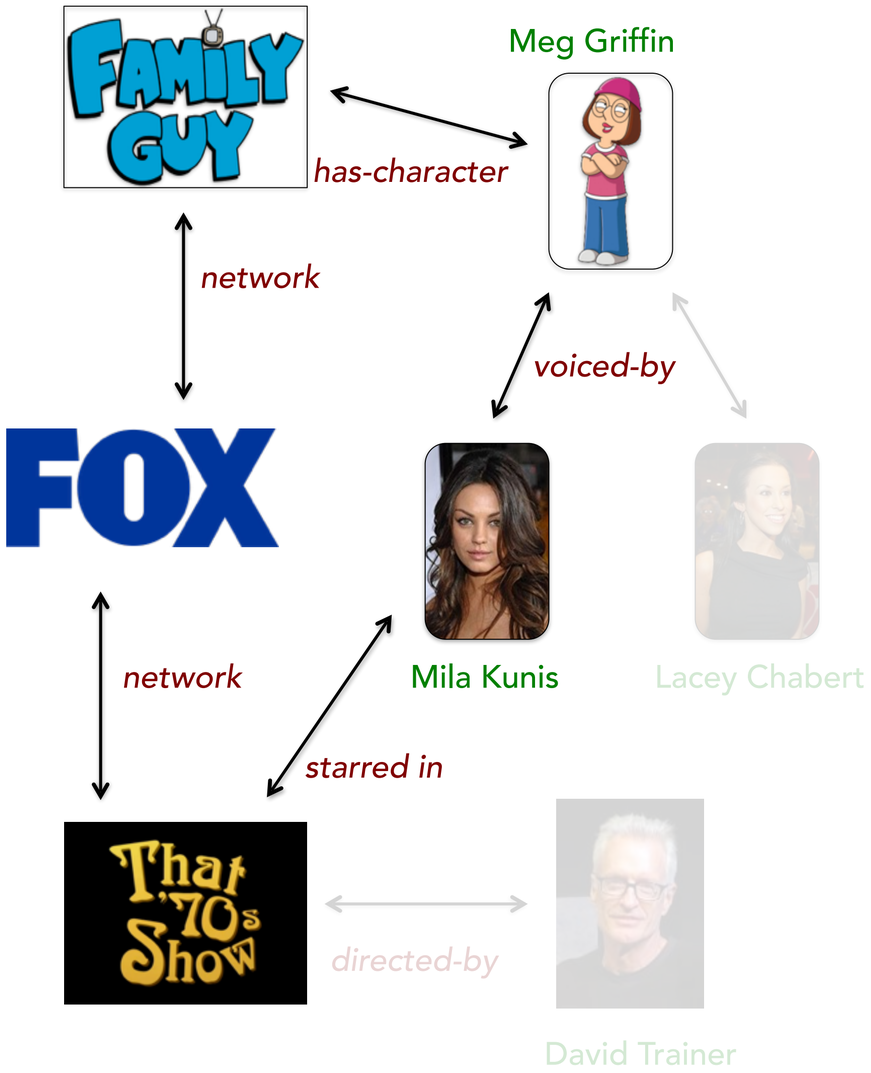
Title and authors removed for anonymity

Multi-Hop Question Answering



Q. *Which TV Networks has Mila Kunis appeared on?*

Multi-Hop Question Answering



Q. *Which TV Networks has Mila Kunis appeared on?*

```
Y1 = {Mila_Kunis}.follow(starred-in).follow(network)

Y2 = {Mila_Kunis}.follow(voiced)....follow(network)

Ans = Y1 UNION Y2
```

A. *Fox*

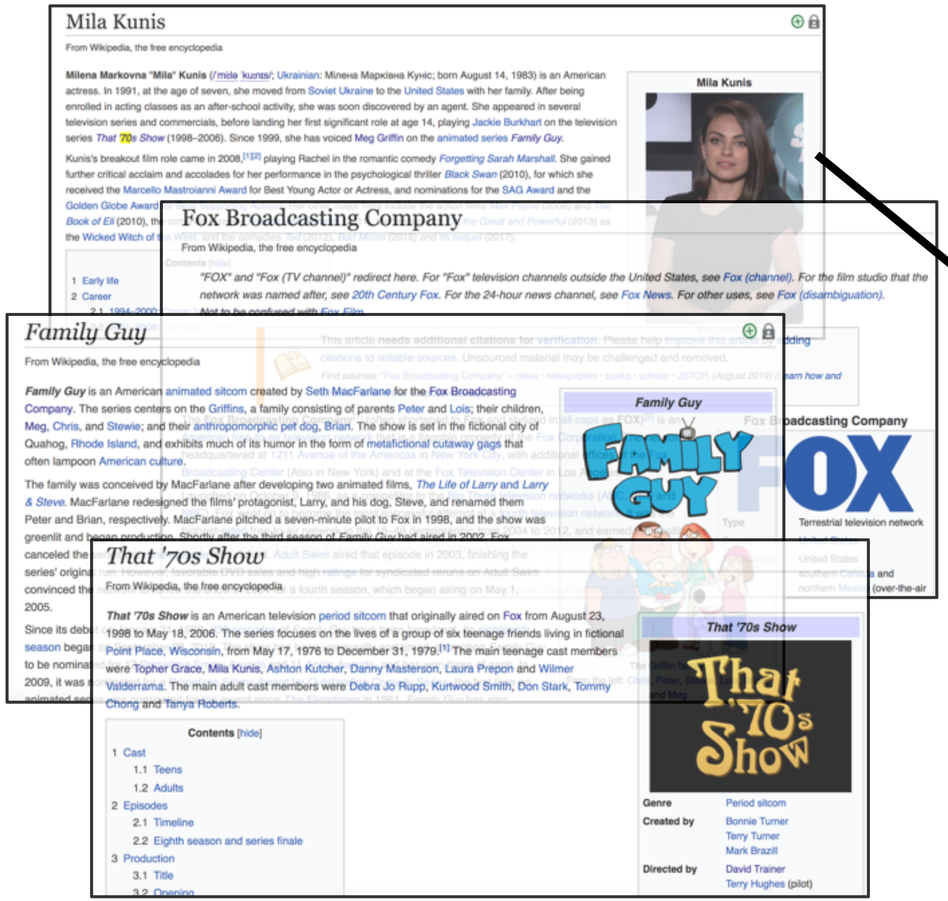
$$Y = X.\text{follow}(R) = \{x' : \exists x \in X \text{ s.t. } R(x, x') \text{ holds}\}$$

Can we do this with a Text Corpus?

Can we do this with a Text Corpus?

Q. *Which TV Networks has Mila Kunis appeared on?*

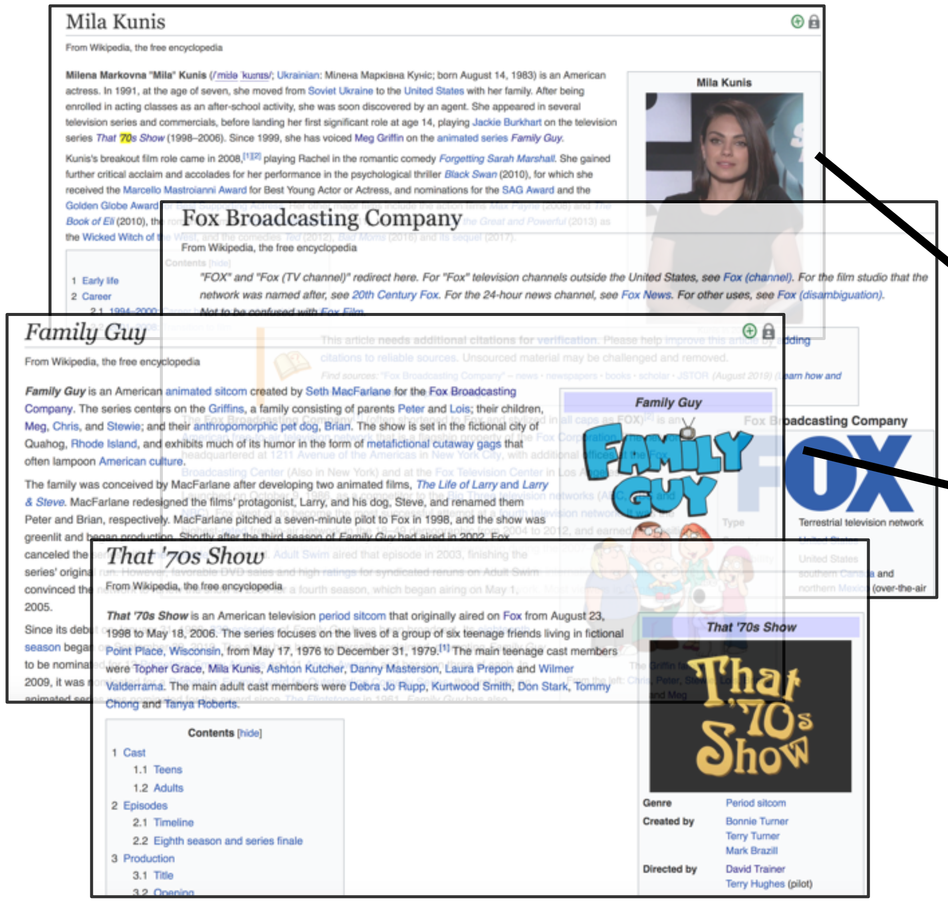
Can we do this with a Text Corpus?



Q. *Which TV Networks has Mila Kunis appeared on?*

```
Since 1999, she has voiced Meg Griffin on the animated
series Family Guy.
```

Can we do this with a Text Corpus?

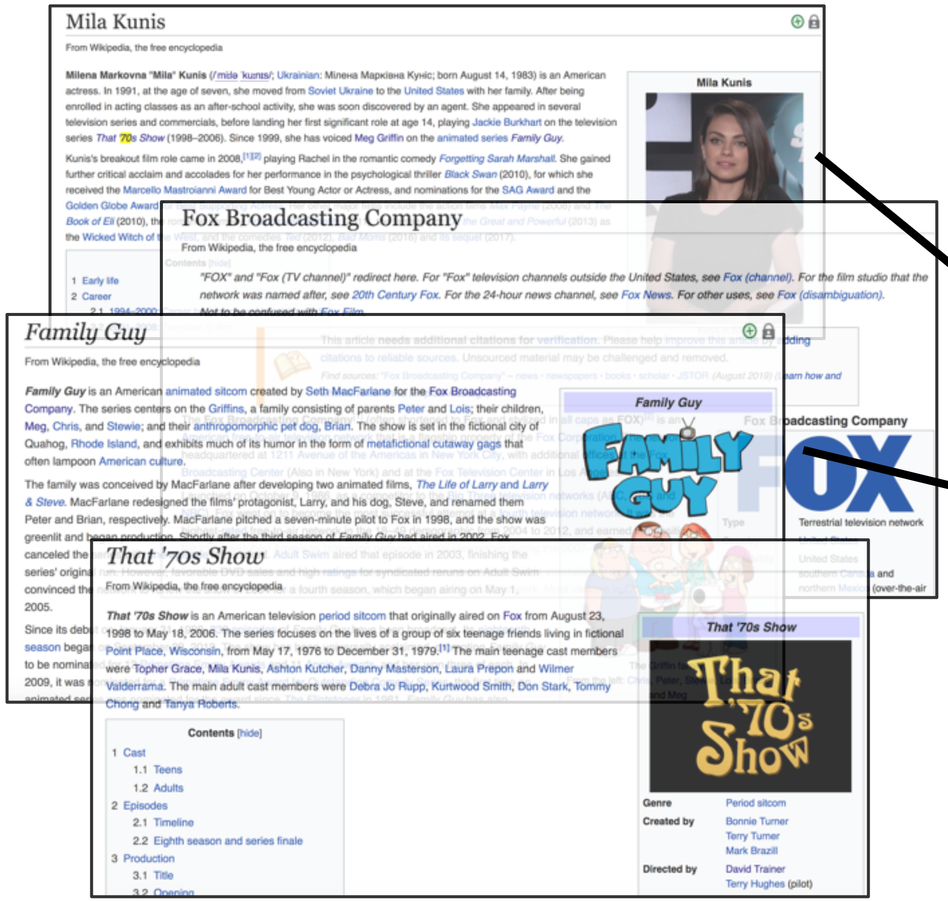


Q. Which TV Networks has Mila Kunis appeared on?

Since 1999, she has voiced Meg Griffin on the animated series Family Guy.

Family Guy is an American animated sitcom created by Seth MacFarlane for the Fox Broadcasting Company.

Can we do this with a Text Corpus?



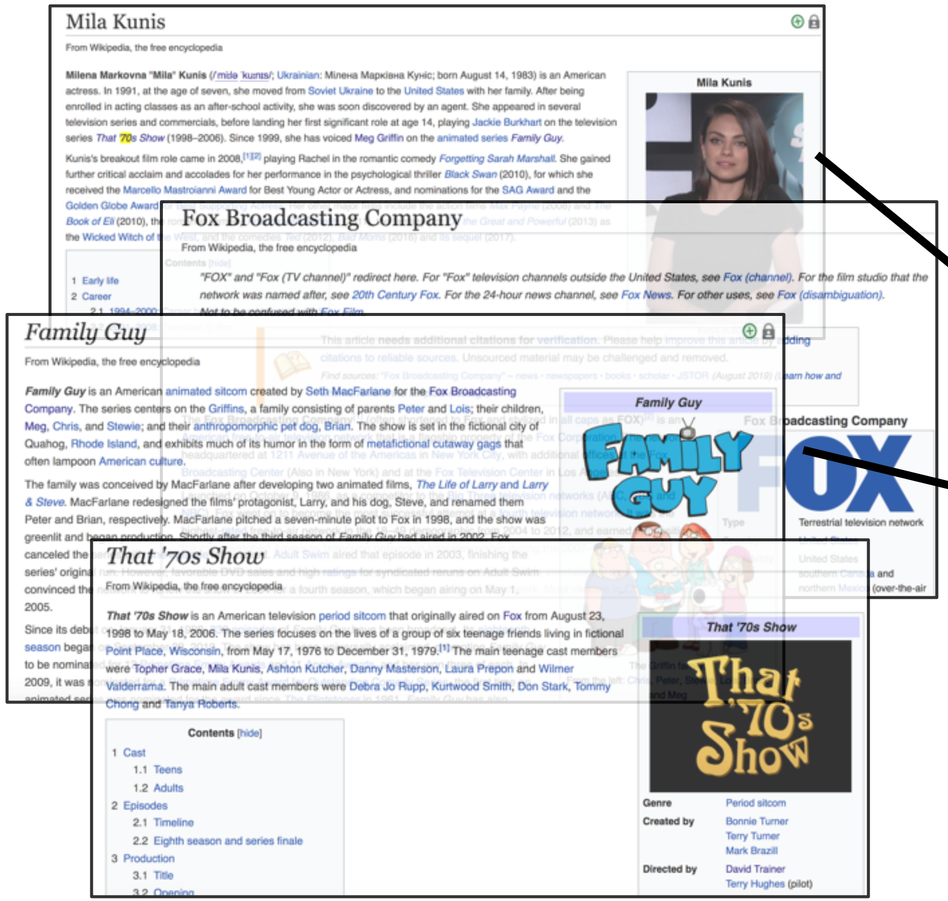
Q. *Which TV Networks has Mila Kunis appeared on?*

```
Since 1999, she has voiced Meg Griffin on the animated
series Family Guy.

Family Guy is an American animated sitcom created by
Seth MacFarlane for the Fox Broadcasting Company.
```

A. *Fox*

Can we do this with a Text Corpus?



Q. *Which TV Networks has Mila Kunis appeared on?*

```
Since 1999, she has voiced Meg Griffin on the animated
series Family Guy.

Family Guy is an American animated sitcom created by
Seth MacFarlane for the Fox Broadcasting Company.
```

A. *Fox*

Information can be spread out across the corpus

MetaQA Benchmark

MetaQA Benchmark

- Expands Wikimovies dataset to 2-hop and 3-hop questions

MetaQA Benchmark

- Expands Wikimovies dataset to 2-hop and 3-hop questions
- 1-hop:
  What movies did Quentin Tarantino direct?

MetaQA Benchmark

- Expands Wikimovies dataset to 2-hop and 3-hop questions
- 1-hop:
  What movies did Quentin Tarantino direct?
- 2-hop:
  Which movies have the same director as that of Pulp Fiction?

MetaQA Benchmark

- Expands Wikimovies dataset to 2-hop and 3-hop questions
- 1-hop:
  What movies did Quentin Tarantino direct?
- 2-hop:
  Which movies have the same director as that of Pulp Fiction?
- 3-hop:
  Who acted in the movies which have the same director as Pulp Fiction?

MetaQA Benchmark

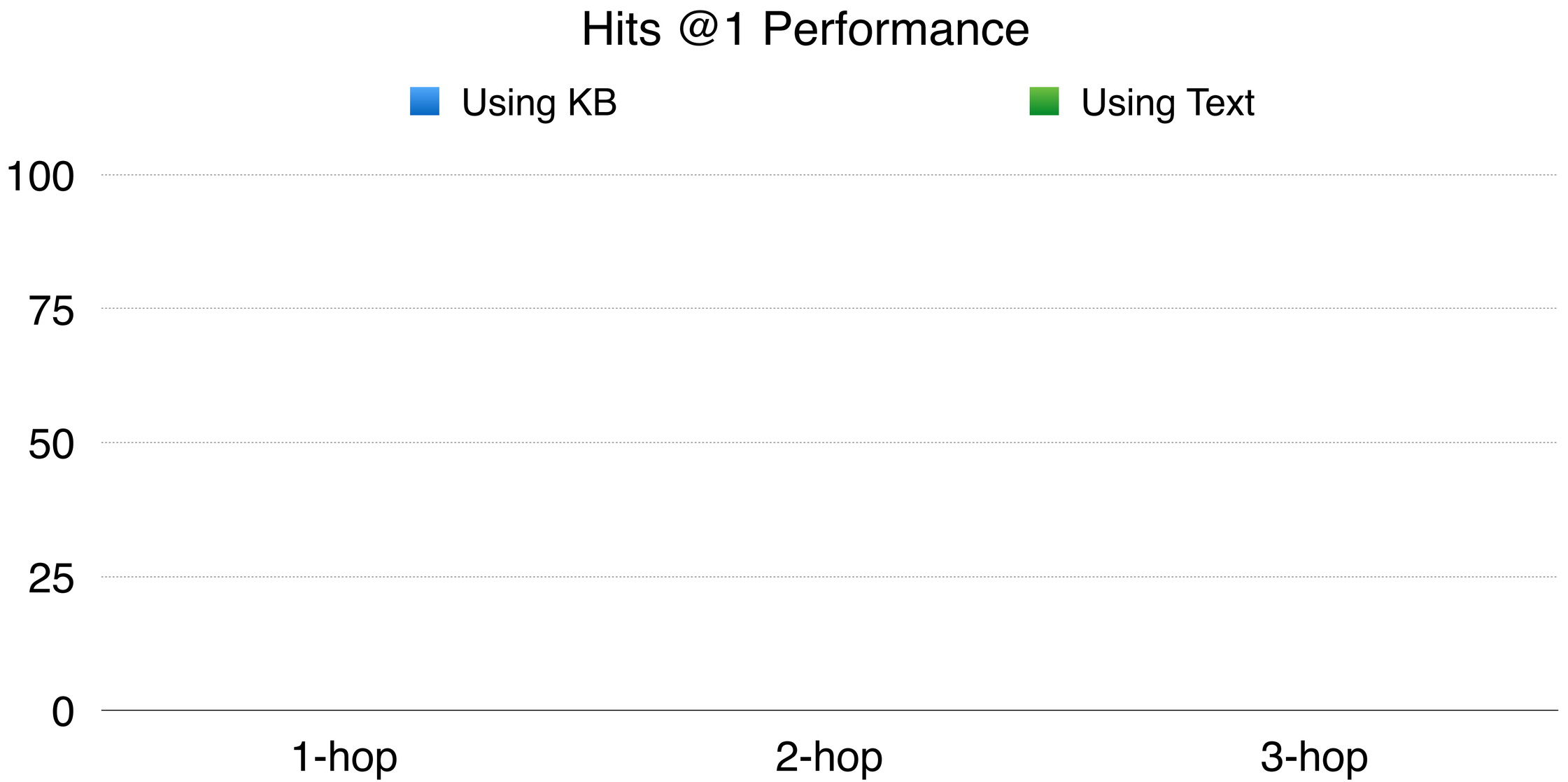
- Expands Wikimovies dataset to 2-hop and 3-hop questions
- 1-hop:
  What movies did Quentin Tarantino direct?
- 2-hop:
  Which movies have the same director as that of Pulp Fiction?
- 3-hop:
  Who acted in the movies which have the same director as Pulp Fiction?
- Text corpus:
  17K first paragraphs from Wikipedia articles about the movies

  Pulp Fiction is a 1994 American crime film written and directed by Quentin Tarantino, who conceived it with Roger Avary. Starring John Travolta, Samuel L. Jackson, Bruce Willis, Tim Roth, Ving Rhames, and Uma Thurman, it tells several stories of criminal Los Angeles. The title refers to the pulp magazines and hardboiled crime novels popular during the mid-20th century, known for their graphic violence and punchy dialogue.

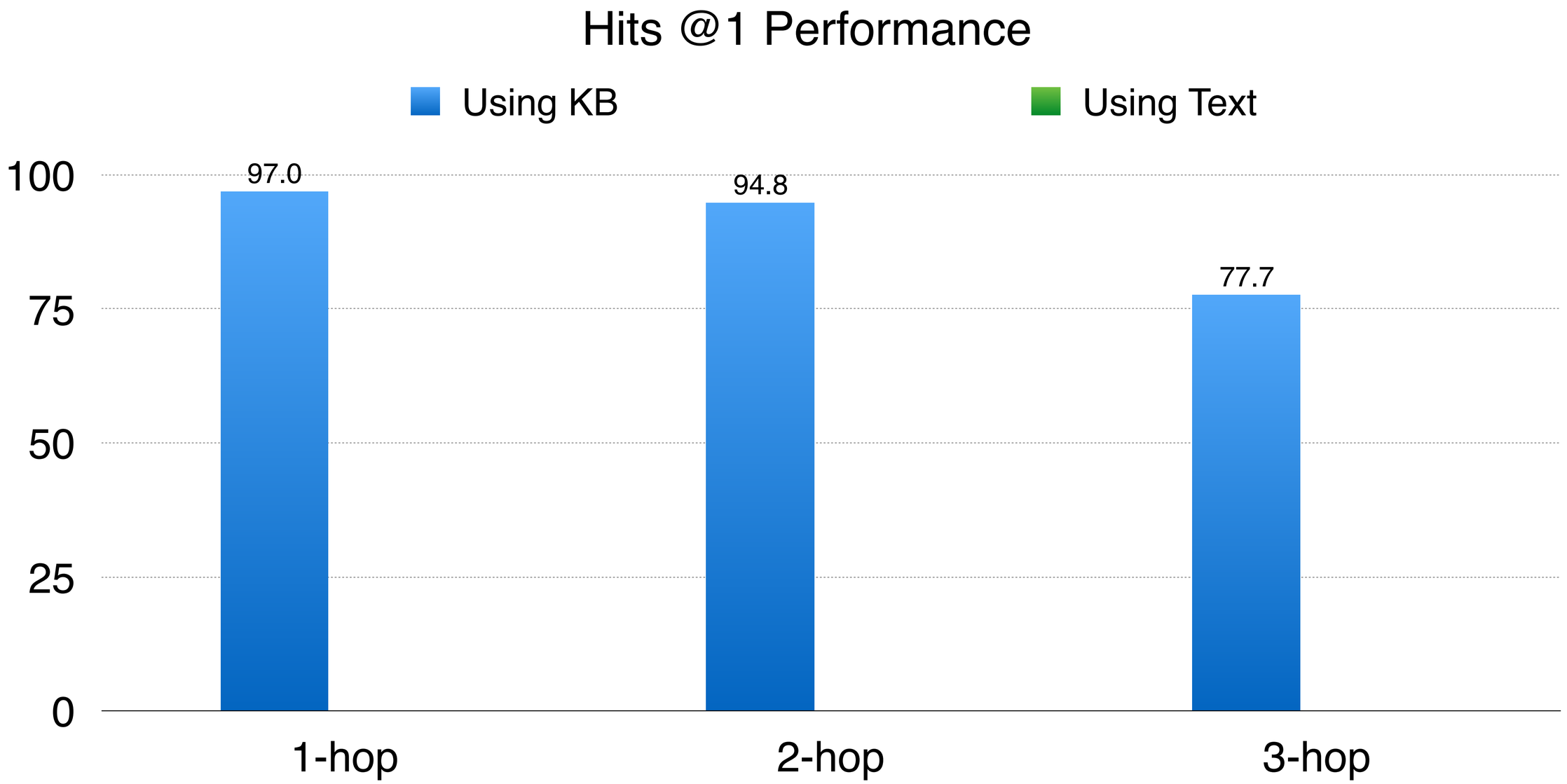
Graph Networks on MetaQA

Hits @1 Performance

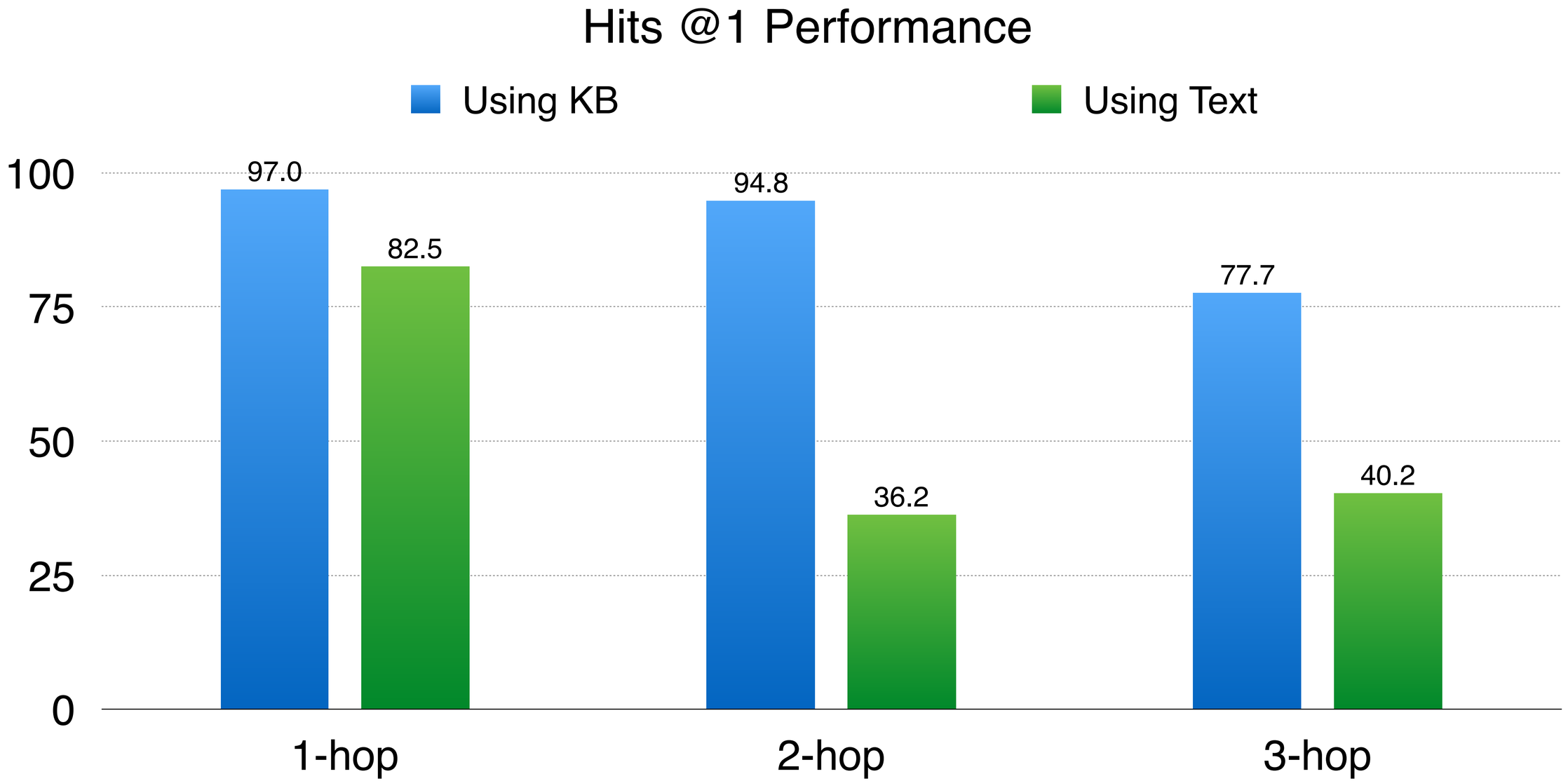
Graph Networks on MetaQA



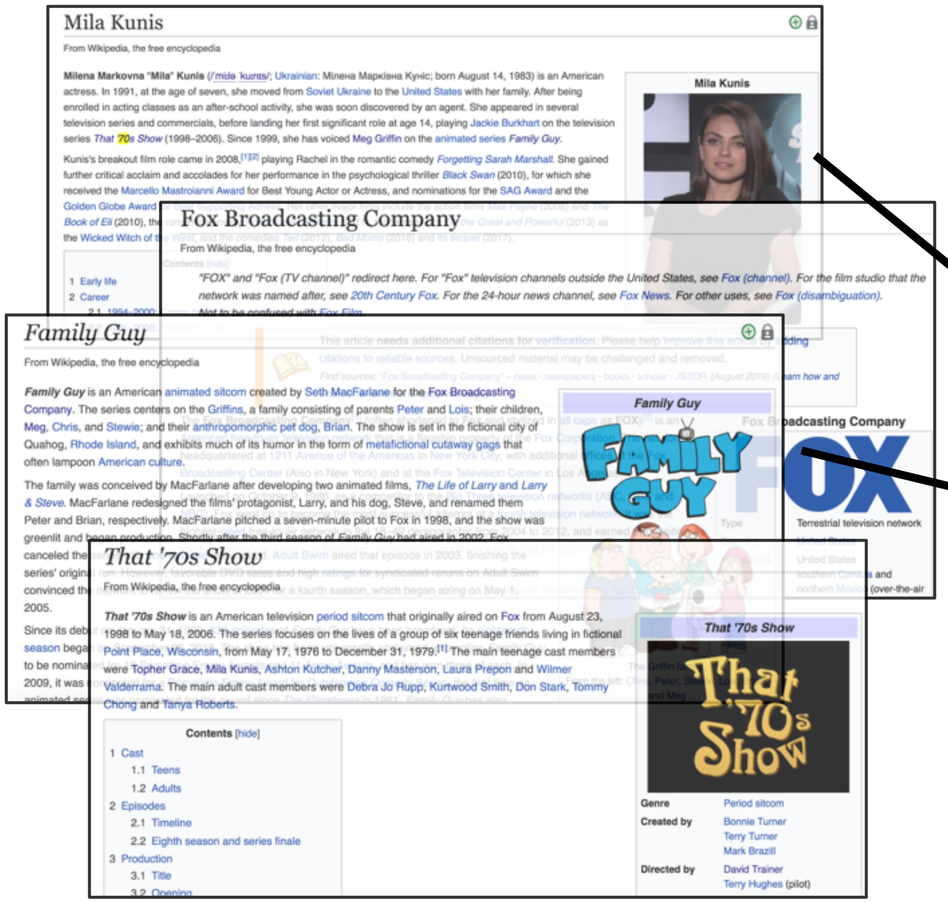
Graph Networks on MetaQA



Graph Networks on MetaQA



Problem: Retrieval



Q. *Which TV Networks has Mila Kunis appeared on?*

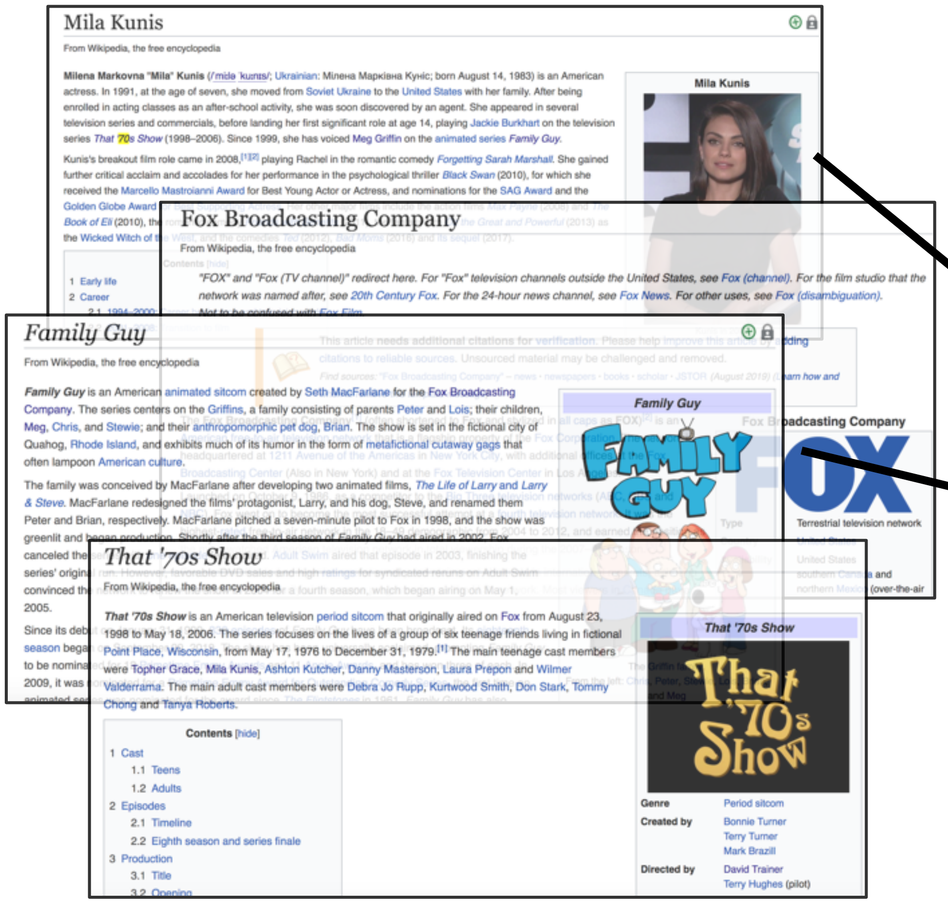
```
Since 1999, she has voiced Meg Griffin on the animated
series Family Guy.

Family Guy is an American animated sitcom created by
Seth MacFarlane for the Fox Broadcasting Company.
```

A. *Fox*

Information can be spread out across the corpus

Problem: Retrieval



Q. *Which TV Networks has Mila Kunis appeared on?*

```
Since 1999, she has voiced Meg Griffin on the animated
series Family Guy.

Family Guy is an American animated sitcom created by
Seth MacFarlane for the Fox Broadcasting Company.
```

A. *Fox*

Information can be spread out across the corpus

- Shallow information retrieval does not work
- Can expand retrieval using e.g. pseudo-relevance-feedback
  - But “reading” large number of documents is expensive

Our Approach: Read offline, Reason online

Our Approach: Read offline, Reason online

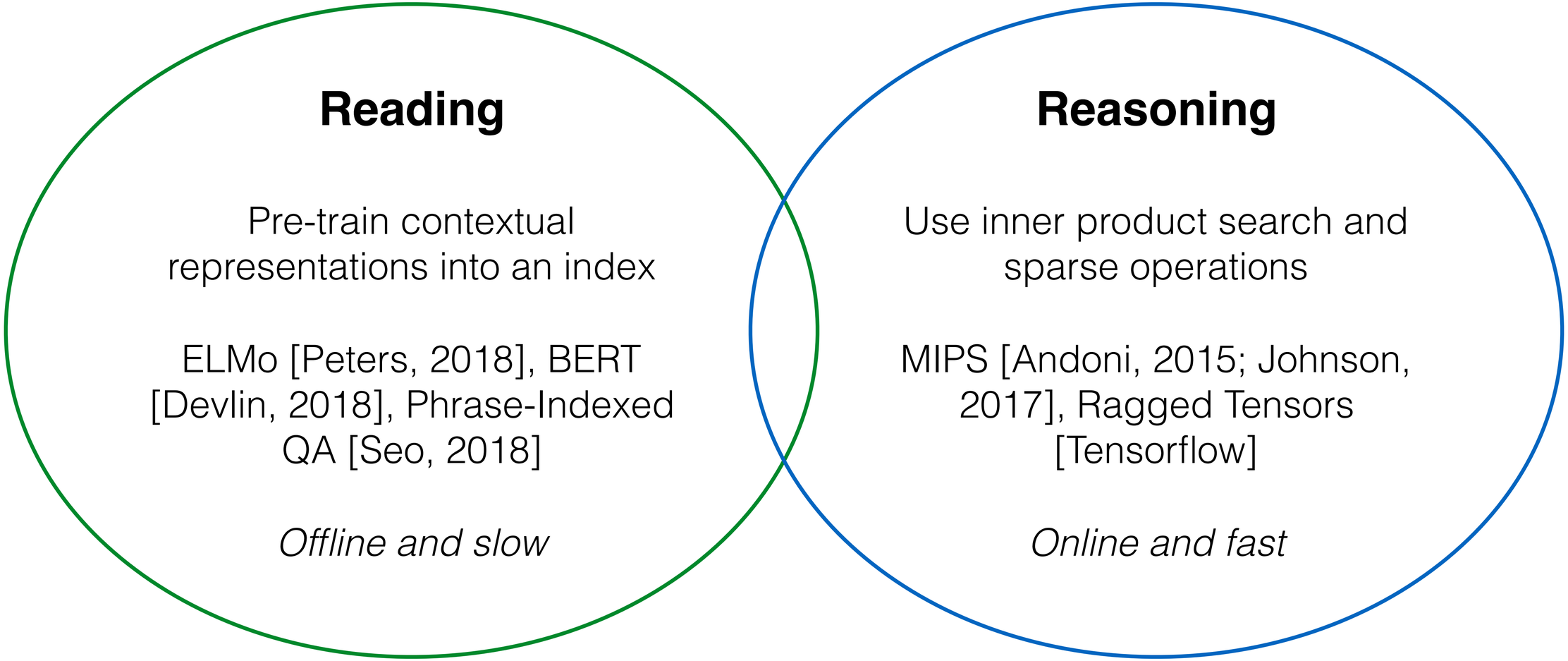
Reading

Pre-train contextual representations into an index

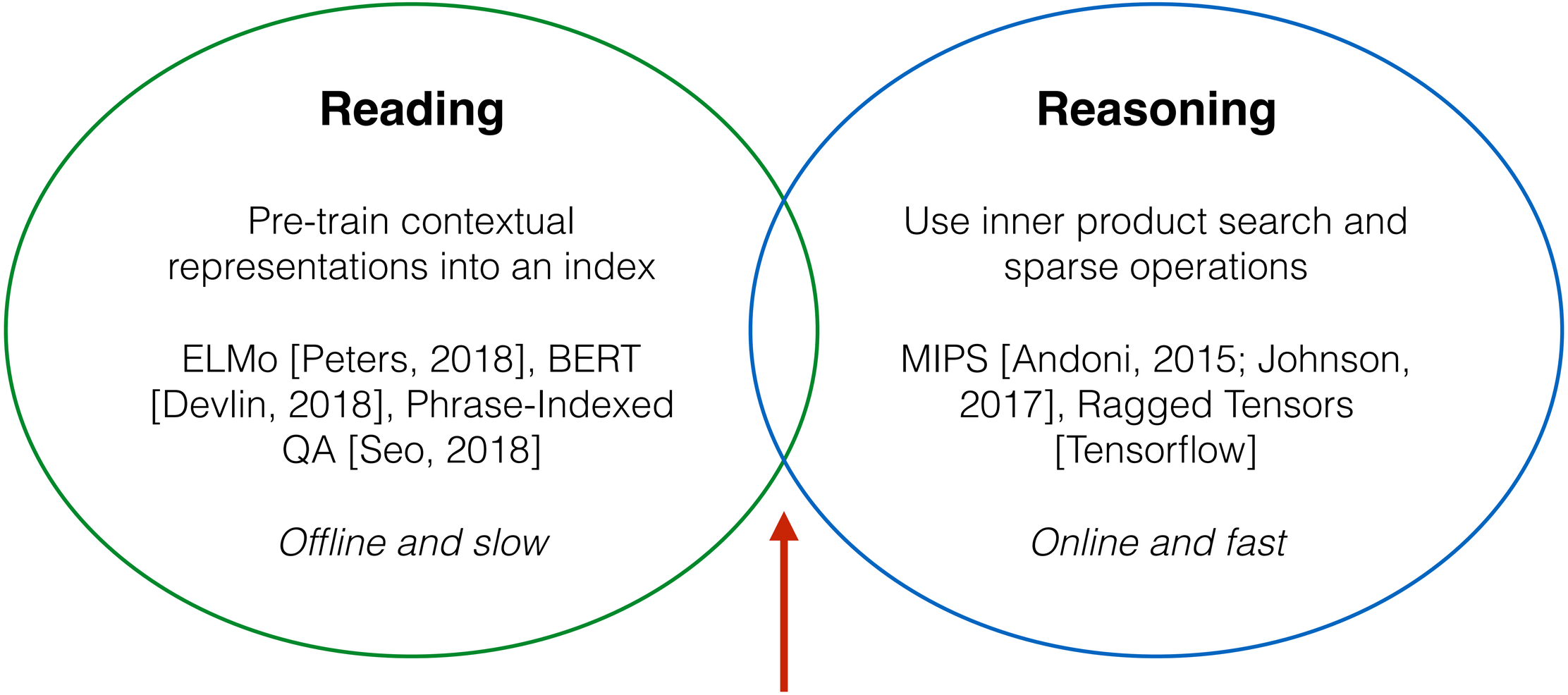
ELMo [Peters, 2018], BERT [Devlin, 2018], Phrase-Indexed QA [Seo, 2018]

Offline and slow

Our Approach: Read offline, Reason online



Our Approach: Read offline, Reason online



Reading: Offline Index of Entity Mentions

Based on Phrase-Indexed QA (Seo et al, EMNLP'18, ACL'19)

Reading: Offline Index of Entity Mentions

Entity Linking

Entity Mentions

… ***Family Guy*** *is an American …*

… their children, ***Meg****, Chris, Stewie …*

… In 1999, ***Kunis*** *replaced Lacey Chabert …*

… created by ***Seth MacFarlene*** *for Fox …*

… cast members were ***Topher Grace****, Mile Kunis …*

… originally aired on Fox from ***August 23*** *…*

… wanted the show to have a ***1970s*** *feel …*

Based on Phrase-Indexed QA (Seo et al, EMNLP'18, ACL'19)

Reading: Offline Index of Entity Mentions

Entity Linking

Contextual Representations

Entity Mentions

… ***Family Guy*** *is an American …*

… their children, ***Meg****, Chris, Stewie …*

… In 1999, ***Kunis*** *replaced Lacey Chabert …*

… created by ***Seth MacFarlene*** *for Fox …*

… cast members were ***Topher Grace****, Mile Kunis …*

… originally aired on Fox from ***August 23*** *…*

… wanted the show to have a ***1970s*** *feel …*

Based on Phrase-Indexed QA (Seo et al, EMNLP'18, ACL'19)

Reading: Offline Index of Entity Mentions

Entity Linking

Entity Mentions

… ***Family Guy*** *is an American …*

… their children, ***Meg****, Chris, Stewie …*

… In 1999, ***Kunis*** *replaced Lacey Chabert …*

… created by ***Seth MacFarlene*** *for Fox …*

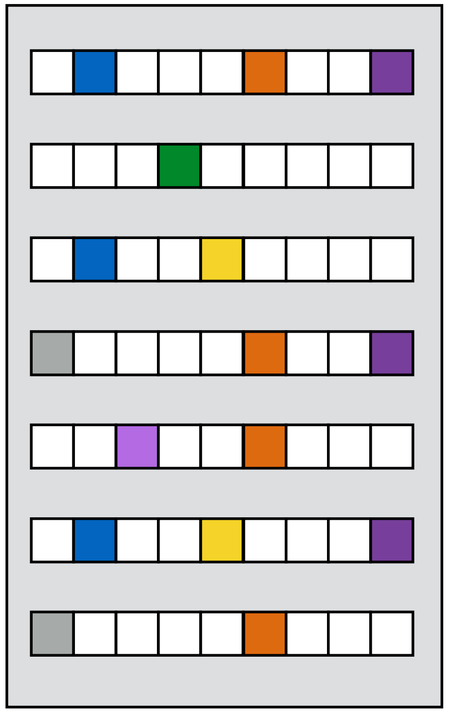
… cast members were ***Topher Grace****, Mile Kunis …*

… originally aired on Fox from ***August 23*** *…*

… wanted the show to have a ***1970s*** *feel …*

Contextual Representations

Sparse



(Fixed)

Based on Phrase-Indexed QA (Seo et al, EMNLP'18, ACL'19)

Reading: Offline Index of Entity Mentions

Entity Linking

Entity Mentions

… ***Family Guy*** *is an American …*

… their children, ***Meg****, Chris, Stewie …*

… In 1999, ***Kunis*** *replaced Lacey Chabert …*

… created by ***Seth MacFarlene*** *for Fox …*

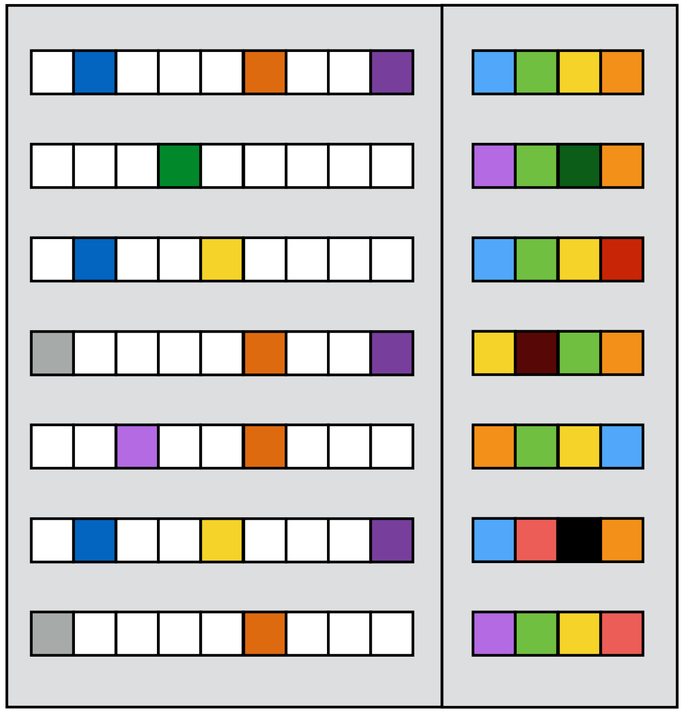
… cast members were ***Topher Grace****, Mile Kunis …*

… originally aired on Fox from ***August 23*** *…*

… wanted the show to have a ***1970s*** *feel …*

Contextual Representations

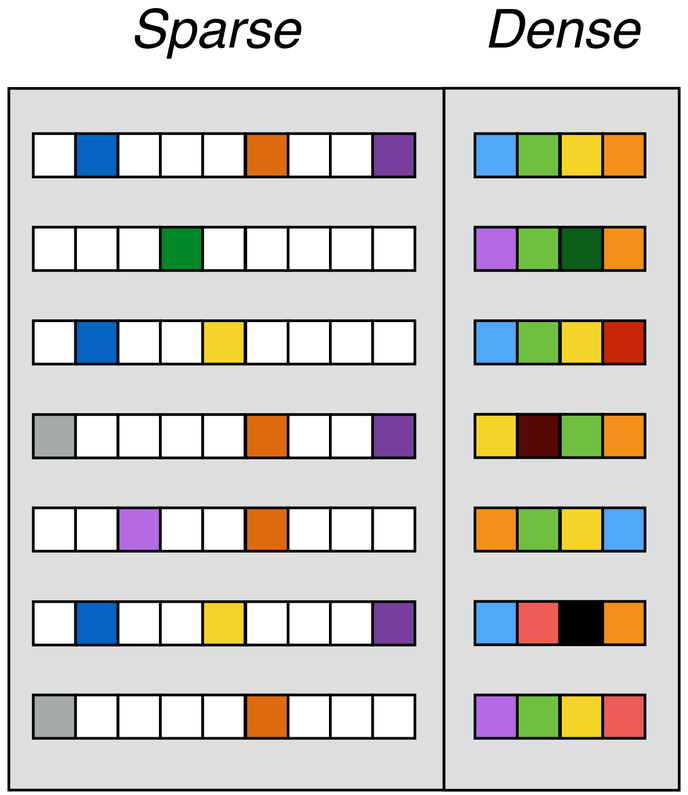
Sparse *Dense*



(Fixed) *(Pre-trained)*

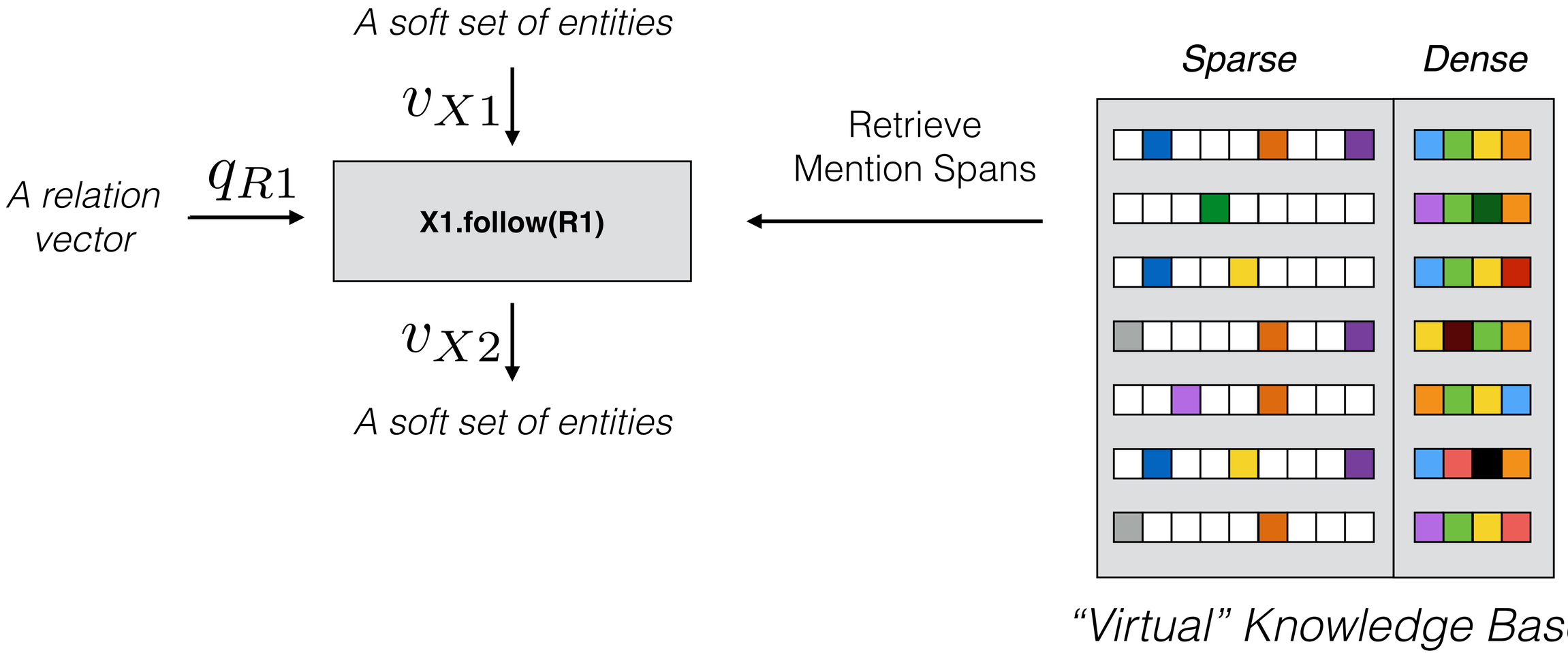
Based on Phrase-Indexed QA (Seo et al, EMNLP'18, ACL'19)

Reasoning: A “Soft” Textual Follow Op



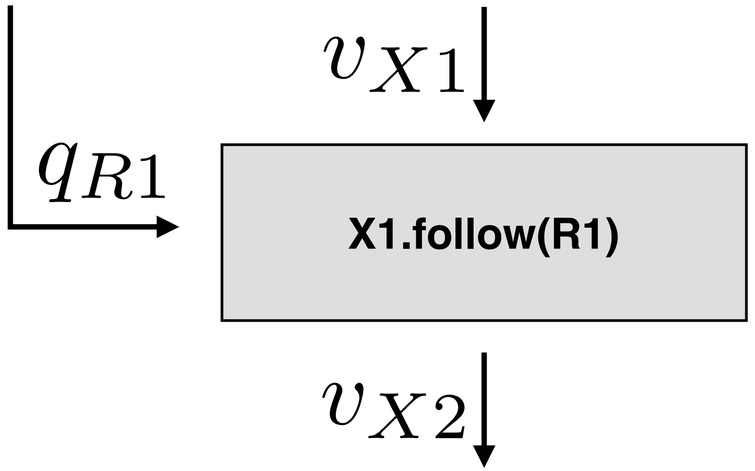
“Virtual” Knowledge Base

Reasoning: A “Soft” Textual Follow Op

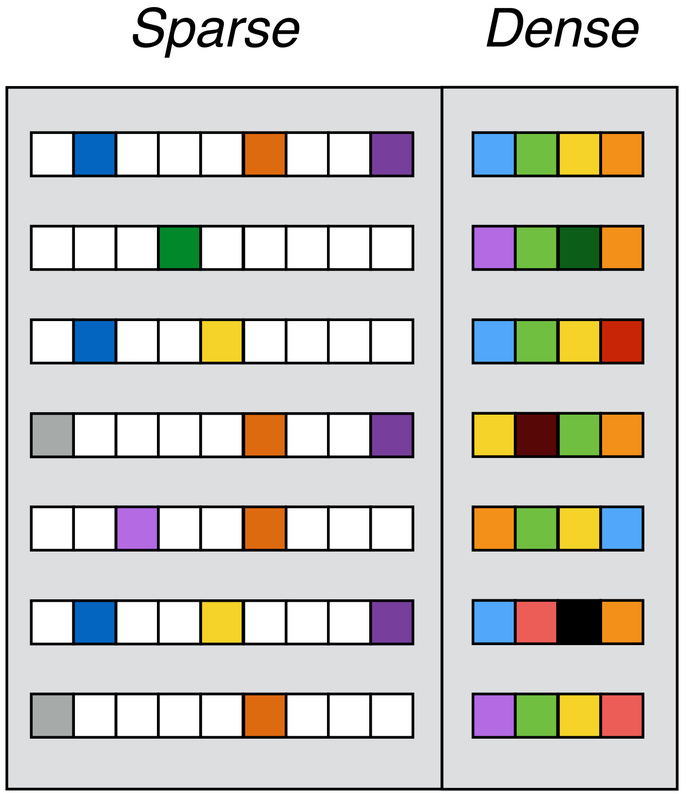


Reasoning: A “Soft” Textual Follow Op

Which TV Networks has Mila Kunis appeared on?

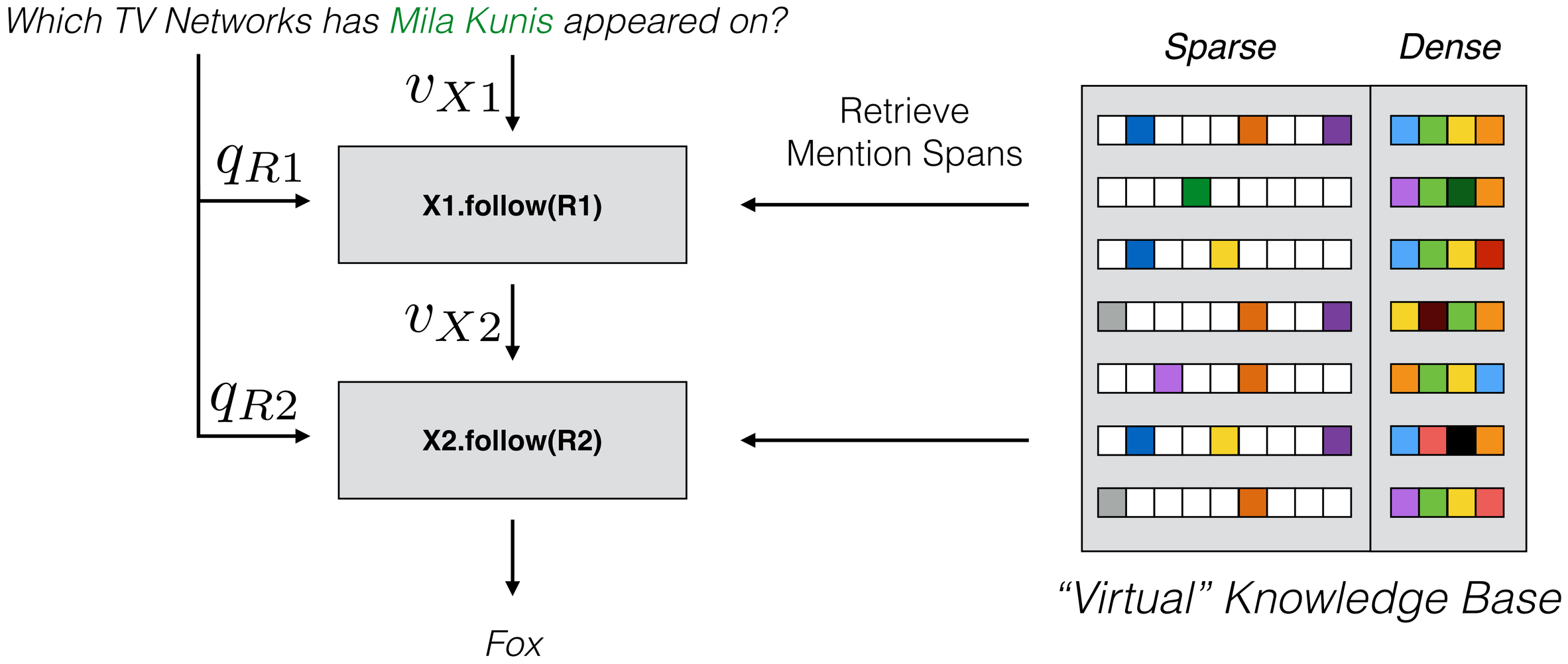


Retrieve
Mention Spans

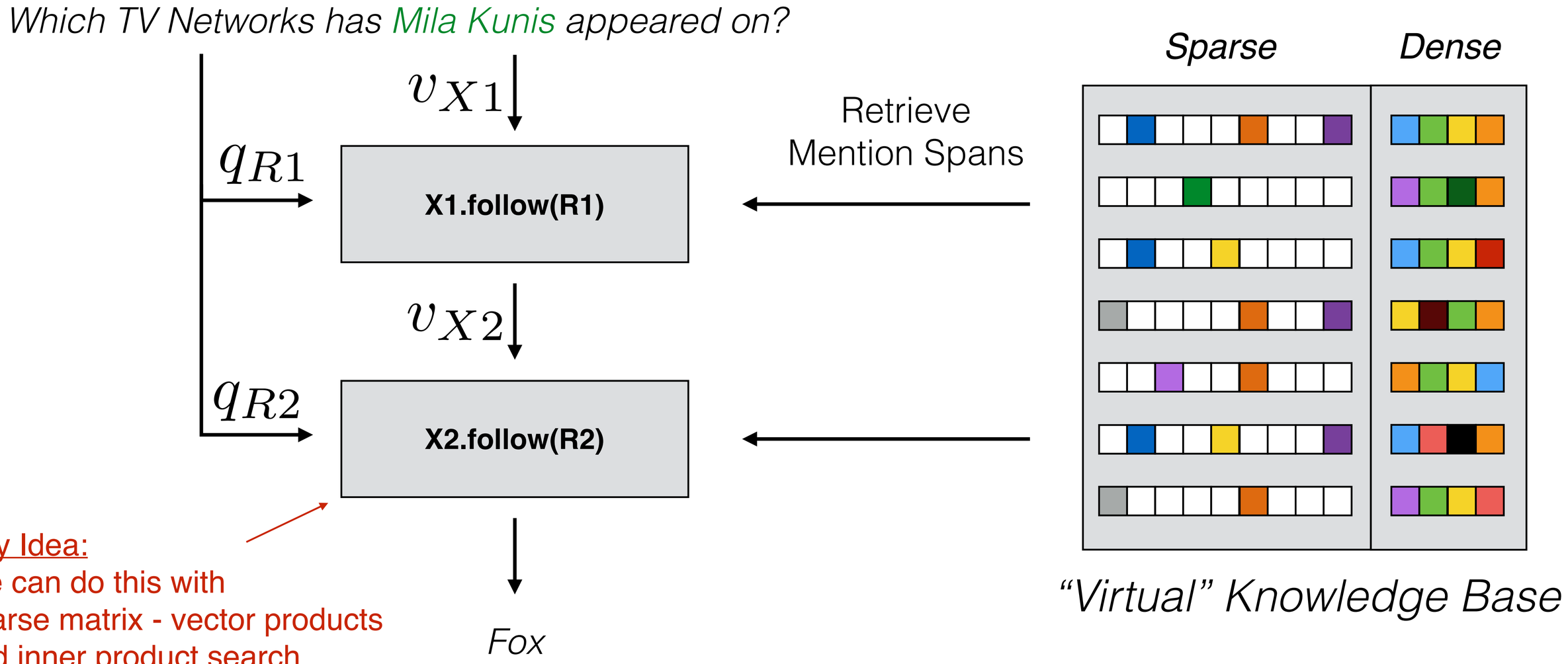


“Virtual” Knowledge Base

Reasoning: A “Soft” Textual Follow Op

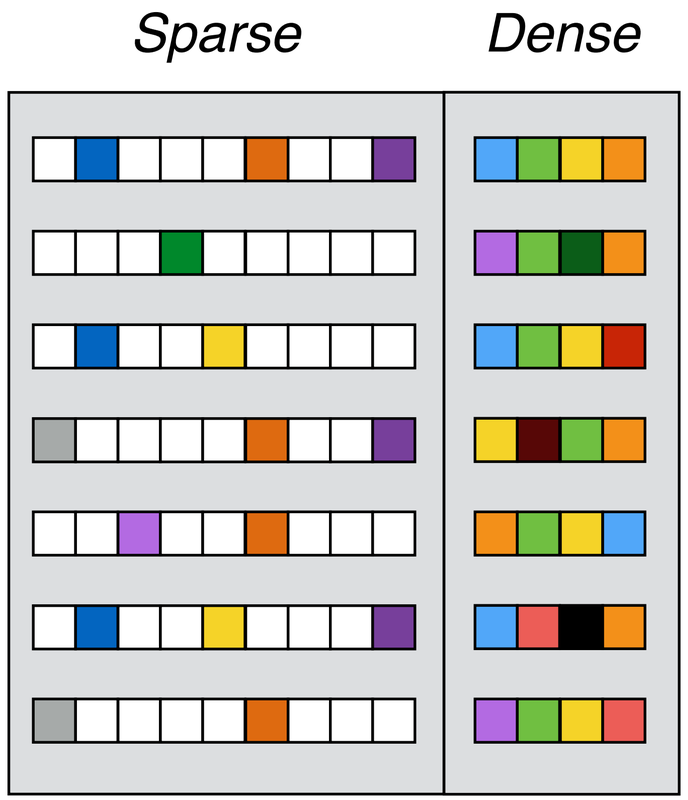


Reasoning: A “Soft” Textual Follow Op

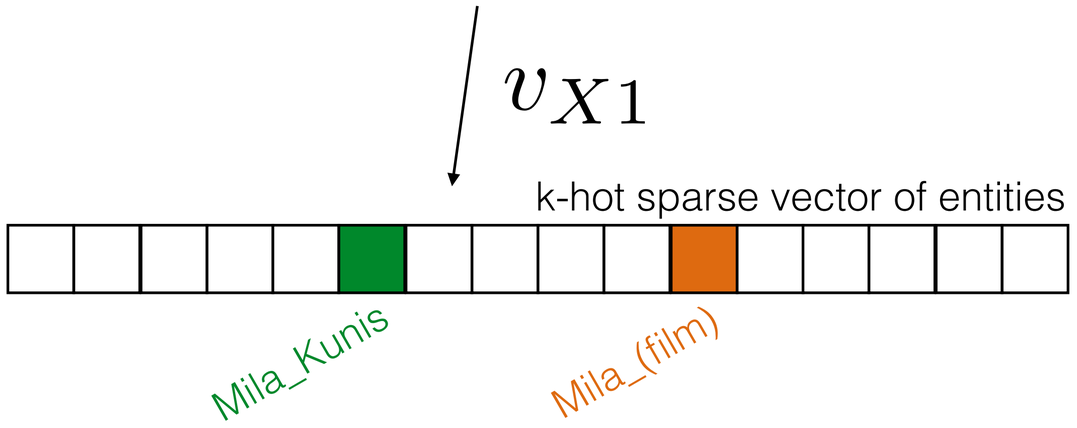


Which TV Networks has Mila Kunis appeared on?

