Reinforcement Learning from Reformulations in Conversational Question Answering over Knowledge Graphs

Magdalena Kaiser, Rishiraj Saha Roy and Gerhard Weikum

Max Planck Institute for Informatics, Germany

SAMPLE 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: ...

CONVERSATIONAL QA IS CHALLENGING

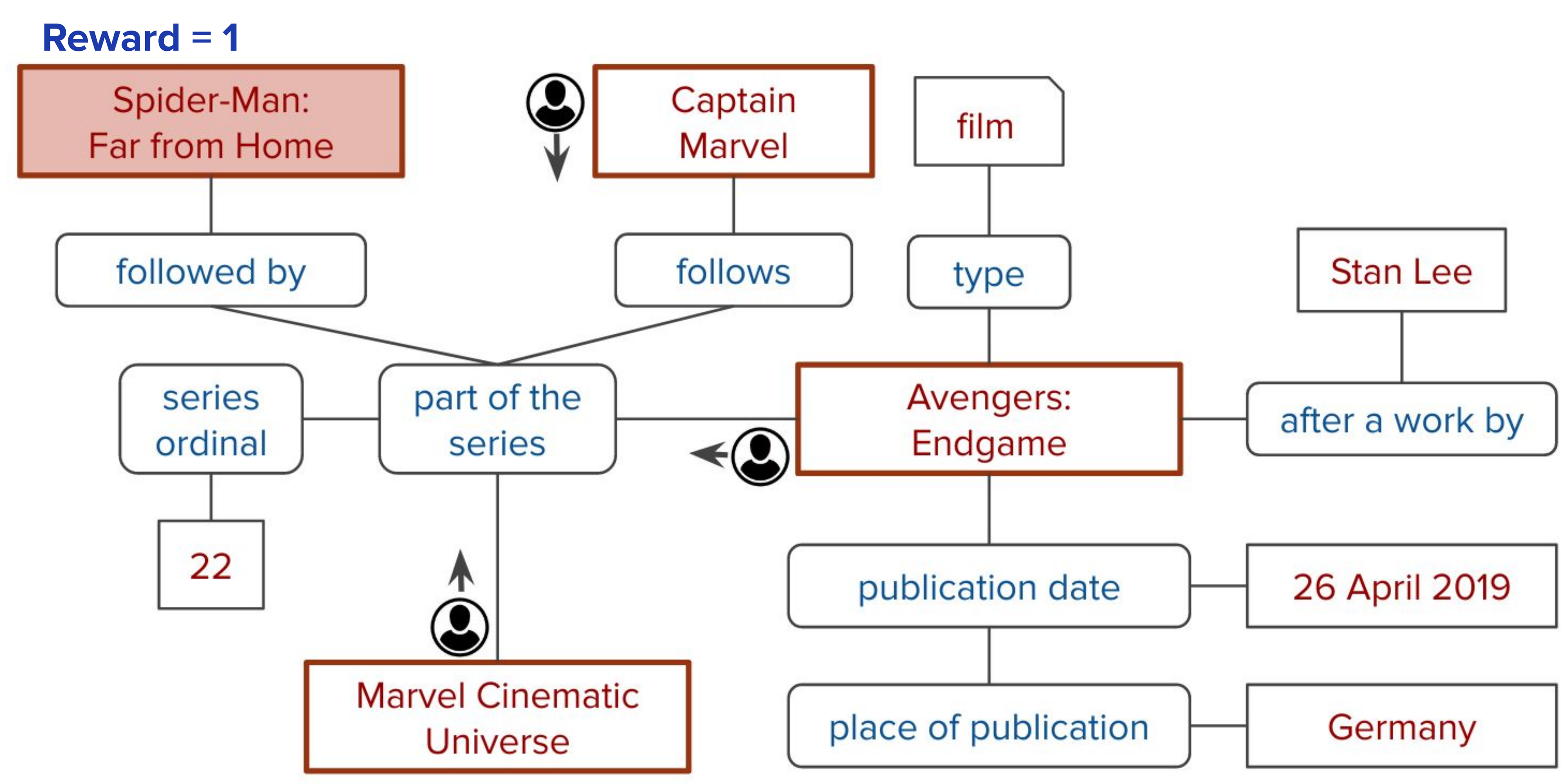
- **Short, incomplete** questions
- **Implicit** context

IMPROVING CONVQA BY LEARNING FROM FEEDBACK

- Current systems learn from gold QA pairs: unrealistic
- **CONQUER**: **Reinforcement learning model** for QA
  - Learns from **conversational stream** in the **absence of gold answers**
  - **Rewards** based on **reformulations: +1** (new intent = correct previous answer), **-1** (reformulation = wrong previous answer)
- **Reformulation detector** based on BERT
- **ConvRef**: Conversational QA **benchmark with reformulations**

CONQUER WORKFLOW

1. Detect **context entities** in conversation = **start points** for **agents' walk** by **scoring KG neighborhood**
2. **Predict path** by applying a **policy network trained** with **REINFORCE** algorithm
3. Generate anwer: follow **sampled path** (during training), **take top scoring paths** and **aggregate answer** (at answering time)
4. **Predict** if **next question** is a **reformulation** by using a **fine-tuned BERT** model and give **reward** accordingly

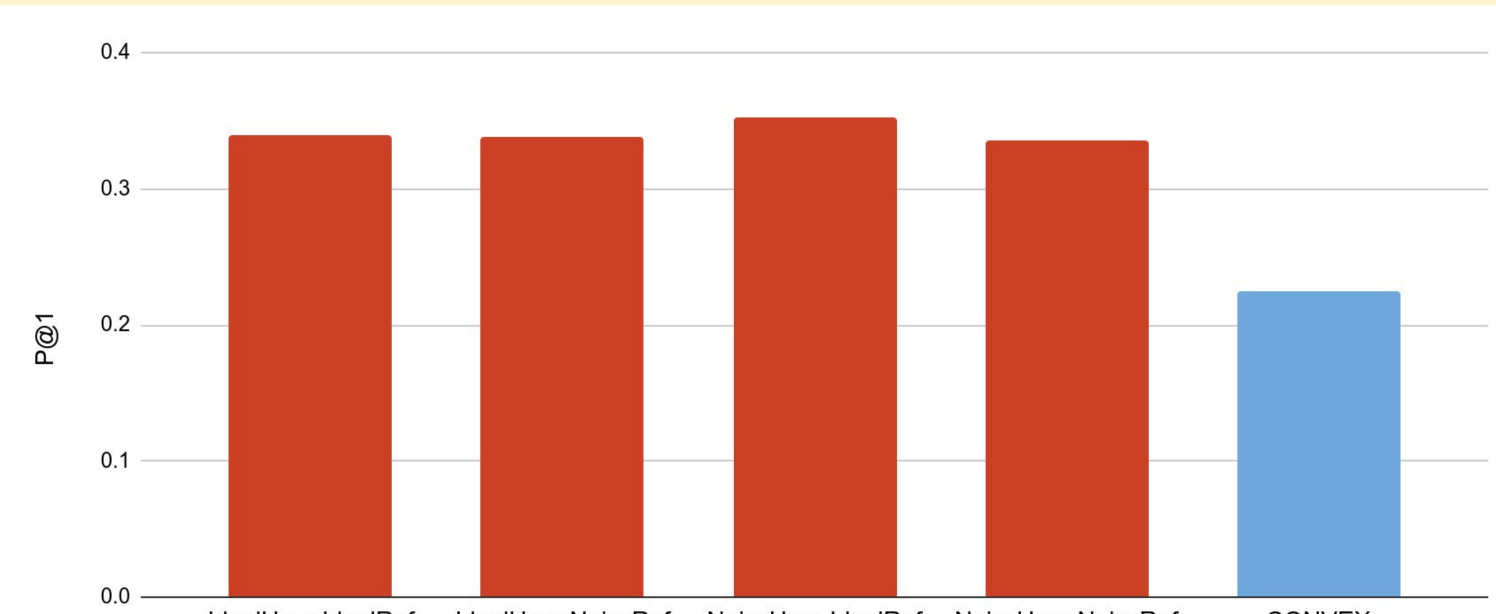


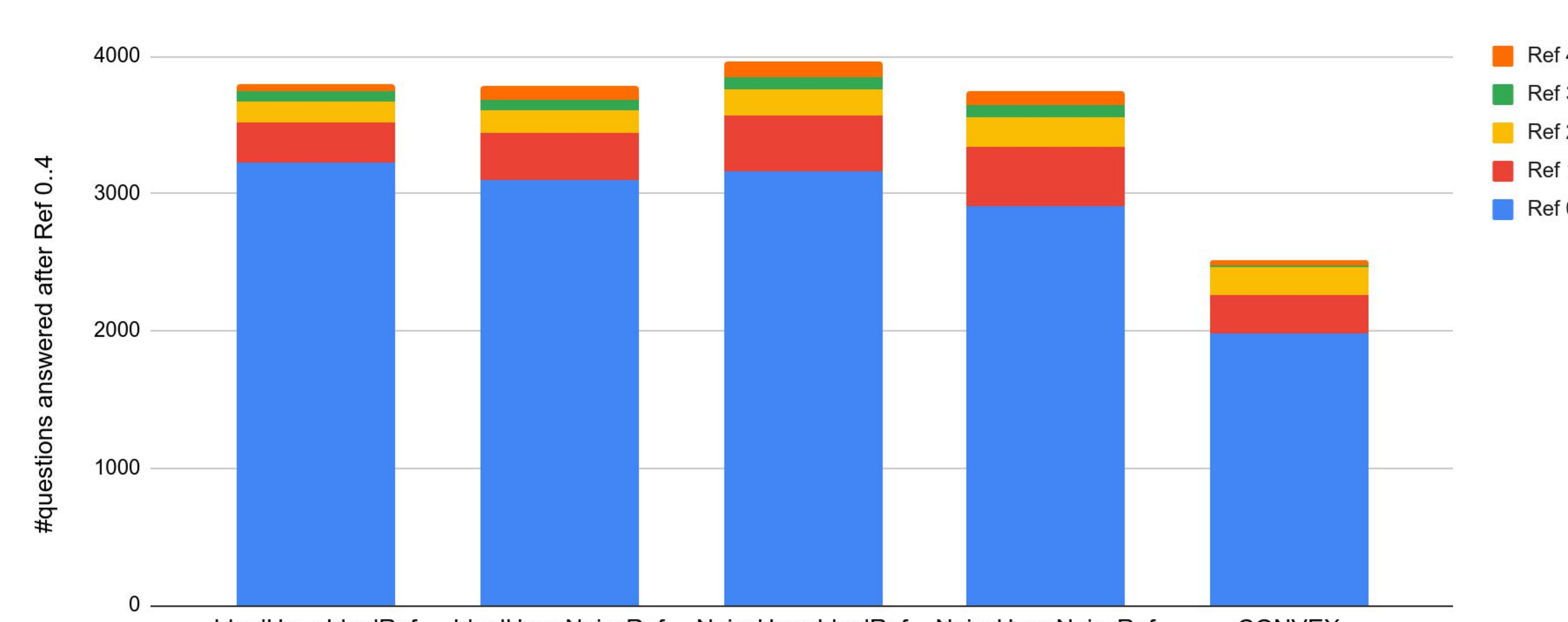
EXPERIMENTAL VARIANTS

- **Four variants** of CONQUER to model **two sources of noise (reformulation predictor** and **user behavior)**:
- **Ideal Reformulation Predictor:**
  - No wrong predictions
- **Noisy Reformulation Predictor:**
  - Fine-tuned BERT model that can make wrong predictions
- **Ideal User Model:**
  - User behaves as in our assumption: reformulates if answer was wrong, otherwise issues new question
- **Noisy User Model:**
  - User can also ask new question even though previous answer was wrong

ConvRef BENCHMARK

- **Builds upon** conversational KG-QA dataset **ConvQuestions** (11k conversations from 5 different domains)
- **User study** to **collect reformulations** by **interacting** with **baseline QA system**
- **Different from paraphrases:** Reformulations **based on conversation history** and system-generated **wrong answer**





CONQUER SUCCESSFULLY LEARNS FROM REFORMULATIONS IN THE PRESENCE OF NOISE

- **CONQUER outperforms** SOTA **baseline** CONVEX **on ConvRef** and **ConvQuestions**
- **Similar performance** of CONQUER **variants**
- **CONQUER answers** more questions **earlier:** requires **less reformulations**

max planck institut informatik

Contact: mkaiser@mpi-inf.mpg.de
Complete Information: conquer.mpi-inf.mpg.de