



# SML Question-Answering System for World History

## Essay Exams at NTCIR-13 QA Lab-3

Yusuke Doi, Takuma Takada, Takuya Matsuzaki, Satoshi Sato  
(Graduate School of Engineering, Nagoya University)

### Summary

We developed a question-answering system for the world history **short essay-type question**. The test run results was not very satisfactory. We analyzed the test run results.

### Short-essay Type Question

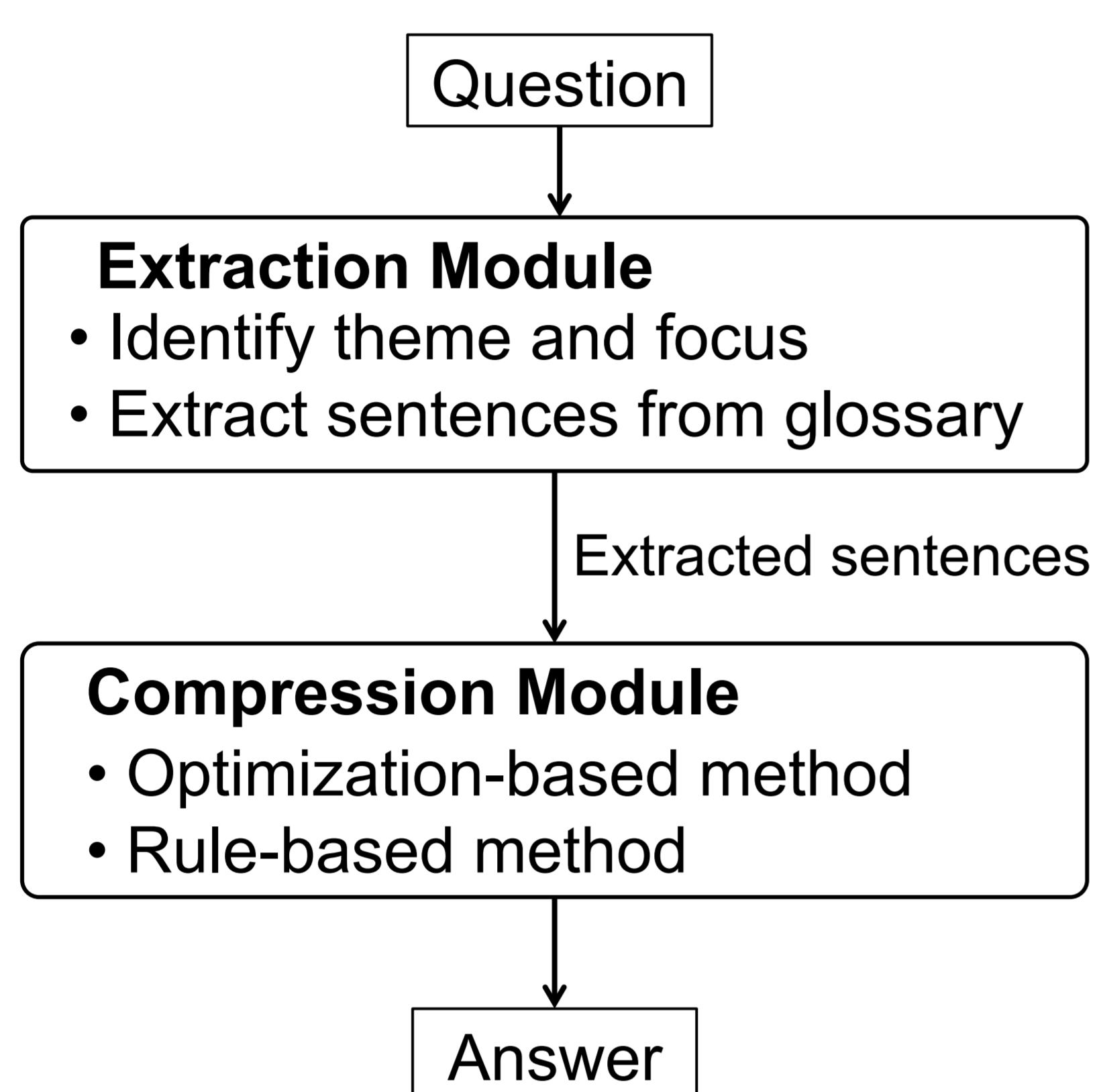
#### Question

ポリスの形成過程を、60字以内で説明しなさい。  
Describe, in no more than 30 English words,  
the process by which the polis were formed.

#### Model Answer

各地で有力貴族の指導下で、集落が連合し、アクロポリスを中心として人々が集住する形でポリスを形成した。  
Under the leadership of powerful nobles, various settlements formed coalitions, and people lived together around the Acropolis, forming poleis.

### Short-essay QA System

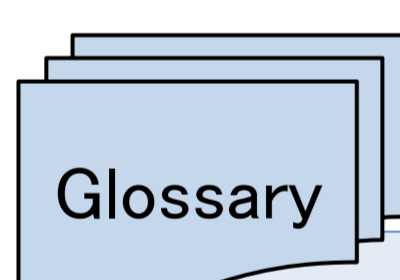


#### Extraction Module

Q. Describe, in no more than 30 English words, the **content** of the **Monroe Doctrine**.

Identify the **theme** and **focus** of question

**Theme : Monroe Doctrine**    **Focus : Content**



#### Monroe Doctrine

- Definition** The Monroe Doctrine was a United States ...
- Content** It stated that further efforts by European ... → Extract
- Content** At the same time, the doctrine noted that ... → Extract

### Compression Module

#### Optimization-based Method [Morita et al. 2011, 2013]

Repeatedly add valid subtrees to answer so that gain of objective function is maximized

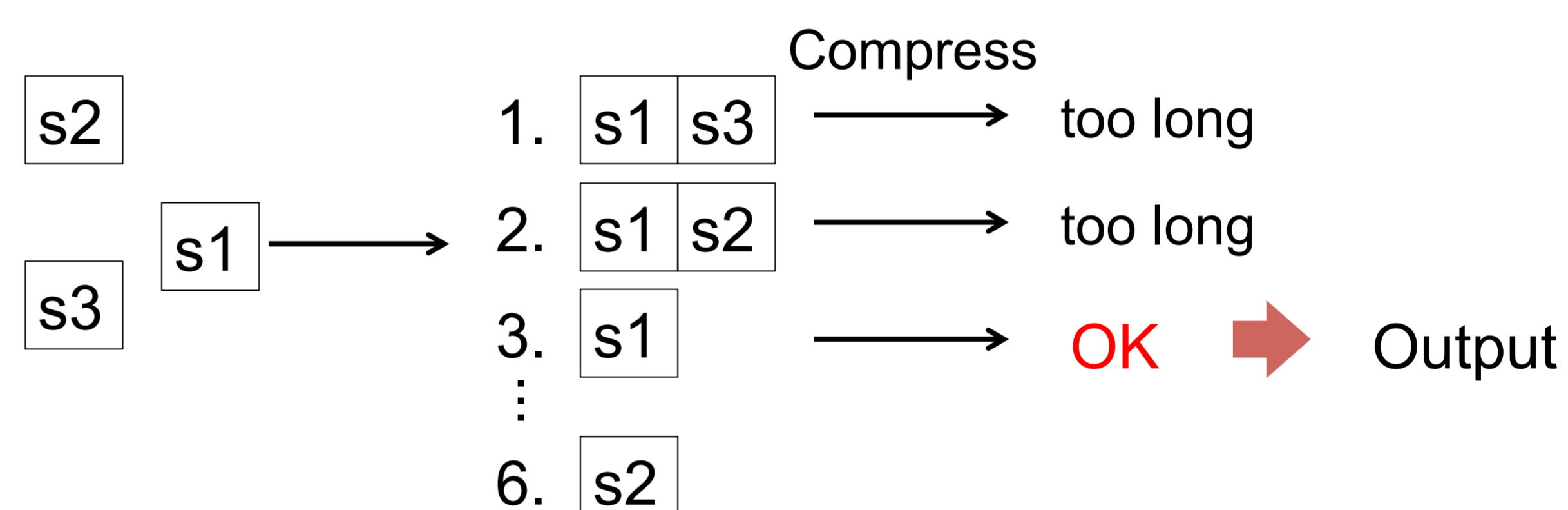
Objective function

$$f(S) = \sum_{w \in \text{words}(S)} \left( \text{qsb}(w) \sum_{i=0}^{\text{count}_S(w)-1} d^i \right) + \gamma \cdot \text{reward}(S)$$

qsb(w) : query relevance score

#### Rule-based Method

1. Extend extracted sentences by concatenating each pair
2. Sort sentences by QSB score
3. Compress sentence in turn by compression rules



#### Compression rules

1. Delete parenthesized phrases
2. Delete conjunctions at the beginning of the sentences
3. Delete adverbs
4. Delete time expressions
5. Delete adnominal phrases one by one from the beginning of the sentence

### Phase-2 Results

Run-1 Rule-based method

Run-2 Optimization-based method

Number of short essay questions is 22.

|       | Nuggets | ROUGE-1 | ROUGE-2 | ROUGE-3 |
|-------|---------|---------|---------|---------|
| Run-1 | 7 / 80  | 0.313   | 0.088   | 0.038   |
| Run-2 | 7 / 80  | 0.312   | 0.091   | 0.039   |

All team's ROUGE-2

1. Forst 0.107
2. SML 0.091
3. KSU 0.072
4. IMTKU 0.052

### Analysis

In **18/22** questions, the extracted sentences included none of the nuggets

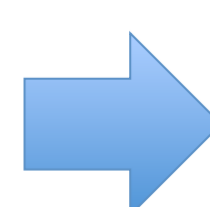


In most cases, the theme was wrongly identified

e.g.

Q. During the middle of the Former Han era **Confucianism**, which up until that point had been merely one of several valid schools of thought, was given a special position of prominence, separate from other schools of thought. Explain, in 30 English words or less, what event led to this.

“Confucianism” was not extracted as the **theme** because it didn't match the rule for identifying the theme



There was no significant difference between the performances of two compression methods