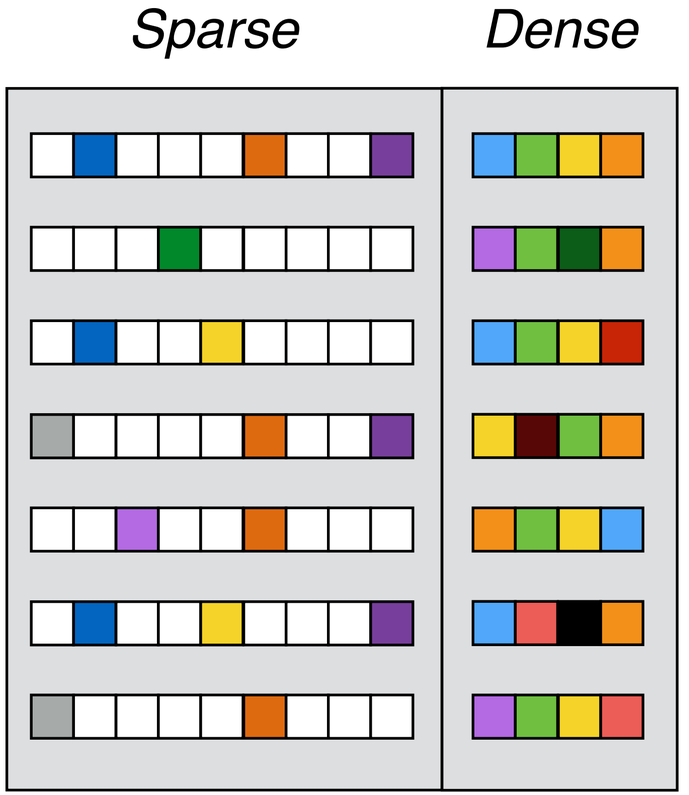
X1.follow(R1)



Which TV Networks has Mila Kunis appeared on?



X1.follow(R1)



Entities are represented as a sparse vector

- Size = # entities in the corpus
- Non-zero values are confidence scores
- E.g. from an entity linking model

Which TV Networks has Mila Kunis appeared on?

v_{X1}

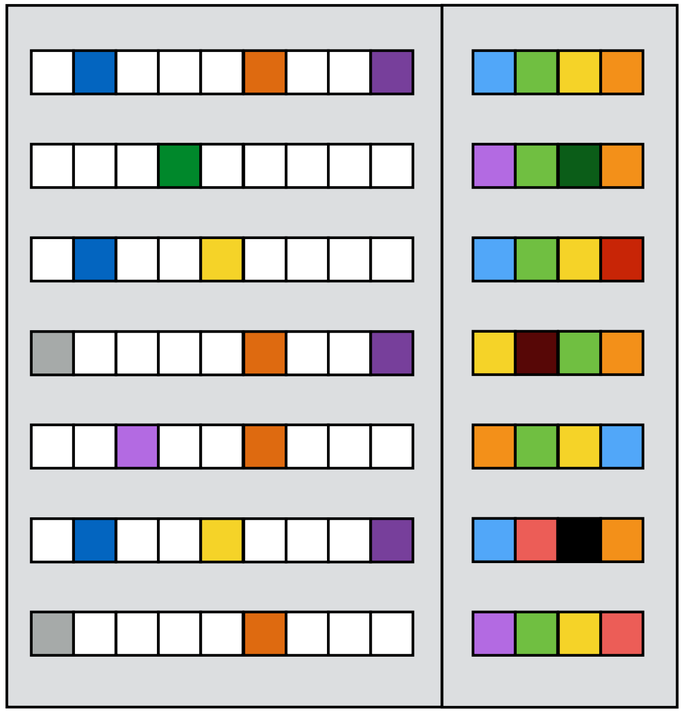
q_{R1}

k-hot sparse vector of entities

dense relation vector

X1.follow(R1)

Sparse *Dense*



Relations are represented as a dense feature vector

- We use a 5-layer Transformer Network over the question

Which TV Networks has Mila Kunis appeared on?

v_{X1}

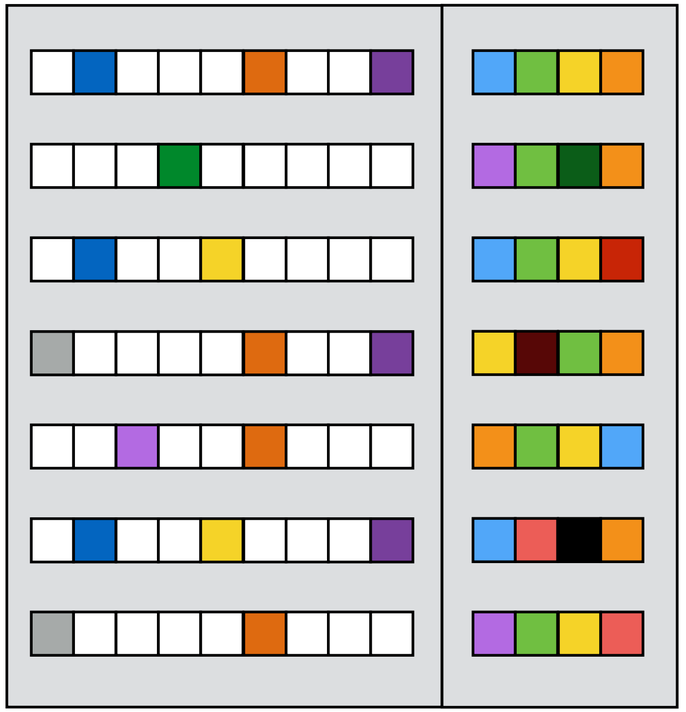
q_{R1}

k-hot sparse vector of entities

dense relation vector

X1.follow(R1)

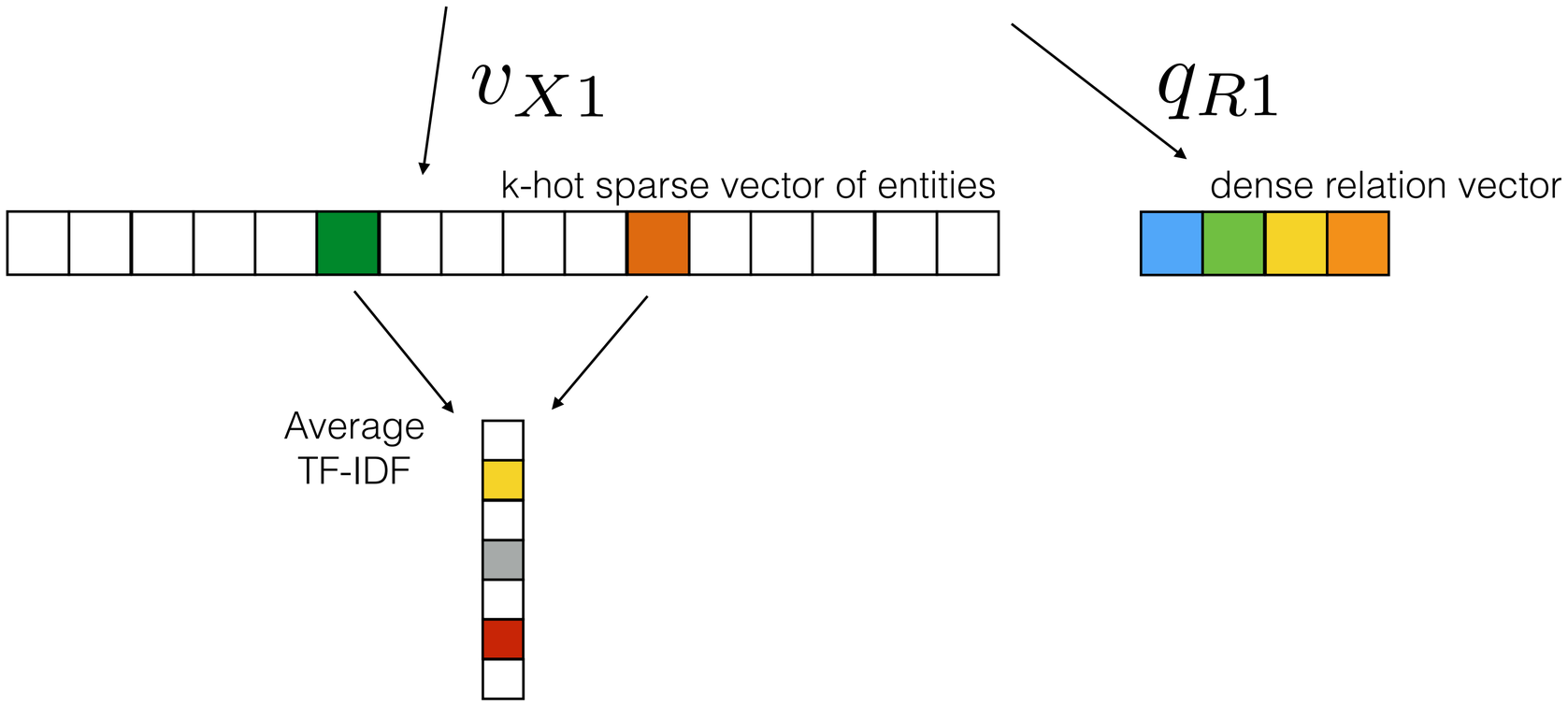
Sparse *Dense*



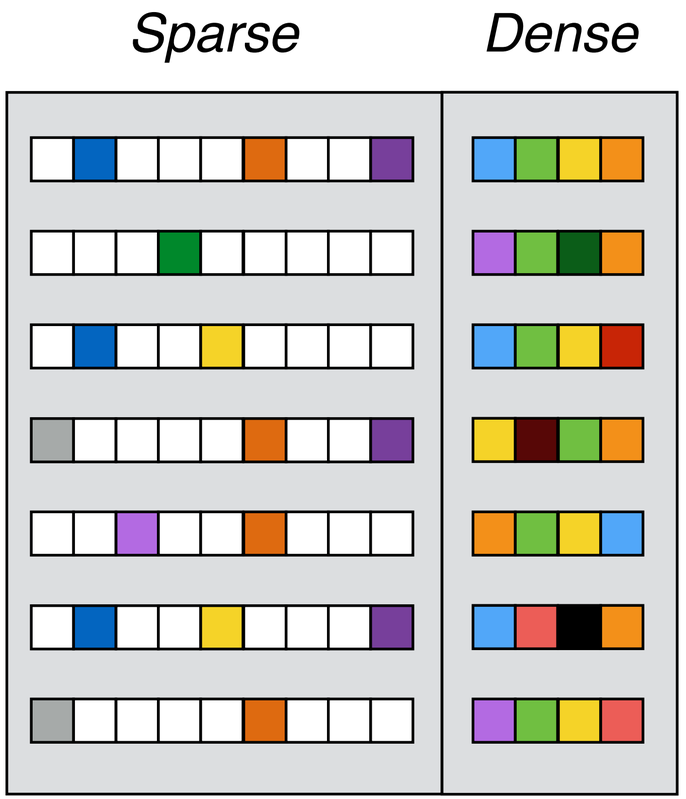
Mentions are retrieved in two steps:

1. TF-IDF against the surface form of entities

Which TV Networks has Mila Kunis appeared on?

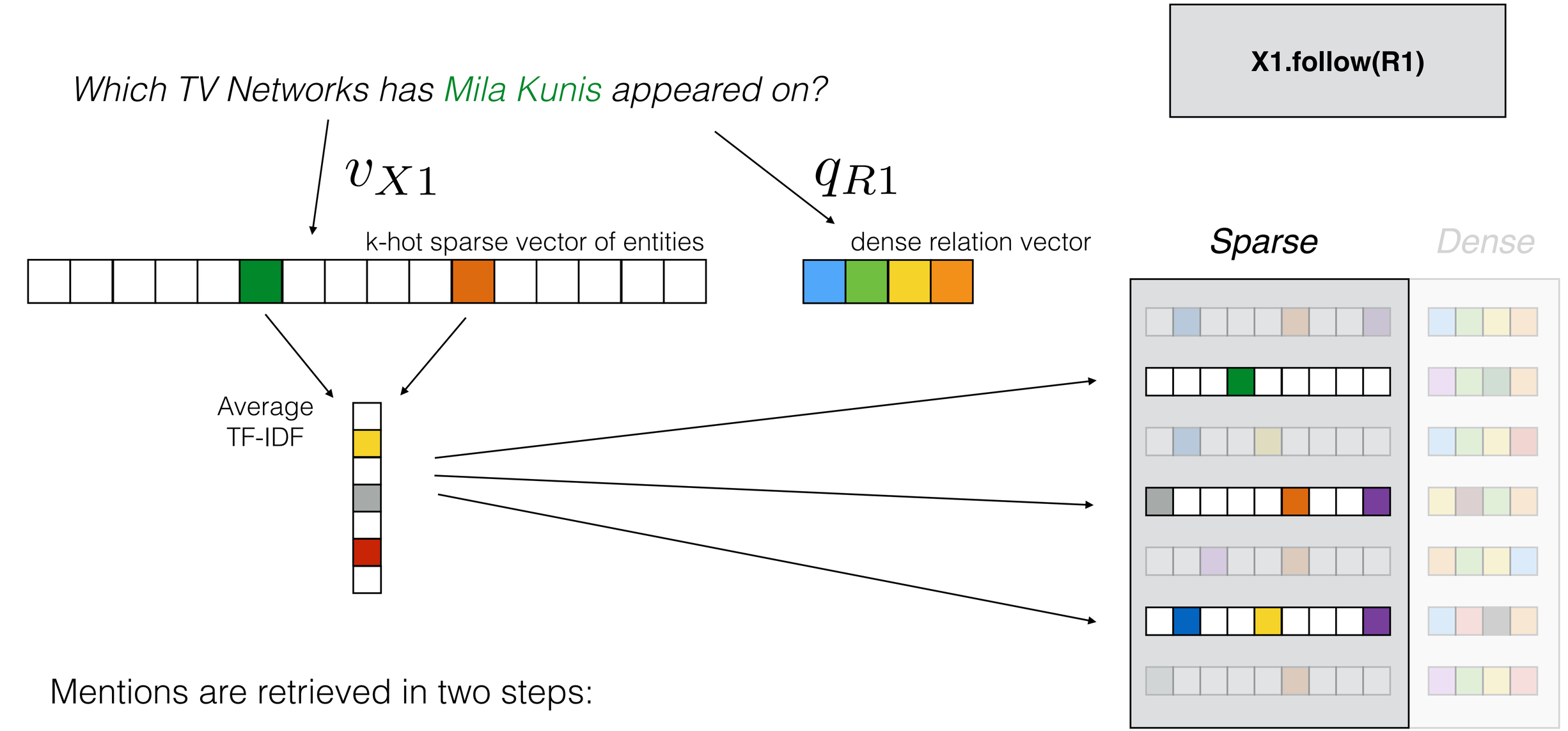


X1.follow(R1)



Mentions are retrieved in two steps:

1. TF-IDF against the surface form of entities



Mentions are retrieved in two steps:

1. TF-IDF against the surface form of entities

X1.follow(R1)

Which TV Networks has Mila Kunis appeared on?

v_{X1}

q_{R1}

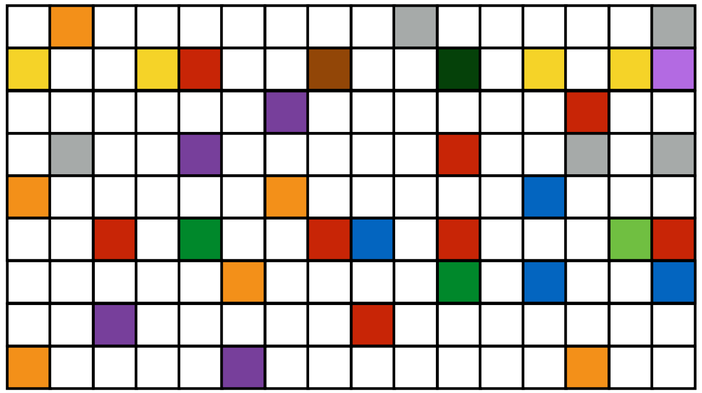
k-hot sparse vector of entities

dense relation vector

Dense

sparse vector
x
sparse matrix

#entities



#mentions

$A_{E \to M}$

This can be pre-computed

If max non-zero entries in each row is bounded (n) we can implement this efficiently using Ragged Tensors[1] —- O(k max(k, n))

[1] (https://www.tensorflow.org/guide/ragged_tensors)

Which TV Networks has Mila Kunis appeared on?

X1.follow(R1)

v_{X1}

q_{R1}

Mentions are retrieved in two steps:

1. TF-IDF against the surface form of entities
2. Dot product against relation vector

Which TV Networks has Mila Kunis appeared on?

X1.follow(R1)

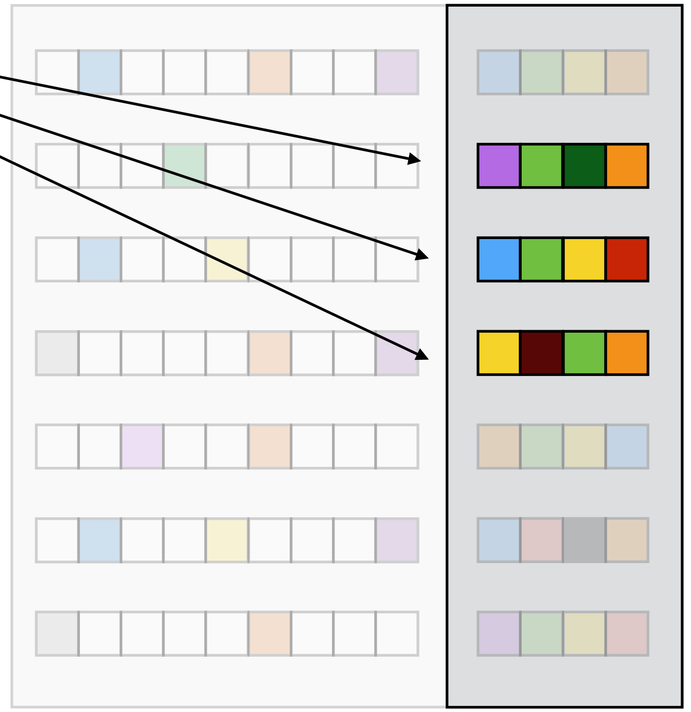
v_{X1}

q_{R1}

k-hot sparse vector of entities

dense relation vector

Sparse *Dense*



Approximate Max Inner Product Search —- O(polylog (#mentions))

Mentions are retrieved in two steps:

1. TF-IDF against the surface form of entities
2. Dot product against relation vector

Which TV Networks has Mila Kunis appeared on?

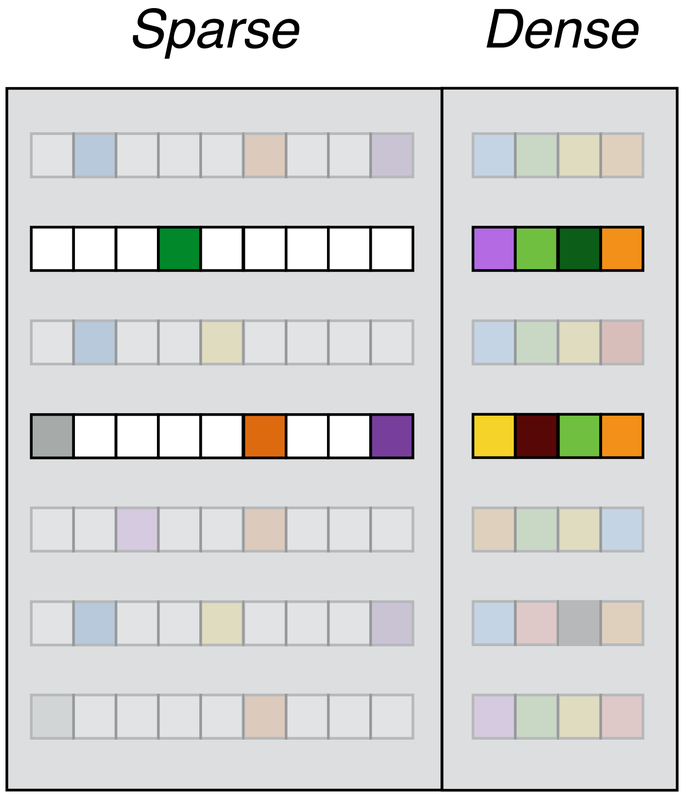
v_{X1}

q_{R1}

k-hot sparse vector of entities

dense relation vector

X1.follow(R1)



Retrieve mentions which score high using both components

- Dense part checks whether *type is correct*
- Sparse part checks whether it *co-occurs with an input entity*

Which TV Networks has Mila Kunis appeared on?

v_{X1}

q_{R1}

k-hot sparse vector of entities

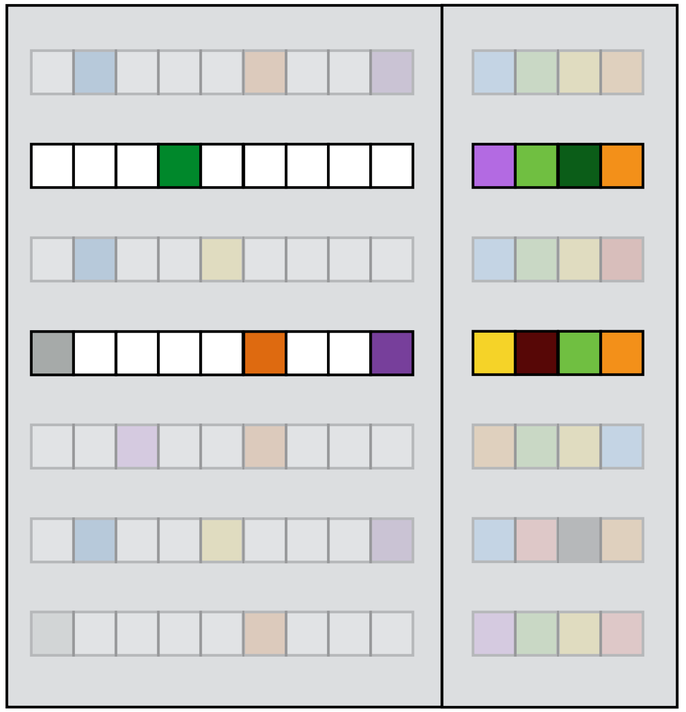
dense relation vector

Aggregate mentions to entities

(take maximum score of all coreferent mentions)

X1.follow(R1)

Sparse *Dense*



Which TV Networks has Mila Kunis appeared on?

X1.follow(R1)

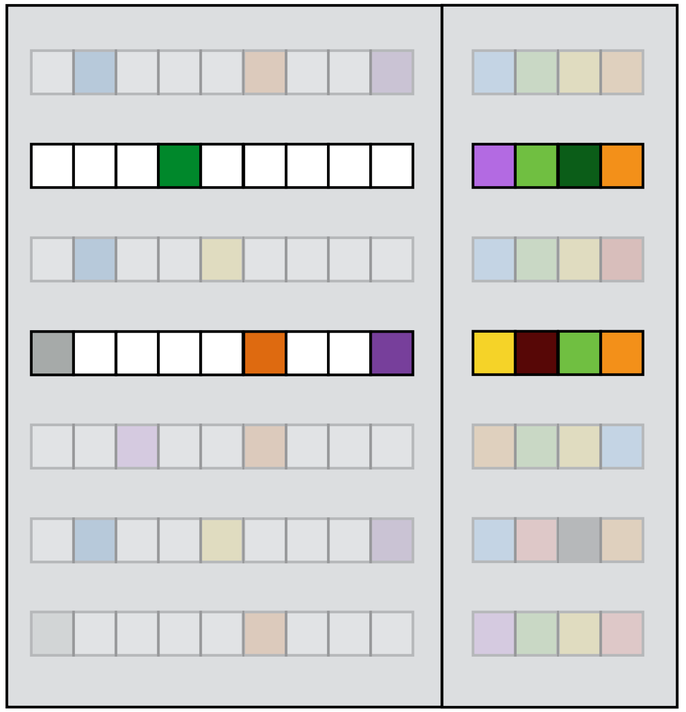
v_{X1}

q_{R1}

k-hot sparse vector of entities

dense relation vector

Sparse *Dense*



Aggregate mentions to entities

(take maximum score of all coreferent mentions)

v_{X2}

Family_Guy

That_70s_Show

Which TV Networks has Mila Kunis appeared on?

v_{X1}

q_{R1}

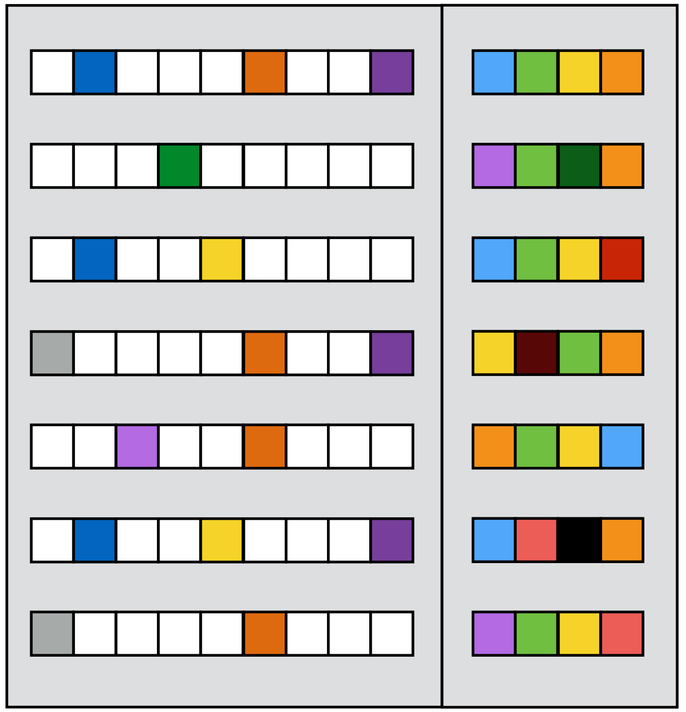
k-hot sparse vector of entities

dense relation vector

X1.follow(R1)

v_{X2}

Sparse *Dense*



Which TV Networks has Mila Kunis appeared on?

v_{X1}

q_{R1}

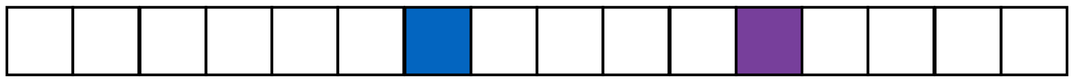
k-hot sparse vector of entities

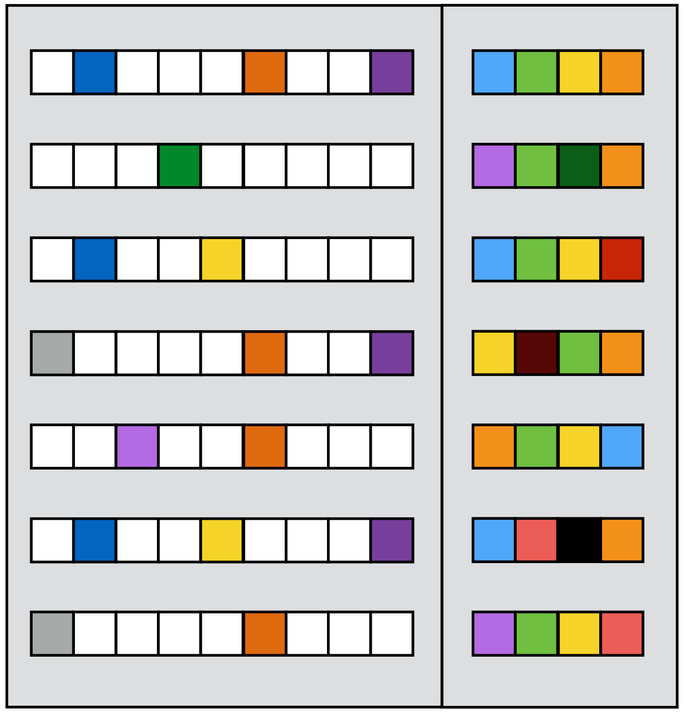
dense relation vector

X1.follow(R1)

v_{X2}

q_{R2}





Which TV Networks has Mila Kunis appeared on?

v_{X1}

q_{R1}

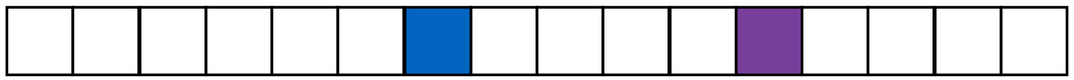
k-hot sparse vector of entities

dense relation vector

X1.follow(R1)

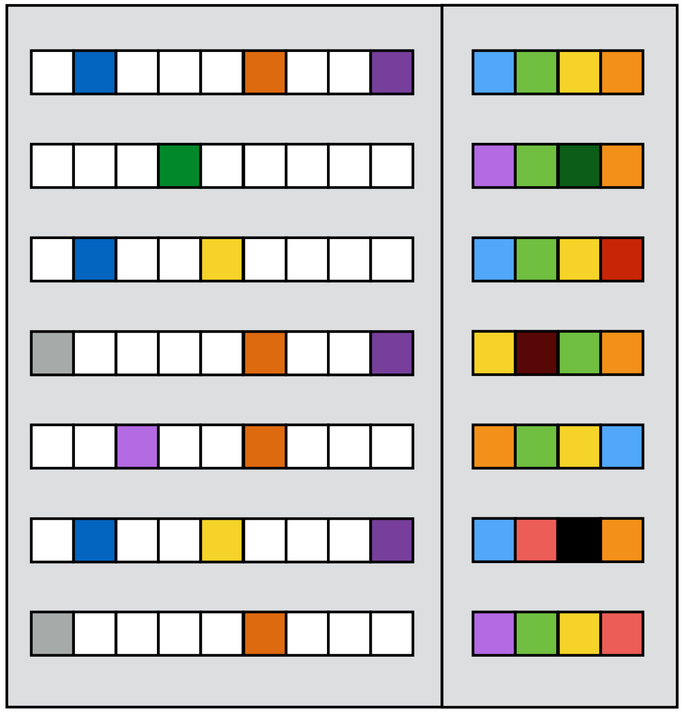
v_{X2}

q_{R2}



X2.follow(R2)

Sparse *Dense*



Summary

Sparse vector of entities

Dense relation vector

$$v_{X.\text{follow}(R)} = [v_X A_{E \to M} \odot \mathcal{T}_K(q_R)] A_{M \to E}$$

Entity -> Mention TF-IDF matrix

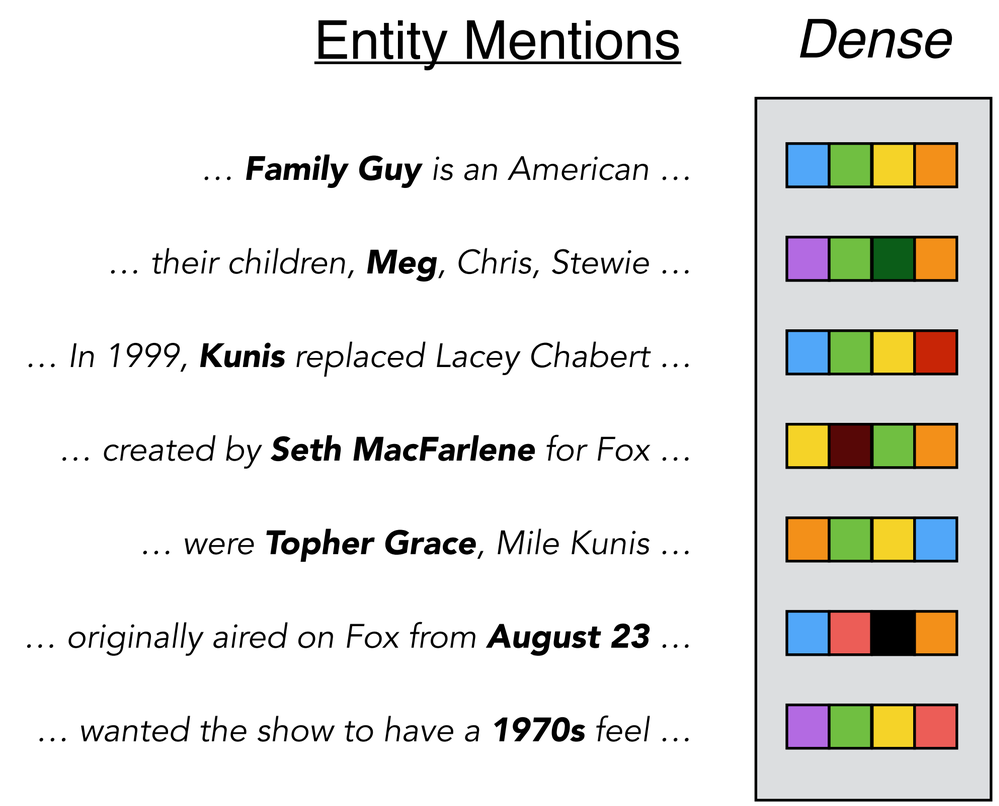
Top-K inner product search

Mention -> Entity Coreference Matrix

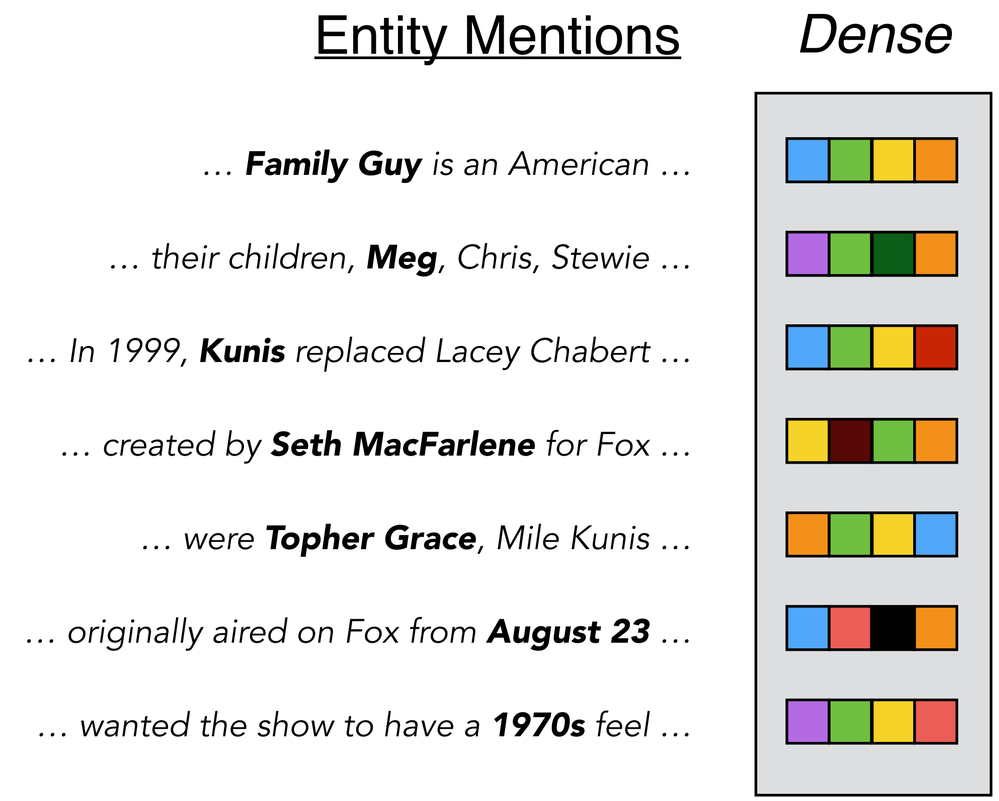
$\odot$ Element-wise Product

- Every operation is differentiable —- can learn from denotations!
- Complexity depends only on **log** of #entities / #mentions!

Pre-training



Pre-training



- Distantly align KB facts to text passages:

Pre-training

Entity Mentions

… ***Family Guy*** *is an American …*

… their children, ***Meg****, Chris, Stewie …*

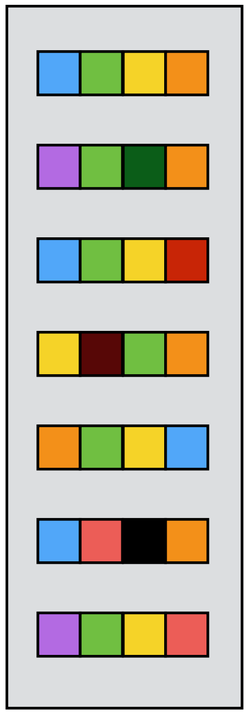
… In 1999, ***Kunis*** *replaced Lacey Chabert …*

… created by ***Seth MacFarlene*** *for Fox …*

… were ***Topher Grace****, Mile Kunis …*

… originally aired on Fox from ***August 23*** *…*

… wanted the show to have a ***1970s*** *feel …*



- Distantly align KB facts to text passages:

KB fact
(Quentin Tarantino, directed, Pulp Fiction)

Pulp Fiction is a 1994 film written and directed by Quentin Tarantino
Text

Pre-training

Entity Mentions

… ***Family Guy*** *is an American …*

… their children, ***Meg****, Chris, Stewie …*

… In 1999, ***Kunis*** *replaced Lacey Chabert …*

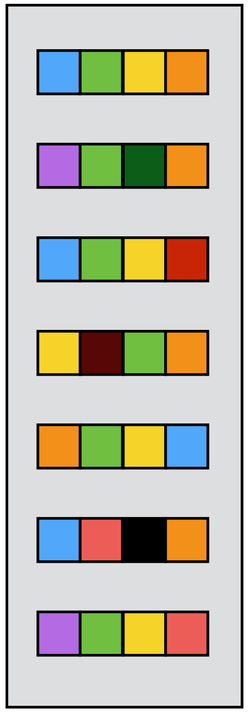
… created by ***Seth MacFarlene*** *for Fox …*

… were ***Topher Grace****, Mile Kunis …*

… originally aired on Fox from ***August 23*** *…*

… wanted the show to have a ***1970s*** *feel …*

Dense



- Distantly align KB facts to text passages:

KB fact
(Quentin Tarantino, directed, Pulp Fiction)

Pulp Fiction is a 1994 film written and directed by Quentin Tarantino
Text

Pre-training

Entity Mentions

… ***Family Guy*** *is an American …*

… their children, ***Meg****, Chris, Stewie …*

… In 1999, ***Kunis*** *replaced Lacey Chabert …*

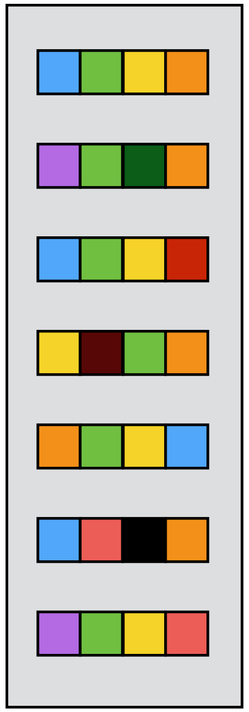
… created by ***Seth MacFarlene*** *for Fox …*

