

## 外 国 語 要 旨

学位論文題目： Building Theory and Corpus for Discourse Relation Recognition for Japanese Text

氏 名： Kimi Kaneko

One of the fundamental goals of natural language processing is to develop a computational system to understand the meaning of texts within contexts. Such meanings involve information related to the structure between sentences, which is not derivable by sentential semantics alone; it is necessary to analyze the semantic relations between sentences and clauses. Such semantic relations have been called discourse relations, and studied in the research field known as discourse theory in both natural language processing and theoretical linguistics. For example, consider the following pairs of sentences:

- (1) a. Hanako fell into a hole.  
    b. Taro helped her.
- (2) a. Hanako fell into a hole.  
    b. Taro pushed her.

The temporal relations between the two events in (1) and two events in (2) are different. In (1), the temporal order with which the events "fell" and "helped" occurred follows the surface order of sentences, whereas in (2), the temporal order with which the events "fell" and "pushed" occurred is the reversed order of the sentences: we can interpret (2) as Taro's pushing Hanako was the cause of Hanako's falling. If the results of semantic parsing of discourse are obtained by composing the results of semantic parsing of sentences, there is no way to distinguish the temporal relations in (1) and (2). Thus, discourse relation recognition plays an important role in capturing the spacio-temporal relations between events within texts, and the faculty for discourse relation recognition is an essential module for natural language understanding.

Sometimes the discourse relations are determined by temporal relations between events. Consider the following sentences:

- (3) a. I opened the window.  
    b. The whole garden was covered with snow.
- (4) a. I opened the window.  
    b. I looked out at the whole garden covered with snow.

The temporal relations between the two sentences in (3) and (4) are also different. A natural interpretation of (3) is that the event "opened" happened where the state of the garden's being "covered with snow" holds, whereas an interpretation of (4) is that the event "looked out at" happened after the event "opened" happened. Since discourse relations and temporal relations have such interactions, we should take it into consideration of the definition of discourse relations. However, dependencies between

discourse relations and temporal relations are not fully analyzed in previous researches on discourse relations.

Based on the above observations, the aim of this study is, firstly, to sort out the dependencies of the intentions of speakers/writers on the discourse relations and secondly to establish a method for uniquely determining the discourse relation for each specific sentence/clause pair. In order to achieve this aim, I establish the decision procedure for judging discourse relations reflecting the dependencies between discourse relations and temporal relations. More concretely, I first analyze the dependencies between eventuality, modality, tense and aspect, temporal relation, and discourse relation, based on which I define the annotation schema and the decision procedure for discourse relations. We performed an annotation experiment following our annotation schema and the decision procedure. By calculating the inter-annotator agreement and conducting error analysis, it is verified whether we could sort out the relations between the intentions of speakers/writers and determine the unique judgment. Finally, I discuss which part of the procedure is left unclear as future work.