



ROYAL UNIVERSITY OF PHNOM PENH

Master of Science in Information
Technology Engineering

Software Development

DEVELOP ALGORITHMS AND SOFTWARE OF AUTOMATIC WORDS SEGMENTATION OF KHMER TEXT

Advisors: Dr. Srun Sovila, Ms. Rim Beanbonyka

Keywords: Text Semantic Analysis, Text Segmentation

Field related: Retrieval Information, Programming Language

Abstract

Differently from most of European sentences, the Khmer sentence is composed by words which are not separated by spaces. As consequences, in web search or document search, the word searching is found the difficulties of recognizing which words are composed and how many of them are appeared. Thus, in order to recognize them, the text/sentence is analyzed and segmented to words. The topic is to develop algorithms to segment the Khmer text and software for demonstration.