… were ***Topher Grace****, Mile Kunis …*

… originally aired on Fox from ***August 23*** *…*

… wanted the show to have a ***1970s*** *feel …*

Dense



- Distantly align KB facts to text passages:

KB fact

(Quentin Tarantino, directed, Pulp Fiction)

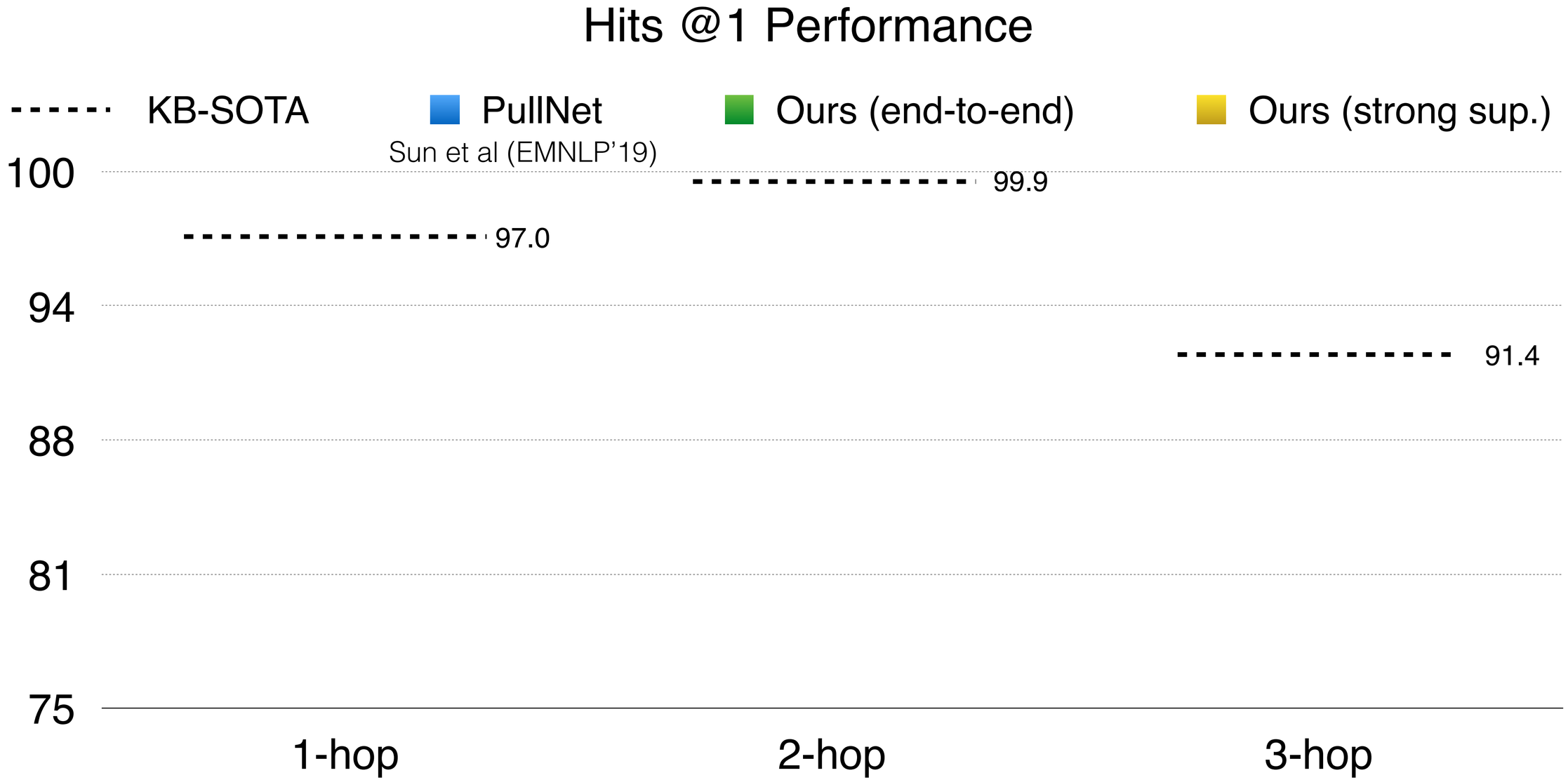
Pulp Fiction is a 1994 film written and directed by Quentin Tarantino

Text

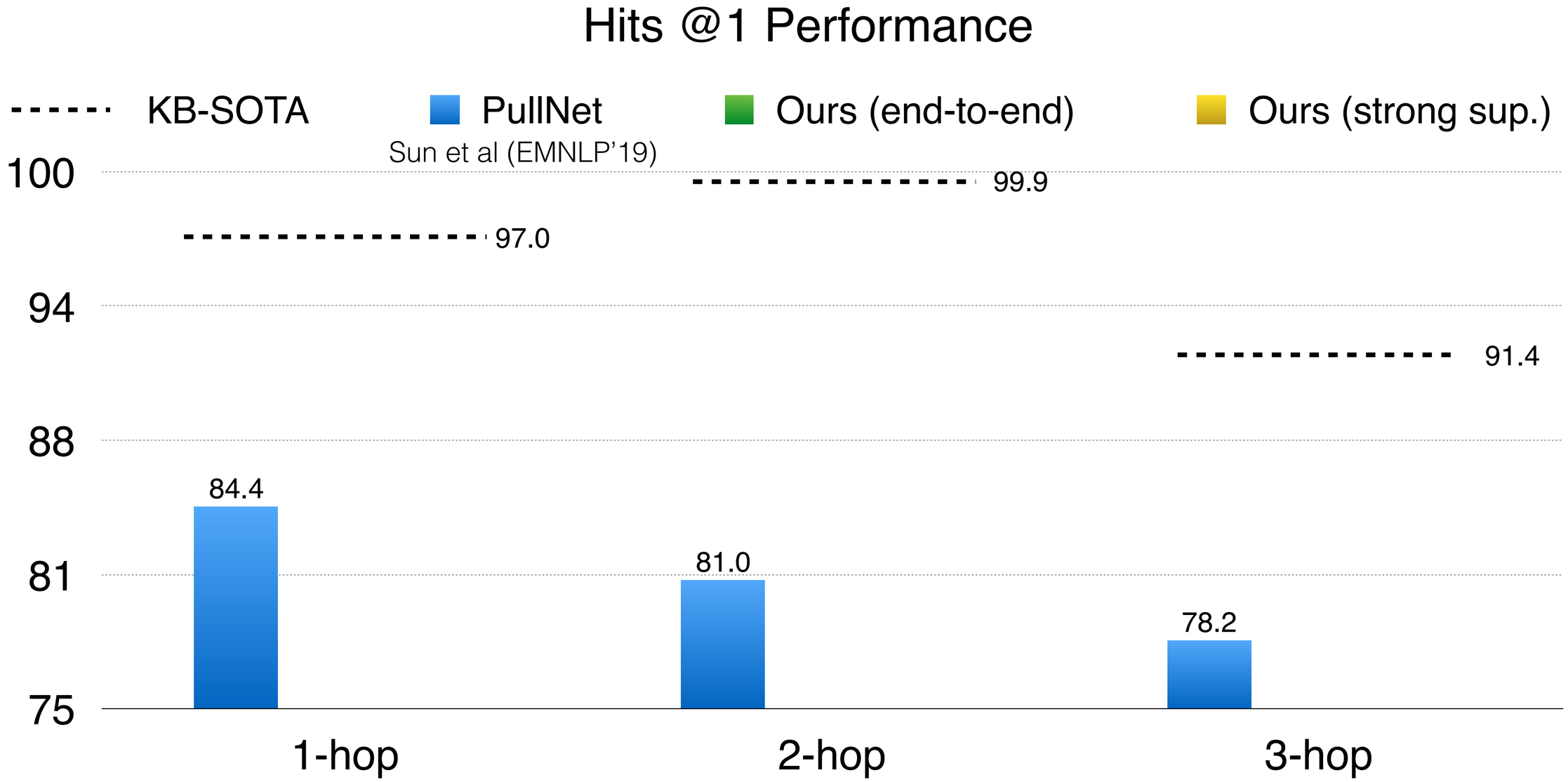
MetaQA Results

Hits @1 Performance

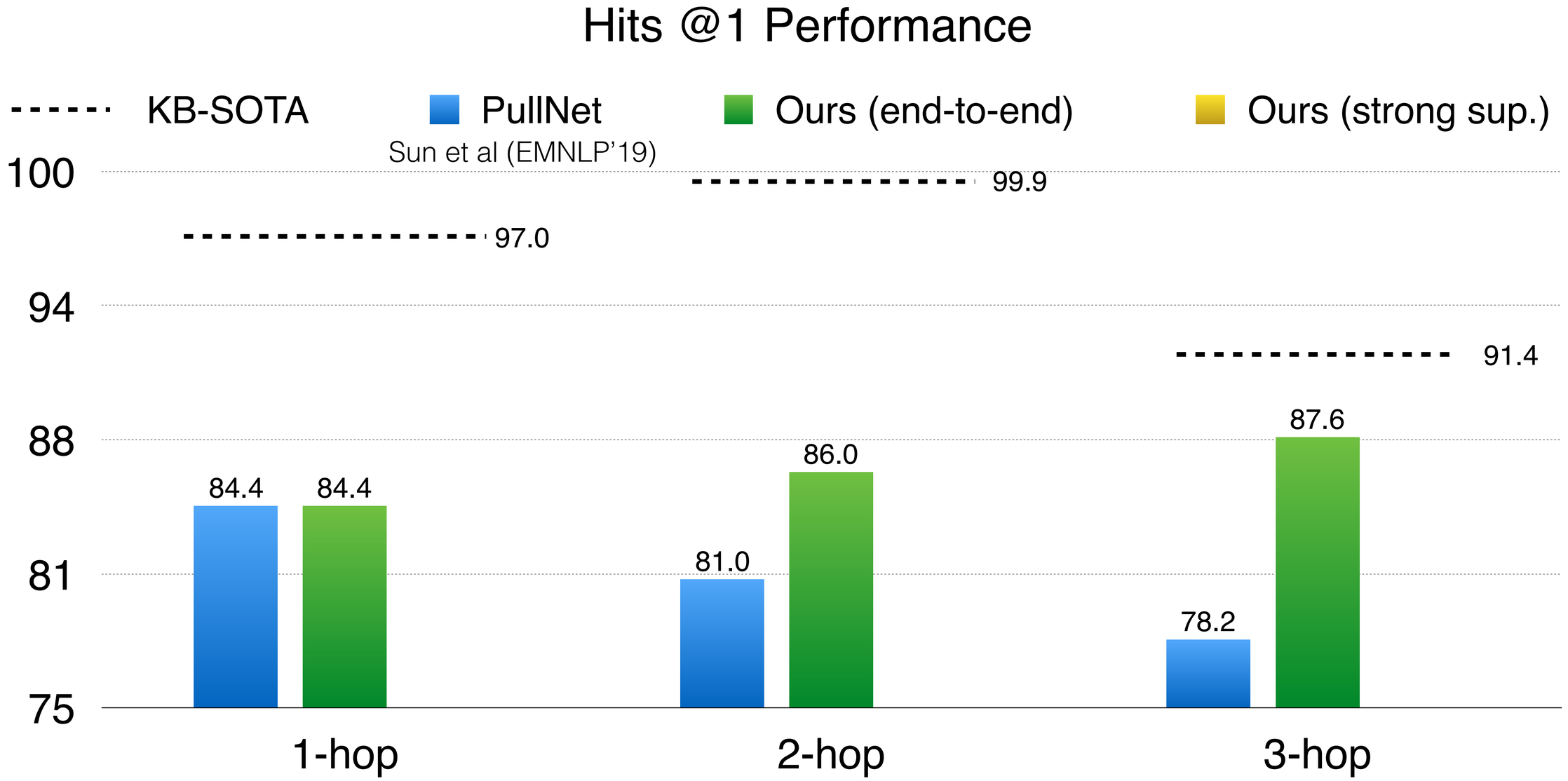
MetaQA Results



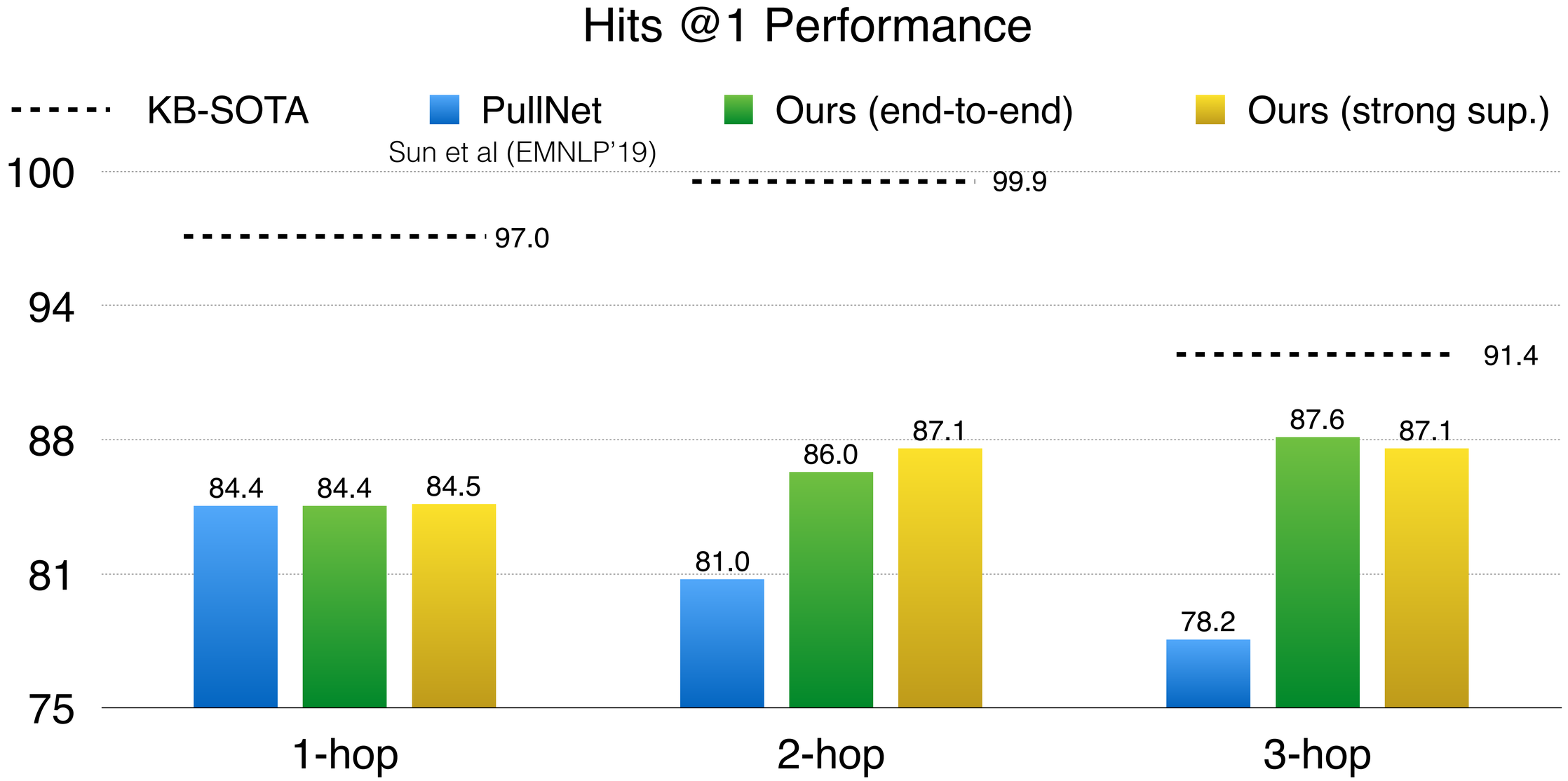
MetaQA Results



MetaQA Results



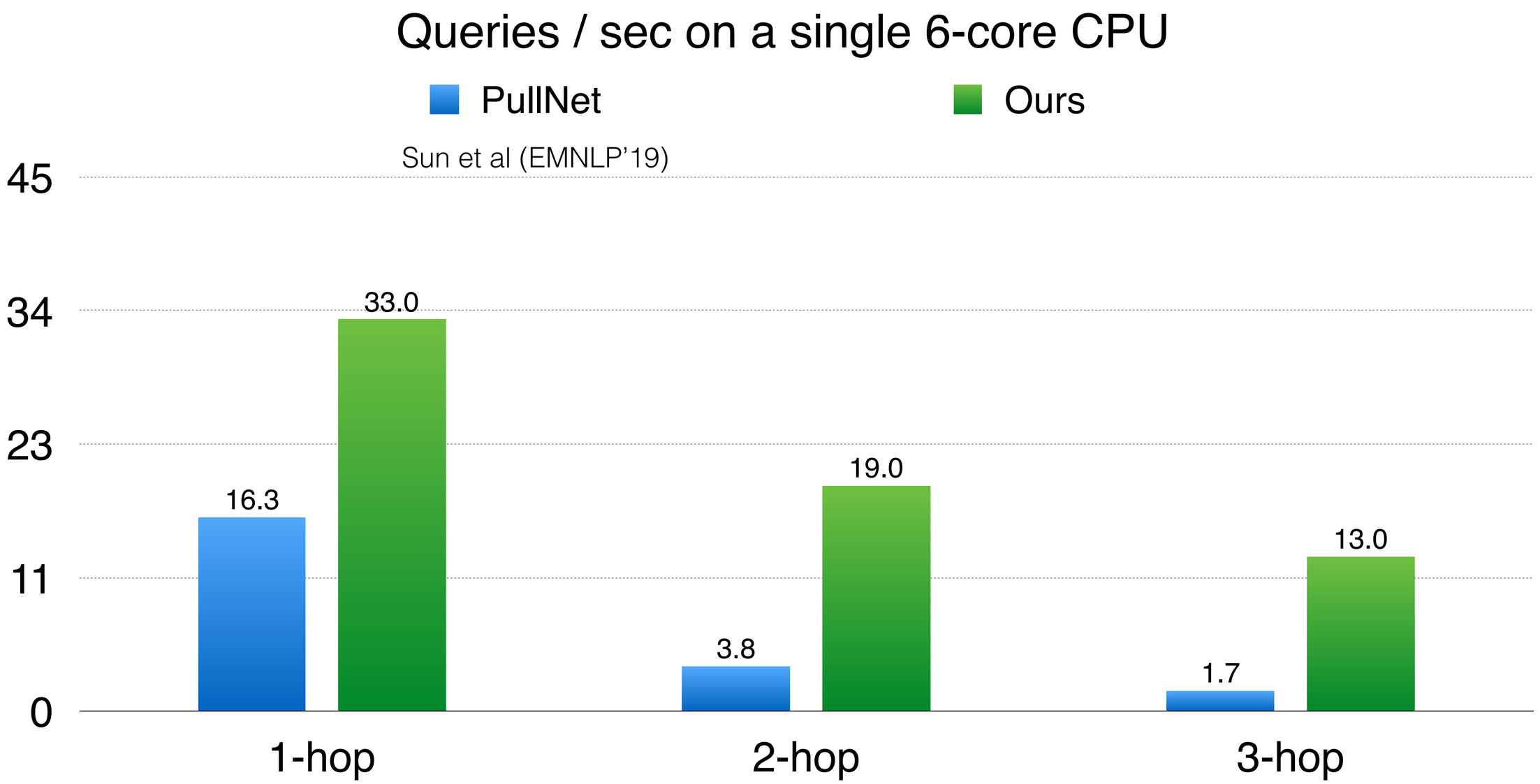
MetaQA Results



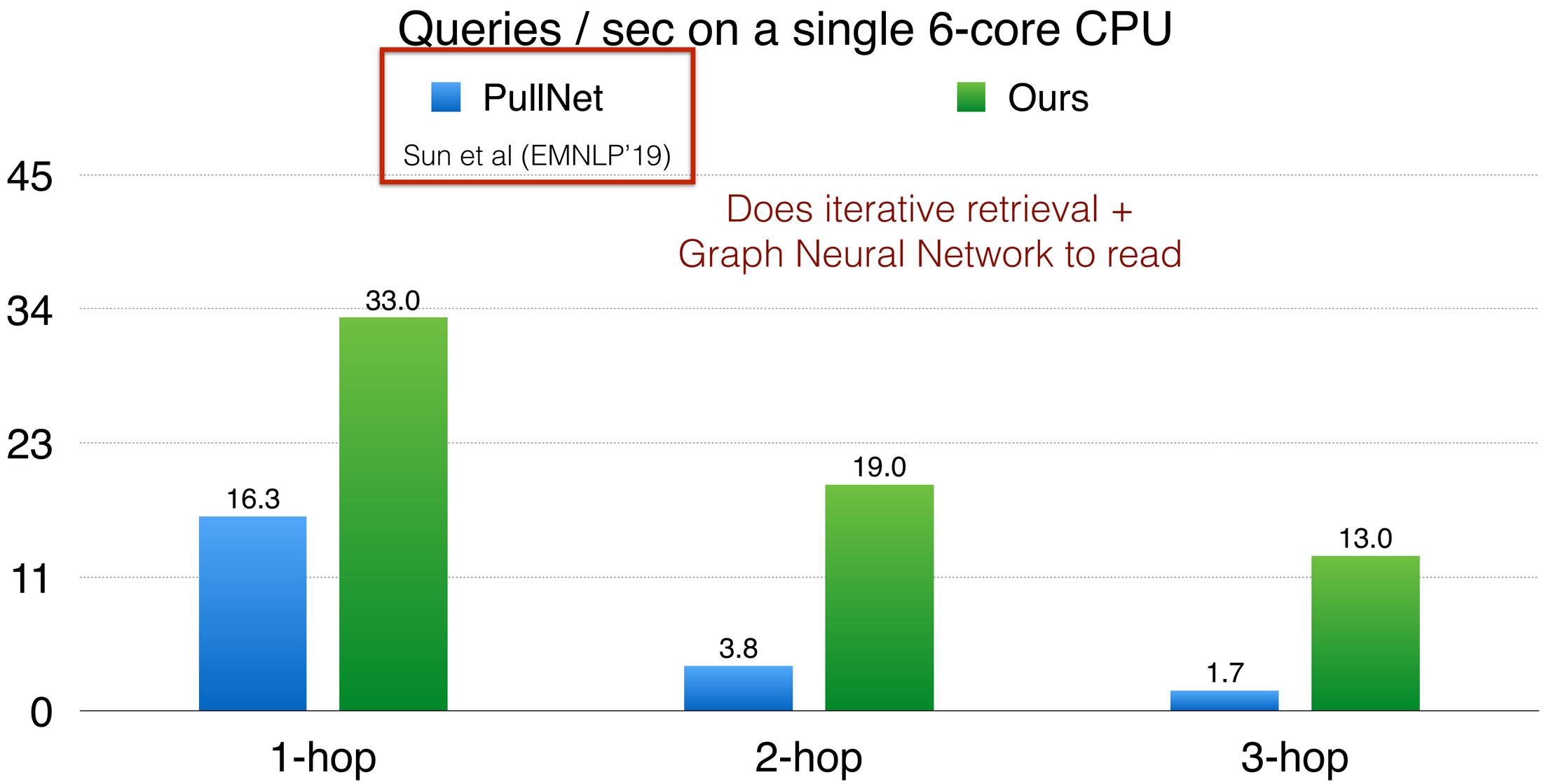
MetaQA Results

Queries / sec on a single 6-core CPU

MetaQA Results



MetaQA Results



New Dataset: WikiData Slot-filling

New Dataset: WikiData Slot-filling

- Constructed by aligning WikiData facts to Wikipedia

New Dataset: WikiData Slot-filling

- Constructed by aligning WikiData facts to Wikipedia
- 1-3 hop semi-synthetic *slot-filling* queries over ~300 relations

  Q. Marcel de Graaff, place of birth, twinned administrative body?
  Ans. Rotterdam -> Antwerp

  Q. Muhammad Sanya, member of political party, chairperson, date of birth?
  Ans. Civic United Front -> Ibrahim Lipumba -> 6 June 1952

New Dataset: WikiData Slot-filling

- Constructed by aligning WikiData facts to Wikipedia
- 1-3 hop semi-synthetic *slot-filling* queries over ~300 relations

  Q. Marcel de Graaff, place of birth, twinned administrative body?
  Ans. Rotterdam -> Antwerp

  Q. Muhammad Sanya, member of political party, chairperson, date of birth?
  Ans. Civic United Front -> Ibrahim Lipumba -> 6 June 1952

- Task is to extract answer from an **unseen** corpus of 10K Wikipedia articles (120K passages)
  - ~200K entities
  - ~1.2M mentions

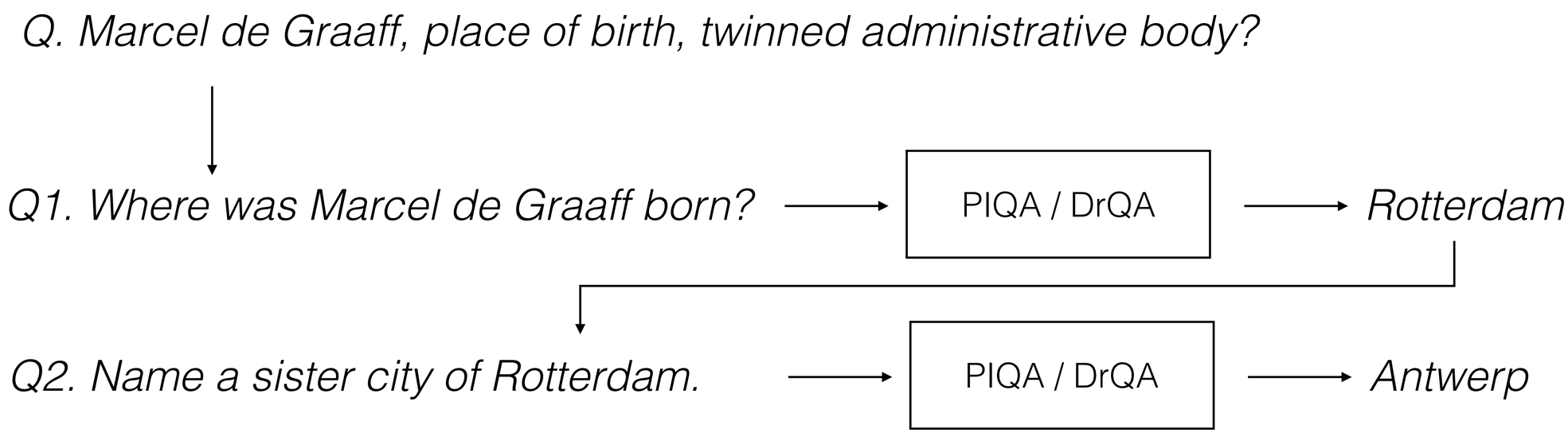
Baselines

- Cascaded versions of two open-domain QA models - PIQA [1] and DrQA [2]
- Relations were converted to natural language questions using templates

[1] Seo, Minjoon, et al. "Real-Time Open-Domain Question Answering with Dense-Sparse Phrase Index." ACL (2019).
[2] Chen, Danqi, et al. "Reading wikipedia to answer open-domain questions." ACL (2017).

Baselines

- Cascaded versions of two open-domain QA models - PIQA [1] and DrQA [2]
- Relations were converted to natural language questions using templates

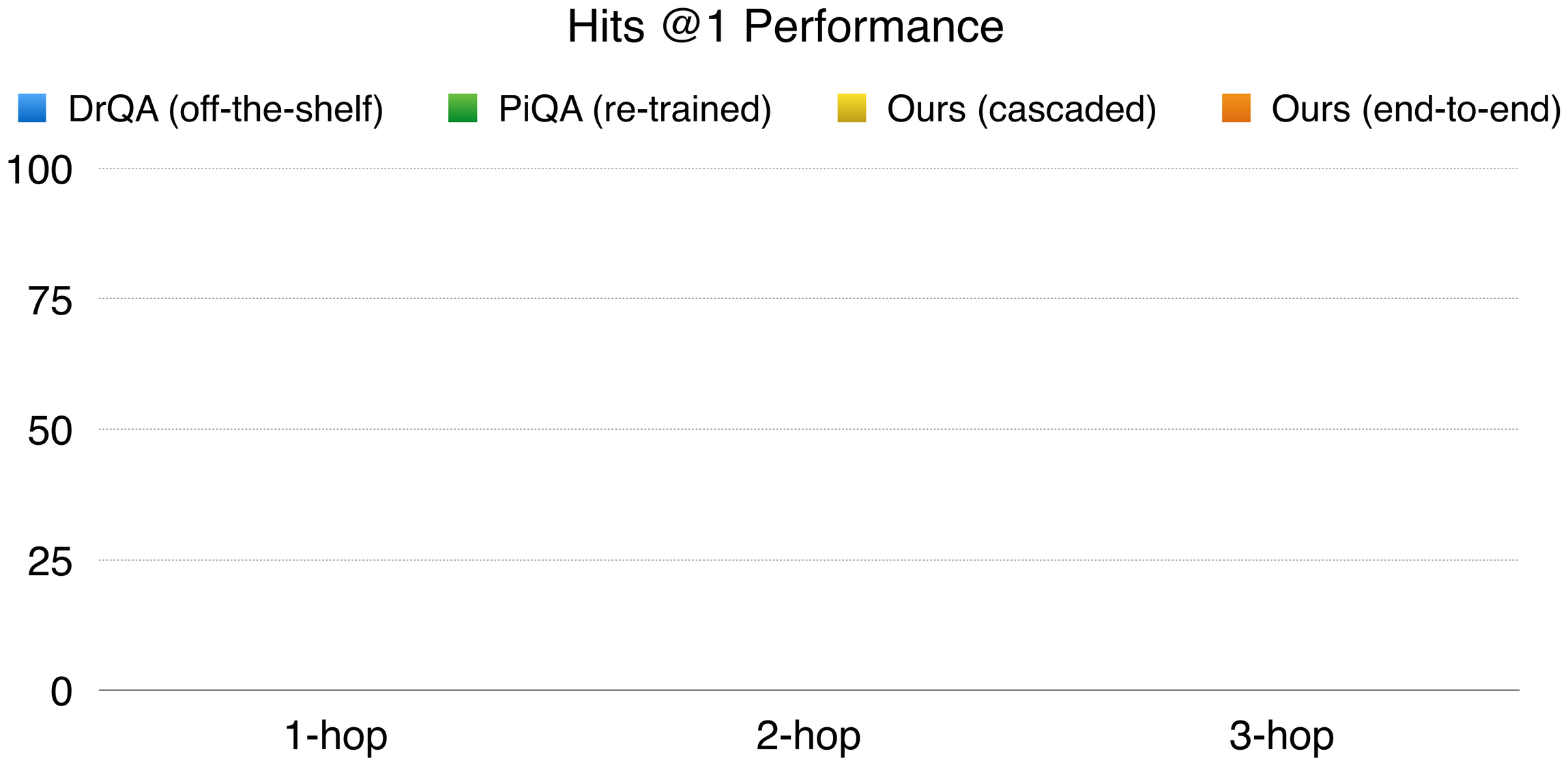


[1] Seo, Minjoon, et al. "Real-Time Open-Domain Question Answering with Dense-Sparse Phrase Index." ACL (2019).
[2] Chen, Danqi, et al. "Reading wikipedia to answer open-domain questions." ACL (2017).

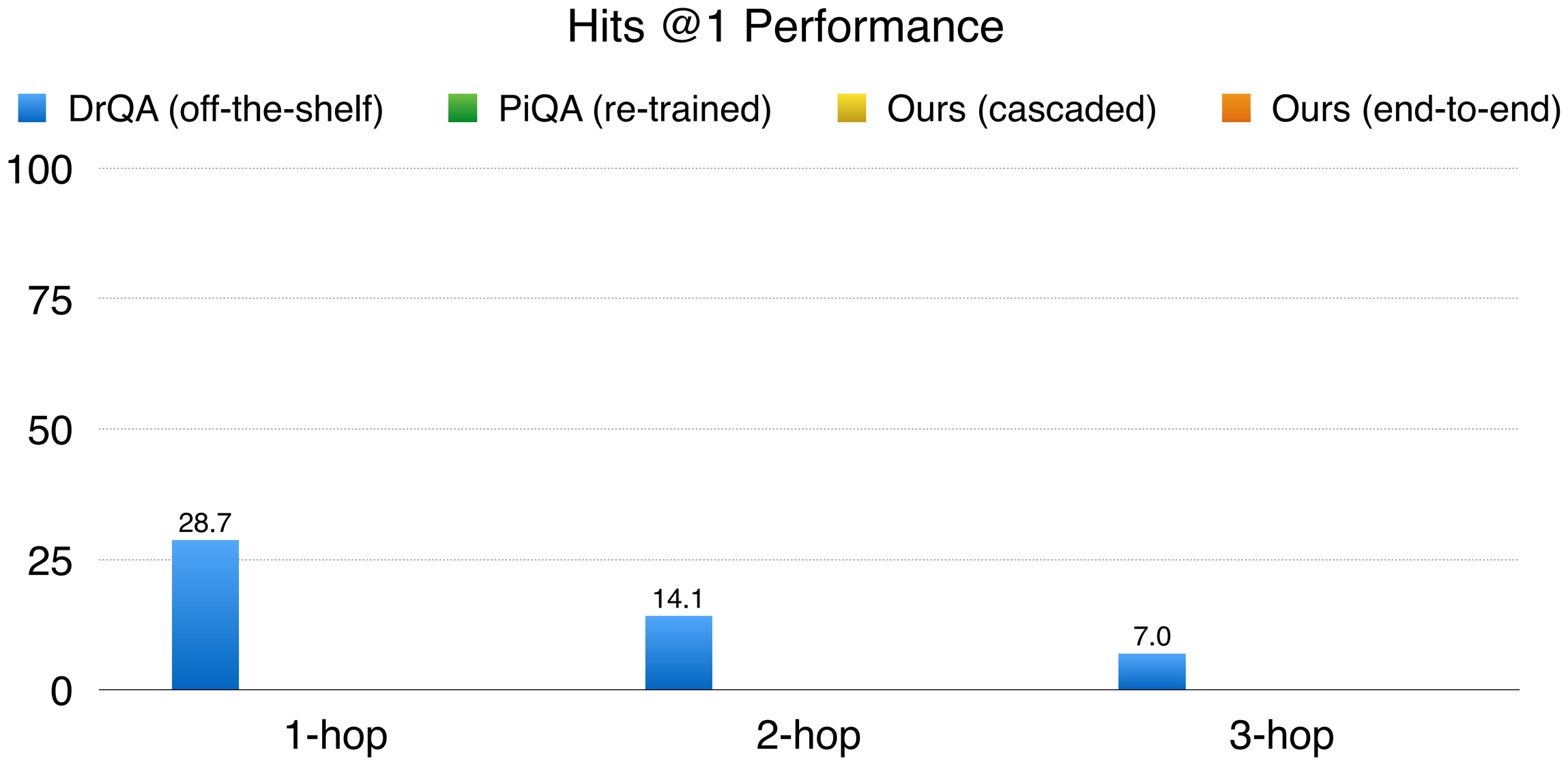
WikiData Slot-Filling Results

Hits @1 Performance

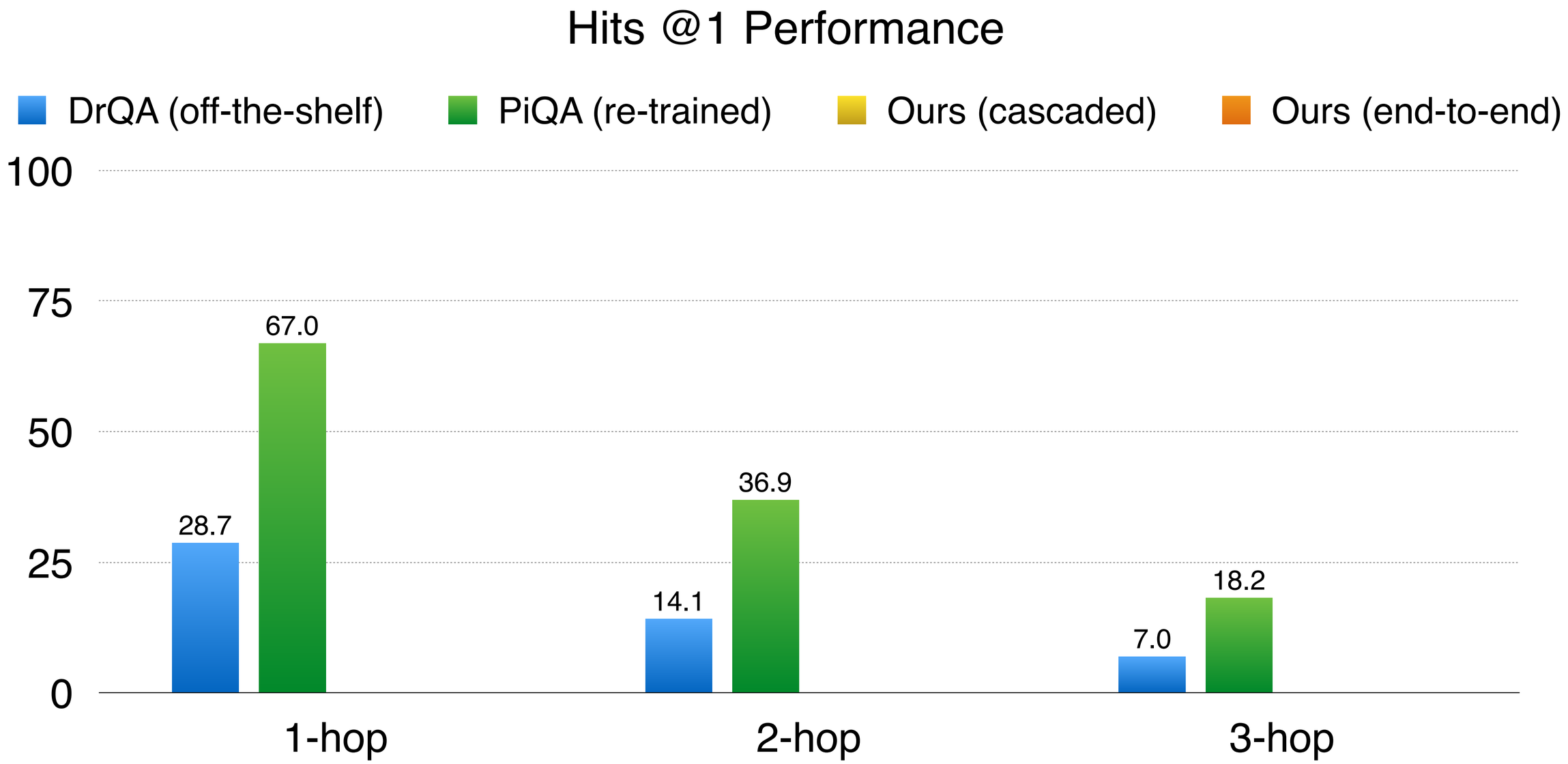
WikiData Slot-Filling Results



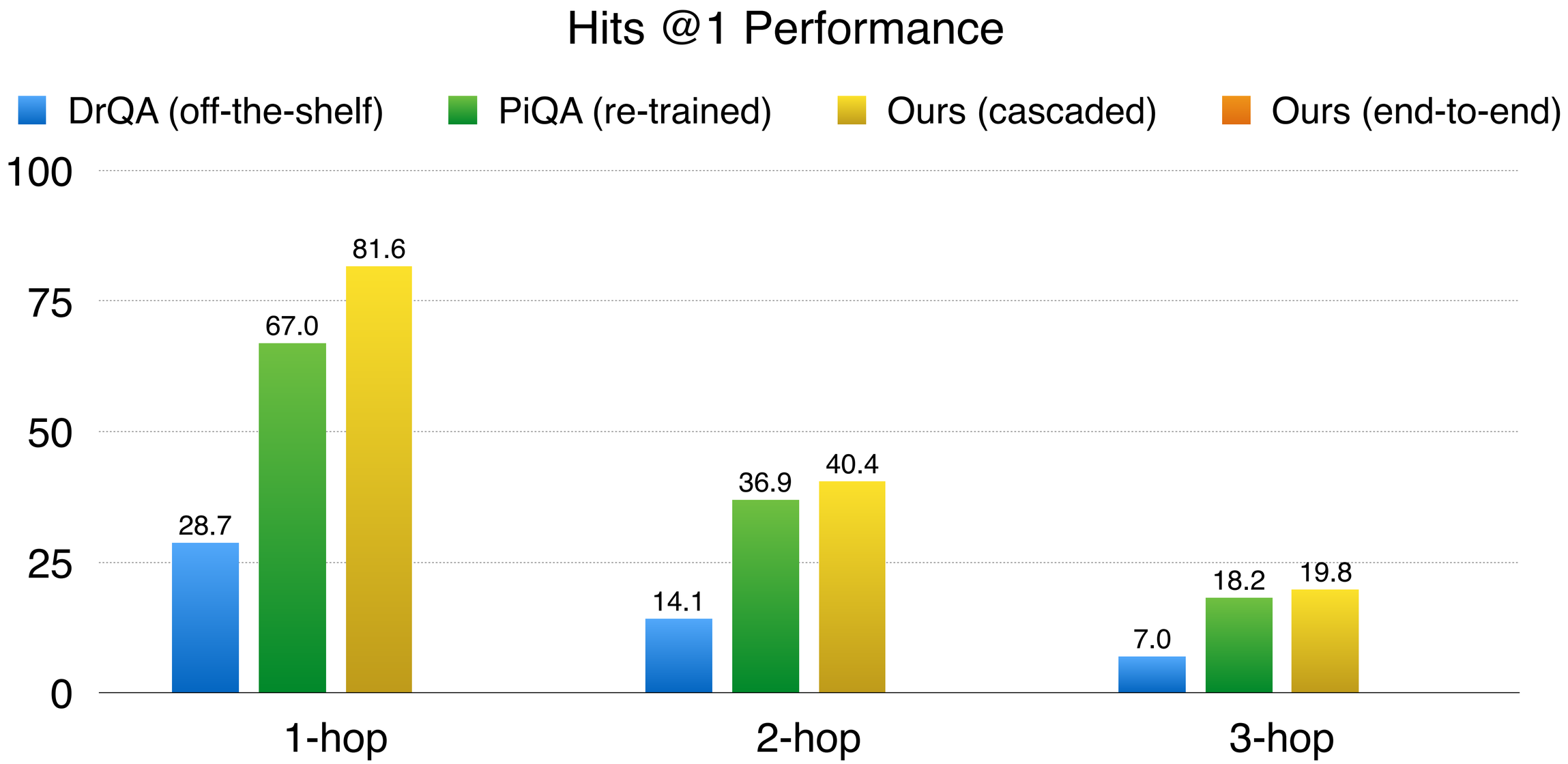
WikiData Slot-Filling Results



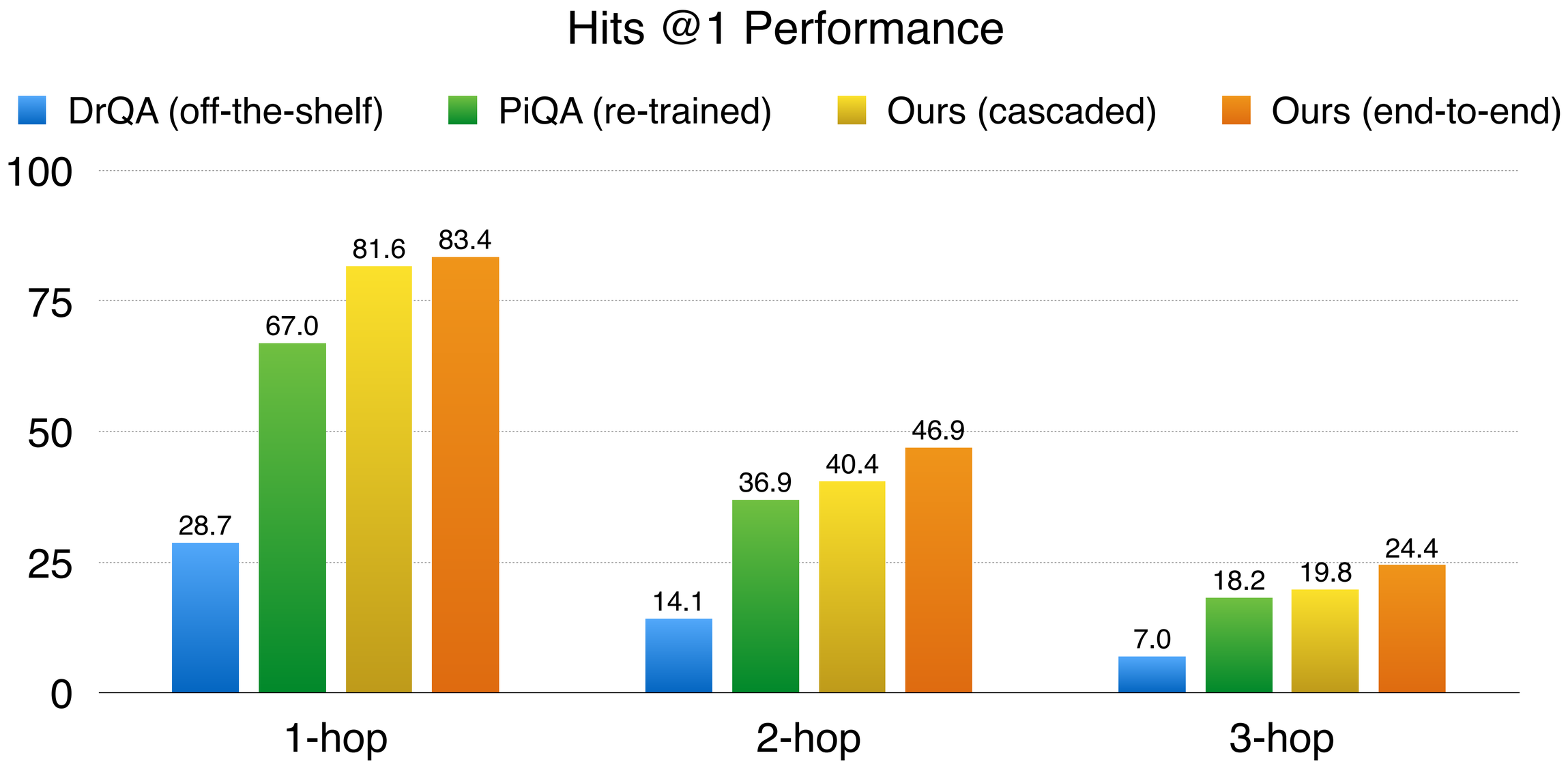
WikiData Slot-Filling Results



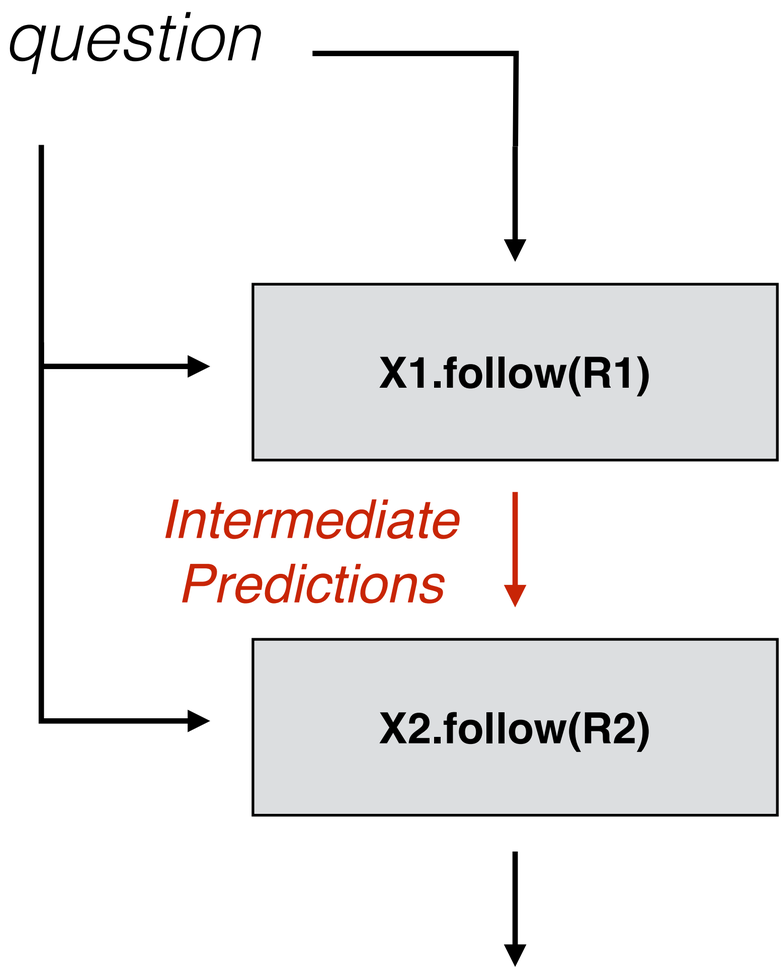
WikiData Slot-Filling Results



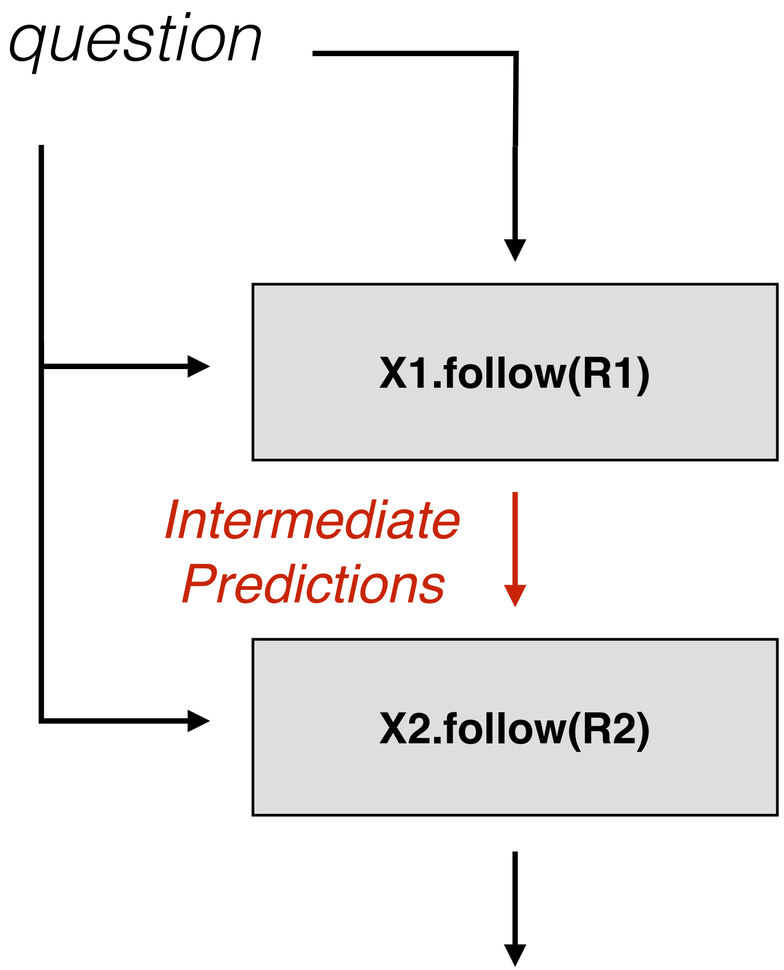
WikiData Slot-Filling Results



Analysis - Intermediate Predictions

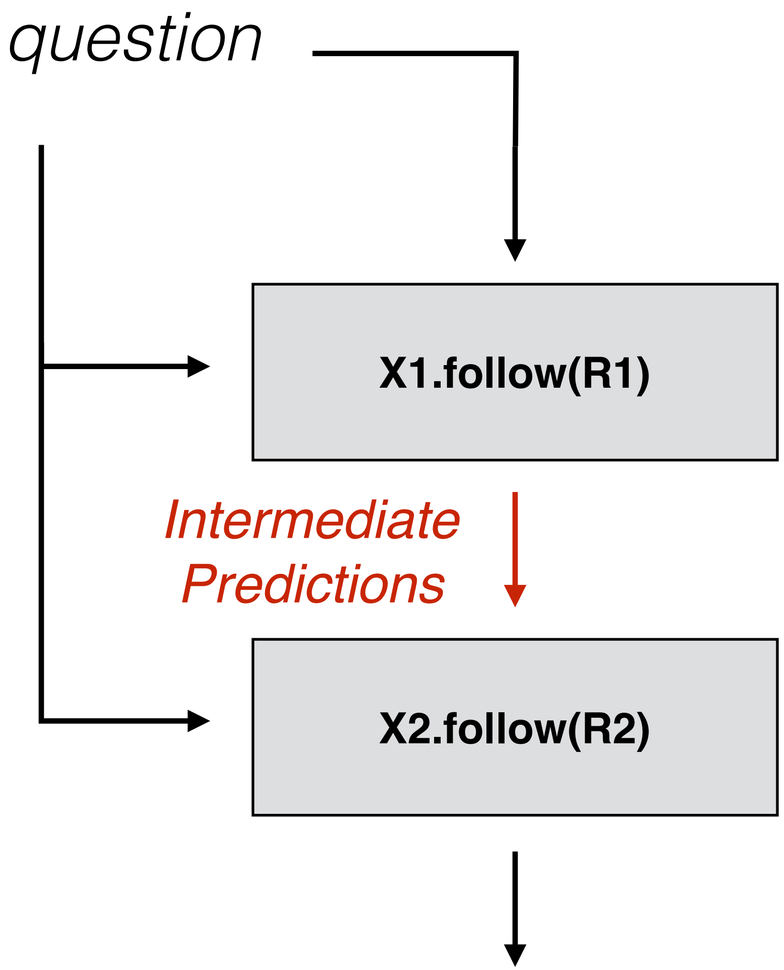


Analysis - Intermediate Predictions



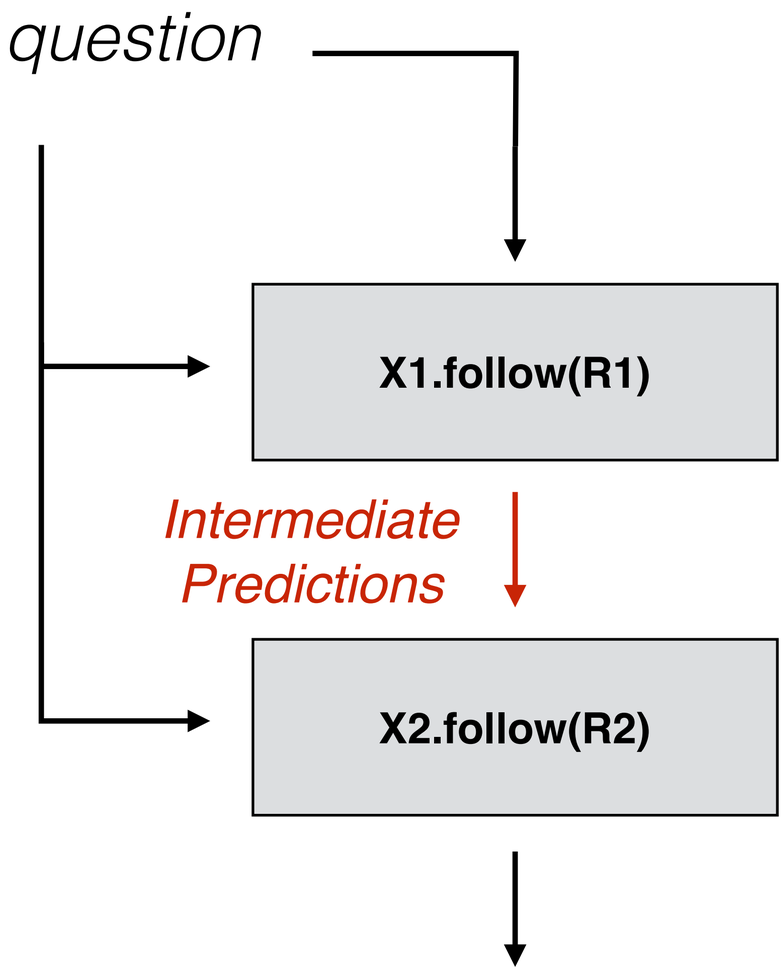
- Inspected 100 **correctly answered** 2-hop questions

Analysis - Intermediate Predictions



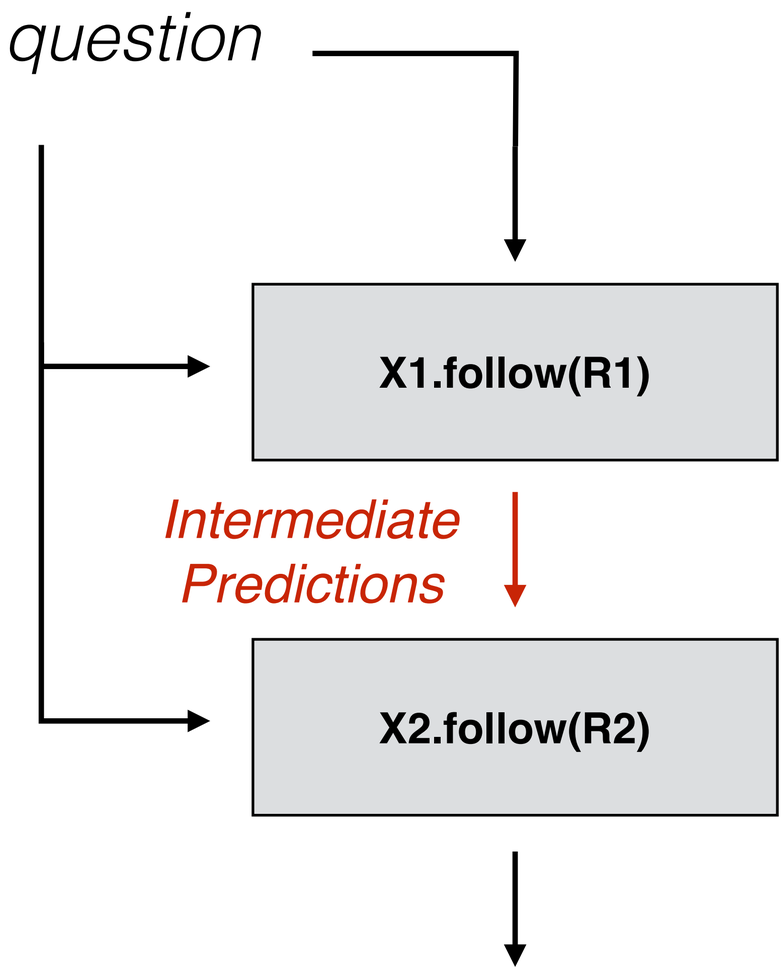
- Inspected 100 **correctly answered** 2-hop questions
- **83%** had at least one correct intermediate prediction

Analysis - Intermediate Predictions



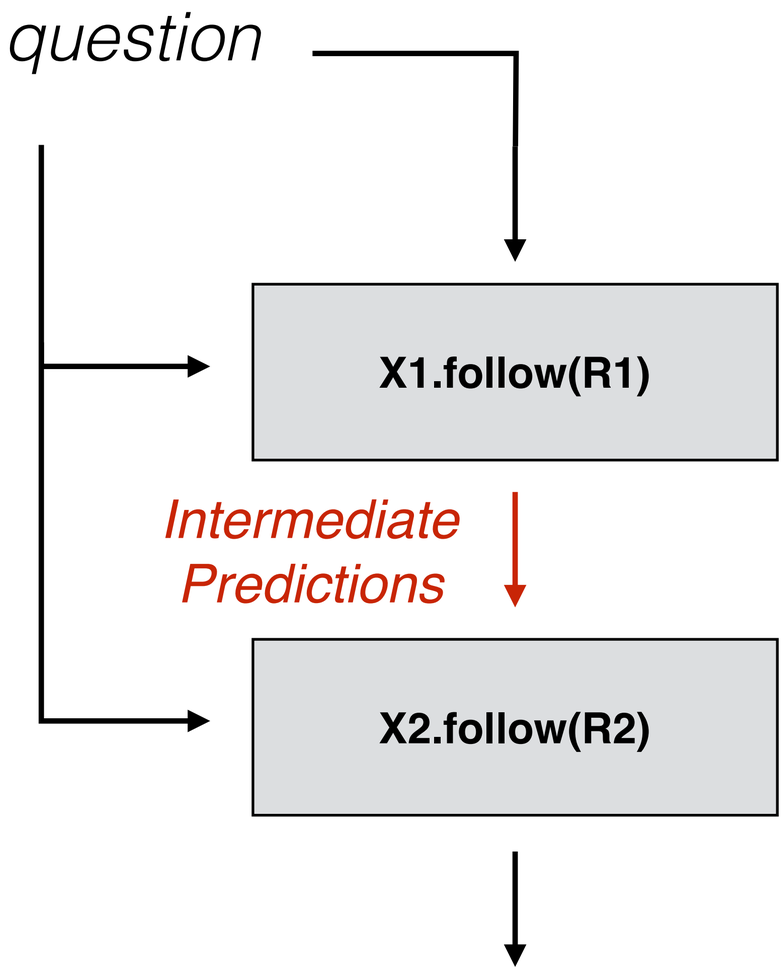
- Inspected 100 **correctly answered** 2-hop questions
- **83%** had at least one correct intermediate prediction
- For **80%** the most confident intermediate prediction was correct
  - This drops to **47%** for **incorrectly answered** questions

Analysis - Intermediate Predictions



- Inspected 100 **correctly answered** 2-hop questions
- **83%** had at least one correct intermediate prediction
- For **80%** the most confident intermediate prediction was correct
  - This drops to **47%** for **incorrectly answered** questions
- Rest **17%** the model learned to answer in a single hop:

Analysis - Intermediate Predictions



- Inspected 100 **correctly answered** 2-hop questions
- **83%** had at least one correct intermediate prediction
- For **80%** the most confident intermediate prediction was correct
  - This drops to **47%** for **incorrectly answered** questions
- Rest **17%** the model learned to answer in a single hop:

What are the genres of the films directed by Justin Simien?
Intermediate = ***drama****, Final =* ***drama***

What genres are the movies acted by Jeremy Lin in?
Intermediate = ***documentary****, Final =* ***documentary***

Conclusion

- Word embeddings provide linguistic knowledge to NLP systems
- Can mention / span embeddings provide world knowledge?

Future Directions

Enriching the Knowledge Base

- Beyond entities and relations
- Beyond English
- Beyond Text

Expanding to tasks other than QA

- Language Modeling
- Dialog
- Translation

Thank You.