

บทสรุป



รูปที่ 12.1 แสดงภาพโดยรวมของงานวิจัยที่ทำ

จากแผนภาพที่แสดง ในการศึกษากการแจ่งส่วนประโยคที่ผิดรูปแบบไวยากรณ์นั้น แทนที่เราจะใช้ไวยากรณ์ขนาดใหญ่ที่ซับซ้อนและใช้เวลาในการประมวลผลมาก เราจะเลือกไวยากรณ์เพียงบางส่วนที่จัดว่าเป็นแกนของภาษาเพื่อนำมาพื้นฐานความรู้เริ่มแรกของตัวแจ่งส่วน ไวยากรณ์ส่วนนี้จะใช้ในการแจ่งส่วนประโยคที่ถูกรูปแบบไวยากรณ์ทั่วไป ส่วนของข้อมูลที่เบี่ยงเบนไปจากไวยากรณ์ที่กำหนดจะถูกจัดการด้วยตัวแจ่งส่วนประโยคที่ผิดรูปแบบไวยากรณ์ และข้อมูลที่เบี่ยงเบนเหล่านี้จะถูกเก็บไว้เป็นคำสถิติ เพื่อนำกลับมาปรับปรุงไวยากรณ์ หรือฐานข้อมูลที่ใช้ให้ทำงานได้อย่างมีประสิทธิภาพมากขึ้น และสามารถกล่าวโดยสรุปถึงงานที่ทำในวิทยานิพนธ์นี้คือ

1. การเสนอแนะวิธีการปรับปรุงตัวแจ่งส่วนแบบจีแอลอาร์ เพื่อให้สามารถจัดการแจ่งส่วนประโยคที่ผิดรูปแบบไวยากรณ์ โดยได้แบ่งประเภทของข้อผิดพลาดออกเป็น 3 ประเภทคือ ข้อผิดพลาดที่เกิดมาจากการแทรกองค์ประกอบบางอย่างลงไปประโยค ข้อผิดพลาดที่เกิดจากการขาดองค์ประกอบบางอย่างของประโยค และ

สุดท้ายคือข้อผิดพลาดที่เกิดจากการแทนที่องค์ประกอบบางตัวของประโยค อันอาจเนื่องมาจากการพิมพ์ผิด หรือการสะกดผิด

การปรับปรุงตัวแฉงส่วนแบบจีแอลลาร์นั้น ได้เสนอวิธีการที่ใช้ในการแฉงส่วนประโยคที่ผิดรูปแบบ ไวยากรณ์ 3 วิธีการด้วยกัน โดยแยกตามประเภทของการพิจารณาหาจุดเริ่มต้นที่จะทำการแก้ไขประโยค คือ

1.1 การแก้ไขประโยค ณ จุดที่ตรวจพบการผิดไวยากรณ์เท่านั้น วิธีการนี้ง่ายกับการเขียนโปรแกรมที่สุด ใช้เวลาในการประมวลผลน้อยที่สุด ครอบคลุมข้อผิดพลาดได้ดีพอสมควร เช่น ประโยค “เขา กับ เธอ ดี ฉัน” ถ้ามีการพิมพ์ผิดเป็น “เขา กัด เธอ ดี ฉัน” การแก้ไขประโยคด้วยวิธีนี้ได้ประโยคออกมาเป็น “เขา กัด เธอ [p] ฉัน” ซึ่งยังไม่เหมาะสมนัก และข้อผิดพลาดที่ “กัด” ก็ยังคงไม่ได้ถูกแก้ไข (อ้างอิงตาม ไวยากรณ์ที่กำหนด) เพราะตัวแฉงส่วนจะตรวจพบข้อผิดพลาดที่คำกริยาคำที่ 2 “ดี” และทำการแก้ไขประโยคที่ ตำแหน่งนั้นเพียงตำแหน่งเดียว

