

Machine Learning based NLP: Experiences and Supporting Tools

Yuji Matsumoto

Nara Institute of Science and Technology

Coprus-based approaches to natural language analysis that utilize recent sophisticated machine learning algorithms have now become to achieve very good performance. In this talk I will overview and categorize machine learning based natural language processing tasks and our experiences of using machine learning to various tasks such as segmentation, POS tagging, phrase and NE chunking, and syntactic parsing. I then discuss pros and cons of machine learning approaches and future issues. Finally, I will introduce an ongoing project of annotated corpus maintenance tools for developing consistent data for corpus and machine learning based NLP research.

