การเชื่อมต่อฐานข้อมูลโดยใช้ภาษาธรรมชาติ

NATURAL LANGUAGE DATABASE INTERFACE

โดย

นายกิตติ บุญยกิจโณทัย

นายมีชัย เจริญด้วยศีล

นายอมรเทพ เลิศทัศนวงศ์

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตร์บัณฑิต

ภาควิชาวิศวกรรมคอมพิวเตอร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2536

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

U33140

ปริญญานิพนธ์ปีการศึกษา 2536

ภาควิชา วิศวกรรมคอมพิวเตอร์

คณะ วิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

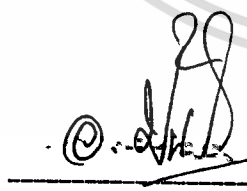
เรื่อง การเชื่อมต่อฐานข้อมูลโดยใช้ภาษาธรรมชาติ

ผู้จัดทำ

นาย กิตติ บุญยกิจโณทัย

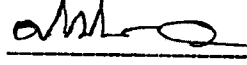
นาย มีชัย เจริญด้วยศีล

นาย อมรเทพ เลิศทัศนวงศ์



อาจารย์ที่ปรึกษา

(ดร. เอื้อน ปิ่นเงิน)



อาจารย์ที่ปรึกษา

(ศ.ดร. ศรีศักดิ์ จามรมาน)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การเชื่อมต่อฐานข้อมูลโดยใช้ภาษาธรรมชาติ

NATURAL LANGUAGE DATABASE INTERFACE

โดย นาย กิตติ บุญยกิจโณทัย

นาย มีชัย เจริญด้วยศีล

นาย อมรเทพ เลิศทัศนวงศ์

อาจารย์ที่ปรึกษา ดร. เอื้อน ปิ่นเงิน

ศ.ดร. ศรีศักดิ์ จามรมาน

ปีการศึกษา 2536

บทคัดย่อ

ปริญญานิพนธ์นี้เป็นการศึกษาภาษาธรรมชาติ และนำมาประยุกต์ใช้งานทางด้านฐานข้อมูล โดยการทำงานจะเป็นลักษณะของการรับอินพุทที่เป็นประโยคภาษาอังกฤษ แล้วนำมาผ่านกระบวนการ 4 ขั้นตอนคือ การแยกคำจากประโยค แล้วนำมาวิเคราะห์ไวยากรณ์ วิเคราะห์ความหมาย และกระบวนการสุดท้ายจะสร้างเป็นภาษากลาง เพื่อใช้ในการดึงข้อมูลจากฐานข้อมูล สำหรับฐานข้อมูลที่ใช้นี้จะเป็นข้อมูลสำหรับการเดินรถไฟ ทั้งนี้ได้กำหนดคำศัพท์ต่าง ๆ ที่ใช้ให้อยู่ในขอบเขตที่จำกัด เพื่อเป็นต้นแบบในการศึกษาเรื่องนี้

ABSTRACT

This thesis is to learn a natural language and apply it for use in database. The method for processing a natural language begins at getting an English sentence for input. Then pass the sentence to 4-step of processing. The processes are lexical analysis (or word seperation), syntactic analysis, semantic analysis and , the last process, code generation to generate meta-language to get data from a database. The database used here is a collection of data for time scheduler of a train. However, the scope of vocabularies used in the program is limitted because it is a prototype program for studying the natural language.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

1 บทนำ 1

2 ทฤษฎีและหลักการของการศึกษาภาษาธรรมชาติ 3

2.1 พื้นฐานภาษาอังกฤษ 3

2.1.1 คำ (word) 3

2.1.2 ส่วนประกอบของวลีของคำนาม (Simple Noun Phrase) 4

2.1.3 ส่วนประกอบของประโยค 6

2.1.4 วลีของบุพบท (Prepositional Phrase) 10

2.1.5 ส่วนเติมเต็ม (Complement) 11

2.1.6 วลีของคุณศัพท์ (Adjective Phrase) 11

2.2 หลักการทำพาร์ซิง (Parsing) ขั้นพื้นฐาน 12

2.2.1 ไวยากรณ์ (Grammar) และโครงสร้างของประโยค 12

2.2.2 ทำอย่างไรถึงจะได้ไวยากรณ์ที่ดีที่สุด 16

2.2.3 วิธีการวิเคราะห์แบบบนลงล่าง (Top-down Parsing method) 17

2.3 ส่วนประกอบและส่วนขยายไวยากรณ์ 22

2.3.1 อ็อกเม็นท์ทรานซิชันเน็ตเวิร์ก (Augmented Transition Network) 22

2.3.2 Useful Feature Systems 28

2.3.3 ส่วนเติมเต็มคำกิริยาและการเซ็ตค่ารีจิสเตอร์ (Verb Complements and Preseting Register) 31

2.4 ไวยากรณ์สำหรับภาษาธรรมชาติ : แฮนดริงมูฟเม็นท์ (Handling Movement) 33

2.4.1 Local Movement* 34

2.4.2 ประโยคคำถามที่ขึ้นต้นด้วย Wh และกระบวนการโฮล (Wh-Questions and the hold Mechanism*) 36

2.5 ซีแมนติคและลอจิคัลฟอร์ม (Semantic and a Logical Form) 40

2.5.1 เหตุผลของการเปลี่ยนไปเป็นลอจิคัลฟอร์ม 40

2.5.2 ชนิดและคุณสมบัติ (Type and Feature) 41

2.5.3 ซีเล็คชันนัลเรสตริกชัน (Selectional Restriction) 44

2.5.4 Case Relation 45

2.5.5 โครงสร้างของคำกิริยา (The Structure of Verbs) 50

2.5.6 ซีแมนติคเน็ตเวิร์ก (Semantics Networks) 53

2.5.7 ลอจิคัลฟอร์ม (The Logical Form) 56

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6 การตีความหมาย (Semantic Interpretation) 61
2.6.1 วิธีการพื้นฐานสำหรับการตีความทางความหมาย 61
2.6.2 อัลกอริทึมการตีความหมาย (The Interpretation Algorithm) 66
2.6.3 ตัวอย่างการกำหนด Case Roles 69
2.6.4 ลำดับชั้นของกฎ (Rule Hierarchies) 72
2.7 Strategies for Semantic Interpretation 76
2.7.1 โดเมนอย่างง่าย (Simple Domain) 76
2.7.2 ไวยากรณ์ทางซีแมนติค (Semantic Grammars) 82
2.7.3 A Simple Interleaved Syntactic and Semantic Analyzer 84
3 การคำนวณและการสร้าง 86
3.1 การออกแบบ ATN (Augment Transition Network) 86
3.2 Flow Chart แสดงการทำงาน 89
3.3 โครงสร้างของ lexicon 91
3.4 โครงสร้างของซีแมนติคดิกชันนารี 92
3.5 ลักษณะของการเก็บข้อมูลในฐานข้อมูล 93
4 การทดลองและผลการทดลอง 94
4.1 การทดลอง 95
4.2 ผลการทดลอง 95
5 บทวิจารณ์และสรุป 97
5.1 บทวิจารณ์ 97
5.2 ปัญหาที่พบ 97
5.3 บทสรุป 97

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ

รูปที่ 1.1 ไดอะแกรมรวมของการประมวลผลภาษาธรรมชาติ 1
รูปที่ 2.1 ระบบสรรพนามในฐานะประธาน 6
รูปที่ 2.2 ระบบสรรพนามในฐานะการแสดงความเป็นเจ้าของ 6
รูปที่ 2.3 ระบบสรรพนามในฐานะของกรรม 6
รูปที่ 2.4 รูปแบบของกิริยาหลักทั้ง 5 7
รูปที่ 2.5 เท็นส์พื้นฐาน 8
รูปที่ 2.6 ลักษณะของโปรเกรสซีฟ 9
รูปที่ 2.7 โครงสร้างแบบต้นไม้แสดงประโยค *John ate the cat.* 12
รูปที่ 2.8 NP อย่างง่าย 15
รูปที่ 2.9 S อย่างง่าย 16
รูปที่ 2.10 ATN ของ S และ NP 18
รูปที่ 2.11 ทรานซิชันเน็ตเวิร์กอย่างง่าย 24
รูปที่ 2..12 เล็กซิคอนอย่างง่าย (assuming third person only) 25
รูปที่ 2.13 ATN อย่างง่าย 25
รูปที่ 2.14 Trace tests and actions used with *The dogs love John.* 27
รูปที่ 2.15 รูปแบบของคำกิริยา 29
รูปที่ 2.16 การวิเคราะห์ประโยคคำถามแบบ wh อย่างง่าย 39
รูปที่ 2.17 Physical Object Hierarchy 42
รูปที่ 2.18 กฎทางซีเมนติก 51
รูปที่ 2.19 ข้อมูลทั่ว ๆ เกี่ยวกับเคสหลัก 52
รูปที่ 2.20 ส่วนของลำดับชั้นของไทพ์ 53
รูปที่ 2.21 แอคชันทั้งหมดมีอนิเมตเอเจนท์ 53
รูปที่ 2.22 ความสัมพันธ์ระหว่างสับพาร์ท 55
รูปที่ 2.23 ซีเมนติคเน็ตเวิร์กสำหรับประโยค *Jack kissed Jill.* 56
รูปที่ 2.24 การแสดงของ *Jack owns car1* 56
รูปที่ 2.25 กฎการตีความหมายสำหรับคำว่า green 61
รูปที่ 2.26 กฎการตีความทางความหมายบางส่วน 65
รูปที่ 2.27 กฎการตีความหมายบางส่วนสำหรับ give 70
รูปที่ 2.28 กฎการตีความหมายบางส่วนของ break 71
รูปที่ 2.29 ระดับชั้นของ RULE BASE 74

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเอกสารทุกครั้งที่มีการนำไปใช้

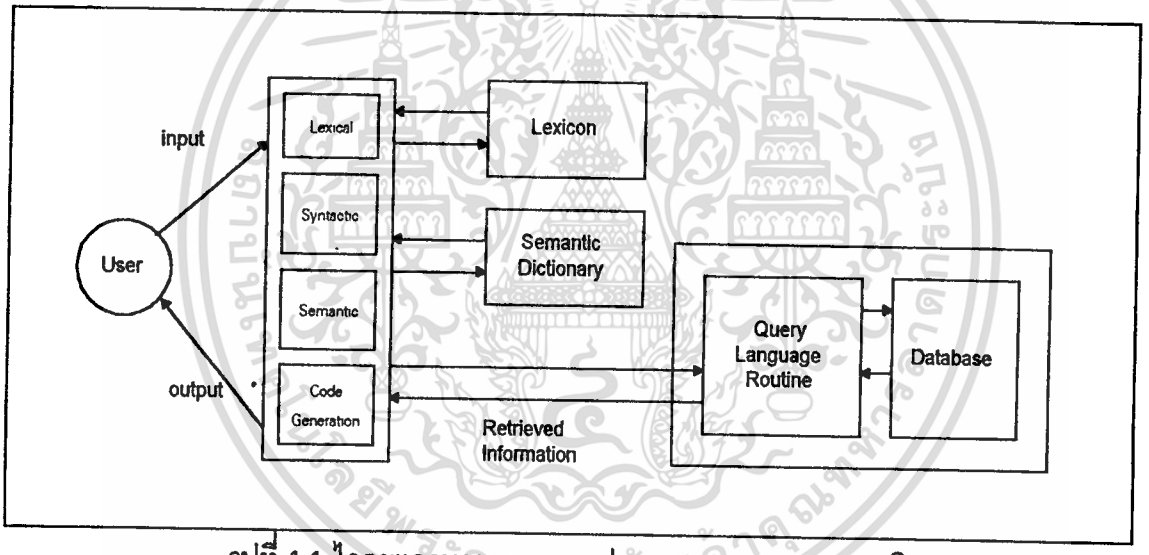
รูปที่ 2.30 กฎของข้อจำกัดสำหรับ burn 74
รูปที่ 2.31 ไวยากรณ์เพิ่มเติม 77
รูปที่ 2.32 โครงสร้างที่เป็นไปได้ของประโยค Can you book a flight to Chicago for me ? 79
รูปที่ 2.33 กฎสำหรับคำกิริยาและคำบุพบท 80
รูปที่ 2.34 กฎสำหรับ wh-terms และ NP 82
รูปที่ 3.1 รูปแสดง ATN ของ S 87
รูปที่ 3.2 รูปแสดง ATN ของ PP 88
รูปที่ 3.3 รูปแสดง ATN ของ NP 88
รูปที่ 3.4 แสดง Flow Chart การทำงานหลัก 89
รูปที่ 3.5 Flow Chart แสดงการทำงานของส่วน Lexical Analysis และ Syntactic Analysis 90
รูปที่ 3.6 Flow Chart แสดงการทำงานของส่วน Semantic Analysis 91
รูปที่ 4.1 หน้าจอแรกของโปรแกรม 94
รูปที่ 4.2 หน้าจอเตรียมพร้อมรับประโยค 94
รูปที่ 4.3 ผลของการทดลองข้อที่ 1 95
รูปที่ 4.4 ผลของการทดลองข้อที่ 2 96
รูปที่ 4.5 ผลของการทดลองข้อที่ 3 96

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

การศึกษาการประมวลผลภาษาธรรมชาติเป็นแขนงหนึ่งของวิชาทางด้านปัญญาประดิษฐ์ (AI:Artificial Intelligent) ซึ่งมีการศึกษากันอย่างกว้างขวางในต่างประเทศ และปัจจุบันงานทางด้านปัญญาประดิษฐ์ในประเทศไทยเอง ได้กลายมาเป็นสิ่งที่ผู้ศึกษาทางคอมพิวเตอร์ให้ความสนใจในการนำมาประยุกต์ใช้งานในด้านต่าง ๆ โดยเฉพาะงานทางด้านสารสนเทศมากขึ้นเป็นลำดับ ซึ่งปริญญานิพนธ์ที่ทำนี้เป็นส่วนหนึ่งของงานทางด้านนี้ด้วย โดยจะนำภาษาธรรมชาติมาใช้ประมวลประโยคภาษาอังกฤษที่เป็นอินพุทเพื่อดึงข้อมูลจากฐานข้อมูลที่เราต้องการ

การประมวลผลภาษาธรรมชาติมีกระบวนการ 4 ขั้นตอนซึ่งสามารถแสดงด้วยบล็อคไดอะแกรมได้ดังรูป



รูปที่ 1.1 ไดอะแกรมรวมของการประมวลผลภาษาธรรมชาติ

การวิเคราะห์เล็กซ์ซิคัล (lexical analysis) คือ ส่วนที่รับประโยคภาษาธรรมชาติแล้วแยกคำต่าง ๆ ในประโยคเป็นโทเคน จากนั้นจะส่งไปให้ส่วนการวิเคราะห์ซินแทคติค (syntactic analysis)

การวิเคราะห์ซินแทคติค คือ ส่วนที่รับโทเคนจากขั้นตอนการวิเคราะห์เล็กซ์ซิคัลและทำการตรวจว่า แต่ละโทเคนที่นำมาต่อเชื่อมกันเป็นประโยคนั้นถูกต้องตามหลักไวยากรณ์ที่กำหนดไว้หรือไม่ จากขั้นตอนนี้เราจะได้ซินแทคติคฟอร์ม (syntactic form) ซึ่งจะเป็นรูปแบบที่ผ่านการตรวจความถูกต้องทางไวยากรณ์แล้ว

การวิเคราะห์ซีแมนติค (semantic analysis) คือ ส่วนที่รับซินแทคติคฟอร์มจากขั้นตอนการวิเคราะห์ซินแทคติคเพื่อนำมาตรวจสอบความถูกต้องตามความหมายของประโยคนั้น ซึ่งขั้นตอนนี้จะสร้างโลจิคัลฟอร์ม (logical form) เพื่อส่งต่อไปยังส่วนต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โค้ดเจนเนอเรชั่น (code generation) คือ ส่วนที่ทำการแปลงโลจิคัลฟอร์มไปเป็นรูปแบบของภาษากลางเพื่อใช้ในการดึงข้อมูลที่ต้องการจากฐานข้อมูล

สำหรับฐานข้อมูลที่จะใช้นั้น เราได้คำนึงถึงปัญหาและความเป็นไปได้ของตัวปริญญานิพนธ์ที่จะเกิดขึ้น ถ้าหากใช้ชนิดและขอบเขตของฐานข้อมูลที่ไม่เหมาะสมอาจทำให้เกิดความซับซ้อนในตัวโปรแกรมเกินไปจนไม่อาจจะยอมรับได้ ดังนั้นเราจึงตัดสินใจใช้ตารางเดินรถไฟเป็นฐานข้อมูล และจำกัดขอบเขตของข้อมูลและคำศัพท์เฉพาะที่เกี่ยวกับเวลาการเดินรถไฟเท่านั้น เนื่องจากจุดมุ่งหมายของปริญญานิพนธ์ต้องการที่จะทำให้เป็นต้นแบบเพื่อศึกษาและพัฒนาต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีและหลักการของการศึกษาภาษาธรรมชาติ

2.1 พื้นฐานภาษาอังกฤษ

ในหัวข้อนี้เราจะพูดถึงหลักการภาษาอังกฤษที่สำคัญ ๆ ทั่ว ๆ ไป เพื่อเป็นพื้นฐานสำหรับการสร้างโปรแกรมเพื่อวิเคราะห์ภาษาธรรมชาติต่อไป

2.1.1 คำ (word)

หน่วยพื้นฐานที่สุดของโครงสร้างภาษาก็คือ คำ เราจะพูดถึงการเปลี่ยนรูปแบบของคำ (morphology) ซึ่งเป็นการศึกษาการสร้างคำจากองค์ประกอบพื้นฐานอย่างหยาบ ๆ โดยองค์ประกอบนี้เราจะเรียกว่า รากศัพท์ (root) ซึ่งสามารถจะนำส่วนอื่น ๆเข้ามาประกอบกันเป็นคำที่แตกต่างกันได้ เช่น คำว่า friend ซึ่งเป็นรากศัพท์ของคำคุณศัพท์ (adjective)*friendly* ซึ่งเป็นการเพิ่มซัฟฟิกซ์ (suffix) *-ly* และคำ ๆ นี้สามารถที่จะแปลงต่อไปเป็น คำคุณศัพท์ *unfriendly* ซึ่งได้จากเพิ่มพรีฟิกซ์ (prefix) *un-* หรือถ้าหากเราจะพิจารณาคำที่ซับซ้อนกว่านี้ เราจะได้คำนาม *friendliness* ซึ่งมาจากรูปแบบคำคุณศัพท์ *friendly* นั่นเอง

เราอาจจะแบ่งคำได้เป็นประเภทต่าง ๆ ตามลักษณะของเหตุการณ์ คือ ประเภทแรก จะเกี่ยวข้องกับคำที่ทำการสนับสนุนวลี (Phrase) และประเภทที่สองจะเกี่ยวข้องกับโครงสร้างไวยากรณ์ (grammar) ของประโยคโดยตรงที่คำนั้นทำหน้าที่อยู่เช่น เราอาจจะจัดให้คำนาม (noun) ใช้เพื่อระบุถึงวัตถุ แนวความคิดหรือสถานที่ที่กำลังกล่าวถึงอยู่ และคำกิริยาวิเศษณ์ (adverb) ซึ่งคำเหล่านั้นจะเป็นการทำให้วัตถุ แนวความคิดหรือสถานที่นั้น มีความหมายชัดเจนขึ้น เช่น คำว่า *green* อาจจะเป็นคำกิริยาวิเศษณ์ และ *book* เป็นคำนาม ถ้าเกิดมันอยู่ในรูปของวลี *the green book* และ *greens books* แต่ว่ามันไม่เป็นการง่ายอย่างที่เห็น เนื่องจาก *green* อาจจะเป็นคำนาม เช่นในประโยค *That green is lighterthan the other.* และ *book* อาจจะเป็นคำขยาย เช่นในวลี *the book worm* ซึ่งจริง ๆ แล้ว ดูเหมือนว่าคำนามทุกตัวจะสามารถทำหน้าที่เป็นส่วนขยาย(modifier) ได้ในบางกรณี บางทีชนิดของคำ อาจจะมีการเหลื่อมล้ำกัน แต่ว่าจะใช้แทนกันไม่ได้ ในทุก ๆ เหตุการณ์ ตัวอย่างเช่น เราจะพิจารณาคำที่จะมาทำให้ประโยค *It's so ...* สมบูรณ์ เราอาจจะใช้ว่า *It's so green, It's so hot, It's so true* หรืออื่น ๆ แต่สำหรับคำว่า *book* ซึ่งมันอาจจะหน้าที่เป็นคำขยายได้ แต่เราไม่สามารถใช้ *It's so book* ได้เลย ดังนั้นจึงมีคำขยาย 2 ชนิดที่แบ่งแยกกันอย่างเห็นได้ชัดเจน คือ ส่วนขยายคำคุณศัพท์(adjective modifiers) และส่วนขยายคำนาม (noun modifier)

ตอนนี้เรามาพิจารณากรณีที่ คำคุณศัพท์ ที่สามารถใช้เป็นคำนาม อย่างเช่น *the green*

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารนี้ทุกครั้งที่มีการนำไปใช้

แต่ว่าคำคุณศัพท์ทุกคำไม่สามารถทำเช่นนี้ได้ เช่น *the hot* หรือ *the angry* จะแสดงถึงกลุ่มของของร้อนหรือกลุ่มของคนที่กำลังโกรธตามลำดับ ซึ่งไม่เหมือนกับ *the green* ที่แสดงถึงสิ่งของสีเขียว ๆ ชิ้นเดียว ดังนั้นเราก็สามารถที่จะแบ่งแยกคำคุณศัพท์ออกเป็นสองประเภทย่อย ก็คือคำคุณศัพท์ที่ใช้อธิบายแนวคิด (concept) ของคำที่มันขยาย ส่วนคำคุณศัพท์ *hot* หรือ *angry* ใช้แสดงถึงลักษณะของคำที่มันขยายอยู่

2.1.2 ส่วนประกอบของวลีของคำนาม (Simple Noun Phrase)

วลีของคำนาม (ต่อไปจะขอกล่าวย่อๆ ว่า NP) จะใช้แสดงถึงสิ่งต่าง ๆ เช่น วัตถุ แนวความคิด สถานที่ เหตุการณ์ คุณภาพ ฯลฯ NP ที่ง่ายที่สุด ก็คือ สรรพนาม (pronoun) ตัวเดียว อย่างเช่น *he, she, they, you, me, it, I, we* ฯลฯ ซึ่งสามารถจะ

แสดงถึงวัตถุที่จับต้องได้ : *It* under the rug.

แสดงถึงเหตุการณ์ : Once I opened the door, I regretted *it* for months.

แสดงถึงลักษณะ : He was so angry ,but he didn't show *it*.

สรรพนามมักจะไม่มีคำขยายใด ๆ ยกเว้นรูปแบบที่ไม่ค่อยพบบ่อย ดังประโยคต่อไปนี้

He who hesitates is lost.

รูปแบบของ NP อีกรูปแบบหนึ่งประกอบไปด้วยชื่อหรือคำเฉพาะ เช่น *John* หรือ *Rochester* ซึ่งคำเหล่านี้จะเห็นว่ามีการใช้ตัวอักษรตัวใหญ่นำหน้าคำ ชื่อนี้อาจจะประกอบด้วยคำหลาย ๆ คำ อย่างเช่น *The New York Times ,Stratford on Avon* และ *King Mongkut's Institute of Technology Ladkrabang*

รูปแบบที่เหลือจะอยู่ในรูปของคำหลักซึ่งเรียกว่า head และคำอื่น ๆ ที่แสดงลักษณะของ head ซึ่ง head ของ NP มักจะเป็นคำนาม ซึ่งคำนามแบ่งเป็น 2 ประเภท ได้แก่ นามที่นับได้ (count noun)ซึ่งใช้อธิบายลักษณะเฉพาะบางอย่างของวัตถุหรือกลุ่มของวัตถุ และนามที่นับไม่ได้ (mass noun) ซึ่งใช้อธิบายองค์ประกอบต่าง ๆ

นามที่นับได้จะอยู่ในรูปของเอกพจน์ (single form) และพหูพจน์ (plural form) ส่วนนามนับไม่ได้ จะอยู่ในรูปของการบอกจำนวนอย่างคร่าว ๆ เช่น *some water,somewheat* หรือ *some sand* ถ้าเราพยายามที่จะนับคำเหล่านี้ เราจะต้องเปลี่ยนความหมายของมัน อย่างเช่น *some wheat* จะหมายถึง จำนวนของเมล็ดข้าวสาลีจำนวนหนึ่ง, *all wheats* จะแสดงถึงเมล็ดข้าวสาลีทั้งหมดทุกชนิด และ*one wheat* แสดงถึงเมล็ดข้าวสาลีชนิดเดียว ซึ่งมันไม่ได้หมายถึงเมล็ดข้าวสาลีหนึ่งเมล็ด นอกจากนั้นนามที่นับไม่ได้ยังใช้ในความหมายของสิ่งนั้น ๆ โดยที่ไม่ต้องใช้รูปพหูพจน์ อย่างเช่น เราจะใช้ประโยค *Water is neccessary for live.* และจะไม่ใช้ *Waters are neccessary for*

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

live.

NP อาจจะมีสเป็คซิไฟเออร์ (Specifier) และควอลิไฟเออร์ (Qualifier) อยู่หน้า head

ควอลิไฟเออร์ จะอธิบายถึงประเภทของวัตถุ (objects) ซึ่งระบุโดย head ในขณะที่ สเป็คซิไฟเออร์ จะบอกถึงจำนวนของวัตถุที่จะถูกอธิบาย วัตถุจะถูกอธิบายอย่างไรนั้นจะขึ้นอยู่กับผู้พูดและผู้ฟัง

สเป็คซิไฟเออร์ จะประกอบขึ้นมาจากคำประเภทต่าง ๆ ดังต่อไปนี้

- ควอนติไฟเออร์ (Quantifiers) คือ คำที่ใช้แสดงจำนวนของวัตถุที่เอ่ยถึง เช่นคำว่า *is,any, all, no, every, many, few, some, neither, both, half* ฯลฯ

- ดีมอนสเตตีฟ (Demonstratives) เป็นคำที่แสดงว่าวัตถุนั้นเกี่ยวข้องกับผู้พูดอย่างไร เช่น *this* หรือ *that,* หรือเป็นคำที่แสดงว่าวัตถุนั้นมีเพียงสิ่งเดียวในเหตุการณ์นั้น อย่างเช่น *the* หรือ *a* และอาจจะหมายถึงว่าผู้ฟังต้องการรายละเอียดของวัตถุนั้นหรือไม่ เช่น *which, what* เป็นต้น

- พอสเซสซีฟ (Possessives) เป็นคำที่ประกอบขึ้นมาจาก NP ชุดอื่นที่ใช้แสดงความเป็นเจ้าของของวัตถุที่อยู่ใน NP ที่กำลังพูดถึง

- ออดินัล (Ordinals) เป็นคำที่อธิบายวัตถุตามลักษณะของการเรียงลำดับอย่างเช่น *first, second, third...*

- คาร์ดินัล (Cardinals) เป็นคำที่แสดงจำนวนของวัตถุ เช่น *one, two, three...*

พอสเซสซีฟจะอยู่ในรูปของการเติมซัฟฟิกซ์ *-'s* ให้กับคำนามหรือชื่อเฉพาะเช่น *King Mongkut's Institute, Joy's Car* สรรพนามแสดงความเป็นเจ้าของ (Possessive Pronouns) จะมีรูปแบบของตัวมันเองเช่น *my, your, our, their...* จะมีข้อยกเว้นเล็กน้อย คือ NP พื้นฐานจะมีเพียงดีมอนสเตตีฟ 1 คำ พอสเซสซีฟ 1 คำ หรือ ควอนติไฟเออร์ 1 คำ แต่จะไม่มีมากกว่า 1 คำ

ควอลิไฟเออร์ ใน NP จะอยู่หลังสเป็คซิไฟเออร์ ถ้ามีสเป็คซิไฟเออร์ และอยู่ก่อน head ซึ่งควอลิไฟเออร์ นี้ประกอบด้วยคำคุณศัพท์และคำนามซึ่งทำหน้าที่เป็นคำขยายดังนี้

- คำคุณศัพท์ เป็นคำที่แสดงลักษณะของวัตถุ เช่น *angry* จะเป็น คำคุณศัพท์ที่แสดงลักษณะของการโกรธอะไรบางอย่าง

- ส่วนขยายคำนาม (Noun Modifier) เป็นนามนับได้หรือนับไม่ได้ก็ได้ ที่ขยายนามตัวอื่น เช่น *the cook book*

พิจารณารูปแบบที่ต่างกันของคำนาม ทั้ง 2 รูปแบบ คือ เอกพจน์และพหูพจน์ สรรพนามจะเป็นคำนามที่ขึ้นอยู่กับบุคคล (บุรุษที่ 1 2 หรือ 3) และเพศ (เพศชาย เพศหญิง และไม่มีเพศ) ดังรูปที่ 2.1, 2.2 และ 2.3 ต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	Person		
Number	First	Second	Third
			he (masculine)
singular	I	you	she (feminine)
			it (neuter)
plural	we	you	they

รูปที่ 2.1 ระบบสรรพนามในฐานะประธาน

	Person		
Number	First	Second	Third
			his
singular	my	your	her
			its
plural	our	your	their

รูปที่ 2.2 ระบบสรรพนามในฐานะการแสดงความเป็นเจ้าของ

	Person		
Number	First	Second	Third
			him
singular	me	you	her
			it
plural	us	you	them

รูปที่ 2.3 ระบบสรรพนามในฐานะของกรรม

2.1.3 ส่วนประกอบของประโยคแบบง่าย ๆ

ประโยค (ในที่นี้จะเรียกย่อ ๆ ว่า "S") บอกเล่า คำถาม หรืออธิบายบางสิ่งบางอย่างบนโลกนี้ ดังนั้นคุณอาจจะบอกว่า เรื่องบางเรื่องเป็นเรื่องจริง หรืออาจจะถามว่าเรื่องนี้เป็นเรื่องจริงหรือเปล่า หรืออาจจะสั่งคนใดให้ทำอะไรก็ตาม ลักษณะของประโยค (mood) ขึ้นอยู่กับประเภทของกริยา ซึ่งลักษณะของประโยคพื้นฐาน 4 ประการแสดงได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Mood	Example
declarative (or assertion)	The cat is sleeping.
yes/no question	Is the cat sleeping?
wh-question	What is sleeping? or
	Which cat is sleeping?
imperative (or command)	Shoot the cat!

กลุ่มของคำกริยา (The verb group)

ประโยคประกอบด้วยส่วนประกอบ 2 ส่วน คือ ส่วน NP ที่ทำหน้าที่เป็นประธานของประโยค และส่วนของ กลุ่มคำกิริยา

กลุ่มคำกิริยาประกอบด้วย กริยาหลัก (head verb) และส่วนกิริยาช่วย (Auxiliary Verb) ซึ่งอาจจะมีหรือไม่มีก็ได้ กริยาหลักอาจจะอยู่ในรูปแบบใดรูปแบบหนึ่งใน 5 รูปแบบดังรูป

Form	Examples	Example Uses
infinitive	hit, cry,	*Hit* the ball!
	go, be	I want to *go.*
simple present	hit, cries,	The dog *cries* every day.
	go, am	I *am* thirsty.
simple past	hit, cried,	I *was* thirsty.
	went, was	I *went* to the store.
present participle	h i t t i n g ,	I'm *going* to the store.
	crying,	
	going, being	*Being* the last in line aggravates me.
past participle	hit, cried,	I've *been* there before.
	gone, been	The cake was *gone.*

รูปที่ 2.4 รูปแบบของกิริยาหลักทั้ง 5

กิริยาช่วย ปรากฏอยู่ในรูปแบบต่าง ๆ คือ รูปแบบของ *be, do* และ *have* รวมถึงกลุ่มของ model auxiliary เช่น *will, can, could* และอื่น ๆ กิริยาช่วยและรูปแบบของกริยาหลักที่มารวมกันจะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทำให้เกิดแบบต่าง ๆ ของเท็นส์ (tense) ขึ้น เท็นส์จะอยู่ในรูปแบบที่ซ้ำซ้อน แต่รูปต่อไปนี้จะแสดงเท็นส์พื้นฐาน

Tense	Structure of Verb Group	Example
simple present	simple present	He walks to the store.
simple past	simple past	He walked to the store.
simple future	*will* + infinitive	He will walk to the store.
past perfect	*have* in present (or present/past) + past participle	He has walked to the store.
future perfect	*will* + *have* in infinitive + past participle	I will have walked to the store.
pluperfect	*have* in past + (or past/past) past participle	I had walked to the store.

รูปที่ 2.5 เท็นส์พื้นฐาน

นอกจากนี้ กริยาอาจจะอยู่ในลักษณะของโปรเกรสซีฟ (Progressive Aspect) ซึ่งโปรเกรสซีฟเท็นส์แต่ละชนิดสร้างขึ้นมาจากโครงสร้างของเท็นส์แบบธรรมดาของ verb to be ตามด้วย present participle ดังรูป

Tense	Structure	Example
present progressive	*be* in present + present participle	He is walking.
past progressive	*be* in past + present participle	He was walking.
future progressive	*will* + *be* in infinitive + present participle	He will be walking.
past perfect progressive	*have* in present + *be* in past participle	He has been walking.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	+ present participle	
future perfect	*will* + *have* in present	He will have been walking.
progressive	+ *be* as past participle	
	+ present participle	
luperfect progressive	*have* in past +	He had been walking.
	be in past participle	
	+ present participle	

รูปที่ 2.6 ลักษณะของโปรเกรสซีฟ

กลุ่มของคำกิริยายังจะต้องขึ้นอยู่กับรูปแบบของประธาน ซึ่งเป็น NP ที่นำหน้า กริยา นั้นมา ดังตัวอย่างต่อไปนี้

	First	Second	Third
singular	I *am*	you *are*	he *is*
	I *walk*	you *walk*	she *walks*
plural	we *are*	you *are*	they *are*
	we *walk*	you *walk*	they *walk*

Transitivity และ Passives ในหัวข้อที่แล้ว เราได้พูดถึงรูปแบบทั่วไปของประโยคซึ่งประกอบด้วย NP ตามด้วยกลุ่มของคำกิริยา เช่น *Jack laughed, he will have been running.* และประโยค *The three largest balloons were rising.* กริยาหลักหลายตัวไม่สามารถที่จะอยู่ในรูปแบบเช่นนั้นได้ เนื่องจากมันต้องการ NP อีกชุดหนึ่ง ตามหลังกลุ่มของคำกิริยาเช่น เราไม่สามารถจะใช้ *Jack found* เนื่องจากเราไม่รู้ว่า Jack พบอะไร แต่เราสามารถที่จะใช้ *Jack found a key.* เพื่อทำให้ประโยคสมบูรณ์แต่กริยาอีกชนิดหนึ่งก็ไม่สามารถที่จะมี NP ชุดที่สองนี้ได้ เช่น เราไม่สามารถที่จะใช้ *Jack laughed a key.* เนื่องจากไม่มีความหมายใด ๆ กริยาที่ต้องการเฉพาะ NP ตัวแรก (ซึ่งก็คือ ประธาน) เรียกว่าอกรรมกิริยา (Intransitive verb) แต่กริยาที่ต้องการ NP ชุดที่สอง (กรรม) เรียกว่าสกรรมกิริยา (Transitive verb) สกรรมกิริยาสามารถที่จะมีรูปแบบของกลุ่มของคำกิริยาอีกชนิดหนึ่งเรียกว่ารูปแบบพาสซีฟ (passive form) ซึ่งใช้ Verb to be ตามด้วย past participle ในรูปของพาสซีฟ กรรมจะมาทำหน้าที่เป็นประธานของประโยค ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

033140

Active Sentence	Passive Sentence
Jack saw the ball.	The ball was seen by Jack.
I will find the clue.	The clue will be found by me.
Jack hit me.	I was hit by Jack.