1.2 การแก้ไขประโยคโดยการแฉงส่วนแบบจีแอลลาร์-ไอจีแอลลาร์ วิธีการนี้ก็สะดวกและง่ายกับการเขียนโปรแกรม ใช้เวลาในการทำงานเป็น สองเท่าของวิธีแรก ครอบคลุมข้อผิดพลาดได้มากขึ้น แต่ก็มีข้อผิดพลาดในบางลักษณะที่จะถูกมองข้ามไป เพราะไม่มีการย้อนกลับไปพิจารณาคำที่แฉงส่วนผ่านมาแล้ว การแฉงส่วนแบบนี้จัดการกับข้อผิดพลาดที่แสดงในข้อ 1.1 ได้ แต่จัดการกับประโยคในลักษณะต่อไปนี้ไม่ได้

ประโยค	: เขา กับ เธอ ดี ฉัน กับ น้อง
ประโยคผิดไวยากรณ์	: เขา กัด เธอ ดี ฉัน กัด น้อง
ประโยคที่แก้ไขแล้ว	: เขา กัด เธอ [p] ฉัน [p] น้อง
	: เขา [p] เธอ [p] ฉัน กัด น้อง

1.3 การแก้ไขประโยคโดยการแฉงส่วนแบบย้อนรอยกองซ้อนโครงสร้างกราฟ วิธีนี้ค่อนข้างซับซ้อนกว่าสองวิธีแรก และใช้เวลาในการประมวลผลมากกว่า แต่มีข้อดีคือสามารถครอบคลุมข้อผิดพลาดได้มากที่สุด (ใน 3 วิธีที่เสนอ) เพราะนอกจากจะแก้ไขข้อผิดพลาด ณ จุดที่ตรวจพบแล้ว ยังมีการย้อนรอยกลับไปแก้ไขประโยคในจุดที่มีโอกาสเกิดข้อผิดพลาดได้ด้วย จึงสามารถจัดการกับประโยคที่แสดงในหัวข้อ 1.2 ได้ และได้ประโยคที่เหมาะสมออกมาคือ “เขา [p] เธอ ดี ฉัน [p] น้อง” และจัดว่าเป็นวิธีที่เหมาะสมกับการใช้งานจริง อย่างไรก็ตาม เนื่องจากการจำกัดขอบเขตของการแก้ปัญหา ดังนั้นก็เป็นไปได้ว่าในการผิดไวยากรณ์บางรูปแบบ การแก้ไขปัญหาดังกล่าวด้วยวิธีการนี้อาจมองข้ามปัญหาในบางจุดไปได้ ตัวอย่างเช่น

จากไวยากรณ์ที่กำหนดให้ ถ้าพิจารณารูปแบบการเกิดของลำดับคำบุพบท (p) กับคำกริยา (v)

ประโยคที่ถูกไวยากรณ์ คือ : เขา กับ เธอ ช่วย ฉัน กับ น้อง (รูปแบบ p-v-p)

n p n v n p n

ประโยคชนิดไวยากรณ์ มีข้อผิดพลาด 3 จุด คือ : เขา กัด เธอ ด้วย ฉัน กัด น่อง (รูปแบบ v-p-v)

n v n p n v n

ประโยคที่ผ่านการแก้ไขแล้วอาจมีได้หลายประโยค แต่โดยสรุปแล้วจะไม่สามารถแก้ไขประโยคให้อยู่ในรูปแบบที่ต้องการได้ เช่น จะแก้ไขประโยคออกมาเป็น : เขา [p] เธอ ด้วย ฉัน กัด น่อง (รูปแบบ p-p-v)

และ เขา กัด เธอ ด้วย ฉัน [p] น่อง (รูปแบบ v-p-p)

(รูปแบบตำแหน่งของ p กับ v ที่ต้องการคือ p-v-p แต่ตัวแฉงส่วนสามารถแก้ไขออกมาได้เพียง 2 รูปแบบ คือ p-p-v และ v-p-p)

2. เสนอแนะวิธีการเก็บข้อมูลทางสถิติ ที่เกิดจากการผิดรูปแบบไวยากรณ์ของประโยคที่นำมาแฉงส่วน เพื่อนำข้อมูลที่ได้เหล่านี้ไปใช้ในการปรับปรุงพจนานุกรม และไวยากรณ์ที่ใช้อยู่ให้เหมาะสมกับการทำงานมากขึ้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย