# Forst: A Challnge to the NTCIR-15 QA Lab-PoliInfo-2 Task

Hiromu Onogi†1, Kiichi Kondo†1, Younghun Lim†1, Xinnan Shen†1, Madoka Ishioroshi†2, Hideyuki Shibuki†2, Tatsunori Mori†1, Noriko Kando†2†3

†1 Yokohama National University, †2 Nationl Institute of Information, †3 SOKENDAI

## Stance Classification Task

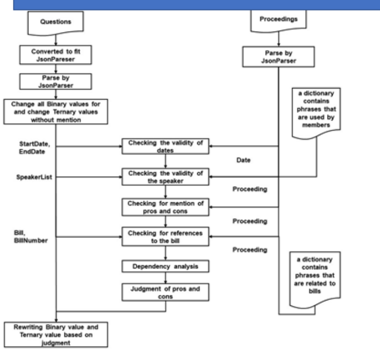
### Approach

We applied a rule-based approach using four rules on the proceedings and questions, and then conducted a dependency analysis on the sentences that satisfied the conditions, and obtained the output on the pros and cons. Because we thought that many of the pros and cons in the remarks could be **identified by a relatively small number of fixed expressions**.

### Related Work

Inoue et al. proposed a method for extracting the favorable and unfavorable opinions from the Web to specific topics such as products and current affairs. Nishimura et al. proposed a method to extract and organize the opinions of specific people from newspaper articles.

### Method




A dictionary that contains phrases that are used by members of each parliamentary group to express their pros and cons to the bill is comstituted of the words such as "都議会自由民主党を代表し", and "都議会生活者ネットワークを代表し", and e.t.c.

A dictionary that contains phrases that are related to bills but do not refer to a specific example is constituted of the words, "全ての議案(all bills)", "すべての議案(all bills)", "全議案(all bills)", "他の議案(other bills)" and "ほかの議案(other bills)".

### Result

| ID | binary result | ternary result |
|---|---|---|
| Forst(ID164) | 0.9382 | 0.144 |
| Forst(ID171) | 0.9388 | 0.852 |
| Forst(ID232) | 0.9391 | - |
| Forst(ID234) | 0.9408 | - |

### Disscussion

We want to check if our system can adapt to other minutes. **We want to create a system that can solve anaphora in the future.** We would also like to consider methods based on machine learning.

## Entity Linking Task

### Abstract

EL model conducts two tasks

1. **Mention detection**: mentions are extracted as spans in a text input.
2. **Disambiguation**: link spans to the corresponding pages on Wikipedia.

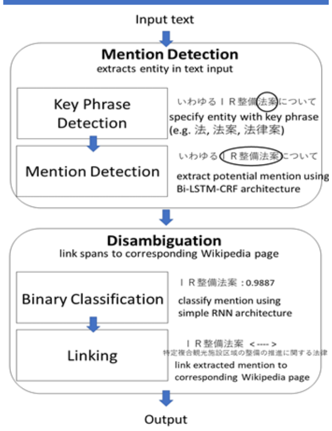
In addition to a typical EL model, It conducts two additional tasks.

1. **Key Phrase Detection** task is to capture a specific phrase in order to specify the entity.
2. **Binary Classification** task is to classify the non-mention entity and entity.

### Approach

The main goal of EL is to extract entities from minutes of the Tokyo Metropolitan Assembly. The category of the mention is restricted only to legislation. Such domain-specific conditions make EL tasks much more difficult. We focused not on the proposal of a new competitive universal EL model but <u>a model for domain-specific conditions</u>.

### Methods




### Result

The F-score of the submitted result is as below.

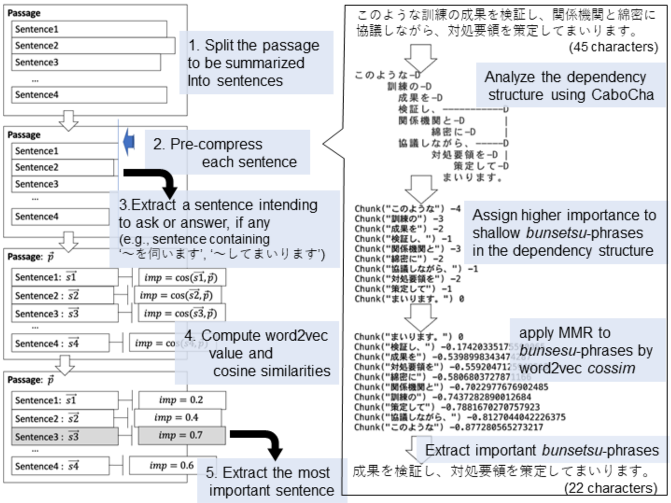
| ID | Method | F-score |
|---|---|---|
| 269 | LSTM-CRF(MD) + Binary Classification | 0.3912 |
| 217 | Rule-based(MD) + Binary Classification | 0.3910 |
| 147 | Rule-based(MD) | 0.3389 |

### Conclusion

We applied two different neural network architectures to the model and trained with only 2k entities. More annotated entity data would improve its accuracy. This model works effectively on the domain-specific condition to capture the entity. However, linking task needs further research.

## Dialog Summarization Task A

We assume that the representative sentence of a passage is similar to the whole passage and so our system extracts such a sentence as a summary. We used cosine similarity of word2vec value. The candidate extraction sentences are pre-compressed not to exceed the character limit. **Our pre-compression method regards depth of *bunsetsu*-phrase in the dependency structure as the basic importance and applies MMR.**




### Result

| Modification | ROUGE |
|---|---|
| method described above | 0.2410 |
| extracting sentences without pre-compression | 0.2275 |
| not applying MMR | 0.2453 |
| not prioritizing sentences intending to ask or answer | 0.1430 |
| (average of all participants) | 0.1850 |

### Discussion

- We have found that simply using the cosine similarity as the importance does not make sentences intending to ask or answer important.
- Sentence extraction with pre-compression improved ROUGE score but applying MMR didn't so.
- However, applying MMR seemed to have made the meaning of the summary easier to understand, and so we will continue to carefully consider whether introducting this.

## Dialog Summarization Task B

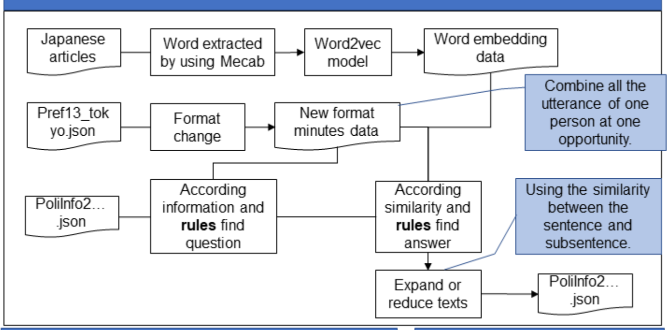
### Related Work

Le et al. proposed that the vector of sentences can represent the distributed expressions. Kimura et al described the tasks of PoliInfo2 and shows the data constructions of the minutes. Kazuki et al. proposed that we can use clue expressions help us find the questions and answers. For instance, **"伺います"** is the clue expression of a question, and **"思っております"** is the clue expression of an answer.

### Approach

According to the information provided by the original answer sheet file, we have **concluded some rules** to find the questions, and then also narrow the range of candidate answer sentences with some rules. **Using the similarity between the questions, "Subtopic" and candidate answers**, we can find the most suitable pairs.

### Methods




### Result

| ID(Forest) | Modification | ROUGE |
|---|---|---|
| 247 | Add a second sentence into the question candidate list | 0.1471 |
| 235 | Increase the importance of "Subtopic" words | 0.1384 |
| 231 | TF-idf | 0.1219 |

### Discussion

Our method can effectively work on a **one-topic-one-question type**, but if the topic needs to be answered from multiple aspects, it cannot extract all questions. One improvement is we may segment the parts of answer more accurately, and then determine whether each sentence can be followed by the question by a machine learning way.