Natural Language Generation Based on Sentence Structural Knowledge

KUMAGAI Kaori

Humans create sentences to convey information in various situations on daily life. In our daily life, when we cook or clean with other people, we write down the contents of the work we want to request and tell them to the other person. In the work scene, in order to communicate the complicated procedures to multiple people, we create sentences explaining each work procedure. Creating a manual consisting of sentences for each procedure, we can accurately convey complex procedures to many people. The ability of humans to generate sentences is an indispensable ability to live smoothly with other people and to proceed with work.

By equipping machines with the ability to generate sentences like humans, we can greatly improve the efficiency of our lives and work. For example, by automatically recording the procedures of daily work such as cooking and cleaning, the time required to tell information to the other people is possible to shorten. In the case of work involving complicated procedures, by taking videos of the work in advance and automatically documenting it, the time of creating a manual can be saved.

In addition, by automatically recording the details of the procedure done unconsciously, you can convey more accurately. Furthermore, the machine automatically generates sentences explaining the surrounding environment for the visually impaired person. The ability can make the lives of visually impaired people more peace of mind. The ability to automatically generate sentences that explain what people want to convey can make our lives and work more efficient and can expand human capabilities.

In this paper, we use the structural rules of sentences as external knowledge and examine the method of generating sentences according to the structural rules of sentences. Sentences with diverse words and phrases can be generated by allowing diversity of words and phrases within the scope of sentence structure rules. As structural rules of sentences, two types of knowledge, grammatical structural rules and semantic structural rules, are used, and the effectiveness of each rule for sentence generation is examined. We adopt a context-free grammar as the grammatical structural rules of sentences, and propose a method to generate a parse tree constructed by applying the context-free grammar. In order to efficiently search the appropriate parse tree, we employ a Monte Carlo tree search algorithm. We adopt case structure labels as semantic rules for sentences. Based on a set of case structure labels, we propose a neural network that predicts the words corresponding to each given case structure label and at the same time predicts the order of the case structure labels so as to have an appropriate semantic structure. Experiments verify the effectiveness of the proposed method. From our experimental results, we show the effectiveness of grammatical structural rules and semantic structural rules for sentence generation methods.