Reinforcement Learning from Reformulations in Conversational Question Answering over Knowledge Graphs

Magdalena Kaiser, Rishiraj Saha Roy and Gerhard Weikum
Max Planck Institute for Informatics, Saarbrücken, Germany
Q1: When was *Avengers: Endgame* released in Germany?

A1: 24 April 2019

Q2: What was next from *Marvel*?

A2: *Spider-Man: Far from Home*

Q3: Release date?

A3: 4 July 2019

Q4: Who played *Spider-Man*?

A4: Tom Holland

Q5: And what about *his girlfriend*?

A5: ...
Ideal Conversation

Q1: When was *Avengers: Endgame* released in Germany?
A1: 24 April 2019

Q2: What was next from Marvel?
A2: Spider-Man: Far from Home

Q3: Release date?
A3: 4 July 2019

Q4: Who played Spider-Man?
A4: Tom Holland

Q5: And what about his girlfriend?
A5: ...

Challenges:
★ Short, incomplete questions
★ Implicit context
Q1: When was *Avengers: Endgame* released in Germany?

A1: 24 April 2019

Q2: What was next from Marvel?

A2: Stan Lee

Q21: I mean, what came next in the series?

A21: Marvel Cinematic Universe

Q22: The following movie in the Marvel series?

A22: Spider-Man: Far from Home

Q3: Release date?

A3: ...

Challenges:
★ Short, incomplete questions
★ Implicit context
Q1: When was *Avengers: Endgame* released in Germany?

A1: 24 April 2019

Q2: What was next from Marvel?

A2: Stan Lee

Q21: I mean, what came next in the series?

A21: Marvel Cinematic Universe

Q22: The following movie in the Marvel series?

A22: *Spider-Man: Far from Home*

Q3: Release date?

A3: ...

Challenges:
- ★ Short, incomplete questions
- ★ Implicit context

Learn from conversational stream:
Reformulation = Wrong answer
New intent = Correct answer
Contributions

⭐ **CONQUER:** Reinforcement learning model for QA
  ○ Learns from *conversational stream* in the *absence of gold answers*
  ○ With *rewards* based on *implicit feedback* in form of question reformulations

⭐ **Reformulation predictor** based on BERT that can classify a follow-up utterance as a reformulation or new intent

⭐ **ConvRef:** ConvQA benchmark with reformulations
Q1: When was **Avengers: Endgame** released in Germany?
A1: 24 April 2019
Q2: What was next from **Marvel**?

**Basic Idea**
Q1: When was **Avengers: Endgame** released in Germany?
A1: 24 April 2019
Q2: What was next from **Marvel**?

![Diagram showing connections between Avengers: Endgame, Spider-Man: Far from Home, Captain Marvel, and Marvel Cinematic Universe]

- **Avengers: Endgame**
  - After a work by **Stan Lee**
  - Publication date: 26 April 2019
  - Place of publication: Germany
- **Spider-Man: Far from Home** followed by
- **Captain Marvel**
- **Marvel Cinematic Universe**
Q1: When was **Avengers: Endgame** released in Germany?
A1: 24 April 2019
Q2: What was next from **Marvel**?
N-ary facts connected via statement-ids:
<AvengersEndgame, partOfSeries, 123>
<123, partOfSeries, MarvelCinematicUniverse>
<123, followedBy, SpiderManFarFromHome>
<123, follows, CaptainMarvel>
<123, seriesOrdinal, 22>
KG Representation

SIMPLE GRAPH MODEL

Main fact

Avengers: Endgame → part of series → Marvel Cinematic Universe

Qualifier

followed by

Spider-Man: Far from Home
**KG Representation**

**SIMPLE GRAPH MODEL**

Main fact: **Avengers: Endgame** is part of the **Marvel Cinematic Universe**.

Qualifier: **Avengers: Endgame** is followed by **Spider-Man: Far from Home**.

**CONQUER GRAPH MODEL**

(a) **Avengers: Endgame** is part of the **Marvel Cinematic Universe**.

followed by **Spider-Man: Far from Home**.
KG Representation

**SIMPLE GRAPH MODEL**

<table>
<thead>
<tr>
<th>Main fact</th>
<th>Entity node</th>
<th>Edge or path label</th>
<th>Entity node</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Avengers: Endgame</strong></td>
<td>![part of series]</td>
<td><strong>followed by</strong></td>
<td><strong>Marvel Cinematic Universe</strong></td>
</tr>
<tr>
<td><strong>followed by</strong></td>
<td><strong>Spider-Man: Far from Home</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**CONQUER GRAPH MODEL**

1. **(a)**
   - **Avengers: Endgame**
   - **part of series**
   - **followed by** **Spider-Man: Far from Home**
   - **Marvel Cinematic Universe**

2. **(b)**
   - **Avengers: Endgame**
   - **followed by** **Marvel Cinematic Universe**
   - **Spider-Man: Far from Home**
Step 1: Context Entity Detection

★ Find entities relevant to current question and its conversational context

Q1: When was **Avengers: Endgame** released in Germany?
A1: 24 April 2019
Q2: What was next from **Marvel**?

![Diagram showing relationships between Marvel Cinematic Universe and Spider-Man: Far from Home, Captain Marvel, and Avengers: Endgame.](image-url)
Step 1: Context Entity Detection

★ Find entities relevant to current question and its conversational context
★ Context entities will be start points for the RL walk

Q1: When was **Avengers: Endgame** released in Germany?
A1: 24 April 2019
Q2: What was next from **Marvel**?
Step 1: Context Entity Detection

★ Find entities relevant to current question and its conversational context
★ Context entities will be start points for the RL walk
★ NED tools do not work well on short, incomplete questions

Q1: When was Avengers: Endgame released in Germany?
A1: 24 April 2019
Q2: What was next from Marvel?

Spider-Man: Far from Home
followed by
Marvel Cinematic Universe

Captain Marvel
follows

part of the series

Avengers: Endgame
Step 1: Context Entity Detection

- Find **entities** relevant to **current question** and its **conversational context**
- **Context entities** will be **start points** for the RL walk
- **NED tools** do **not work** well on **short, incomplete questions**
- Get **initial entities** from first complete question **via NED tool**

Q1: When was **Avengers: Endgame** released in Germany?
A1: 24 April 2019
Q2: What was next from **Marvel**?
Step 1: Context Entity Detection

★ Find **entities** relevant to **current question** and its **conversational context**

★ **Context entities** will be **start points** for the RL walk

★ NED tools do **not work** well on **short, incomplete questions**

★ Get **initial entities** from first complete question **via NED tool**

★ **Score one hop neighborhood** of current context nodes:
  
  o Lexical match
  o Neighbor overlap
  o NED score
  o KG prior

---

Spider-Man: Far from Home

followed by

part of the series

Avengers: Endgame

follows

Marvel Cinematic Universe

Q1: When was *Avengers: Endgame* released in Germany?
A1: 24 April 2019

Q2: What was next from *Marvel*?
Step 2: Path Prediction

Spider-Man: Far from Home followed by part of the series Marvel Cinematic Universe

follows

Captain Marvel

Avengers: Endgame
Step 2: Path Prediction

Spider-Man: Far from Home followed by part of the Avengers: Endgame Marvel Cinematic Universe.

Agent \( S_t \) \(\xrightarrow{A_t}\) Environment \( S_{t+1}, R_t \)

Spider-Man: Far from Home followed by Captain Marvel follows part of the series

Marvel Cinematic Universe
Step 2: Path Prediction

★ **States**: current question, context entity, conversation history (optional)
★ **Actions**: all outgoing paths from the context entity node
★ **Transitions**: entity reached when following selected action, follow-up question, updated conversation history
★ **Rewards**: 1 if next question is a new info need, -1 if reformulation
★ **Policy**: determines which action to select in a given state

[Diagram showing the relationship between Agent, Environment, and Spider-Man: Far from Home, Captain Marvel, and Avengers: Endgame]
When was Avengers: Endgame released in Germany?

What was the next from Marvel?

Avengers: Endgame

genre
director
part of ser. Marvel Cinem. Univ. I followed by
Policy Network

Context question(s) $q_i^{ctx}$
- When was Avengers: Endgame released in Germany?

Question $q_i$
- What was the next from Marvel?

Context entity $e_i^{ctx}$
- Avengers: Endgame

Sentence Embeddings

BERT

KG paths $p$ from context entity $e_i^{ctx}$
- genre
- director
- part of ser. Marvel Cinem. Univ.
  - I followed by
Policy Network

Context question(s) $q_i^{\text{ext}}$

When was Avengers: Endgame released in Germany?

What was the next from Marvel?

Question $q_i$

Context entity $e_i^{\text{ext}}$

Avengers: Endgame

Gene
Director
Part of ser. Marvel Cinem. Univ. I followed by

Sentence Embeddings

BERT

$W_1$

ReLU

$W_2$
Policy Network

Context question(s) $q^\text{ext}_t$
- When was Avengers: Endgame released in Germany?

Question $q_t$
- What was the next from Marvel?

Context entity $e^\text{ext}_t$
- Avengers: Endgame

KG paths $p$ from context entity $e^\text{ext}_t$
- genre
- director
- part of ser. Marvel Cinem. Univ.
- I followed by

Sentence Embeddings

Question Embeddings

Action Embeddings

$$w_1$$

ReLU

$$w_2$$

$$a_1$$

$$a_2$$

...$$a_n$$

$A_s$
Policy Network

Context question(s) $q^{\text{ext}}_t$
- When was Avengers: Endgame released in Germany?

Question $q_t$
- What was the next from Marvel?

Context entity $e^{\text{ext}}_t$
- Avengers: Endgame

KG paths $p$ from context entity $e^{\text{ext}}_t$
- genre
- director
- part of ser. Marvel Cinem. Univ. I followed by

Sentence Embeddings

BERT

Action Embeddings

Action

Multiply and softmax

$P(A_s)$
Policy Network

When was Avengers: Endgame released in Germany?

What was the next from Marvel?

Avengers: Endgame

parameters are updated using the REINFORCE [1] algorithm

Step 3: Answer Generation

★ During training: Sample action
Step 3: Answer Generation

During training: **Sample** action

For answering:

- Take **top actions** and **rank** them
- Main ranking criterion: **prediction score** from policy network, **boosted** if several **agents arrive at same answer** entity

Spider-Man: Far from Home

Captain Marvel

followed by

follows

part of the series

Avengers: Endgame

Marvel Cinematic Universe
Step 4: Reformulation Prediction

★ Determines if two questions are reformulations of each other (reward = -1) or express different intents (reward = 1)
Step 4: Reformulation Prediction

★ Determines if two questions are **reformulations** of each other (reward = -1) or express **different intents** (reward = 1)

★ Fine-tuned BERT-model

Q2: What was next from Marvel?
A2: Spider-Man: Far from Home
Q3: Release date?
ConvRef - Benchmark with Reformulations

★ Builds upon Conversational KG-QA dataset ConvQuestions [2]

ConvRef - Benchmark with Reformulations

- Builds upon Conversational KG-QA dataset **ConvQuestions** [2]
- Up to 4 reformulations per info need, around 205k reformulations in total

ConvRef - Benchmark with Reformulations

- Builds upon Conversational KG-QA dataset ConvQuestions[2]
- Up to 4 reformulations per info need, around 205k reformulations in total
- Data collected in user study with 30 participants

ConvRef - Benchmark with Reformulations

★ Builds upon Conversational KG-QA dataset ConvQuestions [2]
★ Up to 4 reformulations per info need, around 205k reformulations in total
★ Data collected in user study with 30 participants
★ Interacted with baseline system

ConvRef - Benchmark with Reformulations

★ Builds upon Conversational KG-QA dataset ConvQuestions [2]
★ Up to 4 reformulations per info need, around 205k reformulations in total
★ Data collected in user study with 30 participants
★ Interacted with baseline system
★ Participants need to issue a reformulation based on the conversation history and the previously returned wrong answer: differ from simple paraphrases

## ConvRef - Benchmark with Reformulations

<table>
<thead>
<tr>
<th>Nature of reformulation</th>
<th>Percentage</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Words were replaced by synonyms</td>
<td>15%</td>
<td>“When was that released?” - “When was it out?”</td>
</tr>
<tr>
<td>Expected answer types were added</td>
<td>14%</td>
<td>“Who wrote the screenplay?” - “Name of person who wrote the screenplay?”</td>
</tr>
<tr>
<td>Coreferences were replaced by topic entity</td>
<td>24%</td>
<td>“What year did he play in the Summer Olympics?” - “When did Eddie Pope play in the Summer Olympics?”</td>
</tr>
<tr>
<td>Question was rephrased</td>
<td>71%</td>
<td>“Cause of death?” - “Why did Bob Marley die?”</td>
</tr>
<tr>
<td>Words were reordered</td>
<td>5%</td>
<td>“What year did Friends air?” - “Friends aired in year?”</td>
</tr>
<tr>
<td>Completed a partially implicit question</td>
<td>20%</td>
<td>“And what was his sports number there?” - “Number on jersey of Kylian Mbappe in 2018 FIFA world cup?”</td>
</tr>
</tbody>
</table>
Experimental Configurations

★ Four different variations of the CONQUER model stemming from two sources of noise: reformulation predictor and user model
Experimental Configurations

★ Four different variations of the CONQUER model stemming from two sources of noise: reformulation predictor and user model

★ Ideal Reformulation Predictor:
  ○ Always decides correctly whether two questions are reformulations of each other
  ○ We know reformulations based on annotations in ConvRef
Experimental Configurations

★ Four different variations of the CONQUER model stemming from two sources of noise: reformulation predictor and user model

★ Ideal Reformulation Predictor:
  ○ Always decides correctly whether two questions are reformulations of each other
  ○ We know reformulations based on annotations in ConvRef

★ Noisy Reformulation Predictor:
  ○ Fine-tuned BERT model
  ○ Sometimes predictions are incorrect: reformulation is mistaken for new intent and vice versa
Experimental Configurations

★ No access to users during training time, thus **simulate users** using the collected reformulations from ConvRef
Experimental Configurations

★ No access to users during training time, thus simulate users using the collected reformulations from ConvRef

★ Ideal User Model:
  ○ User behaves exactly as in our assumption: reformulates if presented answer was wrong, otherwise issues new question
  ○ Simulated by looping through available reformulations in ConvRef
Experimental Configurations

★ No access to users during training time, thus simulate users using the collected reformulations from ConvRef

★ **Ideal User Model:**
  ○ User behaves exactly as in our assumption: reformulates if presented answer was wrong, otherwise issues new question
  ○ Simulated by looping through available reformulations in ConvRef

★ **Noisy User Model:**
  ○ User can also ask new question even though previous answer was wrong (e.g. out of frustration)
  ○ If no further reformulation available in ConvRef we move to next info need regardless of whether answer was correct or not
Main Results - CONQUER outperforms baseline

★ All CONQUER variants **outperform baseline** CONVEX [2]
★ **Performance** of CONQUER variants **similar** (best variant: NoisyUser-IdealRef)

Main Results - CONQUER answers questions earlier
Main Results - CONQUER answers questions earlier
## Domain-wise Results

<table>
<thead>
<tr>
<th>Method</th>
<th>Movies</th>
<th>TV Series</th>
<th>Music</th>
<th>Books</th>
<th>Soccer</th>
</tr>
</thead>
<tbody>
<tr>
<td>IdealUser-IdealRef</td>
<td>0.320</td>
<td>0.316</td>
<td>0.281</td>
<td>0.449</td>
<td>0.329</td>
</tr>
<tr>
<td>IdealUser-NoisyRef</td>
<td>0.344</td>
<td>0.340</td>
<td>0.303</td>
<td>0.425</td>
<td>0.308</td>
</tr>
<tr>
<td>NoisyUser-IdealRef</td>
<td><strong>0.368</strong></td>
<td><strong>0.367</strong></td>
<td><strong>0.324</strong></td>
<td>0.413</td>
<td><strong>0.329</strong></td>
</tr>
<tr>
<td>NoisyUser-NoisyRef</td>
<td>0.327</td>
<td>0.296</td>
<td>0.300</td>
<td>0.381</td>
<td>0.327</td>
</tr>
<tr>
<td>CONVEX</td>
<td>0.274</td>
<td>0.188</td>
<td>0.195</td>
<td>0.224</td>
<td>0.244</td>
</tr>
</tbody>
</table>
Results on ConvQuestions Benchmark

<table>
<thead>
<tr>
<th>Method</th>
<th>P@1</th>
<th>Hit@5</th>
<th>MRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONQUER (trained with gold labels)</td>
<td>0.263</td>
<td>0.343</td>
<td>0.298</td>
</tr>
<tr>
<td>CONVEX</td>
<td>0.184</td>
<td>0.219</td>
<td>0.200</td>
</tr>
</tbody>
</table>
## Ablation: Context Modeling

<table>
<thead>
<tr>
<th>Context Model</th>
<th>P@1</th>
<th>Hit@5</th>
<th>MRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curr. ques. + cxt. ent.</td>
<td>0.294</td>
<td>0.407</td>
<td>0.346</td>
</tr>
<tr>
<td>Curr. ques. + cxt. ent. + first ques.</td>
<td>0.254</td>
<td>0.370</td>
<td>0.305</td>
</tr>
<tr>
<td>Curr. ques. + cxt. ent. + first ques. + prev. ques.</td>
<td>0.257</td>
<td>0.370</td>
<td>0.307</td>
</tr>
<tr>
<td>Curr. ques. + cxt. ent. + first refs. + prev. refs.</td>
<td>0.262</td>
<td>0.382</td>
<td>0.316</td>
</tr>
</tbody>
</table>
## Ablation: Action Choices

<table>
<thead>
<tr>
<th>Method</th>
<th>P@1</th>
<th>Hit@5</th>
<th>MRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Path</td>
<td>0.294</td>
<td>0.407</td>
<td>0.346</td>
</tr>
<tr>
<td>Context entity + Path</td>
<td>0.293</td>
<td>0.408</td>
<td>0.346</td>
</tr>
<tr>
<td>Path + Answer entity</td>
<td>0.275</td>
<td>0.394</td>
<td>0.329</td>
</tr>
<tr>
<td>Context entity + Path + Answer entity</td>
<td>0.273</td>
<td>0.398</td>
<td>0.328</td>
</tr>
</tbody>
</table>
Ablation: Answer Aggregation

<table>
<thead>
<tr>
<th>Method</th>
<th>P@1</th>
<th>Hit@5</th>
<th>MRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Add scores</td>
<td>0.294</td>
<td>0.407</td>
<td>0.346</td>
</tr>
<tr>
<td>Max scores</td>
<td>0.294</td>
<td>0.406</td>
<td>0.344</td>
</tr>
<tr>
<td>Max scores (ties resolved with majority voting)</td>
<td>0.291</td>
<td>0.405</td>
<td>0.343</td>
</tr>
<tr>
<td>Majority voting (ties resolved with max score)</td>
<td>0.273</td>
<td><strong>0.408</strong></td>
<td>0.334</td>
</tr>
</tbody>
</table>
Performance of Reformulation Predictor

★ Fine-tuned BERT model:
  ○ **Positive** samples: *same intents* from *same conversation* (reformulations)
  ○ **Negative** samples: *different intents* from *same conversation*
Performance of Reformulation Predictor

★ Fine-tuned BERT model:
  - Positive samples: same intents from same conversation (reformulations)
  - Negative samples: different intents from same conversation

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Intent</td>
<td>0.986</td>
<td>0.944</td>
<td>0.965</td>
</tr>
<tr>
<td>Reformulation</td>
<td>0.810</td>
<td>0.948</td>
<td>0.873</td>
</tr>
</tbody>
</table>
Conclusion and Future Work

★ CONQUER model:
  ○ RL-based method for conversational QA
  ○ Leverages noisy implicit feedback coming from reformulations, learns from positive and negative feedback
  ○ Robust to noise

★ Reformulation predictor

★ ConvRef: Benchmark with reformulations
Conclusion and Future Work

★ **CONQUER** model:
  ○ RL-based method for conversational QA
  ○ Leverages noisy implicit feedback coming from reformulations, learns from positive and negative feedback
  ○ Robust to noise

★ Reformulation predictor

★ **ConvRef**: Benchmark with reformulations

Future work may include:
★ Improved modeling of conversational context
★ Context entity detection as part of neural model
★ Further feedback signals
Conclusion and Future Work

★ CONQUER model:
  ○ RL-based method for conversational QA
  ○ Leverages noisy implicit feedback coming from reformulations, learns from positive and negative feedback
  ○ Robust to noise

★ Reformulation predictor
★ ConvRef: Benchmark with reformulations

Future work may include:
★ Improved modeling of conversational context
★ Context entity detection as part of neural model
★ Further feedback signals

Contact: mkaiser@mpi-inf.mpg.de, @mag_kaiser
Benchmark+Demo: https://conquer.mpi-inf.mpg.de
Code: https://github.com/magkai/CONQUER