ABSTRACT

- English Term Question Subtask
- RNN-based Approach
- Explicit Summarization to Reduce Computational Cost
- Discussion About the Effects of Explicit Summarization

CHALLENGES OF THE SUBTASK

- Various format of questions
  - What/Where/Who… for the grand question
  - No What/Where/Who… for the grand question
  - Not for the grand question but another story

Grand Question: During the first half of the 20th century, the principles of world peace and a new ... (B792W10 The University of Tokyo, 2001)
Instruction: What is the treaty referred to in the underlined section (2)? Write the name of the treaty.

Grand Question: Looking back at the history of Western Asia ... (B792W10 The University of Tokyo, 2001)
Instruction: With regard to underlined section (4), from the end of the 19th century there was a political movement among Jews in Europe to return to Palestine. Write the name of this movement.

VALIDATION EXPERIMENTS

Data: Question Sets of Phase 1 and Phase 2

Metrics:
Correct = the system answer contains any word in the gold standard

CONCLUSIONS

Main Issues:
- Wrongly high-rank less important sentences for question answering
- The answer need to be exists exactly the same in the documents (the bottleneck of the RNN of us)

Secondary Issues:
- Lack of specified knowledge base (another possible reason why the results in Japanese are generally better)
- Lack of training data (we have to use less related external data for training)

FUTURE WORKS

- Wrongly high-ranked the meta information
- Revision of the Ranking Function (e.g. Adding Entity Recognition)
- Careful preprocessing
- Knowledge Resource
- Textbook datasets in the future if possible
- More Powerful Machines