Sutee Sudprasert   2010: Design and Development of a Lattice Structure Dependency Parser for Under-Resourced Languages.   Doctor of Engineering (Computer Engineering), Major Field: Computer Engineering, Department of Computer Engineering.   Thesis Advisor: Associate Professor Asanee   Kawtrakul, D.Eng.   80 pages.

Dependency representations have become fashionable again in various Natural Language Processing areas, such as Machine Translation, Information Extraction, Text Summarization, and Ontology.   However, the development of a good dependency parser requires some resources such as training corpora or grammar rules and also morphosyntactic analysis tools as preprocessing tools. Unfortunately, many languages do not have a large training corpus nor reliable morphosyntactic analysis tools.

We present here a corpus-based approach for building a dependency parser especially for under-resourced languages. Dealing with unreliable morphosyntactic analysis tools, we propose a methodology for dependency parsing outputting all possible of morphosyntactic analysis by modifying the Eisner's algorithm ($O(n^3)$), a bottom-up dynamic programming chart parsing algorithm, that does not increase the time complexity order. For computing the parse score, we use Maximum Entropy Models and train the model with a small training corpus (716 sentences). Because the training corpus is very small, we also propose a method for adjusting the parse score by using a Dependency Insertion Grammar (DIG) induced from the corpus. The adjustment will be applied if invalid trees are produced by the statistical model. Moreover, the use of DIG can make it easier for us to observe language behavior and detect annotation errors through the induced DIG rather than looking into the corpus directly. We tested the system by using NAiST Thai Dependency Treebank as training data and the accuracy of the parsing results was 80% if sentences were word-segmented correctly and 85% if the sentences were also part-of-speech tagged.

Student's signature	Thesis Advisor's signature	__ / __ / __