

ชื่อวิทยานิพนธ์ : การวิเคราะห์กระจายคำในประโยคภาษาไทยโดยการ
โปรแกรมเชิงเจนเนติก
ชื่อผู้เขียน : สมนึก สินธุปวน
ชื่อปริญญา : วิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์)
ปีการศึกษา : 2546

งานวิจัยในวิทยานิพนธ์นี้ มีวัตถุประสงค์ที่จะใช้การ โปรแกรมเชิงเจนเนติกในการตัดคำ และทำการวิเคราะห์กระจายคำในประโยคตามหลักไวยากรณ์ภาษาไทย

หลักการของการใช้โปรแกรมเชิงเจนเนติกในการวิเคราะห์กระจายคำคือ การสร้าง ประชากรต้นไม้วีเคราะห์กระจายคำที่เป็นไปได้ของประโยคแล้วทำการพัฒนาประชากรเหล่านี้โดย การข้ามพันธุ์ การกลายพันธุ์ และการคัดลอกพันธุ์ เพื่อให้ได้ประชากรต้นไม้วีเคราะห์กระจาย คำที่มีการตัดคำที่ถูกต้องตามพจนานุกรมและมีโครงสร้างสอดคล้องต้องตามหลักไวยากรณ์ ภาษาไทยจนได้ต้นไม้วีเคราะห์กระจายคำที่ดีที่สุด

ผลการทดลองการใช้โปรแกรมเชิงเจนเนติกในการสร้างต้นไม้วีเคราะห์กระจายคำใน ประโยคพบว่าสามารถสร้างโครงสร้างต้นไม้วีเคราะห์กระจายคำในประโยคได้ถูกต้อง 100% ทั้งในด้านการตัดคำและโครงสร้างต้นไม้ที่สอดคล้องกับหลักไวยากรณ์ของภาษาไทย อย่างไรก็ตาม การใช้การ โปรแกรมเชิงเจนเนติกในการวิเคราะห์กระจายคำในประโยคภาษาไทยยังต้องมีการ ปรับปรุงต่อไป เพื่อเพิ่มประสิทธิภาพในด้านเวลาของการทำงานของโปรแกรมให้ดียิ่งขึ้น

The research in this thesis is aimed to use genetic programming for word segmentation and parsing of Thai sentences based on Thai grammar.

The idea of using genetic programming in Thai sentence parsing is to create a population of possible parsing trees then gradually evolve the population by crossover ,mutation and reproduction operators to get the best parsing tree with correct word segmentation and structure that conforms with the Thai grammar.

The experiment on the genetic programming algorithm showed that it can achieve accuracy of 100% in constructing parse trees. However the algorithm still need some improvement to enhance its runtime performance.