

ตัวแจงส่วนแบบจีแอลอาร์

การแจงส่วนแบบจีแอลอาร์ (General LR parsing, GLR parsing) (Tanaka, 1992, 1993; Tomita, 1991) จัดว่าเป็นการแจงส่วนประโยคที่มีประสิทธิภาพสูง มีความรวดเร็วกว่า การแจงส่วนแบบอื่นๆ การแจงส่วนแบบจีแอลอาร์ เป็นการแจงส่วนที่พัฒนามาจากการแจงส่วนแบบแอลอาร์ (LR parsing) (Aho, 1986) ที่เป็นการแจงส่วนประโยคจากล่างขึ้นบน และทำการแจงส่วนประโยคที่ต้องการวิเคราะห์ จากซ้ายมือของประโยคไปทางขวามือทีละคำ ปัญหาที่พบสำหรับการใช้ตัวแจงส่วนแบบแอลอาร์ (LR parser) ก็คือขอบเขตของภาษาที่สามารถใช้ตัวแจงส่วนแบบแอลอาร์ จัดการได้ค่อนข้างแคบ กล่าวคือจะใช้ได้ดีกับประโยคที่ถูกต้องตามหลักไวยากรณ์แอลอาร์ (LR grammar) เท่านั้น ไม่ได้เตรียมมาให้ใช้จัดการกับประโยคที่ซับซ้อน หรือมีความกำกวมมาก เพราะแต่เดิมตัวแจงส่วนแบบแอลอาร์ ถูกพัฒนาขึ้นมาเพื่อใช้ในการตรวจสอบไวยากรณ์ของภาษาทางคอมพิวเตอร์ ซึ่งไม่ซับซ้อนและไม่มีความกำกวม จุดอ่อนข้อนี้ของตัวแจงส่วนแบบแอลอาร์ทำให้มีปัญหาในการใช้งานกับการประมวลผลภาษาธรรมชาติ ที่ข้อมูลมีความกำกวมเกิดขึ้นตลอดเวลา

ปี ค.ศ. 1985 Masaru Tomita ได้พัฒนาตัวแจงส่วนแบบจีแอลอาร์ (GLR parser) ขึ้น (Tomita, 1991) โดยที่ตัวแจงส่วนแบบนี้ยังคงคุณสมบัติการทำงานที่มีประสิทธิภาพสูงของตัวแจงส่วนแบบแอลอาร์ไว้ ในขณะที่การทำงานของตัวแจงส่วนไม่ถูกจำกัด ให้วิเคราะห์เฉพาะประโยคที่ถูกต้องตามไวยากรณ์แอลอาร์เท่านั้น แต่ได้ขยายการทำงานให้ครอบคลุมประโยคตามไวยากรณ์ไม่ทั้งบริบทด้วย ทำให้ตัวแจงส่วนแบบนี้สามารถแจงส่วนประโยคที่มีความซับซ้อน หรือกำกวมทางไวยากรณ์ได้

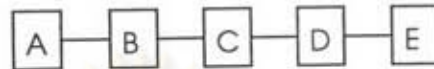
ส่วนประกอบสำคัญของการแจงส่วนแบบจีแอลอาร์ คือส่วนที่เป็นโครงสร้างข้อมูล (data structure) แบบที่เรียกว่า กองซ้อนโครงสร้างกราฟ (Graph-Structured Stack) กับตารางแอลอาร์ (LR table)

3.1 กองซ้อนโครงสร้างกราฟ (Graph-Structured Stack or GSS)

จุดเด่นที่ทำให้การแจงส่วนแบบแอลอาร์ (LR parser) มีประสิทธิภาพ และความเร็วสูงก็คือการแปลงกฎเกณฑ์ไวยากรณ์ต่างๆ และการทำงานให้อยู่ในรูปตารางแอลอาร์ การแจงส่วนแบบจีแอลอาร์ (GLR parser) ก็ยังคงข้อดีในจุดนี้ไว้ และใช้โครงสร้างข้อมูลแบบที่เรียกว่า กองซ้อนโครงสร้างกราฟ เพื่อทำให้ตัวแจงส่วนสามารถทำการแจงส่วนประโยคที่มีความกำกวม หรือประโยคที่มีไวยากรณ์ซับซ้อนได้ โดยไม่เสียเวลาในการประมวลผลมากนัก การทำงานของโครงสร้างข้อมูลแบบนี้แบ่งได้เป็น 3 แบบ (Tomita, 1991) คือ การแยกส่วน (splitting) การรวมกลุ่ม (combining) และการจัดความกำกวม (local ambiguity packing)

1. การแยกส่วน (splitting) เมื่อข้อมูลในกองซ้อน (stack) มีความกำกวม (มีการกระทำได้หลายการกระทำในสถานะนี้) หรือสามารถถูกลดทอน (reduce) เป็นโครงสร้างที่ใหญ่ขึ้นได้ในหลายรูปแบบ ส่วนปลายของกองซ้อน (top of stack) จะถูกแยกออกเป็นหลายปลาย เพื่อให้สามารถทำการลดทอนได้ในทุกรูปแบบที่มี

ตัวอย่างในรูปที่ 3.1 แสดงข้อมูลที่อยู่ในกองซ้อน โดยที่ A เป็นข้อมูลล่างสุดของกองซ้อน (bottom of stack) และ E เป็นข้อมูลบนสุดของกองซ้อน (top of stack)



