



บทสรุปและข้อเสนอแนะ

บทสรุป

งานวิจัยนี้เกิดจากความต้องการสร้างฐานข้อมูลพจนานุกรมไทยที่สามารถเข้าถึงได้หลายครั้งขึ้นที่มีความสามารถในการสืบค้นได้รวดเร็ว มีประสิทธิภาพในเรื่องเนื้อหาที่หน่วยความจำที่ใช้ และสามารถเพิ่มลบคำศัพท์ได้ตลอดเวลาโดยไม่ต้องเสียเวลาสร้างพจนานุกรมทั้งหมดใหม่ จึงได้ศึกษาโครงสร้างข้อมูลทรีเนื่องจากเหมาะสมกับธรรมชาติของการสืบค้นคำ และทรีแถวคู่เป็นทรีที่มีคุณสมบัติตามที่ต้องการทั้งหมด อีกทั้งยังมีโครงสร้างและอัลกอริธึมที่ไม่ซับซ้อน สามารถปรับปรุงแก้ไขได้ง่าย จึงทำการพัฒนาต้นแบบพจนานุกรมไทยขึ้น โดยใช้โครงสร้างทรีแถวคู่

ได้ทำการทดลองทดสอบคุณสมบัติต่างๆเปรียบเทียบกับทรีอัดแน่นที่ปรับปรุงขึ้นเป็นทรีอัดแน่นแบบจลน์ เนื่องจากทรีอัดแน่นแบบจลน์ก็มีคุณสมบัติต่างๆตามต้องการเช่นกัน จากผลการทดสอบพบว่าทรีแถวคู่มีประสิทธิภาพในการเพิ่มลบคำดีกว่า เพราะว่ามีโครงสร้างที่ยืดหยุ่นกว่า

นอกจากความสามารถในการเพิ่มลบคำยังพบว่าทรีแถวคู่มีความสามารถในการท่องสถานะย้อนกลับได้ อีกทั้งยังมีลักษณะที่ตรงกับการเปรียบเทียบ (Pattern Matching) จึงน่าจะนำไปใช้กับการแบ่งคำด้วยพจนานุกรมได้ด้วย

ยิ่งไปกว่านั้นยังพบว่าสามารถเพิ่มโยง (Link) เข้าไปใน BASE ได้โดยง่าย และไม่เสียเนื้อที่เพิ่มขึ้น จึงอาจจะนำไปประยุกต์ใช้เป็นโครงสร้างข้อมูลของฐานข้อมูลอื่นที่ใช้คำศัพท์เป็นกรณีสืบค้น

แต่ทว่าโครงสร้างทรีทรีแถวคู่นี้หากมีอาร์เรย์คู่ (BASE และ CHECK) ขนาดใหญ่ เวลาในการเพิ่มคำศัพท์จะมากขึ้น เนื่องจากอัลกอริธึมในส่วนของฟังก์ชัน X_CHECK เป็นการหา BASE และ CHECK แบบพอดีก่อน (First-fit) ซึ่งหากปรับปรุงอัลกอริธึมในการเพิ่มคำ โดยไม่ต้องทำการย้ายสถานะที่จะต้องถูกแทนที่ในฟังก์ชัน MODIFY แต่ให้เก็บสตริงเดียวใน TAIL ชั่วคราวแทน การเพิ่มคำควรทำได้เร็วขึ้น

ข้อเสนอแนะ

1. บรรณาธิการพจนานุกรม (Dictionary Editor)

นอกจากการพัฒนาฐานข้อมูลเพื่อใช้เป็นคลังศัพท์กลางแล้ว ยังสามารถสร้างบรรณาธิการพจนานุกรมโดยเพิ่มส่วนการติดต่อกับผู้ใช้ เพื่อให้ผู้ใช้มีความสะดวกในการปรับปรุงแก้ไขพจนานุกรมส่วนตัวได้

2. การแบ่งคำ การตัดคำ และการตรวจสอบคำสะกด

ด้วยความสามารถในการย้อนรอย (Backtracking) สถานะต่างๆ ในทรีไค์ จึงอาจนำไปใช้สร้างอัลกอริทึมการแบ่งคำและการตัดคำได้ และยังสามารถประยุกต์ใช้กับการตรวจสอบคำสะกดในการสืบค้นคำที่สะกดใกล้เคียงเมื่อสะกดผิด

3. สารานุกรม (Encyclopedia)

เมื่อเพิ่มโยง (Link) เข้าไปใน BASE ของทรีไค์แวลวู่โดยโยงเข้ากับส่วนข้อมูลอื่นๆ จะสามารถพัฒนาเป็นฐานข้อมูลประเภทต่างๆ ของคำศัพท์

4. อรรถนิธาน (Thesaurus)

สามารถนำเอาอุปสรรครวมที่ถูกยุบของทรีไค์มาจัดเป็นหมวดหมู่เพื่อสร้างอรรถนิธานได้ ตัวอย่าง เช่นหมวดหมู่ของคำที่ขึ้นต้นด้วย "วิชา" คือ วิชาเลขคณิต วิชาประวัติศาสตร์ วิชาภาษาอังกฤษ ฯลฯ หรือที่ขึ้นต้นด้วย "เครื่องพิมพ์" คือ เครื่องพิมพ์เลเซอร์ เครื่องพิมพ์แบบจุด เครื่องพิมพ์แบบชนิดหมึก ฯลฯ

5. ทรีไค์แวลวู่สองปลาย (Bi-Tail Double Array Trie)

จากการที่อาร์เรย์คู่ของทรีไค์แวลวู่ยิ่งใหญ่ขึ้น เวลาที่ใช้ในการเพิ่มคำยิ่งมากขึ้น จึงควรที่จะปรับปรุงทรีไค์แวลวู่นี้โดยยังไม่เพิ่มหรือย้ายสถานะที่อยู่ใน BASE และ CHECK ในทันที แต่เก็บไว้ในอีก TAIL หนึ่ง ซึ่งเมื่อต้องการทำให้เป็นทรีไค์แวลวู่ปกติ ก็ทำการกระจายสตริงเดียวใน TAIL ชั่วครวณนั้นให้เป็นอาร์เรย์คู่ ซึ่งทำให้ไม่เสียเวลาในการเพิ่มคำแต่ละคำ