ระดับของภาษายอมรับได้-เค และความสามารถในการเรียนรู้

นายอนุชิต จิตพัฒนกุล

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรดุษฎีบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2553

A CLASS OF K-ACCEPTABLE LANGUAGES AND ITS LEARNABILITY

Mr. Anuchit Jitpattanakul

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Computer Engineering
Department of Computer Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2010

หัวข้อวิทยานิพนธ์	ระดับของภาษายอมรับได้-เค และความสามารถในการเรียนรู้
โดย	นายอนุชิต จิตพัฒนกุล
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร.อรรถสิทธิ์ สุรฤกษ์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาดุษฎีบัณฑิต

…………………………………………….. คณบดีคณะวิศวกรรมศาสตร์
(รองศาสตราจารย์ ดร.บุญสม เลิศหิรัญวงศ์)

คณะกรรมการสอบวิทยานิพนธ์

……………………………………………ประธานกรรมการ
(ศาสตราจารย์ ดร.ประภาส จงสถิตย์วัฒนา)

…………………………………………….. อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร.อรรถสิทธิ์ สุรฤกษ์)

…………………………………………… กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.เศรษฐา ปานงาม)

…………………………………………… กรรมการภายนอกมหาวิทยาลัย
(รองศาสตราจารย์ ดร.อุทุมพร พลาวงศ์)

…………………………………………… กรรมการภายนอกมหาวิทยาลัย
(ผู้ช่วยศาสตราจารย์ ดร.อานนท์ รุ่งสว่าง)

อนุชิต จิตพัฒนกุล : ระดับของภาษายอมรับได้-เค และความสามารถในการเรียนรู้ (A CLASS OF K-ACCEPTABLE AND ITS LEARNABILITY) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ. ดร.อรรถสิทธิ์ สุรฤกษ์, 59 หน้า

วิทยานิพนธ์ฉบับนี้ได้ศึกษาระดับของภาษารูปนัยที่เรียกว่าภาษายอมรับได้-เค และความสามารถในการเรียนรู้ของระดับของภาษานี้ บนแบบจำลองการเรียนรู้เชิงตัวอย่างที่เรียกว่าการระบุภาษาได้ในขอบเขตจำกัด วิทยานิพนธ์นี้ได้ทำการศึกษาความสามารถในการเรียนรู้ในรูปแบบการนำเสนอตัวอย่างที่แตกต่างกัน 2 รูปแบบ คือ การนำเสนอด้วยตัวอย่างบวกเพียงอย่างเดียว และการนำเสนอด้วยตัวอย่างบวกและตัวอย่างลบ ผลจากการศึกษาเชิงทฤษฎีแสดงให้เห็นว่าระดับภาษายอมรับได้-เคไม่สามารถเรียนรู้ได้ในขอบเขตจำกัดในกรณีที่นำเสนอด้วยตัวอย่างบวกเพียงอย่างเดียว แต่สำหรับในกรณีที่การนำเสนอมีทั้งตัวอย่างบวกและตัวอย่างลบ ระดับภาษายอมรับได้-เคสามารถเรียนรู้ได้ในขอบเขตจำกัด นอกจากนี้งานวิจัยนี้ยังได้ศึกษาถึงประสิทธิภาพของการเรียนรู้อีกด้วย ผลการวิจัยพบว่าระดับของภาษายอมรับได้-เคสามารถเรียนรู้ได้อย่างมีประสิทธิภาพจากเวลาและจำนวนตัวอย่างเชิงพหุนามในกรณีที่การนำเสนอมีทั้งตัวอย่างบวกและตัวอย่างลบ

ภาควิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อนิสิต ....

สาขาวิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก ....

ปีการศึกษา 2553

4771871721 : MAJOR COMPUTER ENGINEERING
KEYWORDS : K-ACCEPTABLE LANGUAGES / LEARNABILITY / K-EDGE DETERMINISTIC FINITE AUTOMATA

ANUCHIT JITPATTANAKUL : A CLASS OF K-ACCEPTABLE LANGUAGES AND ITS LEARNABILITY. ADVISOR : ASST. PROF. ATHASIT SURARERKS, Ph.D., 59 pp.

This thesis studies a class of formal languages called *k*-acceptable languages and its learnability on an explanatory learning model. This model is well known identification in the limit. Two different types of presentation of language information have been investigated to the learnability of this class of languages. First type is presentation with only positive examples. Second type is presentation with both positive and negative examples. The result theoretically shows that the class of *k*-acceptable languages is not learnable in the limit from only positive examples. Conversely, the class of *k*-acceptable languages is learnable in the limit from both positive and negative examples. In addition, the issue of efficiency of learning has been considered in term of learning time and characteristic examples used in the process of learning. Our result shows that the class of *k*-acceptable languages is learnable from polynomial time and data by using both positive and negative examples.

Department: Computer Engineering Student's Signature...

Field of Study: Computer Engineering Advisor's Signature ....

Academic Year: 2010

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความกรุณาจากอาจารย์ที่ปรึกษา ผศ. ดร. อรรถสิทธิ์ สุรฤกษ์ ที่ได้ให้คำแนะนำ ให้แนวทางการแก้ปัญหา แนะนำวิธีคิด เป็นประสบการณ์อันมีค่ายิ่ง หลายสิ่งหลายอย่างที่ได้เรียนรู้จากอาจารย์ทำให้ผู้วิจัยมีวันนี้ได้

ขอขอบคุณกรรมสอบวิทยานิพนธ์ทุกท่าน ได้แก่ ศ. ดร.ประภาส จงสถิตย์วัฒนา ผศ. ดร.เศรษฐา ปานงาม ผศ. ดร.อานนท์ รุ่งสว่าง และ รศ. ดร.อุทุมพร พลาวงศ์ ในการตรวจแก้และให้คำแนะนำอันเป็นประโยชน์ยิ่งต่องานวิจัย

ขอขอบคุณ Prof. Colin de la Higuera สำหรับโอกาสอันดีที่ได้ไปศึกษาวิจัยอยู่ที่ Hubert Curien Laboratoire, University of Saint-Etienne ขอบคุณสำหรับคำแนะนำ และความช่วยเหลือทุก ๆ อย่าง ทั้งขณะที่อยู่ที่ฝรั่งเศสและในเวลาต่อมา

ตลอดระยะเวลาการศึกษา ผู้วิจัยได้รับทุนสนับสนุนจากสำนักงานคณะกรรมการการอุดมศึกษา ทุนพัฒนาอาจารย์สาขาขาดแคลน (คณิตศาสตร์) ตามความต้องการของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ จึงขอขอบคุณมา ณ ที่นี้

วิทยานิพนธ์นี้สำเร็จได้จากความช่วยเหลือของบุคคลหลายท่านที่คอยให้ความช่วยเหลือ ให้คำแนะนำ ข้อคิดเห็นต่างๆ และให้กำลังใจด้วยดีตลอดมา ขอขอบคุณ พี่ ๆ เพื่อน ๆ น้อง ๆ สมาชิกทุกคน ใน ELITE laboratory ตั้งแต่ปี 2547-2554 (สำหรับความเป็นกันเอง กำลังใจ และความช่วยเหลือต่าง ๆ) ขอขอบคุณ เพื่อน ๆ ระดับปริญญาเอกทุกคน (สำหรับความคิดเห็นที่มีประโยชน์และกำลังใจที่ช่วยผลักดันให้ฝ่าฟันอุปสรรคมาได้) ขอขอบคุณ พี่สุบรรณารักษ์ที่น่ารักของภาควิชา (สำหรับความช่วยเหลือด้านเอกสารวิชาการ) ขอขอบคุณ พี่ ๆ เพื่อน ๆ น้อง ๆ ทุกคนที่ไม่สามารถเอ่ยชื่อในที่นี้ได้หมด สำหรับน้ำใจที่มีให้ตลอดมา

สุดท้ายนี้ ขอกราบขอบพระคุณ คุณพ่อ และญาติพี่น้องทุกท่านที่คอยเป็นห่วงเป็นใย ให้กำลังใจ และให้การสนับสนุนในทุกด้านตลอดมา โดยเฉพาะอย่างยิ่งขอขอบคุณ คุณจันทรัตน์ จิตพัฒนกุล ภรรยาที่แสนดี สำหรับกำลังใจที่มีให้ตลอดเวลา

สารบัญ

หน้า

บทคัดย่อภาษาไทย .... ง

บทคัดย่อภาษาอังกฤษ .... จ

กิตติกรรมประกาศ .... ฉ

สารบัญ .... ช

สารบัญภาพ .... ฌ

บทที่

1 บทนำ 1

1.1 ความเป็นมาและความสำคัญของปัญหา 3

1.2 วัตถุประสงค์ของการวิจัย 3

1.3 ขอบเขตของการวิจัย 3

1.4 ขั้นตอนและวิธีดำเนินการวิจัย 3

1.5 ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย 3

1.6 เนื้อหาในวิทยานิพนธ์ 4

1.7 งานตีพิมพ์ 4

2 งานวิจัยที่เกี่ยวข้อง 5

2.1 งานวิจัยที่ศึกษาแบบจำลองการเรียนรู้ภาษา 5

2.2 งานวิจัยที่ศึกษาความสามารถในการเรียนรู้จากตัวอย่างบวกเพียงอย่างเดียว 7

2.3 งานวิจัยที่ศึกษาความสามารถในการเรียนรู้จากตัวอย่างบวกและตัวอย่างลบ 7

2.4 สรุป 9

3 ทฤษฎีพื้นฐานที่ใช้ในงานวิจัย 10

3.1 ภาษารูปนัยและตัวแทนเชิงไวยากรณ์ 10

3.2 การเรียนรู้ภาษา 17

3.3 ความสามารถในการเรียนรู้ภาษา 19

3.3.1 การระบุภาษาได้ในขอบเขตจำกัด 19

3.3.2 การระบุภาษาได้จากข้อมูลที่กำหนดให้ 24

3.3.3 การระบุภาษาได้ในขอบเขตจำกัดจากเวลาและข้อมูลเชิงพหุนาม 26

3.4 สรุป 27

บทที่ หน้า

4 ระดับภาษายอมรับได้-เค 28

4.1 นิยามของภาษายอมรับได้-เค 28

4.2 คุณสมบัติของระดับภาษายอมรับได้-เค 31

4.3 สรุป 35

5 ความสามารถในการเรียนรู้ของระดับภาษายอมรับได้-เค 36

5.1 การเรียนรู้จากตัวอย่างบวกเพียงอย่างเดียว 36

5.2 การเรียนรู้จากตัวอย่างบวกและตัวอย่างลบ 37

5.2.1 อัลกอริทึมการเรียนรู้ภาษายอมรับได้-เค 37

5.2.2 ตัวอย่างลักษณะสำหรับการเรียนรู้ภาษายอมรับได้-เค 45

5.3 สรุป 52

6 สรุปผลการวิจัยและข้อเสนอแนะ 53

รายการอ้างอิง 55

ประวัติผู้เขียนวิทยานิพนธ์ 59

สารบัญภาพ

รูปที่ หน้า

3.1 ลำดับชั้นของภาษารูปนัย 8
3.2 แผนภาพสถานะของออโตมาตาจำกัดเชิงกำหนด 9
3.3 แผนภาพสถานะของ M_{APTA} 11
3.4 กระบวนการเรียนรู้ภาษาจากการอนุมานตัวแทนเชิงไวยากรณ์ของภาษา 12
3.5 แบบจำลองการระบุภาษาได้ในขอบเขตจำกัด 15
3.6 แบบจำลองการระบุภาษาได้ในขอบเขตจำกัดจากเวลาและข้อมูลเชิงพหุนาม 20
4.1 แผนภาพการผ่านของ M_2 27
4.2 แผนภาพการผ่านของ PTA_2 29
5.1 ต้นไม้ยอมรับส่วนหน้าขยายขอบ-เค 40
5.2 การทำงานของอัลกอริทึมเคเอซีแอลไอ 41

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ทฤษฎีภาษารูปนัย (formal language theory) เริ่มต้นขึ้นในช่วงปี ค.ศ. 1950-1960 จากงานวิจัยของนักภาษาศาสตร์ที่ชื่อว่า นอม ชอมสกี ซึ่งพยายามที่จะอธิบายการเรียนรู้ภาษาของมนุษย์ เขาได้เสนอแนวคิดของการมีอยู่ของไวยากรณ์สากล (universal grammars) ในสมองของมนุษย์ และอธิบายกระบวนการเรียนรู้ภาษาของมนุษย์ว่า เป็นกระบวนการกำหนดรูปแบบของกฎเชิงวากยสัมพันธ์ (syntactic rules) ในไวยากรณ์สากลตามข้อมูลของภาษาที่ได้รับมา ซึ่งภาษาแต่ละภาษามีรูปแบบของกฎเชิงวากยสัมพันธ์แตกต่างกัน และรูปแบบของกฎในไวยากรณ์ยังสามารถนำมาแบ่งความซับซ้อนของภาษาได้อีกด้วย โดยชอมสกีได้แบ่งภาษาออกเป็นทั้งหมด 4 ระดับในรูปแบบของลำดับชั้นของระดับภาษาที่สัมพันธ์กัน ก็คือ ระดับภาษาสม่ำเสมอ (regular languages) ระดับภาษาไม่พึ่งบริบท (context-free language) ระดับภาษาพึ่งบริบท (context-sensitive language) ระดับภาษาแจงนับได้แบบวนซ้ำ (recursively enumerable language) แต่อย่างไรก็ตาม ในช่วงเวลานั้นความสามารถของการเรียนรู้ (learnability) ของระดับภาษาต่าง ๆ ยังไม่ได้ถูกศึกษา เนื่องจากในช่วงนั้นยังไม่มีแบบจำลองใดที่สามารถอธิบายกระบวนการเรียนรู้ภาษาได้

การศึกษาเรื่องความสามารถการเรียนรู้ภาษาได้รับความสนใจเป็นอย่างมาก ตั้งแต่ปี ค.ศ. 1967 เมื่อนักวิจัยด้านภาษาศาสตร์ชื่อ โกลด์ ได้เสนอแบบจำลองทางคณิตศาสตร์ที่สามารถอธิบายกระบวนการเรียนรู้ภาษาได้ เพื่อใช้ในการศึกษาความสามารถในการเรียนรู้ของระดับภาษาต่าง ๆ ในลำดับชั้นที่ชอมสกีได้เสนอไว้ แบบจำลองการเรียนรู้ภาษานี้รู้จักกันดีในชื่อของ การระบุภาษาได้ในขอบเขตจำกัด (identification in the limit) โดยเป็นการมองว่าการเรียนรู้ภาษาเป็นกระบวนการหาตัวแทนเชิงไวยากรณ์ (grammatical representation) ของภาษา และเป็นกระบวนการที่ไม่สิ้นสุด (infinite process) ซึ่งการที่จะกล่าวว่าภาษาใด ๆ สามารถเรียนรู้ได้นั้น จะต้องพิจารณาจากการมีอยู่ของอัลกอริทึมการเรียนรู้ (learning algorithm) ที่สามารถระบุไวยากรณ์ของภาษานั้น ๆ ได้ในขอบเขตที่จำกัด โกลด์ได้ใช้แบบจำลองนี้ศึกษาความสามารถในการเรียนรู้ของระดับภาษาต่าง ๆ ในลำดับชั้นชอมสกีในเชิงทฤษฎี และได้เสนอทฤษฎีบทที่มีอิทธิพลต่องานวิจัยด้านการเรียนรู้ภาษาเป็นอย่างมาก 2 ทฤษฎีบท คือ ทฤษฎีบทเชิงบวกที่กล่าวว่าระดับภาษาสม่ำเสมอ ระดับภาษาไม่พึ่งบริบท และระดับภาษาพึ่งบริบท สามารถเรียนรู้ได้จากตัวอย่างบวก (ประโยคหรือสายอักขระที่อยู่ในภาษา) และตัวอย่างลบ (ประโยคหรือสายอักขระที่ไม่อยู่ในภาษา) และทฤษฎีบทเชิงลบที่กล่าวว่า ไม่มีระดับภาษาใดเลยในลำดับชั้นชอมสกีสามารถเรียนรู้ได้จากตัวอย่างบวกเพียงอย่างเดียว ด้วย

ทฤษฎีบทเชิงลบนี้ทำให้นักวิจัยด้านการเรียนรู้ หันไปสนใจศึกษาประเด็นอื่น ๆ จนทำให้ช่วงปี ค.ศ. 1970-1980 ไม่มีงานวิจัยทางด้านการเรียนรู้ภาษามากนัก

การศึกษาเรื่องความสามารถการเรียนรู้ภาษากลับมาได้รับความสนใจอีกครั้งในช่วงต้นของปี ค.ศ. 1980 เมื่อนักวิจัยชื่อ อังกลูอิน ได้แสดงให้เห็นว่า มีระดับย่อยของระดับภาษาสม่ำเสมอที่สามารถเรียนรู้ได้จากตัวอย่างบวกเพียงอย่างเดียว โดยระดับภาษานั้นเรียกว่า ระดับภาษาย้อนกลับได้ (reversible languages) ซึ่งหลังจากนั้น มีระดับภาษาอีกเป็นจำนวนมากได้ถูกพิสูจน์ว่าสามารถเรียนรู้ได้จากตัวอย่างบวกเพียงอย่างเดียว เช่น ระดับภาษาทดสอบได้ (testable languages) ระดับภาษาสม่ำเสมอที่เข้ารหัส (code regular languages) ระดับภาษาสม่ำเสมอแบบเข้ม (strictly regular languages) เป็นต้น โดยระดับภาษาเหล่านี้ส่วนใหญ่เป็นระดับย่อยของภาษาสม่ำเสมอ ที่นิยามขึ้นมาใหม่ตามตัวแทนเชิงไวยากรณ์ในรูปแบบของออโตมาตาชนิดต่าง ๆ

การนำผลที่ได้จากงานวิจัยด้านการเรียนรู้ภาษา ไปประยุกต์ใช้กับการแก้ปัญหาด้านการรู้จำ (recognition problems) เป็นประเด็นหนึ่งที่น่าสนใจและท้าทายเป็นอย่างมาก เนื่องจากนักวิจัยส่วนใหญ่เชื่อว่า ปัญหาการรู้จำส่วนใหญ่มีความซับซ้อนอยู่ในระดับสม่ำเสมอ ดังนั้นนักวิจัยส่วนหนึ่งได้สนใจที่จะหาอัลกอริทึมการเรียนรู้ที่มีประสิทธิภาพของระดับภาษาสม่ำเสมอ รวมถึงระดับย่อยของภาษาสม่ำเสมอ เนื่องจากระดับภาษาสม่ำเสมอนี้ได้ถูกพิสูจน์แล้วว่าสามารถเรียนรู้ได้ในขอบเขตจำกัด นั่นคือมีอัลกอริทึมการเรียนรู้ที่มีประสิทธิภาพโดยใช้ออโตมาตาเป็นตัวแทนเชิงไวยากรณ์ของภาษา แต่อย่างไรก็ตาม ปัญหาหนึ่งที่เกิดจากการใช้ออโตมาตาเป็นตัวแทนเชิงไวยากรณ์ คือ ขนาดของออโตมาตาที่คำนวณได้จากอัลกอริทึมการเรียนรู้มีขนาดใหญ่ในกรณีที่ภาษาถูกนิยามด้วยสัญลักษณ์เป็นจำนวนมาก ซึ่งเป็นสาเหตุให้ไม่สามารถนำไปประยุกต์ใช้งานจริงได้

ในปี ค.ศ. 2006 นักวิจัยด้านการอนุมานเชิงไวยากรณ์ที่ชื่อว่า อิกูเอรา ได้นิยามออโตมาตาชนิดพิเศษขึ้นมาใหม่ ออโตมาตาตัวนี้เรียกว่า ออโตมาตาจำกัดเชิงกำหนดขอบ-เค (*k*-edge deterministic finite automata) ซึ่งนิยามบนเซตของสัญลักษณ์ที่มีอันดับ ออโตมาตาตัวนี้มีคุณสมบัติที่น่าสนใจคือ ขนาดของออโตมาตาไม่ขึ้นอยู่กับจำนวนสัญลักษณ์ที่ใช้นิยามภาษา แต่ขึ้นอยู่กับค่า เค ที่ระบุไว้ ซึ่งเหมาะสำหรับการรู้จำภาษาที่นิยามจากสัญลักษณ์ที่มีลำดับที่มีอยู่เป็นจำนวนมากในภาษาที่พบได้ทั่วไป เช่น ภาษาของทำนองเพลงที่ตัวโน้ตมีลำดับ เป็นต้น แต่จนถึงปัจจุบันยังไม่มีงานวิจัยใดศึกษาคุณสมบัติการเรียนรู้ภาษาของระดับภาษาที่รู้จำได้โดยออโตมาตาจำกัดเชิงกำหนดขอบ-เค

สำหรับงานวิจัยนี้ได้ศึกษาความสามารถการเรียนรู้ของระดับภาษาที่รู้จำได้โดยออโตมาตาจำกัดเชิงกำหนดขอบ-เค ซึ่งเราได้เรียกระดับภาษานี้ว่า ระดับภาษายอมรับได้-เค (*k*-acceptable languages) โดยทำการศึกษาบนแบบจำลองการเรียนรู้ของโกลด์ โดยศึกษาใน

สถานการณ์การเรียนรู้ที่แตกต่างกันสองแบบ คือ แบบแรกพิจารณาในกรณีที่มีตัวอย่างบวกเพียงอย่างเดียว และแบบที่สองพิจารณาในกรณีที่มีทั้งตัวอย่างบวกและตัวอย่างลบ โดยทำการศึกษาเชิงทฤษฎีเพื่อเสนอผลการศึกษาในรูปแบบของทฤษฎีบทของความสามารถการเรียนรู้ของระดับภาษายอมรับได้-เค ของทั้งสองรูปแบบของการนำเสนอข้อมูล และยืนยันผลการศึกษาโดยวิธีการพิสูจน์ทางคณิตศาสตร์ นอกจากนั้นในกรณีที่พิสูจน์แล้วว่าระดับยอมรับได้-เคสามารถเรียนรู้ได้ งานวิจัยนี้จะเสนออัลกอริทึมการเรียนรู้ที่มีประสิทธิภาพ โดยพิจารณาจากเวลาและขนาดของตัวอย่างที่ใช้ในการเรียนรู้

1.2 วัตถุประสงค์ของการวิจัย

เพื่อพิสูจน์ความสามารถในการเรียนรู้ของระดับภาษายอมรับได้-เค และเสนออัลกอริทึมการเรียนรู้ที่มีประสิทธิภาพ

1.3 ขอบเขตของการวิจัย

1. ศึกษาความสามารถในการเรียนรู้ของระดับภาษายอมรับได้-เค จากการนำเสนอข้อมูล 2 รูปแบบ คือ การนำเสนอด้วยตัวอย่างบวกเพียงอย่างเดียว และการนำเสนอทั้งตัวอย่างบวกและตัวอย่างลบ
2. ศึกษาการเรียนรู้ของระดับภาษายอมรับได้-เค ผ่านตัวแทนเชิงไวยากรณ์ของภาษาที่เรียกว่า ออโตมาตาจำกัดเชิงกำหนดขอบ-เค
3. ความสำเร็จของการเรียนรู้ภาษา พิจารณาจากเกณฑ์การเรียนรู้ในขอบเขตจำกัดของโกลด์
4. ประสิทธิภาพของอัลกอริทึมการเรียนรู้ พิจารณาจากเวลา และขนาดของตัวอย่างที่ใช้ในการเรียนรู้

1.4 ขั้นตอนและวิธีดำเนินการวิจัย

1. ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง
2. วิเคราะห์และออกแบบอัลกอริทึมการเรียนรู้
3. พัฒนาประสิทธิภาพของขั้นตอนการเรียนรู้
4. ตั้งทฤษฎีบทและพิสูจน์
5. วิเคราะห์และสรุปผล
6. จัดทำเอกสารงานวิจัย

1.5 ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. ได้ทฤษฎีบทของความสามารถในการเรียนรู้ภาษาในระดับของภาษายอมรับได้-เค

2. ได้อัลกอริทึมในการเรียนรู้ที่มีประสิทธิภาพสำหรับภาษายอมรับได้-เค ในกรณีที่พิสูจน์แล้วว่าเรียนรู้ได้

1.6 เนื้อหาในวิทยานิพนธ์

รายละเอียดต่าง ๆ ในวิทยานิพนธ์จะนำเสนอเป็นลำดับดังนี้ ในบทที่ 2 จะกล่าวถึงกล่าวถึงงานวิจัยที่เกี่ยวข้องกับการเรียนรู้ภาษา บทที่ 3 จะแสดงความรู้พื้นฐานที่ใช้ในงานวิจัย ได้แก่ นิยามและสัญลักษณ์ที่เกี่ยวข้อง แบบจำลองการเรียนรู้เชิงตัวอย่างของโกลด์ เกณฑ์วัดความสามารถการเรียนรู้ภาษาที่เรียกว่าการระบุภาษาได้ในขอบเขตจำกัด บทที่ 4 จะนำเสนอนิยามของภาษายอมรับได้-เค และออโตมาตาจำกัดเชิงกำหนดขอบ-เค และผลการศึกษาคุณสมบัติพื้นฐานของระดับภาษานี้ ส่วนบทที่ 5 จะแสดงผลการศึกษาความสามารถในการเรียนรู้ของระดับภาษายอมรับได้-เค โดยแบ่งการศึกษาออกเป็น 2 กรณี คือ ในกรณีที่นำเสนอด้วยตัวอย่างบวกเพียงอย่างเดียว และในกรณีที่นำเสนอด้วยตัวอย่างบวกและตัวอย่างลบ ซึ่งจะแสดงผลการศึกษาในรูปแบบของทฤษฎีบท และบทพิสูจน์ ส่วนสุดท้ายจะทำการสรุปผลการศึกษาวิจัยและแนวทางที่น่าจะศึกษาต่อจากงานวิจัยนี้

1.7 งานตีพิมพ์

ในระหว่างการศึกษาได้มีการตีพิมพ์ผลงานวิจัยดังนี้

วารสารวิชาการ

Anuchit Jitpattanakul, and Athasit Surarerks, “Characteristic Sets for Learning *k*-Acceptable Languages”, ECTI-CIT Journal. (Submitted)

รายงานการประชุมวิชาการ

Anuchit Jitpattanakul, and Athasit Surarerks, “Characteristic Sets for Learning *k*-Acceptable Languages”, In Proceeding of The 7th Annual International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI), Chiang Mai,Thailand, May 19-21 2010.

Anuchit Jitpattanakul, and Athasit Surarerks, “An Algorithm for Learning *k*-DFA from Informant”, In Proceeding of the 13th International Annual Symposium on Computational Science and Engineering: ANSCSE13, Kasetsart University, Bangkok, Thailand, March 25-27, 2009.

บทที่ 2

งานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงงานวิจัยที่ศึกษาเกี่ยวกับความสามารถในการเรียนรู้ภาษา โดยแบ่งออกเป็น 3 หัวข้อคือ งานวิจัยที่ศึกษาแบบจำลองการเรียนรู้ภาษา งานวิจัยที่ศึกษาความสามารถในการเรียนรู้ภาษาจากตัวอย่างบวกเพียงอย่างเดียว และงานวิจัยที่ศึกษาความสามารถในการเรียนรู้ภาษาจากตัวอย่างบวกและตัวอย่างลบ

2.1 งานวิจัยที่ศึกษาแบบจำลองการเรียนรู้ภาษา

จุดประสงค์ในช่วงแรกของการศึกษาความสามารถในการเรียนรู้ภาษา คือ ต้องการที่จะเข้าใจกระบวนการเรียนรู้ภาษาของมนุษย์ โดยเริ่มต้นในปี ค.ศ. 1956 นอม ชอกสกี [1] ได้เสนอแนวคิดของการมีอยู่ของไวยากรณ์สากล ทุกภาษามีระบบไวยากรณ์พื้นฐานที่เหมือนกัน แต่สิ่งที่ต่างกันคือ การตั้งค่าของระบบไวยากรณ์ให้สอดคล้องกับภาษา ดังนั้นกระบวนการการเรียนรู้ภาษา ก็คือ การกระบวนการหาระบบไวยากรณ์ที่ตั้งค่าเหมาะสำหรับภาษานั้น ๆ จากข้อมูลภาษาที่ได้รับมานั่นเอง นอกจากนี้เขายังได้เสนอการแบ่งลำดับความซับซ้อนของภาษาตามระบบไวยากรณ์ ที่เรียกว่า ลำดับชั้นของชอมสกี (Chomsky’s hierarchy) [2] โดยประกอบด้วยระดับภาษาสม่ำเสมอ ระดับภาษาไม่พึ่งบริบท ระดับภาษาพึ่งบริบท ระดับภาษาแจงนับได้แบบวนซ้ำ โดยภาษามนุษย์นั้นอยู่เหนือกว่าระดับภาษาไม่พึ่งบริบท [3] เนื่องจากไม่สามารถใช้ไวยากรณ์ไม่พึ่งบริบทบรรยายได้ แต่จะอยู่ในระดับชั้นใดนั้น จนถึงปัจจุบันยังไม่มีบทพิสูจน์ใดที่สามารถยืนยันได้

แบบจำลองทางคณิตศาสตร์เป็นสิ่งหนึ่งที่จำเป็นในการศึกษาวิจัยความสามารถการเรียนรู้ภาษา โดยแบบจำลองนี้ต้องสามารถที่จะอธิบายกระบวนการเรียนรู้ภาษา และมีกฏเกณฑ์ที่ใช้บอกความสำเร็จของการเรียนรู้ ตั้งแต่ปี ค.ศ. 1960 จนถึงปัจจุบัน งานวิจัยส่วนใหญ่ได้ทำการศึกษาการเรียนรู้ภาษาโดยใช้แบบจำลองการเรียนรู้สามแบบ [4-9] คือ แบบจำลองการเรียนรู้ภาษาเชิงตัวอย่าง (explanatory learning) ของโกลด์ [10] แบบจำลองการเรียนรู้ภาษาจากการซักถาม (active learning) ของอังกลูอิน [11] และการเรียนรู้ด้วยความถูกต้องโดยการประมาณที่น่าจะเป็นไปได้ (probably approximately correct learning) ที่เสนอโดยวาเลียน [12] ทั้งสามแบบจำลองนี้แตกต่างกันที่รูปแบบของการนำเสนอข้อมูลภาษา และเกณฑ์การพิจารณาความสำเร็จของการเรียนรู้ภาษา สำหรับวิทยานิพนธ์นี้ใช้แบบจำลองการเรียนรู้เชิงตัวอย่างของโกลด์เป็นพื้นฐานในงานวิจัย ดังนั้นในหัวข้อนี้จะกล่าวเฉพาะงานวิจัยที่เกี่ยวข้องกับแบบจำลองการเรียนรู้แบบนี้เท่านั้น

แบบจำลองการเรียนรู้เชิงตัวอย่างของโกลด์นั้นเป็นการเรียนรู้ในลักษณะที่ผู้เรียนรับข้อมูลภาษาได้ทางเดียว (passive learning) และเป็นกระบวนการที่ไม่สิ้นสุด (infinite process) ความสำเร็จของการเรียนรู้ของแบบจำลองนี้พิจารณาได้จากเกณฑ์ที่เรียกว่า "การระบุภาษาได้ในขอบเขตจำกัด" นั่นคือ การเรียนรู้ภาษาจะสำเร็จเมื่อผู้เรียนสามารถระบุตัวแทนเชิงไวยากรณ์ของภาษานั้นได้ในเวลาที่จำกัด โกลด์ได้ใช้แบบจำลองนี้ศึกษาความสามารถการเรียนรู้ของภาษาต่าง ๆ ในลำดับชั้นชอมสกี ซึ่งโกลด์ได้เสนอทฤษฎีบทที่มีผลกระทบต่อวงการวิจัยเป็นอย่างมาก นั่นคือ ทฤษฎีบทเชิงบวกของโกลด์ ซึ่งกล่าวว่า ระดับภาษาวนซ้ำสามารถเรียนรู้ได้ในขอบเขตจำกัดโดยใช้ตัวอย่างบวก (สายอักขระที่อยู่ในภาษา) และตัวอย่างลบ (สายอักขระที่ไม่อยู่ในภาษา) และทฤษฎีบทเชิงลบของโกลด์ ที่กล่าวว่า ทุกระดับภาษาในลำดับชั้นของชอมสกีไม่สามารถเรียนรู้ภาษาได้ในขอบเขตจำกัดโดยใช้ตัวอย่างบวกเพียงอย่างเดียว ต่อมาในปี ค.ศ. 1978 โกลด์ [13] ได้พัฒนาแบบจำลองการเรียนรู้ภาษาอีกแบบหนึ่งที่เรียกว่า การระบุภาษาได้จากข้อมูลที่ได้รับ (identification from given data) โดยพิจารณาความสำเร็จของการเรียนรู้จากข้อมูลที่จำกัด ในงานวิจัยนี้โกลด์ได้เน้นไปที่ระดับภาษาสม่ำเสมอ ซึ่งต่อมาในปี ค.ศ.1996 ดูปองท์ [14] ได้พิสูจน์ว่า แบบจำลองการเรียนรู้ของโกลด์ทั้งสองแบบนั้นสมมูลกันในทางทฤษฎี นั่นคือ ระดับภาษาใดสามารถเรียนรู้ได้ในขอบเขตจำกัด ก็ต่อเมื่อ ระดับภาษานั้นสามารถเรียนรู้ได้จากข้อมูลที่ได้รับ

