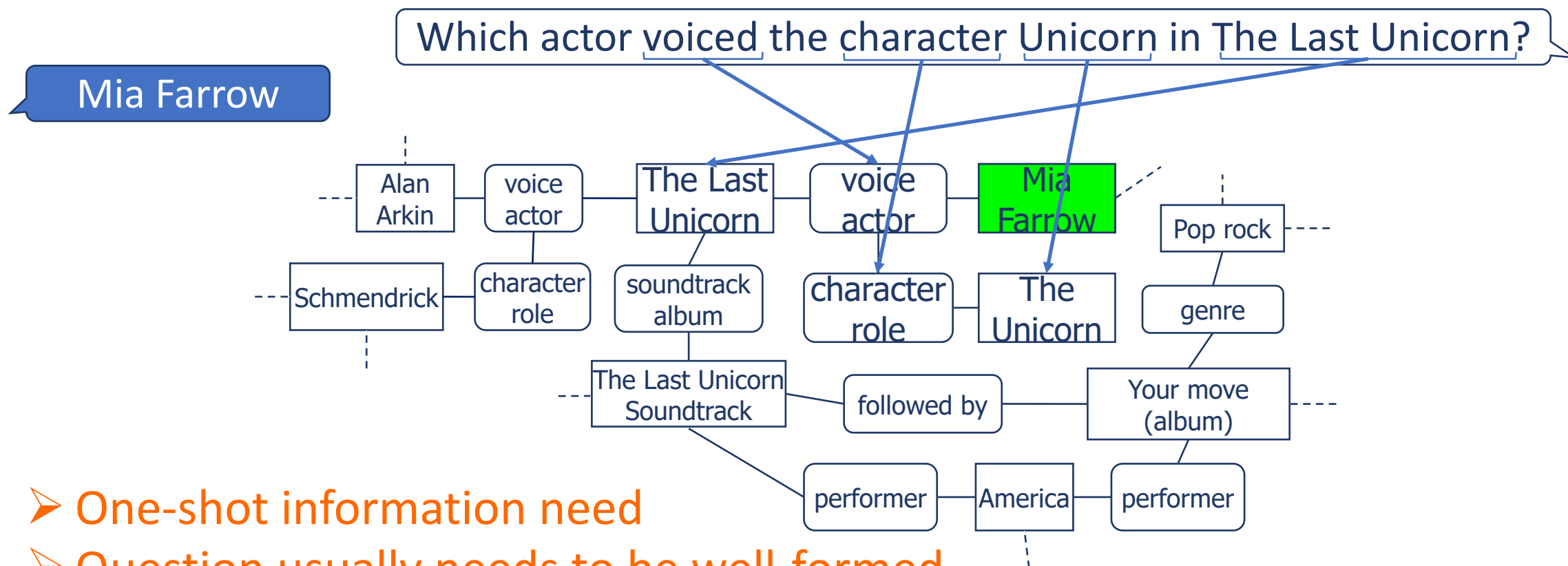


Look before you Hop: Conversational Question Answering over Knowledge Graphs Using Judicious Context Expansion

Philipp Christmann, Rishiraj Saha Roy, Abdalghani Abujabal, Jyotsna Singh, Gerhard Weikum

Max Planck Institute for Informatics, Saarbrücken, Germany

Question Answering over KGs



- One-shot information need
- Question usually needs to be well-formed

[Abujabal et al. (2018), Diefenbach et al. (2019), Huang et al. (2019)]

Conversational KG-QA

Mia Farrow

Schmendrick

America

Folk rock

Jules Bass

Which actor voiced the character Unicorn in The Last Unicorn?

And Alan Arkin was behind . . . ?

So who performed songs?

Genre of this band?

By the way, who directed the movie?

- Information needs rarely one-shot
- Natural mode of interaction

Conversational KG-QA

Complete

{ Which actor voiced the character Unicorn in The Last Unicorn?

Incomplete

{ And Alan Arkin was behind . . . ?

{ So who performed songs?

{ Genre of this band?

{ By the way, who directed the movie?

Conversational KG-QA

➤ Ad hoc

And Alan Arkin was behind . . . ?

So who performed songs?

Genre of this band?

By the way, who directed the movie?

Conversational KG-QA

➤ Ad hoc

➤ Ungrammatical

And Alan Arkin was behind . . .?

So who performed songs?

Genre of this band?

By the way, who directed the movie?

Conversational KG-QA

➤ Ad hoc

And Alan Arkin was behind . . . ?

➤ Ungrammatical

So who performed songs?

➤ Information left out

Genre of this band?

By the way, who directed the movie?

Related Work

Question Completion [Kumar et. al (2017)]

- Reformulation as complete, self-contained question

Large-scale benchmark CSQA [Saha et. al (2018)]

- Question created semi-automatically using templates
- Artificial conversation flow

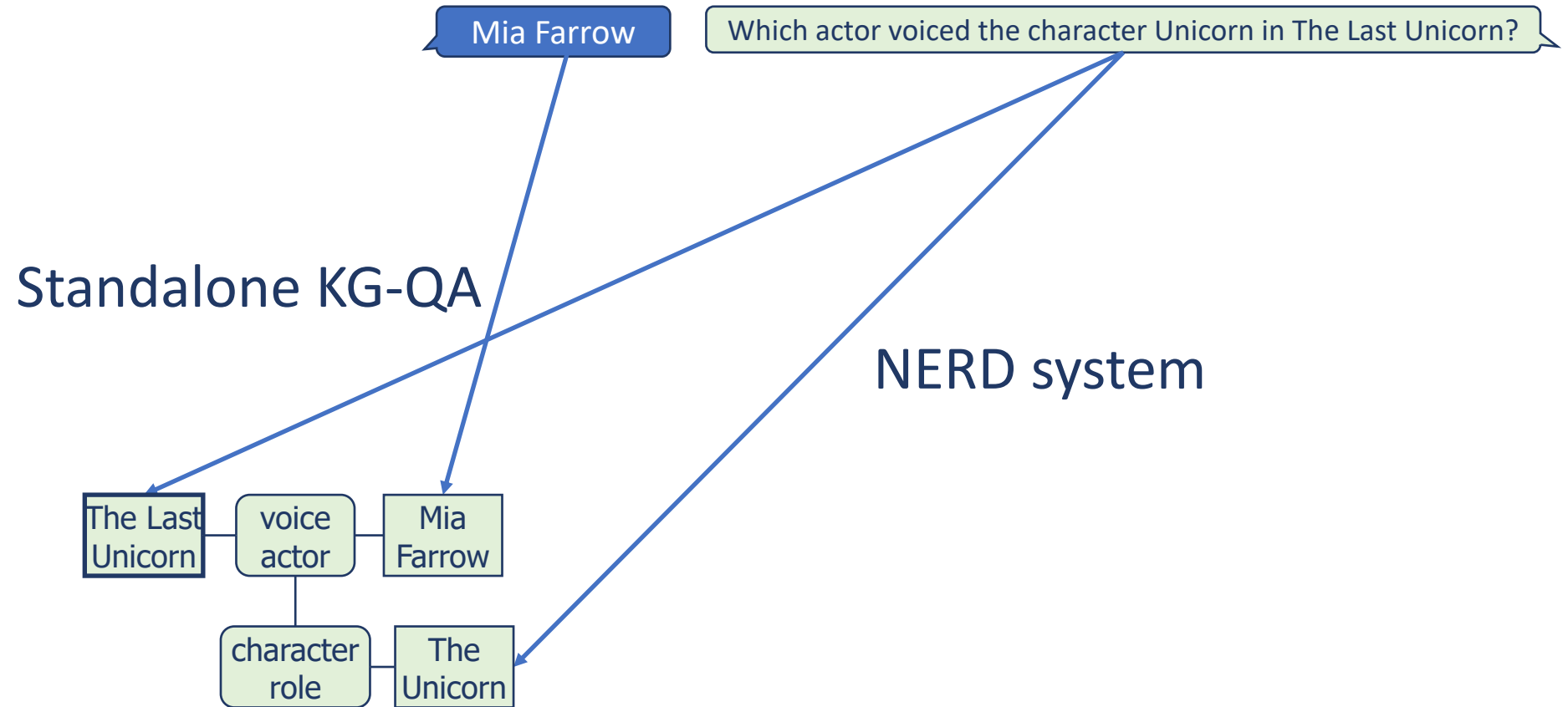
CSQA—method and Dialog-to-Action [Guo et. al (2018)]

- Seq-2-Seq learning methods
- Training data sparse
- Ad-hoc formulations key problem

Desiderata and Contributions

- Large topic jumps in conversations are rare
 - Conversations establish localized context in KG
- Harness underlying KG-connectivity
 - Expand context with relevant entities and predicates in neighborhood
- *CONVEX: CONVersational KG-QA using judicious context EXpansion*
 - Completely unsupervised!
- CONVEX works on top of any KG-QA system to handle conversations

Initial Context

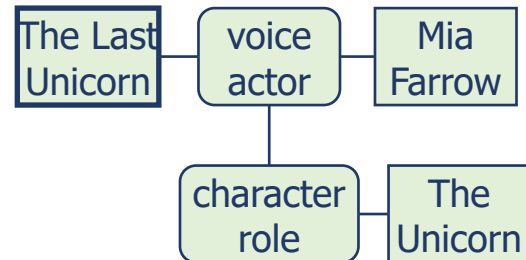


Initial Context

Mia Farrow

Which actor voiced the character Unicorn in The Last Unicorn?

And Alan Arkin was behind . . . ?



How to expand the context?