รูปที่ 3.1 แสดงตัวอย่างข้อมูลในกองซ้อนที่มี A เป็นข้อมูลล่างสุด และ E เป็นข้อมูลบนสุด

ถ้าข้อมูลในกองซ้อนถูกลดทอนได้โดยกฎการผลิต 3 ข้อคือ

F → DE

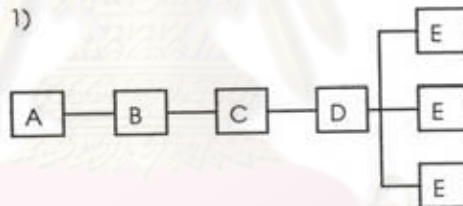
G → DE

H → CDE

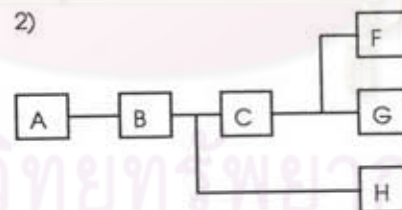
การทำงานของข้อมูลคือแยก E ออกเป็น 3 ปลาย จากนั้นก็ลดทอนข้อมูลในกองซ้อนตามกฎที่กำหนด

ทั้ง 3 ข้อ

1)



2)

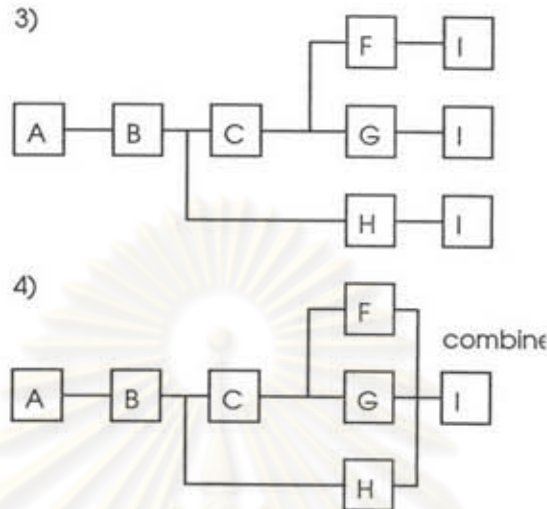


รูปที่ 3.2 แสดงการแยกส่วน ที่ส่วนปลายของกองซ้อน เพื่อทำการลดทอนกฎทั้ง 3 แบบที่กำหนดมา

เนื่องจากโครงสร้างข้อมูลกองซ้อนที่ทำงานอยู่ เป็นโครงสร้างข้อมูลแบบกราฟ ดังนั้นจึงมีส่วนปลายของกองซ้อนได้หลายค่า จากตัวอย่างในรูปที่ 3.2 F, G และ H เป็นข้อมูลบนสุดของกองซ้อนทั้ง 3 ค่าการทำงานในขั้นต่อไปจะต้องเริ่มจากค่าปลายทั้ง 3 นี้

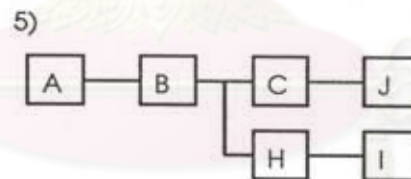
2. การรวมกลุ่ม (combining) การรวมกลุ่มมักเกิดหลังการกระทำการปรับเปลี่ยน (shift action) ถ้ากองซ้อนมีการปรับเปลี่ยนไปยังค่าเดียวกันทุกปลายของกองซ้อน เราจะทำการรวมปลายข้อมูลใหม่เข้าด้วยกัน จากนั้นการทำงานในขั้นต่อไปจะเหลือการทำงานกับปลายของกองซ้อนเพียงปลายเดียวทำให้การทำงานรวดเร็วและใช้เนื้อที่

ในหน่วยความจำน้อยลง จากตัวอย่างที่แสดงในรูปที่ 3.3 ถ้ากำหนดให้การทำงานในขั้นต่อไป คือการปรับเปลี่ยนค่า I ลงกองซ้อน การทำงานจะมีขั้นตอนดังนี้



รูปที่ 3.3 แสดงการรวมกลุ่มค่าปลายของกองซ้อนทั้ง 3 สาย

3. การจัดความกำกวม (local ambiguity packing) ในบางครั้ง บางส่วนของข้อมูลอาจมีความกำกวมเกิดขึ้น เช่น $J \rightarrow FI$ หรือ $J \rightarrow GI$ จากตัวอย่างข้างบน รูปที่ 3.3 ขั้นตอนที่ 4 จะถูกยุบรวมในขั้นตอนที่ 5 คือ



รูปที่ 3.4 การจัดความกำกวมของข้อมูลในกองซ้อน ที่ค่าปลาย FI กับ GI

โดยที่ A-B-C-J สามารถกระจายได้เป็น 2 ทาง แต่ถูกจัด (pack) รวมไว้ด้วยกันเพื่อว่าเวลาทำงานจะได้จัดการเหมือนกับว่ามีกองซ้อนที่ต้องดูแลเพียงสายเดียว

3.2 ตารางแอลอาร์ (LR table)

