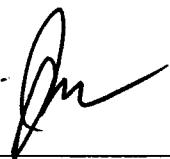


บทคัดย่อ

**T162625**

ปัญหาที่สำคัญในการประยุกต์ใช้คอมพิวเตอร์กับการประมวลผลภาษาธรรมชาติคือ การตัดพยางค์ ซึ่งปรากฏในภาษาของโลกตะวันตกหลายภาษารวมถึงภาษาไทยด้วย เมื่อพิจารณาวิธีและลักษณะการเขียนในภาษาโลกตะวันตกหลายภาษา เช่น ภาษาอังกฤษจะมีการเขียนเว้นวรรคระหว่างคำ ในขณะที่ภาษาไทยไม่มีลักษณะดังกล่าว ดังนั้น จึงมีหลายงานวิจัยได้เสนอวิธีการแก้ปัญหาเรื่องการตัดพยางค์ งานวิจัยที่ผ่านมามีอยู่ 2 วิธีหลักได้แก่ การตัดพยางค์โดยรูปแบบของหลักภาษาและการตัดพยางค์โดยฐานข้อมูล สำหรับวิธีการแรกจะดำเนินการตัดพยางค์เมื่อพยางค์นั้นมีการเขียนครบทั้งพยัญชนะ สระ และวรรณยุกต์ สำหรับวิธีการหลังพยางค์จะถูกตัดเมื่อพบพยางค์ที่ต้องการตัดในฐานข้อมูล ในงานวิจัยนี้จะใช้วิธีการตัดพยางค์โดยฐานข้อมูล อย่างไรก็ตาม ฐานข้อมูลของงานวิจัยนี้จะเก็บเฉพาะคู่พยัญชนะที่ไม่สามารถเขียนติดต่อกันได้ ซึ่งมีผลทำให้ขนาดของฐานข้อมูลมีขนาดเล็ก สำหรับทฤษฎีของงานวิจัยนี้ให้ประสิทธิภาพด้านความถูกต้องถึง 98.95 เปอร์เซ็นต์

(วิทยานิพนธ์มีจำนวนทั้งสิ้น 143 หน้า)



ประธานกรรมการที่ปรึกษาวิทยานิพนธ์

## Abstract

**TE162625**

One of the crucial problems to apply natural language processing to computer utilization is syllable segmentation, which is occurred by many eastern languages, including Thai one. Regarding method of writing, there is a space between words when it is written in English while there is none when it is written in Thai. Consequently, there have been several researches about how to solve Thai syllable segmentation problem. Based on such researches, there are two main theories. The first one is syllable segmentation based on pattern of writing and the other one is syllable segmentation based on dictionary. According to the first theory, the syllable segmentation will be done when consonant, vowel, and tonal mark are spelled perfectly. According to second theory, the syllable segmentation will be done when the word matches what is in the dictionary. Based on second theory, this research is presented a method that syllable segmentation will be done where certain strange combination is found. By this approach, the experimental results show that it can segment syllable correctly more than 98.95 %

(Total 143 Pages)



---

Chairperson