ข้อควรสังเกตเกี่ยวกับเรื่องรูปแบบของพาสซีฟก็คือ การนำสรรพนามที่เป็นกรรมขึ้นมาเป็นประธาน อย่างเช่น จะต้องใช้ *I was hit.* ไม่ใช่ *Me was hit.* มีกริยาบางประเภทที่อนุญาติให้มี NP 2 ชุดตามหลังมันในประโยคนั้น เช่น *Jack gave Sue a book.* หรือ *Jack found me a key.* ในลักษณะนี้ มันจะมีประโยคที่เทียบเท่ากับมัน ซึ่งมี NP อยู่ติดกับกลุ่มของคำกิริยา เรียก NP ชุดนี้ว่ากรรมรอง (Indirect object) ตามหลังคำบุพบท (preposition) ในประโยคแบบนั้น NP อีกชุดหนึ่งจะทำหน้าที่เป็นกรรมที่แท้จริงของโครงสร้างประโยคที่ได้กล่าวมาแล้ว ซึ่งจะเรียกมันว่ากรรมตรง (direct object) ดังตัวอย่างต่อไปนี้

Sentence with Indirect	Object Equivalent Form
Jack gave *Sue* a book.	Jack gave a book *to Sue.*
Jack found *me* a book.	Jack found a book *for me.*

พาร์ติเคิล (Particle)

รูปแบบของคำกริยาบางประเภท จะอยู่ในรูปของคำกริยาและคำที่เพิ่มขึ้นมาซึ่งเรียกว่าพาร์ติเคิล ซึ่งพาร์ติเคิลส่วนมากจะไปซ้ำกับคำประเภทที่เป็นบุพบท เช่นคำว่า *up, out, over, in* ซึ่งกริยาต่าง ๆ เช่น *look, take* หรือ *put* จะต้องใช้ควบคู่กับพาร์ติเคิล ยกตัวอย่างเช่น *look up, look out, look over* และอื่น ๆ ในประโยคบางประโยค ถ้าเราใช้พาร์ติเคิลหรือคำบุพบทจะให้ความหมายที่แตกต่างกัน ดังเช่น *look over the paper* ถ้าเราให้ over เป็นพาร์ติเคิล (จะได้กริยา look over) ประโยคนี้จะหมายถึง อ่านหนังสือพิมพ์ แต่ถ้าให้ over เป็นคำบุพบท ประโยคนี้จะหมายถึง การมองสิ่งของข้างหลังหรือเหนือหนังสือพิมพ์

2.1.4 **วลีของบุพบท** (Prepositional Phrase)

วลีของบุพบทในรูปแบบที่ง่ายที่สุดจะประกอบด้วยคำบุพบท ตามด้วย NP ที่ทำหน้าที่เป็นกรรมของคำบุพบทนั้น เช่น ประโยค *I gave the book to Sue.* วลีของบุพบทก็คือ to Sue จะทำหน้าที่ติดต่อโครงสร้างของกริยา *give* เพราะมันระบุว่าให้ใคร ยังมีรูปแบบของวลีของบุพบทที่ซับซ้อนมากกว่านี้ อย่างเช่น *out of the oven* จะเห็นว่ามีคำบุพบท 2 ตัว หรือบางทีเราอาจจะเห็นคำบุพบทอยู่แยกกับกรรมของมัน อย่างเช่น *we saw the river that Sam Patch died in.* สุด

เอกสารนี้มีไว้สำหรับใช้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ท้ายนี้เราอาจจะพิจารณาคำต่าง ๆ ต่อไปนี้ว่ามันเป็นวลีของบุพบทได้โดยไม่มี คำบุพบท เลย คือ *here, their, afterward, later, upstair, today* หรือ *home* เนื่องจากคำเหล่านี้สามารถที่จะใช้ในที่ที่วลีของบุพบทใช้ได้

2.1.5 **ส่วนเติมเต็ม** (Complement)

ส่วนเติมเต็มจะอยู่ในรูปต่าง ๆ โดยที่จะทำหน้าขยาย head word ของ NP ให้มีความหมายชัดเจนขึ้น

2.1.6 **วลีของคุณศัพท์** (Adjective Phrase)

คำคุณศัพท์สามารถที่จะเป็นตัวเดี่ยวอยู่ใน NP เช่น *the Heavy book* หรือมีอยู่หลายตัวใน NP เช่น *the (heavy) (red) book* ซึ่งจะเห็นว่าคำคุณศัพท์เหล่านี้จะขยาย Head นอกจากนั้นวลีของคุณศัพท์ยังสามารถที่จะหน้าที่เป็น ส่วนเติมเต็มของกริยาพวก *be,look* หรือ *see* ดังประโยค *The book is red.* หรือ ประโยค *The book look heavy.* วลีของคุณศัพท์ อาจจะอยู่ในรูปที่ซับซ้อนกว่าที่กล่าวมาแล้ว คืออาจจะมีตัวบอกลักษณะมาขยายคำคุณศัพท์ ตัวบอกลักษณะดังกล่าวอาจจะเป็นคำคุณศัพท์ตัวอื่น อย่างเช่น *Bright red* ในวลี *the (bright red) book.* หรืออาจจะเป็นคำกิริยาวิเศษณ์อย่างเช่น *compeatly red* และ *terribly heavy* หรืออาจจะเป็นการปฏิเสธ (Negative) เช่น *not red* คำคุณศัพท์อาจอยู่ในรูปของวลีของบุพบท ช่น *hungry for learning* และ *yellow with age* หรือเป็นประธาน เช่นประโยค *careful to step on every crack, afraid thatjoy would lose, easy to get along with,* และ *proud to be a frog.*

คำคุณศัพท์อีกรูปแบบหนึ่งซึ่งเรียกว่าดีกรีคลอส (Degree Clause) จะเป็นการบอกลักษณะที่อยู่ในรูปแบบที่ตายตัวอย่างเช่นในวลี *as slow as a dead horse* และ *so tired that he passed out*

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 หลักการทำพาร์ซิง (Parsing) ขั้นพื้นฐาน

การที่เราจะรู้ว่าประโยคถูกตามโครงสร้างหรือไม่ เราจะต้องพิจารณาถึงองค์ประกอบ 2 ประการต่อไปนี้ คือ

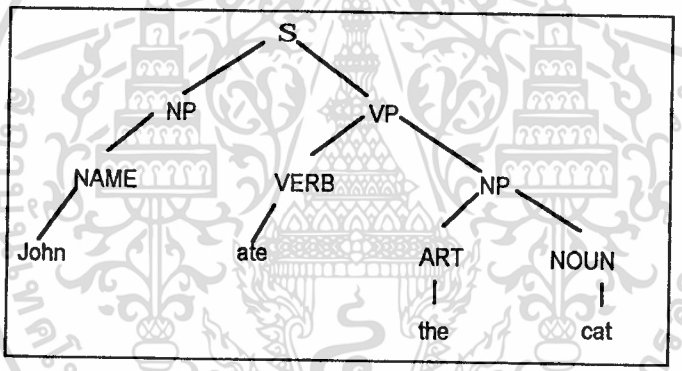
- หลักไวยากรณ์ (grammar) ซึ่งเป็นข้อบังคับของโครงสร้างของภาษานั้น ๆ

- เทคนิคการทำพาร์ซิง (Parsing Technique) คือการวิเคราะห์ประโยค โดยทดสอบโครงสร้างของประโยคว่าถูกต้องตามหลักไวยากรณ์หรือไม่

2.2.1 ไวยากรณ์ (Grammar) และโครงสร้างของประโยค

การแสดงโครงสร้างของประโยค

วิธีที่แสดงโครงสร้างของประโยคที่ง่ายที่สุดคือการใช้รูปแบบต้นไม้ (Treelike) ซึ่งจะแสดงประโยคว่าแบ่งออกเป็นส่วนประโยคหลัก ๆ อะไรบ้างและส่วนหลักเหล่านั้นแยกย่อยออกเป็นอะไรอีก พิจารณาตัวอย่างประโยค *Thep eat cat* จะแสดงได้ดังรูปต่อไปนี้



รูปที่ 2.7 โครงสร้างแบบต้นไม้แสดงประโยค *John ate the cat.*

ซึ่งสามารถจะอ่านได้ดังนี้

ประโยค (S) ประกอบไปด้วย NP และ VP ซึ่ง NP ก็คือชื่อ *John* ส่วน VP นั้นประกอบไปด้วยคำกิริยา *eat* ตามด้วย NP ซึ่งมี ART (คือ Article) *the* และคำนาม *cat* ซึ่งแสดงได้ดังนี้

```
(S (NP (NAME John))
   (VP (VERB ate)
       (NP (ART the)
           (NOUN cat))))
```

Context-Free Grammar

รีไรท์รูล (Rewrite Rule) ใช้ในการอธิบายว่าโครงสร้างต้นไม้ใดสามารถจะมีได้ในไวยากรณ์นั้น พิจารณารูปข้างบน จะประกอบไปด้วยไวยากรณ์ดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

S ← NP+VP

VP ← VERB+NP

NP ← NAME

NP ← ART+NOUN

กฎดังกล่าวอธิบายได้ดังนี้

S ประกอบไปด้วย NP ตามด้วย VP, VP ประกอบไปด้วย VERB (คำกิริยา) ตามด้วย NP, NP ประกอบไปด้วย NAME (ชื่อ) หรือประกอบไปด้วย ART ตามด้วย NOUN (คำนาม)

ไวยากรณ์ประกอบไปด้วยกฎซึ่งอยู่ในรูป $<symbol> \leftarrow <symbol>_1 \ldots <symbol>_n$ โดยที่ $n \geq 1$ ซึ่งเรียกว่าคอนเท็กซ์ฟรีแกรมมาร์(Context-Free Grammar - CFG) CFG เป็นไวยากรณ์ที่สำคัญมากเนื่องจากเหตุผล 2 ประการคือ รูปแบบของมันสามารถที่จะอธิบายโครงสร้างของภาษาได้เกือบทั้งหมด และยังมีขอบเขตเพียงพอที่จะสร้างพาร์เซอร์ (Parser) เพื่อวิเคราะห์โครงสร้างของประโยคได้ ซึ่งไวยากรณ์สามารถแยกเป็นสัญลักษณ์ต่าง ๆ อย่างเช่น NOUN ART และ VERB ในตัวอย่างที่แล้วเรียกว่าเทอร์มินัลซิมโบล (Terminal Symbol) และสัญลักษณ์ที่เหลือคือ NP,VP และ S เรียกว่านอนเทอร์มินัลซิมโบล (Nonterminal Symbol)

เทอร์มินัลซิมโบลจะเป็นกลุ่มของคำและโครงสร้างของคำเรียกว่าเล็กซิคัล (Lexical) ซึ่งจะเก็บรายชื่อของคำทั้งหมดซึ่งมีหลายคำที่สามารถอยู่ได้หลายกลุ่ม เช่น *can* สามารถเป็นได้ทั้ง VERB และ NOUN

เทคนิคการทำพาร์ซิงอย่างง่าย ๆ 2 วิธีคือ การวิเคราะห์แยกคำจากบนลงล่าง (Top-Down Parsing) และการวิเคราะห์แยกคำจากล่างขึ้นบน (Bottom-up Parsing) โดยการวิเคราะห์แยกคำจากบนลงล่างจะเริ่มต้นที่ S และกระจายเป็น VP และ NP และจะกระจายเรื่อย ๆ ตามรีไรท์รูลแต่ละข้อจนถึงเทอร์มินัลซิมโบล (เช่น NOUN) แต่ละตัว จะถูกเขียนเป็นคำ (เช่น cat) ซึ่งจะเป็นคำนามในเล็กซิคัล จะเห็นว่าการวิเคราะห์แยกคำจากบนลงล่างจะใช้กฎจากทางด้านขวาและเขียนใหม่ให้เป็นกฎทางด้านซ้าย ส่วนการวิเคราะห์แยกคำจากล่างขึ้นบนจะเริ่มต้นจากคำแต่ละคำและแทนที่ด้วยกลุ่มไวยากรณ์ของมัน โดยการใช้รีไรท์รูลเราจะสามารถแทนที่รูปแบบที่กำลังวิเคราะห์อยู่ ไปเป็นแบบที่มีขนาดเท่ากันหรือเล็กกว่า ในกรณีนี้ทางด้านซ้ายของกฎจะถูกนำมาใช้เพื่อเขียนลำดับของสัญลักษณ์ทางด้านขวา

ตัวอย่างการวิเคราะห์คำแบบบนลงล่างของประโยค *Thep ate the cat*

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

S → NP VP

→ NAME VP (rewrite NP)

→ Thep VP (rewrite NAME)

→ Thep VERB NP (rewrite VP)

→ Thep ate NP (rewrite VERB)

→ Thep ate ART NOUN (rewrite NP)

→ Thep ate the NOUN (rewrite ART)

→ Thep ate the cat (rewrite NOUN)

ตัวอย่างการวิเคราะห์คำแบบล่างขึ้นบนของประโยค *Thep ate the cat*

→ NAME ate the cat (rewrite Thep)

→ NAME VERB the cat (rewrite ate)

→ NAME VERB ART cat (rewrite the)

→ NAME VERB ART NOUN (rewrite cat)

→ NP VERB ART NOUN (rewrite NAME)

→ NP VERB NP (rewrite ART NOUN)

→ NP VP (rewrite VERB NP)

→ S (rewrite NP VP)

เราจะมาพิจารณาไวยากรณ์ที่ใหญ่กว่านี้อีกเล็กน้อย โดยจะแสดงไวยากรณ์ได้ดังต่อไปนี้

1. S ← NP+VP
2. NP ← ART+NOUN
3. NP ← NAME
4. PP ← PREP+NP
5. VP ← VERB
6. VP ← VERB+NP
7. VP ← VERB+NP+PP
8. VP ← VERB+PP

ตัวอย่างของประโยคที่สามารถใช้กับไวยากรณ์นี้อย่างเช่น

Thep saw the cat by the pond.

The dog barked in the house.

แต่ว่าไวยากรณ์นี้ก็ไม่ได้บอกว่ากริยาตัวใดที่จะมีกรรมตามได้, พรีโพซิชัน (Preposition) จะเหมาะสมกับ กริยา ตัวนั้น , ดังนั้นไวยากรณ์นี้จะยอมรับประโยคต่อไปนี้ด้วย

The dog allow the house.

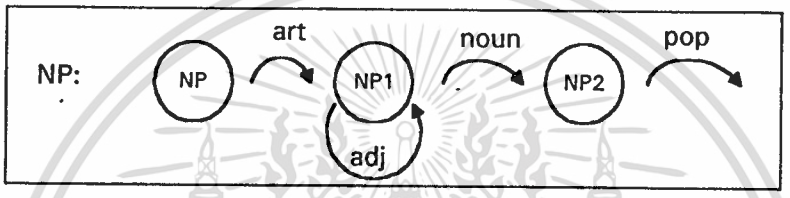
Thep barked the cat by the pond.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โครงสร้างต้นไม้ที่แสดง ดังรูปที่ 2.7 สามารถที่จะมองเป็นเรคอร์ด (record) ของกฎ CFG ที่จะสร้างขึ้นมาเป็นประโยค หรือพูดอีกนัยหนึ่งก็คือ ถ้าเราเก็บเรคอร์ดของวิธีการทำพาร์ซิงเราก็ได้สิ่งที่คล้ายโครงสร้างต้นไม้ซึ่งจะกล่าวถึงในโอกาสต่อไป

ทรานซิชันเน็ตเวิร์กอย่างง่าย (Simple Transition Network)

ทรานซิชันเน็ตเวิร์กเป็นการแทนรูปแบบไวยากรณ์อีกแบบหนึ่งซึ่งจะสะดวกในการมองเห็นรูปแบบของ ไวยากรณ์ซึ่งประกอบไปด้วยโหนด (node) และลาเบลอาร์ค (labeled arc) (คือลูกศร ที่เชื่อมโหนดแต่ละโหนด) พิจารณารูปที่ 2.8



รูปที่ 2.8 NP อย่างง่าย

แต่ละอาร์คจะถูกแสดงเป็นชนิดของคำ โดยเริ่มต้นที่โหนดที่กำหนด เราสามารถที่จะไปยังอาร์คนั้นได้ ถ้าคำที่อยู่ในประโยคคำนั้นเป็นคำชนิดเดียวกับที่อยู่บนอาร์ค วลีที่ถูกต้องของ NP จะต้องมีทางจากโหนด NP ไปยังป๊อปอาร์ค (pop arc) (คืออาร์คที่มีชื่อว่า pop) ซึ่งตามรูปเราสามารถที่จะเขียน CFG ได้ดังนี้

NP ← ART NP1

NP1 ← ADJ NP1

NP1 ← NOUN

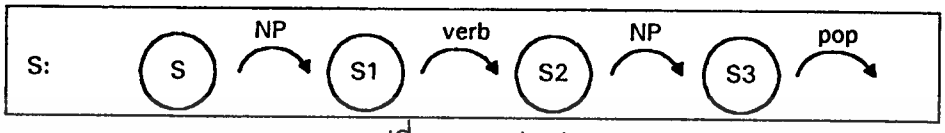
พิจารณาการวิเคราะห์ Noun Phrase : *a purple cow* เริ่มต้นที่ node NP เราสามารถที่จะไปตามทางของ arc ที่มีชื่อว่า art เนื่องจากคำ ๆ นี้คือคำอาร์ติเคิล (article) ที่ชื่อว่า *a* จากโหนด NP1 จะมีอาร์คที่ชื่อว่า adj ซึ่งเราสามารถที่จะใช้คำคุณศัพท์ : *purple* และท้ายสุด จากโหนด NP1 เราสามารถที่จะไปตามอาร์คที่มีชื่อว่า NOUN โดยใช้คำนาม *cow* ดังนั้นเราจะถึงป๊อปอาร์คซึ่งจะเห็นว่า *a purple cow* เป็น NP ที่ถูกต้อง

รีเคอร์ซีฟทรานซิชันเน็ตเวิร์ค Recursive Transition Network

จากหัวข้อที่แล้วเราได้รู้จักกับทรานซิชันเน็ตเวิร์กอย่างง่ายซึ่งยังไม่เพียงพอในการอธิบาย CFG ดังนั้น การที่จะอธิบาย CFG ได้ เราจะต้องมีส่วนที่แสดงถึงการรีเคอร์ชัน (recursion) ในเน็ต - เวิร์คนั้น รีเคอร์ซีฟทรานซิชันเน็ตเวิร์ก เรียกย่อ ๆ ว่า RTN มีลักษณะเหมือนกับทรานซิชัน-

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เน็ตเวิร์กอย่างง่าย ยกเว้นว่ามันสามารถที่จะอ้างถึงเน็ตเวิร์ก ในอาร์คได้ด้วย (ในทรานซิชัน-เน็ตเวิร์กอย่างง่าย ในอาร์คจะอ้างได้เฉพาะประเภทของคำเท่านั้น) ดังนั้นจากไวยากรณ์ในรูปที่ 2.8 เราจะสามารถสร้างประโยคภาษาอังกฤษแบบง่าย ๆ ดังแสดงในรูปที่ 2.9 โดยที่ชื่อที่เป็นตัวอักษรตัวพิมพ์ใหญ่จะแสดงถึงเน็ตเวิร์กในขณะที่ตัวอักษรตัวพิมพ์เล็กจะแสดงถึงชนิดของคำ



รูปที่ 2.9 S อย่างง่าย

นอกจากนี้ RTN ยังอนุญาติให้เน็ตเวิร์กที่อยู่ในอาร์คนั้นเป็นตัวมันเองได้อีกด้วย ซึ่งจะเห็นได้ว่าเป็นการเรียกตัวเองที่แท้จริง

ต่อไปนี้จะสรุปชนิดของอาร์คทั้งหมดรวมถึงรูปแบบที่จะใช้ต่อไปหลังจากนี้

Arc Type	Example	How Used
CAT	noun	succeeds only if current word is of the named category
WRD	of	succeeds only if current word is identical to the label
PUSH	NP	succeeds only if the named network can traversed be successfully
JUMP	jump	always succeeds
POP	pop	succeeds and signals the successful end of the network

2.2.2 ทำอย่างไรถึงจะได้ไวยากรณ์ที่ดีที่สุด

เมื่อเราสร้างไวยากรณ์ของภาษาใด ๆ ขึ้นมา เราก็จะสนใจว่าประโยคที่ไวยากรณ์ของเราจะต้องวิเคราะห์ให้ถูกต้องนั้นจะต้องครอบคลุมถึงแค่ไหน และประโยคที่ไม่ถูกต้องในไวยากรณ์จะทำให้เกิดปัญหาหรือไม่ และจะต้องทำให้ไวยากรณ์เรียบง่ายและเข้าใจได้ง่าย

ในไวยากรณ์ขนาดเล็ก การวิเคราะห์ประโยค ๆ หนึ่ง จะเห็นว่าเข้าใจได้ง่ายกว่าไวยากรณ์ที่ยุ่งยากกว่า ถ้าเราพยายามจะเพิ่มความสามารถของไวยากรณ์ให้วิเคราะห์ให้มากขึ้น เรามักจะพบว่ามักจะเพิ่มได้ง่ายแต่จะแก้ไขปรับปรุงยาก ดังนั้นเราควรจะวิเคราะห์ได้ว่า เราจะต้องคงความเรียบง่ายเอาไว้ ถ้าเราต้องการความสามารถที่เพิ่มขึ้นมาจริง ๆ

ถ้าเราต้องการที่จะแบ่งกลุ่มของคำเพื่อที่จะสร้างส่วนประกอบหลัก ๆ ขึ้นมา และพยายามที่จะสร้างประโยคใหม่ ซึ่งใช้กลุ่มของคำพวกนี้ รวมกับกลุ่มของคำที่มีอยู่เดิม เราจะต้องทำการทดสอบอย่างดีว่าเป็น กลุ่มคำชนิดเดียวกันหรือไม่ที่นำมารวมกัน อย่างเช่น ประโยคต่อไปนี้สามารถที่จะยอมรับกลุ่มคำที่นำมารวมกันได้

NP-NP: I ate a hamburger and a hotdog.

VP-VP: I will eat the hamburger and throw away the hotdog.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

S-S: I ate a hamburger and John ate a hotdog.

PP-PP: I saw the hotdog in the bag and on the stove.

ADJP-ADJP: I ate a cold and well burned hotdog.

ADVP-ADVP: I ate the hotdog slowly and carefully.

NOUN-NOUN: I ate a hamburger and hotdog.

VERB-VERB: I will cook and burn a hamburger.

AUX-AUX: I can and will eat the hotdog.

ADJ-ADJ: I ate the very cold and burned hotdog (that is, very cold and very burned).

ส่วนประโยคต่อไปนี้เป็นการเชื่อมต่อกันที่ยอมรับไม่ได้

*I ate a hamburger and on the stove.

*I ate a cold hotdog and well burned.

*I ate the hotdog slowly and a hamburger.

การทดสอบอีกอย่างหนึ่งก็คือ การแทรกส่วนประกอบที่เหมือนกันเข้าไปในประโยคอื่นที่มีส่วนประกอบเดียวกันนี้ เช่น วลี *Thep's hitting of Pui* เป็น NP ของประโยค *Thep's hitting of Pui alarmed Ben.* ดังนั้น ก็ควรจะใช้ NP นี้ในประโยคอื่นได้เช่นกัน จริง ๆ แล้วมันก็เป็นความจริง เพราะว่า NP นี้สามารถที่จะเป็นกรรมของกริยาได้เช่นในประโยค *Nut can't explain Thep's hitting of Pui.* หรือสามารถที่จะทำรูปพาสซีฟดังเช่นประโยค *Ben was alarmed by Thep's hitting of Pui.* ต่อไปเราจะพิจารณา 2 ประโยคที่มาเชื่อมกันว่ามีโครงสร้างเหมือนกันหรือไม่ อย่างเช่น ประโยค *Nut looked up Thep's phone number and Nut looked up Thep's chimney.* แต่ถ้าเราพยายามที่จะเชื่อมวลี *up Thep's phonenumber* เข้ากับวลีของบุพบทอื่น อย่างเช่น ในประโยค *Nut looked up Thep's phone number and in her cupboards.* ซึ่งดูน่าแปลกประหลาด แต่ถ้าประโยค *Nut looked up Thep's chimney and in her cupboards.* จะเห็นได้ว่า มันยอมรับได้อย่างมาก ดังนั้น การที่คิดว่า *up Thep's phone number* เป็นวลีของบุพบทนั้นไม่ถูกต้อง

2.2.3 วิธีการวิเคราะห์แบบบนลงล่าง (Top-down Parsing method)

วิธีการวิเคราะห์แบบบนลงล่างจะเริ่มต้นจากประโยคและแบ่งแยกออกเป็นส่วนย่อย ๆ จนกระทั่งเราสามารถที่จะแปลงเป็นกลุ่มของคำเฉพาะที่สามารถจะตรวจกับประโยคที่เป็นอินพุทต์จริง ๆ ในหัวข้อนี้เราจะพูดถึงการวิเคราะห์แบบบนลงล่าง 2 วิธีคือ RTN parser และ CFG parser ซึ่งทั้งหมดจะใช้อัลกอริทึมเดียวกันแต่ว่ารายละเอียดส่วนย่อยจะต่างกันเพื่อที่จะใช้การแทนไวยากรณ์ที่แตกต่างกันได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำการวิเคราะห์แบบบนลงล่างโดยใช้ RTN

พิจารณาการพาร์ซิงด้วย RTN เราจะต้องรู้ข้อมูลในการทำดังต่อไปนี้

current position หมายถึงเรคอร์ดที่บอกว่าส่วนไหนของประโยคที่ยังไม่ถูกพาร์ส (parse)

current node หมายถึงโหนดที่อยู่ในขณะนี้ในเน็ตเวิร์ก

returnpoint หมายถึง ถ้าเรายังอยู่ในเน็ตเวิร์กเพราะว่าถูกเรียกจากเน็ตเวิร์ก

ก่อนอื่นเราจะต้องรู้ว่าเราจะต้องกลับไปยังโหนดใดในเน็ตเวิร์กที่เรียกเรามาถ้าเราป๊อปออกจากเน็ตเวิร์กที่เราอยู่

อัลกอริทึมที่จะทำ Top-down Parsing มีดังต่อไปนี้

Case 1: If arc names a word category and next word in sentence is in that category,

Then (1) Update *current position* to start at the next word;

(2) Update *current node* to the destination of the arc.

Case 2: If arc is a push arc to a network N,

Then (1) add the destination of the arc onto *return points;*

(2)update *current node* to the starting node in network N.

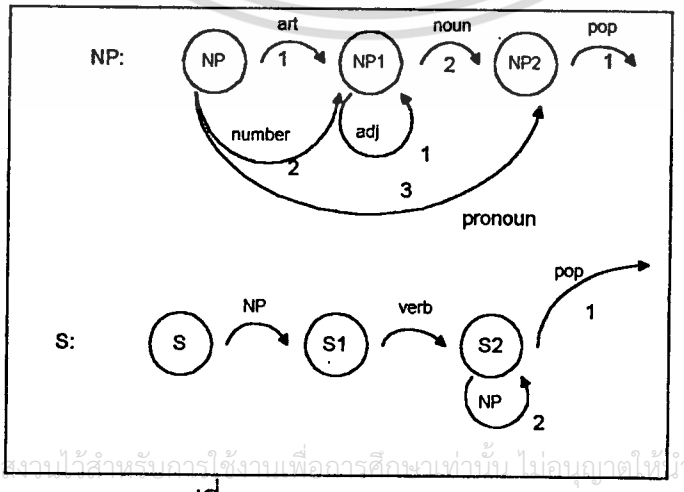
Case 3: If arc is a pop arc and *return points* list is not empty,

Then (1) remove the first return point and make it *current node.*

Case 4: If arc is a pop arc and the *return points list is empty and there are no words left,*

Then (1) parse completes successfully.

เราจะใช้ไวยากรณ์ตามรูปที่ 2.10 และเล็กซิคอนต่อไปนี้



รูปที่ 2.10 ATN ของ S และ NP

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

art: the,a

number: one

pronoun: one

adj: wild,green

noun: dogs,man,saw,green

verb: cried,saw,broke,faded,man

ตัวเลขที่อยู่บนของอาร์คจะบอกถึงลำดับที่จะใช้อาร์คนั้น เมื่อมีอาร์คมากกว่า 1 อาร์คออกมาจากโหนดนั้น ประโยค $_1$*The* $_2$ *wild* $_3$ *dogs* $_4$ *cried.*$_5$ สามารถที่จะพาร์สได้โดยใช้อัลกอริทึมข้างต้น โดยที่ตัวเลขระหว่างคำแต่ละคำใช้เพื่อบอกว่าตอนนี้ตำแหน่งปัจจุบันอยู่ที่ไหน โดยจะพาร์สได้ดังต่อไปนี้

Step	Current Node	Current Position	Return Points	Arc Followed	Comments
1.	(S,	1,	Nil)	S/1	followed push arc to NP network, to return to S1 later
2.	(NP,	1,	(S1))	NP/1	followed art arc to NP1
3.	(NP1,	2,	(S1))	NP1/1	followed adj arc to NP1 again
4.	(NP1,	3,	(S1))	NP1/2	followed noun arc to NP2
5.	(NP2,	4,	(S1))	NP2/1	the pop arc gets us back to S1
6.	(S1,	4,	NIL)	S1/1	followed verb arc to S2
7.	(S2,	5,	NIL)	S2/1	parse succeeds on pop arc from S2

ในประโยคนี้ การแยกคำสำเร็จ แต่อย่างไรก็ตามถ้าเรามีประโยค *The green faded* เข้ากับอัลกอริทึมนี้ มันจะผิดเนื่องจาก มันจะเห็นว่า green เป็นคำคุณศัพท์และไม่สามารถที่จะหาคำนามตามหลังมันได้ดังนั้นการที่เราจะแก้ไขการผิดพลาดเช่นนี้เราจะต้องจำตำแหน่งในเน็ตเวิร์ก ที่สร้างขึ้นมาจากการตัดสินใจและอาจจะผิดได้ เนื่องจากเราสามารถที่จะรู้ข้อมูลทั้ง 3 อย่างคือตำแหน่งปัจจุบัน (current position), โหนดปัจจุบัน (current node) และจุดรีเทิร์น (return point) ดังนั้นจึงเป็นการง่ายมากที่เราจะจำตำแหน่งเมื่อเราเดินไปตามอาร์คและมีอาร์คอื่นอยู่ในโหนดปัจจุบันที่เรายังไม่ได้ทดสอบเราก็จะเก็บตำแหน่งนั้นไว้ ถ้าการวิเคราะห์ไม่สามารถจะทำต่อไปจนจบได้ เราก็เพียงแค่นำตำแหน่งที่เก็บไว้ขึ้นมา และวิเคราะห์อาร์คอื่น ๆ ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พิจารณาการใช้เทคนิคนี้ได้จากประโยคต่อไปนี้

$_1$*One* $_2$ *saw* $_3$ *the* $_4$ *man.*$_5$

จะได้ดังการทำ parser ต่อไปนี้

Step	Current State	Arc Followed	Backup States
1.	(S,1,NIL)	S/1	NIL
2.	(NP,1,(S1))	NP/1(& NP/3 for backup	NIL
3.	(NP1,2,(S1))	NP1/2	(NP2,2,(S1))
4.	(NP2,3,(S1))	NP2/1	(NP2,2,(S1))
5.	(S1,3,NIL)	no arc can be followed	(NP2,2,(S1))
6.	(NP2,2,(S1))	NP2/1	NIL
7.	(S1,2,NIL)	S1/1	NIL
8.	(S2,3,NIL)	S2/2	NIL
9.	(NP,3,(S2))	NP/1	NIL
10.	(NP1,4,(S2))	NP1/2	NIL
11.	(NP2,5,(S2))	NP2/1	NIL
12.	(S2,5,NIL)	S2/1	NIL
13.	the parse succeeds		

การกระทำจะเหมือนกับตัวอย่างที่แล้ว ยกเว้นอยู่ 2 ที่คือ ที่สเต็ปที่ 2 มีอาร์คที่ออกจากโหนด NP 2 อาร์ค ซึ่งยอมรับคำว่า *one* โดยอาร์ค NP/1 แบ่งให้ *one* เป็นตัวเลขและสร้างสเตต (state) ต่อไปและอาร์ค NP/3 ให้ *one* เป็นสรรพนาม (pronoun) และสร้างสเตตสำรองซึ่งสเตตสำรองนี้จะได้ใช้ในสเต็ปที่ 6 เมื่อมันพบว่าไม่มีอาร์คใดออกจากโหนด S1 ที่สามารถจะรับคำ *the*

การทำการวิเคราะห์แบบบนลงล่างโดยใช้ CFG

เราสามารถที่จะสร้างการวิเคราะห์แบบบนลงล่างจาก CFG โดยใช้เทคนิคที่ได้อธิบายไปแล้วในหัวข้อที่แล้ว แต่มีข้อแตกต่างกันในเรื่องการอธิบายสถานะของการพาร์ส CFG ซึ่งมีองค์ประกอบอยู่ 2 ตัว คือตำแหน่งปัจจุบัน และสตริง (string) ของสัญลักษณ์ทางไวยากรณ์ ซึ่งดัดแปลงมาจากรีไรท์รูลของไวยากรณ์ซึ่งแต่ละอันจะพิจารณาจากสัญลักษณ์ทางด้านซ้ายสุด ถ้ามันมีชื่อตรงกับประเภทของคำต่อไป ก็จะเอาสัญลักษณ์นั้นออกมาและแก้ไข ตำแหน่งปัจจุบัน เราจะใช้กฎ 8 กฎที่กล่าวมาข้างต้น และประโยค $_1$*The* $_2$ *dog* $_3$ *cried.*$_4$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เราจะทำพาร์ซิงได้ดังต่อไปนี้

Step	Current State	Backup States	Position	Comment
1.	(S)		1	initial position
2.	(NP VP)		1	rewriting S by rule 1
3.	(ART NOUN VP)		1	rewriting NP by rules 2 & 3
		(NAME VP)	1	
4.	(NOUN VP)		2	matching ART with *the*
		(NAME VP)	1	
5.	(VP)		3	matching NOUN with *dogs*
6.	(VERB)		3	rewriting VP by rules 5-8
		(VERB NP)	3	
		(VERB NP PP)	3	
		(VERB PP)	3	
		(NAME VP)	1	
7.				the parse succeeds as VERB is matched to *cried* ,leaving an empty grammatical symbol list with an empty sentence

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 ส่วนประกอบและส่วนขยายไวยากรณ์ (Features and Augmented Grammars)

กระบวนการต่างๆที่ได้กล่าวในบทที่แล้ว เราจำกัดทางเดินของการวิเคราะห์ประโยค กล่าวคือ สิ่งที่เราสามารถทำได้ ก็เพียงแต่ตัดสินได้ว่า ประโยคผ่านกฎ (rules) ที่กำหนดหรือไม่ สามารถใช้กับไวยากรณ์ที่กำหนดได้ (reject) ซึ่งส่วนที่ยังขาดคือ ความชัดเจนในการจำแนกประธาน-กิริยา ,กริยา ตัวใดต้องมีกรรม ตัวใดไม่ต้องมีกรรมมารับ และนอกจากนั้นยังมีปัญหาที่เกี่ยวข้องกับข้อเท็จจริงอื่น ๆ อีก ซึ่งในหัวข้อนี้ เราจะเพิ่มส่วนขยาย (augmented) ซึ่งจะจัดการกับส่วนดังกล่าวนี้

2.3.1 อ๊อกเม็นท์ทรานซิชันเน็ตเวิร์ก (Augmented Transition Network)

ในหัวข้อนี้เราจะกล่าวถึงกระบวนการสำหรับการรวบรวมรายละเอียดและส่วนประกอบที่ใช้ทดสอบในขณะที่ทำพาร์ซิงโดยการใช้ RTN framework ซึ่งเรียกว่าอ๊อกเม็นท์ทรานซิชันเน็ตเวิร์ก (เรียกย่อ ๆ ว่า ATN)

<u>การเก็บโครงสร้างประโยคในขณะทำการพาร์ซิง (Recording the Sentence Structure While Parsing)</u>

สมมติว่าเราต้องการที่เก็บโครงสร้างของประโยคที่ถูกต้องที่จะใช้ในการวิเคราะห์ โครงสร้างที่เรารู้จักกันแล้วก็คือซินแทคติคพาร์สทรี (ซินแทคติค parse tree) ซึ่งอยู่ในรูปของ CFG วิธีที่ดีและมีประสิทธิอีกวิธีหนึ่งก็คือ การใช้สลอต-ฟิลเลอร์ (slot-filler) ซึ่งมีฟังก์ชั่นของกฎในประโยคเพิ่มขึ้นมากล่าวคือ เราจะแทนส่วน NP ในประโยคด้วยซินแทคติค ซับเจ็ค (ซินแทคติค ประธาน) (ต่อไปจะเรียกย่อ ๆ ว่า SUBJ) และแทนส่วน VP ด้วยซินแทคติคซับเจ็คของคำกิริยา (OBJ) ภายใน NP เราอาจจะกำหนดส่วนประกอบเพิ่มเติมเช่น ดีเทอร์ไมเนอร์ (determiner) (คือ a, an, the), คำคุณศัพท์, head noun และอื่น ๆ ดังนั้นประโยค

Jack found a dime.

