

## บทที่ 7

### บทสรุปและข้อเสนอแนะ

#### 7.1 บทสรุป

ระบบการค้นคืนข้อความเป็นเครื่องมือสำคัญของมนุษย์ในการค้นหาข้อมูลเพื่อเพิ่มความรวดเร็ว โดยส่วนใหญ่แล้วระบบการค้นคืนมักนิยมจัดทำดัชนีเตรียมไว้ก่อนเพื่อความเร็วในการค้นคืนภายหลัง และการจัดทำดัชนีนิยมนำส่วนของข้อมูลมาเป็นคำหลักเพื่อใช้เป็นคีย์ในการค้นหา ซึ่งการเลือกคำหลักของข้อมูลภาษาไทยมีอาจทำได้ง่าย ๆ จำเป็นต้องอาศัยการแบ่งคำที่ชัดเจน ดังนั้นในวิทยานิพนธ์ฉบับนี้จึงได้นำเสนอโครงสร้างข้อมูลต้นไม้เพื่อแก้ไขปัญหาดังกล่าว โดยการนำข้อมูลมาจัดแบ่งเป็นข้อมูลย่อยๆ ตามแต่หน่วยความจำที่มีอยู่อย่างจำกัด พร้อมทั้งค้นหาชิสตริงของข้อมูลด้วย จากนั้นนำชิสตริงเหล่านี้ไปสร้างต้นไม้เพื่อค้นหาแพ็ทอะเรย์ ผลที่ได้คือชิสตริงถูกเรียงตามลำดับ แล้วนำแพ็ทอะเรย์ที่ได้มาผสานเป็นแพ็ทอะเรย์ฉบับใหม่เพื่อใช้เป็นดัชนีในการค้นหาข้อมูล แต่ส่วนใหญ่แล้วแพ็ทอะเรย์มักมีขนาดใหญ่มากเกินกว่าที่จะเก็บได้ในหน่วยความจำจึงจำเป็นต้องสร้างแพ็ทอะเรย์สั้นเพื่อเป็นดัชนีคลุมแพ็ทอะเรย์เก็บไว้ในหน่วยความจำอีกต่อหนึ่ง แล้วจึงนำไปเป็นดัชนีเพื่อการค้นคืนแบบเดิมหน้ากับแบบบูล

#### 7.2 ข้อเสนอแนะ

1. ควรปรับปรุงกฎทางภาษาศาสตร์เบื้องต้น ให้มีความถูกต้องและเหมาะสมมากขึ้น ทั้งนี้ต้องพิจารณาถึงความต้องการในการค้นหาข้อมูลด้วย เนื่องจากทำให้จำนวนดัชนีที่ใช้ในการค้นหาลดจำนวนลง (Recall ลดลง) แต่ช่วยลดเนื้อที่ที่ใช้ในการจัดเก็บข้อมูลได้
2. ควรปรับปรุงให้สามารถประมวลผลกับคุณสมบัติของชิสตริงอื่นๆ ได้ เช่น ชิสตริงที่มีความสำคัญต่างๆ กัน
3. ควรปรับปรุงให้สามารถประมวลผลการค้นหาข้อมูลได้หลากหลายแบบมากขึ้น เช่น การค้นหาแบบใกล้เคียง หรือการค้นหาแบบช่วง เป็นต้น