

| | |
|-----------------------|---|
| Thesis Title | Thai Word Separator Using Hashing Dictionary |
| Student | Mr.Sorasak Thaitae |
| Thesis Advisor | Assistant Professor Dr.Boontee Kruatrachue |
| Level of Study | Master of Science program in Information Technology |
| Degree | Faculty of Information Technology |
| | King Mongkut's Institute of Technology Ladkrabang |
| Year | 1998 |

ABSTRACT

Thai alphabet are written adjacent, without any special alphabet between words. Therefore, it is difficult to separate each word in a line, or to detect any misspelling word. Separate words automatically will help both spelling correction and also document indexing. This research present search algorithm and the new Thai word separator using Hashing dictionary, and design the new separator algorithm using iterative instead of recursive with double stack data structure. We can also make word suggestion to the incorrect word with little modification to the dictionary. Hashing dictionary is also easier to add new words. The new algorithm and dictionary has been compared with softwares Word 6.0 and Word 7.0 of Microsoft in both the running time and correctness. The word separable result is correct as long as the words are in the dictionary. The separable speed has been improved by noticing the max number of character per word in the dictionary is 25. The dictionary also mark ended words so that the separator can stop before 25 characters when the ended words are detected. The speed can be improve with trie structure with trade of new words addition flexibility and word suggestion feature.