สามารถแสดงโครงสร้างได้ดังนี้

```
(S SUBJ (NP NAME Jack)
   MAIN-V found
   TENSE PAST
   OBJ (NP DET a
           HEAD dime))
```

ตัวอย่างโครงสร้างข้างบนนี้ สร้างมาจากพาร์เซอร์ของ RTN โดยอนุญาติให้แต่ละเน็ตเวิร์กมีเซ็ทของรีจิสเตอร์ในแต่ละครั้งที่มีการใส่เน็ตเวิร์กใหม่จะการสร้างเซ็ทของรีจิสเตอร์ซึ่งจะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ว่างอยู่เมื่อเน็ตเวิร์กถูกป๊อป รีจิสเตอร์ก็จะหายไป สำหรับรีจิสเตอร์เราสามารถที่จะให้ค่ามันได้ และค่านี้เราสามารถที่จะดึงออกจากรีจิสเตอร์ได้ด้วย ในกรณีนี้รีจิสเตอร์จะมีชื่อของสลอตที่ใช้ในแต่ละโครงสร้างของชินแทคติคดังนั้น เน็ตเวิร์กของ NP จะมีชื่อของ รีจิสเตอร์ ต่อไปนี้คือ DET,ADJS, HEAD และ NUM รีจิสเตอร์ จะถูกเซ็ทโดยแอคชั่น *แอคชั่น)* ซึ่งมันสามารถกำหนดลงบน อาร์คเมื่อเรา follow ไปตาม อาร์ค, แอคชั่นต่าง ๆ ที่เกี่ยวข้องกับมันจะถูกเอ็กซีคิวต์, แอคชั่นที่ธรรมดาที่สุดจะเกี่ยวกับการเซ็ทรีจิสเตอร์ให้เป็นค่าที่ถูกต้อง ส่วนแอคชั่นอื่น ๆ จะกล่าวถึงในภายหลัง เมื่อเราไล่ไปตามป๊อปอาร์ค รีจิสเตอร์ทั้งหมดในเน็ตเวิร์กปัจจุบันจะถูกรวบรวมโดยอัตโนมัติ เป็นรูปแบบโครงสร้างที่ประกอบด้วยชื่อเน็ตเวิร์ก ตามด้วยลิสต์ (list) ของรีจิสเตอร์และค่าของมัน RTN กับรีจิสเตอร์, ส่วนเทสต์ (tests) และแอคชั่นบนรีจิสเตอร์เหล่านี้ก็คือ ATN นั่นเอง

ปัญหาก็คือแอคชั่นทำอย่างไรที่จะได้ค่าซึ่งจะนำมาเซ็ทในรีจิสเตอร์ มีกลไกซึ่งจะทำให้มันง่ายขึ้นโดยเมื่อเรา ไล่ไปตามแคทอาร์ค (cat arc) ซึ่งเป็นคำกริยา คำซึ่งเป็นอินพุทต์จะถูกใส่เข้ามาในตัวแปรพิเศษที่มีชื่อว่า " * " ดังนั้นเมื่อแอคชั่นซึ่งอยู่บนอาร์คจาก S1 ไป S2 ในรูปที่ 2.11 จะถูกเซ็ทคำกิริยา รีจิสเตอร์ให้เป็นคำปัจจุบันหรือเขียนได้ดังนี้

verb ← *

พุชอาร์ค (push arc) อย่างเช่น NP จะถูกทำต่างกัน ซึ่งคำหลายคำที่ใช้ในเน็ตเวิร์กจะถูกเรียกโดยพุชอาร์ค รวมทั้งเน็ตเวิร์กซึ่งใช้ในพุชจะต้องเซ็ทรีจิสเตอร์ซึ่งจะดักโครงสร้างของส่วนประกอบ ที่ถูกพาร์สเข้ามา การที่จะทำให้มันสามารถเข้าถึงเน็ตเวิร์กได้ โครงสร้างซึ่งถูกสร้างโดยพุชเน็ตเวิร์กจะส่งค่ากลับมาสู่ใน " * " ดังนั้นแอคชั่นบนอาร์ค จาก S ไป S1 ในรูปที่ 2.11 จะเป็น

SUBJ ← *

ถ้าประโยคเริ่มต้นด้วยวลี *The purple cow ...* แอคชั่นนี้จะให้ผลลัพธ์ในรีจิสเตอร์ SUBJ ซึ่งจะถูกเซ็ทในโครงสร้างที่ได้รับค่ากลับมาในเน็ตเวิร์กของ NP ดังนี้

(NP ART the

ADJS (purple)

HEAD cow)

<u>การตรวจสอบองค์ประกอบโดยการทดสอบ (Checking Features with Tests)</u>

นอกจากจะใช้รีจิสเตอร์รวบรวมโครงสร้าง รีจิสเตอร์ยังสามารถที่จะใช้ในการทำให้ไวยากรณ์มีทางเลือกที่ดีกว่าโดยไม่ทำให้เน็ตเวิร์กยุ่งยากสับสนโดยไม่จำเป็น ในภาษาอังกฤษ จำนวนของ NP ชุดแรกจะขึ้นอยู่กับจำนวนของ คำกิริยา คือ เราไม่สามารถที่จะใช้ NP ชุดแรกเป็นเอกพจน์กับกริยาพหูพจน์ได้ เช่น ประโยค *The dog are sick.* ที่สำคัญไปกว่านั้น มีคำจำนวนมากที่ทำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

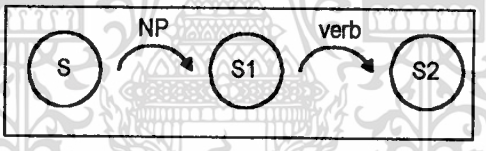
ให้ประโยคเกิดความกำกวม พิจารณาประโยค 2 ประโยคต่อไปนี้

Flying planes is dangerous. หมายความว่า การทำกิจกรรมในการขับเครื่องบินมีอันตราย

Flying planes are dangerous. หมายความว่า เครื่องบินหลาย ๆ ลำซึ่งบินอยู่มีอันตราย

วิธีแก้ปัญหาที่ดีกว่าการใช้ RTN (ซึ่ง RTN จะทำให้ต้องเพิ่มไวยากรณ์ ขึ้นมาอีก 2 เท่า) ก็คือ การกำหนดให้คำและโครงสร้างทางไวยากรณ์มีองค์ประกอบ (feature) ซึ่งการทำเช่นนี้สามารถจะทำได้โดยสร้างที่เก็บข้อมูลที่มีประโยชน์เกี่ยวกับคำนั้น ๆ ในโครงสร้างข้อมูลที่เรียกว่าเล็กซิคอน รูปที่ 2.12 จะแสดงเล็กซิคอนซึ่ง 3s หมายถึง เอกพจน์ และ 3p หมายถึง พหูพจน์

เราจะทำการขยายความสามารถของ RTN โดยการเพิ่มการทดสอบ (ส่วนเทสต์) ไปในแต่ละอาร์คของ เน็ตเวิร์ก, เทสต์เป็นฟังก์ชั่นซึ่งจะสำเร็จถ้ามันรีเทิร์นค่ากลับมา และจะล้มเหลวเมื่อมันรีเทิร์นเซ็ตว่างหรือค่านิล (nil) ถ้าส่วนเทสต์ล้มเหลว ก็จะไม่ไปทางอาร์คนั้น หรือจะพูดอีกนัยหนึ่งก็คือส่วนเทสต์จะถูกประยุกต์ใช้หลังจากที่ค่าที่รับเข้ามาถูกต้อง ตามแคทอาร์ค หรือหลังจากที่มีการเรียกตัวเอง (recursive call) เกิดขึ้นและสำเร็จบนพุชอาร์ค ในระบบจริง ๆ มักจะพบว่า จะใช้ส่วนเทสต์ก่อนที่จะพุช



รูปที่ 2.11 ทรานซิชันเน็ตเวิร์กอย่างง่าย

Word	Representation
dogs	(NOUN ROOT DOG NUM {3p})
dog	(NOUN ROOT DOG NUM {3s 3p})
the	(ART ROOT THE NUM {3s 3p})
a	(ART ROOT A NUM {3s})
cried	(VERB ROOT CRY NUM {3s 3p})

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

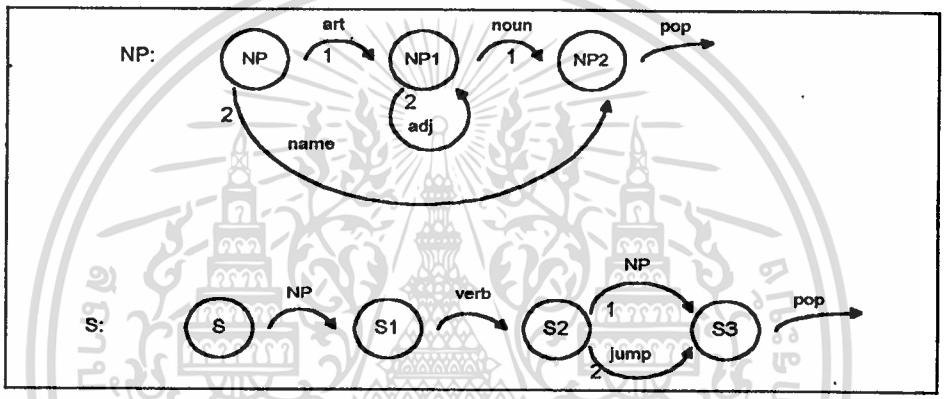
loves (VERB ROOT LOVE
NUM {3s})

love (VERB ROOT LOVE
NUM {3p})

wild (ADJ ROOT WILD)

John (NAME ROOT JOHN)

รูปที่ 2..12 เล็กซิคอนอย่างง่าย (assuming third person only)



Arc	Test	Action
NP/1	none	DET ← *
		NUM ← NUM *
NP1/1	NUM ∩ NUM *	HEAD ← *
		NUM ← NUM ∩ NUM *
NP1/2	none	ADJS←Append(ADJS, *)
NP/2	none	NAME ← *
		NUM ← NUM *
S/1	none	SUBJ ← *
S1/1	NUM_{SUBJ} ∩ NUM *	MAIN-V ← *
		NUM←NUM_{SUBJ} ∩ NUM *
S2/1		OBJ ← *

รูปที่ 2.13 ATN อย่างง่าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำงานของเน็ตเวิร์ก S

Step	Node	Position	Arc	Followed	Register
1.	S	1	S/1	attempted (for recursive call see trace below)	
5.		3	S/1	succeeds	SUBJ ← (NP DET the HEAD dogs NUM {3p})
	S1	3	S1/1	(tests whether {3p} ∩ {3p}	MAIN-V ← love NUM ← {3p}
6.	S2	4	S2/1	(for recursive call trace, see below)	OBJ ← (NP NAME John NUM {3p})
9.	S3	5	S3/1	succeeds since no words left	returns (S SUBJ (NP DET the HEAD dogs NUM {3p}) VERB love NUM {3p} OBJ (NP NAME John NUM {3s})

การทำงานของการเรียก NP ชุดแรก : Arc 6

Step	Node	Position	Arc	Followed	Register
2.	NP	1	NP/1		DET ← the NUM ← {3s 3p}
3.	NP1	2	NP/1	(check if {3s 3p} ∩ {3p} not empty)	HEAD ← dogs NUM ← {3p}
4.	NP2	3	NP2/1		returns (NP DET the

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

HEAD dogs

NUM {3p})

การทำงานของการเรียก NP ชุดที่สอง : Arc 8

Step	Node	Position	Arc	Followed	Register
7.	NP	4	NP/2		NAME ← John
8.	NP2	5	NP2/1		returns (NP NAME John)
					NUM {3s})

รูปที่ 2.14 Trace tests and actions used with *The dogs love John.*

ตัวอย่าง

ขณะนี้เรามีกลไกที่จะอธิบาย ATN parser สำหรับประโยคโครงสร้างภาษาอังกฤษบางส่วน รูปที่ 2.13 จะยอมรับประโยคซึ่งเป็นไปได้ทั้งกริยามีกรรมและไม่มีกรรม และ simple nounphase ซึ่งเป็นชื่อเฉพาะหรือเป็นคำที่มี คำคุณศัพท์ มาขยาย ไวยากรณ์ ตามรูปที่ 2.13 ประกอบกับ lexicon ตามรูปที่ 2.12 สามารถที่จะยอมรับประโยคเหล่านี้ได้

The wild dog cried.

The dogs love John.

John loves the wild dogs.

แต่ว่าไม่ยอมรับประโยคเหล่านี้

* A wild dogs cried. (no number agreement in NP)

* John love a dog. (no number agreement in SUBJ/คำกิริยา)

พิจารณาตัวอย่างสักหนึ่งตัวอย่าง เมื่อเราจะวิเคราะห์ประโยค *The dogs love John.* ATN ในรูปที่ 2.13 จะตรวจสอบ agreement ระหว่าง *The* กับ *dogs* เป็นอย่างแรก *The* จะเป็นได้ทั้งเอกพจน์และพหูพจน์ แต่ *dogs* จะเป็น พหูพจน์ ดังนั้น NP ที่ประกอบขึ้นมาจะเป็นพหูพจน์ (ซึ่งเป็นผลลัพธ์จาก NP1/1) NP นี้จะถูกกำหนดให้เป็น SUBJ รีจิสเตอร์ และต่อไปก็จะตรวจสอบ agreement ของ คำกิริยา ซึ่งจะเป็นไปตามรูปที่ 2.14 ดังนั้น S/1 ซึ่งเป็น อาร์ค อันแรกที่ออกจาก S node ประโยคจะแบ่งตำแหน่งของคำได้ดังต่อไปนี้

1 The 2 dogs 3 love 4 John 5.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.2 Useful Feature Systems

ในหัวข้อนี้เราจะทดสอบโครงสร้างของภาษาอังกฤษหลาย ๆ แบบ ดังต่อไปนี้

ความเข้ากันได้ของประธาน-กิริยา

จากหัวข้อที่แล้วเราได้เห็นการวิเคราะห์ของความเข้ากันได้ระหว่างประธานและคำกิริยา และนอกจากที่กล่าวมาแล้ว ประธานและคำกิริยาจะต้องมีความสอดคล้องกันในด้านของบุคคล (person) ในประโยคภาษาอังกฤษบุคคลจะถูกระบุโดยการใช้คำสรรพนาม ซึ่งจะประกอบไปด้วย บุรุษที่ 1 (I, we), บุรุษที่ 2 (you) และ บุรุษที่ 3 (he, she, it, they) และประธานของประโยคที่ไม่ได้เป็นสรรพนามทั้งหมดจะเป็นบุรุษที่ 3 พิจารณาวลีข้างล่างนี้

I love	we love
you love	you love
he/she loves	they love

จะเห็นได้ว่า love มีรูปแบบที่แตกต่างกันถึง 5 รูปแบบ สัญลักษณ์ที่ใช้แสดงบุคคลและจำนวนของ ฟีเจอร์ (feature) ของคำกิริยาจะประกอบไปด้วยลิสต์ซึ่งมีรูปแบบที่เป็นไปได้ทั้งหมดสำหรับเอกพจน์และพหูพจน์ ดังนั้น love จะถูกแบ่งประเภทออกเป็น {1s 2s 1p 2p 3p} ซึ่งบอกว่ามันจะเข้ากันได้กับประธานที่เป็นเอกพจน์บุรุษที่ 1 และ 2 หรือประธานที่เป็นพหูพจน์ทั้งหมด อีกด้านหนึ่ง loves อาจจะเป็น { 3s } ซึ่งจะบอกว่ามันจะเข้ากันได้กับประธานที่เป็นเอกพจน์บุรุษที่ 3

ด้วยระบบฟีเจอร์นี้เราสามารถที่จะกำหนดการทดสอบที่เรียกว่า Agr(feature1, feature2) ซึ่งจะมี 2 ฟีเจอร์ และคำนวณการอินเตอร์เซ็คชั่น (intersection) ของมัน ถ้าการอินเตอร์เซ็คชั่นได้ผลลัพธ์เป็นค่านุล (null) หรือว่างเปล่า การทดสอบจะล้มเหลว นอกจากนั้นการทดสอบจะสำเร็จและจะส่งค่าที่ได้จากการอินเตอร์เซ็คชั่นกลับมา เช่น Agr({3s 3p},{1s 2s 1p 2p 3p}) ประสบผลสำเร็จและรีเทิร์น {3p} Agr({2s 2p},{1s 2s 1p 2p 3p}) สำเร็จและรีเทิร์น {2s 2p} Arg({3s},{2s 1p 2p 3p}) ล้มเหลว

ความเข้ากันได้ของ กิริยาช่วย-คำกิริยา (Auxiliary-verb Agreement)

ต่อไปนี้จะเป็นประโยคซึ่ง มีกลุ่มของคำกริยา ประกอบด้วย คำกิริยา ช่วย ซึ่งมีตั้งแต่ 0 ตัวขึ้นไป ตามด้วย คำกิริยา หลัก ดังต่อไปนี้

I can see the house.

I will have seen the house.

I was watching the movie.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

I should have been watching the movie.

I will be seen at the house.

คำกิริยาช่วยพวกนี้ จะทำให้ คำกิริยา ที่ตามมันมีรูปแบบเฉพาะดังเช่นในรูปที่ 2.15

Form	Feature Name	Examples
infinitive	inf	go, be, say, decide
present	pres	go,goes,am, is, say, says,decide
present participle	ing	going, being, saying, deciding
past participle	en	goen, been, said, decided
past	past	went, was, said, decided

รูปที่ 2.15 รูปแบบของคำกิริยา

พิจารณาว่าคำกิริยาช่วยเหล่านี้มีกฎเกณฑ์บังคับคำกิริยาที่ตามมันเช่น *have* จะต้องตามด้วย past participle และ *be* จะต้องตามด้วย present participle หรืออยู่ในรูปของพาสซีฟ ซึ่งจะเป็นรูปของ past participle , *do* มักจะปรากฏเดี่ยว ๆ แต่มันก็ยังสามารถที่จะยอมรับ infinitive เช่น *I did eat my carrots!*

Auxiliary ที่นอกเหนือจากนี้ (จะเรียกว่าโมดอล (Modal)) มักจะตามด้วยรูปอินฟินิทีฟ (infinitive) ซึ่งกฎเหล่านี้จะนำมารวมกันเป็นโพรซีเดอร์ (procedure) AuxAgree(verb, aux-list) ซึ่งจะนำคำกิริยาตัวต่อไปมาตรวจสอบว่า มันถูกต้องตามกฎหรือไม่ ซึ่งโพรซีเดอร์ นี้จะเป็นดังต่อไปนี้

```
AuxAgree(v, aux-list)
IF aux-list = nil
    Then if FORM_V = pres or past
            then succeed and return T
            else fail;
Let L = LAST(aux-list)
If ROOT_L = BE & FORM_V = ing
    then succeed and return PROGRESSIVE;
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

If $ROOT_L$ = BE & $FORM_V$ = en

then succeed and return PASSIVE;

If $ROOT_L$ = HAVE & $FORM_V$ = en

then succeed and return PERFECT;

If $TYPE_L$ is a modal & $FORM_V$ = inf

then succeed and return MODAL;

Fail;

โพรซีเดอร์นี้สามารถที่จะใช้ตรวจสอบวลี *Can be seen* ได้ดังต่อไปนี้

AuxAgree(*can*,nil) returns T

AuxAgree(*be*,*(can)*) returns MODAL

AuxAgree(*seen*,*(can be)*) returns PASSIVE

ส่วนเติมเต็มของกิริยา (Verb Complement)

โครงสร้าง complement ของ คำกิริยา ประกอบด้วย NP และอนุประโยค ซึ่งตามหลังคำกริยา เราสามารถที่จะแบ่งประเภทของ คำกิริยา โดยพิจารณาจากองค์ประกอบซึ่งตามหลังคำกิริยาได้หลายชนิดดังต่อไปนี้

NONE	no complements allowed(intransitive)
OBJ	single NP allowed to follow verb(transitive)
IOBJ+OBJ	two NPs allowed to follow verb(bitransitive)
ADJ	takes an adjective complement
OBJ+ADJ	takes an object followed by an adjective phrase
OBJ+PP	
FOR-TO-INF	- as in *I prayed for the doctor to come in time.*
OBJ+TO-INF	- as in *I presuaded him to do it.*
TO-INF	- as in *I tried to do it.*
OBJ+INF	- as in *I saw him do it.*
INF	- as in *I helped do it.*
THAT	- as in *I know that Jack left.*

I know Jack left.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

OBJ+THAT - as in *Jack told Mary that he had lost his bicycle.*

WH-COMP - as in *We doube what John said.*

OBJ+WH-COMB - as in *They asked Jack whether it was raining.*

เราไม่สามารถจะใช้ประโยคเหล่านี้ได้

* Jack was sad for me to get my clothes clean.

* I was destined that my canary died.

การที่จะแสดงการทดสอบ ไวยากรณ์ เหล่านี้ให้รัดกุม พิจารณาฟังก์ชั่นต่อไปนี้ ซึ่งเหมาะสมสำหรับลักษณะขั้นต้น

Intrans(f) = f { NONE ADJ FOR-TO-INF to-INF INF THAT WH-COMP }

Trans(f) = f { OBJ IOBJ+OBJ OBJ+To-INF OBJ+INF OBJ+ADJ OBJ+PP
OBJ+THAT OBJ+WH-COMP }

2.3.3 **ส่วนเติมเต็มคำกิริยาและการเซ็ตค่ารีจิสเตอร์** (Verb Complements and Preseting Register)

พิจารณา คำกิริยา อีกประเภทหนึ่ง (อย่างเช่น *want* และ *pray*) ซึ่งจะรับส่วนเติมเต็มโดยอยู่ในรูปของ คำกิริยาอินฟินิทีฟ ซึ่งจะตามหลังคำว่า *to* ซึ่งประกอบไปด้วย หลักเกณฑ์ต่อไปนี้ TO-INF Mary wants to have a party.

OBJ+TO-INF Mary wants John to have a party.

FOR-TO-INF I prayed for John to leave the party.

จะสังเกตเห็นว่า ในรูปแบบนี้จะเหมือนกับโครงสร้างประโยคทั่ว ๆ ไปแต่ไม่มี NP ซึ่งเป็นประธานและไม่มี auxiliary และคำกิริยาจะต้องอยู่ในรูปของอินฟินิทีฟ ดังนั้นจึงเป็นการง่ายที่จะนำ S เน็ตเวิร์ก มาใช้ใหม่แทนที่จะเขียนเน็ตเวิร์กขึ้นมาใหม่

การที่จะทำอย่างนี้ได้เราจำเป็นต้องสร้างจุดเริ่มต้นใน S เน็ตเวิร์ก ในโหนดอื่น ดังนั้นเราจะเห็นว่าโหนดที่ต้องการคือโหนด S2 และยิ่งไปกว่านั้น พิจารณาข้อแตกต่างระหว่างประโยค 2 ประโยคต่อไปนี้ *Mary wants to have a party.* และ *Mary wants John to have a party.* ในประโยคแรก หมายความว่า Mary เป็นผู้ที่จะต้องไปในงาน party แต่ว่าในประโยคหลัง John จะเป็นผู้ไปใน party เราสามารถที่จะนำประโยคที่มีความหมายแตกต่างกันนี้ไปใส่ในส่วน การวิเคราะห์ความหมาย (semantic analysis) ก็ได้ แต่อาจจะทำให้ส่วนการวิเคราะห์ความหมายยุ่งยากเกินไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นมันก็เป็นเหตุผลที่ดีที่เราจะใส่ความเข้าใจในประโยคเหล่านี้ไว้ในประโยคซินแทคติคเลย อย่างเช่นพิจารณาประโยคต่อไปนี้

Mary wants to dress herself in the closet.

Mary wants John to dress himself in the closet.

* Mary wants to dress himself in the closet.

* Mary wants John to dress herself in the closet.

จะเห็นว่า 2 ประโยคสุดท้ายนั้นไม่สามารถจะยอมรับได้เนื่องจาก เพราะว่าเพศของสรรพนามที่แสดงถึงตัวเอง ไม่ตรงกับประธานในส่วนเติมเต็มนั้น การที่จะบังคับให้ไวยากรณ์เข้าใจเพศได้ถูกต้องก็ต้องใส่เพศลงใน SUBJ ซึ่งอยู่ในส่วนเติมเต็มด้วย

ดังนั้นจะเป็นการดีที่จะพรีเซ็ต (preset) รีจิสเตอร์ SUBJ ใน เน็ตเวิร์ก S อันใหม่ของเราดังนั้นพุชอาร์ค จะอยู่ในรูปดังต่อไปนี้

PUSH(N,<register in N> <register in current network>)

ตัวอย่างเช่น

PUSH(S2,SUBJ SUBJ)

ซึ่งหมายความว่าจะพุชลงไปในโหนด S2 ใน เน็ตเวิร์ก S และพรีเซ็ต รีจิสเตอร์ SUBJ ใน เน็ตเวิร์ก S อันใหม่ให้เป็นค่าของ รีจิสเตอร์ SUBJ ใน เน็ตเวิร์ก S อันปัจจุบัน ซึ่งจะสามารถอธิบายประโยค *Mary wants to have a party.* ส่วนในกรณีที่ 2 *Mary wants John to have a party.* จะต้องใช้รูปแบบ

PUSH(S2,SUBJ OBJ)

ปัญหาสุดท้ายก็คือจะทำอย่างไรกับคำ *to,* auxiliary หรือกริยาตัวแรกที่ตามหลัง *to* จะต้องอยู่ในรูปอินฟินิทีฟ ดังนั้นเราจึงพิจารณาให้ *to* เป็น auxiliary modal verb

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4 ไวยากรณ์สำหรับภาษาธรรมชาติ : แฮนดริงมูฟเม็นท์ (Handling Movement)

ในหัวข้อนี้เราจะขยายไวยากรณ์ของภาษาออกไปอีก ซึ่งจะมีโครงสร้างของประโยคพื้นฐานที่แตกต่างจากที่เคยเห็นแล้ว คือในบางกรณีคำหรือวลีอาจจะจัดลำดับใหม่ได้ภายในประโยค บางวลีอาจจะถูกเคลื่อนที่ย้ายออกไปจากจุดที่ควรจะอยู่ในประโยคปกติ นอกจากนั้น ยังรวมถึงการ แสดงตัวอย่างของการเพิ่มและการลบวลีบางอย่างด้วย

พิจารณาตัวอย่างของโครงสร้างของประโยคคำถามแบบ Yes-No ต่อไปนี้

John went to the store.

Did John go to the store?

He will run in the marathon next year.

Will he run in the marathon next year?

Jack is giving Sue a back rub.

Is Jack giving Sue a back rub?

Henry goes to school every day.

Does Henry go to school every day?

จะสังเกตเห็นว่าคำถามแบบ Yes-No จะเหมือนกับโครงสร้างเดิมของมันยกเว้น NP ที่เป็นประธานและ auxiliary ตัวแรกจะสลับที่กัน ถ้าประโยคนั้นดั้งเดิมไม่มี auxiliary ก็จะใช้ auxiliary ประเภท *do* ประกอบกับเท็นส์ที่เหมาะสมการที่จัดลำดับของประธานและ auxiliary ใหม่นี้เรียกว่า subject-aux inversion

โดยปกติแล้วเราสามารถจะคิดได้ว่า การแปลงคำถามแบบ Yes-No จากประโยคดั้งเดิมของมันนั้น จะทำโดยการเคลื่อนย้ายในลักษณะที่ได้กล่าวไปแล้วในตัวอย่างข้างต้นซึ่งการเคลื่อนย้ายแบบมีข้อจำกัด (bounded movement) แต่มีการเคลื่อนย้ายอีกประเภทหนึ่งซึ่งจะสลับตำแหน่งภายในองค์ประกอบต่าง ๆ ซึ่งเรียกว่า การเคลื่อนที่แบบไม่มีข้อจำกัด (unbounded movement) ซึ่งจะพบในประโยคคำถามที่ขึ้นต้นด้วย Wh (Wh-Questions) ในกรณีของการเคลื่อนย้ายแบบไม่มีข้อจำกัด ส่วนประกอบของประโยคจะเคลื่อนที่ไกลจากจุดเริ่มต้นของมันมาก

พิจารณาตัวอย่างประโยคซึ่งเกี่ยวพันกับประโยคดั้งเดิมนี้

The fat man will angrily put the book in the corner.

ถ้าเราสนใจว่าใครเป็นคนทำเราอาจสามารถจะถามใหม่ได้ดังประโยคต่อไปนี้

Which man will angrily put the book in the corner?

Who will angrily put the book in the corner?

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หรืออีกนัยหนึ่งถ้าเกิดเราสนใจว่ามันถูกทำอย่างไรเราก็จะถามได้ดังประโยคข้างล่างต่อไปนี้

How will the fat man put the book in the corner?

In what way will the fat man put the book in the corner?

หรืออาจจะถามแบบอื่น ๆ ได้ดังต่อไปนี้

What will the fat man angrily put in the corner?

Where will the fat man angrily put the book?

In what corner will the fat man angrily put the book?

What will the fat man angrily put the book in?

ประโยคคำถามแต่ละประโยคมีรูปแบบจากประโยคดั้งเดิมเหมือนประโยคเดียวกัน ยกเว้นส่วนที่จะนำมาถาม ซึ่งจะถูกย้ายขึ้นมาและแทนที่ด้วย Wh-phrase ที่ต้นประโยคนั้น

จากตัวอย่าง Wh-Questions การเคลื่อนที่แบบหลัก ๆ เป็นแบบอย่างที่แน่นอนตามหลักไวยากรณ์ ซึ่งจะรับรูปแบบของประโยคได้มากกว่า CFG และ RTN

2.4.1 Local Movement *

ในส่วนนี้เราจะสนใจการย้ายอยู่ 2 กรณี คือประโยคคำถามแบบ Yes-No และประโยคพาสซีฟและจะแสดงให้เห็นว่า ATN ที่เรามีอยู่จะสามารถนำมาใช้วิเคราะห์ประโยคเหล่านี้ได้อย่างไร

ประโยคคำถามแบบ Yes-No (Yes-No Questions)

จุดประสงค์ของเราก็คือขยายไวยากรณ์ของ ATN ที่ตอนนี้มันสามารถจะรับประโยคแบบปกติให้สามารถรับ ประโยคคำถามแบบ Yes-No ได้โดยเราจะพยายามใช้ไวยากรณ์ให้มากที่สุดเท่าที่จะทำได้ ซึ่งจะทำได้ง่าย ๆ โดยขยายความสามารถในการยอมรับให้ auxiliary สามารถอยู่ก่อน NP ตัวแรกและใช้โครงสร้างประโยคตาม เน็ตเวิร์ก หลังจาก NP ตัวแรกตามเดิม การที่จะทำเช่นนี้ได้เราจะต้องกำหนดรีจิสเตอร์เพื่อว่าสามารถที่จะวิเคราะห์ผล เช่นเดียวกับที่สามารถวิเคราะห์ในประโยคพื้นฐานได้ ถ้าเราเซ็ทรีจิสเตอร์บางตัว (จะเรียกว่าMOOD) ให้มีค่าเป็น Yes-No-Q ในประโยคคำถามแบบ Yes-No และให้มีค่าเป็น Assertion ในกรณีอื่น ๆ เราก็สามารถจะคงรูปแบบประโยคไว้ได้โดยไม่เสียความหมาย

พาสซีฟ

กริยาเกือบทั้งหมดยกเว้นอกรรมกริยาสามารถทำให้อยู่ในรูปพาสซีฟได้ รูปแบบนี้จะเกี่ยวกับการนำ NP ที่อยู่ในตำแหน่งของกรรมมาเป็น NP ชุดแรกซึ่งทำหน้าที่เป็นประธานและ NP ที่เคยทำหน้าที่เป็นประธานของประโยคอาจจะถูกตัดทิ้ง หรือจะแปลงให้อยู่ในรูปวลีของคำบุพบท

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยใช้คำบุพบท *"by"* รูปพาสซีฟนี้จะต้องมีการเพิ่ม กิริยาช่วยในรูปของ *be* ตามด้วย past participle และการแสดงเท็นส์ในรูปแบบพาสซีฟทั้งหมดจะแสดงอยู่ในรูปของ กิริยาช่วยนี่คือ อย่างประโยคที่เป็นแอคทีฟ

I will hide my hat in the drawer.

I hid my hat in the drawer.

I had hid my hat in the drawer.

I was hiding my hat in the drawer.

I had been hiding my hat in the drawer.

สามารถที่จะแสดงเป็นประโยคในรูปพาสซีฟ ได้ดังนี้

My hat will be hidden (by me) in the drawer.

My hat was hidden (by me) in the drawer.

My hat had been hidden (by me) in the drawer.

My hat was being hidden (by me) in the drawer.

My hat had been being hidden (by me) in the drawer.

เราสามารถที่จะกระทำกับประโยคแบบพาสซีฟ โดยการใช้รีจิสเตอร์ VOICE ให้มีค่าเป็น แอคทีฟ หรือ พาสซีฟ ตามที่เกี่ยวข้องและกำหนดให้ NP ตัวแรกในประโยคให้กับค่ารีจิสเตอร์ FIRST-NP และกำหนดค่าให้กับรีจิสเตอร์ SUBJ หรือ OBJ ตามความเหมาะสม หลังจากที่ทำการวิเคราะห์กลุ่มของกริยาแล้ว

ไวยากรณ์ไม่ควรจะเซ็ทค่า SUBJ เมื่อพบวลีของคำบุพบท ซึ่งมีคำบุพบทเป็น *by* เพราะว่ามันไม่แน่ใจว่ามันจะเป็นการบอกถึงประธานของประโยคจริง ๆ ถ้าไม่มีการตรวจสอบข้อมูลทางด้านความหมายเกี่ยวกับชนิดของ NP ถ้า NP เป็นการบอกถึงการกระทำมันอาจจะเป็นประธานของประโยคได้ ในกรณีอื่น ๆ มันอาจจะเป็นคำขยายธรรมดา ๆ เช่น

The hat was hidden by three o'clock.

พิจารณาว่าการแก้ปัญหาแบบนี้จะทำได้ดีเพียงไรในรูปแบบอื่น ๆ อย่างเช่น ถ้าเรารู้ไวยากรณ์ที่ใช้ในประโยคปกติซึ่งมีส่วนเติมเต็ม เช่นประโยค

Jack wants Sue to kiss George.

และการแก้ปัญหาสำหรับรูปแบบพาสซีฟ ที่เราออกแบบมาสามารถที่จะวิเคราะห์ประโยคได้ถูกต้องเช่น

Sue was wanted to kiss George by Jack.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.2 ประโยคคำถามที่ขึ้นต้นด้วย Wh และกระบวนการโฮล (Wh-Questions and the hold Mechanism *)

ในตอนต้นเราได้กล่าวถึงประโยคคำถามแบบ Wh ไปบ้างแล้วประโยคคำถามแบบ Wh จะมีโครงสร้างประโยคส่วนมากเหมือนกับประโยคคำถามแบบ Yes-No คือจะสลับที่ประธานและ auxiliary สำหรับประโยคคำถามแบบ Wh ได้ ปัญหาที่ใหญ่ที่สุดก็คือ เราจะทำอย่างไรกับองค์ประกอบบางส่วนที่หายไปจากประโยคในบางส่วนเช่น พิจารณาวลีของคำกิริยาในประโยค

What will the fat man angrily put in the corner?

ประโยคนี้เป็นประโยคที่ต้องยอมรับได้ แต่ *angrily put in the corner* ไม่สามารถจะยอมรับได้ ถ้าเราใช้ วลีของคำกิริยา ในไวยากรณ์ เดิม เพราะเราไม่สามารถยอมรับประโยคเช่น *I angrily put in the corner.* ซึ่งวลีของคำกิริยาอย่างนี้จะยอมรับได้ในประโยคคำถามแบบ Wh เท่านั้น และนอกจากนี้ มันจะต้องยอมรับแต่เฉพาะวลีของคำกิริยา ที่สอดคล้องกับองค์ประกอบของประโยคคำถามแบบ Wh อย่างเช่น

What will the fat man angrily put in the corner?

สามารถยอมรับได้ แต่ประโยคต่อไปนี้จะยอมรับไม่ได้

* *Where will the fat man angrily put in the corner?*

จะเห็นว่าไวยากรณ์ ที่มีอยู่ไม่สามารถจะนำมาใช้ได้ ถ้าไม่ทำการเพิ่มเติมสัญลักษณ์บางส่วนให้กับไวยากรณ์

เราจะใช้กลไกในการเก็บองค์ประกอบซึ่งจะถูกใช้ในการเติมองค์ประกอบที่ขาดหายไป จากนั้นเราก็ให้ไวยากรณ์ ตรวจสอบองค์ประกอบที่เก็บไว้นั้นว่ามันไปแทนส่วนที่ช่องว่างได้อย่างเหมาะสมหรือไม่ ถ้ามันพบว่า องค์ประกอบที่เก็บไว้นั้นมีคุณสมบัติเหมาะสมไวยากรณ์ ก็จะทำการตรวจสอบต่อไปโดยทำเหมือนกับว่า องค์ประกอบนั้นได้เข้าไปอยู่ในช่องว่างเรียบร้อยแล้ว ดังนั้นการวิเคราะห์วลีของคำกิริยาในประโยคดังเช่น

What will the fat man angrily put in the corner?

สามารถจะถูกวิเคราะห์ผ่านโดยเหมือนกับว่ากำลังวิเคราะห์ *angrily put what in the corner* และวลีของคำกิริยาในประโยค

What will the fat man angrily put the book in?

จะถูกนำมาวิเคราะห์เหมือนกับว่าวลีของคำกิริยา คือ *angrily put the book in what* นอกจากประโยคที่เห็นข้างต้นแล้ว จะพบว่าการทดสอบกับรูปแบบของความเข้ากันได้ของประธาน-กริยา, รูปแบบในกรณีของสรรพนาม (who,whom), และ สกรรมกิริยาซึ่งอยู่ในรูปของ Wh-term

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สามารถที่จะเติมลงในช่องว่างได้ เช่น ตัวอย่างพิจารณา สกรรมกิริยาจากคำถาม

what did you put in the cupboard?

จะยอมรับได้ถึงแม้ว่า *put* เป็น สกรรมกิริยา และมันต้องการกรรม ซึ่งกรรมนั้นจะถูกแทนที่ด้วย Wh-term ซึ่งตรงกับข้อบังคับของสกรรมกิริยานอกจากนั้นประโยคที่มีกรรมอยู่แล้วจะไม่สามารถยอมรับได้เช่น

* *What did you put the buttle in the cupboard?*

ประโยคนี้ยอมรับไม่ได้เนื่องจากมีกรรมอยู่สองตัวให้กับ คำกิริยา *put* ซึ่งมันจะมีค่าเท่ากับประโยคต่อไปนี้

* *You put what the buttle in the cupborad?*

จะเห็นได้ว่าการทดสอบกับการมีกรรมแบบธรรมดาจะสามารถทำงานได้ ถ้าเราสมมติว่า Wh- เริ่มต้นสามารถที่จะแทรกหรือย้ายถูกเข้าไปในตำแหน่งของ NP ที่เป็นกรรม

<u>กระบวนการโฮลใน ATN (The hold Mechanism in ATNs)</u>

ความสามารถในการเก็บองค์ประกอบเพื่อที่จะใช้ภายหลังในระบบของ ATN จะต้องทำการปรับปรุง ATN หลายแบบ ขั้นแรกเราจะต้องมีโครงสร้างข้อมูลซึ่งเรียกว่าโฮลลิสส์ (hold list) ซึ่งจะจัดการกับองค์ประกอบที่จะนำเข้ามาเก็บ เราสามารถเก็บองค์ประกอบได้หลายตัวไว้ในโฮลลิสส์ในเวลาเดียวกันแต่ในประโยคคำถามที่เรากำลังจะพิจารณากันนี้เราจะเก็บองค์ประกอบเพียงแค่ตัวเดียวในโฮลลิสส์ซึ่งการที่องค์ประกอบจะถูกใส่เข้าไปในโฮลลิสส์ ได้นั้นจะต้องมีแอคชั่นใหม่ที่เกิดขึ้นบนอาร์คซึ่งเรียกว่าโฮลแอคชั่น ซึ่งจะนำองค์ประกอบเข้ามาใส่ใน โฮลลิสส์

โฮลแอคชั่นสามารถที่จะเก็บองค์ประกอบที่อยู่ในรีจิสเตอร์ในขณะนั้น (อย่างเช่น โฮล SUBJ เก็บองค์ประกอบซึ่งอยู่ในรีจิสเตอร์ SUBJ) หรือมันสามารถจะเก็บองค์ประกอบที่สร้างขึ้นมาใหม่ (ตัวอย่างเช่น โฮลซึ่งไม่มี อาร์กิวเม็นต์ จะเก็บองค์ประกอบภายใต้การสร้างโดยเน็ต-เวิร์กในขณะนั้น)

การที่เราจะสามารถวิเคราะห์บางประโยคได้อย่างถูกต้อง และไม่ยอมรับประโยคอื่น ๆ ที่ผิด เราจะต้องเก็บค่าขององค์ประกอบเพื่อที่จะเติมในช่องว่างไว้ตลอดเวลา นอกจากนั้นประโยค เช่น *What did you put the bottle in the cupboard?* จะถูกยอมรับโดยไม่ใช้ค่า Wh-term ที่เก็บไว้เลย (ซึ่งประโยคนั้นจะผิด) ข้อบังคับนี้จะไม่อนุญาตให้ผ่านป๊อปอาร์คไปได้ถ้าไม่ใช้องค์ประกอบทุกตัวที่เก็บไว้โดยแอคชั่นที่อยู่บนอาร์ค หรืออีกนัยหนึ่งก็คือ องค์ประกอบที่ถูกเก็บจะต้องใช้ในการเติมส่วนที่ขาดหายไป ในองค์ประกอบที่ยอมรับในขณะนั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ท้ายที่สุดเราต้องการกลไกบางอย่างเพื่อที่จะตรวจสอบและเติมช่องว่าง กลไกของ ATN ที่ใช้ในอาร์ค ใหม่จะเรียกว่า VIR (เป็นคำย่อของ virtual) ซึ่งใช้ชื่อขององค์ประกอบเป็นอาร์กิวเม็นต์ เราสามารถที่จะไล่ไปตามอาร์คของโฮลลิสล์ขณะนั้น ถ้าเราสามารถไล่ไปตามอาร์คได้สำเร็จ องค์ประกอบจะถูกนำออกมาจากโฮลลิสล์และส่งค่ากลับเป็นค่าของอาร์ค ซึ่งจะเท่ากับค่าที่ส่งกลับมาพุชไวยากรณ์ของ ATN ที่ใช้ในการจัดการประโยคคำถามแบบ Yes-No จะสามารถที่จะนำมาใช้ขยายเพื่อที่จะจัดการกับประโยคคำถามแบบ Wh ได้เช่นกัน กลไกที่ใช้ในการเก็บนี้จะถูกนำมาใช้ทุกๆ กรณีของประโยคคำถามแบบ Wh แม้แต่เมื่อประธานจะเป็นตัวที่เป็นคำถามเอง เช่น ในประโยค *Who won the race?* ในกรณีนี้จะจัดการแยกต่างหากเนื่องจากมันเป็นกรณีเดียวที่ไม่ได้มีการสลับประธานและ auxiliary (sub-aux inversion)

พิจารณาตัวอย่างจากรูปที่ 2.16 ไวยากรณ์ควรจะสามารถจัดการกับช่องว่างซึ่งเป็นของวลีของคำบุพบท, วลีของคำกิริยา และวลีของคำกิริยาวิเศษณ์ได้อย่างเช่นในประโยคดังต่อไปนี้

In which town were you born?

How fast did you run?

What did you do?

ในประโยคแรกวลีของคำบุพบท*in which town* จะถูกเก็บไว้และใช้ในการเติมช่องว่างหลังคำกริยาเหมือน กับเรากำลังวิเคราะห์ประโยค *You were born in which town?* ในประโยคที่สองวลีของคำกิริยาวิเศษณ์ *how* คำกิริยา *fast* จะถูกเก็บและนำเข้าไปในตำแหน่งที่คำกิริยาวิเศษณ์จะอยู่ได้เช่นในประโยค *you ran how fast?* ในกรณีที่ 3 *what* จะถูกเก็บและแทนที่ในตำแหน่งที่เหลือของวลีของคำกิริยา เช่นในประโยค*You did what?*

(S WH-QUERY (NP_1 PRO who
TYPE WH)
MOOD WH-Q
SUBJ -> NP_1
AUXS (is)
VERB carrying
OBJ (NP DET the
HEAD baby))

การวิเคราะห์ประโยค *Who is carrying the baby ?*

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(S WH-QUERY (NP_1 PRO who

TYPE WH)

MOOD WH-Q

AUXS (is)

SUBJ (NP DET the

HEAD baby)

VERB carrying

OBJ -> NP_1

การวิเคราะห์ประโยค *Who is the baby carrying ?*

(S WH-QUERY (NP_1 PRO who

TYPE WH)

MOOD WH-Q

SUBJ -> NP_1

VERB is

OBJ (NP DET the

HEAD baby))

การวิเคราะห์ประโยค *Who is the baby ?*

รูปที่ 2.16 การวิเคราะห์ประโยคคำถามแบบ wh อย่างง่าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5 ซีแมนติคและลอจิคัลฟอร์ม (Semantic and a Logical Form)

จากโครงสร้างทางซินเทคทิคของประโยคที่เรามีอยู่จะเป็นขั้นแรกซึ่งนำไปสู่โมเดลของขบวนการเข้าใจภาษา สิ่งที่สำคัญก็คือความหมายของโครงสร้างประโยคนั้น มันเป็นการยากที่จะกำหนดอย่างแน่นอนว่าประโยคที่เรารู้สองประโยคนั้นมีความหมายเหมือนกันหรือไม่ อย่างเช่นประโยคต่อไปนี้

I gave the contribution to the boyscouts.

The boyscouts receive a donation from me.

ความหมายของประโยคที่มีเป็นคู่ ๆ เหล่านี้ ก็ยังคงถูกนำมาใช้ในประโยคเกี่ยวกับคำถามให้ตอบ

John drove to the store เมื่อเราถามว่า Did John get into a car? คำตอบก็คือ yes การที่จะตอบเช่นนี้ได้ เราจะต้องรู้การกระทำซึ่งจะอธิบายความหมายของคำว่า drive ซึ่งเป็นการกระทำที่จะต้องทำภายในรถ

ดังที่ได้กล่าวไปแล้ว ก็จะทำให้เกิดปัญหาขึ้นมาสองอย่าง ขั้นแรกจะรู้ความหมายที่เหมาะสมของแต่ละคำได้อย่างไรและความหมายเหล่านี้จะประกอบกันเป็นรูปของลอจิคัลฟอร์ม (ต่อไปจะเรียกย่อ ๆ ว่า LF) หลังจากนั้น LF ก็จะตีความตาม Contextual Knowledge ซึ่งจะเกิดผลสรุปว่าประโยคนั้นคืออะไรได้

2.5.1 เหตุผลของการเปลี่ยนไปเป็นลอจิคัลฟอร์ม (Why derive a logical form ?)

เราต้องการการแสดงซีเมนติคเพราะว่า มันจะช่วยแก้ปัญหาสำคัญหลัก ๆ สองประการได้คือ ปัญหาประการแรก จะเกี่ยวข้องกับความเข้าใจความหมายของคำ และประโยคที่มีความหมายกำกวม จะมีคำหลายคำที่มีลักษณะทางซินเทคทิคได้หลายแบบ แต่ละคำในแต่ละส่วนของซินเทคทิคก็จะมีความหมายแตกต่างกัน อย่างเช่นคำกิริยา go อีก 50 แบบ ถึงแม้ว่าคำในประโยคจะมีได้หลายความหมายแต่ถ้านำคำมารวมกันเป็นวลีก็มักจะมีความหมายได้ความหมายเดียว เนื่องจากเนื่องจากแต่ละคำจะมีกฎมาบังคับซึ่งกันและกัน กฎที่เข้ามาบังคับรวมกันอยู่นี้จะเป็นกระบวนการที่ซับซ้อน มีค่าแก่การศึกษา

ปัญหาที่สองจะเกี่ยวกับการใช้ความรู้ที่มีอยู่จริง ๆ และการแสดงคอนเท็กซ์เพื่อบ่งบอกลำดับต่าง ๆ ที่เหมาะสมของประโยค อย่างเช่น ถ้าคุณได้ยินประโยค The president has resigned คุณจะต้องใช้ความรู้ในเหตุการณ์ขณะนั้นเพื่อที่จะระบุได้ว่าใครคือ ประธาน และเป็นประธานในองค์กรไหน และจะต้องสรุปต่อไปได้ว่า บุคคลนี้ไม่บริหารองค์กรนั้นอีกต่อไปแล้ว และนอกจากนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คุณอาจจะยังต้องสรุปผลอื่น ๆ ที่ตามมาขึ้นอยู่กับโอกาสต่าง ๆ อย่างเช่น จะต้องมีการหาประธานคนใหม่มาแทน ในหลักความจริงผลที่ตามมาอาจจะมีได้มากจนนึกไม่ถึง ดังนั้นการพยายามที่จะทำให้มันเป็นเรื่องที่มีรูปแบบที่ชัดเจนขึ้นจะขึ้นอยู่กับการแปลความ หมายของประโยคสุดท้าย ในรูปแบบบางรูปแบบทางลอจิก (logic) และจัดโครงสร้างของข้อสรุปต่าง ๆ ที่แตกต่างกันให้อยู่ในรูปของการอนุมานทางตรรกศาสตร์

กระบวนการแรกที่ได้อธิบายไปนั้น รู้จักกันในรูปแบบของความหมายที่ไม่กำกวมการจัดโครงสร้างกระบวนการดีดักทีฟ (deductive process) จะทำได้ไม่ง่ายนัก ส่วนประกอบต่าง ๆ ของความหมายซึ่งก็คือ ความหมายของคำ ความหมายของวลีและอื่น ๆ มักจะสอดคล้องกับส่วนๆ หนึ่งในทางลอจิก เช่น ชื่อเพรดดิเคทชนิดและอื่น ๆ นอกเหนือจากนี้หลักการอนุมานไม่ได้กำหนดไว้ อย่างเช่นเมื่อเราวิเคราะห์ความหมายของประโยคต่อไปนี้ คุณอาจต้องใช้เพรดดิเคท BREAK ในการจัดโครงสร้างของความหมายของคำกริยานี้

1. John broke the window with a hammer.
2. The hammer broke the window.
3. John broke the window.
4. The window broke.

ยิ่งไปกว่านั้นถ้าเราสมมุติว่า BREAK จะมีอาร์กิวเม็นต์ได้ 3 ตัวคือ ผู้ทำกริยา break, สิ่งที่ถูก break และสิ่งที่ใช้ในการ break การจะวิเคราะห์ประโยคเหล่านี้เราจะต้องรู้ข้อมูลเกี่ยวกับการใช้งานของซีเมนติคของ NP ต่าง ๆ คือ John, the hammer, the windows ซึ่งอยู่ในรูปของซินเทคทิคในประโยค การที่จะตัดสินใจว่าประธาน NP สอดคล้องกับอาร์กิวเม็นต์ตัวที่ 3 ในประโยคที่ 2 และในประโยคที่ 1 และประโยคที่ 3 จะสอดคล้องกับอาร์ กิวเม็นต์ตัวแรก ส่วนอาร์กิวเม็นต์ตัวที่สองจะสอดคล้องกับอาร์กิวเม็นต์ตัวที่ 4 ตรงนี้จะยังไม่แน่ชัดว่าเราสามารถจะสร้างโครงสร้างของกระบวนการดีดักทีฟได้อย่างไร

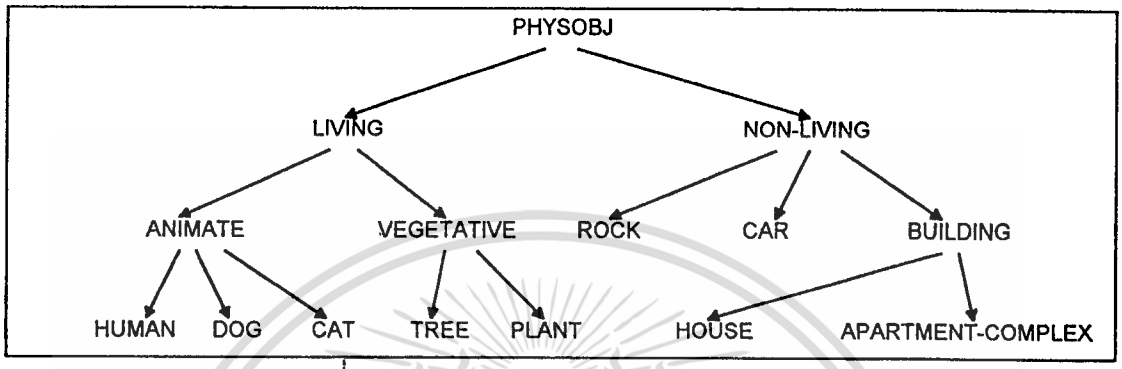
2.5.2 ชนิดและคุณสมบัติ (Type and Feature)

ก่อนที่เราจะระบุความหมายของคำและประโยค เราจะต้องรู้วิธีการบางอย่างในการแสดงความรู้เกี่ยวกับโครงสร้างที่มีอยู่จริงในโลก ส่วนที่สำคัญที่สุดของความรู้นี้ก็คือวิธีที่ออปเจ็คในโลกจะถูกแบ่งออกเป็นกลุ่มขึ้นอยู่กับคุณสมบัติของมัน คุณสมบัติที่เป็นพื้นฐานที่สุดตัวหนึ่งของออปเจ็คใด ๆ ก็คือ ชนิดของมัน ซึ่งเป็นการบอกว่าออปเจ็คนั้นเป็นอะไร อย่างเช่นออปเจ็คบางตัวจะถูกแบ่งให้เป็นสุนัข และออปเจ็คอื่น ๆ อาจจะถูกแบ่งให้เป็นแมว หรือถ้าเกิดจะมองให้ใหญ่ไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กว่านั้นออปเจ็คตัวเดียวกันนี้อาจจะถูกแบ่งให้เป็นสัตว์ หรือเป็นสิ่งมีชีวิตก็ได้

เราสามารถจะสร้างลำดับชั้นของชนิด (type hierachies) ซึ่งใช้อธิบายสิ่งที่มีอยู่จริงดังแสดงในรูปที่ 2.17 สังเกตว่า มีเพียงสองสามสิ่งเท่านั้นใน hierachies ตัวอย่างนี้



รูปที่ 2.17 Physical Object Hierarchy

สิ่งสำคัญก็คือจะต้องพิจารณาว่า เราจะนำไทพ์ไหนมาใช้ อย่างเช่น คุณอาจจะต้องการแบ่งแยกความแตกต่างระหว่างออปเจ็คซึ่งสามารถจะถูกคนหิ้วไปได้ และออปเจ็คที่ถูกคนหิ้วไม่ได้ จะเห็นได้ชัดว่า ถ้าคุณใส่ไทพ์แบบไม่มีข้อจำกัด โครงสร้างไทพ์ของคุณก็จะยุ่งยากมากและคุณจะเสียประโยชน์ในการใช้งาน ยิ่งไปกว่านั้นจะทำให้เกิดปัญหาใหม่ขึ้นมาในการตีความของซีเมนติค ถ้าคุณจงใจที่จะสร้างไทพ์ใหม่เพื่อที่จะแก้ปัญหาหนึ่งในประโยคนั้น ทางที่ดีเราควรหลีกเลี่ยงความยุ่งยากเหล่านี้ โดยเราจะอนุญาตให้เพียงแต่ไทพ์ที่มีความสอดคล้องกับคำเดี่ยว ๆ อย่างเช่น DOG จะยอมรับเป็นไทพ์ได้ ในขณะที่ "ความสามารถที่จะให้คนหิ้ว (carryable by one person)" จะยอมรับไม่ได้ เนื่องจากคุณสามารถที่จะสร้างไทพ์ของ "carryable" ซึ่งจะไม่เจาะจงว่าจะต้องทำโดยคนหนึ่งคนสิบคน หรือว่าใช้รถเครนในการยก คำจำกัดความเหล่านี้ไม่รวมถึงไทพ์อย่างพวก KID ซึ่งหมายถึง ลูกแพะและลูกคนด้วย ซึ่งพวกนี้จะอยู่ในคลาสอื่น ๆ แล้ว อย่างเช่น ประโยค I have to kids and Gorge has three. อาจจะตีความหมายได้ว่า ทั้ง Gorge และ I เป็นคนเลี้ยงแพะ หรือตีความหมายได้ว่า พวกเรามีลูก แต่ว่ามันจะเป็นความหมายรวมทั้งสองอย่างไม่ได้ (ก็คือเป็นไปไม่ได้ที่ฉันจะมีแพะและ Gorge จะมีลูก) กฎเดียวกันนี้จะถูกนำมาใช้ในกรณีของกริยาได้เช่นเดียวกัน ดังประโยค *I ran last year and Gorge did too.* จะหมายความว่า เราทั้งสองคนชนะการเลือกตั้ง หรือเราอาจจะไปวิ่งในการแข่งขันบางที่ (run หมายถึง ได้คะแนนต่อเนื่องกันหรือหมายถึงวิ่งแข่งก็ได้) ถ้าไทพ์นั้นสอดคล้องกับคำที่ทำให้การทดสอบล้มเหลว คำ ๆ นั้นก็จะเป็นคำที่กำกวม และความหมายจะขึ้นอยู่กับที่ระบุไว้ในไทพ์นั้น

โดยการข้อกำหนดที่กล่าวไว้ในเบื้องต้นไทพ์ใด ๆ ที่มีสับไทพ์จะมีความกำกวม แต่ว่าสิ่งก็ไม่

เอกสารนี้ ... ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทุกกรณีอย่างเช่นไทพ์ HORSE จะถูกแบ่งเป็นชนิดต่าง ๆ คือ MARE (ม้าตัวเมีย) COLT (เด็กหนุ่ม) TROTER (สัตว์ที่มีลักษณะเหมือนม้า) และอื่น ๆ ซึ่งเราจะยอมรับได้ ถ้าพูดว่า I have one hot and Gorge has too. ถึงแม้ว่าจะเป็น I have a colt and Gorge has mares. เนื่องจาก hot เป็น reasonable type

ลักษณะอื่น ๆ ก็จะมีความเกี่ยวข้องกับไทพ์ทั้งนั้น ในความจริงคำกริยาทุกตัวจะสามารถแบ่งเป็นไทพ์ได้เช่นกัน ในกรณีนี้ไทพ์จะถูกแบ่งเป็น เหตุการณ์และสถานะ (event and state) ในคลาส EVENT จะมีคำกิริยาอย่างเช่น WALK, DONATE, FALL, SLEEP และอื่น ๆ ในขณะที่คลาส STATE จะมีคำกิริยาพวก BE (อย่างเช่น BE HAPPLY), BELIEVE, WANT และ OWN คลาสของคำกิริยาเหล่านี้จะแตกต่างกันทางการทดสอบทางภาษาโดยเฉพาะอย่างยิ่ง stative คำกิริยาซึ่งใช้ในการอธิบายสถานะจะไม่สามารถใช้ในรูปของ progressive อย่างเช่นคุณไม่สามารถพูดว่า

*I am beliveing that Jack caught a rabbit.

*I am owning a car.

Nonstative Verb ซึ่งใช้แสดงเหตุการณ์ สามารถที่จะใช้ progressive แต่จะต้องใช้เฉพาะใน simple present ถ้าการกระทำซ้ำ ๆ กันนั้นกำลังถูกอธิบาย อย่างเช่น I walk to the store. จะหมายความว่า I walk to the store regularly. คือจะต้องทำเป็นประจำ subclass ที่สำคัญของ event ก็คือแอคชั่นที่ใช้อธิบายโดย agent หรือสิ่งที่ทำให้เกิดเหตุการณ์นั้น

มันไม่ยากที่จะหาตัวอย่างที่แสดงไทพ์ที่ไม่สามารถจัดให้อยู่ใน hierachies ธรรมดาได้ อย่างเช่น คำว่า pet ซึ่งอาจจะจัดให้อยู่ใน type PET ซึ่งจะเห็นว่ามันไม่สามารถที่จะเข้าไปอยู่ในรูปใดในรูปที่ 2.17 ได้เลย ตัวอย่างเช่น สุนัขส่วนมากจะเป็นสัตว์เลี้ยงแต่สุนัขบางตัวเช่นสุนัขจรจัดและหมาป่าจะไม่เป็นสัตว์เลี้ยง ดังนั้น PET จึงไม่สามารถจะนำไปใส่ไว้เหนือหรือใต้ DOG ได้เลย โดยทั่วไปแล้วความสัมพันธ์ตามอำเภอใจ จะถูกกำหนดให้อยู่ระหว่างไทพ์ซึ่งมันอาจจะอินเตอร์เซ็คกันอย่างเช่น PET กับ DOG หรือดิสจอยท์ (disjoint) กัน อย่างเช่น LIVING และ NON-LIVING หรือบางไทพ์อาจจะอินเตอร์เซ็คกับไทพ์อื่น ๆ หลาย ๆ ไทพ์เช่น ถ้าคุณมีไทพ์ BABY ADOUT ดังนั้นความหมายหนึ่งของ KID อาจจะอินเตอร์เซ็คกับ BABY ของ HUMAN และอีกความหมายหนึ่งอาจจะไปอินเตอร์เซ็คกับ BABY ของ GOAT สิ่งที่สำคัญที่สุดในที่นี้ ก็คือไทพ์บางชุดที่มีอยู่จะมีซับเซ็ทใหญ่มากและเซ็ทเหล่านี้สามารถที่จะจัดให้เป็นโครงสร้างใหม่ได้

ตัวอย่างของไทพ์ที่ซับซ้อนซึ่งมีประโยชน์คือไทพ์ของ ORG ซึ่งเป็นไทพ์ที่รวมออปเจ็คใด ๆ ที่มีความสามารถในการเป็นตัวแทนหรือการกระทำ ดังนั้นไทพ์นี้ก็จะมีออปเจ็คของไทพ์ ANIMATE, ออปเจ็คของ LEGAL-ENTITY อย่างเช่น บริษัทและออปเจ็คของไทพ์ AUTOMATA อย่างเช่น เครื่อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขายของอัตโนมัติ มีคลาสใหญ่ ๆ บาง คลาสที่เป็นนามธรรมก็จะถูกจัดให้เป็นไทพ์ด้วยอย่างเช่น ไทพ์ของ COLOR, TIME, LOCATION, IDEAS, GROUPS และอื่น ๆ ซึ่งจะกล่าวถึงในโอกาสต่อไป

2.5.3 **ซีเล็คชันนัลเรสตริกชัน** (Selectional Restriction)

นอกจากการแบ่งความหมายของคำต่าง ๆ แล้ว คุณจะต้องกำหนดด้วยว่าคำและวลีจะติดต่อกับคำอื่นและวลีอื่นเพื่อที่จะสร้างเป็นซีเมนติคของวลีที่ใหญ่กว่าได้อย่างไร พิจารณางานง่าย ๆ ในการเดาความหมายของคำที่มีความหมายกำกวมทำได้โดยการพิจารณาว่า มันเกี่ยวข้องกับคำที่อยู่ใกล้ ๆ มันอย่างไรบ้าง ข้อจำกัดเหล่านี้ซึ่งจะเรียกว่า **ซีเล็คชันนัลเรสตริกชันซึ่งจะแสดงให้เห็นถึงเซ็ทของรูปแบบที่เหมือนกับโครงสร้างของซินเทคทิค**

ยกตัวอย่างพิจารณาคำคุณศัพท์ *green* ถ้า *green* ใช้เป็นคำคุณศัพท์ มันจะขยายฟิสิคัลออบ-เจ็ค *The green cow* เป็นรูปแบบของ NP แต่ *The green belief* ไม่ใช่ NP ข้อจำกัดนี้จะถูกแปลงให้อยู่ในรูปแบบของเล็กซิคัล! ของ *green* โดยการขยายเล็กซิคัลนี้เพื่อให้สามารถจัดการกับเซ็ทของรูปแบบที่อธิบายกฎเกณฑ์ทางความหมายของวลีที่เก็บคำนี้ อย่างเช่น ข้อจำกัดอาจจะแสดงได้ในรูปแบบนี้ (NP ADJS (green)

HEAD + physobj)

รูปแบบนี้จะเหมือนกับโครงสร้างที่เก็บคำ green ในลิสต์ของ ADJS โดยที่คำนี้มีความหมายสอดคล้องกับ type PHYSOBJ ในส่วนของ HEAD ข้อจำกัดของไทพ์ที่ซับซ้อนจะมีได้ในแพทเทิร์น (pattern) ต่อไปนี้โดยเฉพาะอย่างยิ่ง ข้อจำกัดของ *+physobj-animate* จะมีได้ในความหมายใด ๆ ซึ่งเป็นของ physobj แต่ไม่เป็นของ animate

อีกความหมายหนึ่งของ green คือ เขียว อย่างเช่นใน *The green recruits* ความหมายนี้ก็จะต้องมีใน lexicon ซึ่งมี Selection Restriction ต่างกัน และบังคับให้คำนำหลักต้องเป็นคนอย่างเช่น

(NP ADJS (green)

HEAD + human)

Selection Restriction ของคำกิริยาอาจจะจำกัดชนิดของ NP ของประธาน การที่จะทำเช่นนี้ได้คุณจะต้องมีกลไกบางอย่างที่แปลงมาจากไทพ์สำหรับองค์ประกอบที่ใหญ่กว่าคำ ๆ เดียวอย่างเช่น สมมติว่าไทพ์ขององค์ประกอบที่จะมาเป็นไทพ์ของคำหลัก (head word) ของมัน (ซึ่งก็คือ HEAD ของ NP PREP ของ PP และ MAIN-V ของ S) โดยใช้กลไกอันนี้เราสามารถที่จะแสดงซีเล็ค-ชันนัลเรสตริกชันของคำกิริยา die ซึ่งประธานจะต้องมีชีวิต (animate)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(S SUBJ + animate

MAIN-V die)

กฎนี้รวมกับกฎของ *green* จะยอมรับประโยค *The green sole died* ซึ่งจะถูกวิเคราะห์และกำหนดความหมายของคำที่เหมาะสมให้ โดยเฉพาะอย่างยิ่งคำนาม *sole* จะมีความหมายของการไม่มีชีวิต (the sole of a shoe-ส้นรองเท้า) และมีความหมายของสิ่งมีชีวิต (ปลาลิ้นหมา) ซึ่งการวิเคราะห์ NP *The green sole* จะไม่มีความแตกต่างระหว่างความหมายเหล่านี้เพราะว่า ทั้งสองตัวก็เป็น physobj เหมือนกัน แต่เมื่อมีข้อจำกัดของคำกิริยาขึ้นมาก็จะบังคับให้ NP ชุดนี้จะต้องเป็นสิ่งมีชีวิต ดังนั้น sole ในความหมายของปลา จะถูกกำหนดขึ้นมา

ไม่ใช่มีเพียงแต่ adj และคำกิริยาเท่านั้นที่มีซีเล็คชันนัลเรสตริกชัน คำนาม และคำบุพบท ก็มีได้เช่นกัน อย่างเช่น คำนาม president จะต้องการ pobj ซึ่งเป็นของ PP มาขยายด้วยคำบุพบท *of* โดยเป็น type+legal_entity หรืออีกนัยหนึ่งคำบุพบท *in* ถ้าใช้ในความหมายของการบรรจุ ก็จะต้องการ POBJ ที่เป็น +physobj

2.5.4 Case Relation

significant generalization จะเป็นเรื่องที่เกี่ยวข้องกับ NP มีความสัมพันธ์ทางความหมายกับคำกิริยาและ adj ในประโยคอย่างไร สิ่งที่มีประโยชน์ในการคำนวณความสามารถเช่นนั้นก็คือ case grammar และตัวตาม (successor) และส่วนขยายของมัน จุดประสงค์ในขณะนี้ case grammar ได้วางโครงสร้างให้ครอบคลุมกฎเกณฑ์ทางซีเมนติคซึ่งเรียกว่าเคสซึ่งอาจจะใช้ NP ร่วมกับคำกิริยาหรือ adj สิ่งที่น่าสนใจอย่างหนึ่งของ case grammar ก็คือจำนวนของความเป็นไปได้ในความสัมพันธ์ของซีเมนติคมีน้อยมากถึงแม้ว่ามันมีความสอดคล้องกันน้อยซึ่งประกอบกันอยู่ในเซ็ทเล็ก ๆ นี้ วิธีการที่กำลังพูดถึงนี้จะสมมติว่า เซ็ทของซีเมนติคเคสเป็นไปได้อยู่ในรูปแบบที่กำหนดโครงสร้างไว้ดีแล้ว และจำนวนของสับเคส (subcase) ในขณะนั้นของเคสใด ๆ จะถูกจำกัดให้มีจำนวนน้อย ๆ โครงสร้าง (frameworks) ที่กำหนดไว้นี้จะมีเซ็ทของเคสที่แตกต่างกันมากมาย ซึ่งสามารถอธิบายความจริงที่ว่าพวกมันทั้งหมดจะมีพื้นฐานในการวิเคราะห์ซีเมนติคในชั้นที่แตกต่างกันของลำดับชั้น

เนื่องจากจุดประสงค์ในการวิเคราะห์ของเรานั่นคือ การทำการแยกแยะความหมายของประโยค การวิเคราะห์ประโยคที่มีความหมายแตกต่างกันจะให้ผลลัพธ์ที่แตกต่างกัน ประโยคที่มีโครงสร้างทางซินเทคติคที่แตกต่างกันแต่มีความหมายเหมือน จะต้องถูกนำมาแมพเป็นโครงสร้างที่มีลักษณะเหมือนกัน อย่างเช่น พิจารณากลุ่มของประโยคต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

John broke the window with a hammer.

The hammer broke the window.

The window broke.

John, The hammer และ *The window* จะมีกฎทางซีเมนติคเหมือนกันในแต่ละประโยค John เป็นผู้กระทำ The window เป็นกรรม และ The hammer เป็นเครื่องมือที่ใช้ในการทำกิริยา "breaking of the window"

การที่จะทำสิ่งนี้ให้ชัดเจนขึ้น เราจะกำหนดเคสเฉพาะขึ้นมาเพื่อที่จะใช้ต่อไป บางทีวิธีที่ง่ายที่สุดในการกำหนดก็คือ AGENT case NP จะทำให้ AGENT case สมบูรณ์ ถ้ามันอธิบายถึงผู้เร่งเร้า (instigator) ในการกระทำของประโยคนั้น ยิ่งไปกว่านั้นเคสนี้จะแสดงลักษณะของความตั้งใจ การฝ่าฝืนหรือการตอบสนองใน แอคชั่นที่เอเจ้นท์ (agent) อธิบายนั้น การทดสอบ AGENT-hood ทำได้โดยเพิ่มวลีอย่างเช่น *intentionally* หรือ *in order to* ในประโยคที่เป็นรูปแบบแอคทีฟถ้าประโยคที่เป็นผลลัพธ์อยู่ในรูปที่ถูกต้อง NP ที่เป็นประธานก็ทำให้ AGENT สมบูรณ์ ตัวอย่างเช่น ประโยคต่อไปนี้จะยอมรับได้

John intentionally broke the window.

John broke the window in order to let in some air.

แต่สองประโยคนี้จะไม่ยอมรับ

*The hammer intentionally broke the window.

*The window broke in order to let in some air.

ดังนั้น NP *John* จะทำให้ AGENT case ในสองประโยคแรกข้างต้นสมบูรณ์

สิ่งมีชีวิตที่เป็น NP บางตัวถึงแม้ว่าจะอยู่ในตำแหน่งของประธานไม่สามารถทำให้ AGENT สมบูรณ์ อย่างเช่น คุณไม่สามารถที่จะพูดได้ดังนี้

*John intentionally died.

*Mary remembered her birthday in order to get some presents.

แน่นอนว่าการเพิ่ม *intentionally* ในประโยคเหล่านี้ อาจจะเป็นไปได้ที่ประโยคนี้จะถูกต้อง (John อาจจะฆ่าตัวตายเอง) แต่ผลลัพธ์นี้จะเนื่องมาจากการเพิ่มเติมความหมายดั้งเดิมซึ่งก็คือ *John died*

NP ที่อธิบายบางอย่างที่กำลังเปลี่ยนแปลงหรือกำลังกระทำจะทำให้เคสใหม่อีกเคสหนึ่งซึ่งเรียกว่า THEME สมบูรณ์ สิ่งนี้มักจะสอดคล้องกับออปเจ็คของซินเทคทิคสำหรับสกรรมกิริยา X ซึ่งเป็นคำตอบของคำถาม "What was Xed?" อย่างเช่น ให้ประโยค The gray eagle saw the

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

mouse. NP ที่ชื่อว่า the mouse เป็น THEME และเป็นคำตอบของคำถาม "What was seen?" สำหรับสกรรมกิริยา THEME จะใช้สำหรับ NP ที่เป็นประธานและไม่ได้เป็น AGENT อย่างเช่น *The clouds appeared over the horizon* NP ชื่อ *the clouds* จะทำให้ THEME สมบูรณ์ ตัวอย่างต่อไปนี้จะขีดเส้นใต้ที่ NP ที่เป็น THEME

The rock broke.

John broke the rock.

I gave John the book.

เคสต่อไปที่จะแนะนำอีก 3 ตัวคือ AT, TO, FROM ที่จะต้องใช้งานร่วมกับคำอธิบายสถานะของออปเจ็คและการเปลี่ยนแปลงของสถานะนั้น ๆ พิจารณาตัวอย่างของตำแหน่ง (location) คุณสามารถจะกำหนดตำแหน่ง ของออปเจ็คบางตัวหรือกิจกรรมบางอย่างโดยใช้วลีของคำบุพบทดังเช่น

Harry walked along the road.

We stayed by the door.

นี่คือเคสของการอธิบายว่าจะเดินไปที่ไหนและเราอยู่ที่ไหน ซึ่งกรณีทั้งหมดนี้จะทำให้ AT-LOC เป็นเคสที่สมบูรณ์

ประโยคอื่น ๆ ที่อธิบายการเปลี่ยนแปลงของ location ได้แก่

I walked from here to school yesterday.

It fell to the ground.

The birds flew from the lake along the river gorge.

วลีของคำบุพบทที่อธิบายว่าอะไรบางสิ่งมาจากไหน จะทำให้ FROM-LOC สมบูรณ์ (หรือ SOURCE) หรือในขณะที่อีกตัวหนึ่งอธิบายว่าสิ่งนั้นจะไปไหน จะทำให้ TO-LOC (หรือ DESTINATION) ประโยคอาจจะมีเคสของ ตำแหน่งครบทั้ง 3 กรณีอย่างเช่น

They drove from the farm to the lake on route 71.

case location ทั้ง 3 นี้ สามารถจะแปลงให้เป็นเคสอื่น ๆ ตามสถานะต่าง ๆ ที่มีอยู่ซึ่งเรียกว่า AT case และเคสของแต่ละสถานะที่เปลี่ยนแปลง (FROM และ TO case) ดังนั้น AT-LOC ก็คือรูปแบบพิเศษของ AT case, FROM-LOC ก็คือรูปแบบพิเศษของ FROM case, TO-LOC ก็คือรูปแบบพิเศษของ TO case คุณจะเห็นรูปแบบพิเศษอื่น ๆ ของเคสเหล่านี้เมื่อคุณพิจารณาประโยคที่แสดงความเป็นเจ้าของดังเช่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

I threw the ball to John.	(the TO-LOC case)
I gave a book to John.	(the TO-POSS case)
I caught the ball from John.	(the FROM-LOC case)
I borrowed a book from John.	(the FROM-POSS case)
The box contains a ball.	(the AT-LOC case)
John owns a book.	(the AT-POSS case)

คุณอาจจะกำหนด AT-TIME, TO-TIME, FROM-TIME ได้เช่นเดียวกันกับในสถานะทั่วไปและการเปลี่ยนสถานะเช่นใช้กับอุณหภูมิ (temperature) ดังต่อไปนี้

The temperature remains at zero.	(AT-VALUE)
The temperature rose from zero.	(FROM-VALUE)

ดังนั้นสัญลักษณ์ของค่าทั่วไปและการเปลี่ยนแปลงค่าในหลาย ๆ แบบ จะถูกสนับสนุนในวิธีการเดียวกันกับ REALIZE case เหล่านี้ในประโยค

John believed that it was raining.

THEME case จะถูกทำให้สมบูรณ์ด้วยอนุประโยค that it was raining เนื่องจากนี่คือสิ่งที่ถูกเชื่อ John ไม่สามารถจะเป็น AGENT ในที่นี้เนื่องจากไม่มีความตั้งใจในการเชื่อในบางสิ่ง ดังนั้นคุณจะต้องสร้างเคสใหม่ซึ่งจะเรียกว่า EXPERIENCER ซึ่งจะถูกทำให้สมบูรณ์โดยออปเจ็คที่มีชีวิตที่อยู่ในสถานะที่อธิบายฟิโซโลจี (phychology) หรือได้รับกระบวนการทางฟิโซโลจีบางอย่างอย่างเช่น ความรู้สึก (perception) ดังในประโยคที่แล้วและในประโยคต่อไปนี้

John saw the unicorn.

เคสอื่น อีกเคสหนึ่งก็คือ BENIFICIARY case ซึ่งจะถูกทำให้สมบูรณ์โดยบุคคลที่มีชีวิตสำหรับใครก็ตามที่ถูกจัดการโดยบางเหตุการณ์

I rolled on the floor for Lucy.

Find me the papers!

I gave the book to Jack for Susan.

ในตัวอย่างสุดท้ายจะแสดงถึงความต้องการในข้อแตกต่างของ TO-POSS case (ก็คือ *to Jack*) จาก BENIFICIARY case

INSTRUMENT case จะอธิบายถึงวัสดุ วัตถุดิบหรือแรงงานเพื่อใช้สร้างเหตุการณ์ขึ้นบางเหตุการณ์

Jack saw the ship with the telescope.

Jack used the telescope to see the ship.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Harry broke the glass with the telescope.

The telescope broke the glass.

I used some flour to make a cake.

I made a cake with some flour.

ขึ้นอยู่กับคำกิริยาแต่บางครั้ง INSTRUMENT case สามารถจะใช้เป็น surface subject เมื่อไม่ได้ระบุ AGENT case ไว้แรงตามธรรมชาติก็ยังสามารถให้มีคุณสมบัติของ INSTRUMENT ได้ในที่นี้ ถึงแม้ว่าคุณอาจจะไม่เห็นด้วยในการวิเคราะห์แบบที่แตกต่างกัน ดังนั้นประโยคต่อไปนี้ก็เป็นตัวอย่างของ INSTRUMENT case ด้วยเช่นกัน

The sun dried the apples.

Jack used the sun to dry the apples.

AGENT และ INSTRUMENT case สามารถใช้ร่วมกันและจะเรียกว่า CAUSAL-AGENT

เคสอื่น ๆ ต้องการถูกระบุก่อนที่ประโยคนั้นจะถูกวิเคราะห์อย่างเช่น บางประโยคที่อธิบายสถานการณ์เมื่อคนสองคนกระทำซึ่งกันและกันอย่างเช่น

Henry lifted the piano with Jack.

เพื่อที่จะจัดการกับประโยคเช่นนี้ จะต้องมี case CO-AGENT เพื่อที่จะบันทึก PP with Jack

เคสที่ซับซ้อนกว่านี้จะปรากฏในประโยคที่เกี่ยวกับการแลกเปลี่ยนหรือการกระทำโต้ตอบกันที่ซับซ้อน อย่างเช่นพิจารณาประโยค

Jack paid $1 to the man for the book.

Jack bought the book from the man for $1.

ประโยคทั้งสองนี้จะอธิบายเหตุการณ์เมื่อ Jack ให้เงิน the man ไป 1 ดอลล่าร์และรับหนังสือมาเป็นการแลกเปลี่ยน ในประโยคแรกตัว *$1* เป็น THEME และไม่มีเคสที่จะบันทึก *the book* ในประโยคที่สองสถานการณ์จะกลับกัน *the book* เป็น THEME และ *$1* จะไม่ถูกบันทึก เพื่อที่จะจัดการกับเคสเหล่านี้ คุณจะต้องเพิ่ม case CO-THEME สำหรับออปเจ็คตัวที่สองของการแลกเปลี่ยน

วิธีการแก้ปัญหาที่ง่ายกว่านี้ก็คือ จะวิเคราะห์ประโยคที่อธิบายสองเหตุการณ์โดยเหตุการณ์แรก (primary event) คือสิ่งหนึ่งที่คุณจะต้องพิจารณาต่อไป แต่เหตุการณ์ที่สอง (second event) ก็จะยังคงอยู่ ในการวิเคราะห์คุณอาจจะวิเคราะห์ประโยคแรกได้ดังต่อไปนี้ (เหตุการณ์แรกก็คือ การที่ Jack จ่ายเงินและ เหตุการณ์ที่สอง คือ Jack ได้รับหนังสือ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5.5 โครงสร้างของคำกิริยา (The Structure of Verbs)

ดังที่ได้เห็นไปแล้วคำกิริยาสามารถจะถูกแบ่งได้เป็นเคสตามความหมายตามที่ต้องการ การที่จะแบ่งสิ่งเหล่านี้ได้อย่างถูกต้อง คุณจะต้องแยกแยะข้อแตกต่างระหว่างเคสแต่ละเคสซึ่งมีความสัมพันธ์กันอย่างใกล้ชิดกับคำกริยาและคำที่ไม่ใช่กริยานั้น อย่างเช่นคำกิริยาที่เป็น Past tense เกือบทั้งหมด จะยอมรับ case AT-TIME ซึ่งรู้มาจากคำกิริยาวิเศษณ์ *yesterday* ซึ่งในกรณีอย่างนี้จะขึ้นอยู่กับความเหมาะสมของ VP มากกว่าที่จะมาพิจารณาคำกิริยาเดี่ยว ๆ เพียงตัวเดียว อย่างไรก็ตามในเคสอื่น ๆ จะรู้ได้โดยพิจารณาองค์ประกอบที่อยู่ใน subcategorizes ของคำกิริยา อย่างเช่น คำว่า *put* จะมีสับแคทสำหรับ PP และตัวอื่น ๆ PP ตัวนี้จะรู้ได้โดยใช้ case TO-LOC ในการแบ่งคำกิริยาโดยแยกแยะตามประเภทของเคสนั้นสำคัญมาก และเคสเหล่านี้จะเรียกว่า inner case ของคำกิริยา

ในการทดสอบอื่นขึ้นอยู่กับการสำรวจว่าคำกิริยาทั้งหมด จะสามารถมี NP ได้อย่างมากที่สุดหนึ่งตัวใน inner case ถ้ามันต้องการ NP หลายตัวมันจะต้องเชื่อมต่อกันด้วยคำสันธาน (conjuction) ดังที่จะได้เห็นต่อไปนี้

John and I ran to the store.

แต่ว่าประโยคต่อไปนี้จะผิด

*John I ran to the store.

และเช่นเดียวกันคุณสามารถจะกล่าวได้ว่า

I ran to the store and to the bank.

แต่ประโยคนี้ก็จะผิดเช่นกัน

*I ran to the store to the bank.

ดังนั้น case AGENT และ TO-LOC สำหรับคำกิริยา *run* จะเป็น inner case ลองเปรียบเทียบกับประโยคนี้

Case	Sub-Cases	Common Names	Definition
CAUSAL-AGENT			- the object that caused the event to happen
	AGENT		- intentional causation
	INSTRUMENT		- force or tool used in causing the event
THEME			- the thing that was affected by the event; with a transitive verb X , it is the thing in answer to the question "What was Xed?"

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

EXPERIENCER			- the person who is involved in perception or in a psychological state
BENEFICIARY			- the person for whom some act is done
AT			- the state/value on some
	AT-LOC	LOCATION	dimension where the event
	AT-POSS	POSSESSOR	occured
	AT-VALUE		
TO			- final value in a state change
	TO-LOC	DESTINATION	- final location
	TO-POSS	RECIPIENT	- final possessor
	TO-VALUE		- final temperature
FROM			- original value in a state change
	FROM-LOC	SOURCE	- original location
	FROM-POSS		- original possessor
	FROM-VALUE		- original temperature
CO-AGENT			- a secondary agent in an action
CO-THEME			- a secondary theme in an exchange

รูปที่ 2.18 กฎทางซีเมนติค

Jack	- AGENT of both PRIMARY and SECONDARY event
$1	- THEME of PRIMARY event
the man	-TO-POSS of PRIMARY ,FROM-POSS of SECONDARY
the book	- THEME of SECONDARY event

รูป.18 แสดงข้อสรุปของเคสที่ใช้มากที่สุดและลักษณะโครงสร้างความสัมพันธ์ระหว่างเคสแต่ละตัว

I saw the play in Rochester at the Eastman Theater.

ซึ่ง case AT-LOC ทั้งสองอันจะใช้ได้อย่างถูกต้อง เราสามารถที่จะระบุคำกิริยาให้มี inner case ได้ถึง 3 inner case และจะต้องมีอย่างน้อย 1 เคสที่จะถูกรับรู้ในประโยคโดยใช้คำกิริยาในบางครั้งเคสเฉพาะบางตัวจะต้องเป็น present (อย่างเช่น TO-LOC ที่ใช้กับคำกิริยา *put*) และนอกจากนั้น THEME case จะต้องมีข้อบังคับเช่นเดียวกัน ในขณะที่ AGENT case จะเป็นตัวเลือกสำหรับคำกิริยาตัวใด ๆ ที่ยอมรับรูปแบบ passive

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มีข้อกำหนดทางไวยากรณ์สำหรับที่จะวิเคราะห์เคสหลายแบบเช่นเดียวกัน รูป 2.19 แสดงตัวอย่างของวิธีการที่เคสแต่ละเคสจะถูกวิเคราะห์ในประโยคที่ต่างกัน

Case	Realization
AGENT	- as SUBJECT in active sentence
	- preposition *by* in passive sentence
INSTRUMENT	- as SUBJECT in active sentences with no AGENT
	- preposition *with*
EXPERIENCER	- as SUBJECT in active sentences with no AGENT
BENEFICIARY	- as IOBJ with transitive verbs
	- preposition *for*
AT-LOC	- prepositions *in, on beyond, etc.*
AT-POSS	- possessive NP
	- SUBJECT of sentence if no AGENT
TO-LOC	- preposition *to, into*
TO-POSS	- preposition to, IBOJ with certain verbs
FORM-LOC	- preposition *from, out of, etc.*
FORM-POSS	- preposition *form*

รูปที่ 2.19 ข้อมูลทั่ว ๆ เกี่ยวกับเคสหลัก

ต่อไปนี้เป็นตัวอย่างประโยคซึ่ง PP,NP และอนุประโยคได้ถูกแบ่งโดยเคสที่เกิดขึ้นของมัน

Jack ran.	AGENT only
Jack ran with a crutch.	AGENT + INSTRUMENT
Jack ran with a crutch for Susan.	AGENT + INSTRUMENT + BENEFICIARY
Jack destroyed the car.	AGENT + THEME
Jack put the car through the wall.	AGENT + THEME + TO-LOC
Jack sold Henry the car.	AGENT + TO-POSS + THEME
Henry pushed the car from Jack's house to the junkyard.	AGENT + THEME + FROM-LOC + TO-LOC
Jack is tall.	THEME
Henry believes that Jack is tall.	EXPERIENCER + THEME
Susan owns a car.	AT-POSS + THEME
I am in the closet.	THEME + AT-LOC
The ice melted.	THEME
Jack enjoyed the play.	EXPERIENCER + THEME
The ball rolled down the hill to the water.	THEME + AT-LOC + TO-LOC

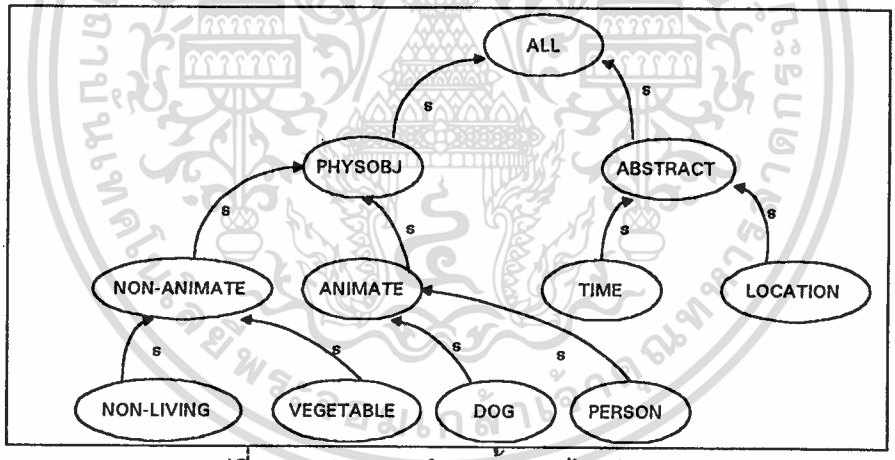
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5.6 ซีแมนติคเน็ตเวิร์ก (Semantics Networks)

คำว่าซีแมนติคเน็ตเวิร์ก จะใช้ในแสดงรูปแบบได้แตกต่างกันและใช้ในจุดประสงค์ที่แตกต่างกันได้หลายแบบ เราจะใช้รูปแบบที่เป็นทางการเพื่อแสดงข้อมูลทางสถิติอย่างเช่น การแสดง word sense และเลือกข้อจำกัดระหว่าง sense โดยเฉพาะพวกที่เป็นเคสของความสัมพันธ์ สัญลักษณ์แรกจะกำหนดสำหรับการแสดงข้อมูลทั่ว ๆ ไปและต่อมาก็จะขยายเพื่อให้แสดงประโยคที่มีลักษณะเฉพาะมากขึ้น

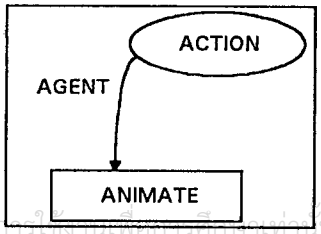
Representing general information about word senses

ซีแมนติคเน็ตเวิร์กคือโครงสร้างที่เป็นรูปกราฟที่มีสัญลักษณ์ที่เป็น link เชื่อมต่อกันระหว่างโหนดแต่ละโหนดและในโหนดก็จะแสดง word sense และเชื่อมต่อกันกับซีแมนติคที่มีความสัมพันธ์ในระหว่าง sense นั้น ๆ อย่างเช่น โครงสร้างของไทพ์จะกำหนดอย่างง่าย ๆ ในซีแมนติคเน็ตเวิร์กโดยใช้ลิงค์ (link) ที่มีชื่อว่า "s" เพื่อที่จะแสดงความสัมพันธ์ของ subtype รูปที่ 2.20 จะกำหนดส่วนหนึ่งของโครงสร้างของไทพ์ที่ประกอบด้วยกลุ่มของ ฟิสิคัลออปเจ็ค (วัตถุที่เป็นรูปธรรม) และแอบสแตก(abstract)ออปเจ็ค (วัตถุที่เป็นนามธรรม)



รูปที่ 2.20 ส่วนของลำดับชั้นของไทพ์

ข้อมูลของเคสจะสามารถเก็บไว้ในเน็ตเวิร์กได้โดยใช้อาร์คซึ่งคุณจะพูดได้ว่าแอคชั่นทั้งหมดจะมี AGENT case ซึ่งทำให้อนิเมต (animate) ออปเจ็คสมบูรณ์โดยใช้เน็ตเวิร์กในรูป 2.21



รูปที่ 2.21 แอคชันทั้งหมดเป็นอนิเมตเอเจนท์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชนิดของโหนดตัวใหม่ซึ่งเรียกว่าเอ็กซิสเทนเชียลโหมด (existential node) ซึ่งจะแสดงอยู่ในกรอบสี่เหลี่ยมซึ่งเป็นตัวแสดงของค่าของเคสในกรณีนี้ AGENT case จะมีข้อบังคับให้เป็นออปเจ็คของ type ANIMATE มีการแสดงคล้าย ๆ อย่างนี้อีกหลายแบบที่จะอนุญาติให้มีการเก็บข้อมูลเกี่ยวกับค่าของเคสเช่นเดียวกัน

ซีแมนติคที่เป็นทางการจะถูกกำหนดให้เป็นส่วนหนึ่งของซีแมนติคเน็ตเวิร์กโดยการแมพโครงสร้างแต่ละตัวไปที่หลักการที่เหมือนกัน (equivalent axiom) ใน FOPC อย่างเช่นโหนดของไทพ์จะสามารถแมพไปที่อันนารี (unary) เพรดดิเคทและต่อไปที่ลิงค์ของสับไทพ์ระหว่างโหนดสองโหนดคือ ANIMATE และ PHYSOBJ ซึ่งจะเป็นการแทรกข้อมูลที่เท่ากันโดย ANIMATE object ทั้งหมดจะเป็นฟิสิคัลออปเจ็คดังนี้คือ

x.ANIMATE(x) PHYSOBJ(x)

เช่นเดียวกันโหนดที่เป็นค่าของเคสจะสามารถถูกแมพไปที่ตัวแปรที่แสดงปริมาณที่มีอยู่ และลิงค์ของเคสที่เป็นไบนารี (binary) เพรดดิเคทอย่างเช่น เน็ตเวิร์ก ในรูปที่ 2.21 อาจจะแมพไปที่ข้อมูลที่ว่าแอคชั่นทั้งหมดจะมี ANIMATE AGENT และ AGENT ทั้งหมดของแอคชั่นจะเป็น ANIMATE

x.ACTION(x) a.AGENT(x,a) & ANIMATE(a)

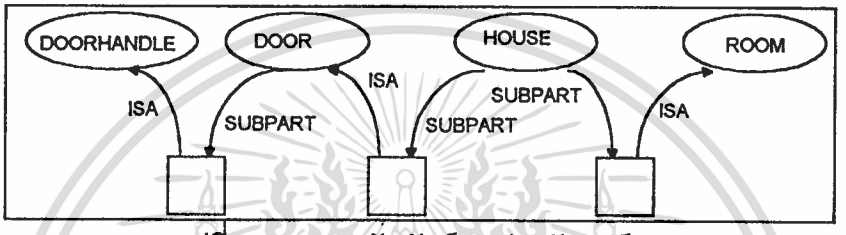
a,x.ACTION(x) & AGENT(x,a) ANIMATE(a)

ซีเมนติคที่ให้มานี้เราสามารถที่จะกำหนดการกระทำที่เรียกว่าอินเฮอริแทนซ์ (inheritance) ซึ่งสับไทพ์จะถูกถ่ายทอดคุณสมบัติทั้งหมดของซุปเปอร์ไทพ์ (supertype) ทั้งหมดของมันเข้ามาด้วยการแมพไปที่ FOPC รูปแบบอย่างง่ายของอินเฮอริแทนซ์จะแสดงเป็นรูปแบบพิเศษของลอจิคัลอิมพลิเคชัน (logical implication) รูปที่ 2.22 จะแสดงข้อสรุปของข้อมูลของเคสสำหรับกลุ่มของ verb sense ที่เป็นแอคชั่นของสับเคสทั้งหมด โดยการใช้กลไกอินเฮอริแทนซ์นี้เราจะพบคลาสของแอคชั่น TRANFER-ACTION ที่จะยอมรับ case AGENT, AT-TIME และ AT-LOC ซึ่งทั้งหมดจะถ่ายทอดมาจากคลาสแอคชั่นส่วน case THEME และ INSTRUMENT จะถ่ายทอดมาจาก คลาส OBJ/ACTION และ case TO-POSS ซึ่งถูกกำหนดเอาไว้สำหรับ TRANFER-ACTION โดยตรง สังเกตุว่าการแทนค่าอย่างนี้เป็นการพยายามที่จะกำหนดเคสที่ยอมรับได้ทั้งหมดซึ่งไม่ใช่เฉพาะ inner case เท่านั้น

ซีเมนติคเน็ตเวิร์กยังใช้ในการแสดงรูปแบบอื่น ๆ ของความรู้ทั่วไปเกี่ยวกับโครงสร้างต่าง ๆ ในโลกนอกจากจะใช้กับโครงสร้างของไทพ์ดังที่ได้กล่าวมาแล้ว ระดับชั้นที่สำคัญอีกอันหนึ่งก็คือ part-of ซึ่งจะแสดงออปเจ็คที่มีความสัมพันธ์กับสับพาร์ท (subpart) ของมันเพื่อที่จะใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประโยชน์ของความเหมือนกันระหว่างเรื่องของสับพาร์ท และเคส เราสามารถที่จะแสดงสับพาร์ท โดยใช้ existential square node ซึ่งเราได้ใช้มาแล้วในกรณีของเคสตัวอย่างนี้จะแสดงข้อมูลของไทพ์สำหรับ existential node โดยใช้ลิงค์ใหม่ที่มีชื่อว่า ISA ซึ่งจะมีความหมายเท่ากับสัญลักษณ์ของไทพ์ที่อยู่ในบล็อก (block) และสองสัญลักษณ์นี้ จะสามารถใช้แทนที่กันได้ดังนั้นคุณสามารถที่จะแสดงว่า a house has room and door เป็นสับพาร์ทและ door นั้นจะมี handle ดังแสดงในรูปที่ 2.22



รูปที่ 2.22 ความสัมพันธ์ระหว่างสับพาร์ท

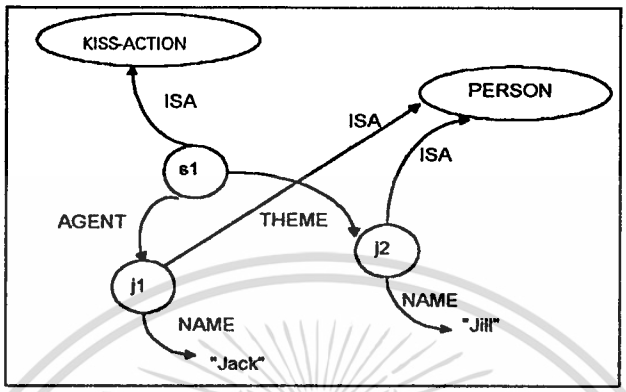
ระบบที่ใช้แสดงที่สมบูรณ์จะต้องสามารถบ่งบอกได้ว่าสับพาร์ทนี้เป็นยูนิค (unique) ออปเจ็ค (the handle on a door) หรือเป็นเซ็ทของออปเจ็ค (the rooms in a house) นอกจากนี้มันจะต้องมีความสามารถที่จะแสดงการกระจายและความสัมพันธ์อื่น ๆ ที่สับพาร์ทแต่ละตัวมีต่อกัน ซึ่งก็คือถ้าร่างกายของมนุษย์ประกอบด้วย หัว ตัว แขนและขา ดังนั้น คุณจะต้องแสดงว่าหัวเป็นยูนิคและต่อกับตัวและจะต้องมีแขนสองแขนต่อกับตัวเช่นนี้ไปเรื่อย ๆ ซึ่งทั้งหมดนี้จะต้องเป็นส่วนของสับพาร์ทที่กำหนดอยู่ใน BODY

การแสดงความหมายของประโยค : โทเคน

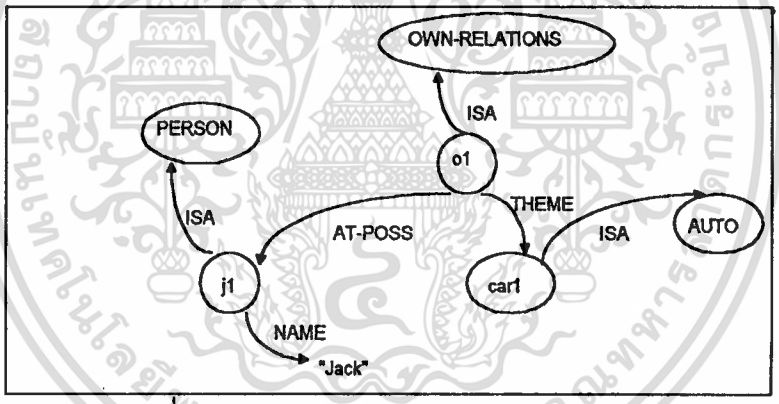
ต่อไปนี้เราจะพัฒนาซีเมนติคเน็ตเวิร์กเพื่อที่จะใช้แสดงความหมายของประโยค การอนุญาติให้เน็ตเวิร์ก ของโครงสร้างเหมือนกับที่ได้กล่าวไปแล้ว แต่ระบุการถ่ายทอดที่เฉพาะเจาะจงมากกว่าที่จะเป็นไทพ์เราจะต้องสร้างคลาสของโหนดขึ้นมาใหม่เรียกว่าโทเคน (token) ซึ่งจะใช้ในการแสดงออปเจ็คและแอคชั่นที่เฉพาะเจาะจง อย่างเช่น การ แสดงเนื้อหาของประโยค *Jack kissed Jill* ซึ่งได้แสดงในรูปที่ 2.23

การถอดความเน็ตเวิร์กนี้จะเป็นดังนี้คือ จะมี KISS-ACTION, s1 ที่มี AGENT case ที่ทำให้ j1 ซึ่งเป็นบุคคลสมบูรณ์และมี THEME case ที่ทำให้ j2 ซึ่งเป็นบุคคลสมบูรณ์ นอกจากนี้ j1 จะถูกกำหนดชื่อไว้เป็น Jack และ j2 จะถูกกำหนดชื่อไว้เป็น Jill เน็ตเวิร์กสามารถที่จะใช้แสดงประโยคที่ได้อธิบายขั้นตอนไปแล้วได้ รูปที่ 2.24 แสดงความจริงที่ว่า Jack owns a particular car. (ซึ่งแสดงในโหนด car1)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.23 ซีเมนติคเน็ตเวิร์กสำหรับประโยค *Jack kissed Jill.*



รูปที่ 2.24 การแสดงของ *Jack owns car1*

ซีเมนติคเน็ตเวิร์กที่เป็นทางการหลายตัวจะเสนอเพียงความสามารถที่ถูกจำกัดสำหรับแสดงสูตรที่มีการกำหนดจำนวนและสูตรที่เกี่ยวกับ ข้อขัดแย้งและข้อปฏิเสธ ข้อจำกัดเหล่านี้จะไม่เป็นคุณสมบัติที่ถูกถ่ายทอด

2.5.7 ลอจิคัลฟอร์ม (The Logical Form)

ซีเมนติคเน็ตเวิร์คที่ได้กล่าวมาในหัวข้อที่แล้วจะสามารถใช้เป็นจุดเริ่มต้นสำหรับการเปลี่ยนแปลงไปเป็น ลอจิคัลฟอร์ม เน็ตเวิร์ก ในรูปที่2.23 สามารถที่จะแสดงในรูปของลอจิคัลฟอร์ม (LF)ได้ดังต่อไปนี้

(PAST s1 KISS-ACTION (AGENT s1 (NAME j1 PERSON "Jack"))

(THEME s1 (NAME j2 PERSON "Jill"))

เอกสารนี้มีลิขสิทธิ์ ห้ามนำไปใช้ประโยชน์ด้านการค้า ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งใน LF แต่ละส่วนจะถูกสร้างขึ้นมาจากส่วนประกอบต่าง ๆ ดังนี้

- operator จะเป็นการบอกชนิดของโครงสร้างในที่นี้ PAST หมายถึง การแสดงค่านี้จะ เป็นการแทนความรู้ว่ามีบางอย่างเกิดขึ้นในอดีต
- name ใช้สำหรับออปเจ็คที่กำลังอธิบาย ในกรณีนอินสแตนท์ของ KISS-ACTION จะถูกเรียกใช้โดย s1
- ไทพ์ของออปเจ็คก็คือ KISS-ACTION
- modifier ของออปเจ็คอาจจะเป็นรายการของโครงสร้างของ LF; ในกรณีนี้ modifier จะประกอบไปด้วย เคสสองเคส คือ AGENT และ THEME

โอเปอร์เรเตอร์ (operator) อื่น ๆ สำหรับโครงสร้างประโยคจะประกอบด้วย PRES (simple present tense) ,FUT (simple future) และ INF (สำหรับ infinitive clause)

ลอจิคัลฟอร์มของ NP

ภาษานี้จะถูกขยายเพื่อที่แสดง NP ดังต่อไปนี้ โอเปอร์เรเตอร์สำหรับ NP อย่างง่ายจะถูกใช้เพื่อที่จะบอกถึงข้อมูลที่ใช้ในการตัดสินเพื่อมันจะถูกนำไปใช้ในการประมวลผลขั้นต่อไป การกำหนด NP สามารถจะนำมาใช้ประกอบกันได้ดังนี้

DEF/SING	- definite singular reference (*the boy*)
DEF/PL	- definite plural reference (*the boys*)
INDEF/SING	- indefinite singular reference (*a boy*)
INDEF/PL	- indefinite plural reference (*boys*)

name และไทพ์ของ simple NP จะเป็นไปตามที่คาดไว้ และส่วนขยายจะประกอบไปด้วย quantification ที่สร้างโดยคำคุณศัพท์หรือวลีขยายอื่น ๆ ดังนั้น LF สำหรับ NP *the last boy* จะเป็น

(DEF/SING b1 BOY (LARGE b1))

quantify NP จะแมพไปที่โครงสร้างที่คล้ายกันโดยใช้เครื่องหมายที่ต่างกันตามความเหมาะสม ดังนั้นประโยค *Each boy ate a large pizza.* จะแสดงได้ดังต่อไปนี้

(PAST a1 EAT-EVENT
  (AGENT a1 (EACH b2 BOY))
  (THEME P1 (INDEF/SING p1 PIZZA (LARGE p1))))

สำหรับ ประโยคคำถามแบบ WH ประโยคในเทอมของ WH (WH-term) จะต้องมีเครื่องหมายพิเศษอยู่ใน LF ซึ่งเราสามารถใช้ได้ 3 รูปแบบคือ WH, WH/PL และ WH/SING ขึ้นอยู่กับได้ระบุจำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ในเชิงพาณิชย์ การค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น ไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารนี้ทุกครั้งที่มีการนำไปใช้

นวนไว้เป็นพหูพจน์หรือ เอกพจน์ตามลำดับ อย่างเช่น NP *who* จะมี LF

(WH p1 PERSON)

ในขณะที่ NP *which dog* จะมี LF

(WH/PL d1 DOG)

รูปแบบพิเศษสองตัวซึ่งใช้สำหรับ NP ประกอบด้วยชื่อเฉพาะ (proper name) และสรรพนาม เราได้เห็นรูปแบบไปแล้ว คือโอเปอร์เรเตอร์ NAME เพื่อที่จะบอกว่า NP นั้นอยู่ในรูปแบบของชื่อเฉพาะและชื่อนั้นจะถูกแสดงในตำแหน่งที่อยู่หลังไทพ์เราสามารถที่จะเพิ่มส่วนขยายเข้าไปได้ถ้าต้องการ อย่างเช่น NP *big bad John* จะถูกแมพเป็น

(NAME j3 PERSON "John" (BIG j3)(BAD j3))

สำหรับคำสรรพนามก็จะมีรูปแบบที่คล้ายกันโดยจะกำหนดเครื่องหมายไว้เป็น PRO ดังนั้น NP *he* จะแสดงได้ดังนี้

(PRO h1 MALE "he")

ลอจิคัลฟอร์มสำหรับ NP ที่อธิบายเหตุการณ์ซึ่งที่เหมือนกันในการแสดงของเหตุการณ์ในประโยค ยกเว้น เครื่องหมายจะเป็นการบ่งบอกการชี้เฉพาะมากกว่าที่จะเป็นเครื่องหมายของเท็นส์ อย่างเช่น NP *the arrival of George at the station* จะแสดงได้ดังนี้

(DEF/SING a1 ARRIVE-EVENT

(AGENT a1 (NAME g1 PERSON "George"))

(TO-LOC a1 (DEF s4 STATION)))

โดยที่ประโยค George arrivaled at the station จะแสดงอยู่ในรูปของลอจิคัลฟอร์มได้คือ

(PAST a2 ARRIVE-EVENT

(AGENT a2 (NAME g1 PERSON "George"))

(TO-LOC a2 (DEF s6 STATION)))

สำหรับข้อสงวนของเหตุการณ์ที่เป็น inner case คุณสามารถที่จะละเว้นรูปแบบของความสัมพันธ์และใช้คำย่อโดยอยู่ในเครื่องหมาย "[ ]" ซึ่งถ้าใช้อย่างนี้ลอจิคัลฟอร์มของประโยค George arrivaled at the station จะเขียนใหม่ได้เป็น

(PAST a2 ARRIVE-EVENT [AGENT (NAME g1 PERSON "George")]

[TO-LOC (DEF s6 STATION)])

ถ้ามีความกำกวมเกิดขึ้นในลอจิคัลฟอร์ม มันจะสามารถแสดงโดยเขียนรายการที่เป็นไปได้ในเครื่องหมาย "{ }" อย่างเช่น NP *the ball* จะมีความกำกวมระหว่างการเป็นลูกกลม ๆ และการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เต้นรำในสังคม ซึ่งเราสามารถจะแสดงโดยใช้รูปแบบที่ระบุอย่างใดอย่างหนึ่งได้ดังต่อไปนี้

(DEF/SING b1 {SPHERE,DANCE})

ลอจิคัลฟอร์มของประโยค (The Logical Form of Sentences)

ยังมีรายละเอียดอีกเล็กน้อยที่เกี่ยวข้องกับประโยค โดยเฉพาะอย่างยิ่งยังไม่มีอะไรใน LF ที่กล่าวถึงเกี่ยวกับประโยค declarative, imperative หรือ interrogative และยังไม่มีอะไรที่พูดเกี่ยวกับประโยคเชิงซ้อนและประโยคเชิงประกอบ

ชนิดของประโยคใหญ่ 4 ตัวจะมีรูปของลอจิคัลฟอร์มที่ใช้การตีความประโยคเป็นอาร์กิวเม็นต์ อย่างเช่น สำหรับประโยคบอกเล่าอย่างประโยค *the man ate a peach.* จะมี LF คือ

(ASSERT (PAST e1 EAT [AGENT (DEF/SING m1 MAN)]
[THEME (INDEF/SING p1 PEACH)]))

สำหรับ Y/N-Q อย่างเช่น Did the man eat a peach? จะมี LF คือ

(Y/N-QUERY (PAST e1 EAT [AGENT (DEF/SING m1 MAN)]
[THEME (INDEF/SING p1 PEACH)]))

สำหรับ WH-Q อย่างเช่น what did the man eat? มี LF คือ

(WH-QUERY w1 (PAST e1 EAT [AGENT (DEF/SING m1 MAN)]
[THEME (WH w1 PHYSOBJ)]))

สุดท้ายประโยคคำสั่งอย่างเช่น eat the peach มี LF คือ

(COMMAND (INF e1 EAT [THEME (DEF/SING p1 PEACH)]))

ประโยคเชิงซ้อนอย่างอนุประโยคสามารถที่จะจัดการได้ในวิธีการเดียวกันกับประโยคอื่น ๆ พอยนเตอร์ ใด ๆ ในโครงสร้างเชิงซ้อน จะถูกวิเคราะห์แบบธรรมดาโดยการ แทรกชื่อของโครงสร้างที่สร้างไว้สำหรับส่วนประกอบที่ถูกเอ่ยถึง เช่นประโยค I want to leave จะมีรูปชินเทคติคคือ

(S SUBJ (NP1 PRO I)
MAIN-V want
TENSE pres
COMP (S SUBJ -> NP1
MAIN-V leave))

และมีรูป LF คือ

(ASSERT (PRES w1 WANT [EXPERIENCER (PRO i1 PERSON "I")]
[THEME (INF I1 LEAVE [AGENT i1])]))

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประเด็นสุดท้ายก็คือประโยคเชิงประกอบจะมีเนื้อหาที่ค่อนข้างซับซ้อนและจะเอ่ยถึงต่อไป โอเปอร์เรเตอร์ของประโยค อย่างเช่น คำสันธาน (and หรือ but) จะแนะนำเมื่อต้องการใช้ อย่างเช่น ประโยค I wanted to leave but i lose the keys จะมีรูปแบบของ LF ดังนี้

(ASSERT (BUT (PAST w1 WANT [EXPERIENCER

(PRO i1 PERSON "I")]

[THEME (INF I1 LEAVE

[AGENT i1])])

(PAST I1 LOSE [AGENT (PRO i2 PERSON "I")]

[THEME (DEF/PL k1 KEY)])))

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6 การตีความหมาย (Semantic Interpretation)

ในบทนี้เราจะมากำหนดวิธีการใช้ข้อมูลทางซีแมนติคที่ได้อธิบายไปในบทที่แล้วเพื่อที่จะสร้างลอจิคัลฟอร์ม ตัวตีความหมาย (interpreter) ตัวนี้จะทำงานแยกต่างหากจากตัววิเคราะห์ทางซินแทกซ์ (syntax) โดยเด็ดขาด ซึ่งก็หมายความว่าอินพุทต์ของมันคือผลลัพธ์ที่ได้มาจากการวิเคราะห์ทางไวยากรณ์ที่ได้ทำไปในเทอมที่แล้วนั่นเอง

2.6.1 วิธีการพื้นฐานสำหรับการตีความทางความหมาย

การตีความทางซีแมนติคจะทำสำเร็จได้โดยใช้กฎในการตีความที่ยอมรับรูปแบบของ selection restriction ที่ได้กล่าวไว้ในหัวข้อ 2.5 และใช้ลอจิคัลฟอร์มบางส่วน ในการสร้างลอจิคัลฟอร์มบางส่วน เราจะเรียกว่า partial decription (PD) ซึ่ง PD นี้จะสร้างขึ้นมาโดยใช้ฟังก์ชั่นดังต่อไปนี้

* - ใช้เป็นเครื่องหมายสำหรับชื่อที่ชื่อเดียวที่รับเข้ามาใหม่เพื่อที่จะแสดงออปเจ็คที่อธิบายอยู่ใน ลอจิคัลฟอร์ม

I(slotname) - ใช้ส่งค่าของชนิดของซีแมนติคที่เป็นส่วนประกอบในสล็อตที่ได้กำหนดไว้กลับมาให้

V(slotname) - ส่งค่าการตีความซีแมนติคทั้งหมดของส่วนประกอบในสล็อตที่ได้กำหนดกลับมาให้

ยกตัวอย่างเช่น ในการตีความกฎสำหรับคำคุณศัพท์ green แสดงได้ดังรูปที่ 2.25

```
(green.1)   (NP ADJS (green)
                HEAD +physobj) → (? * I(HEAD)(COLOR * GR137))
(green.2)   (NP ADJS (green)
                HEAD +fruit)   → (? * I(HEAD)(UNRIPE * ))
(green.3)   (NP ADJS (green)
                HEAD +person)  → (? * I(HEAD)(NOVICE * ))
```

รูปที่ 2.25 กฎการตีความหมายสำหรับคำว่า green

กฎ green.1 จะจับคู่กับ NP ตัวใดก็ได้กับสลอตของ ADJS ที่เก็บคำ green และในสลอตของ HEAD ที่มีชนิดของซีแมนติคเป็น PHYSOBJ กฎนี้จะสร้าง LF บางส่วนที่ประกอบด้วยเครื่องหมายที่ไม่ได้ระบุไว้, ชนิดของซีแมนติค ขององค์ประกอบในสลอตของ HEAD และส่วนขยาย (COLOR * GR137) (GR137 เป็นตัวที่สมมุติให้เป็นค่าคงที่ที่ใช้ในการแสดงสีเขียว) ในสภาวะใด

ที่กฎนี้ถูกนำมาใช้จะมีการสร้างชื่อที่ไม่ซ้ำ (unique name) สำหรับองค์ประกอบที่ใช้ในการตีความ และใช้เพื่อแทนที่ " * " ที่เกิดขึ้นทั้งหมด อย่างเช่นให้ NP เป็น

(NP DET a

ADJ (green)

HEAD apple)

กฎ green.1 จะสร้าง PD

(? c1 APPLE (COLOR c1 GR137)

ซึ่งในกฎ green.2 จะสร้าง PD

(? c1 APPLE (UNRIPE c1))

กฎ green.3 จะทำไม่สำเร็จเพราะไม่สามารถจับคู่กับ NP นี้ได้ เนื่องจากในเล็กซิคอนของ apple ไม่ได้บอกว่ามันเป็น PERSON สำหรับองค์ประกอบย่อยที่ซับซ้อน ฟังก์ชั่น V ใช้สำหรับจัดการเรื่องการวิเคราะห์ความหมายขององค์ประกอบนั้น ค่านี้จะถูกนำมาเก็บในสล็อตใหม่ในองค์ประกอบซึ่งเรียกว่า SSEM เช่นหลังจากที่ทำขบวนการทาง semantic สำหรับ NP the man สำเร็จแล้ว โครงสร้างของประโยค The man laughed จะเป็นดังนี้

(S SUBJ (NP DET the

HEAD man

NUM {3s}

SEM (DEF/SING M1 MAN))

MAIN-V laugh

TENSE PAST)

ต่อไปเรามาพิจารณาการตีความที่ใช้กฎสำหรับ laugh

(laugh.1) (S SUBJ +animate

MAIN-V laugh) → (? * LAUGH [AGENT I(SUBJ)])

กฎนี้จะใช้ได้เนื่องจากชนิดของ register SUBJ (ซึ่งก็คือ MAN) เป็นชนิดย่อยของ ANIMATE ดังนั้นจึงมีการสร้าง PD ต่อไปนี้

(? I1 LAUGH [AGENT (DEF/SING m1 MAN)])

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Merging

การที่จะสร้าง LF ที่สมบูรณ์แบบเราจะต้องรวม PD แต่ละตัวที่สร้างขึ้นมา ซึ่งการทำเช่นนี้เรียกว่า Merging ซึ่งสามารถจะมองได้ว่าเป็นการรวมตัวกันที่มีหลากหลายรูปแบบ ในกรณีที่ง่ายที่สุดสำหรับการจะรวม PD สองตัว ถ้า สับพาร์ท (subpart) แต่ละตัวของมันเข้ากันได้หรือมีตัวใดตัวหนึ่งที่ไม่มีค่าของสับพาร์ท อย่างเช่น กฎของอาร์ติเคิล a คือ

(a.1) (np DET a) → (INDEF/SING * ? ?)

เมื่อมาประยุกต์ใช้กับโครงสร้างทางไวยกรณ์ของ NP a green apple ก็จะสร้าง PD

(INDEF/SING c1 ? ?)

PD นี้สามารถจะรวมกับ PD อื่น ๆ ที่ได้เปลี่ยนแปลงมาจากกฎสำหรับ green ที่ได้กล่าวถึงก่อนหน้านี้ เพื่อที่จะสร้างการตีความที่สมบูรณ์แบบสำหรับ NP ทั้งตัวซึ่งมันมีหนทางที่เป็นไปได้สองหนทางมีชื่อเรียกดังต่อไปนี้

(INDEF/SING c1 APPLE (COLOR c1 GR137))

(INDEF/SING c1 APPLE (UNRIPE c1))

ความซับซ้อนในการรวมขึ้นอยู่กับส่วนขยายในส่วนธรรมดาที่สุดของมันจะเป็นการตัดสินใจว่า PD สองตัวจะสามารถรวมกันได้หรือไม่ ซึ่งเป็นเรื่องเกี่ยวกับการตรวจสอบว่าส่วนขยายของทั้งสองตัวซึ่งนำมาพิจารณาประกอบกันนี้เข้ากันได้ตามหลักตรรกศาสตร์หรือไม่ PD ทั้งสองตัวจะนำมารวมกันถ้าส่วนขยายของ inner case เข้ากันได้ ดังที่ได้กล่าวไว้แล้วว่าใน inner case ใด ๆ จะมีค่าได้มากที่สุดเพียงค่าเดียว ดังนั้นในกรณีนี้ เมื่อ PD สองตัวมีส่วนขยายที่ระบุค่าสำหรับแต่ละ inner case ของมัน ค่าทั้งสองนี้จะต้องนำมารวมกันให้ได้อย่างสมบูรณ์ อย่างเช่น พิจารณา PD สองตัวต่อไปนี้

(? r1 RUN1 [THEME (NAME j1 PERSON"John")])

(? r1 RUN1 [THEME (DEF/SING p1 PRINT-PRESS)])

ไม่สามารถนำมารวมกันได้เนื่องจาก theme case ทั้งสองรวมกันไม่ได้หรืออีกกรณีหนึ่ง PD ทั้งสองตัวต่อไปนี้จะรวมกันได้

(? r1 RUN1 [AGENT (NAME j1 PERSON"John")])

(? r1 RUN1 [THEME (DEF/SING p1 PRINT-PRESS)])

และให้ผลลัพธ์เป็น PD ต่อไปนี้

(? r1 RUN1 [AGENT (NAME j1 PERSON"John")]
[THEME (DEF/SING p1 PRINT-PRESS)])

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งอาจจะเป็น PD ที่สร้างขึ้นมาในขณะที่กำลังวิเคราะห์ประโยค John ran the printing press

ถ้า PD มีค่ากำกวม (ก็คือมีรายการอยู่ใน [ ]) ค่านั้นจะถูกนำมารวมกันโดยใช้การอินเตอร์เซ็ค (intersect) ค่าระหว่าง PD ทั้งสอง ตัวอย่างของค่ากำกวมแบบธรรมดาที่สร้างขึ้นจากการตีความกฎสำหรับ ball ที่มีความหมายทางนัยของ physobj (คือ ชนิดแบบ SPHERE) และในความหมายทางนัยของเหตุการณ์ทางสังคม (ซึ่งก็คือชนิดของ DANCE) การตีความกฎสำหรับ ball อาจจะเป็นดังนี้

(NP HEAD ball) → (? * {SPHERE,DANCE} ?)

ซึ่งอาจจะสร้าง PD ต่อไปนี้ โดยใช้การวิเคราะห์ทางไวยกรณ์ของ NP a colorful ball

(? b2 {SPHERE,DANCE} ?)

เมื่อรายการทางเล็กซิคัลมีได้หลายความหมายจะมีเพียงความหมายเดียวที่ยอมรับได้ในข้อกำหนดของชนิดในกฎของการตีความในกรณีนี้ ถ้าชนิดของสิ่งนั้นถูกนำมาใช้เพื่อสร้าง PD เฉพาะความหมายที่นำมาใช้นั้นที่จะเหมาะสม อย่างเช่น ถึงแม้ว่า ball จะมีความหมายได้สองนัยดังที่ได้กล่าวไปแล้ว ใน NP the green ball จะวิเคราะห์ตามกฎ green.1 ซึ่งจะเข้ากันได้กับความหมายของทรงกลม (sphere) เนื่องจากมันต้องการ physobj ดังนั้นกฎ green.1 จะสร้าง PD ต่อไปนี้ขึ้นมา

(? b3 SPHERE (COLOR b3 GR137))

สำหรับกรณีที่ซับซ้อนมากขึ้นจะเกิดขึ้นเมื่อการตีความองค์ประกอบทั้งหมดมีความกำกวม อย่างเช่น ใน NP the white dishwasher อาจจะหมายถึงคนที่ถูกจ้างให้เป็นผู้ล้างจานซึ่งเป็นคนผิวขาว หรืออาจจะหมายถึงเครื่องใช้สำหรับล้างจานที่เป็นสีขาว ดังนั้นในการตีความทางความหมายของ NP นี้จะมีผลลัพธ์อยู่สองแบบซึ่งทั้งสองแบบนี้จะถูกเก็บอยู่ในสลอต SEM แต่ถ้าให้ประโยค The white dishwasher laughed การวิเคราะห์ประโยคจะต้องเป็นดังนี้

```
(s SUBJ (NP DET the
           ADJS (white)
           HEAD dishwasher
           SEM ((DEF/SING d1 PERSON-TYPE-43
                           (RACE d1 WHITE)),
                (DEF/SING d1 APPLIANCE-TYPE-3
                           (COLOR d1 WHITE))))
    MAIN-V laugh
    TENSE PAST)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งการวิเคราะห์ข้างต้นยังไม่ได้ใช้กฎ laugh.1 และเมื่อนำกฎมาใช้จะมีเพียงความหมายแรกของ NP เท่านั้นที่เหมาะสมกับข้อจำกัดของ type + animate ทำให้ผลลัพธ์ของ PD เป็นดังนี้

(? I3 LAUGH [AGENT (DEF/SING d1 PERSON-TYPE-43
(RACE d1 WHITE))])

แต่สำหรับคำกิริยาอื่น ๆ ความกำกวมอาจจะยังมีอยู่ในโครงสร้างของ S อย่างเช่น กฎสำหรับ wash อาจจะสร้าง PD ต่อไปนี้ เมื่อนำมาใช้กับโครงสร้างทางไวยกรณ์ของประโยค The pot was washed by the white dishwasher.

(? w1 WASH [AGENT {(DEF/SING d1 PERSON-TYPE-43
(RACE d1 WHITE)),
(DEF/SING d1 APPLIANCE-TYPE-3
(COLOR d1 WHITE))}]
[THEME (DEF/SING p1 POT)])

ความกำกวมนี้จะถูกแก้ไขโดยการรวมกับ PD อื่น ๆ ต่อไปหรืออาจจะยังอยู่ในการตีความขั้นสุดท้ายและทิ้งให้เป็นหน้าที่ของส่วน contextual analyzer มาทำหน้าที่วิเคราะห์ต่อไป

```
(past.1)        (S TENSE PAST)      → (PAST ? ? ?)
(laugh.1)       (S SUBJ +animate
                   MAIN-V laugh)    → (? * LAUGH [AGENT I(YSUBJ)])
(dishwasher.1)  (NP HEAD dishwasher) → (? * (PERSON-TYPE-43
                                              APPLIANCE-TYPE-3) ?)
(the.1)         (NP DET the
                   NUM (3s))        → (DEF/SING * ? ?)
(white.1)       (NP ADJS (white)
                   HEAD +physobj-human) → (? * I(HEAD)
                                              (COLOR * WHITE))
(white.2)       (NP ADJS (white)
                   HEAD +human)     → (? * I(HEAD)(RACE * WHITE))
```

รูปที่ 2.26 กฎการตีความทางความหมายบางส่วน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6.2 อัลกอริทึมการตีความหมาย (The Interpretation Algorithm)

อัลกอริทึมในการตีความหมายทำได้โดยพิจารณาโครงสร้างทางไวยกรณ์ โดยในขั้นแรกจะพยายามตีความส่วนนอกสุดของ S และต่อมาก็ตีความองค์ประกอบย่อยแต่ละตัวของมันเข้ามาเรื่อย ๆ ตามที่ต้องการเพื่อที่จะให้การทดสอบความหมายที่ทำในรูปแบบที่ใช้นั้นสมบูรณ์ เพื่อที่จะเห็นภาพได้ชัด เรามาพิจารณาการตีความในประโยค The diskwasher laughed โดยใช้กฎการตีความในรูปที่ 2.26 ซึ่งโครงสร้างทางไวยกรณ์เริ่มต้นจะได้ดังนี้

(S SUBJ (NP DET the
NUM {3s}
HEAD dishwasher)
MAIN-V laugh
TENSE PAST)

สำหรับกฎ laugh.1 ที่จะสำเร็จได้สล็อตของ SUBJ จะต้องเป็นซนิด ANIMATE เพื่อที่จะตรวจสอบเช่นนี้ได้ ในขั้นแรก NP จะต้องถูกวิเคราะห์ลึกเข้าไป ในกฎ the.1 และ dishwasher.1 จะเข้ากันได้อย่างเหมาะสมและสร้าง PD สองตัวดังนี้

(DEF/SING d1 ? ?)

(? d1 {PERSON-TYPE-43,APPLIANCE-TYPE-3} ?)

เมื่อทำการรวม PD ทั้งสองเพื่อจะสร้าง LF สำหรับ NP จะต้องนำมันเข้าไปใส่ในสล็อตของ SEM ของมันเอง ดังนั้นโครงสร้างจะเป็นดังนี้

(S SUBJ (NP DET the
NUM {3s}
HEAD dishwasher)
SEM (DEF/SING d1 {PERSON-TYPE-43
APPLIANCE-TYPE-3}))
MAIN-V laugh
TENSE PAST)

กฎ laugh.1 จะทำการทดสอบประธานของประโยคในด้านความหมายของมัน และเมื่อสำเร็จจะสร้าง PD ต่อไปนี้ซึ่งมีชนิดของซีแมนติคของ NP ที่เป็นประธานดังนี้

(? I1 LAUGH [AGENT (DEF/SING d1 PERSON-TYPE-3)])

เนื่องจากกฎ past.1 ถูกใช้ในการสร้าง PD (PAST ? ? ?) ในการตีความขั้นสุดท้ายจะสร้างโดย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การรวม PD สองตัวเข้าด้วยกันเป็น

(PAST I1 LAUGH [AGENT (DEF/SING d1 PERSON-TYPE-3)])

Interpreting a Single Constituent

เมื่อเรามีโครงสร้างที่ได้เอ่ยไปแล้ว ซึ่งองค์ประกอบย่อยของมันถูกตีความตามที่ต้องการ คุณสามารถที่จะตรวจสอบรายละเอียดของอัลกอริทึมสำหรับองค์ประกอบเดี่ยว ๆ ซึ่งประกอบด้วยขั้นตอนสองขั้นตอนใหญ่ ๆ คือการวิเคราะห์คำหลัก (head analysis) การวิเคราะห์ส่วนขยาย (Modifier analysis) การวิเคราะห์คำหลักจะเป็นการตีความคำหลักของวลี (ซึ่งก็คือ MAIN-V ใน S, HEAD ใน NP หรือ PREP ใน PP) เข้าด้วยกัน ด้วยองค์ประกอบย่อยของสับแคทของมัน การวิเคราะห์ส่วนขยายจะตีความต่อจากนั้น โดยจะตีความส่วนที่เหลือขององค์ประกอบนั้น

คุณจะทำการวิเคราะห์คำหลักโดยการดึงกฎของการตีความทั้งหมดที่อยู่ในเล็กซิคัลในตำแหน่งของคำหลักและใช้มันในการสร้างเซ็ตของ PD ต่อจากนั้นคุณก็จะรวม PD เหล่านี้เข้าด้วยกันเพื่อที่จะสร้างการตีความสำหรับ head และองค์ประกอบของมัน ในกรณีที่ต้องมีการตีความหลายครั้งการตีความแต่ละอันจะรายงานผลของเซ็ตขององค์ประกอบย่อยแตกต่างกันเล็กน้อย

ขั้นต่อไปคุณจะทำขั้นตอนที่สองคือการวิเคราะห์ส่วนขยายสำหรับ การตีความ head แต่ละตัวที่ได้ทำมาในขั้นแรก ในขั้นตอนนี้คุณจะต้องหาองค์ประกอบย่อยที่ยังไม่ถูกรายงานจากขั้นการวิเคราะห์คำหลัก และคุณจะต้องวิเคราะห์ลึกลงไปแต่ละตัวเพื่อที่จะสร้าง PD ขึ้นมารวมกับขั้นที่ได้จากการวิเคราะห์คำหลักซึ่งจะกล่าวโดยละเอียดต่อไป

การวิเคราะห์คำหลัก

ในขั้นตอนของการวิเคราะห์คำหลักจะต้องสร้างเซ็ตของ PD ที่ประกอบขึ้นมาจากการตีความที่แตกต่างกันของคำหลักและวลีย่อยของมัน และรายการขององค์ประกอบที่ได้ถูกรายงานไว้สำหรับมัน (เพื่อที่จะผ่านไปให้ในขั้นของการวิเคราะห์ส่วนขยาย) รายการขององค์ประกอบสามารถที่จะแสดงออกมาว่าองค์ประกอบแต่ละตัวถูกใช้ในรูปแบบใดของกฎที่ใช้ในการตีความคำหลัก คุณสามารถที่จะแยกแยะองค์ประกอบได้โดยใช้ชื่อของสล็อตของมัน เว้นแต่ว่าในกรณีที่มีค่าที่เป็นรายการของสมาชิก ในกรณีนี้รายการเฉพาะแต่ละตัวที่อยู่ในลิสนั้นจะต้องถูกรายงานขึ้นมาโดยอ้างถึงแบบตรง ๆ พิจารณาตัวอย่างง่าย ๆ ผลลัพธ์ของการวิเคราะห์คำหลัก สำหรับ NP *The white diskwasher* จะเป็น PD ตัวเดียว (ซึ่งดัดแปลงมาจากกฎ diskwasher.1),(? d1 (PERSON-

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

TYPE-43,APPLIANCE-TYPE-3) ?) และลิสต์ขององค์ประกอบที่ถูกแสดงมาสำหรับมันจะเป็นคำหลักแบบง่าย ๆ (นั่นก็คือสลอตของ slot) ซึ่งหมายความว่าองค์ประกอบนั้นจะต้องอยู่ในสล็อต DET, NUM และ ADJS จะต้องถูกแสดงเพื่อใช้ในขั้นการวิเคราะห์ส่วนขยายสำหรับตัวอย่างอื่นอีกอันหนึ่ง โดยการวิเคราะห์ในประโยค The diskwasher laughed การวิเคราะห์คำหลักจะตีความองค์ประกอบเกือบจะทั้งหมดเนื่องจากกฎ laugh.1 จะแสดงองค์ประกอบทั้งของ SUBJ และ MAIN-V slot เฉพาะสลอต TENSE เท่านั้นที่จะต้องถูกแสดงในขั้นการวิเคราะห์ส่วนขยาย

การวิเคราะห์ส่วนขยาย

คุณจะสร้าง PD สำหรับการวิเคราะห์ส่วนขยาย โดยการหากฎของการตีความในองค์ประกอบย่อยแต่ละตัวที่ต้องการที่จะให้แอคเคานท์ (account) ถ้าองค์ประกอบย่อยเป็นไอเท็ม (item) ของเล็กซิคัล หรือเป็นค่าธรรมดา (PAST) คุณสามารถที่จะเข้าถึงกฎสำหรับไอเท็มนั้นได้โดยตรง ถ้ามันเป็นองค์ประกอบอย่างเช่น PP คุณจะต้องใช้กฎสำหรับ head ขององค์ประกอบนั้น ซึ่งในกรณีนี้ก็คือ PREP

เนื่องจากคุณได้สร้าง PD คุณจะต้องสร้างเซ็ตของการตีความชุดสุดท้ายโดยรวมการตีความของ head แต่ละตัว โดยใช้การตีความทางใดทางหนึ่งที่เป็นไปได้สำหรับส่วนขยายแต่ละตัว

ตัวอย่างเช่น ถ้ามี NP *The white dishwasher* คุณจะต้องแอคเคานท์ให้กับสล็อต DET, NUM และ ADJS หลังจากนั้นคุณจึงสร้าง PD ต่อไปนี้ ในขั้นตอนของการวิเคราะห์ส่วนขยาย

for DET and NUM: (DEF/SING d1 ? ?)

for ADJS: (? d1 APPLIANCE-TYPE-43 (COLOR d1 WHITE))

(? d1 PERSON-TYPE-3 (RACE d1 WHITE))

คุณสามารถที่จะสร้างการตีความขั้นสุดท้ายได้สองแบบ แต่ละแบบจะใช้การตีความ head ของ *dishwasher* ที่สร้างมาก่อนหน้านี้ PD ที่สร้างมาจาก DET slot และ PD ตัวหนึ่งที่สร้างมาจากสลอต ADJS ก็คือ

(DEF/SING **d1** APPLIANCE TYPE-43 (COLOR **d1** WHITE))

(DEF/SING **d1** PERSON-TYPE-3 (RACE **d1** WHITE))

โดยทั่วไปแล้วคุณจะสร้างเซ็ตของการตีความขั้นสุดท้ายโดยค้นหาจากเซ็ตที่เป็นไปได้ทั้งหมดในการ รวม PD เข้าด้วยกัน คุณจะสร้างการตีความครั้งสุดท้ายที่ยอมรับได้โดยการรวมเซ็ตของ PD (เรียกว่า PD-SET) ก็ต่อเมื่อเกิดเหตุการณ์ทั้งสองเหตุการณ์นี้พร้อมกัน

1.องค์ประกอบย่อยทั้งหมดถูกแอคเคานท์โดยกฎอย่างน้อยที่สุดหนึ่งกฎเพื่อใช้ในการสร้าง PD-SET

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. PD-SET จะต้องเล็กที่สุดเท่าที่เป็นไปได้ ซึ่งก็คือจะไม่มี PD ใน PD-SET ที่สามารถย้ายออกไปและมี PD-SET ที่ยังคงแอคเคานท์ สำหรับองค์ประกอบทั้งหมด

คอมโพซิชันนอลลิตี (Compositionality)

เนื่องจากคุณได้สร้างการตีความขององค์ประกอบโดยใช้การตีความองค์ประกอบย่อยของมัน ซึ่งรูปแบบการตีความชนิดนี้เรียกว่าเป็นคอมโพซิชันนอล (compositional) โดยหลักการแล้วคุณสามารถสร้างการตีความหมายแบบล่างขึ้นบนได้ ในกรณีนี้คุณจะต้องวิเคราะห์องค์ประกอบย่อยแยกกันต่างหากจากองค์ประกอบที่ใหญ่ที่สุดที่เก็บมันไว้ ในขณะที่เรามีความต้องการใช้คอมโพซิชันนอล มันจะต้องอนุโลมในบางกรณีอย่างเช่น ในวลีที่เป็นสำนวนจะถูกกำหนดการตีความที่เป็นอิสระกับการตีความองค์ประกอบย่อยที่มีอยู่แล้วของมัน อย่างเช่น ในประโยค *John is having a ball* ซึ่งหมายถึง John กำลังสนุกกับตัวเอง ซึ่งเป็นสำนวน ดังนั้นเราจะไม่ต้องทำอะไรกับ *ball* ในความหมายของ SPHERE หรือ DANCE คุณสามารถที่จะเก็บข้อแตกต่างนี้ในการทำงานปัจจุบันโดยใช้กฎ เช่น

(S SUBJ +animate
  MAIN-V have
  OBJ (NP ART a
    HEAD ball)) → (? * ENJOY [EXPERIENCERV(SUBJ)])

ให้สังเกตุว่าเนื่องจากไม่มีการทำการทดสอบทางความหมายของ NP ที่เป็นกรรม ดังนั้นการตีความจะไม่เรียกตัวเองเพื่อที่จะตีความวลีของ NP นั้น

ผลลัพธ์ที่ได้ก็คือการตีความนี้ไม่เป็นคอมโพซิชันนอลแบบเข้มงวด มันเป็นคอมโพซิชันนอลในรูปของการอนุโลมซึ่งเป็นกรณีที่ได้กล่าวมาแล้ว ซึ่งรูปของไวยกรณ์ขององค์ประกอบย่อยจะตรงกับกฎเลยทีเดียว องค์ประกอบย่อยนั้นจะไม่มีการตีความทางความหมาย

2.6.3 ตัวอย่างการกำหนด Case Roles

พิจารณาวิธีการที่คุณสามารถใช้ในการตีความนี้เพื่อที่จะวิเคราะห์ในกรณีของคำกิริยาดังที่ดังกล่าวไปแล้ว ใน inner case จะต้องถูกวิเคราะห์ในขั้นของการวิเคราะห์คำหลักโดยกฎที่กำหนดไว้สำหรับคำกิริยาในขณะที่ outer case จะถูกวิเคราะห์ในขั้นของการวิเคราะห์ส่วนขยายโดยกฎที่กำหนดไว้สำหรับ preposition อย่างเช่น AT-LOC case ซึ่งสามารถที่จะขยายคำกิริยาได้เกือบทั้งหมด จะสร้างขึ้นโดยกฎสำหรับ at.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(at.1) (s MODS (PP PREP at

POBJ +location))→(? * (AT-LOC * V(MODS PBOJ)))

ในกฎนี้จะมีการแนะนำสัญลักษณ์เล็กน้อย ในกฎที่รูปแบบได้ขยายออกมามากกว่าหนึ่งองค์ประกอบ ฟังก์ชั่น V และ I จะใช้ลำดับของชื่อของสล็อตเป็น argument ดังนั้น V (MODS POBJ) เป็นค่าทาง semantic ของ POBJ slot ของ MODS slot ที่เข้ากันกับรูปแบบนั้น

inner case จะถูกกำหนดเอาไว้สำหรับ verb แต่ละตัวอย่างเช่น ส่วนของเซ็ตของกฎการตีความสำหรับ give ได้ถูกกำหนดไว้ในรูปที่ 2.27

(give.1) (S MAIN-V give

SUBJ +animate) → (? * GIVE1 [AGENTV(SUBJ)])

(give.2) (S MAIN-V give

OBJ +physobj) → (? * GIVE1 [THEMEV(OBJ)])

(give.3) (S MAIN-V give

IOBJ +org) → (? * GIVE1 [TO-POSSV(IOBJ)])

(give.4) (S MAIN-V give

MODS (PP PREP to

POBJ +org)) → (? * GIVE1 [TO-POSSV(MODS POBJ)])

รูปที่ 2.27 กฎการตีความหมายบางส่วนสำหรับ give

พิจารณาการตีความประโยค The man gave the company a car at the fairgrounds. โครงสร้างทางไวยากรณ์เริ่มต้น (ซึ่งรายละเอียดของการวิเคราะห์ NP ได้แสดงไว้ด้วย) เป็นดังนี้

(S SUBJ (NP "the man")

MAIN-V give

TENSE PAST

OBJ (NP "a car")

IOBJ (NP "the company")

MODS ([PP PREP at

POBJ (NP "the fairground")]))

ในขั้นตอนของการวิเคราะห์คำหลักกฎของ give ใช้ในการสร้าง PD สามตัว (จากกฎ give.1, give.2 และ give.3) ซึ่งเมื่อนำมารวมกันแล้วจะเป็นรูปของ single head analysis

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(? g1 GIVE1 [AGENT (DEF/SING m1 MAN)]

[THEME (INDEF/SING c1 AUTO)]

[TO-POSS (DEF/SING c2 COMPANY-ORG)])

สล็อตที่ไม่ได้ถูกแอคเคานท์ account เอาไว้ในขั้นนี้คือ TENSE และ MODS คุณจะต้องสร้าง PD อันหนึ่งสำหรับสล็อตเหล่านี้แต่ละตัว จากกฎ past.1 และ at.1

(PAST ? ? ?)

(? g1 ? (AT-LOC g1 (DEF/SING f1 FAIRGROUND)))

เมื่อสิ่งเหล่านี้ได้ถูกนำเข้ามารวมกัน การตีความขั้นสุดท้ายของประโยคจะเป็น

(PAST g1 GIVE1 [AGENT (DEF/SING m1 MAN)]

[THEME (INDEF/SING c1 AUTO)]

[TO-POSS (DEF/SING c2 COMPANY-ORG)]

[AT-LOC g1 (DEF/SING f1 FAIRGROUND)])

จำไว้ว่าข้อแตกต่างระหว่าง inner และ outer case จะเน้นให้เห็นโดยใช้เครื่องหมาย [] เฉพาะสำหรับ inner case เท่านั้น

พิจารณาการวิเคราะห์ที่ซับซ้อนกว่านี้อีกเล็กน้อยที่เกี่ยวข้องกับ verb break ที่อาจจะใช้ในวิธีต่าง ๆ ต่อไปนี้

(break.1) (S MAIN-V break
SUBJ +animate) → (? * BREAK [AGENTV(SUBJ)])

(break.2) (S MAIN-V break
OBJ +physobj) → (? * BREAK [THEMEV(OBJ)])

(break.3) (S MAIN-V break
SUBJ +physobj-animate
OBJ +physobj) → (? * BREAK [INSTV(SUBJ)] [THEMEV(OBJ)])

(break.4) (S MAIN-V break
SUBJ +physobj-animate
OBJ nil) → (? * BREAK [THEMEV(SUBJ)])

(break.5) (S MAIN-V break
MODS (PP PREP with
POBJ +physobj-animate)) → (? * BREAK [INSTV(MODS POBJ)])

รูปที่ 2.28 กฎการตีความหมายบางส่วนของ break

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

The amn broke the window with a hammer.

The hammer broke the window.

The window broke.

กฎสำหรับ *break* ได้แสดงไว้แล้วใน รูปที่ 2.28 กฎเหล่านี้จะตีความประโยคสามประโยคข้างต้นได้สำเร็จเช่น การวิเคราะห์ *The hammer broke the window* จะใช้ PD ที่สร้างขึ้นมาจากกฎ break.2 และ break.3 และจะสร้างการวิเคราะห์คำหลักต่อไปนี้

(? b1 BREAK [INST (DEF/SING h1 HAMMER)]

[THEME (DEF/SING w1 WINDOW)])

สังเกตุว่ากฎ break.4 จะไม่เข้ากับประโยคนี้เนื่องจาก OBJ slot ไม่ว่าง การวิเคราะห์ *The window broke* จะใช้กฎ break.4 ในการสร้าง PD ต่อไปนี้

(? b1 BREAK [THEME (DEF/SING w1 WINDOW)])

กฎ break.1 จะใช้กับประโยคนี้ไม่ได้เนื่องจากประธานไม่ใช่ animate และกฎ break.2 และ break.3 ใช้ไม่ได้เพราะว่า ไม่มีค่าอยู่ใน OBJ slot NP *the window* ถูกให้ค่าเป็น THEME case ใน instant แต่ละตัวตามที่ต้องการ

2.6.4 ลำดับชั้นของกฎ (Rule Hierarchies)

ในการตีความความหมายจะทำงานโดยใช้กฎที่ร่วมกันโดยเล็กซิคัลไอเท็ม (lexical item) และค่าทางคุณสมบัติที่เหมาะสม คุณสามารถจะสร้างโครงสร้างของกฎในการตีความได้ตามความพอใจ โดยมีกฎที่จะต้องเป็นพื้นฐานสำหรับคลาสทั้งหมดของคำ เช่น กฎ laugh.1, break.1 และ give.1 ทั้งหมดนี้ จะมีความสัมพันธ์ระหว่างสลอตของ SUBJ และ AGENT case เช่นเดียวกัน ซึ่งจริง ๆ แล้วเป็นความสัมพันธ์ซึ่งโดยปกติจะครอบคลุมคำกิริยาที่เป็นการกระทำทั้งหมด สิ่งที่คุณต้องการในที่นี้ก็คือ กฎ ๆ เดียวที่ได้รวบรวมความปกติดังต่อไปนี้

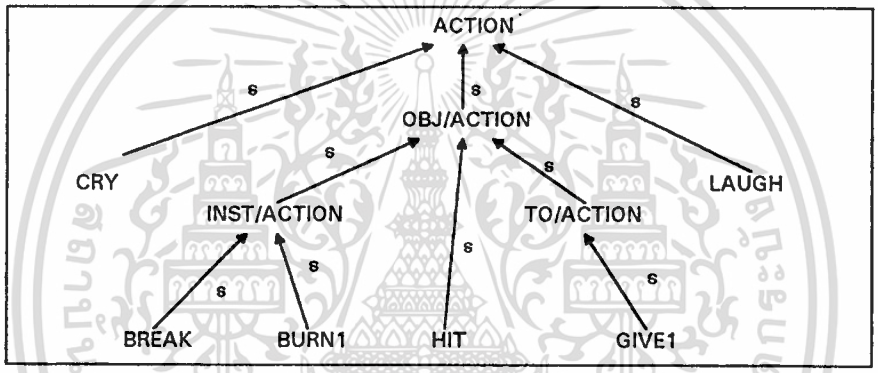
(action.1) (S MAIN-V +action

SUBJ +animate) → (? * I(MAIN-V) [AGENT V(SUBJ)])

คุณสามารถจัดโครงสร้างกฎเหล่านี้ใน semantic network โดยใช้กลไกของ inheritance เพื่อว่ากฎทั้งหมด จะถูกประยุกต์ใช้ให้เป็น type T ซึ่งใช้ได้กับสับไทพ์ (subtype) ทุก ๆ ตัวและอินสแตนซ์ (instance) ของ T เช่น คุณสามารถที่จะแสดงเซ็ตของกฎสำหรับ laugh, give, break โดยใช้ type hierarchy และการตีความกฎที่ได้แสดงในรูปที่ 2.29 ในขณะที่จำนวนของกฎแต่ละตัว โดยการใช้การจัดโครงสร้างเช่นนี้ไม่เล็กลงมากไปกว่าใช้กฎในคำกิริยาแต่ละตัวโดยตรง แต่คุณจะเห็นข้อได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เปรียบ ถ้าคุณเพิ่มคิกิริยาเข้าไปใหม่อย่างเช่น คุณจะมีกฎในการตีความสำหรับ cry ได้โดยการแบ่งประเภทของมันให้เป็นอินสแตนซ์ของ ACTION คุณจะได้รับกฎของ hit โดยการแยกมันให้เป็น OBJ/ACTION คุณจะได้รับกฎของ burn โดยการแบ่งให้มันเป็น INST/ACTION และอื่น ๆ อีก ดังนั้นคุณสามารถขยายเล็กซิคอนของคุณได้อย่างง่ายดายโดยใช้กฎโครงสร้างของ hierarchy ที่มีอยู่ ปัญหาทางเทคนิคอันหนึ่งที่ทำให้การใช้ hierarchy เหล่านี้ซับซ้อนมากกว่าครั้งแรกที่มันปรากฏในปัญหาที่เฉพาะเจาะจงลงไปจะเพิ่มขึ้นเรื่อย ๆ เมื่อกฎนั้นได้รับการสืบทอดเข้ามาแต่มันธรรมดาจนเกินไป เช่นในความหมายกิริยา BURN1 สืบทอดให้กับกฎที่เหมาะสมจาก INST/ACTION เพื่อว่าประโยคอย่างเช่น Jack burn the meat with a blowtorch จะได้รับการตีความตามความเหมาะสมดังนี้



(action.1) (S MAIN-V +action
SUBJ +animate) → (? * I(MAIN-V)[AGENT V(SUBJ)])

(obj/action.1) (S MAIN-V +obj/action
OBJ +anything) → (? * I(MAIN-V)[THEME V(OBJ)])

(to/action.1) (S MAIN-V +to/action
IOBJ +org) → (? * I(MAIN-V)[TO-POSS V(IOBJ)])

(to/action.2) (S MAIN-V +to/action
MODS (PP PREP to
POBJ +org)) → (? * I(MAIN-V)[TO-POSS V(MODS POBJ)])

(inst/action.1) (S MAIN-V +inst/action
SUBJ +physobj-animate
OBJ +physobj) → (? * I(MAIN-V)[INST V(SUBJ)]
[THEME V(OBJ)])

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(inst/action.2) (S MAIN-V +inst/action

SUBJ +physobj-animate

OBJ nil) → (? * I(MAIN-V)[THEME V(SUBJ)])

(inst/action.3) (S MAIN-V +inst/action

MODS (PP PREP with

POBJ +physobj-animate)) → (? * I(MAIN-V)[INST V(MODS POBJ)])

รูปที่ 2.29 ระดับชั้นของ RULE BASE

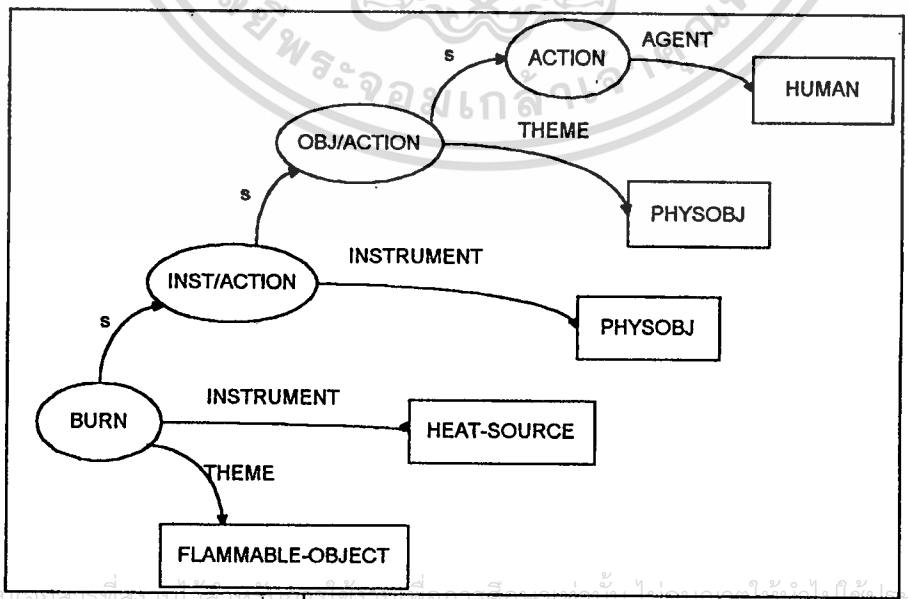
(PAST b1 BURN1 [AGENT (NAME j1 PERSON "Jack")]

[THEME (DEF/SING m1 MEAT)]

[INSTRUMENT (INDEF/SING b2 TOOL-TYPE-7)])

แต่กฎที่ใช้จะธรรมดามากเกินไปสำหรับ burn มันไม่สามารถจะเก็บข้อบังคับที่ว่า THEME จะต้องเป็น burnable object และ instrument จะต้องเป็นแหล่งกำเนิดความร้อน ดังนั้นประโยคอย่างเช่น Jack burn the rock with some water จะถูกต้องตามความหมายตามกฎนี้

ผลลัพธ์ก็คือคุณไม่สามารถที่จะแสดงถามหาความหมายของคำได้อย่างสมบูรณ์ โดยใช้กฎการตีความที่ได้รับการสืบทอด คุณจะต้องใช้กลไกอื่น ๆ ที่สามารถระบุคำ โดยการใช้ semantic network อย่างเข้มงวด คุณสามารถใช้ข้อมูลของ selectional restriction ที่จะทำสิ่งนี้ได้ ดังแสดงในรูปที่ 2.30



รูปที่ 2.30 กฎของข้อจำกัดสำหรับ burn

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ด้วยความสามารถส่วนขยาย application ของกฎการตีความจะต้องประกอบไปด้วยสองส่วน คือ

1. การจับคู่ (matching)

2. การสร้าง PD

และหลังจากนั้นแอปพลิเคชันของ case restriction จะตรวจสอบว่า PD ที่ได้มานั้นถูกต้องตามซีแมนติค หรือไม่ ในประโยคที่มีแบบเฉพาะ Jack burn the rock with the water จะมีโครงสร้างทางไวยากรณ์ดังนี้

```
(S SUBJ (NP NAME John)
   MAIN-V burn
   TENSE PAST
   OBJ (NP DET the
         HEAD rock)
   MODS(PP PREP with
         POBJ (NP DET the
               HEAD water)))
```

โดยใช้กฎ action.1, obj/action.1, inst/action.3 ที่ทั้งหมดจะเป็นกฎที่ถูกสืบทอดมาสำหรับ burn head analysis จะสร้าง PD ต่อไปนี้

```
(? b1  BURN [AGENT (NAME j1 PERSON "John")])
(? b1  BURN [THEME (DEF/SING s1 STONE)])
(? b1  BURN [INSTYY (DEF/SING w1 WATER)])
```

หลังจากนั้นมันจะตรวจสอบ PD เหล่านี้กับ selectional restriction สำหรับ BURN และไม่ยอมรับสองอันหลังเนื่องจากมันผิดปกติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.7 Strategies for Semantic Interpretation

ในหัวข้อที่แล้วเราได้พูดถึงวิธีการตีความทางซีแมนติค ซึ่งเกี่ยวข้องกับส่วนของการจัดการในเรื่องซินแทคติคและซีแมนติคในหัวข้อนี้เราจะพิจารณาวิธีอื่น ๆ ของการจัดการโครงสร้างของการตีความทางซีแมนติคที่พบในเอกสาร บทความหรือวรรณคดีต่าง ๆ วิธีการเหล่านี้รวมทั้งวิธีการในการเข้ารหัสข้อมูลทางซีแมนติค (semantic information) ไปเป็นไวยากรณ์ทางซินแทค, เทคนิคของการโมดูลแต่ทำงานในลักษณะเพิ่มขึ้น, เทคนิคการตีความทางซีแมนติค ในรูปของสัญลักษณ์บน grammatical rule และเทคนิคที่จำเป็นสำหรับ semantically driven และใช้ข้อมูลทางซินแทคติค (syntactic information) ให้น้อยที่สุด

โครงสร้างทั้งหมดเหล่านี้มีข้อดีและข้อเสียเฉพาะโครงสร้างนั้น ๆ ซึ่งมันแปรเปลี่ยนไปตามการจัดโครงสร้างของไวยากรณ์และข้อมูลของการตีความทางซีแมนติค (semantic interpretation information) ซึ่งมันเป็นเรื่องปกติสำหรับการทำงานกับไวยากรณ์ที่ใหญ่ เพื่อที่จะทำให้การเปรียบเทียบง่ายขึ้น ในหัวข้อนี้เราจะแสดงตัวอย่างง่าย ๆ ของโครงสร้างที่ได้กล่าวไว้ในหัวข้อที่แล้ว ตัวอย่างเดียวกันนี้จะนำมาใช้อีกเมื่อพิจารณาโครงสร้างอีกวิธีการหนึ่ง

เทคนิคหลักสำหรับการนำข้อมูลของซีแมนติคมารองรับเท่าที่จะเป็นไปได้จะเป็นการลดความกำกวมลงไป ซึ่งเราก็ได้เห็นตัวอย่างประโยคที่มีความกำกวมซึ่งจริง ๆ แล้วมันจะไม่กำกวมเมื่อเราพิจารณาข้อมูลทางซีแมนติค ตัวอย่างเช่นประโยค We broke the window with a hammer. ซึ่งเป็นประโยคที่ถูกซินแทคแต่กำกวม ซึ่ง PP Phrase ตัวสุดท้ายได้ทำหน้าที่ขยาย NP (the window) หรือกริยาหลัก อย่างไรก็ตามการใช้ข้อมูลทางซีแมนติค เราอาจไม่สามารถหาข้อวิเคราะห์ของ NP (the window with a hammer) เราสามารถที่จะวิเคราะห์ breaking with a hammer ได้ง่าย แม้ว่าในที่นี้ hammer ที่เติม Instrument เป็นกรณีของ breaking event เช่นเดียวกับประโยค *The sun rose from behind the toolbox with a hammer.* ขณะที่ syntac มันกำกวมเช่นเดียวกับตัวอย่างก่อนหน้านี้ซึ่งสามารถตีความทางซีแมนติคด้วย PP ขยาย NP ในประโยคที่มีส่วนที่กำกวมหลายส่วน ถ้าจำนวนของ syntac มากเป็นไปได้ที่สามารถตีความได้อย่างมากมาย

2.7.1 โดเมนอย่างง่าย (Simple Domain)

เราจะอธิบายไวยากรณ์เล็ก ๆ และเซ็ตของกฎของการตีความหมายซึ่งจะใช้สำหรับเปรียบเทียบผลที่ได้ ไวยากรณ์จัดการกับเซ็ตของคำถาม ซึ่งสามารถถามเกี่ยวกับตารางเวลาของสายการบิน กฎของไวยากรณ์และการตีความเฉพาะเรื่องสามารถใช้กับตัวอย่างต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

When does the flight from Boston to Chicago leave?

At what time does the flight to Chicago leave Boston?

Can you book a flight to Chicago for me?

Rules	Equation
1. S ← AUX NP VP	
2. S ← PP AUX NP VP/PP	$TYPE_{PP}$ = WH
	$HELD_{VP/PP}$ = PP
	$ROOT_{AUX}$ = DO
3. VP ← SIMPLE-VP	
4. VP ← SIMPLE-VP PPS	
5. VP/PP ← SIMPLE-VP PPS/PP	$HELD_{PPS/PP}$ = HELD
6. SIMPLE-VP ← VERB	$SUBCAT_V$ = NONE
7. SIMPLE-VP ← VERB NP	$SUBCAT_V$ = OBJ
8. NP ← ART NOUN	TYPE = $TYPE_{ART}$
9. NP ← ART NOUN PPS	TYPE = $TYPE_{ART}$
10. NP ← PROPER-NAME	
11. NP ← PRO	
12. NP/PP ← ART NOUN PPS/PP	$HELP_{PPS/PP}$ = HELD
13. PPS ← PP	
14. PPS ← PP PPS	
15. PPS/PP ←	
16. PPS/PP ← PPS	
17. PP ← PREP NP	TYPE = $TYPE_{NP}$
18. PP ← WH-WORD	TYPE = WH
	HEAD = WH-WORD

รูปที่ 2.31 ไวยากรณ์เพิ่มเติม

รูปที่ 2.31 แสดงไวยากรณ์สำหรับคำถามเหล่านี้ในรูปแบบของ CFG โดยใช้เทคนิคที่อธิบายในหัวข้อที่ 5 ในการจัดการกับคำถาม annotation ที่แสดงเป็นการจัดการกับ Wh-movement และ Verb Subcatetorization ด้วย structure-building ที่เหมาะสม ในกรณีนี้จะถูกแสดงตามการจัดการจากประโยคแรก *When does the flight from Boston to Chicago leave?* จะมีโครงสร้างดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(S TYPE WH-Q

WH-QUERY (PP1 TYPE WH HEAD when)

SUBJ (NP DET the

HEAD flight

MODS ((PP PREP from

POBJ (NP NAME Boston))

(PP PREP to

PBOJ (NP NAME Chicago))))

MAIN-V leave

MODS (→PP1))

ประโยคที่สอง *At what time does the flight to Chicago leave Boston?* ก็มีโครงสร้างคล้ายกันยกเว้นแต่ว่ามีออปเจ็ค Boston ในโครงสร้าง S ด้วย ประโยคที่สาม *Can you book a flight to Chicago for me?* มีไวยากรณ์ที่กำกวมเพราะว่า มี PP Phrase สองที่ที่ตอนท้ายของประโยค ที่อาจจะเชื่อมกับคำกิริยาหรือ NP ตัวสุดท้ายก็ได้ ด้วยความเป็นไปได้จากตัวอย่างนี้เราสามารถคิดว่ามันอาจจะเป็นไปได้ 3 รูปแบบ ดังรูปที่ 2.32

1. (S TYPE Y-N-Q

SUBJ (NP PRO you)

AUXS (can)

MAIN-V book

OBJ (NP DET a

HEAD flight

MODS ((PP PREP to

POBJ (NP NAME Chicago))

(PP PREP for

POBJ (NP PRO me)))))

2. (S TYPE Y-N-Q

SUBJ (NP PRO you)

AUXS (can)

MAIN-V book

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

OBJ (NP DET a

HEAD flight

MODS ((PP PREP to

POBJ (NP NAME Chicago))

(PP PREP for

POBJ (NP PRO me))))

3. (S TYPE Y-N-Q

SUBJ (NP PRO you)

AUXS (can)

MAIN-V book

OBJ (NP DET a

HEAD flight

MODS ((PP PREP to

POBJ (NP NAME Chicago))))

MODS((PP PREP for

POBJ (NP PRO me))))

รูปที่ 2.32 โครงสร้างที่เป็นไปได้ของประโยค Can you book a flight to Chicago for me.?

กฎของการตีความหมายต้องการโดเมนนี้รวมทั้งกฎที่แสดงในรูปที่ 2.33 แฟลกเมนต์ (fragment) ของระดับชั้นของคำกิริยา มีความสัมพันธ์กัน สำหรับตัวอย่างนี้แสดงดังรูปที่ 2.33 ในขณะที่ส่วนที่สูงกว่าของ hierarchy แสดง Class ทั่ว ๆ ไปของคำกิริยาเช่น ACTION และ EVENT ส่วนที่ต่ำลงมาของ hierarchy คือส่วนที่ระบุถึงคำกิริยาที่ใช้ในโดเมนนี้ สำหรับ *depart* ที่กำหนดในที่นี้ใช้สำหรับการบิน(Flights)เท่านั้น TO และ FROM location (TO-LOC,FROM-LOC) จะต้องถูกตามด้วยเมือง (CITY) ต่าง ๆ กฎที่ใช้สำหรับการตีความ NP และ wh-terms แสดงดังรูปที่ 2.34 คำนามถูกสมมติให้มีโครงสร้างคล้ายกับระดับชั้นของคำกิริยา ในตัวอย่างนี้กฎ (anything.1) เท่านั้นซึ่งใช้กับคำนามที่เป็นชื่อ คือไม่มี DET หรือ PRO นำหน้าในกฎ

(action.1) (S VERB + action

SUBJ + human) → (? * I(VERB)[AGENT V(SUBJ)])

(obj/action.1) (S VERB + obj/action

OBJ + anything) → (? * I(VERB)[THEME V(OBJ)])

(event.1) (S VERB + event

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                OBJ + physobj-human) → (?  *  I(VERB)[THEME V(SUBJ)])
(leave.1)       (S VERB + departing
                OBJ + location) → (?  *  I(VERB)[FROM-LOC V(OBJ)])
(to.1)          (S MAIN-V + vehicle
                MODS(PP PREP to
                      POBJ + location)) → (?  *  I(MAIN-V)
                                          (TO-LOC  *  V(MODS POBJ)))
(to.2)          (S HEAD + vehicle
                MODS(PP PREP to
                      POBJ + location)) → (?  *  I(HEAD)
                                          (TO-LOC  *  V(MODS POBJ)))
(from.1)        (S MAIN-V + vehicle
                MODS(PP PREP to
                      POBJ + location)) → (?  *  I(MAIN-V)
                                          (FROM-LOC  *  V(MODS POBJ)))
(from.2)        (NP HEAD + vehicle
                 MODS(PP PREP from
                       POBJ + location))
                            → (?  *  I(HEAD)
                              (FROM-LOC  *  V(MODS POBJ)))
(for.1)         (S MIAN-V + ben/action
                MODS (PP PREP for
                      POBJ + org))
                            → (?  *  I(MAIN-V)
                              (BENFICIARY  *  V(MODS POBJ)))
```

รูปที่ 2.33 กฎสำหรับคำกิริยาและคำบุพบท

การประยุกต์กฎเหล่านี้ไปเป็นโครงสร้างของไวยากรณ์ (syntactic structure) เราสามารถวิเคราะห์สองประโยคแรกที่กล่าวมาแล้วดังต่อไปนี้

When does the flight from Boston to Chicago leave?

```
(PRES d1 DEPARTING
      (THEME(DEF/SING f1 FLIGHT
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(FROM-LOC f1 (NAME b1 CITY "Boston"))

(TO-LOC f1 (NAME c1 CITY "Chicago"))))]

[AT-TIME(WH/SING t1 TIME)])

At what time does the flight to Chicago leave Boston?

(PRES I2 DEPARTING

[THEME(DEF/SING f2 FLIGHT

(TO-LOC f2 (NAME c1 CITY "Chicago"))))]

[FROM-LOC f1 (NAME b1 CITY "Boston")]

[AT-TIME(WH/SING t2 TIME)])

ในประโยคที่สาม *Can you book a flight to Chicago for me?* ซึ่งมีกรณีที่เป็นไปได้ในการวิเคราะห์ไวยากรณ์สามกรณี ดังแสดงในรูปที่ 2.32 จากที่เห็นมีเพียงอันเดียวที่สามารถตีความทางซีแมนติคได้ถูกต้องในรูปที่ 2.32 ข้อ 1 ไม่สามารถวิเคราะห์ได้เนื่องจากกฎสำหรับ PP ที่ตามด้วย PREP for ต้องการให้มันขยาย BEN/ACTION verb ไม่ใช่คำว่า flight เช่นเดียวกันสำหรับกฎ PP ตามด้วย PREP to สามารถขยาย MOTION/VERB หรือ VEHICLE แต่ไม่ใช่แอคชั่นของ RESERVING ดังนั้นสามารถการตีความในข้อ 2 ก็ไม่ต้องพิจารณา การตีความในข้อ 3 จะสามารถวิเคราะห์ได้ดังนี้

(PRES b1 RESERVING

[AGENT(PRO y1 HUMAN "you")]

[THEME(DEF/SING f3 FLIGHT

(TO-LOC f3 (NAME c1 CITY "Chicago"))))]

[BENEFICIARY (PRO m1 HUMAN "me")])

(when.1) (S MODS

(PP HEAD when)) → (? * ? (AT-TIME * (WH/SING * * TIME)))

(at.1) (S MODS

(PP PREP at

POBJ + time)) → (? * ? (AT-TIME * V(POBJ)))

(wh.1) (NP DET + wh

NUM {3p}

HEAD + anything) → (WH/PL * I(HEAD))

(wh.2) (NP DET + wh

NUM {3s}

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                    HEAD + anything) → (WH/SING * I(HEAD))
(name.1)        (NP NAME + name) → (NAME * I(NAME) V(NAME))
(anything.1)    (NP HEAD + anything) → (? * I(HEAD))
(the.1)         (NP DET the
                    NUM (3s)) → (DEF/SING * ?)
(the.2)         (NP DET the
                    NUM (3p)) → (DEF/PL * ?)
(pro.1)         (NP PRO + pro) → (PRO * I(PRO) V(PRO))
(a.1)           (NP DET a) → (INDEF/SING * ?)
```

รูปที่ 2.34 กฎสำหรับ wh-terms และ NP

เหตุผลสำหรับการตัดสินใจรวมการวิเคราะห์ซินแทคติคและซีแมนติคก็เพื่อกำจัดการตีความที่ผิด ๆ จากไวยากรณ์ให้มากที่สุดเท่าที่จะเป็นไปได้ และเป็นการลดความกำกวมในระหว่างการทำ parsing ด้วย

2.7.2 ไวยากรณ์ทางซีแมนติค (Semantic Grammars)

เทคนิคแรกสำหรับการรวมการประมวลผลของซินแทคติคและซีแมนติคเพื่อที่จะทำให้มันเป็นซิงเกิลยูนิฟอร์มแฟรมเวอร์ค (single uniform framework) หรือไม่ก็ CFG หรือ ATN ผลลัพธ์ที่ได้จะเรียกว่าไวยากรณ์ทางซีแมนติคซึ่งคล้าย ๆ กับ regular grammar ยกเว้นแต่ว่ามันใช้ชนิดของซีแมนติคสำหรับ terminal symbol ตัวอย่างเช่นจะมีกฎดังนี้

VP ← SIMPLE-VP PPS

ไวยากรณ์ทางซีแมนติคจะมีกฎที่แยกกันสำหรับแต่ละคลาสของ semantic Verb ดังนั้นเราอาจจะแทนกฎเหล่านี้ด้วยกฎต่อไปนี้ในรูป simple airline query domain:

(sem.1) RES-VP ← RESERVING RES-MODS

(sem.2) DEP-VP ← DEPARTING DEP-MODS

กฎ (sem.1) ซึ่งอาจจะเป็นจริงสำหรับเฟส book a flight for me ซึ่งจะเกี่ยวข้องกับคำกิริยาเช่น book ผนวกกับส่วนเติมที่ให้มีได้ (allowable complement) กฎ (sem.2) สามารถมองอีกนัยหนึ่งคือคำกิริยาของ departure และอาจจะเป็นจริงสำหรับเฟส leaves for Chicago

เช่นเดียวกันกฎสำหรับ SIMPLE-VP และ PPS จะถูกแทนที่โดยไวยากรณ์ทางซีแมนติคที่ equivalent สำหรับ RESERVING verb เราจำเป็นต้องใช้กฎต่อไปนี้

(sem.3) RESERVING ← RESERVE-VERB FLIGHT

(sem.4) RES-MODS ← for PERSON

เอกสารนี้มีสำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(sem.5) RES-MODS ←

จะสังเกตุได้ว่าส่วนขยายของ PP เหล่านั้น ที่สามารถตีความทางซีแมนติคได้นั้น ตอนนี้อนุญาตให้ใช้โดย grammar fragment นี้ โดยเฉพาะส่วนขยายของ PP เช่น for me จะยอมรับได้ในเฟสของ book it for me อย่างไรก็ตาม PP to Boston จะไม่สามารถยอมรับได้

สำหรับ DEPARTING verb เราต้องใช้กฎต่อไปนี้

(sem.6) DEPARTING ← DEPART-VERB

(sem.7) DEPARTING ← DEPART-VERB SOURCE-LOCATION

(sem.8) DEP-MODS ← DEP-MOD DEP-MODS

(sem.9) DEP-MODS ←

(sem.10) DEP-MOD ← to DEST-LOCATION

(sem.11) DEP-MOD ← from SOURCE-LOCATION

ส่วนขยาย PP เดียวเท่านั้นที่ยอมรับได้ทางซีแมนติคตอนนี้จะยอมรับได้โดยไวยากรณ์นี้ด้วย

ดังนั้นซินแทคติค rules ทั้ง 6 (3,4,6,7,13,14) ได้ถูกแทนที่ด้วยไวยากรณ์ทางซีแมนติค ทั้ง 11 กฎ เพื่อให้ครอบคลุมชนิดของคำกิริยาสองชนิด การเพิ่มชนิดของคำกิริยาให้มากขึ้น สามารถต่อเนื่องด้วยการขยายไวยากรณ์ทางซีแมนติคซึ่งจะได้กฎใหม่ประมาณ 6 กฎต่อชนิดของคำกิริยาดังนั้นไวยากรณ์ทางซีแมนติค จึงมีแนวโน้มที่จะมีขนาดใหญ่กว่าไวยากรณ์ทางซีแมนติคที่สอดคล้องกัน ดังนั้นประสิทธิภาพของพาร์เซอร์อาจจะลดการพิจารณาลงไปได้ อันที่จริงแล้วสิ่งนี้ไม่ได้เกิดขึ้นในหลาย ๆ กรณีเนื่องจากจำนวนของคำในแต่ละ terminal caterogies มีจำนวนน้อยกว่ามาก (ตัวอย่างเช่น เปรียบเทียบจำนวนของคำในชนิดของ DEPART-VERB ด้วยจำนวนใน syntactic caterogies VERB) ในที่นี้หมายถึงว่าประโยคใด ๆ ที่มีเพียงซับเซ็ตเล็ก ๆ ของไวยากรณ์เท่านั้นที่จะใช้ได้

เหตุผลข้อหนึ่งที่ไวยากรณ์มีขนาดใหญ่ก็คือ ความหลากหลายของหลักภาษาซึ่งบางครั้งต้องใช้วิธีการจัดการแบบบรูต-ฟอร์ค (brute-force) ตัวอย่างเช่น ในการจัดการกับประโยคในรูปแบบพาสซีฟเราต้องเพิ่มกฎสำหรับรูปแบบพาสซีฟทุก ๆ ประเภทของ semantic Verb ที่ใช้ (ข้อหนึ่งสำหรับ DEPARTIVE อีกข้อหนึ่งสำหรับ RESERVING และอื่น ๆ) เมื่อเราพยายามที่จะหาวิธีแก้ปัญหาที่ดีขึ้น เราต้องเคลื่อนย้ายไวยากรณ์

ตัวอย่าง พิจารณาว่าอะไรจะเกิดขึ้นถ้าเราต้องการให้เก็บ hold list solution สำหรับประโยคคำถามแบบ WH วิธีการบรูต-ฟอร์คจะ duplicate rule 2 สำหรับคำกิริยาแต่ละชนิด ถ้าเราพิจารณาเพียงชนิดของ DEPARTING เราต้องใช้กฎต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

a	S	←	WH-DEP-MOD AUX FLIGHT HELD = WH-DEP-MOD DEP-VP/DEP-MOD
(sem.13)	WH-DEP-MOD	←	when
(sem.14)	WH-DEP-MOD	←	to what DEST-LOCATION
(sem.15)	WH-DEP-MOD	←	from what SOURCE-LOCATION
(sem.16)	WH-DEP-MOD	←	at what TIME
(sem.17)	DEP-VP/DEP-MOD	←	DEPARTING HELD= $HELD_{DEP\text{-}VP/DEP\text{-}MOD}$ DEP-MODS/DEP-MOD
(sem.18)	DEP-MODS/DEP-MOD	←	FIRST = HELD
(sem.19)	DEP-MODS/DEP-MOD	←	DEP-MODS FIRST = HELD

ถ้าเราจำเป็นต้องใช้กฎเช่นเดียวกันเพื่อจัดการกับ RESERVING verb

ถ้าเราต้องการที่จับคอนเซ็ปต์นี้ในวิธีทั่ว ๆ ไป เราจะต้องผ่อนผันเป้าประสงค์ (goal) ซึ่งเฉพาะประโยคที่ตีความทางซีแมนติคสามารถที่จะพาร์สโดยไวยากรณ์ โดยเฉพาะกฎ (sem.12) และคำจำกัดความของ WH-DEP-MOD ในกฎ (sem.13) ถึง (sem.16) จำเป็นในการบังคับให้ส่วนของ query สามารถมีส่วนขยายของคำกิริยาที่ถูกต้อง ซึ่งสิ่งนี้เป็นการประกันว่า เมื่อส่วนประกอบ HELD ถูกใช้โดยกฎ (sem.18) หรือ (sem.19) ก็จะสร้างโครงสร้างการตีความทางซีแมนติคได้ ถ้าเราใช้กฎทั่ว ๆ ไปในการจัดการกับ WH-TERM ใด ๆ ที่อยู่ในส่วนเริ่มต้นของ ประโยคคำถามแบบ WH เราจะไม่มีอะไรเป็นสิ่งประกันได้เลยว่าส่วนประกอบ HELD จะสามารถแทรกเป็น DEP-MOD และสร้างประโยคที่ถูกต้องขึ้นมาได้

ในทางปฏิบัติไวยากรณ์ทางซีแมนติคจะใช้ได้ประโยชน์มากที่สุดสำหรับการสร้างความถูกต้องของการเชื่อมต่อของภาษาธรรมชาติ (natural language interface) ในโดเมนที่จำกัด ซึ่งสามารถที่จะดำเนินการวิเคราะห์ขอบเขตของปัญหาได้ สำหรับระบบทั่ว ๆ ไป งานที่เกี่ยวกับการสร้างไวยากรณ์ดูเหมือนว่าจะยากมากจนยอมรับไม่ได้นอกจากเราจะอาศัยประเภทของการสร้างไวยากรณ์จากการกำหนดรายละเอียดซินแทคติคและซีแมนติคของผู้ใช้โดยอัตโนมัติ

2.7.3 A Simple Interleaved Syntactic and Semantic Analyzer

ข้อดีของการแยกซินแทคติคและขบวนการทางซีแมนติคก็คือ มันยอมให้มีพาร์ซิงโมเดล (parsing model) เนื่องจากว่า ส่วนของซินแทคติคระบบเป็นโดเมนที่อิสระดังนั้นแอพพลิเคชั่นใหม่แต่ละตัวสามารถที่จะสร้างงานที่กล่าวก่อนหน้านี้บนแอพพลิเคชั่นอื่น ๆ พูดอีกนัยหนึ่งก็คือการรวมขบวนการทางซินแทคติคและซีแมนติคแบบ single stage เช่นในไวยากรณ์ทางซีแมนติคสามารถใช้ข้อมูลทางซีแมนติคเพื่อกำจัดการตีความโดยใช้ซินแทคติคจำนวนมากซึ่งมีซีแมนติคที่ผิดความหมายในหัวข้อนี้จะเป็นรายละเอียดของระบบง่าย ๆ ซึ่งจะทำให้เราได้ประโยชน์มากมาย approaches ทั้งสองนี้ ขบวนการทางซีแมนติคและซินแทคติคจะแยกจากกัน ซึ่งข้อมูลทางซีแมนติคจะนำไปสู่การวิเคราะห์ซินแทคติค

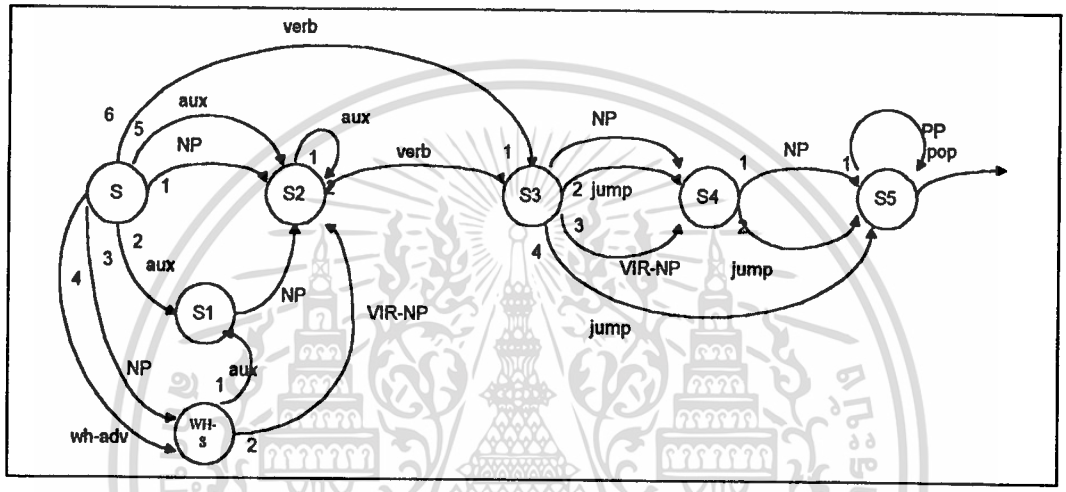
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พิจารณาการวิเคราะห์ซีแมนติคของโครงสร้าง NP ซึ่งในตัวอย่างเกือบทั้งหมดของการตีความของ NP ได้กระทำอย่างอิสระของคอนเท็กซ์ (โครงสร้างข่อง S หรือ PP) ที่มันพบในสถานการณ์นี้ เราสามารถเรียกใช้การตีความทางซีแมนติคได้ทันที เมื่อโครงสร้างของ NP ถูกกำหนดโดยซินแทคแล้ว ถ้า NP ถูกตีความสำเร็จการตีความนี้จะถูกบันทึกและซินแทคติคก็จะพาร์สอย่างต่อเนื่องกันตามปกติ ถ้าการตีความทางซีแมนติคไม่สามารถตรวจพบได้ ตัวประมวลผลทางซินแทคติค (syntactic processor) จะเฟล (fail) ตอนที่ตรวจ NP (นั่นคือ มันจะแบ็คเทรกหรือเฟล ในการที่จะเพิ่ม NP ไปยังชาร์ทที่จะใช้ต่อจากนั้น ซึ่งขึ้นอยู่กับอัลกอริทึมในการทำ parsing ที่ถูกใช้ด้วย)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
การคำนวณและการสร้าง

ในการคำนวณและการสร้างงาน เราเริ่มจากการออกแบบ ATN (Augment Transition Network) ซึ่งเป็นไดอะแกรมที่แสดงสถานะ (state) ของการวิเคราะห์ประโยค โดยจะประกอบด้วย ATN 3 รูปคือ ATN ของประโยครวม (S:Sentence), ATN ของส่วนวลีของคำนาม (NP:Noun Phrase) และ ATN ของส่วนวลีของบุพบท (PP:Preposition Phrase) แสดงดังรูป



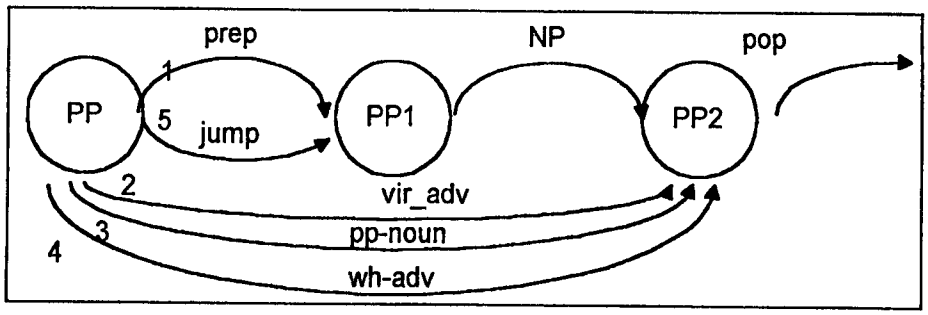
Arc	Test	Actions
S/1	——	FIRST-NP ← *
		MOOD ← DECL
S/2	——	AUXS ← Append(AUXS, *)
		FIRST-V ← *
		MOOD ← Yes-No-Q
S/3	TYPE = WH	HOLD *
		MOOD ← WH-Q
		WH-QUERY ← *
S/4	TYPE = WH	HOLDADV *
		MOOD ← WH-Q
		WH-QUERY ← *
S/5	Form * = inf	First-NP ← "You"
		Append(Auxs, *) ← AUXS
		First-V ← *
		MOOD ← Command

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

S/6	Form * = int	MAIN-V ← *
		VOICE ← ACTIVE
		SUBJ ← "You"
		MOOD ← Command
S1/1	——	FIRST-NP ← *
S2/1	AuxAgree(* ,AUXS)	If FIRST-V is nil
		then FIRST-V ← *
		AUXS ← Append(AUXS, *)
S2/2	AuxAgree(* ,AUXS)	MAIN-V ← *
	If FIRST-V not nil	If $ROOT_{LAST(AUXS)}$ = BE
	then Agr($NUM_{FIRST-NP}$,$NUM_{FIRST-V}$)	& FORM * = en
	else Agr($NUM_{FIRST-NP}$,NUM *)	then VOICE ← PASSIVE
		OBJ ← FIRST-NP
		else VOICE ← ACTIVE
		SUBJ ← FIRST-NP
S2/3	First-V = BE	MAIN-V ← First-V
	Agr($NUM_{First-NP}$,$NUM_{First-V}$)	VOICE ← ACTIVE
		First-NP ← SUBJ
S3/1	Trans($SUBCAT_{MAIN-V}$)	FIRST-OBJ ← *
	& VOICE = ACTIVE	
S3/2	VOICE = PASSIVE	
S3/3	Trans($SUBCAT_{MAIN-V}$)	FIRST-OBJ ← *
	& VOICE = active	
S3/4	Intrans($SUBCAT_{MAIN-V}$)	
S4/1	$SUBCAT_{MAIN-V}$=IOBJ+OBJ	OBJ ← *
		IOBJ ← FIRST-OBJ
S4/2	——	OBJ ← FIRST-ONJ
S5/1	——	MODS ← Append(MODS, *)
WH-S/1	——	AUXS ← Append(AUXS, *)
		FIRST-V ← *
WH-S/2	——	FIRST-NP ← *

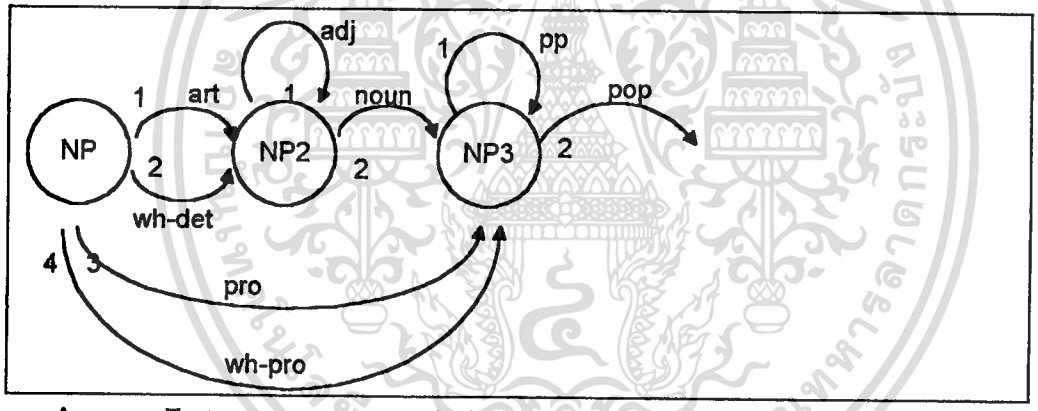
รูปที่ 3.1 รูปแสดง ATN ของ S

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



Arc	Test	Actions
PP/1	—	PREP ← *
PP/2	—	HEAD ← *
PP1/1	—	POBJ ← *
PP/3	—	HEAD ← *
PP/4	—	HEAD ← *
PP/5	MOOD = Wh-q-adv	—

รูปที่ 3.2 รูปแสดง ATN ของ PP

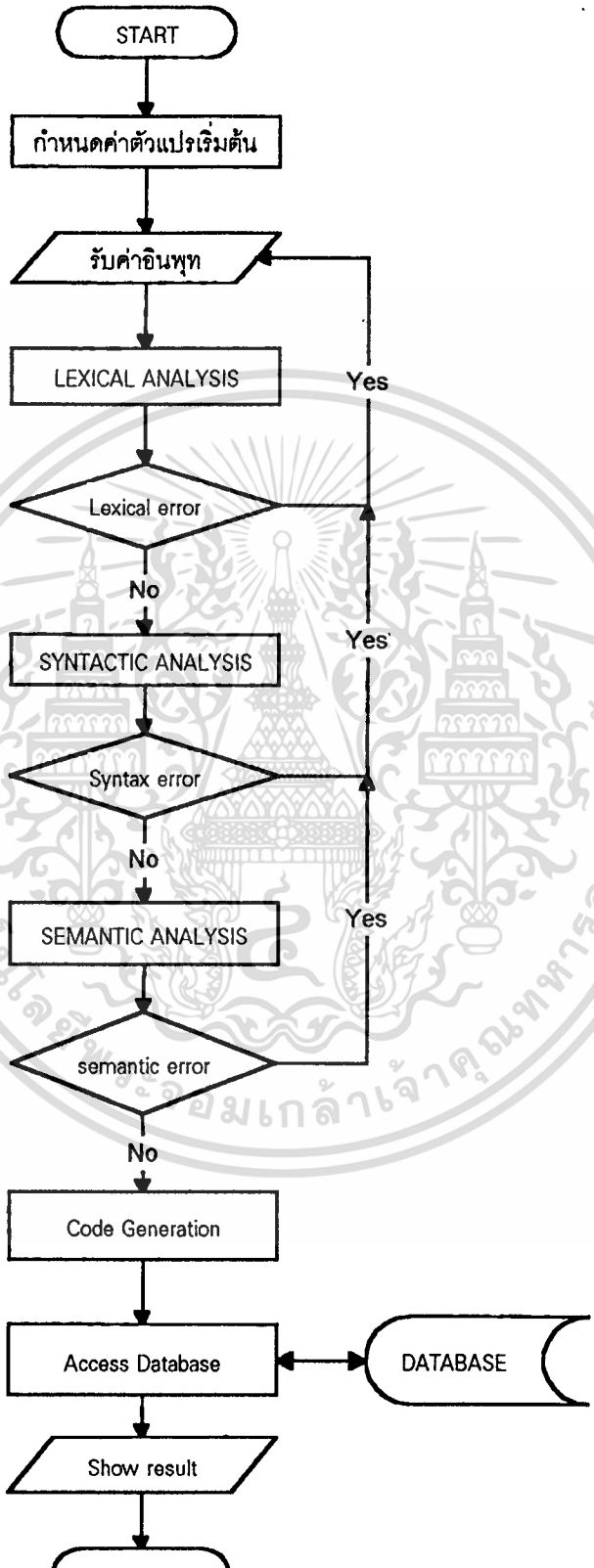


Arc	Test	Actions
NP/1	—	DET ← * NUM ← NUM *, TYPE ← Simple
NP/2	—	DET ← * TYPE ← WH
NP/3	—	PRO ← *, TYPE = Simple
NP/4	—	PRO ← * TYPE ← WH
NP2/1	—	ADJS ← Append(ADJS, *)
NP2/2	Agr(NUM,NUM *)	HEAD ← * NUM ← Agr(NUM,NUM *)
NP3/1	—	NMODS ← Append(NMODS, *)

รูปที่ 3.3 รูปแสดง ATN ของ NP

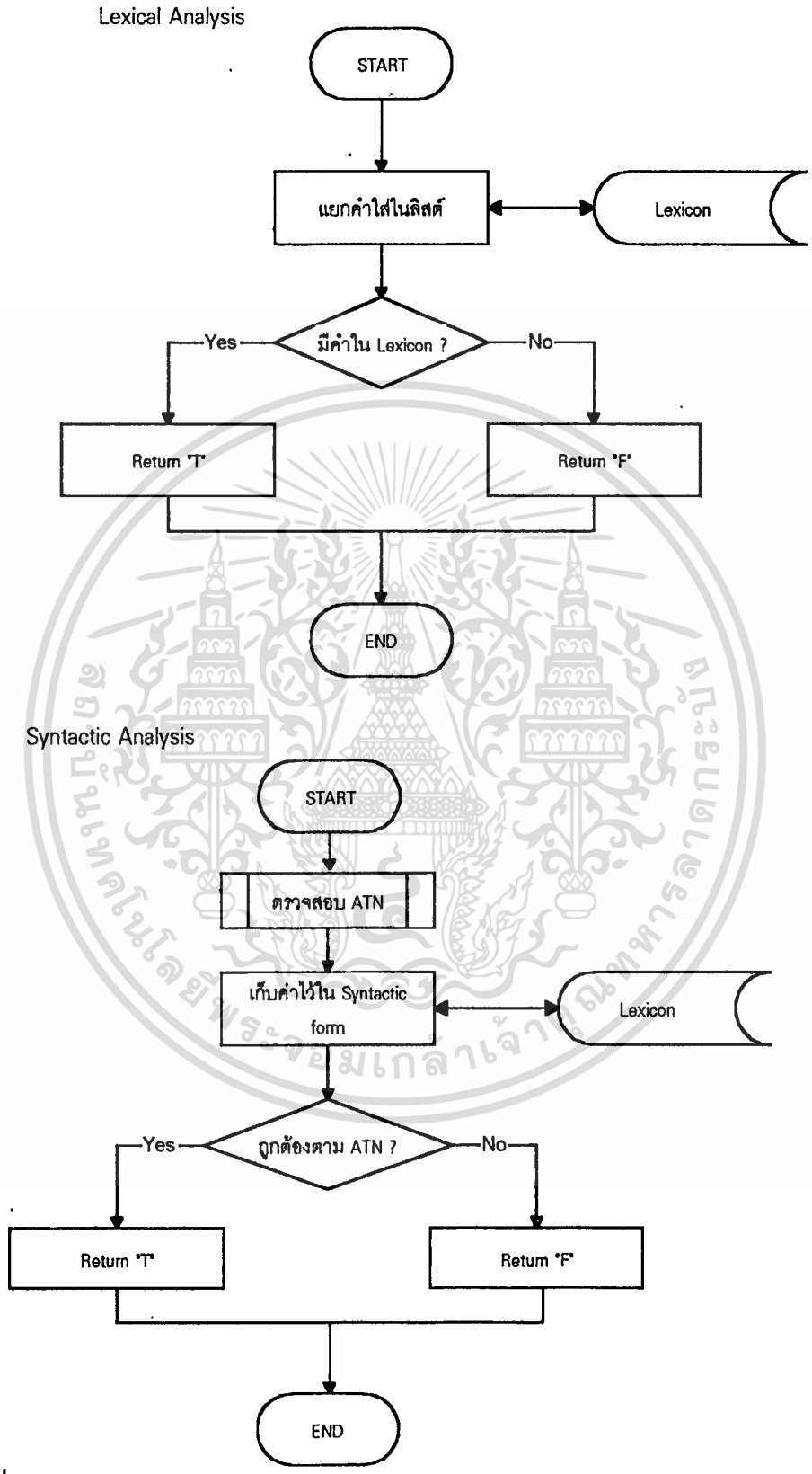
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และมี Flow Chart แสดงการทำงานแต่ละขั้นตอนดังนี้



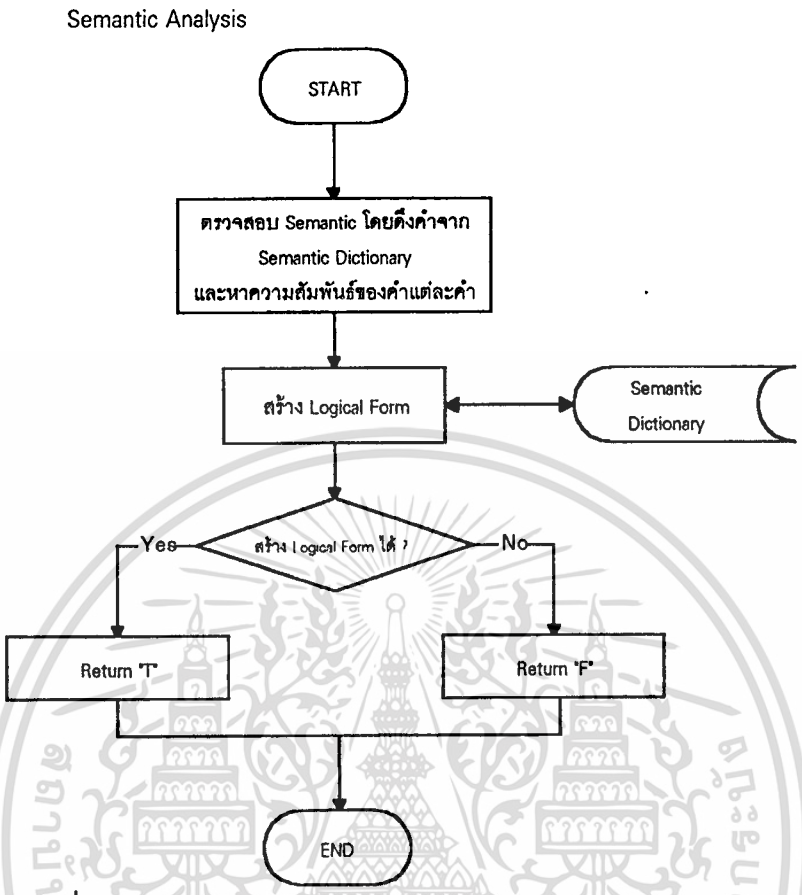
รูปที่ 3.4 แสดง Flow Chart การทำงานหลัก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.5 Flow Chart **แสดงการทำงานของส่วน** Lexical Analysis **และ** Syntactic Analysis

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.6 Flow Chart แสดงการทำงานของส่วน Semantic Analysis

โครงสร้างของ lexicon เป็นดังนี้

ชื่อของคำศัพท์ เก็บเป็นข้อมูลชนิด character ขนาด 20 ตัวอักษร

ชนิดของคำศัพท์ เก็บเป็นข้อมูลชนิด integer ดังนี้

type {art,adj,noun,pro,aux,verb,wh_pro,wh_det,prep,pp_noun,wh_adv}

0 1 2 3 4 5 6 7 8 9 10

พจน์ของคำศัพท์ เก็บเป็นข้อมูลแบบบิทของ integer โดยมีบิทเรียงตามกันดังนี้

- **บิทบน 7-6 ไม่ใช้**
- **บิท 5 หมายถึง เป็นพหูพจน์บุรุษที่ 3**
- **บิท 4 หมายถึง เป็นพหูพจน์บุรุษที่ 2**
- **บิท 3 หมายถึง เป็นพหูพจน์บุรุษที่ 1**
- **บิท 2 หมายถึง เป็นเอกพจน์บุรุษที่ 3**
- **บิท 1 หมายถึง เป็นเอกพจน์บุรุษที่ 2**
- **บิท 0 หมายถึง เป็นเอกพจน์บุรุษที่ 1**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รากศัพท์ เก็บเป็นข้อมูลชนิด character ขนาด 20 ตัวอักษร

รูปแบบในกรณีของคำกิริยา เก็บเป็นข้อมูลแบบบิทชนิด integer โดยมีบิทเรียงตามลำดับดังนี้

- บิท 7-5 ไม่ใช้
- บิท 4 หมายถึง infinitive
- บิท 3 หมายถึง progressive
- บิท 2 หมายถึง participle
- บิท 1 หมายถึง past tense
- บิท 0 หมายถึง present tense

(LONG,1,63,LONG,0,{ TO-INF FOR-TO-INF THAT })

(TRAIN,2,4,TRAIN,0,{})

(IT,3,4,IT,0,{})

(DO,4,59,DO,17,{})

(GOES,5,4,GO,1,{})

(WHAT,6,63,WHAT,0,{})

(WHICH,7,63,WHICH,0,{})

(FROM,8,63,FROM,0,{})

(QUICKLY,9,63,QUICK,0,{})

(WHEN,10,63,WHEN,0,{})

โครงสร้างของซีแมนติคดิกชันนารี มีรูปแบบดังนี้

รากศัพท์ กระบวนการที่เกี่ยวข้อง

รากศัพท์ เก็บเป็นข้อมูลแบบ character ขนาด 20 ตัวอักษร

กระบวนการที่เกี่ยวข้องจะเก็บเป็นวิธีการเพื่อแมพไปยังโครงสร้างดังนี้

- เลขที่ขบวนรถไฟ เป็นข้อมูลแบบ integer
- ชื่อของขบวนรถไฟ เป็นข้อมูลแบบ character ขนาด 40 ตัวอักษร
- สถานีต้นทาง (From-Loc) เป็นข้อมูลแบบ character ขนาด 20 ตัวอักษร
- สถานีปลายทาง (To-Loc) เป็นข้อมูลแบบ character ขนาด 20 ตัวอักษร
- เวลาออกจากต้นทาง (From-Time) เป็นข้อมูลแบบ integer
- เวลาถึงปลายทาง (To-Time) เป็นข้อมูลแบบ integer
- วันที่เดินทาง (date) เป็นข้อมูลแบบ character ขนาด 4 ตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลักษณะของการเก็บข้อมูลในฐานข้อมูล

ไฟล์ข้อมูลหลัก (Header file) เป็นไฟล์ที่มีส่วนขยายเป็น .HDR เพื่อเก็บข้อมูลของจำนวนสถานี และจำนวนของขบวนรถไฟ รวมทั้งชื่อสถานีและรายละเอียดของขบวนรถไฟดังนี้

- จำนวนสถานี เก็บเป็นข้อมูลแบบ integer
- จำนวนของขบวนรถไฟ เก็บเป็นข้อมูลแบบ integer
- ลิสต์ของรายการชื่อสถานี เก็บเป็นข้อมูลแบบ character ขนาด 30 ตัวอักษร
- ลิสต์ของรายละเอียดขบวนรถไฟ เก็บเป็นข้อมูลแบบ structure

```
struct TR_NAME {
                int Tid;                    (Train ID#)
                char Tname[2*LOC_LEN];   (Train name)
              };
```

ไฟล์เวลาเดินรถ (Item file) เป็นไฟล์ที่มีส่วนขยายเป็น .ITM เพื่อเก็บเวลาเดินรถ เรียงตามลำดับของสถานีและขบวน โดยข้อมูลแต่ละตัวจะเก็บเป็น integer

ไฟล์เพิ่มเติม (Extension file) เป็นไฟล์ที่เก็บข้อมูลปลีกย่อยของการเดินรถ เช่น ไม่มีขบวนนี้ในวันเสาร์-อาทิตย์ เป็นต้น โดยข้อมูลแต่ละตัวเก็บเป็น structure ดังนี้

```
struct TR_EXT {
                int Tid;       (Train ID#)
                int Type;     (Train type)
              };
```

โดยที่ train type มีค่าดังนี้

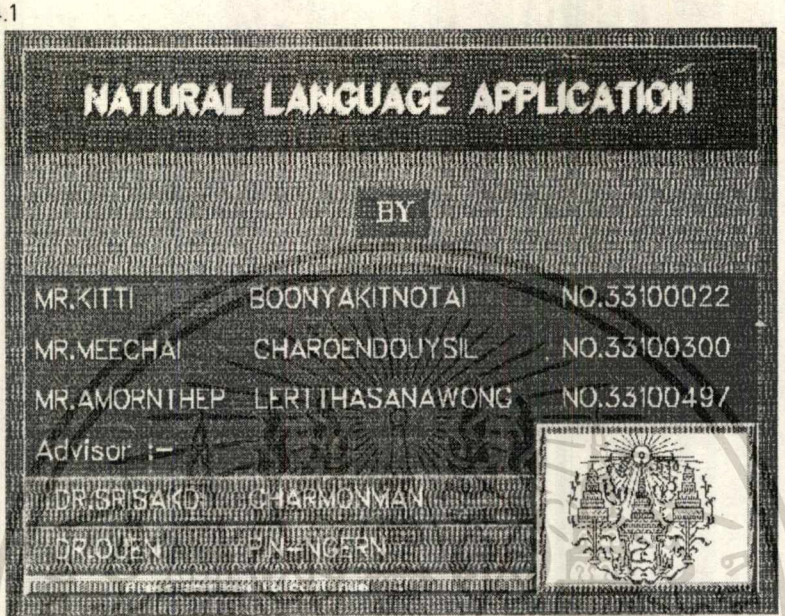
- 0 หมายถึง ขบวนรถไฟที่ไม่เดินทุกวันเสาร์-อาทิตย์ และวันหยุดนักขัตฤกษ์
- 1 หมายถึง ขบวนรถไฟที่มีเดินเฉพาะวันเสาร์-อาทิตย์ และวันหยุดนักขัตฤกษ์วันสุดท้าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

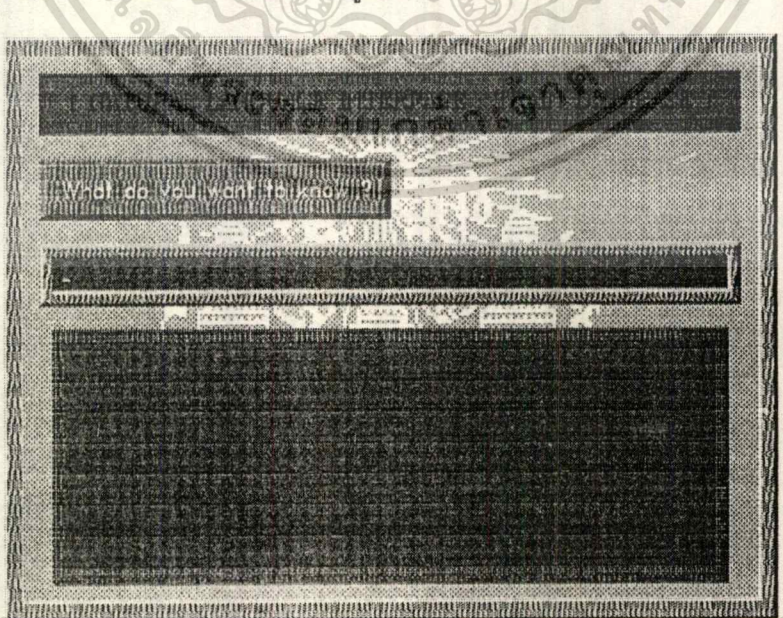
การทดลองและผลการทดลอง

เมื่อเรียกโปรแกรมขึ้นมาใช้งานจะมีหน้าจอแสดงชื่อปริญญานิพนธ์ ชื่อนักศึกษา ตลอดจนชื่ออาจารย์ที่ปรึกษาดังรูปที่ 4.1



รูปที่ 4.1 หน้าจอแรกของโปรแกรม

เมื่อกดปุ่มใด ๆ โปรแกรมก็จะเปลี่ยนหน้าจอไปเป็นหน้าจอรับประโยคภาษาอังกฤษ โดยประโยคหนึ่งจะมีได้ไม่เกิน 65 ตัวอักษร ซึ่งส่วนที่รับประโยคภาษาอังกฤษนั้นจะอยู่กลางจอภาพ ถัดจากส่วนรับข้อมูลลงมาก็จะเป็นพื้นที่ว่างสำหรับคำตอบของประโยคที่ป้อนเข้ามา ดังรูปที่ 4.2



รูปที่ 4.2 หน้าจอเตรียมพร้อมรับประโยค

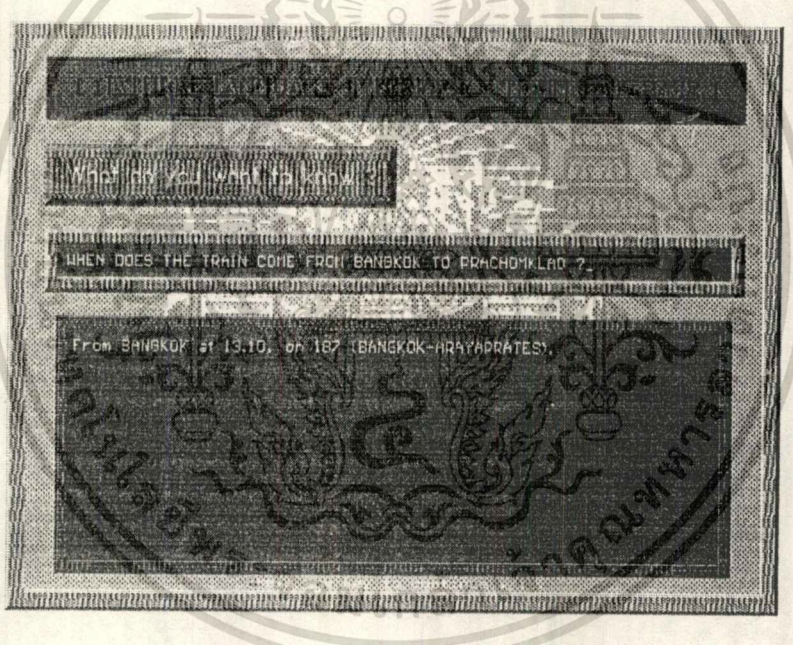
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทดลองมีดังต่อไปนี้

1. ทดลองป้อนประโยค "WHEN DOES THE TRAIN COME FROM BANGKOK TO PRACHOMKLAO ?"
2. ทดลองป้อนประโยค "WHEN DOES THE PLANE COME FROM BANGKOK TO PRACHOMKLAO ?"
3. ทดลองป้อนประโยค "WHEN THE TRAIN COME FROM BANGKOK TO PRACHOMKLAO ?"

ผลการทดลอง

1. จากการทดลองข้อ 1 โปรแกรมจะตอบว่า "From BANGKOK at 13:10, on 187 (BANGKOK-ARANYAPRATES) " ดังรูปที่ 4.3 เนื่องจากประโยคนี้ถูกหลักไวยากรณ์และถูกความหมาย คือต้องการจะรู้เวลาที่รถไฟออกจากสถานีกรุงเทพ ฯ ไปสถานีพระจอมเกล้า รวมทั้งคำศัพท์ทุกคำก็มีอยู่ในดิกชันนารี ดังนั้นโปรแกรมจึงสามารถวิเคราะห์และเข้าใจคำถามและค้นหาคำตอบโดยคิดจากเวลาที่ผู้ใช้ขณะนั้นถาม ได้อย่างถูกต้อง



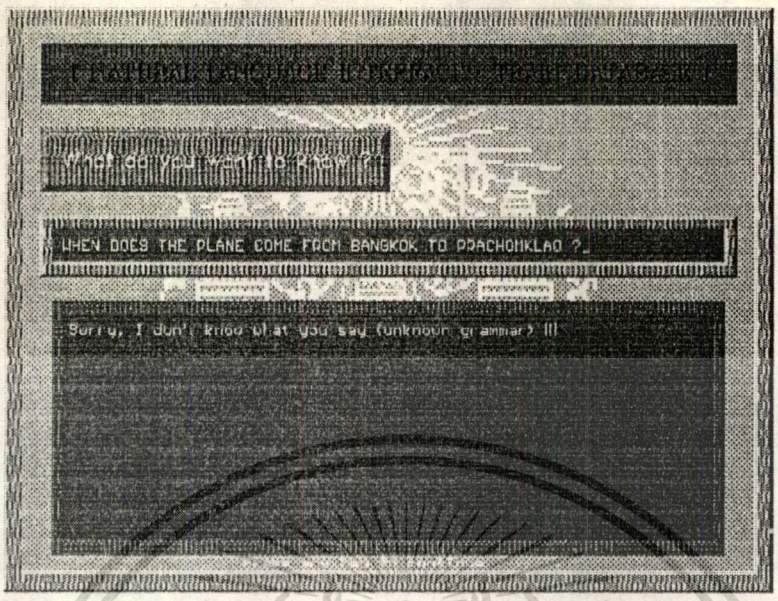
รูปที่ 4.3 ผลของการทดลองข้อที่ 1

2. จากการทดลองข้อ 2 จะเห็นว่าประโยคนั้นมีคำศัพท์ที่ไม่ได้อยู่ในขอบเขตที่กำหนดไว้ในดิกชันนารีคือ "PLANE" ดังนั้นโปรแกรมจึงตอบว่า

"Sorry, I don't know what you say (unknow grammar) !!!"

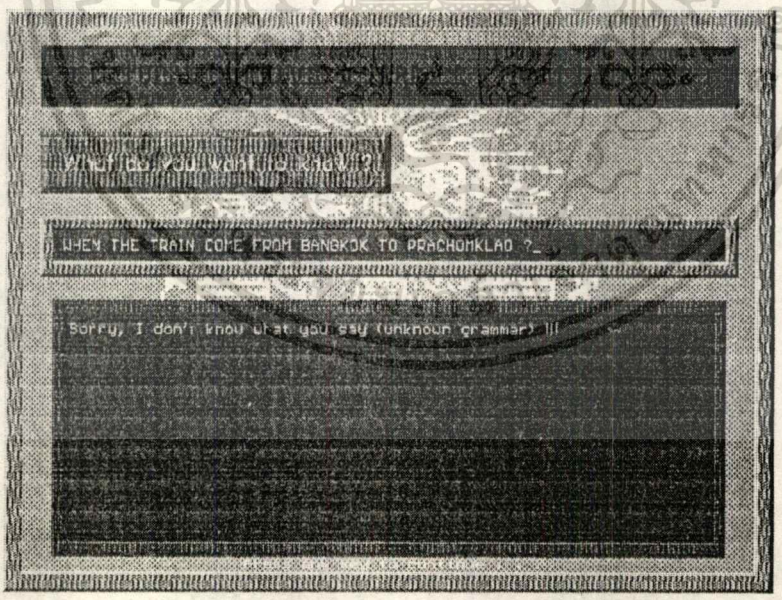
ผลการทดลองแสดงดังรูปที่ 4.4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.4 ผลของการทดลองข้อที่ 2

3. จากการทดลองข้อ 3 จะเห็นว่าประโยคที่ป้อนผิดตามหลักไวยากรณ์ คือไม่มีคำว่า "DOES" นำหน้าคำว่า "THE TRAIN" ซึ่งทำให้ประโยคขาดส่วนของกิริยาช่วยไป ดังนั้นประโยคนี้จึงผิดในขั้นตอนของการตรวจหลักไวยากรณ์ (Syntactic Analysis) ผลการทดลองแสดงดังรูปที่ 4.5



รูปที่ 4.5 ผลของการทดลองข้อที่ 3

นอกจากตัวอย่างดังกล่าวนี้แล้วโปรแกรมยังสามารถรับประโยคภาษาอังกฤษอื่น ๆ ได้อีก ทั้งนี้ประโยคที่ป้อนต้องอยู่ในขอบเขตของคำศัพท์และไวยากรณ์ที่กำหนดไว้ด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
บทวิจารณ์และสรุป

บทวิจารณ์

จากการทดลองใช้งานโปรแกรมที่สร้างขึ้น ผลปรากฏว่า สามารถใช้งานได้น่าพอใจระดับหนึ่ง ซึ่งโปรแกรมสามารถโต้ตอบกับผู้ใช้ในเรื่องเกี่ยวกับตารางเวลารถไฟได้ดีพอสมควร อาจจะมีติดขัดอยู่บ้างตรงที่ ผู้ทดสอบซึ่งเป็นคนไทยไม่คล่องไวยากรณ์ภาษาอังกฤษมากนักและคำศัพท์ที่ใช้ในโปรแกรมก็ยังมีอยู่น้อย

ปัญหาที่พบ

1. หลังจากที่ได้ออกแบบ ATN เป็นที่เรียบร้อยแล้ว เราได้เริ่มเขียนโปรแกรม โดยที่ได้รับคำแนะนำจากอาจารย์ที่ปรึกษาว่า ควรใช้ภาษาโปรล็อกในการอิมพลีเม้นต์ และจากการศึกษาภาษาโปรล็อก ก็เห็นว่าเป็นภาษาที่เหมาะสมกับการเขียนโปรแกรมทางด้านนี้ จึงตกลงเริ่มทำการอิมพลีเม้นต์ด้วยภาษาโปรล็อกในเทอมแรก โดยทำในส่วนของ Lexical Analysis และ Syntactic Analysis จนสำเร็จตามขอบเขตที่กำหนดไว้ จากนั้นในเทอมที่ 2 จึงได้พัฒนาต่อในส่วนของ Semantic Analysis ผลปรากฏว่า เกิดปัญหาหน่วยความจำไม่พอ และปัญหานี้เป็นข้อจำกัดของคอมไพเลอร์ของภาษาโปรล็อก จึงไม่สามารถที่จะพัฒนาโปรแกรมต่อไปได้ ดังนั้นเราจึงตัดสินใจใช้ภาษา C++ ในการอิมพลีเม้นต์แทนทั้งหมด โดยได้รับความเห็นชอบจากอาจารย์ที่ปรึกษา

2. เนื่องจากการหันมาใช้ภาษา C++ ในการอิมพลีเม้นต์แทน เราจึงต้องเขียนโค้ดเพื่อทำงานส่วนของการ backtrack และการจัดการเกี่ยวกับลิสล์ซึ่งในภาษาโปรล็อกทำได้ง่ายกว่า

บทสรุป

การเชื่อมต่อข้อมูลโดยใช้ภาษาธรรมชาติ เป็นการศึกษาโดยนำภาษาธรรมชาติมาใช้กับคอมพิวเตอร์เพื่อให้สามารถไปดึงข้อมูลจากฐานข้อมูลตามที่ผู้ใช้ต้องการ กระบวนการทำงานจะเริ่มต้นจากการที่ผู้ใช้ป้อนประโยคภาษาอังกฤษที่ถูกต้องตามหลักไวยากรณ์ที่กำหนดไว้ จากนั้นประโยคจะถูกส่งไปผ่านกระบวนการ 4 ขั้นตอนคือการแยกคำ (Lexical Analysis), การตรวจไวยากรณ์ (Syntactic Analysis), การวิเคราะห์ความหมาย (Semantic Analysis) และ การสร้างโค้ดเพื่อไปดึงข้อมูลจากฐานข้อมูล (Code Generation) แล้วตอบสนองไปยังผู้ใช้ตามความหมายของประโยคอินพุทที่ผู้ใช้ต้องการ

จากการศึกษาและดำเนินงานคาดว่าการนำภาษาธรรมชาติมาประยุกต์ใช้งานกับคอมพิวเตอร์อาจจะมีความเหมาะสมสำหรับบุคคลทั่วไปที่ไม่มีความรู้ทางการใช้คอมพิวเตอร์เลย สามารถที่จะโต้ตอบกับคอมพิวเตอร์ได้โดยง่าย เนื่องจากผู้ใช้เพียงแต่มีความสามารถในการป้อน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประโยคภาษาธรรมชาติที่ใช้กันอยู่ในชีวิตประจำวันให้ถูกต้องตามหลักของภาษาเท่านั้น โปรแกรมก็จะสามารถตอบสนองสิ่งที่ผู้ใช้ต้องการตามความสามารถที่มีอยู่นั้นได้

ปัจจุบันการพัฒนาโดยการใช้ภาษาไทยทางด้านนี้ยังไม่ได้รับความสนใจเท่าที่ควร ซึ่งจะพบเห็นได้น้อยมากจนแทบจะไม่มีเลย ทั้งนี้เพราะว่าภาษาไทยเราจะเขียนติดกันในแต่ละคำ การที่จะทำให้คอมพิวเตอร์รู้และเข้าใจจนสามารถแยกแต่คำออกมานั้นทำได้ยากมาก ซึ่งปัจจุบันยังไม่มีอัลกอริทึมที่สามารถทำได้สมบูรณ์แบบจนเป็นที่ยอมรับได้

จากที่กล่าวมาจะเห็นได้ว่าการนำการประมวลผลภาษาธรรมชาติมาประยุกต์ใช้งานจึงมีประโยชน์ควรศึกษาเป็นอย่างยิ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ขอขอบคุณ ดร.เอื้อน ปิ่นเงิน ที่ได้ดูแลและให้คำแนะนำระหว่างการทำปริญญานิพนธ์มาด้วยดี และขอขอบคุณ ศ.ดร.ศรีศักดิ์ จามรมาน ที่กรุณามาเป็นที่ปรึกษาร่วมในการทำปริญญานิพนธ์นี้

ขอขอบคุณพี่ปู (นภัทร) และเพื่อน ๆ ร่วมภาคและเพื่อนร่วมสถาบันที่เคยมาให้กำลังใจและเสนอแนะในบางส่วน ท้ายที่สุดนี้ขอขอบคุณ บิดา มารดา ที่ให้ความอุปการะและส่งเสริมการศึกษามาโดยตลอด ขอบคุณครับ

นาย กิตติ บุญยกิจโณทัย
นาย มีชัย เจริญด้วยศีล
นาย อมรเทพ เลิศทัศนวงศ์
16 มีนาคม 2537

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หนังสืออ้างอิง

1. CHRISTOPHER F. CHABRIS, "Artificial Intelligence & Turbo C", Dow Jones-Irwin., 385 p., 1991
2. DANIEL H. MARCELLUS, "Expert Systems Programming in Turbo Prolog", Prentice-Hall., 227 p., 1989
3. HARRY TENNANT, "Natural Language Processing", Petrocelli Books , 275 p., 1981
4. JAMES. ALLEN, "Natural Language Understanding", The Benjamin / Cummings Publishing, ,1989

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้