Neighborhood of
Mia Farrow

Neighborhood of
The Last Unicorn

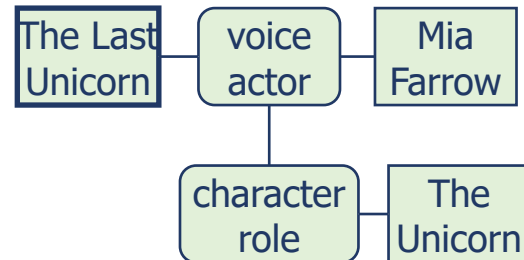
Neighborhood of
The Unicorn

Judicious Context Expansion

Mia Farrow

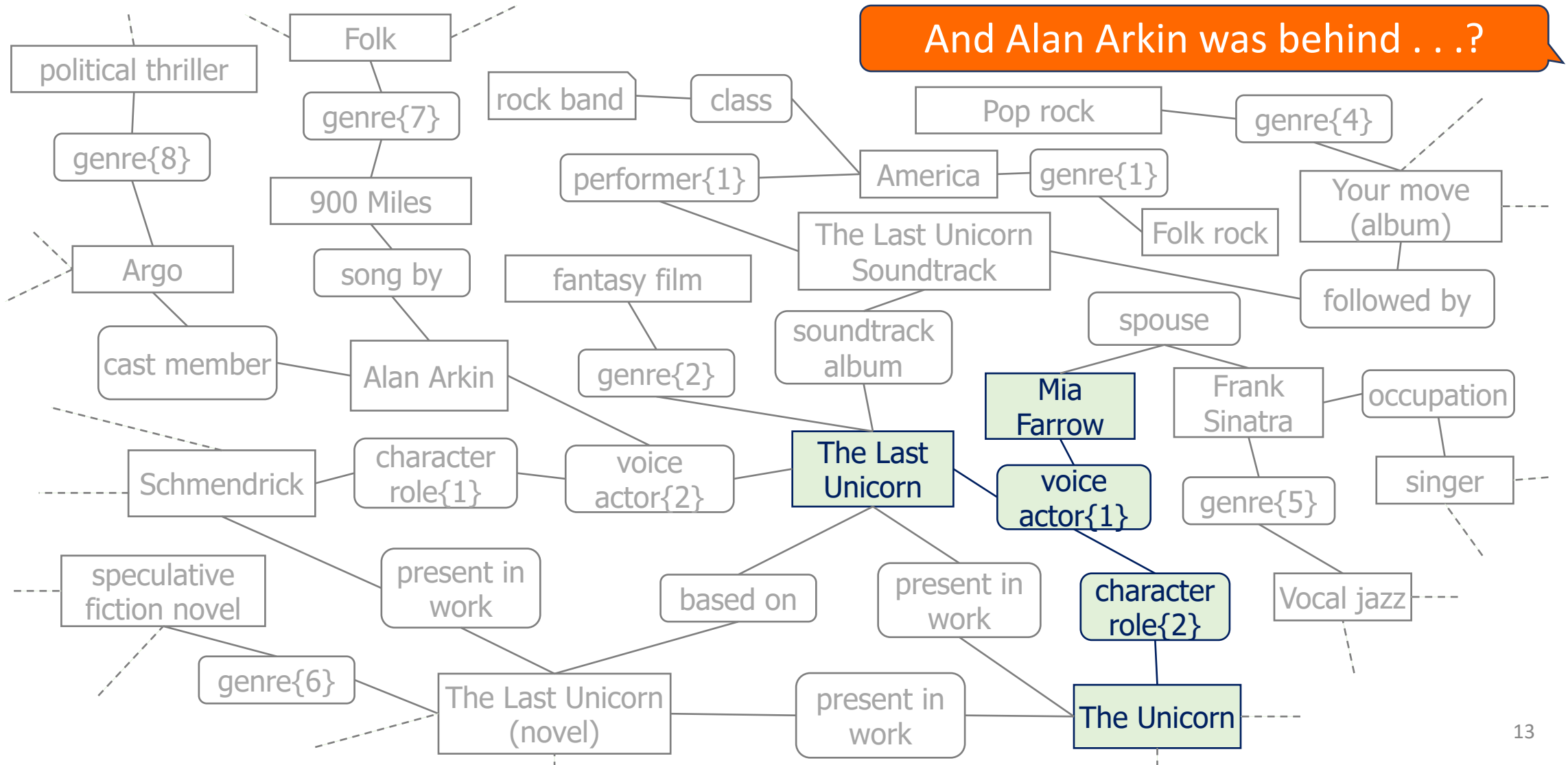
Which actor voiced the character Unicorn in The Last Unicorn?

And Alan Arkin was behind . . .?

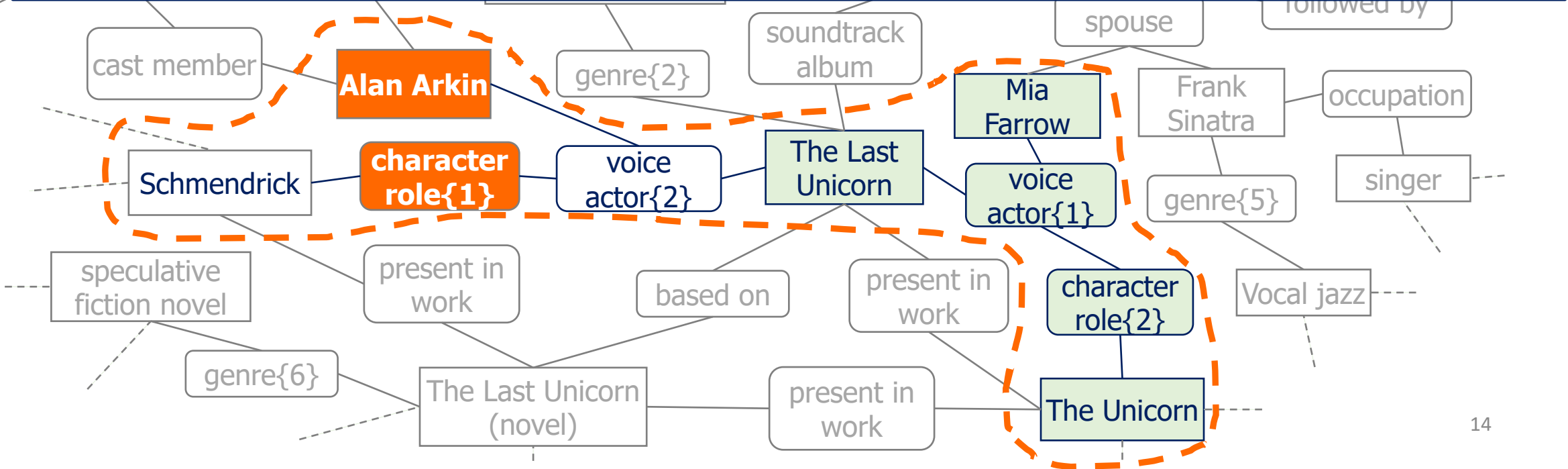


Do not expand with the
complete neighborhood!

Problem: Exploring Context Neighborhood



Determine Frontier nodes to describe an expansion border

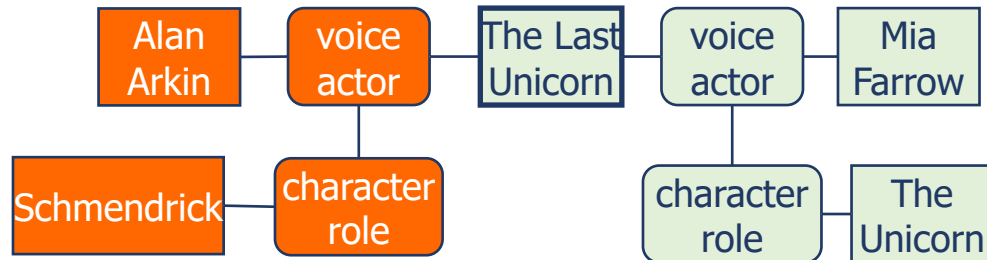


Context Graph

Mia Farrow

Which actor voiced the character Unicorn in The Last Unicorn?

And Alan Arkin was behind . . . ?



Expand graph accordingly!

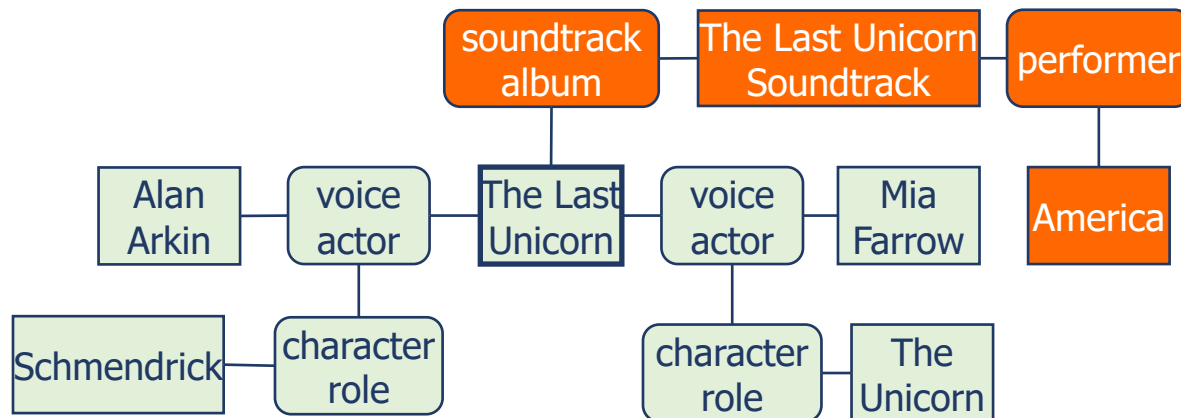
Context Graph

Mia Farrow

Which actor voiced the character Unicorn in The Last Unicorn?

And Alan Arkin was behind . . . ?

So who performed songs?



Graph expanded with
relevant facts only

Context Graph

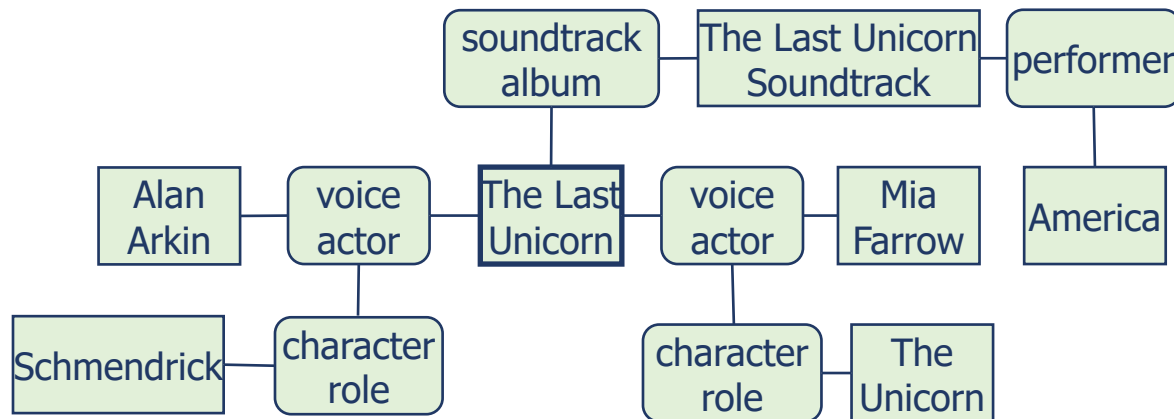
Mia Farrow

Which actor voiced the character Unicorn in The Last Unicorn?

And Alan Arkin was behind . . . ?

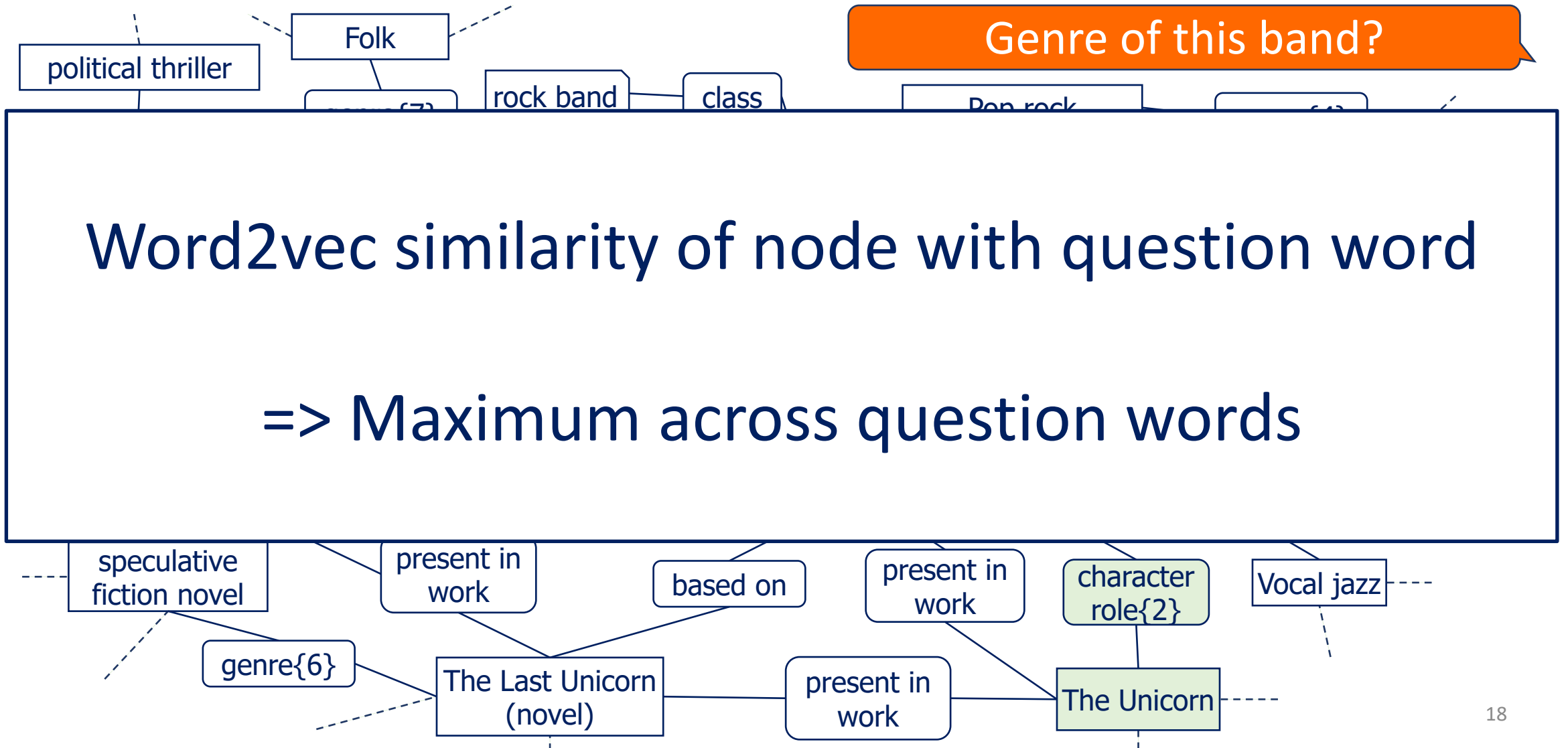
So who performed songs?

Genre of this band?



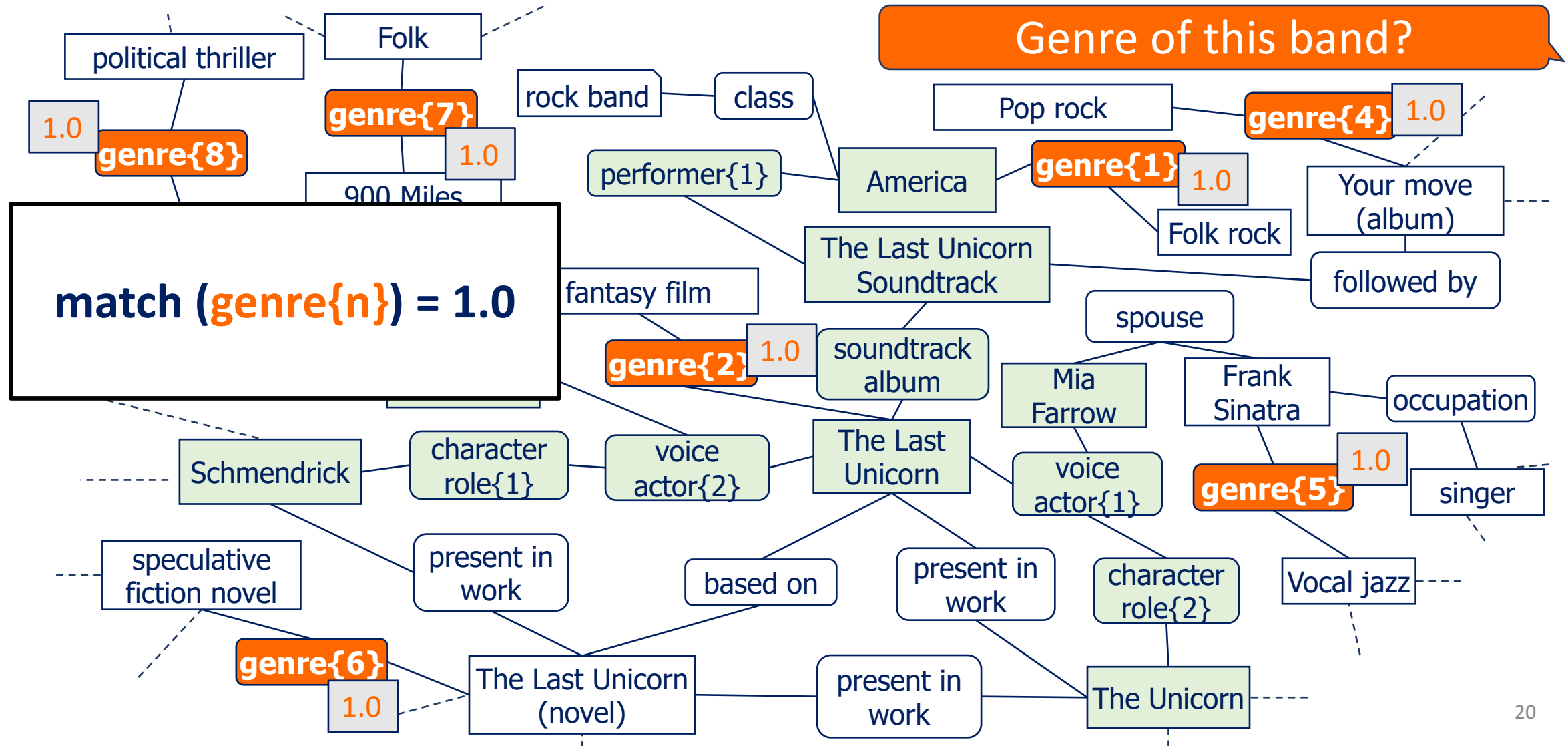
How to determine
Frontier nodes?

Relevance to the Question

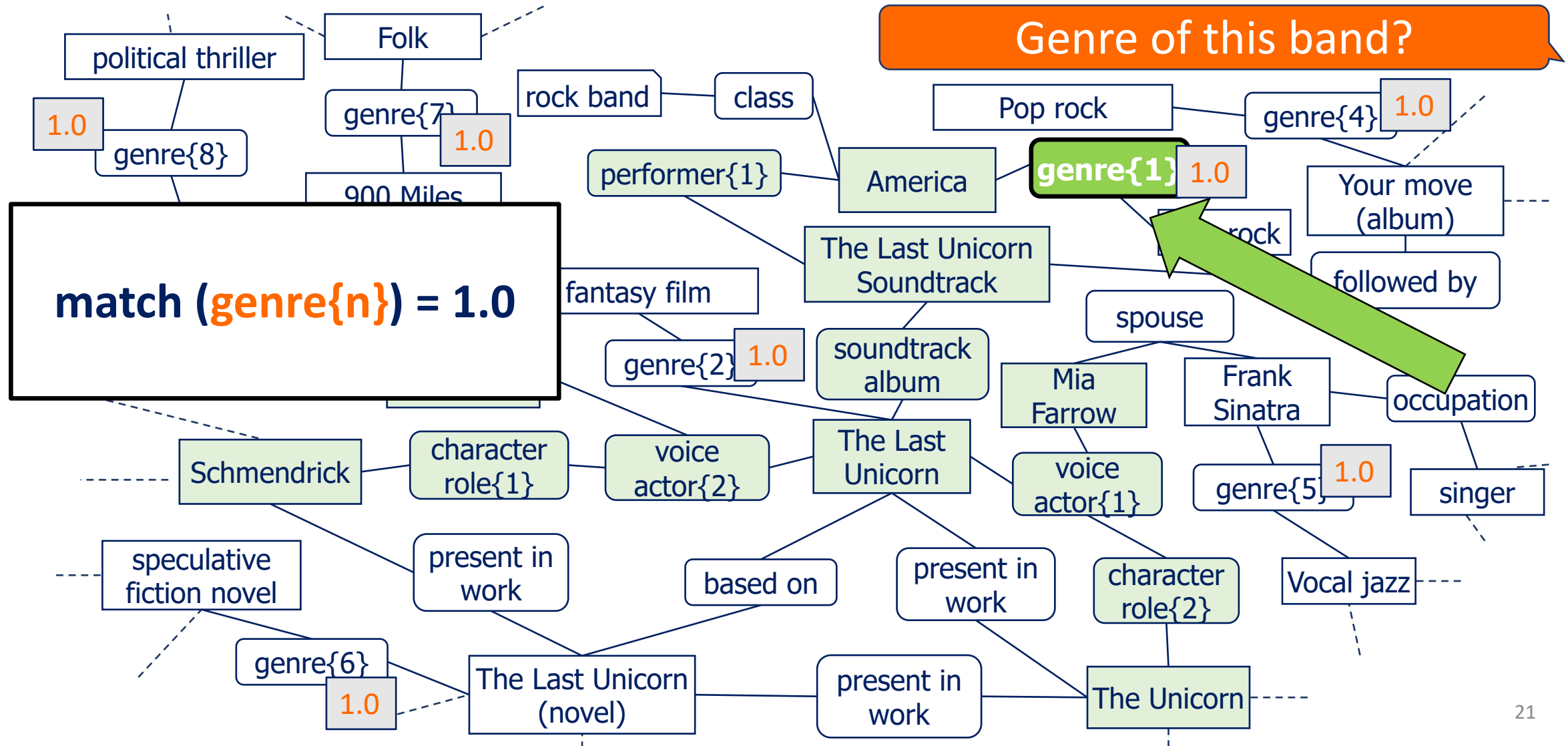




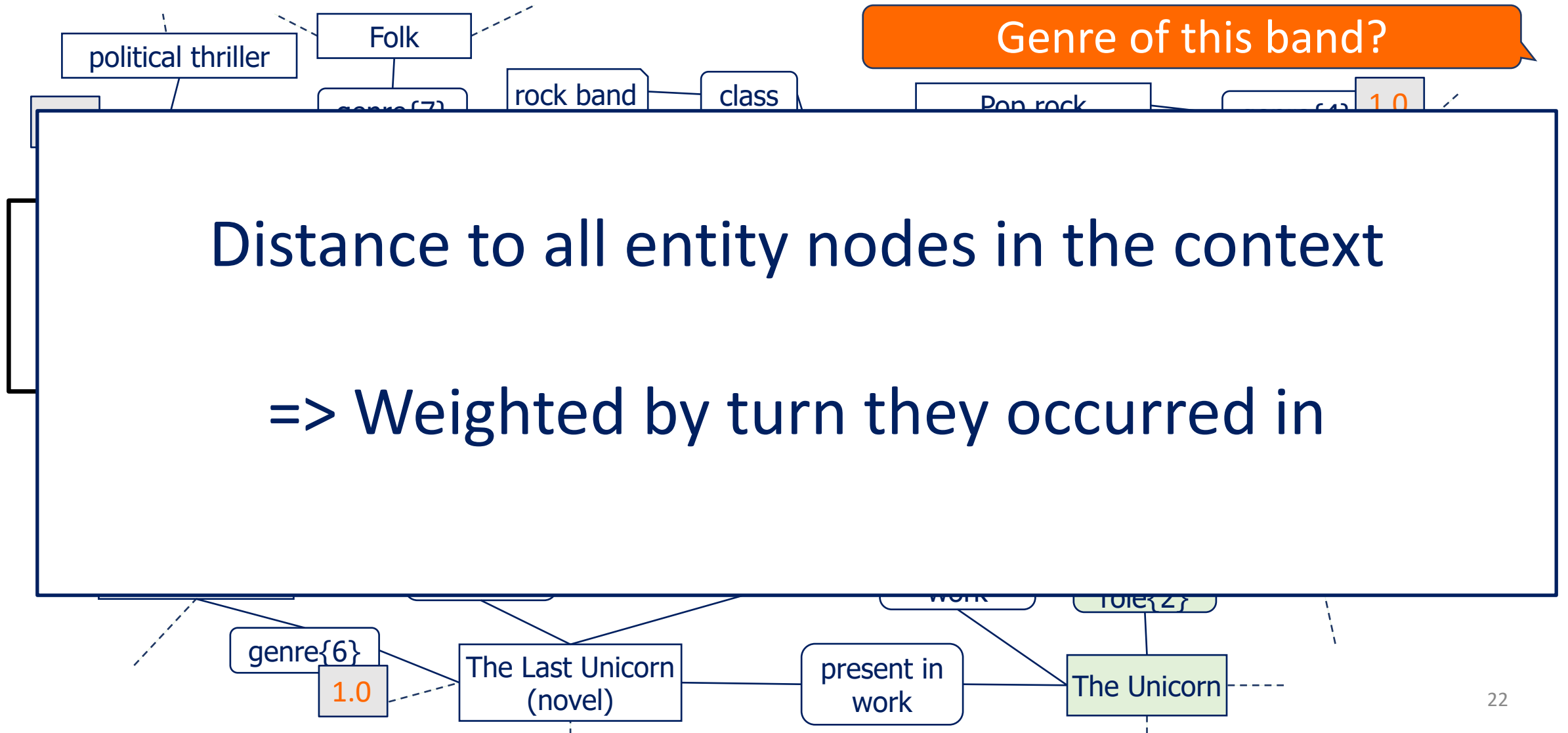
Relevance to the Context



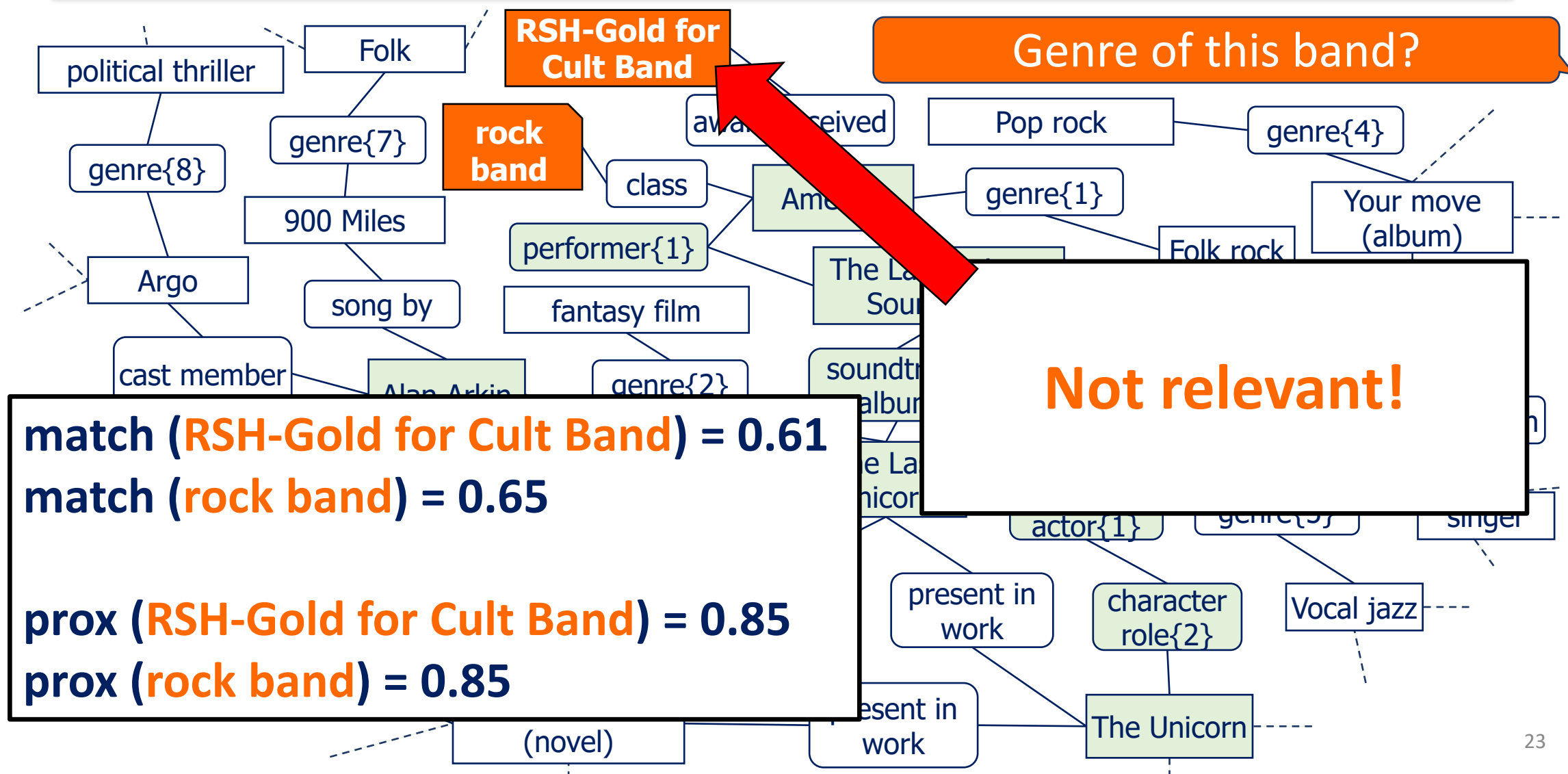
Relevance to the Context



Relevance to the Context



KG priors



KG priors

RSH-Gold for

Genre fiction level 12

Prioritize the more frequent/prominent entities and predicates

=> Normalize the value with maximum frequency

prox (RSH-Gold for **Carl Band**) = 0.85

prox (**rock band**) = 0.85

(novel)

present in
work

work

The Unicorn

Frontier Score

Matching similarity

match (candidate c)

Context relevance

prox (candidate c)

KG priors

prior (candidate c)


$$\text{frontier_score}(\text{candidate } c) = h_1 \cdot \text{match}(c) + h_2 \cdot \text{prox}(c) + h_3 \cdot \text{prior}(c)$$

With hyperparameters h_1, h_2, h_3

Frontier Nodes

Matching similarity

<i>Candidate</i>	<i>Match</i>
genre{1}	1.00
genre{2}	1.00
...	...
folk rock band	0.89
RSH-Gold for Cult Band	0.87
fantasy film	0.36
...	...

Context relevance

<i>Candidate</i>	<i>Prox</i>
genre{1}	0.91
folk rock band	0.86
RSH-Gold for Cult Band	0.86
...	...
genre{2}	0.34
fantasy film	0.36
...	...

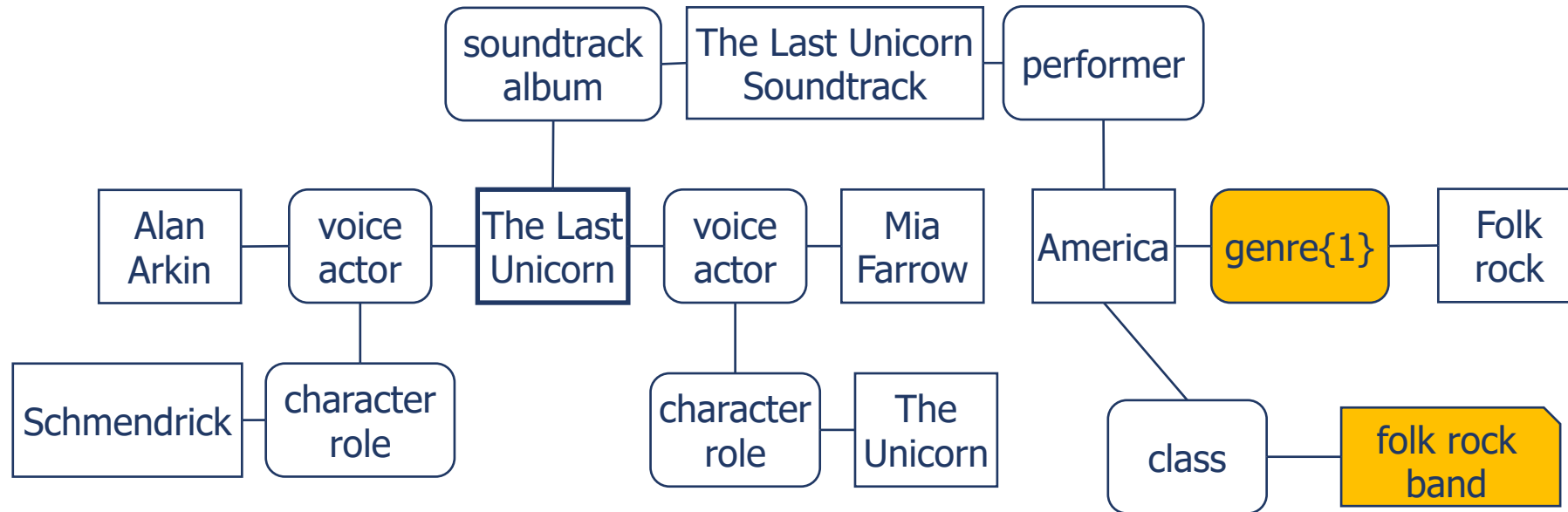
KG priors

<i>Candidate</i>	<i>KG priors</i>
...	...
genre{1}	0.56
genre{2}	0.56
...	...
folk rock band	0.34
...	...
RSH-Gold for Cult Band	0.01

Fagin's Threshold Algorithm to retrieve top- k
ranked nodes according to frontier score

Frontier Nodes

Genre of this band?

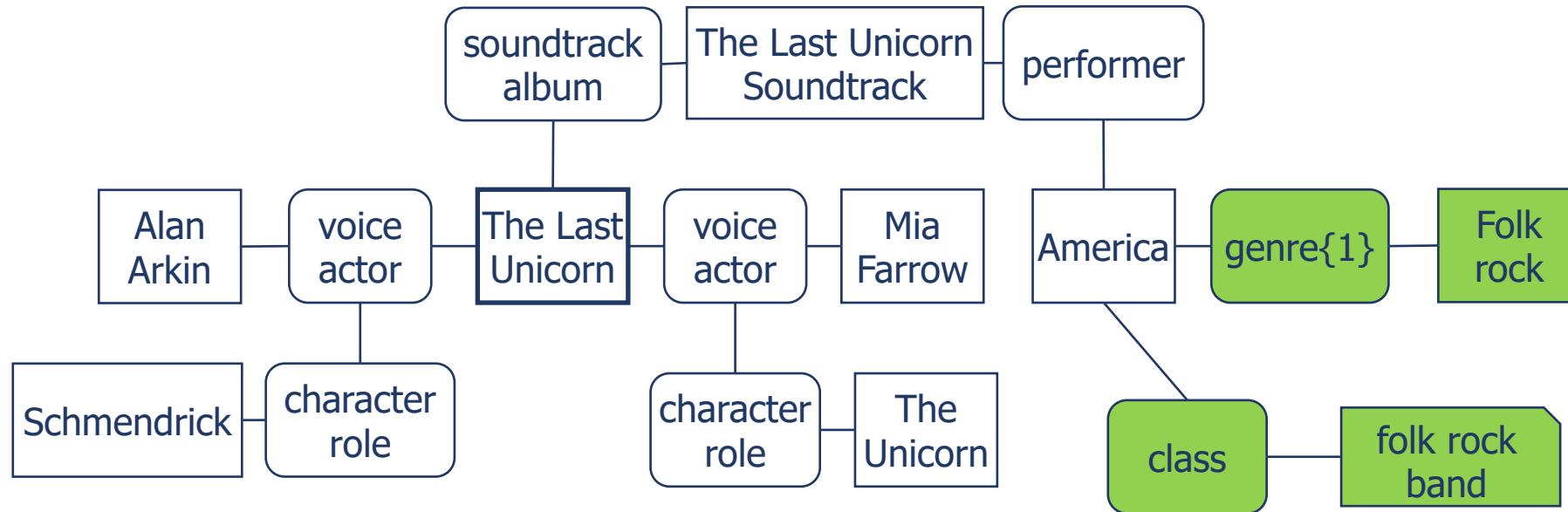


Frontier nodes

- Output of **Fagin's Threshold Algorithm**
 - ⇒ Top-ranked candidates according to Frontier score

Look before you Hop!

Genre of this band?



Answer to the Question?

Genre of this band?

- Distance to **Frontier nodes**
 - Weighted by the frontier score
 - distance_F

=> Explicit part

- Distance to all nodes in **context graph X**
 - Weighted by the turn they occurred in
 - distance_X

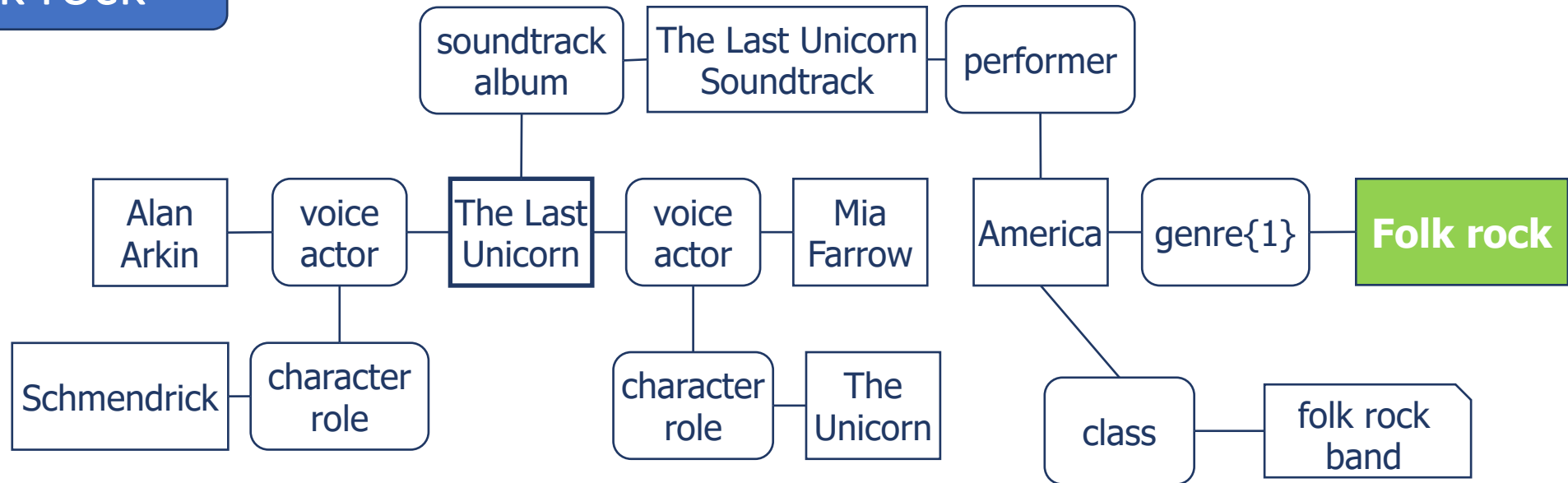
=> Implicit part

$$\text{answer_score}(\text{candidate } c) = h_4 \cdot \text{distance_F} + h_5 \cdot \text{distance_X}$$

Answer Detection

Genre of this band?

Folk rock



- Top-ranked node according to *answer_score*

Answering Steps

1. Define expansion border
 - ⇒ Determine most relevant nodes in neighborhood of context
 - ⇒ *Frontier nodes*
2. Expand context according to frontier nodes
3. Detect answer in expanded graph

Experimental Dataset: ConvQuestions

- 11,200 distinct conversations
- 5 utterances per conversation
 - Initial question + 4 follow-up questions
- Domains: Books, Movies, Music, TV Series, Soccer
- Gathered via crowdsourcing

Benchmark Properties

- Realistic benchmark
 - Questions created **by humans** from Amazon Mechanical Turk
 - In topic of their choice
- **Natural flow** of conversations
 - Conversations were **not interleaved**
 - Order of utterances was **not permuted**

ConvQuestions Examples

Books

What is the name of the second book?

★ Ordinal questions

ConvQuestions Examples

Books

What is the name of the second book?

Movies

The director's first wife?

ConvQuestions Examples

Books

What is the name of the second book?

Movies

The director's first wife?

Music

First album?

★ Incomplete cues

ConvQuestions Examples

Books

What is the name of the second book?

Movies

The director's first wife?

Music

First album?

TV Series

How many creators has the TV series with less episodes?

★ Comparatives

ConvQuestions Examples

Books

What is the name of the second book?

Movies

The director's first wife?

Music

First album?

TV Series

How many creators has the TV series with less episodes?

Soccer

Did they win the Super Cup the previous year?

★ Temporal questions

Experimental Setup

- Underlying KG
 - Wikidata
- Standalone KG-QA systems
 - **QAnswer**, Platypus, Naïve, Oracle
- Metrics
 - P@1, **MRR**, Hit@5
- Hyperparameters h_1, \dots, h_5 tuned on small dev set

Experimental Setup

- **Baselines**

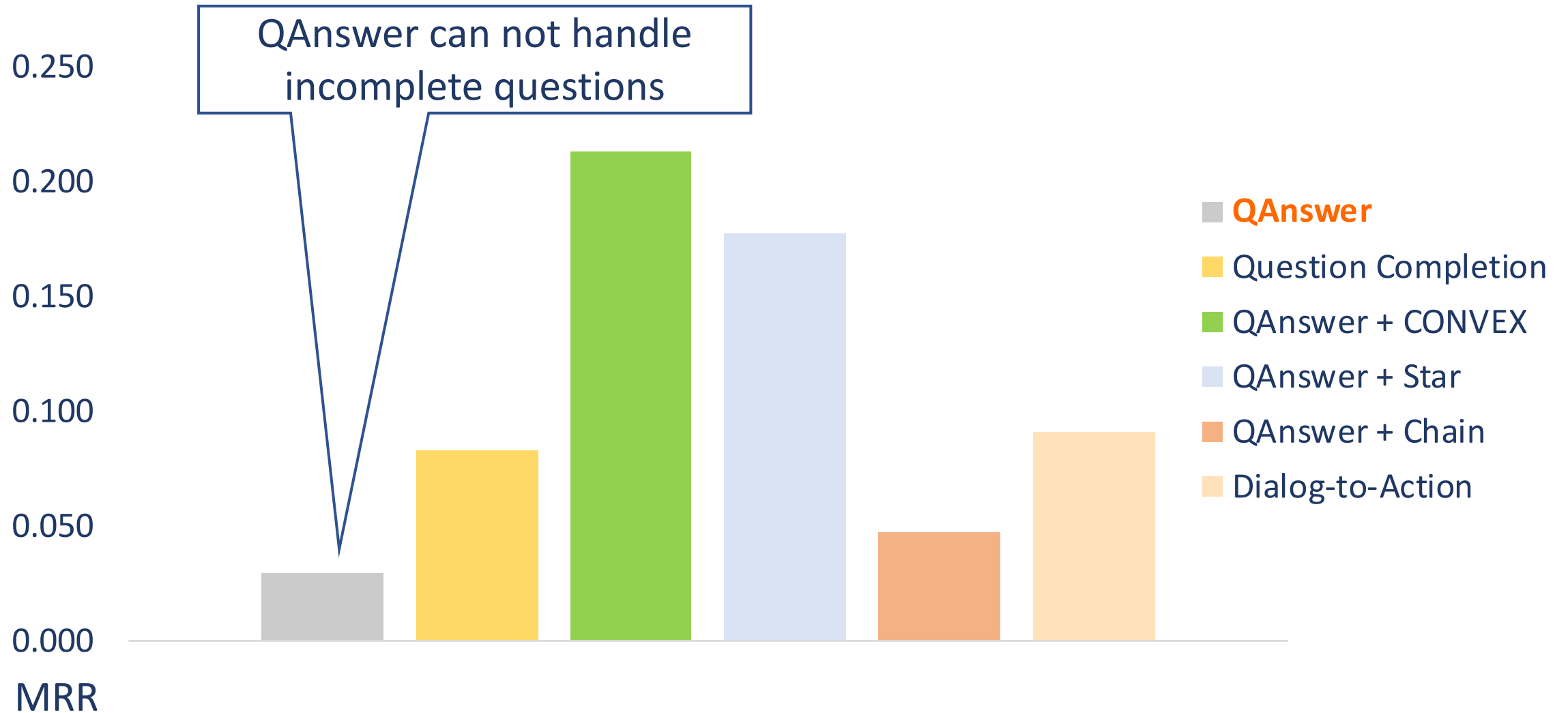
- Star model

- Chain model

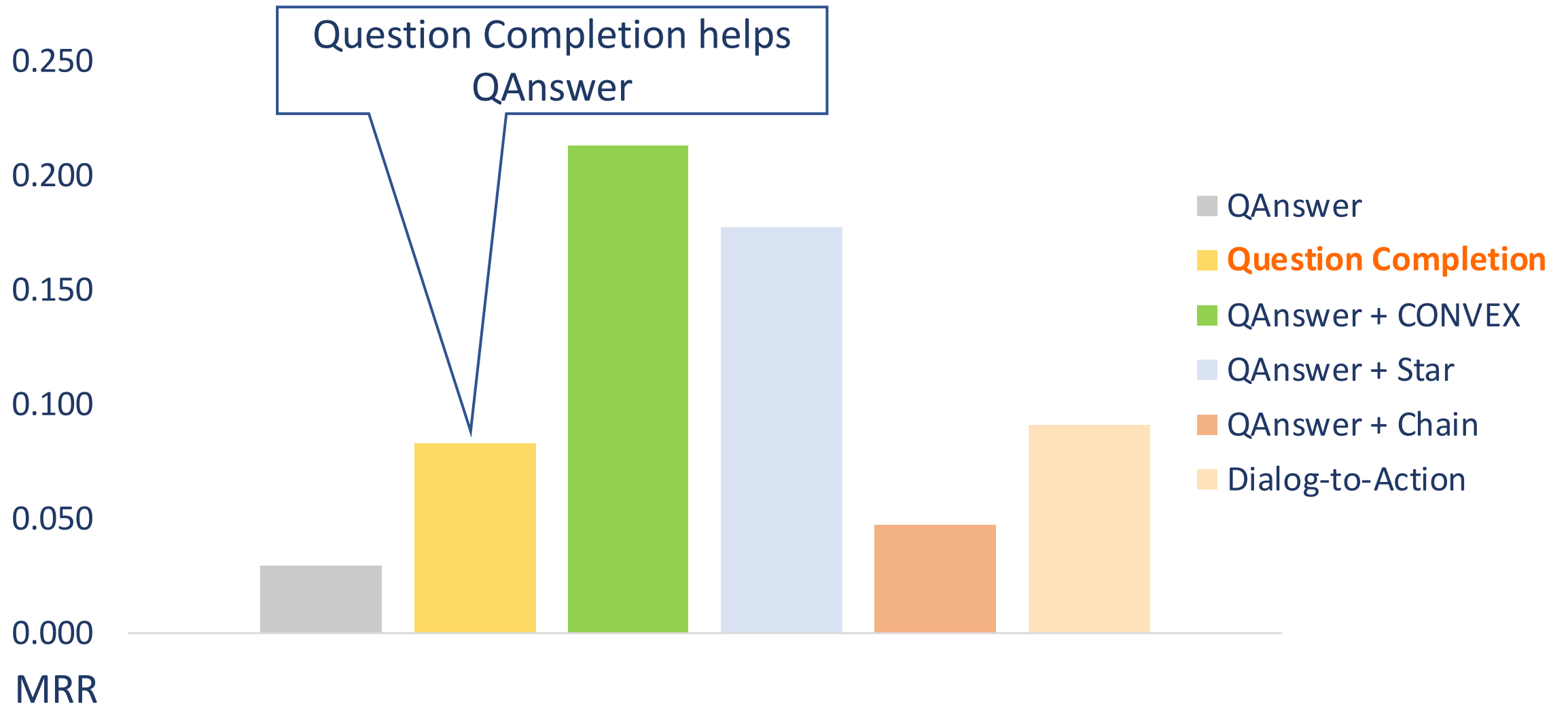
- Question Completion

- Dialog-to-Action (*Guo et al., NeurIPS 2018*)

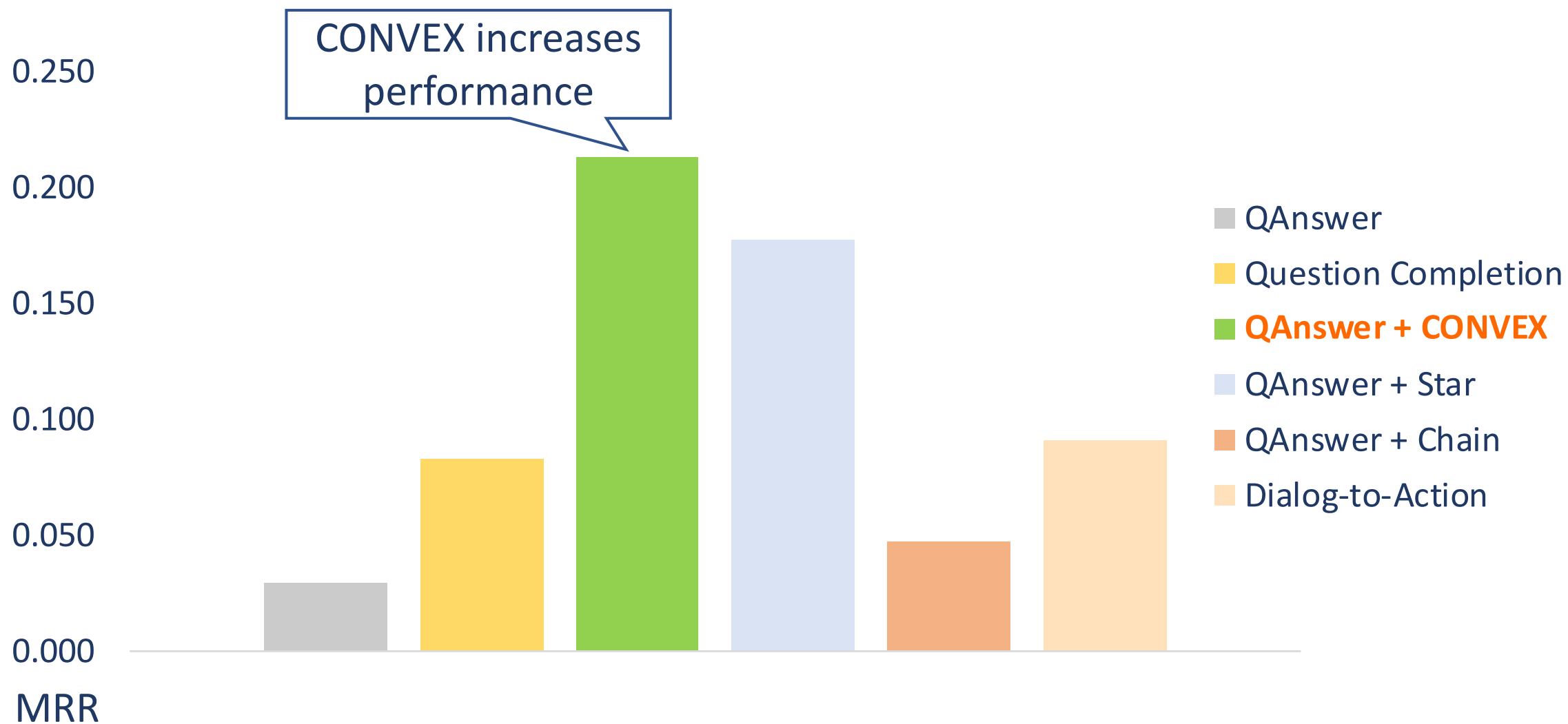
Main Results



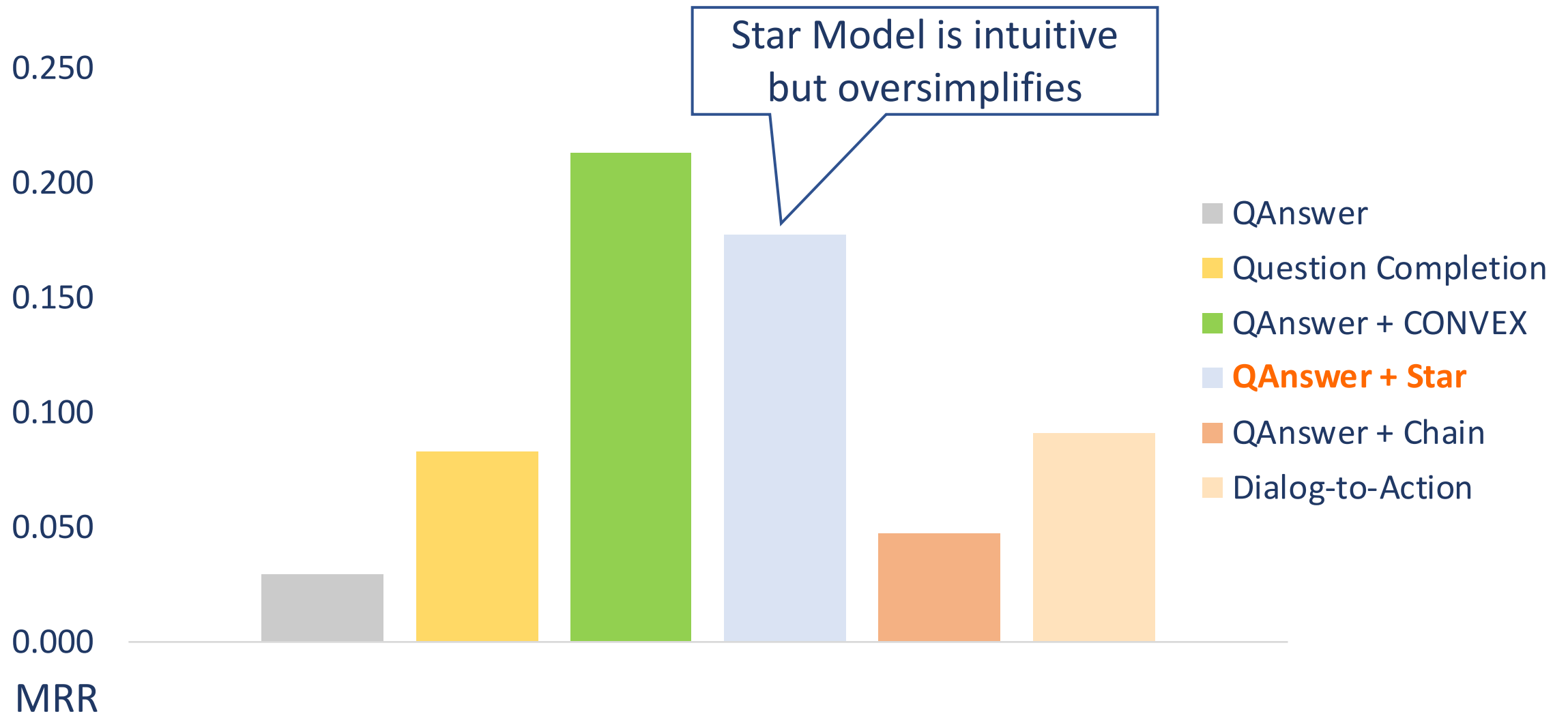
Main Results



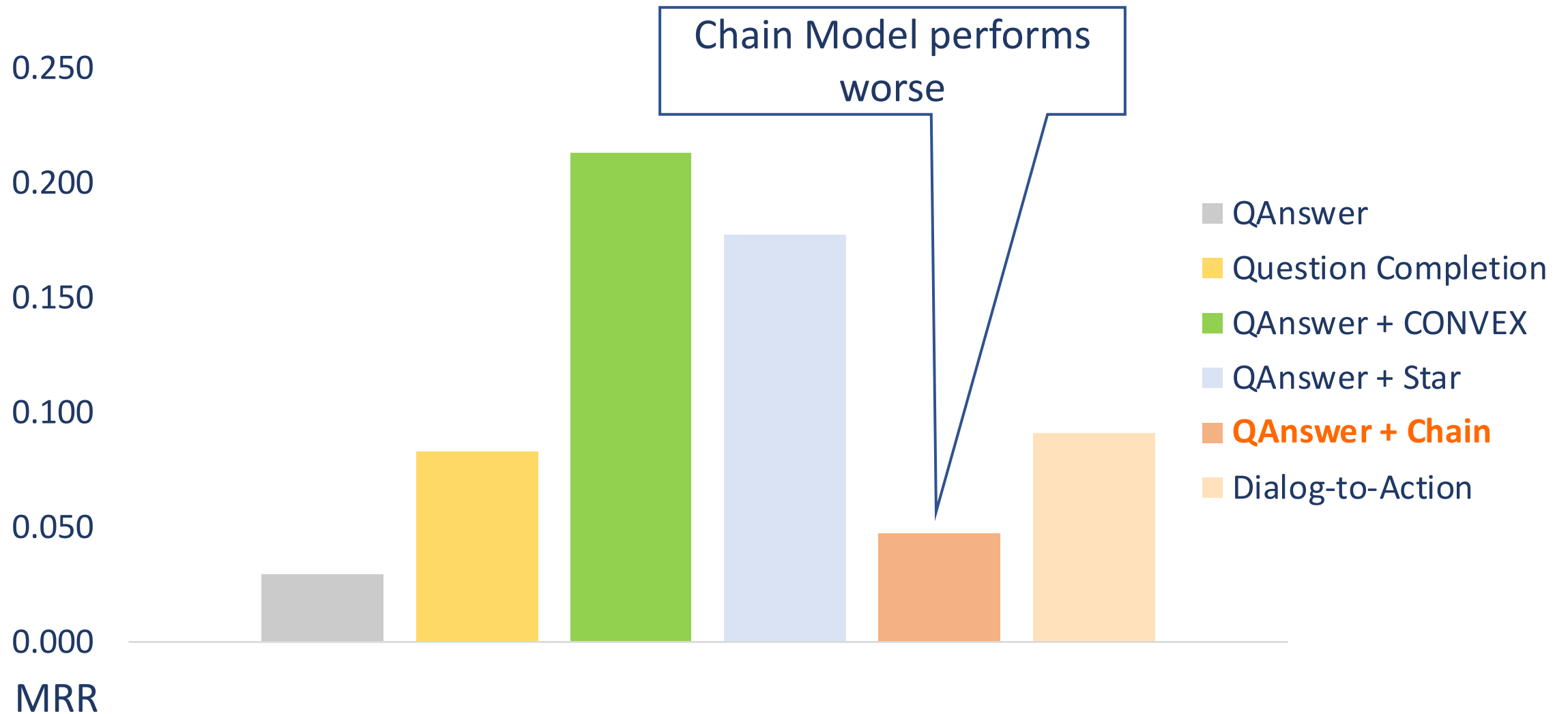
Main Results



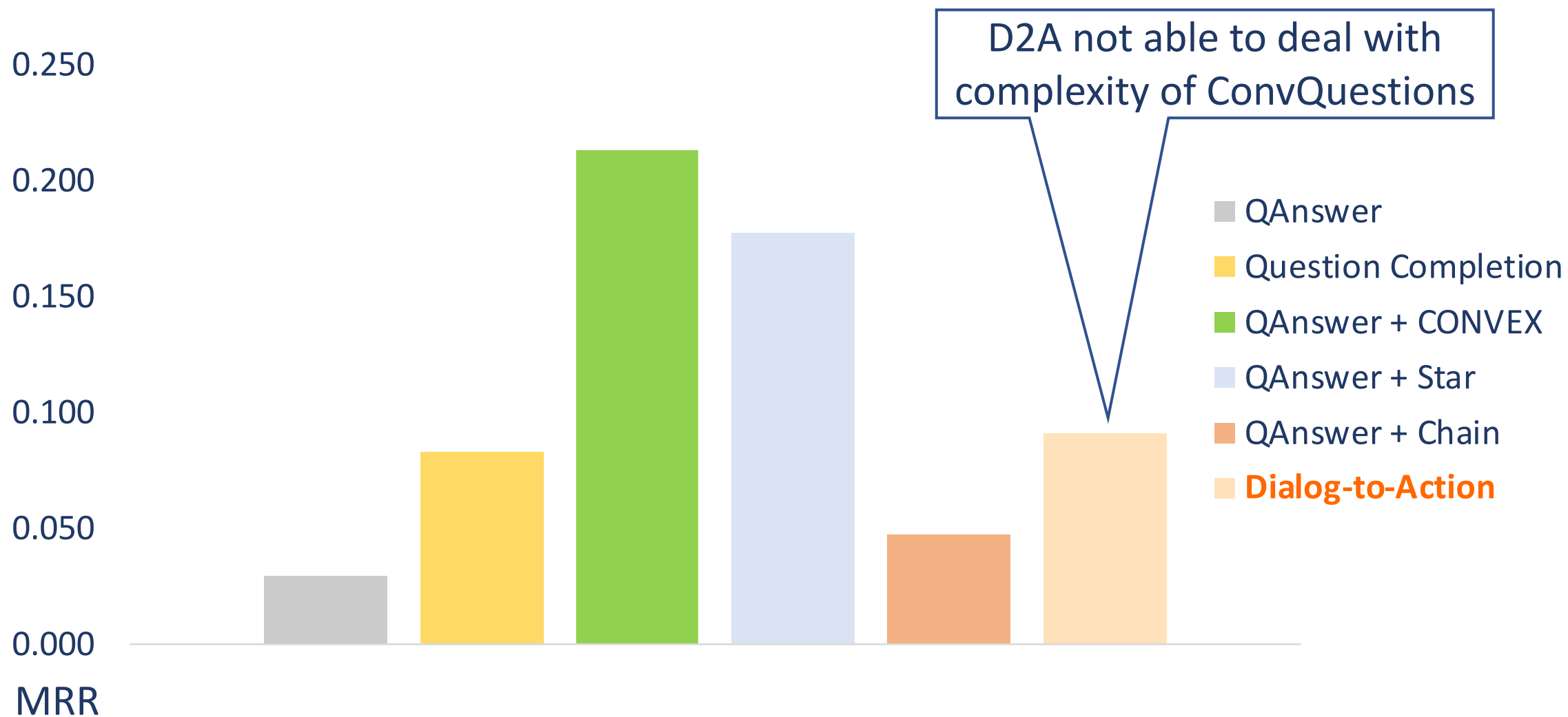
Main Results



Main Results



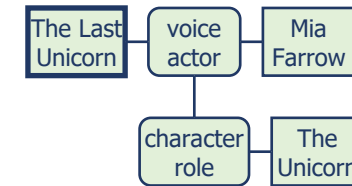
Main Results



Conclusion

➤ **CONVEX**

- First **unsupervised** system
- Enables any standalone KG-QA with **conversational support**
- Based on **judicious context expansion**



➤ **ConvQuestions**

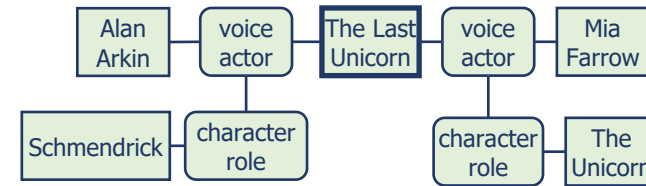
- First **realistic benchmark** on Conversational KG-QA
- 11,200 conversations from 5 domains

Data and Code:
qa.mpi-inf.mpg.de/convex

Conclusion

➤ **CONVEX**

- First **unsupervised** system
- Enables any standalone KG-QA with **conversational support**
- Based on **judicious context expansion**



➤ **ConvQuestions**

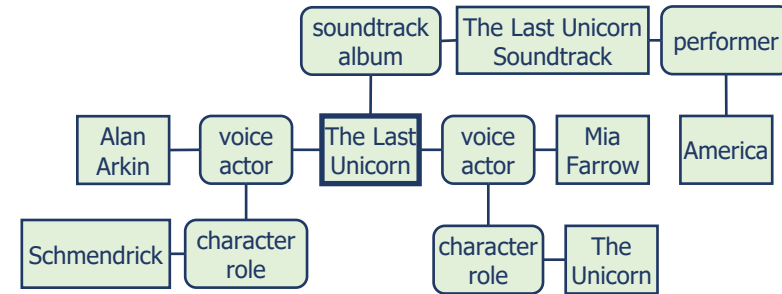
- First **realistic benchmark** on Conversational KG-QA
- 11,200 conversations from 5 domains

Data and Code:
qa.mpi-inf.mpg.de/convex

Conclusion

➤ **CONVEX**

- First **unsupervised** system
- Enables any standalone KG-QA with **conversational support**
- Based on **judicious context expansion**



➤ **ConvQuestions**

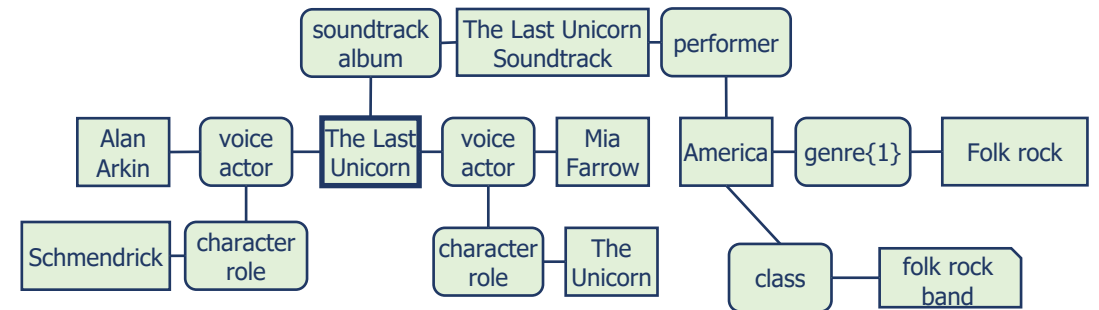
- First **realistic benchmark** on Conversational KG-QA
- 11,200 conversations from 5 domains

Data and Code:
qa.mpi-inf.mpg.de/convex

Conclusion

➤ **CONVEX**

- First **unsupervised** system
- Enables any standalone KG-QA with **conversational support**
- Based on **judicious context expansion**



➤ **ConvQuestions**

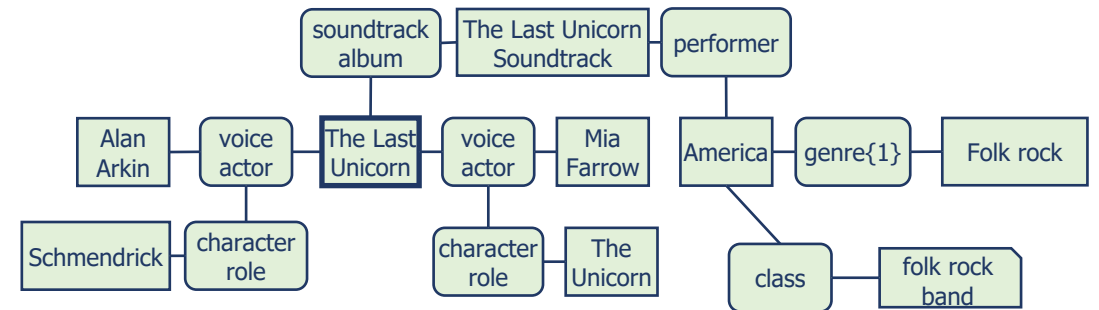
- First **realistic benchmark** on Conversational KG-QA
- 11,200 conversations from 5 domains

Data and Code:
qa.mpi-inf.mpg.de/convex

Conclusion

➤ **CONVEX**

- First **unsupervised** system
- Enables any standalone KG-QA with **conversational support**
- Based on **judicious context expansion**



➤ **ConvQuestions**

- First **realistic benchmark** on Conversational KG-QA
- 11,200 conversations from 5 domains

Data and Code:
qa.mpi-inf.mpg.de/convex

Thanks for SIGIR student travel grant!

**Thank
you!**