ในการศึกษาวิจัยเรื่องการเรียนรู้ภาษานั้น ประสิทธิภาพในการเรียนรู้ภาษาเป็นประเด็นที่นักวิจัยด้านนี้สนใจ มีงานวิจัยหลายงานได้เสนอเกณฑ์ในการวัดประสิทธิภาพโดยพิจารณาจากปัจจัยที่แตกต่างกัน [15] เช่น เวลาที่ใช้ในการเรียนรู้เพียงอย่างเดียว จำนวนความผิดพลาดที่เกิดขึ้นในกระบวนการเรียนรู้ จำนวนของการเปลี่ยนแปลงรูปแบบของตัวแทนเชิงไวยากรณ์ เป็นต้น แต่เกณฑ์ในการวัดประสิทธิภาพที่เป็นที่ยอมรับกันในงานวิจัยด้านการอนุมานเชิงไวยากรณ์ คือ ขนาดของตัวอย่างลักษณะ (ตัวอย่างที่ทำให้การอนุมานไวยากรณ์สำเร็จ) ที่ใช้ในการเรียนรู้ และเวลาที่ผู้เรียนใช้ในการหาไวยากรณ์ภาษา ทั้งสองตัวชี้วัดนี้ต้องเป็นอยู่ในรูปแบบของฟังก์ชันพหุนาม แบบจำลองการเรียนรู้ภาษาที่รวมสองตัวชี้วัดนี้ก็คือ การระบุได้ในขอบเขตจำกัดจากเวลาและข้อมูลเชิงพหุนาม (identification in the limit from polynomial time and data) เสนอโดย อิกูเอรา [16] ในปี ค.ศ. 1997 ในงานนี้เขาได้ศึกษาการมีอยู่จริงของตัวอย่างลักษณะ และได้พิสูจน์ว่า ออโตมาตาจำกัดเชิงไม่กำหนด (non-deterministic finite automata) และ ไวยากรณ์ไม่พึ่งบริบท (context-free grammar) ไม่สามารถเรียนรู้ได้ในขอบเขตจำกัดด้วยเวลาและจำนวนข้อมูลเชิงพหุนาม แต่ออโตมาตาจำกัดเชิงกำหนดสามารถเรียนรู้ได้ในขอบเขตจำกัดด้วยเวลาและจำนวนข้อมูลเชิงพหุนาม แบบจำลองการระบุภาษาได้ในขอบเขตจำกัดจากเวลาและข้อมูลเชิงพหุนาม ได้ถูกใช้กันอย่างแพร่หลายในงานวิจัยด้านการเรียนรู้ภาษา และถูกอ้างอิงเป็นจำนวนมาก ในวิทยานิพนธ์นี้ใช้แบบจำลองดังกล่าวในการศึกษาประเด็นของประสิทธิภาพของการเรียนรู้ภาษา

2.2 งานวิจัยที่ศึกษาความสามารถในการเรียนรู้จากตัวอย่างบวกเพียงอย่างเดียว

จากทฤษฎีบทเชิงลบของโกลด์ทำให้นักวิจัยจำนวนหนึ่งตีความว่า ระดับของภาษามนุษย์ซึ่งเป็นระดับย่อยของระดับภาษาวนซ้ำก็ไม่สามารถเรียนรู้ได้ในขอบเขตจำกัดโดยใช้ตัวอย่างบวกเพียงอย่างเดียว การตีความลักษณะนี้ทำให้นักวิจัยเหล่านี้เปลี่ยนความสนใจไปศึกษาประเด็นที่อื่น ๆ จนทำให้ช่วงปี ค.ศ. 1970-1979 เป็นยุคมืดของการศึกษาด้านการเรียนรู้ภาษา [17] แต่ต่อมาในปี ค.ศ. 1982 อังกลูอิน [18] ได้พิสูจน์ว่า มีภาษาที่เรียกว่า ภาษาย้อนกลับได้-เค (*k*-reversible languages) สามารถเรียนรู้ได้จากตัวอย่างบวกเพียงอย่างเดียว งานวิจัยชิ้นนี้เองทำให้การศึกษาความสามารถในการเรียนรู้ภาษาจากตัวอย่างบวกเพียงอย่างเดียวกลับมาได้รับความสนใจอีกครั้ง หลังจากนั้นได้มีระดับของภาษาจำนวนหนึ่งถูกพิสูจน์ว่าเรียนรู้ได้จากตัวอย่างบวก ได้แก่ระดับของภาษาต่อไปนี้ *k*-testable languages [19], Szilard languages [20], *k*-variable pattern languages [21], strictly deterministic languages [22], code regular languages [23], *k*-piecewise testable languages [24], deterministic even linear languages [25], uniquely terminating regular languages [26], terminal distinguishable languages [27], residual languages [28], very simple languages [29], biRFSA languages [30], substitutable context-free languages [31], *k*,*l*-substitutable context-free languages [32], multidimensional substitutable multiple context-free languages [33], strict prefix deterministic languages [34] เป็นต้น จากระดับของภาษาเหล่านี้จะสังเกตได้ว่า งานวิจัยส่วนใหญ่จะศึกษาภาษาในระดับย่อยของระดับภาษาสม่ำเสมอที่นิยามขึ้นมาเฉพาะ