สร้างมาจากการแปลงกฎการผลิตต่างๆ ให้อยู่ในรูปตาราง เพราะฉะนั้น ตัวแ่งส่วนแบบนี้จึงไม่สามารถทำการแ่งส่วนประโยคที่ผิดรูปแบบไปจากไวยากรณ์ไม่พึงบริบทที่กำหนดไว้ได้ จากตัวอย่างไวยากรณ์ไม่พึงบริบทในรูปที่ 3.5 จะสามารถสร้างตารางแอลอาร์ ได้ดังตารางที่ 3.1

- | | |
|----------------------------|---------------------------|
| (1) $S \rightarrow NP VP$ | (5) $PP \rightarrow p NP$ |
| (2) $S \rightarrow S PP$ | (6) $VP \rightarrow v NP$ |
| (3) $NP \rightarrow n$ | (7) $VP \rightarrow v PP$ |
| (4) $NP \rightarrow NP PP$ | |

รูปที่ 3.5 ไวยากรณ์ไม่พึ่งบริบทแบบง่ายสำหรับภาษาไทย

State	Action Field				Goto Field			
	n	v	p	\$	NP	PP	VP	S
0	sh3				2			1
1			sh5	acc		4		
2		sh8	sh5			7	6	
3		re3	re3	re3				
4			re2	re2				
5	sh3				9			
6			re1	re1				
7		re4	re4	re4				
8	sh3		sh5		10	11		
9		re5	re5/sh5	re5		7		
10			re6/sh5	re6		7		
11			re7	re7				

ตารางที่ 3.1 ตารางแอลอาร์ สำหรับไวยากรณ์ไม่พึ่งบริบท จากรูปที่ 3.5

ตารางแอลอาร์จะประกอบด้วยส่วนสำคัญ 2 ส่วนคือ ส่วนของเขตการกระทำ (Action field) และส่วนของเขตไปยัง (Goto field) โดยที่

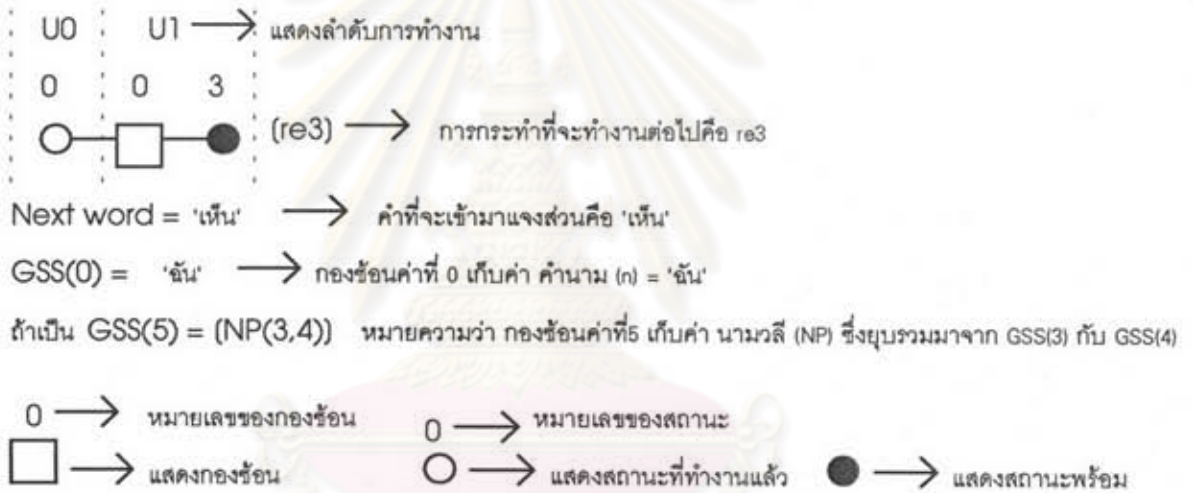
เขตการกระทำ จะเป็นตัวกำหนดการกระทำ ที่จะต้องทำในลำดับต่อไปของสถานะต่างๆ โดยขึ้นอยู่กับสถานะพร้อม (active state) ในขณะนั้น กับค่าชนิดของคำที่เข้ามาแฉงส่วน ตัวอย่างเช่น ถ้าสถานะพร้อมในขณะนั้นเป็นสถานะที่ 0 และชนิดของคำที่เข้ามาแฉงส่วนคือคำนาม (n) การกระทำที่จะต้องทำต่อไป ก็คือการปรับเปลี่ยนไปยังสถานะที่ 3 (sh3)

ส่วนของเขตไปยัง จะเป็นตัวบอกสถานะพร้อมลำดับต่อไป ที่จะทำงานต่อหลังจากที่มีการลดทอนข้อมูลในกองซ้อนให้เป็นองค์ประกอบที่ใหญ่ขึ้น โดยขึ้นอยู่กับสถานะพร้อมในขณะนั้นกับกฎที่ใช้ในการลดทอน ตัวอย่างเช่น ถ้าสถานะพร้อมในขณะนั้นเป็นสถานะที่ 0 แล้วมีการลดทอนข้อมูลในกองซ้อนเป็นส่วนนามวลี (NP) สถานะที่จะเป็นสถานะพร้อมในการทำงานครั้งต่อไปคือสถานะที่ 2

ความหมายของสัญลักษณ์ที่ใช้ในตารางแอลฮาร์

- sh[No.] : การปรับเปลี่ยนไปยังสถานะที่ No. เช่น sh4 คือการปรับเปลี่ยนไปยังสถานะที่ 4
- re[No.] : การลดทอนตามกฎข้อที่ No. เช่น re6 คือการลดทอน คำกริยา (v) กับส่วนนามวลี (NP) เป็นส่วนกริยาวลี (VP) ตามกฎข้อที่ 6 (VP -> v NP).
- acc : การยอมรับว่าประโยคที่ทำการแฉ่งส่วนเป็นประโยคที่ถูกต้องตามไวยากรณ์
- \$: เป็นสัญลักษณ์ปิดท้ายประโยค แสดงการสิ้นสุดของประโยคที่ส่งเข้ามาแฉ่งส่วน
- / : เป็นสัญลักษณ์ที่บอกว่าการกระทำที่สถานะปัจจุบันมีการกระทำได้มากกว่า 1 ทาง เช่น re6/sh5 คือตัวแฉ่งส่วนต้องกระทำทั้งการลดทอน re6 และการปรับเปลี่ยน sh5

ความหมายของสัญลักษณ์ที่ใช้ในการอธิบายขั้นตอนการแฉ่งส่วนประโยค

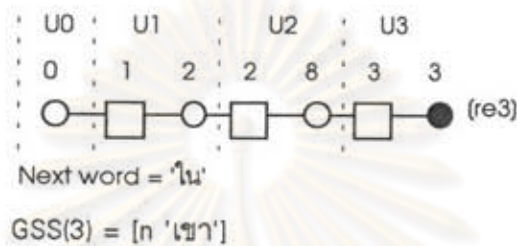


รูปที่ 3.6 แสดงความหมายของสัญลักษณ์ที่ใช้ในการอธิบายขั้นตอนการแฉ่งส่วนประโยค

การทำงานของตัวแฉ่งส่วนแบบจีแอลฮาร์นั้น เมื่อนำคำจากประโยคข้อมูลที่ต้องการวิเคราะห์เข้ามาทำการแฉ่งส่วน ตัวแฉ่งส่วนจะทราบ ชนิดคำ (category) ของคำที่เข้ามาโดยการให้พจนานุกรม เช่นตัวอย่างประโยค "ฉัน เห็น เขา ใน สวน" มีชนิดคำเรียงตามลำดับคือ "n v n p n \$" ('\$' เป็นสัญลักษณ์ แสดงการจบประโยค) แล้วนำชนิดคำที่ได้ มาหาขั้นตอนการทำงานต่อไปใน เขตการกระทำ (Action field) ตามสถานะ (state) ที่ทำงานอยู่ขณะนั้น จากตารางแอลฮาร์ (ตัวแฉ่งส่วนจะเริ่มต้นการแฉ่งส่วนที่สถานะ 0)

การทำงานของตัวแฉ่งส่วนแบบจีแอลฮาร์ ตามการกระทำ (action) ที่กำหนด จะมีอยู่ 4 ลักษณะด้วยกัน (Aho, 1986; Tanaka, 1993; Tomita, 1991) คือ

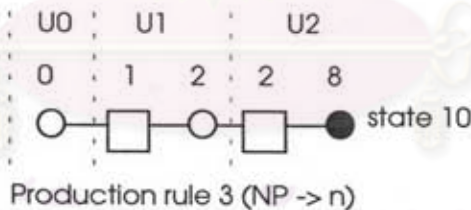
1. ลดทอน (reduce : re[No.]) จะเป็นการยุบรวม คำ หรือ กฎ ในกองซ้อนโครงสร้างกราฟ ให้เป็น กฎ หรือ สัญลักษณ์ไม่ปลายทาง ในระดับที่สูงขึ้น (No. คือ หมายเลขของกฎที่นำมาใช้) คำหรือองค์ประกอบที่นำมายุบรวมกัน (สัญลักษณ์ทางขวามือของกฎการผลิต) จะถูกดึงออกจากกองซ้อน (pop stack) ตามจำนวนที่สอดคล้องกับจำนวนของสัญลักษณ์ทางขวามือของกฎที่ใช้ในการยุบ เช่น ถ้าการกระทำที่กำหนด เป็นการลดทอนด้วยกฎการผลิตข้อที่ 3 คือ $NP \rightarrow n$ ตัวแฉ่งส่วนก็จะดึงค่าค่านาม (n) ออกจากกองซ้อน 1 คำ เพื่อยุบรวมเป็นส่วนนามวลี (NP) แล้วเก็บค่านามวลีนี้ลงกองซ้อน ตัวอย่างเช่น



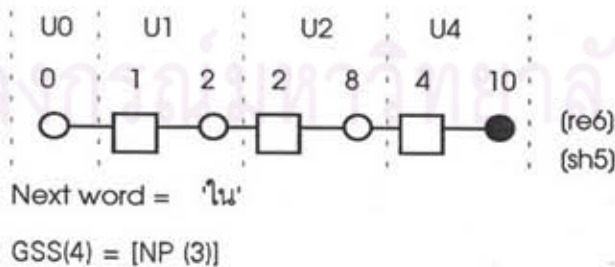
รูปที่ 3.7 สถานะพร้อม 3 คำที่เข้าตรวจสอบคือคำบุพบท 'ใน' จากตารางแอลอาร์ การกระทำที่ได้คือ re3

กฎการผลิตที่ใช้ในการลดทอนคือกฎข้อ 3 ($NP \rightarrow n$) ดังนั้นตัวแฉ่งส่วนจะดึงค่าจากกองซ้อนออกมา 1 คำ คือค่า $GSS(3) = [n \text{ 'เขา'}]$ มายุบเป็นองค์ประกอบทางซ้ายมือของกฎการผลิต (NP) ดังนั้นผลลัพธ์จากการลดทอน คือ NP (3) จะถูกใส่ลงในกองซ้อนลำดับที่ 4 ($GSS(4) = [NP (3)]$)

ก)



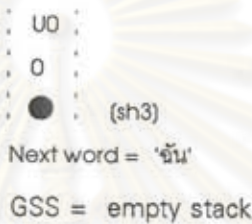
ข)



รูปที่ 3.8 แสดงขั้นตอนการทำงานหลังจากดึงค่า GSS(3) ออกจากกองซ้อนแล้ว

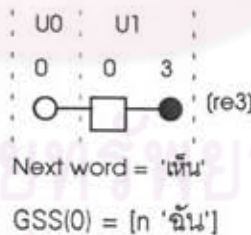
รูปที่ 3.8 ก. ขณะที่มีการลดทอนกฎ และดึงค่า GSS(3) ออกจากกองซ้อน สถานะพร้อมทำงานจะอยู่ที่สถานะที่ 8 ส่วนสถานะที่จะทำงานต่อไปจะดูไปจากเขตไปยัง คอลัมน์ NP ซึ่งเป็นคอลัมน์ของกฎที่ใช้ในการลดทอน (ทุกครั้งที่มีการลดทอนกฎ การทำงานในสถานะต่อไปจะดูจากเขตไปยัง ถ้าเป็นการปรับเลื่อนการทำงานของตัวแฉงส่วน จะพิจารณาค่าสถานะจากเขตการกระทำเท่านั้น) ค่าที่ได้คือสถานะ 10 ดังนั้นตัวแฉงส่วนจะเก็บค่า GSS(4) ลงกองซ้อน และเปลี่ยนสถานะพร้อมเป็นสถานะที่ 10 ตามตัวอย่างที่แสดงในรูป 3.8 ข.

2. ปรับเลื่อน (shift : sh[No.]) เป็นการเปลี่ยนสถานะ ไปยังสถานะที่กำหนดพร้อมทั้งเก็บค่าลงกองซ้อน โครงสร้างกราฟ (No. เป็นหมายเลขของสถานะ ที่จะทำงานต่อไป) ตัวอย่างเช่น



รูปที่ 3.9 สถานะพร้อมคือ 0 ค่าที่ตรวจสอบคือ คำนาม (n) จากตารางแอลอาร์ จะได้รับการกระทำ sh3

การทำงานของตัวแฉงส่วนคือการใส่ค่า คำนาม 'จัน' ลงกองซ้อน GSS(0) = [n'จัน'] และเปลี่ยนสถานะพร้อมไปเป็นสถานะที่ 3 เมื่อมีการปรับเลื่อนค่าคำที่แฉงส่วนลงในกองซ้อน ตัวแฉงส่วนจะต้องดึงค่าใหม่จากประโยคที่วิเคราะห์เข้ามา ดังนั้นค่าที่จะวิเคราะห์ต่อมา คือคำกริยา 'เห็น' ตามตัวอย่างที่แสดงต่อไป



รูปที่ 3.10 ผลที่ได้หลังการทำงานตามขั้นตอน U0

3. ผิดพลาด (error) จะมีข้อผิดพลาดเกิดขึ้น ในกรณีที่ผลการเปรียบเทียบชนิดคำที่ได้ กับตารางแอลอาร์ แล้วไม่มีการทำงานในสถานะพร้อมนั้น จะสรุปได้ว่า ประโยคข้อมูลที่ต้องการวิเคราะห์ เป็นประโยคที่มีรูปแบบจากไวยากรณ์ที่กำหนด

4. ยอมรับ (accept : acc) ถ้าพบการกระทำนี้ จะสามารถสรุปได้ว่าประโยคข้อมูลที่ต้องการวิเคราะห์ เป็นประโยคที่ถูกต้องตามไวยากรณ์

ตัวอย่างการแจงส่วนประโยค

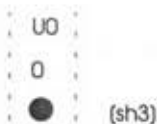
การแจงส่วนประโยค "ฉัน เห็น เขา ใน สวน" โดยใช้ตารางแอลอาร์ตารางที่ 3.1 จะมีขั้นตอนการแจงส่วนประโยค ดังนี้

ประโยคที่ทำการแจงส่วน : ฉัน เห็น เขา ใน สวน

ชนิดของคำที่ใช้ในการแจงส่วน : n v n p n

1) - ขั้นตอนแรก (U0) ตัวแจงส่วน จะเริ่มทำงานที่ สถานะ 0

- คำที่เข้ามา คือ 'ฉัน' มีชนิดคำเป็น คำนาม (n)
- จาก สถานะ 0 , คอลัมน์ (n) จะได้ การกระทำ sh3
- กองซ้อนโครงสร้างกราฟ (GSS) ว่างอยู่

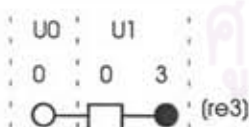


Next word = 'ฉัน'

GSS = empty stack

2) - ต่อมา (U1) ใส่คำ (push) คำนาม 'ฉัน' ลงกองซ้อนโครงสร้างกราฟ $GSS(0) = [n \text{ 'ฉัน'}]$

