

Active Reading Comprehension: A dataset for learning the Question-Answer Relationship strategy

Diana Galvan¹ (dianags@ecei.tohoku.ac.jp)

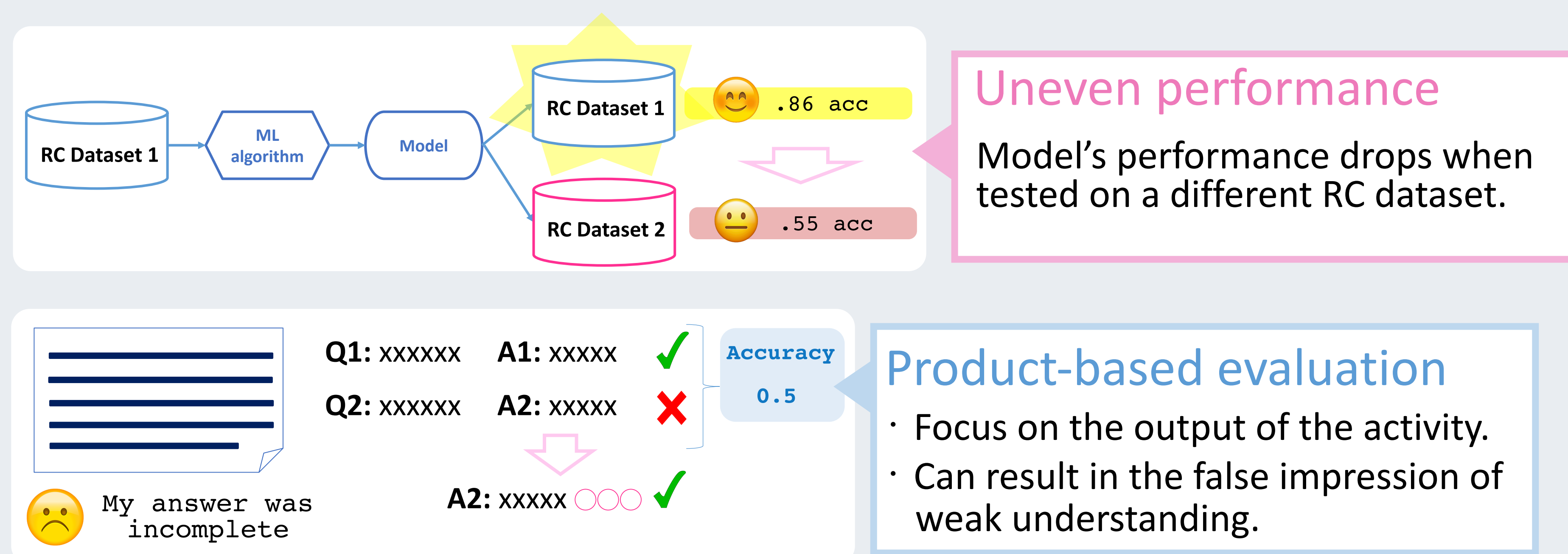
¹Graduate School of Information Sciences, Tohoku University

Overview

- ✓ **Goal:** Create a general reading comprehension model.
- ✓ **How:** Importing the QAR strategy from the educational literature to the NLP domain.
- ✓ **Contribution:** Novel evaluation method.
- ✓ **Plan:** Rely on previous RC work to create a new dataset.

Background

Current Reading Comprehension (RC) setting



Research proposal

Key Idea: Evaluate the overall comprehension process

Interviewer: Mrs. Choy, would you like to tell us something about your background before the Japanese invasion?

Elizabeth Choy: 1. Oh, it will go back quite a long way, you know, because I came to Singapore in December 1929 for higher education. 2. I was born in North Borneo which is Sabah now. 3. My ancestors were from China. 4. They went to Hong Kong, and from Hong Kong, they came to Malaysia. 5. They started plantations, coconut plantations, rubber plantations. 6. My parents and grandparents were more advanced for their times and when they could get on a bit, **they wanted their daughters to be educated** too. 7. So, **we were sent to schools away from home**. 8. First, we went to Jesselton which is Kota Kinabalu now. 9. There was a girls' school run by English missionaries.

Word match (question & text)

Q2: What were the consequences of Elizabeth's parents and grandparents being "more advanced for their times"?

A2: They wanted their daughters to be educated

so they sent them to schools away from home.

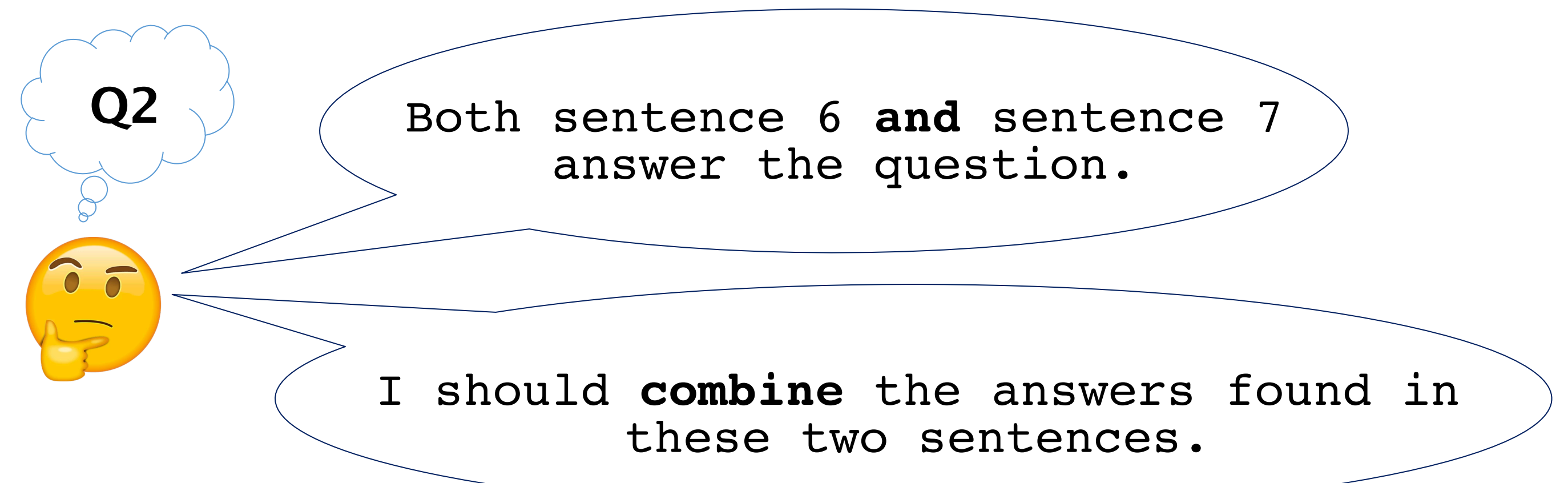
Answer location (sentence 6)

Missing piece of information

Question-answer relationship strategy

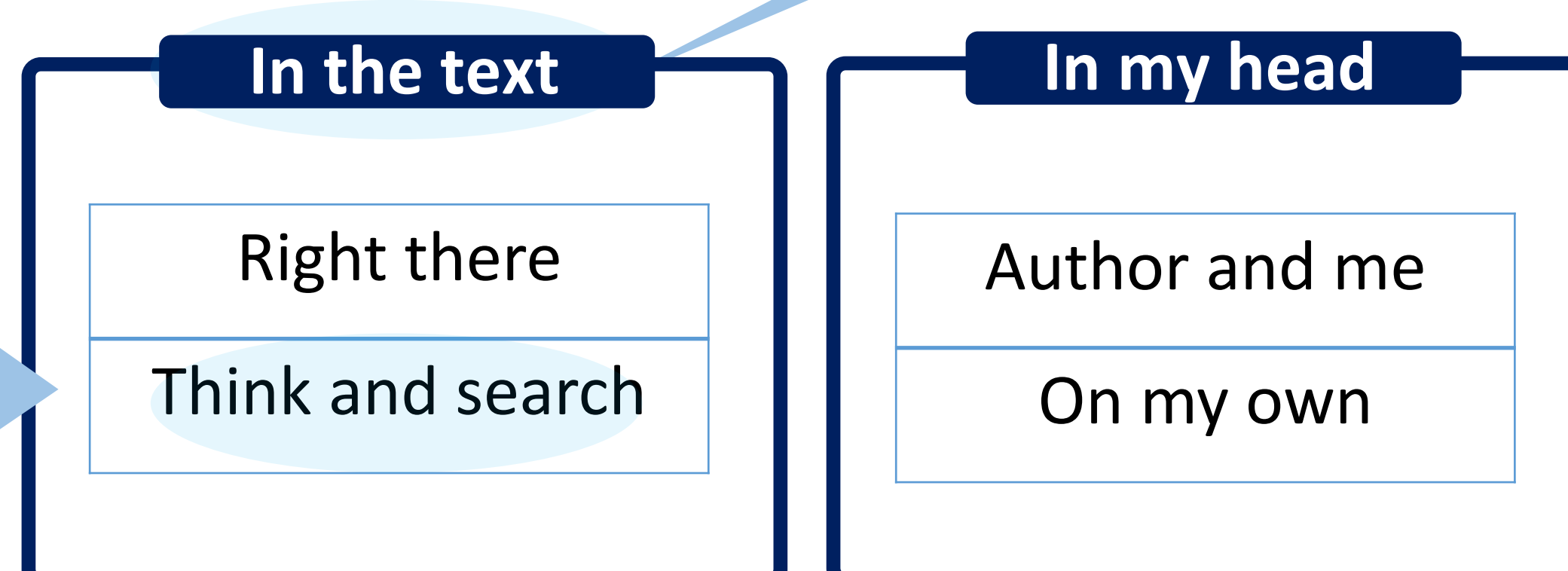
- ◆ Proposed by Raphael (1982) as a way of improving children reading performance **across grades** and **subject areas**.
- ◆ Emphasizes the relationship between the **text**, **questions** and the reader's **prior knowledge**.

Step 2 Identify the information source



Step 3 Identify the type of reasoning

Step 1 Question classification



Research plan

1 Create a new dataset for a QAR classification task

Sentence	Answer	Sentence Needed	Sentence	Answer	QAR category
1 Mary moved to the bathroom			1 Mary moved to the bathroom		
2 John went to the kitchen			2 John went to the kitchen		
3 Mary took the football there			3 Mary took the football there		
4 How many objects is Mary carrying?	one	3	4 How many objects is Mary carrying?	one	1
5 John went back to the kitchen			5 John went back to the kitchen		
6 Mary dropped the ball			6 Mary dropped the ball		
7 How many objects is Mary carrying?	none	3, 6	7 How many objects is Mary carrying?	none	2

New annotations

$$QARcategory = \begin{cases} 1, & \text{for } n = 1 \\ 2, & \text{for } n > 1 \end{cases}$$

Where n is the number of sentences

Task: Classify RC questions into one of the QAR categories.

2 Evaluate state-of-the-art RC models on the QAR classification task

- ❑ Knowledgeable reader (Mihaylov and Frank, 2018)
- ❑ Knowledge-enhanced reader (Mihaylov et al., 2018)
- ❑ TriAN (Wang et al., 2018)
- ...

