

Improving Name Recognition Tool in Japanese

Anh Nguyen

Students learning Japanese face a unique challenge because there are no spaces between words in Japanese sentences, which makes it more difficult to understand the grammar, and name entities may be mistaken for other kinds of vocabulary. I have developed Sentence Analyzer as a tool to help students quickly recognize name entities, new vocabulary that appear in Japanese texts and improve their comprehension of the grammatical structure of Japanese sentences. Utilizing Dynamic Programming approach, I obtain the initial morphological analysis of a Japanese text from *MeCab*, a part-of-speech analyzer in Japanese, and process it with a corpus of over 20,000 Japanese last names that I've built and an algorithm that detects name entities that aren't recognized in the initial result. The final analysis can be displayed in either Japanese or English, together with the meanings of each word in the original text translated into English to further help with students' learning process. The primary focus of this project is to enhance human's name entity recognition ability, but it can also be extended to work with other kinds of name entities or name entities in different languages.