- คำต่อมา คือ 'เห็น' มีชนิดคำเป็น คำกริยา (v)
- กระทำ [sh3]
- กระโดดไป สถานะ 3, คอลัมน์ (v) จะได้ การกระทำ re3



Next word = 'เห็น'

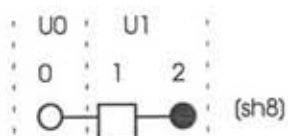
$GSS(0) = [n \text{ 'ฉัน'}]$

3) - เมื่อพบ การกระทำลดทอน (reduce) จะมีการดึงองค์ประกอบบางตัวออกจากกองซ้อน กรณีนี้ลดทอนด้วยกฎการผลิตข้อ 3 คือ $NP \rightarrow n$ ดังนั้นจะดึงองค์ประกอบออกจากกองซ้อน 1 คำ คือ $GSS(0)$ (คำนาม 'ฉัน') ออกมาเพื่อยุบเป็นองค์ประกอบที่ใหญ่ขึ้น คือ NP ใน $GSS(1)$ จากนั้นสถานะที่พร้อม คือ สถานะ 0

- ลดทอนกฎ ข้อ 3 (NP) ใช้การกระทำจากเซตไปยัง (Goto field) สถานะ 0, คอลัมน์ (NP) จะได้สถานะพร้อมคือ สถานะที่ 2

- ใส่ค่าองค์ประกอบใหม่ลงในกองซ้อนโครงสร้างกราฟ ตามกฎที่ใช้ คือ $GSS(1) = [NP (0)]$

- เลื่อนไปสถานะ 2, คอลัมน์ (v) จะได้รับการกระทำในสถานะใหม่คือ sh8



Next word = 'เห็น'

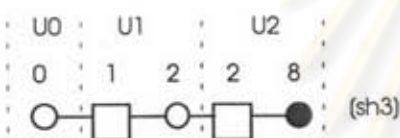
$GSS(1) = [NP (0)]$

4) - ต่อมา (U2) ใส่คำกริยา 'เห็น' ลงกองซ้อนโครงสร้างกราฟ $GSS(2) = [v 'เห็น']$

- คำต่อมา คือ 'เขา' ชนิดคำ เป็นคำกริยา (v)

- กระทำ [sh8]

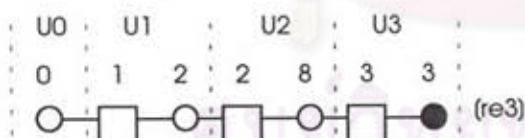
- กระโดดไป สถานะ 8, คอลัมน์ (n) จะได้ การกระทำ : sh3



Next word = 'เขา'

$GSS(2) = [v 'เห็น']$

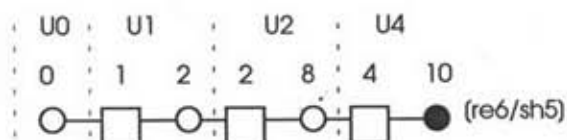
5) เมื่อทำตาม sh3 จะได้ U3



Next word = 'ใน'

$GSS(3) = [n 'เขา']$

6) เมื่อทำตาม re3 จะได้ผลลัพธ์ในขั้นตอน U4 การกระทำในขั้นตอนต่อไปคือ re6/sh5

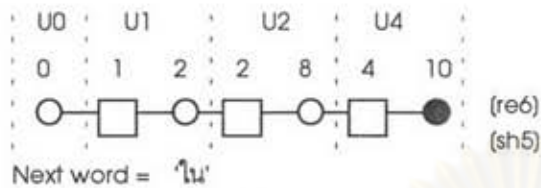


Next word = 'ใน'

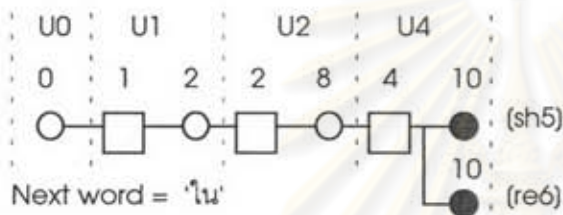
$GSS(4) = [NP (3)]$

7) มาถึงขั้นตอนนี้ จะเห็นว่ามีการกระทำ ที่ต้องทำอยู่ 2 ทาง เราจะต้องทำการแจกแจงส่วนต่อทั้ง 2 ทางเลือก โดยการแยกส่วนปลายของกองซ้อนออกเป็น 2 สาย และแยกการกระทำที่ต้องทำออกไป (รูป ข.)

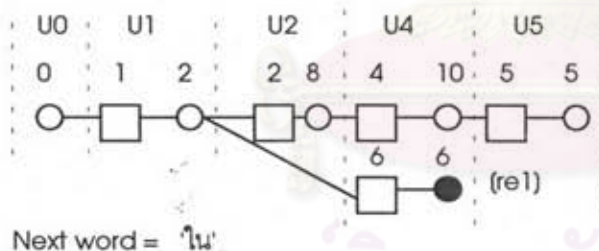
ก)



ข)

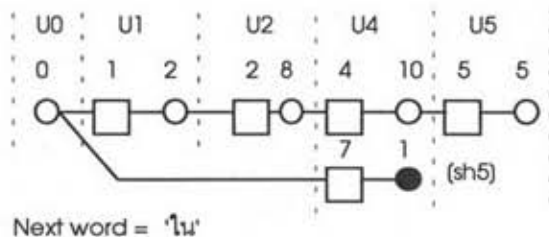


8) หลังจากกระทำ re6 และ sh5 ในข้อที่ 7 แล้ว ต่อมาเราจะให้สถานะที่เป็น การลดทอน ได้รับสิทธิ์ทำงานก่อน เพราะการทำลดทอน คำต่อมา (Next word) จะไม่เลื่อนไป จนกว่าจะพบการปรับเปลี่ยนทั้ง 2 ทางเลือก แล้วจึงทำงานไปพร้อม ๆ กัน



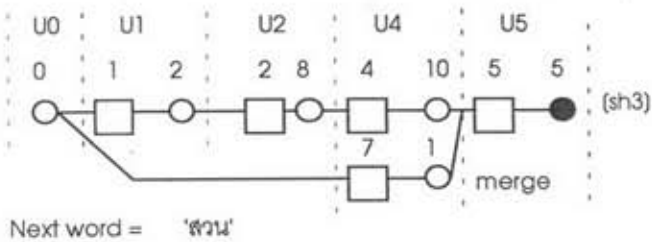
GSS(5) = [p 'ใน'], GSS(6) = [VP (2,4)]

9) หลังจากขั้นตอนที่ 8 จะเห็นว่าทางเลือกทั้ง 2 ทาง ไปที่สถานะ 5 ทั้งคู่ และคำที่เข้ามาคือ 'สวน' เหมือนกัน ขั้นตอนต่อไป เราจึงสามารถรวมทางเลือก 2 ทางที่มีอยู่เข้าด้วยกันได้

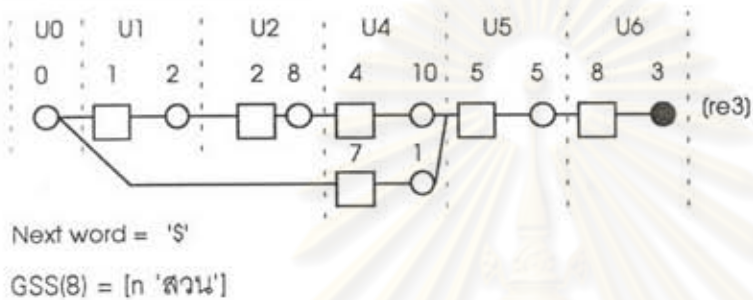


GSS(7) = [S (1,6)]

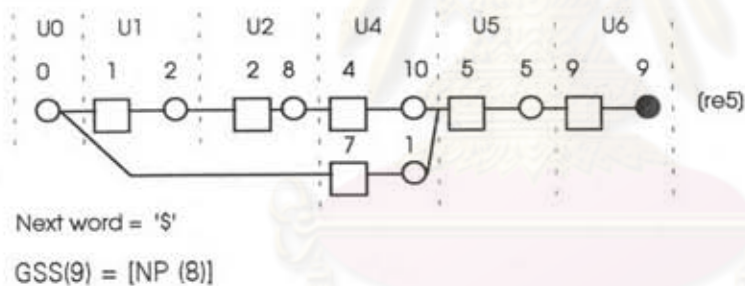
10) เมื่อรวมทางเลือกทั้ง 2 เข้าด้วยกันสถานะต่อไปที่จะทำงานต่อ คือ สถานะที่ 5



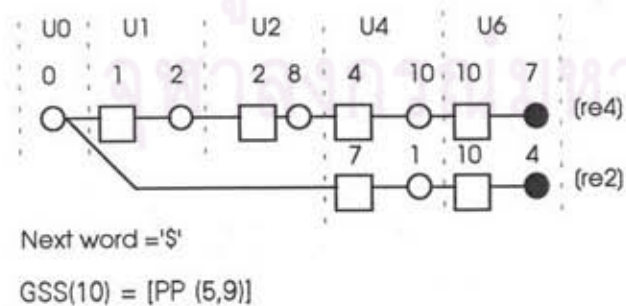
11) เมื่อทำตาม sh3 จะได้ U6



12) ทำ re3 จะได้

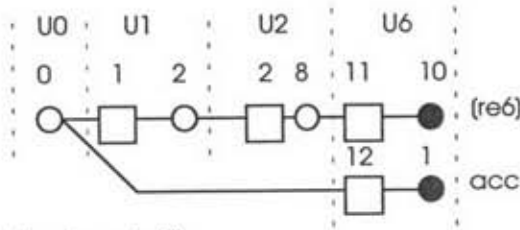


13) เมื่อทำ re5 เสร็จแล้ว จะได้ทางเลือก 2 ทางคือ re4 กับ re2 ทำทั้งสองทางพร้อม ๆ กัน (เป็นการลดทอน ทั้งคู่และค่าที่จะเข้ามาตรวจสอบหมดแล้วจึงทำงานพร้อมกัน ทั้ง 2 สถานะ)



14) ทำการแจงส่วนต่อไปจนจบประโยค

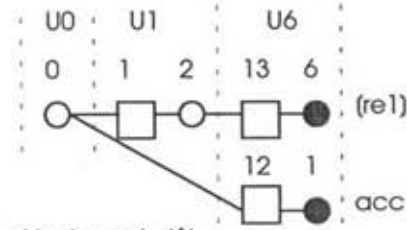
ก)



