

บทที่ 9

บทสรุปและแนวทางการพัฒนาต่อ

การตัดคำเป็นงานที่จำเป็นสำหรับงานด้านการประมวลผลภาษาและธรรมชาติสำหรับภาษาไทย ทำให้ได้มีการพัฒนาการตัดคำต่อเนื่องกันมาเป็นเวลานาน และจากงานวิทยานิพนธ์นี้ได้มีการนำคุณลักษณะต่างๆ เข้ามาใช้ในการตัดคำ ซึ่งจะสามารถแก้ปัญหาเรื่องความกำกวม และเรื่องคำศัพท์ที่ไม่ปรากฏในพจนานุกรม

9.1 ประสิทธิภาพการนำคุณลักษณะมาใช้ในการแก้ปัญหาการตัดคำ

สำหรับปัญหาเรื่องความกำกวม ภายในงานวิทยานิพนธ์นี้ได้แบ่งประเภทความกำกวมออกเป็น 2 ประเภท คือความกำกวมประเภทที่ต้องใช้บริบทในการแก้ปัญหา และความกำกวมที่ไม่จำเป็นต้องใช้บริบทในการแก้ปัญหา เมื่อนำการตัดคำโดยใช้คุณลักษณะเข้ามาช่วยนั้นสามารถจะแก้ปัญหาความกำกวมทั้ง 2 แบบได้ และให้ประสิทธิภาพดีกว่าการตัดคำที่ผ่านมา คือการตัดคำโดยใช้โมเดลไตรแกรม และการตัดคำโดยเลือกแบบเหมือนมากที่สุด

ส่วนปัญหาเรื่องคำศัพท์ที่ไม่ปรากฏในพจนานุกรมนั้นในวิทยานิพนธ์นี้ได้มีการแบ่งประเภทของคำศัพท์ที่ไม่ปรากฏในพจนานุกรมออกเป็น 2 ประเภทหลักๆ คือ คำศัพท์ที่ไม่ปรากฏในพจนานุกรมอย่างชัดเจน และคำศัพท์ที่ไม่ปรากฏในพจนานุกรมอย่างซ่อนเร้น และสำหรับคำศัพท์ที่ไม่ปรากฏในพจนานุกรมอย่างซ่อนเร้นนั้นยังแบ่งออกเป็นอีก 2 ประเภทย่อย ๆ คือ คำศัพท์ที่ไม่ปรากฏในพจนานุกรมอย่างซ่อนเร้นบางส่วน และ คำศัพท์ที่ไม่ปรากฏในพจนานุกรมอย่างซ่อนเร้นทุกส่วน ซึ่งในงานวิทยานิพนธ์สามารถแก้ปัญหาคำศัพท์ที่ไม่ปรากฏในพจนานุกรมได้ทุกประเภท แต่สำหรับการแก้ปัญหาคำศัพท์ที่ไม่ปรากฏในพจนานุกรมอย่างซ่อนเร้นทุกส่วนนั้น จะแก้ได้ดีกว่าแบบอื่น เนื่องจากคำศัพท์ประเภทนี้จะทำการค้นหาได้ยากกว่าแบบอื่น

ดังนั้นการนำคุณลักษณะเข้ามาใช้ในการแก้ปัญหาคำศัพท์ที่ไม่ปรากฏในพจนานุกรมนั้นจะมีประสิทธิภาพมากกว่า การตัดคำโดยใช้โมเดลไตรแกรม และการตัดคำโดยเลือกแบบที่เหมือนมากที่สุด

9.2 ข้อเสนอแนะ

วิทยานิพนธ์นี้ได้นำคุณลักษณะเข้ามาแก้ปัญหาทั้งความกำกวม และปัญหาคำศัพท์ที่ไม่ปรากฏในพจนานุกรม สำหรับการแก้ปัญหาความกำกวมแบบที่ไม่ต้องใช้บริบทนั้น ในทางปฏิบัติถ้ามีรายการข้อความที่กำกวมประเภทนี้ ก็สามารถจะนำไปใส่ไว้ในพจนานุกรมได้เลย และต้องมีข้อมูลบอกว่าควรจะตัดคำเป็นอย่างไร ซึ่งเมื่อบรรจุไว้ในพจนานุกรมแล้วจะช่วยให้การตัดคำสามารถทำงานได้เร็วขึ้น เนื่องจากไม่ต้องตัดคำที่เป็นไปได้ทุกแบบสำหรับข้อความกำกวมนั้น

สำหรับการแก้ปัญหาคำศัพท์ที่ไม่ปรากฏในพจนานุกรมภายในวิทยานิพนธ์นี้ จากการทดลองแสดงให้เห็นว่าคำศัพท์ที่ไม่ปรากฏในพจนานุกรมอย่างซ่อนเร้นทุกส่วนนั้น จะแก้ปัญหาได้ดีกว่าแบบอื่นๆ เนื่องจากการค้นหาคำศัพท์ประเภทนี้จะทำได้ยากกว่าแบบอื่นๆ ดังนั้นในการพัฒนาประสิทธิภาพในการแก้ปัญหาของคำศัพท์ประเภทนี้จึงควรจะต้องปรับปรุงวิธีการค้นหาบริเวณที่เกิดคำศัพท์ประเภทนี้



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย