Question Answering and Dialogue over Structured Knowledge using a Neural Network Model with External Memories

MURAYAMA Yuri

Various approaches to connect natural language with knowledge have been studied. As the common problem for their studies, they aim to convert natural language into query language in order to access from natural language expression to knowledge represented in the formal form such as Linked Open Data. Based on the approaches of prior studies, in this study, we propose a method that generates SPARQL queries by iteratively referencing to knowledge compiled as Linked Open Data. In the experiment, our method was able to generate correct SPARQL queries for each question type.

Nowadays Transformer-based studies are commonplace in the natural language processing, yet there are some problems. The Differentiable Neural Computer (DNC) solved one of the problems that Transformer cannot store data over a long time. Another problem is that a general-purpose language model such as Transformer is able to solve various natural language processing tasks by pre-training the large scale text data, whereas the ability such as knowledge using and arithmetic processing is supposed to be obtained implicitly from training vast amount of corpora and the correctness is not guaranteed. Therefore, our research question is that we aim to improve the performance on natural language processing tasks requiring knowledge and the operation by incorporating architectures for knowledge and arithmetic processing, and Transformer into DNC. Our objective is to build a model incorporating Transformer and architectures for knowledge and arithmetic processing into DNC, a neural network modeling of the Turing machine which is the principle of the computer, and verify whether the model can deal with natural language processing tasks requiring knowledge and the operation.

Before building the proposed model, we conducted a preliminary study applying DNC to Japanese dialogue. There are many works on dialogue systems considering context using the neural network. In our work, to realize a robot cafeteria, we applied DNC to the process of ordering dialogue. We created a Japanese ordering dialogue dataset and conducted an experiment on it. As the result, we obtained the low mean test error rate.

It is essential for dialogue to understand context and use knowledge. However, researches dealing with these issues have not progressed much yet. In addition, nowadays sequence-to-sequence models such as Seq2Seq, T5 and BART are commonplace for dialogue architectures. On the other hand, it is necessary for more natural and intellectual dialogue to understand context and use knowledge. However, scientists have argued that such models are limited in their ability to store data over a long time. To retain the long-term information, neural network models with an external memory such as DNC have been proposed. These models learned to process more complex information compared to standard ones by adding the memory, and achieved a highly accurate performance in dialogue considering context. In our work, we enrich the DNC architecture and propose a method using both context and structured knowledge by the distributed representation. We conducted an experiment on the CSQA dataset composed of a series of coherently linked question answering that require a large scale knowledge graph. Even though the overall test error rate of our proposed method was lower than that of DNC, our method performed better at five out of ten items. In an experiment with the Dialog bAbI dataset, our improved DNC, rsDNC and DNC-MD outperformed their original models. In particular, each model obtained an improvement of approximately 14%, 20% and 7% respectively on a task requiring knowledge. In the Movie Dialog dataset, our improved rsDNC and DNC-DMS also yield better performance than their original models.

DNC, a neural network model with an addressable external memory, can solve algorithmic and question answering tasks. As improved versions of DNC, rsDNC and DNC-DMS have been proposed. However, how to integrate structured knowledge and calculation into these DNC models remains a challenging research question. We incorporate architectures for knowledge and calculation into such DNC models, i.e. DNC, rsDNC and DNC-DMS, to improve the ability to generate correct answers for questions with multi-hop reasoning and calculation over structured knowledge. Our improved rsDNC model achieved the best performance for the mean top-1 accuracy and our improved DNC-DMS model scored the highest for top-10 accuracy in GEO dataset. In addition, our model improving rsDNC outperformed the other models with the mean top-1 accuracy and the mean top-10 accuracy in augmented GEO dataset.

Keywords: Natural Language Processing, Linked Open Data, Deep Learning