Next word = '\$'

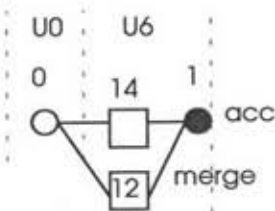
GSS(11) = [NP (4,10)] , GSS(12) = [S (7,10)] , GSS(13) = [VP (2,11)]

ข)



Next word = '\$'

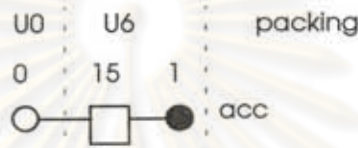
ค)



Next word = '\$'

GSS(14) = [S (1,13)], GSS(15) = [S (7,10) (1,13)]

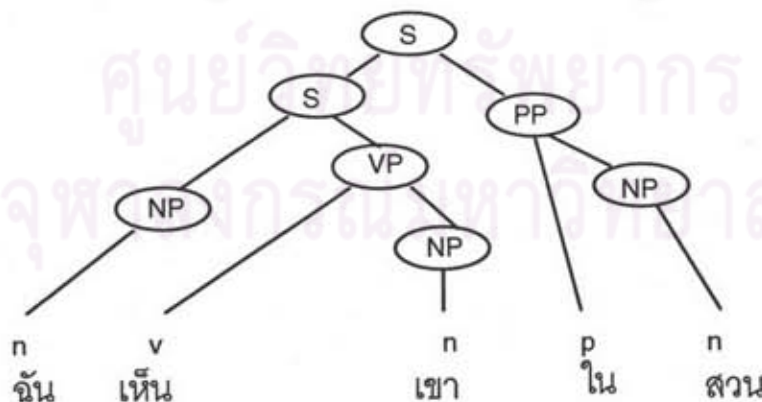
ง)



รูปที่ 3.11 แสดงการแจงส่วนประโยค 'ฉัน เห็น เขา ใน สวน' (ขั้นตอนที่ 1 ถึง 14)

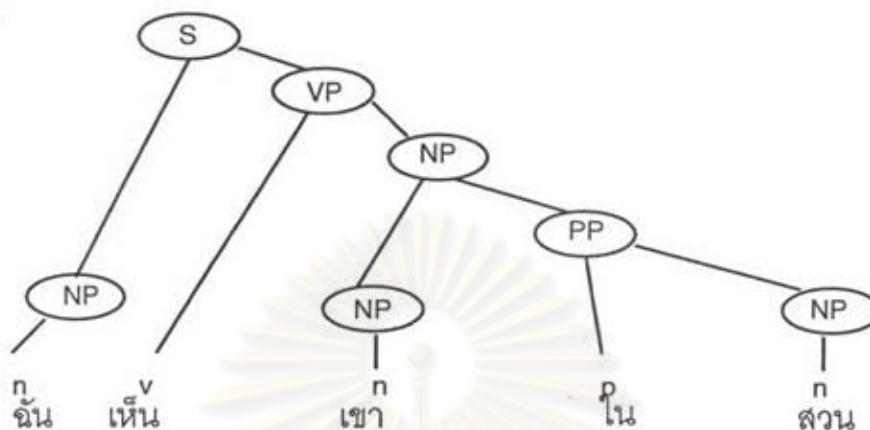
เมื่อนำข้อมูลในกองซ้อนโครงสร้างกราฟ มากระจายเป็น ต้นไม้แจงส่วน (แผนภาพแสดงความสัมพันธ์ระหว่าง องค์ประกอบของประโยคกับคำต่างๆ ในประโยคที่ต้องการวิเคราะห์, parsed tree) เราจะได้ประโยคที่มีโอกาสเป็นไปได้ ตามไวยากรณ์ไม่พ้องบริบทที่กำหนด 2 รูปแบบด้วยกัน ดังแสดงในรูปที่ 3.12

ก)



รูปที่ 3.12 ก) แสดงต้นไม้แจงส่วนในรูปแบบที่ 1 จากการกระจายกองซ้อนโครงสร้างกราฟ

ข)



รูปที่ 3.12 ข) แสดงต้นไม้แจงส่วนในแบบที่ 2 จากการกระจายกองซ้อนโครงสร้างกราฟ

การทำงานของตัวแจงส่วนแบบจีแอลอาร์ที่กล่าวมาในบทนี้ เป็นการทำงานในแบบที่ใช้วิธีการแบบง่าย หรือที่เรียกว่า วิธีการแจงส่วนแบบ Simplified LR (SLR parsing algorithm) ซึ่งเป็นวิธีการที่ใช้ในการทำวิทยานิพนธ์นี้ แต่โดยทั่วไปแล้วยังมีตัวแจงส่วนแบบจีแอลอาร์ในแบบอื่นอีก เช่น Canonical LR (CLR) (Aho, 1986) และ Lookahead-LR (LALR) (Aho, 1986) ถึงแม้ว่าตัวแจงส่วนในสองแบบหลังนี้จะมีประสิทธิภาพสูงกว่า การแจงส่วนแบบ SLR แต่ก็มีคามยุ่งยากมากกว่าในการใช้งานและการใช้ตารางแอลอาร์ที่ซับซ้อนขึ้น อย่างไรก็ตามหลักการ ทำงานโดยทั่วไปแล้วก็ยังคงคล้ายคลึงกัน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย