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

CIKM 2019, Beijing
Which actor voiced the character Unicorn in The Last Unicorn?

- Mia Farrow voiced the character Unicorn.

One-shot information need

Question usually needs to be well-formed

[Abujabal et al. (2018), Diefenbach et al. (2019), Huang et al. (2019)]
Which actor voiced the character Unicorn in The Last Unicorn?

- Mia Farrow
- Schmendrick
- America
- Folk rock
- Jules Bass

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?

And Alan Arkin was behind . . .?

Incomplete

So who performed songs?

Genre of this band?

By the way, who directed the movie?
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?
And Alan Arkin was behind . . .?

So who performed songs?

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
Which actor voiced the character Unicorn in The Last Unicorn? Mia Farrow

Standalone KG-QA

NERD system

The Last Unicorn

voice actor

Mia Farrow

character role

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

And Alan Arkin was behind . . .?

How to expand the context?
And Alan Arkin was behind...?

Which actor voiced the character Unicorn in The Last Unicorn?

Mia Farrow
Which actor voiced the character Unicorn in The Last Unicorn?

And Alan Arkin was behind . . .?

The Last Unicorn
character role
Mia Farrow
voice actor
The Unicorn

Do not expand with the complete neighborhood!

Mia Farrow
Problem: Exploring Context Neighborhood

And Alan Arkin was behind . . .?
Determine Frontier nodes to describe an expansion border
And Alan Arkin was behind . . .?

Which actor voiced the character Unicorn in The Last Unicorn?

Expand graph accordingly!
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

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?

Mia Farrow
Word2vec similarity of node with question word

=> Maximum across question words
Genre of this band?
The Last Unicorn

character
role{1}

voice
actor{2}

Mia Farrow

The Unicorn

performer{1}

soundtrack
album

The Last Unicorn Soundtrack

America

game

rock band

class

genre{7}

Folk rock

Your move (album)

followed by

Frank Sinatra

singer

Vocal jazz

Genre of this band?

match (genre{n}) = 1.0

political thriller

Folk

900 Miles

genre{8}

rock band

class

fantasy film

genre{2}

The Last Unicorn

soundtrack
album

The Last Unicorn

character
role{2}

based on

present in work

present in work

The Last Unicorn (novel)

The Unicorn

genre{6}

speculative
fiction novel

900 Miles

1.0

1.0

1.0

1.0

1.0

1.0

1.0

1.0

1.0

1.0

1.0
Relevance to the Context

**Genre of this band?**

match (genre{n}) = 1.0
Distance to all entity nodes in the context

=> Weighted by turn they occurred in
match (RSH-Gold for Cult Band) = 0.61
match (rock band) = 0.65
prox (RSH-Gold for Cult Band) = 0.85
prox (rock band) = 0.85
Prioritize the more frequent/prominent entities and predicates

=> Normalize the value with maximum frequency

prox (RSH-Gold for Cult Band) = 0.85
prox (rock band) = 0.85
## Frontier Score

<table>
<thead>
<tr>
<th>Matching similarity</th>
<th>Context relevance</th>
<th>KG priors</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{match (candidate } c ) )</td>
<td>( \text{prox (candidate } c ) )</td>
<td>( \text{prior (candidate } c ) )</td>
</tr>
</tbody>
</table>

\[
\text{frontier\_score(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

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Match</th>
</tr>
</thead>
<tbody>
<tr>
<td>genre{1}</td>
<td>1.00</td>
</tr>
<tr>
<td>genre{2}</td>
<td>1.00</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>folk rock band</td>
<td>0.89</td>
</tr>
<tr>
<td>RSH-Gold for Cult Band</td>
<td>0.87</td>
</tr>
<tr>
<td>fantasy film</td>
<td>0.36</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

### Context relevance

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Prox</th>
</tr>
</thead>
<tbody>
<tr>
<td>genre{1}</td>
<td>0.91</td>
</tr>
<tr>
<td>folk rock band</td>
<td>0.86</td>
</tr>
<tr>
<td>RSH-Gold for Cult Band</td>
<td>0.86</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>genre{2}</td>
<td>0.34</td>
</tr>
<tr>
<td>fantasy film</td>
<td>0.36</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

### KG priors

<table>
<thead>
<tr>
<th>Candidate</th>
<th>KG priors</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>genre{1}</td>
<td>0.56</td>
</tr>
<tr>
<td>genre{2}</td>
<td>0.56</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>folk rock band</td>
<td>0.34</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>RSH-Gold for Cult Band</td>
<td>0.01</td>
</tr>
</tbody>
</table>

**Fagin’s Threshold Algorithm** to retrieve top-\(k\) ranked nodes according to frontier score.
Output of Fagin’s Threshold Algorithm
⇒ Top-ranked candidates according to Frontier score
Look before you Hop!

Genre of this band?

- Alan Arkin: voice actor in The Last Unicorn
- Schmendrick: character role
- Mia Farrow: voice actor in The Last Unicorn
- America: performer
- The Last Unicorn Soundtrack: soundtrack album
- The Unicorn: performer
- America: genre folk rock
- Folk rock band: class
- Folk rock: genre

The genre of the band for The Last Unicorn Soundtrack is folk rock.
Answer to the Question?

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

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

\[
answer\_score(\text{candidate } c) = h_4 \cdot \text{distance}_F + h_5 \cdot \text{distance}_X
\]
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

Movies

Music

TV Series

What is the name of the second book?

The director’s first wife?

First album?

How many creators has the TV series with less episodes?

Comparatives
## ConvQuestions Examples

<table>
<thead>
<tr>
<th>Category</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Books</td>
<td>What is the name of the second book?</td>
</tr>
<tr>
<td>Movies</td>
<td>The director’s first wife?</td>
</tr>
<tr>
<td>Music</td>
<td>First album?</td>
</tr>
<tr>
<td>TV Series</td>
<td>How many creators has the TV series with less episodes?</td>
</tr>
<tr>
<td>Soccer</td>
<td>Did they win the Super Cup the previous year?</td>
</tr>
</tbody>
</table>

🌟 Temporal questions
Experimental Setup

- Underlying KG
  - Wikidata
- Standalone KG-QA systems
  - QAnswer, Platypus, Naïve, Oracle
- Metrics
  - $P@1, MRR, Hit@5$
- Hyperparameters $h_1, ..., h_5$ tuned on small dev set
Experimental Setup

- **Baselines**
  - Star model
  - Chain model

- Question Completion

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

QAnswer can not handle incomplete questions

- QAnswer
- Question Completion
- QAnswer + CONVEX
- QAnswer + Star
- QAnswer + Chain
- Dialog-to-Action
Main Results

Question Completion helps QAnswer

- QAnswer
- Question Completion
- QAnswer + CONVEX
- QAnswer + Star
- QAnswer + Chain
- Dialog-to-Action

MRR
Main Results

CONVEX increases performance

![Graph showing MRR performance for different models and methods]

- QAnswer
- Question Completion
- QAnswer + CONVEX
- QAnswer + Star
- QAnswer + Chain
- Dialog-to-Action

43
Main Results

Star Model is intuitive but oversimplifies
Main Results

Chain Model performs worse

MRR

QAnswer
Question Completion
QAnswer + CONVEX
QAnswer + Star
QAnswer + Chain
Dialog-to-Action
Main Results

D2A not able to deal with complexity of ConvQuestions
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

Thank you!