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

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: ...

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

Realistic Conversation

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

Realistic Conversation

New intent:

Q1: When was **Avengers: Endgame released** in **Germany**?

A1: 24 April 2019

New intent:

Q2: What was **next** from **Marvel**?

A2: **Stan Lee**

Reformulation:

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

A21: Marvel Cinematic Universe

Reformulation:

Q22: The **following movie** in the **Marvel series**?

A22: Spider-Man: Far from Home

New intent:

Q3: Release date?

A3: ...

Challenges:

- ★ Short, incomplete questions
- ★ Implicit context

Learn from conversational stream:

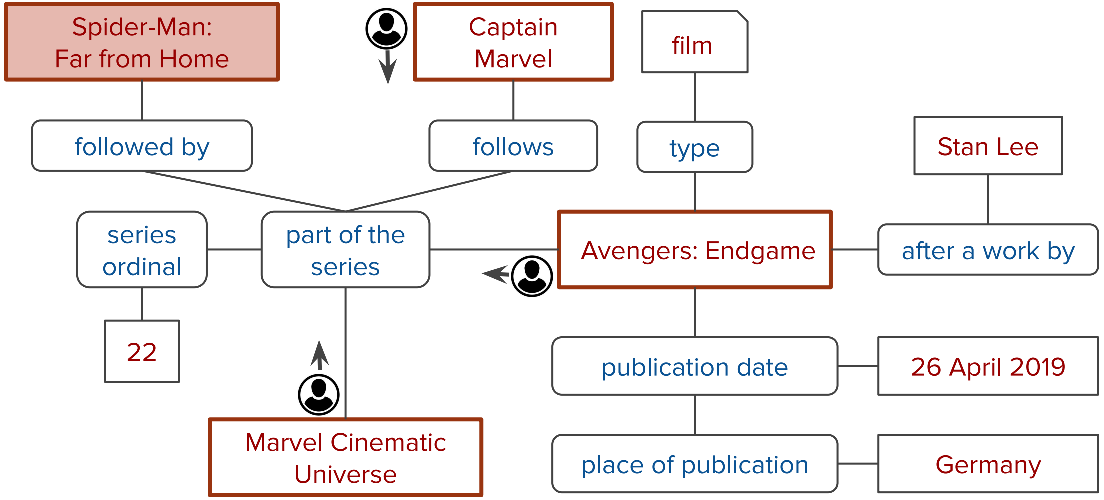
Reformulation = Wrong answer
New intent = Correct answer

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**

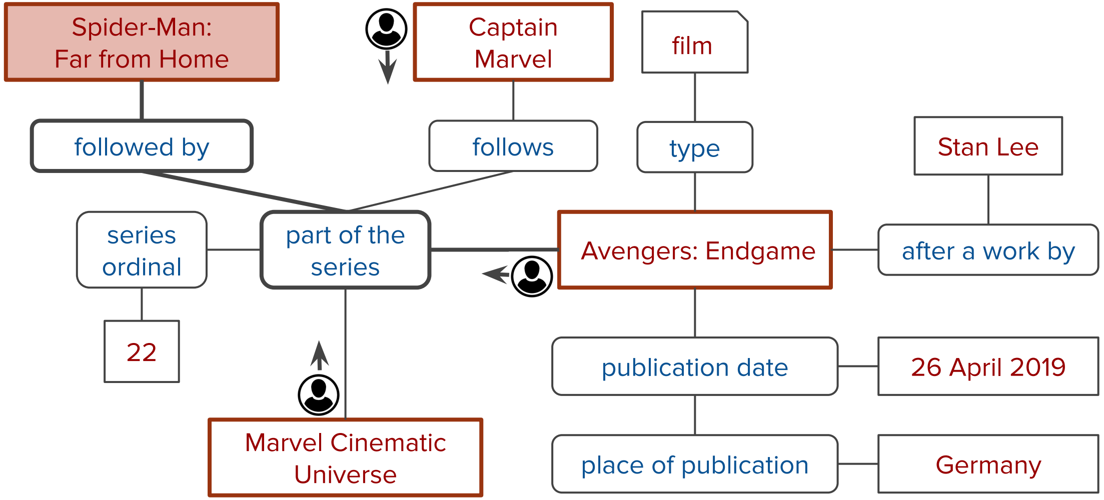
Basic Idea

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



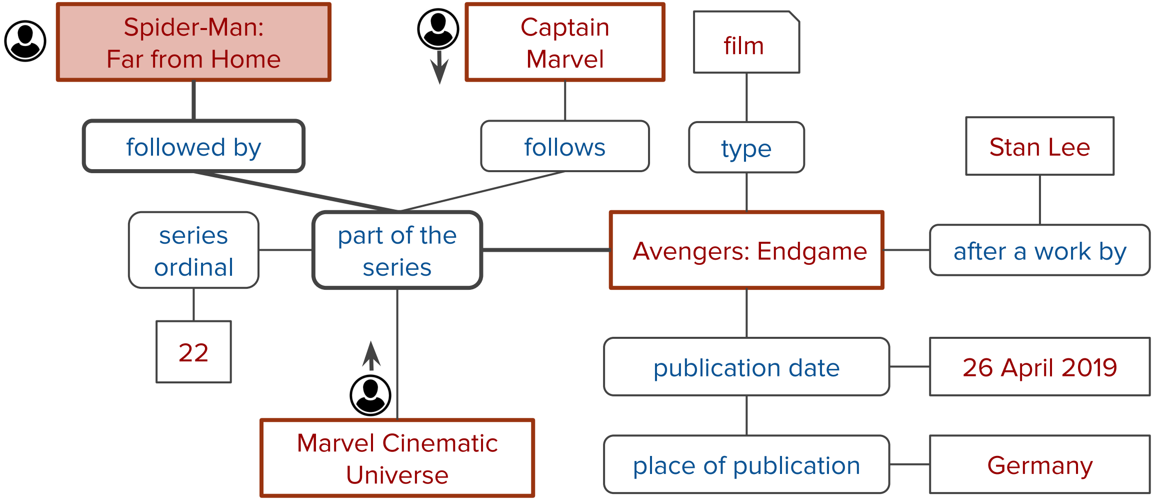
Basic Idea

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



Basic Idea

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



KG Representation

<u>N-ary facts connected via statement-ids:</u>

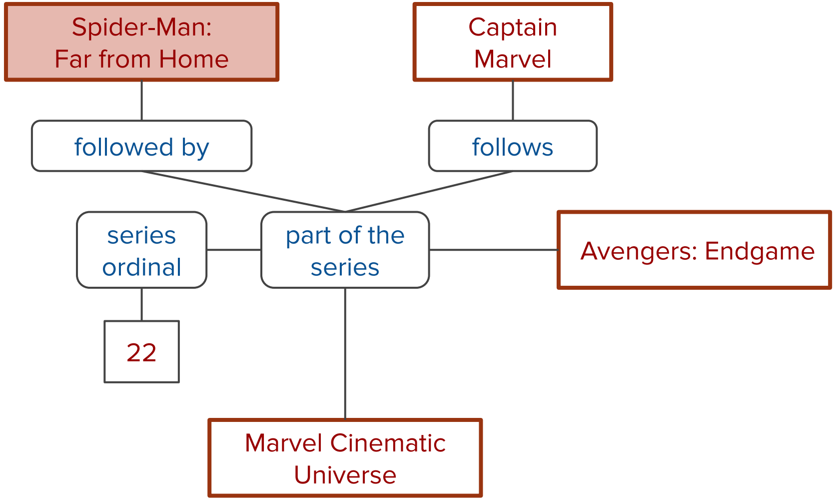
<AvengersEndgame, partOfSeries, 123>

<123, partOfSeries, MarvelCinematicUniverse>

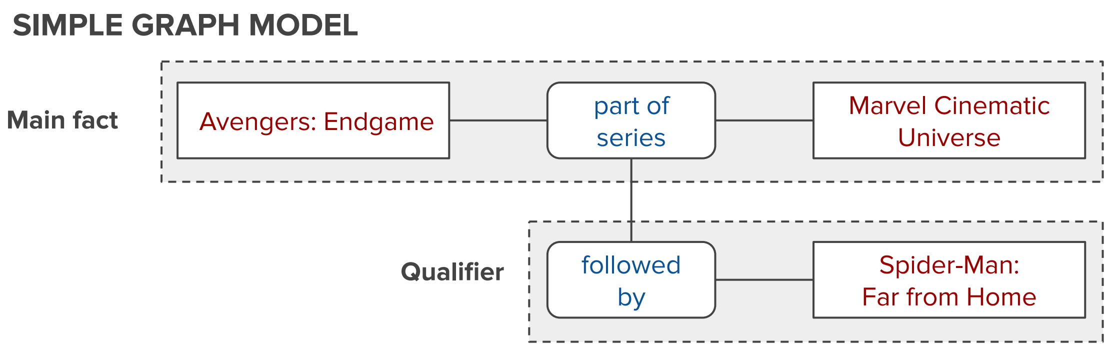
<123, followedBy, SpiderManFarFromHome>

<123, follows, CaptainMarvel>

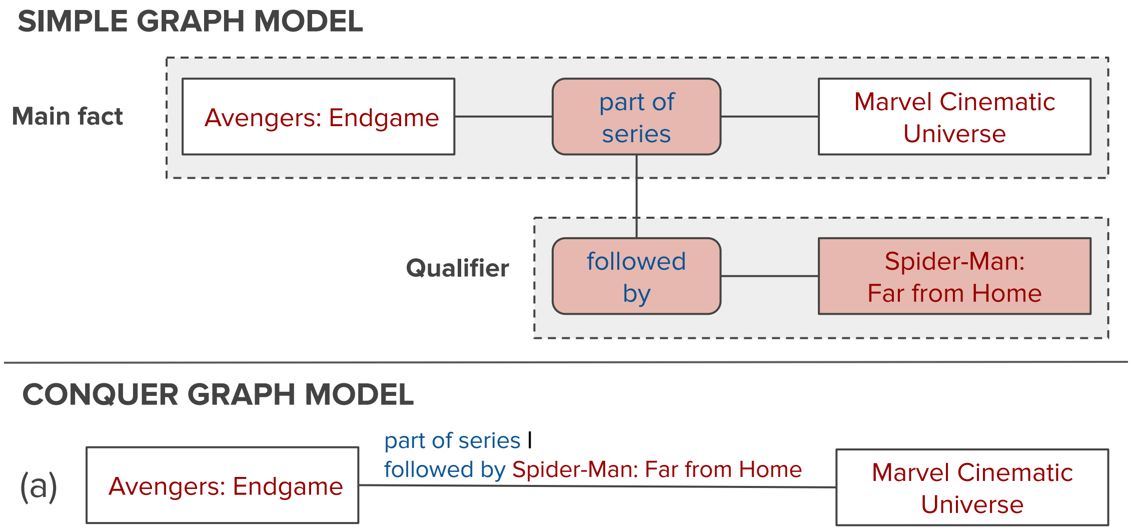
<123, seriesOrdinal, 22>



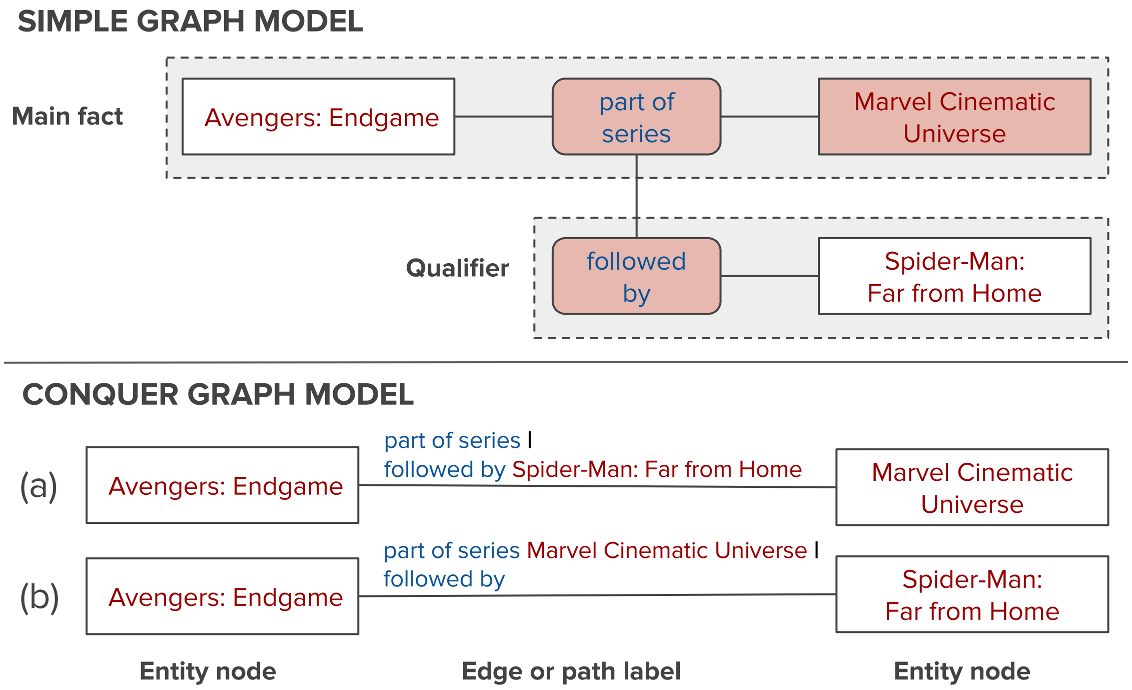
KG Representation



KG Representation



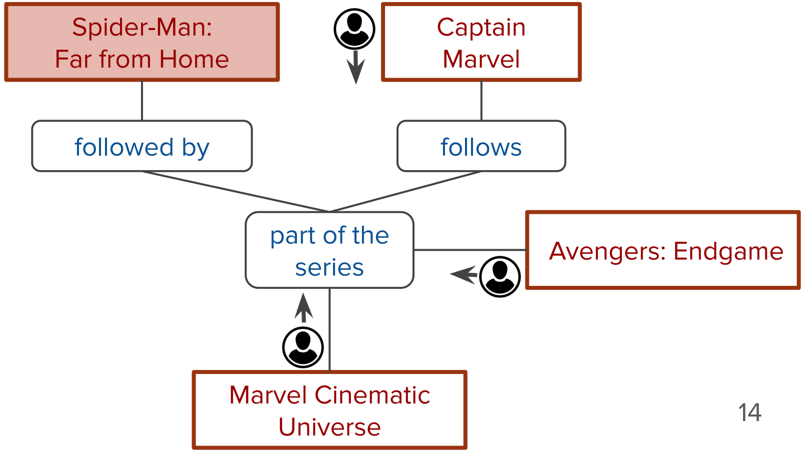
KG Representation



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**?*

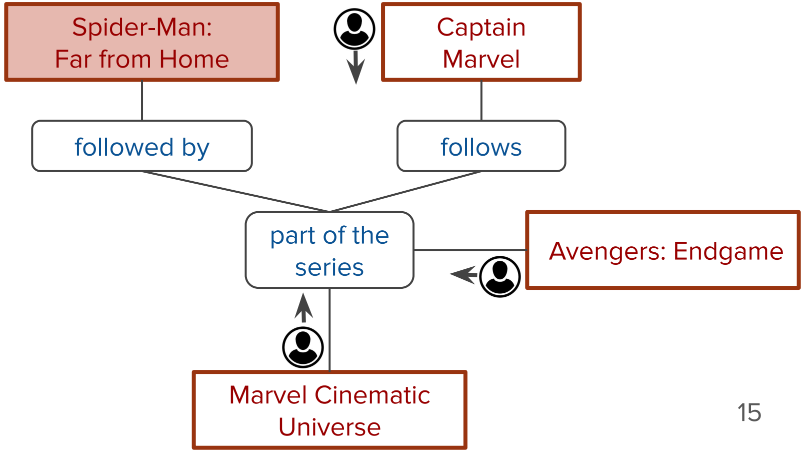


Step 1: Context Entity Detection

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

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

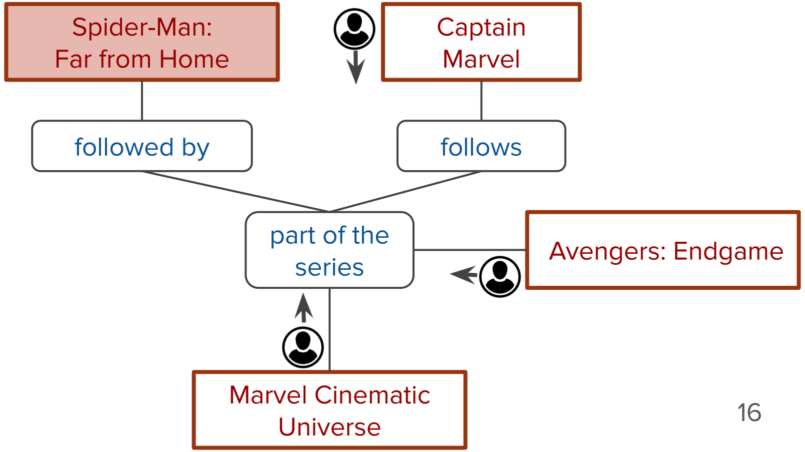
*Q1: When was **Avengers: Endgame** released in Germany?*
A1: 24 April 2019
*Q2: What was next from **Marvel**?*



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**?*



Step 1: Context Entity Detection

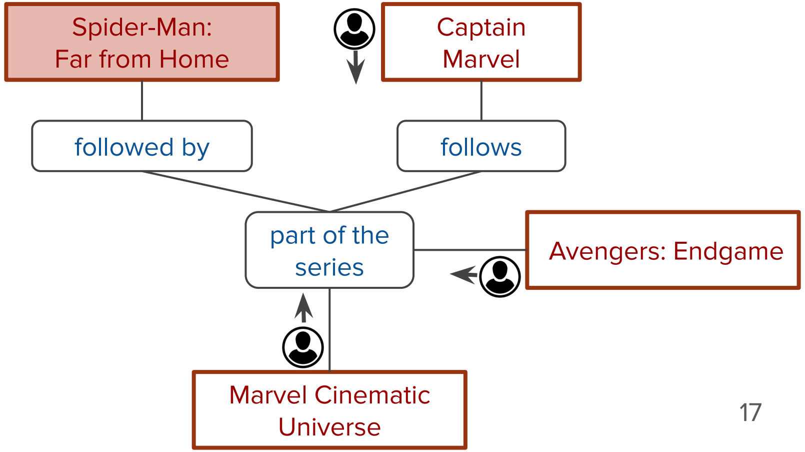
★ Find **entities** relevant to **current question** and its **conversational context**

★ **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**

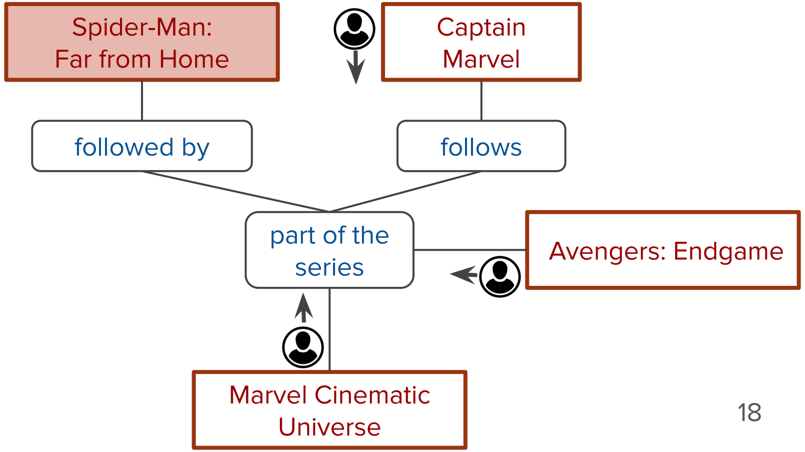
*Q1: When was **Avengers: Endgame** released in Germany?*
A1: 24 April 2019
*Q2: What was next from **Marvel**?*



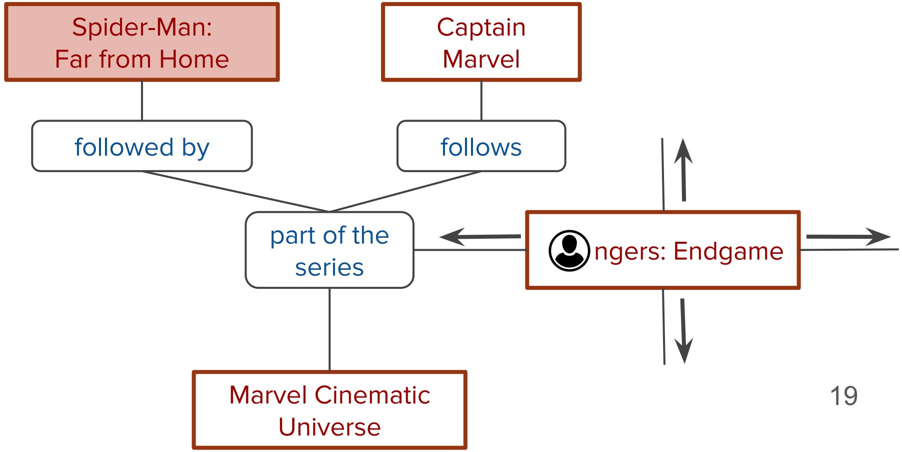
Step 1: Context Entity Detection

- ★ 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:
  - Lexical match
  - Neighbor overlap
  - NED score
  - KG prior

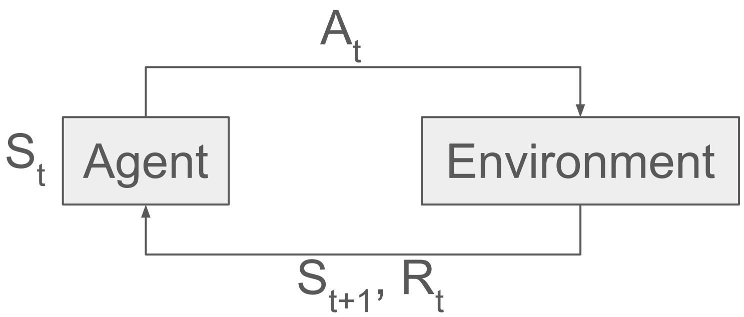
*Q1: When was **Avengers: Endgame** released in Germany?*
A1: 24 April 2019
*Q2: What was next from **Marvel**?*

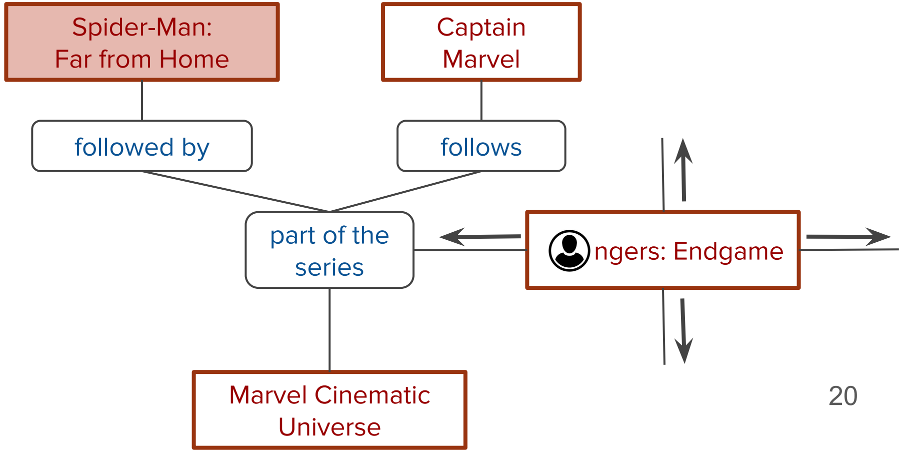


Step 2: Path Prediction



Step 2: Path Prediction





Step 2: Path Prediction

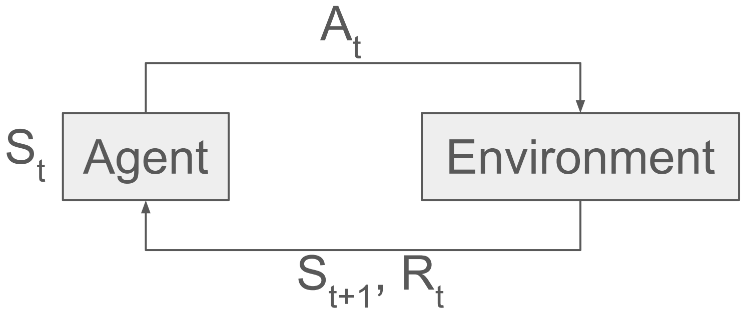
★ **States:** current question, context entity, conversation history (optional)

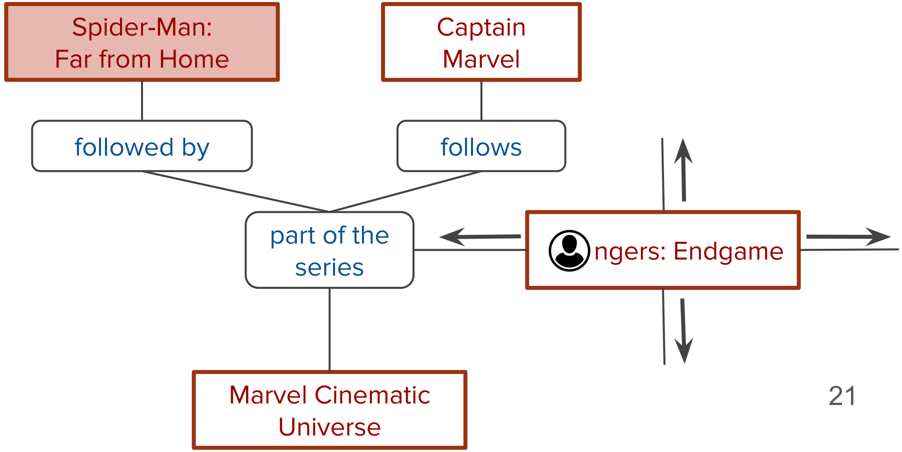
★ **Actions:** all outgoing paths from the context entity node

★ **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

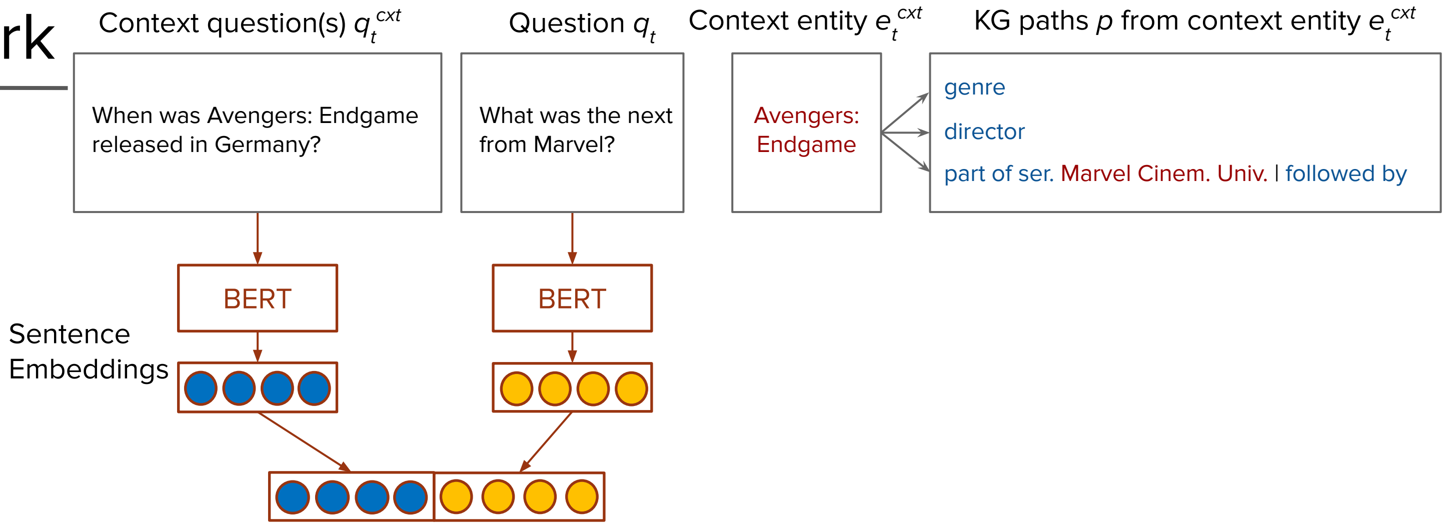




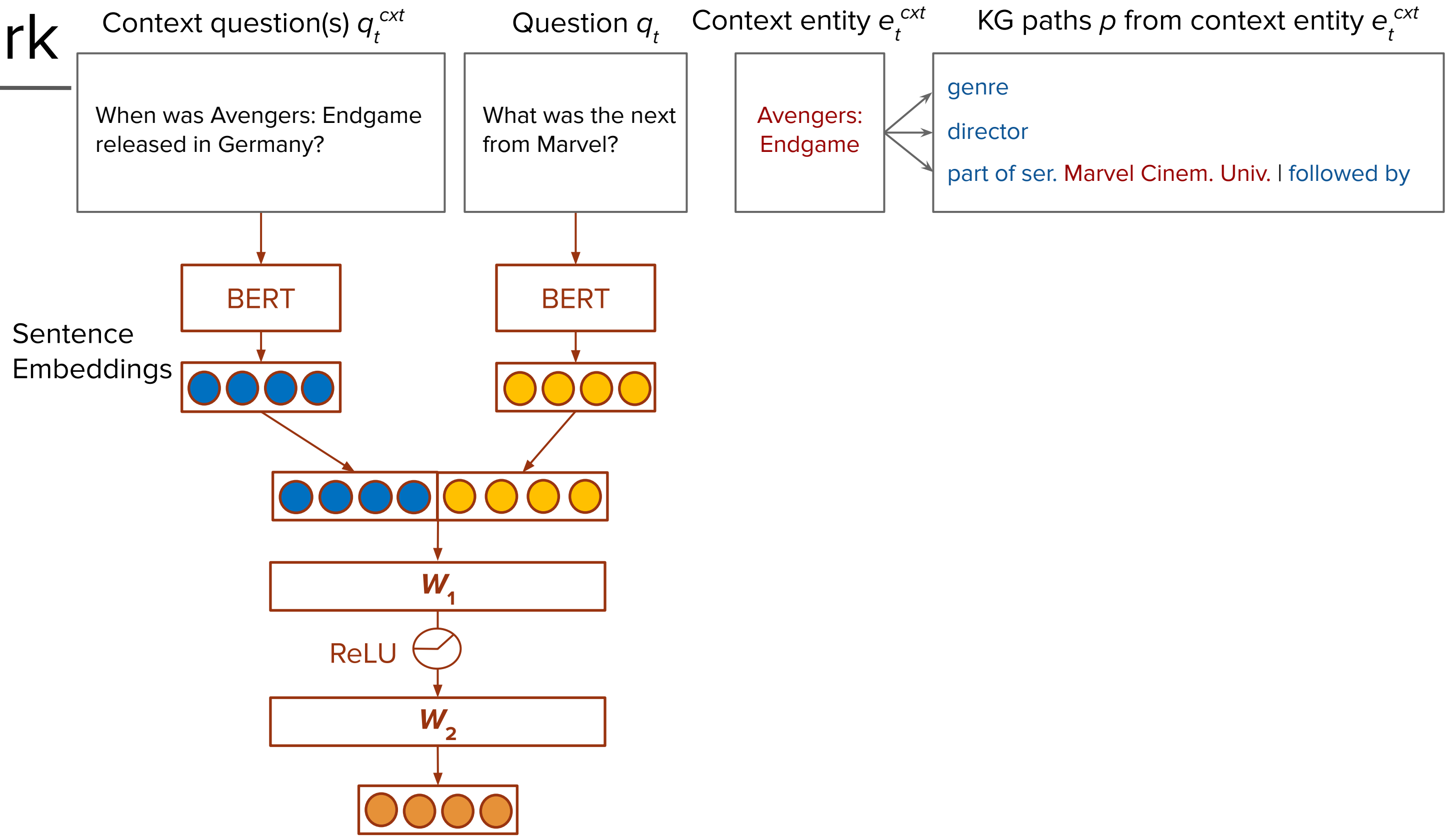
Policy Network

Context question(s) q_t^{cxt}	Question q_t	Context entity e_t^{cxt}	KG paths p from context entity e_t^{cxt}
When was Avengers: Endgame released in Germany?	What was the next from Marvel?	Avengers: Endgame	genre director part of ser. Marvel Cinem. Univ. | followed by

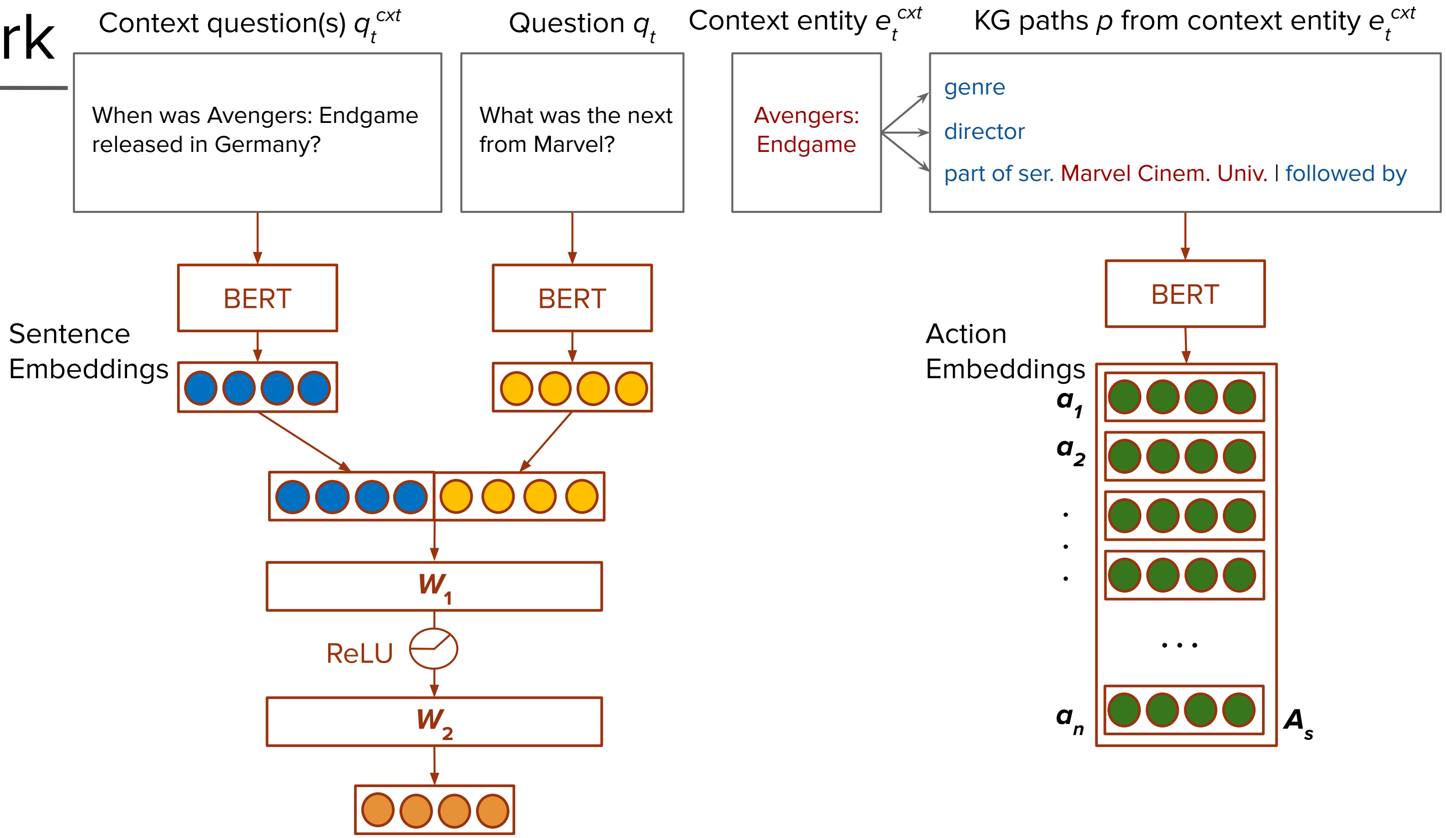
Policy Network



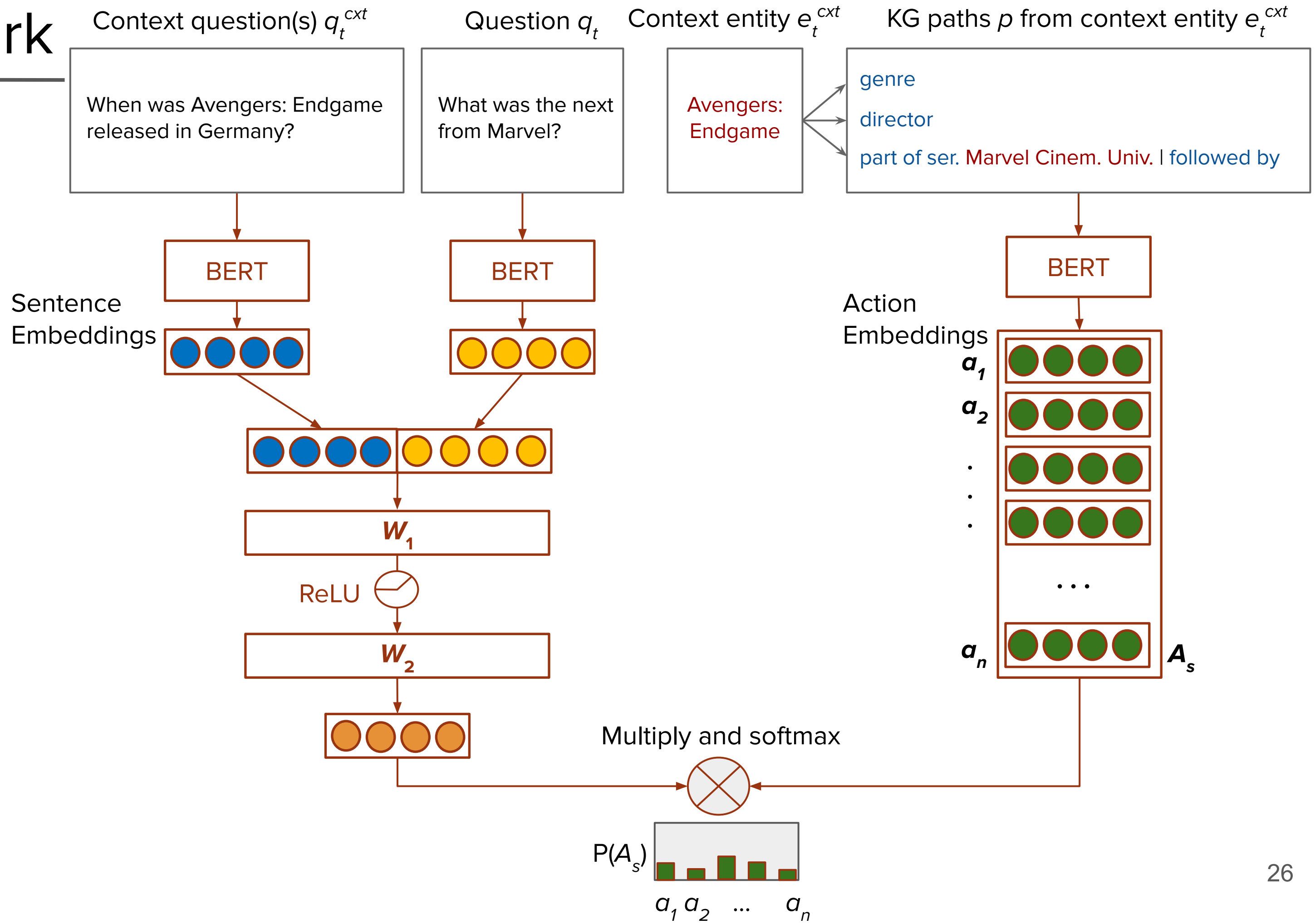
Policy Network



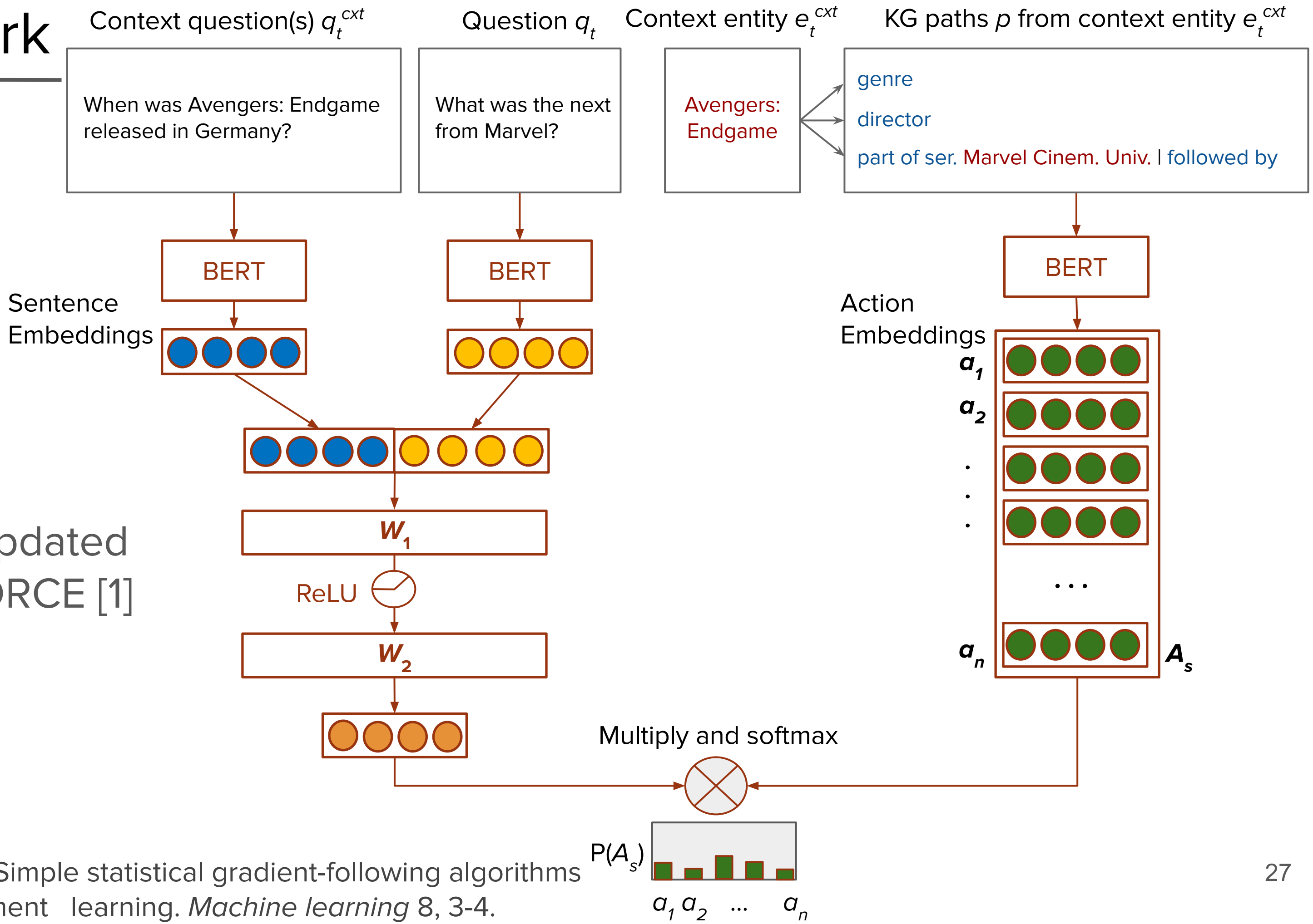
Policy Network



Policy Network



Policy Network



Parameters are updated using the REINFORCE [1] algorithm

[1] Ronald J Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning* 8, 3-4.

Step 3: Answer Generation

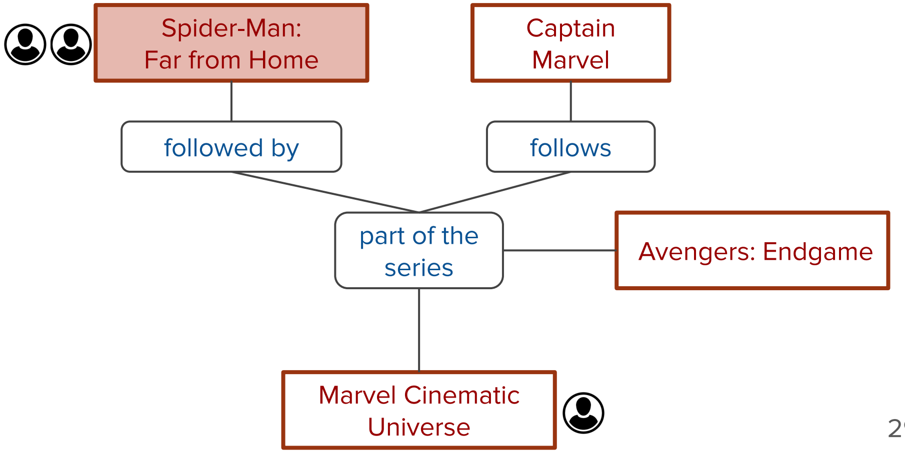
★ During **training**: **Sample** action

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



Step 4: Reformulation Prediction

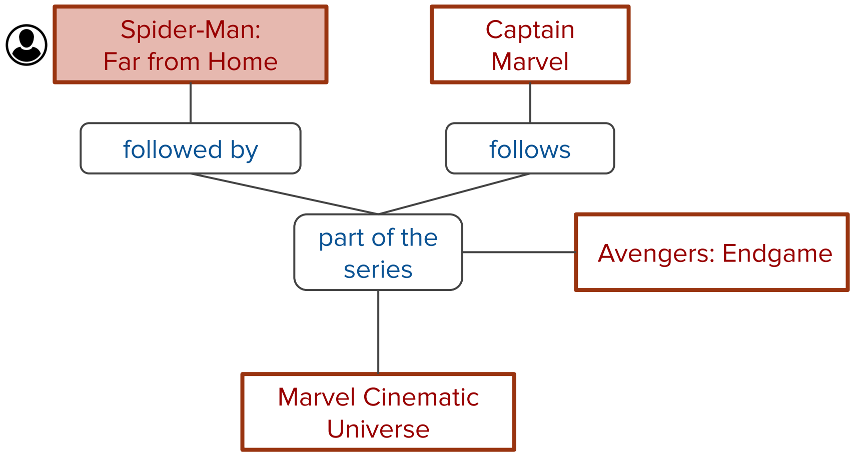
★ Determines if two questions are **reformulations** of each other (reward = -1) or express **different intents** (reward = 1)

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?

Reward = 1



ConvRef - Benchmark with Reformulations

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

[2] Philipp Christmann, Rishiraj Saha Roy, Abdalghani Abujabal, Jyotsna Singh, and Gerhard Weikum. 2019. Look before you Hop: Conversational Question Answering over Knowledge Graphs Using Judicious Context Expansion. In CIKM.

ConvRef - Benchmark with Reformulations

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

[2] Philipp Christmann, Rishiraj Saha Roy, Abdalghani Abujabal, Jyotsna Singh, and Gerhard Weikum. 2019. Look before you Hop: Conversational Question Answering over Knowledge Graphs Using Judicious Context Expansion. In CIKM.

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

[2] Philipp Christmann, Rishiraj Saha Roy, Abdalghani Abujabal, Jyotsna Singh, and Gerhard Weikum. 2019. Look before you Hop: Conversational Question Answering over Knowledge Graphs Using Judicious Context Expansion. In CIKM.

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

[2] Philipp Christmann, Rishiraj Saha Roy, Abdalghani Abujabal, Jyotsna Singh, and Gerhard Weikum. 2019. Look before you Hop: Conversational Question Answering over Knowledge Graphs Using Judicious Context Expansion. In CIKM.

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**

[2] Philipp Christmann, Rishiraj Saha Roy, Abdalghani Abujabal, Jyotsna Singh, and Gerhard Weikum. 2019. Look before you Hop: Conversational Question Answering over Knowledge Graphs Using Judicious Context Expansion. In CIKM.

ConvRef - Benchmark with Reformulations

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

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

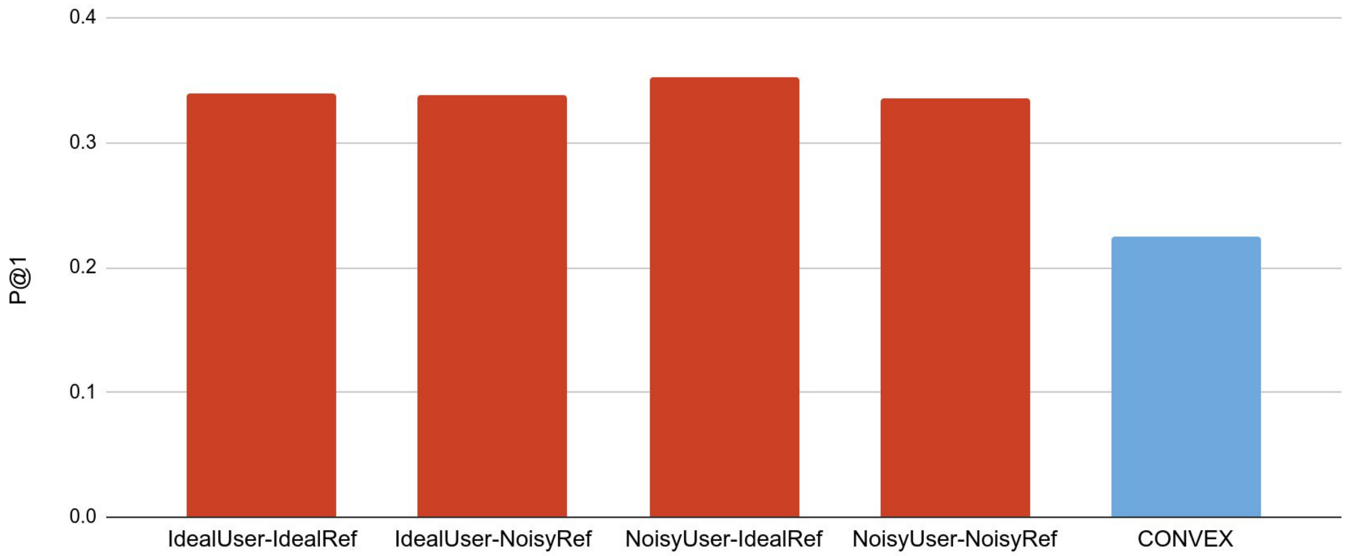
- ★ No access to users during training time, thus **simulate users** using the collected reformulations from ConvRef
- ★ **Ideal User Model:**
  - User behaves exactly as in our assumption: reformulates if presented answer was wrong, otherwise issues new question
  - Simulated by looping through available 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

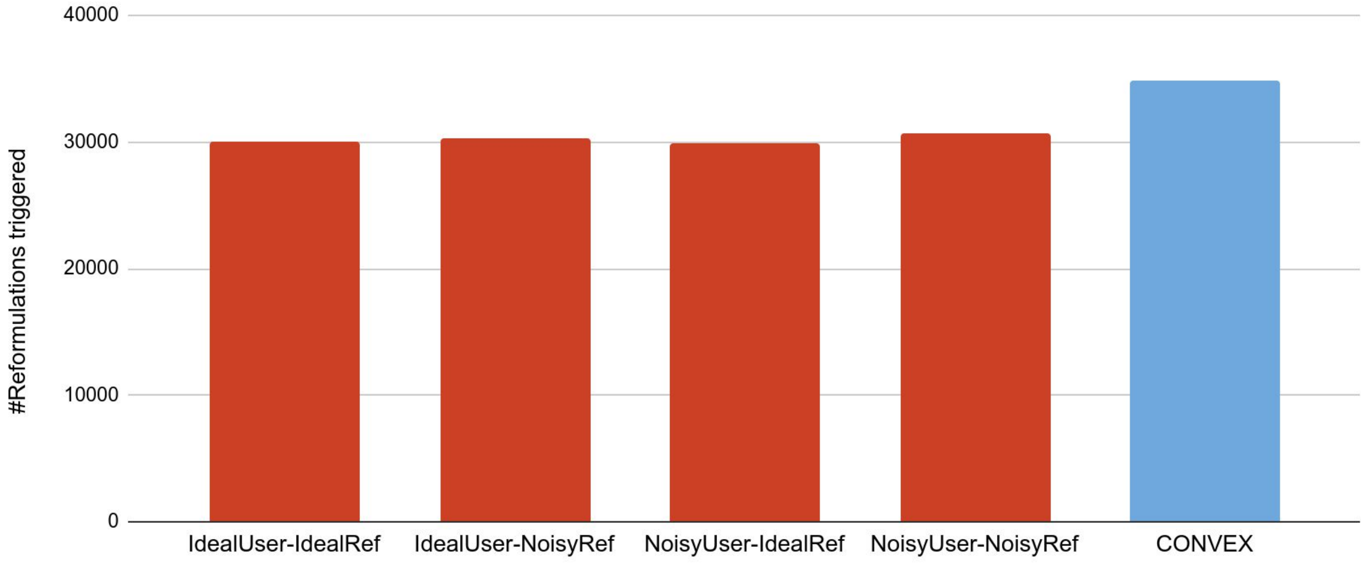
Main Results - CONQUER outperforms baseline

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

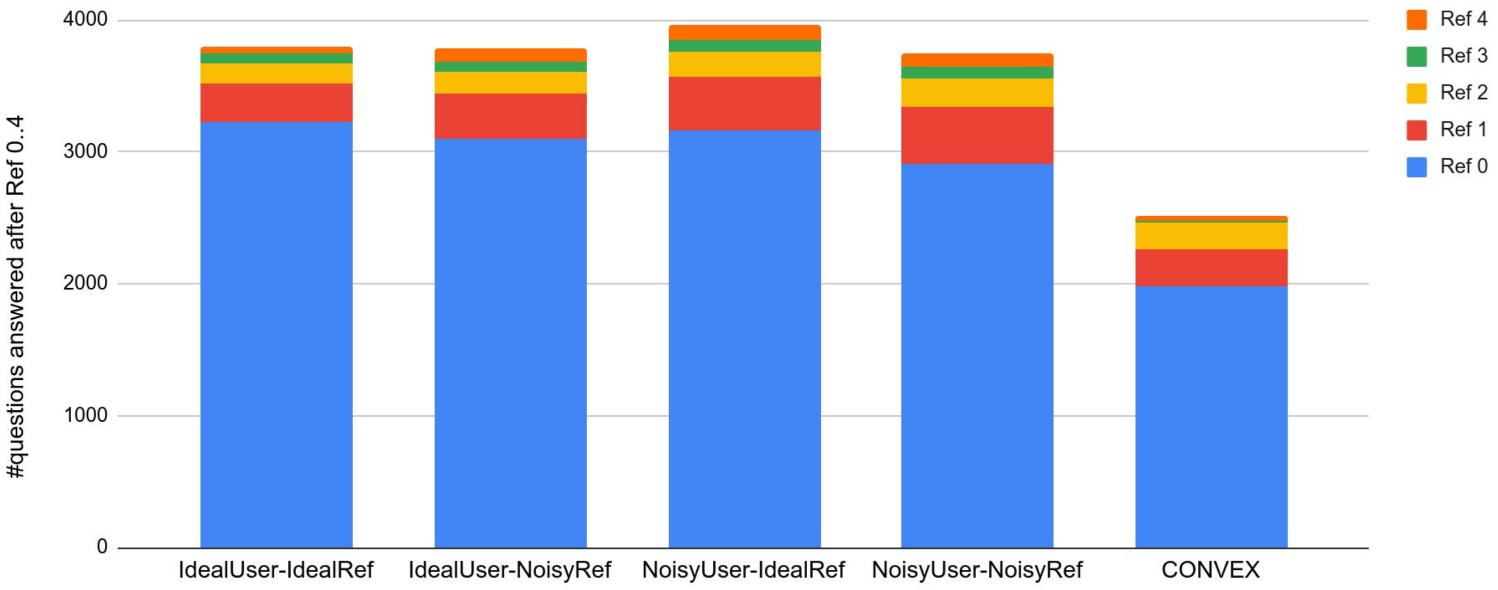


[2] Philipp Christmann, Rishiraj Saha Roy, Abdalghani Abujabal, Jyotsna Singh, and Gerhard Weikum. 2019. Look before you Hop: Conversational Question Answering over Knowledge Graphs Using Judicious Context Expansion. In CIKM.

Main Results - CONQUER answers questions earlier



Main Results - CONQUER answers questions earlier



Domain-wise Results

Method	**Movies**	**TV Series**	**Music**	**Books**	**Soccer**
IdealUser-IdealRef	0.320	0.316	0.281	**0.449**	**0.329**
IdealUser-NoisyRef	0.344	0.340	0.303	0.425	0.308
NoisyUser-IdealRef	**0.368**	**0.367**	**0.324**	0.413	**0.329**
NoisyUser-NoisyRef	0.327	0.296	0.300	0.381	0.327
CONVEX	0.274	0.188	0.195	0.224	0.244

Results on ConvQuestions Benchmark

Method	**P@1**	**Hit@5**	**MRR**
CONQUER (trained with gold labels)	**0.263**	**0.343**	**0.298**
CONVEX	0.184	0.219	0.200

Ablation: Context Modeling

Context Model	**P@1**	**Hit@5**	**MRR**
Curr. ques. + cxt. ent.	**0.294**	**0.407**	**0.346**
Curr. ques. + cxt. ent. + first ques.	0.254	0.370	0.305
Curr. ques. + cxt. ent. + first ques. + prev. ques.	0.257	0.370	0.307
Curr. ques. + cxt. ent. + first refs. + prev. refs.	0.262	0.382	0.316

Ablation: Action Choices

Method	P@1	Hit@5	MRR
Path	**0.294**	0.407	**0.346**
Context entity + Path	0.293	**0.408**	**0.346**
Path + Answer entity	0.275	0.394	0.329
Context entity + Path + Answer entity	0.273	0.398	0.328

Ablation: Answer Aggregation

Method	**P@1**	**Hit@5**	**MRR**
Add scores	**0.294**	0.407	**0.346**
Max scores	**0.294**	0.406	0.344
Max scores (ties resolved with majority voting)	0.291	0.405	0.343
Majority voting (ties resolved with max score)	0.273	**0.408**	0.334

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**

	Precision	**Recall**	**F1**
New Intent	0.986	0.944	0.965
Reformulation	0.810	0.948	0.873

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