

Pseudo Ambiguous and Clarifying Questions Based on Sentence Structures Toward Clarifying Question Answering System

Yuya Nakano, Seiya Kawano, Koichiro Yoshino,
Katsuhito Sudoh and Satoshi Nakamura

Nara Institute of Science and Technology (NAIST)
Guardian Robot Project, RIKEN



奈良先端大
知能コミュニケーション研究室
Augmented Human
Communication Laboratory,
NAIST



Question Answering (QA)

◆ QA systems generate an answer to a given question

- Retrieval approach is widely used to realize QA system
- The problem is defined as a two-stage task
 - Document retrieval to find documents that may contain answers to questions
 - Span prediction to predict

Q. Who is the inventor of Volleyball?

Volleyball is a team sport in which two teams of six players are separated by a net.
William G. Morgan, created a new game...



QA model



A. **William G. Morgan**

Question & Document

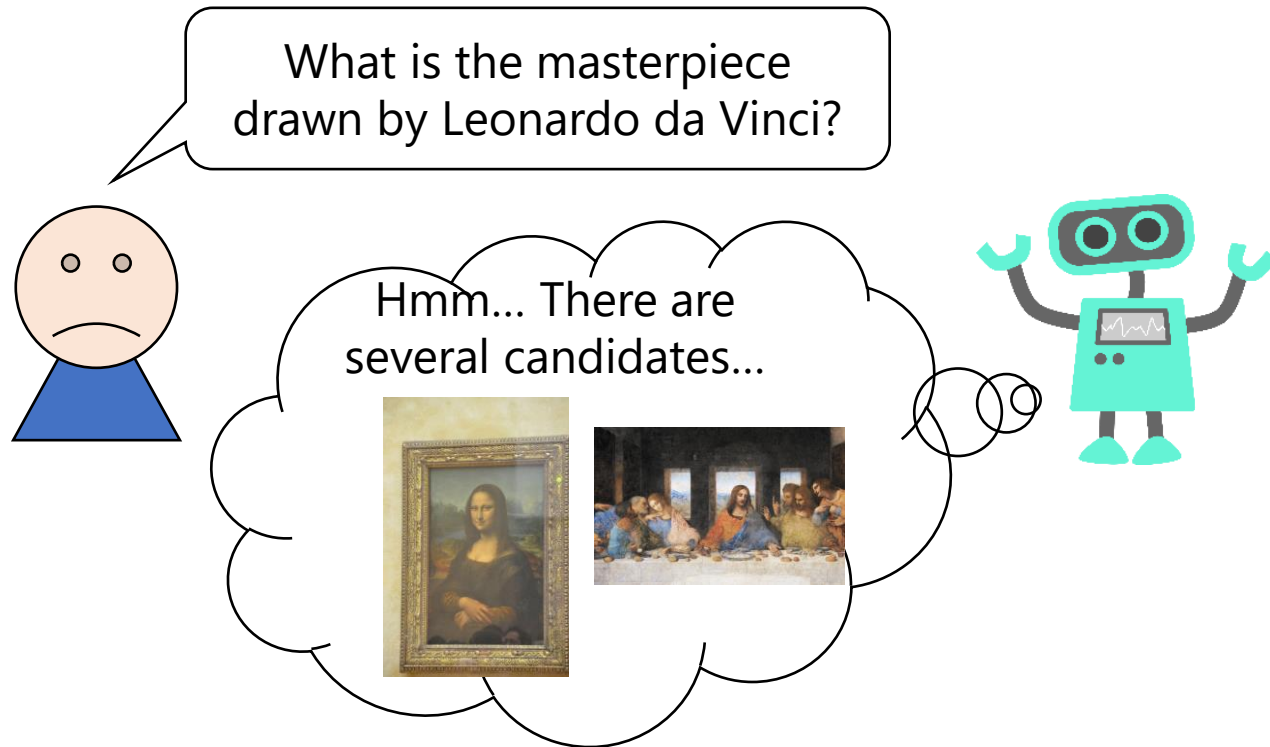
answer



Problem of existing QA systems

◆ User questions are often ambiguous

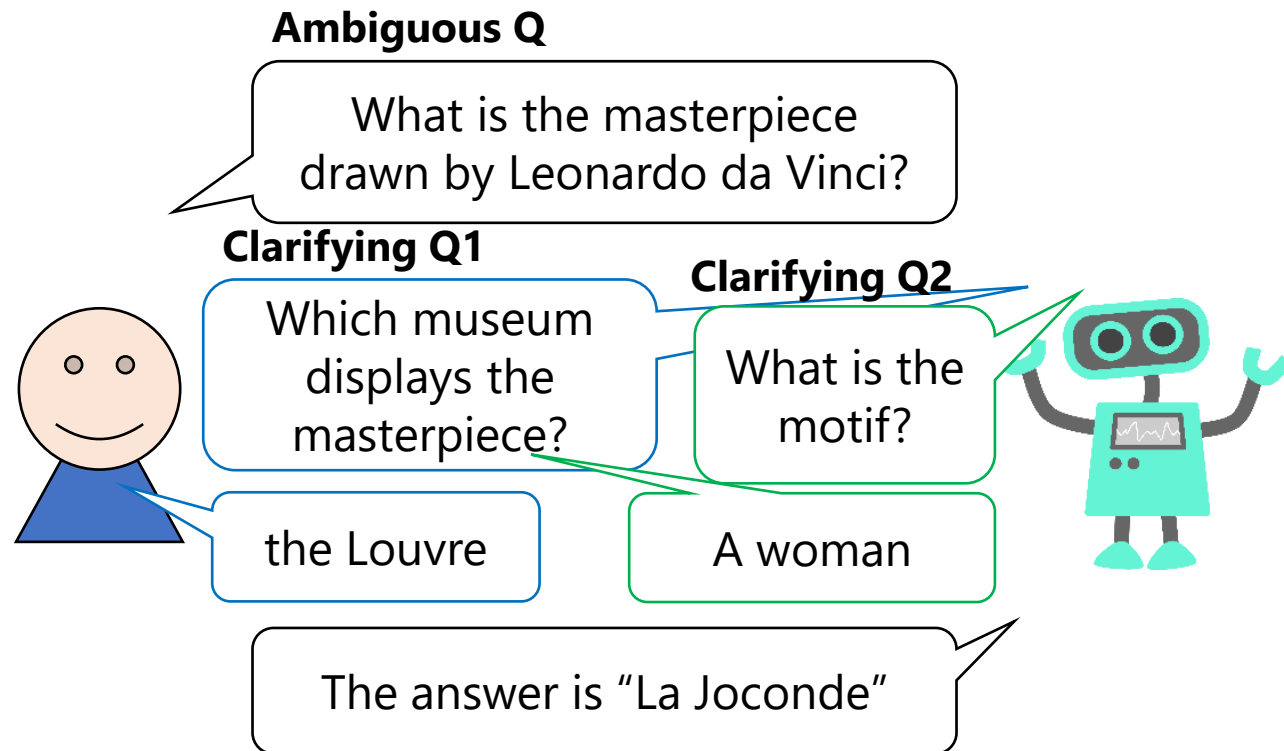
- It is hard to predict one answer



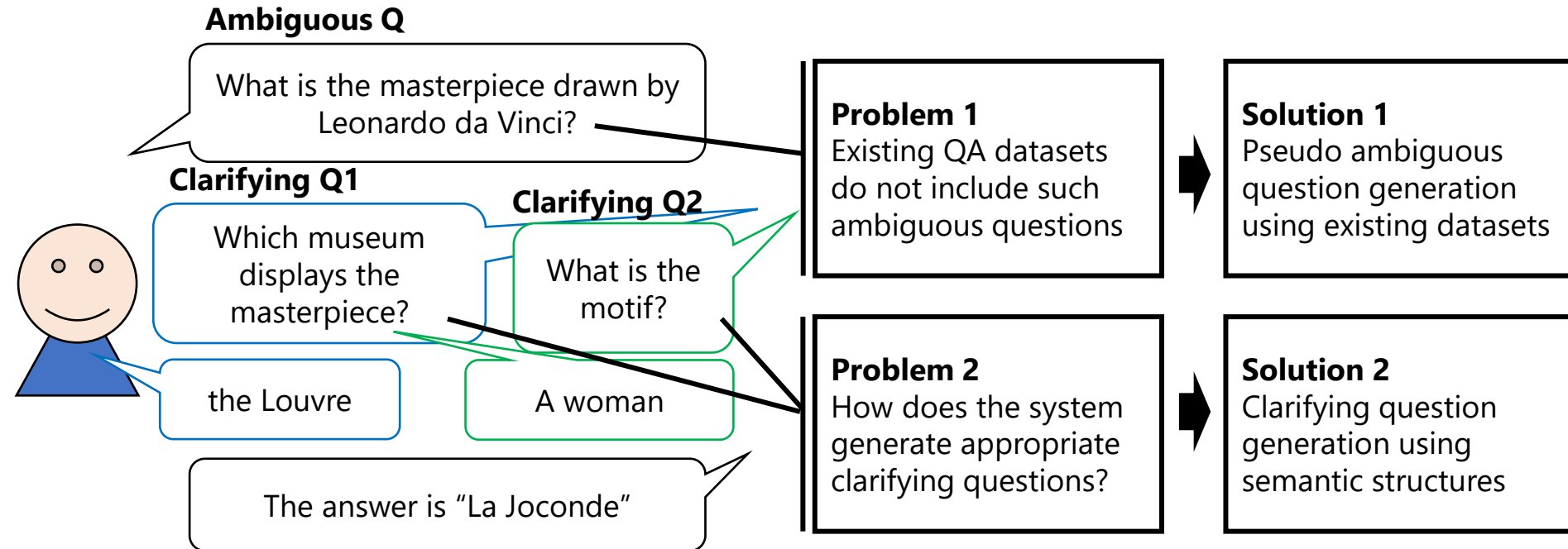
Solution: clarifying QA

◆ The system can ask the user clarification questions

- Additional information is extracted to find the answer



Paper focus



◆ This paper focus on pseudo ambiguous question generation and clarifying question generation



Related works

- ◆ Rephrasing ambiguous question [Otsuka19]
- ◆ Clarification using target document [Rosenbaum20]



Pseudo ambiguous question to extend existing dataset

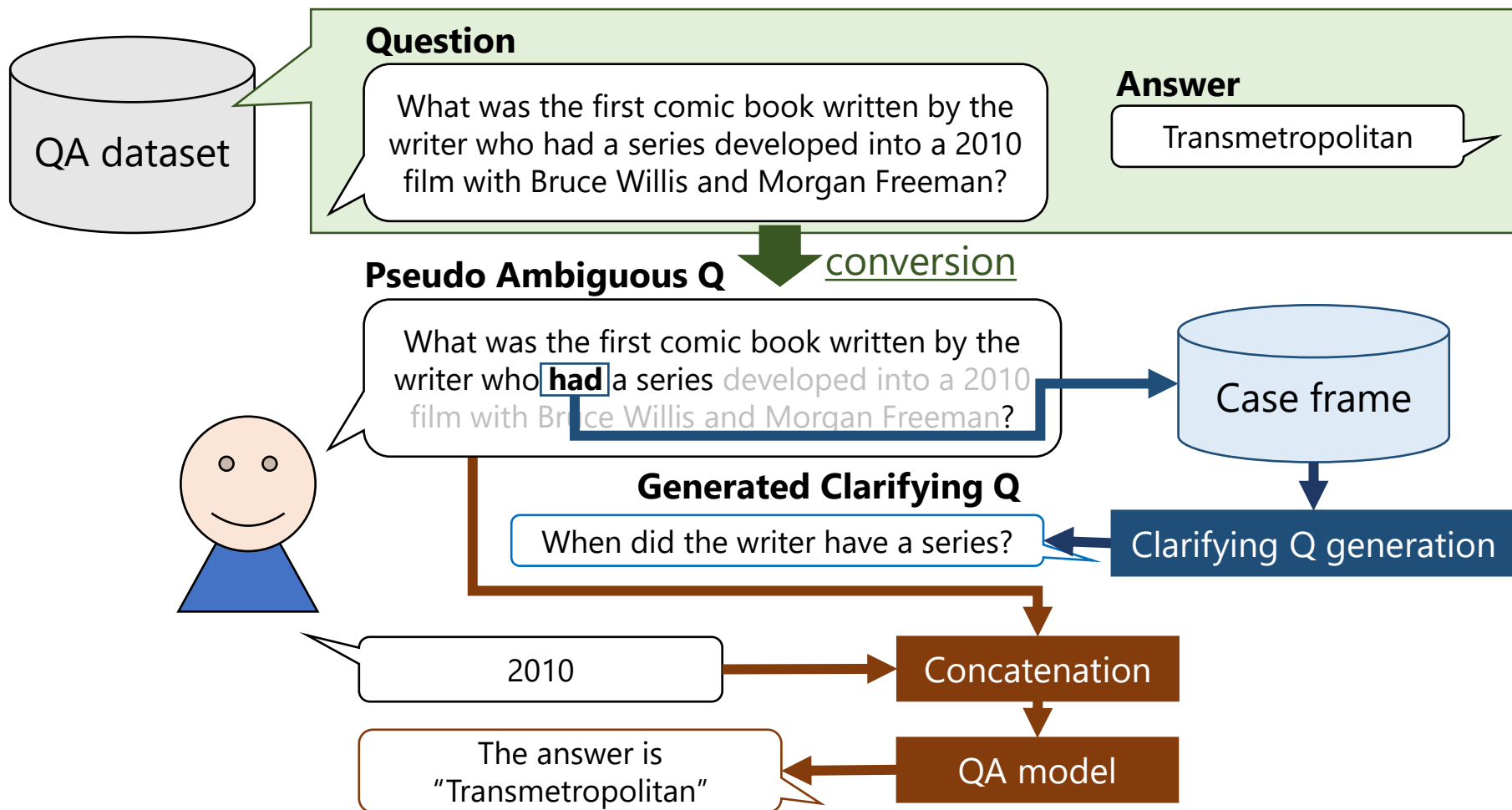
- ◆ Question generation based on sentence structure [Furukawa18]



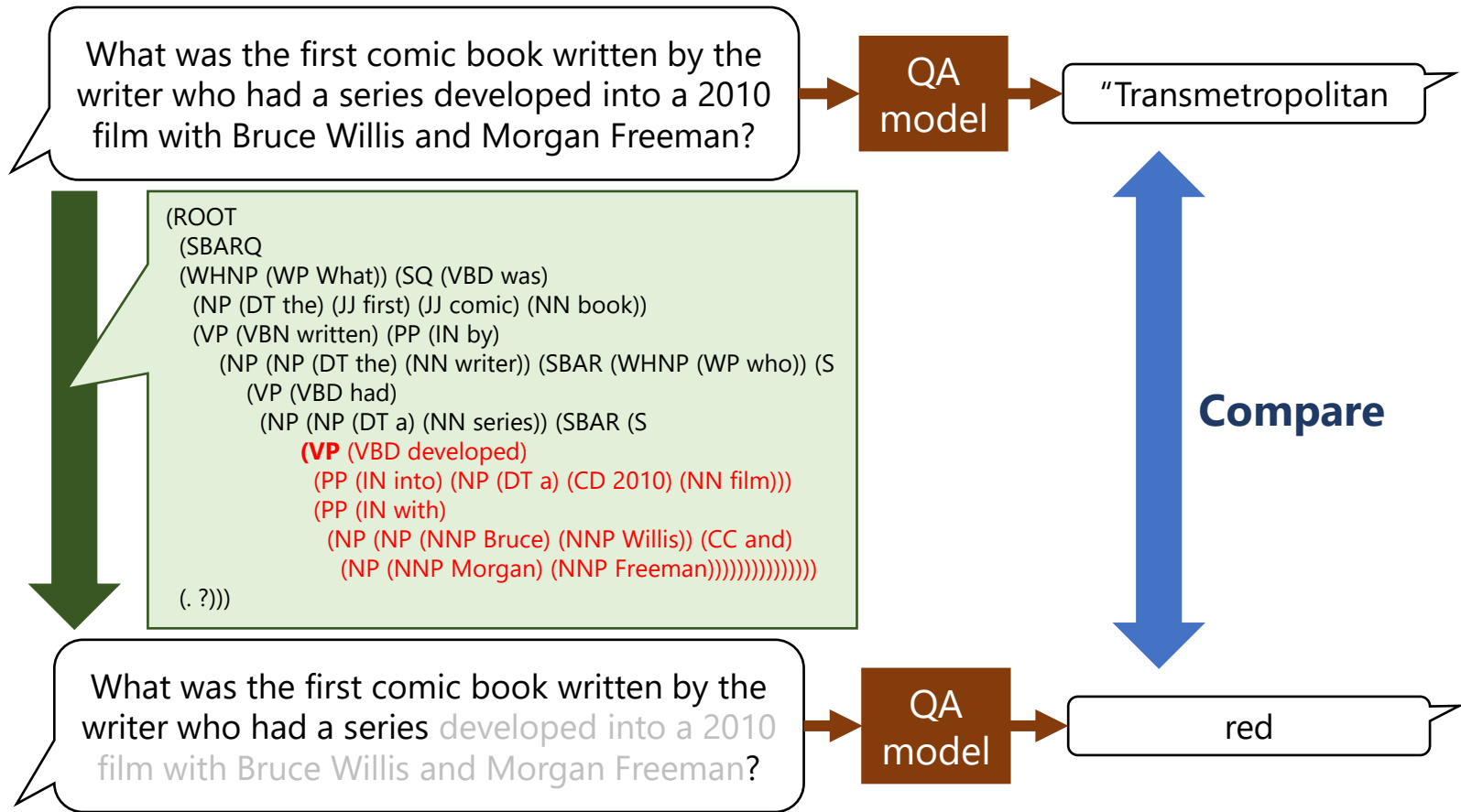
Predicting clarifying part based on sentence structure



System overview

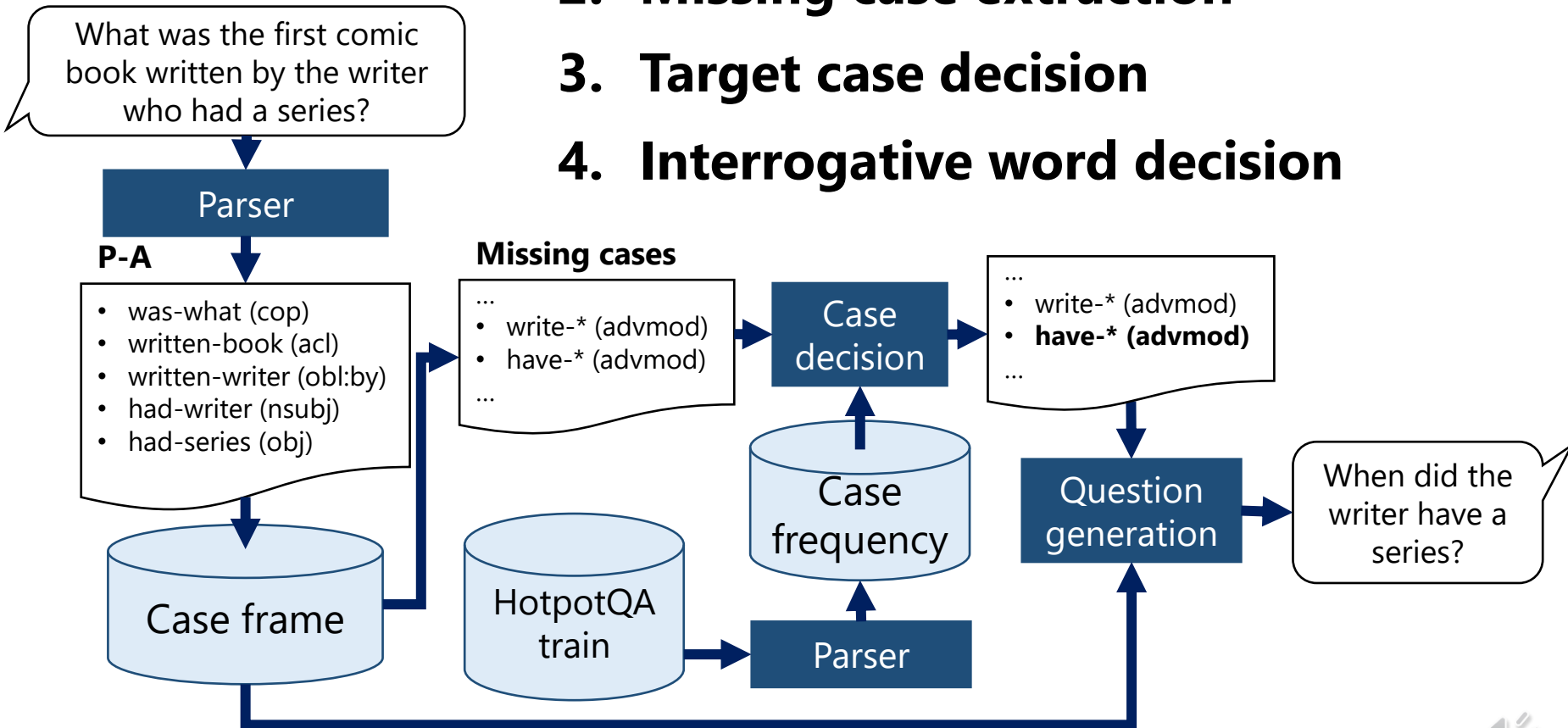


Pseudo ambiguous question

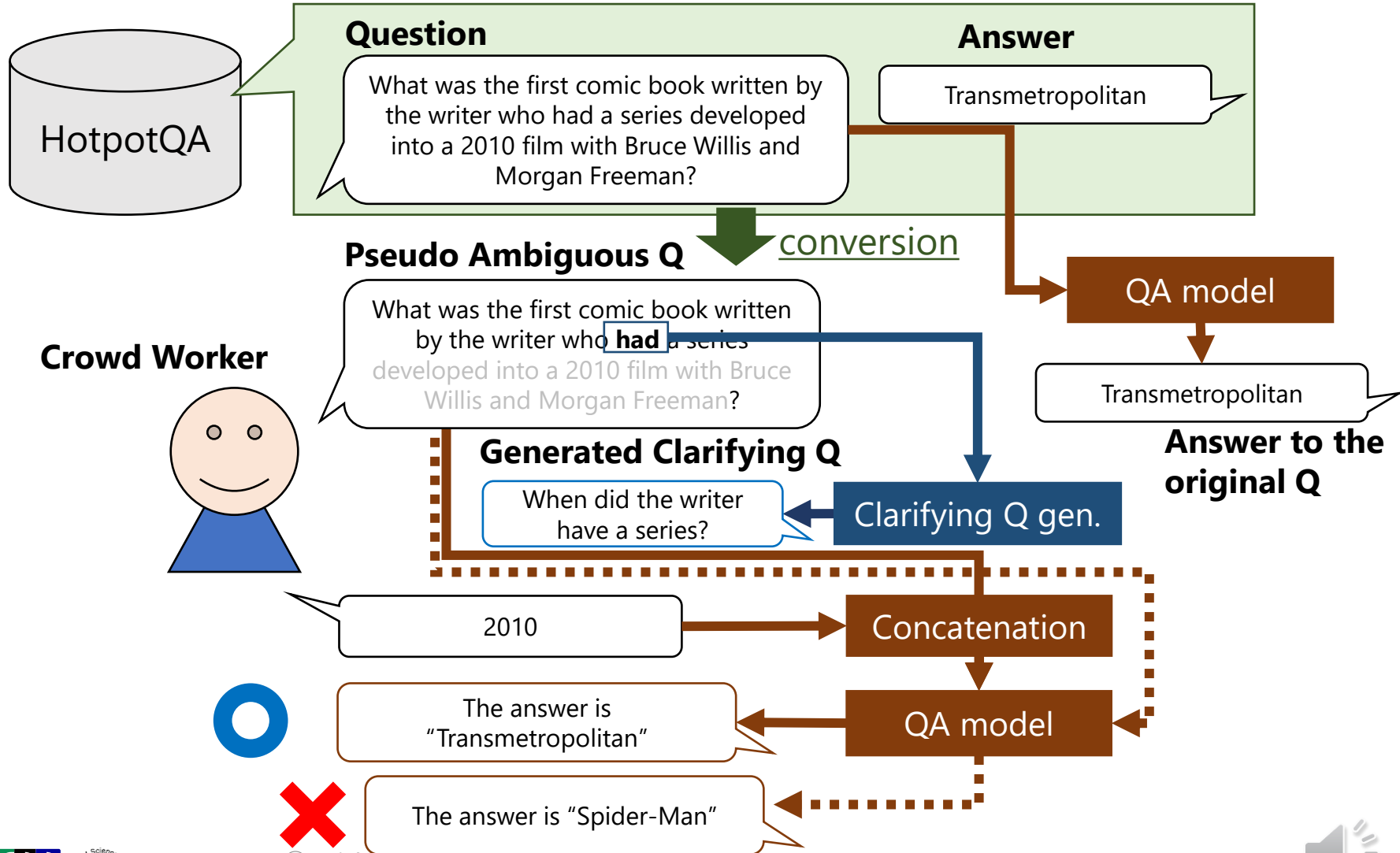


Clarifying question generation

1. Predicate identification
2. Missing case extraction
3. Target case decision
4. Interrogative word decision



Evaluation



Experimental setting

- ◆ **We prepared the eval data in the following procedure**
 - Prepare ambiguous questions by the pseudo ambiguous question generation. We selected samples by:
 - the eliminated words are less than 50% of the original Q
 - the eliminated words do not contain interrogative
 - QA model results are changed from correct to incorrect
- ◆ **Original dataset:**
 - HotpotQA
- ◆ **QA model used in the evaluation:**
 - BERT-based model
- ◆ **Answers to the generated clarifying Q:**
 - collected by crowdsourcing (AMT)



Results

Category	EM	F1	#q	#eval
w/o Clarify	0.0	0.0	850	4,250
Yes+No	49.52	57.28	850	4,250
Yes	50.21	57.82	486	2,430

◆ We also have asked crowd-workers to evaluate the clarifying questions itself

- Yes: appropriate, No: not appropriate

◆ Half of samples we prepared were recovered

- Clarifying questions can get additional information to improve the QA accuracy
- Crowd-workers often provide useful information even if the clarifying question is not appropriate



Examples

Original Question

What is the name of the executive producer of the film that has a score composed by Jerry Goldsmith?

Pseudo Ambiguous Q (A)

What is the name of the executive producer of the film that has a score composed?

Clarifying Q (C)

which composed?

Resp.

Jerry Goldsmith

Answer to A+R

Ronald Shusett

Answer to A

Jerry Goldsmith



Original Question

Lt Col. Stewart Francis Newcombe ~~was a British army officer~~ and associate of a military officer that was given what title?

Pseudo Ambiguous Q (A)

Lt Col. Stewart Francis Newcombe and associate of a military officer that was given what title?

Clarifying Q (C)

which was the Newcombe and associate given?

Resp.

a military officer

Answer to A+R

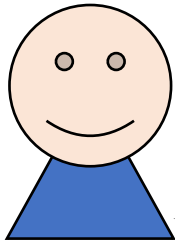
Lawrence of Arabia

Answer to A

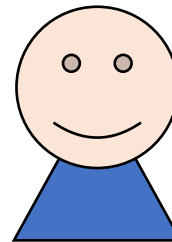
British archaeologist, military officer, diplomat, and writer



Worker



Worker



Summary

- ◆ **We built pseudo ambiguous question generation and clarifying question generation system toward clarifying question answering system**
 - Important to consider sentence structure in generation
 - Predicate (event)-oriented NLP is important to generate clarification questions
 - Even if the clarification question generation failed, the clarification will be a help to input additional information by users
- ◆ **We have to build better algorithms to use the pseudo data**
- ◆ **We also have to evaluate the system with real ambiguous questions by users**