2.3 งานวิจัยที่ศึกษาความสามารถในการเรียนรู้จากตัวอย่างบวกและตัวอย่างลบ

จากทฤษฎีบทเชิงบวกในงานวิจัยของโกลด์ [4] ที่กล่าวว่า ระดับของภาษาวนซ้ำ (recursive languages) สามารถเรียนรู้ได้ในขอบเขตจำกัดจากตัวอย่างบวกและตัวอย่างลบ ทำให้ทราบว่าระดับภาษาสม่ำเสมอ ระดับภาษาไม่พึ่งบริบท ระดับภาษาพึ่งบริบท ซึ่งเป็นระดับย่อยของระดับภาษาวนซ้ำ สามารถเรียนรู้ได้ในขอบเขตจำกัดจากตัวอย่างบวกและตัวอย่างลบด้วย แต่ในงานวิจัยนี้ได้พิสูจน์ให้เห็นเชิงทฤษฎีเท่านั้น ไม่ได้นำเสนออัลกอริทึมการเรียนรู้ที่มีประสิทธิภาพแต่อย่างใด

ในช่วงตั้งแต่ปี ค.ศ. 1973 ถึงปัจจุบันได้มีงานวิจัยจำนวนมากที่ศึกษาเกี่ยวกับความสามารถในการเรียนรู้จากตัวอย่างบวกและตัวอย่างลบ แต่ส่วนใหญ่จะศึกษาที่ระดับภาษาสม่ำเสมอ อัลกอริทึมการเรียนรู้แรกถูกเสนอขึ้นใน ปี ค.ศ. 1973 โดย แทรคเทนบอร์ท และบาร์ซดิน [35] อัลกอริทึมนี้เริ่มต้นจากการสร้างต้นไม้ยอมรับจากตัวอย่างบวก และใช้การเทคนิคการผสานสถานะ (state-merging technique) และใช้ออโตมาตาจำกัดเชิงกำหนดเป็นตัวแทน

เชิงไวยากรณ์ แต่อัลกอริทึมนี้มีข้อจำกัดในการใช้งานคือ ตัวอย่างสำหรับการเรียนรู้ภาษานั้นจะต้องสมบูรณ์ที่ความยาวเท่ากับจำนวนสถานะของออโตมาตาจำกัดเชิงกำหนด ซึ่งทำให้ต้องใช้ตัวอย่างเป็นจำนวนมากในกระบวนการเรียนรู้ และในปี ค.ศ. 1978 โกลด์ ได้เสนออัลกอริทึมการเรียนรู้สำหรับภาษาสม่ำเสมอ แต่อัลกอริทึมนี้สามารถเรียนรู้ได้ในกรณีตัวอย่างที่ใช้ในการเรียนรู้มีตัวอย่างลักษณะอยู่เท่านั้น

หลังจากนั้นในปี ค.ศ. 1992 อนซิน่า และการ์เซีย [36] ได้เสนออัลกอริทึมการเรียนรู้ภาษาสม่ำเสมอที่เรียกว่าอาร์พีเอ็นไอ (Regular Positive Negative Inference) ซึ่งเป็นอัลกอริทึมที่ทำงานในเวลาเชิงพหุนาม และได้พิสูจน์ว่าอัลกอริทึมอาร์พีเอ็นไอสามารถใช้เรียนภาษาสม่ำเสมอได้อย่างมีประสิทธิภาพ โดยไม่ขึ้นกับรูปแบบตัวอย่าง และยังพิสูจน์อีกด้วยว่าเมื่อใดก็ตามที่ตัวอย่างสำหรับการเรียนรู้ภาษามีตัวอย่างลักษณะอยู่ จะเรียนรู้ออโตมาตาจำกัดที่ถูกต้องของภาษาได้ อัลกอริทึมอาร์พีเอ็นไอเริ่มต้นด้วยการสร้างต้นไม้ยอมรับจากตัวอย่างบวกและใช้เทคนิคผสานสถานะ จากนั้นทำการพิจารณาการผสานคู่ของสถานะใด ๆ โดยพิจารณาตามอันดับเชิงพจนานุกรมโดยพิจารณาความยาวรวมด้วย (lexicographic-length order) การอนุญาตให้ผสานสถานะหรือไม่จะพิจารณาจากตัวอย่างลบ โดยถ้าออโตมาตายอมรับตัวอย่างลบก็จะปฏิเสธการผสาน

ในปีเดียวกัน แลง [37] ได้พัฒนาอัลกอริทึมของแทรคเทนบอร์ท และบาร์ซดิน แต่ตัวอย่างสำหรับการเรียนรู้ไม่จำเป็นต้องสมบูรณ์ เขาได้ทดลองกับตัวอย่างข้อมูลภาษาที่เบาบาง (sparse data) และได้รายงานผลที่น่าสนใจว่าการผสานสถานะแนวกว้าง จะให้ผลดีกว่าการผสานสถานะแนวลึก

ต่อมาในปี ค.ศ. 1996 อิกูเอรา และคณะ [38] ได้นำอัลกอริทึมอาร์พีเอ็นไอมาพัฒนา และเสนอขั้นตอนการเลือกสถานะโดยพิจารณาจากค่าความถี่ของข้อมูลที่ผ่านสถานะนั้น ๆ ขั้นตอนนี้เรียกว่า ดีดีดีไอ (Data-Driven DFA Inference) ซึ่งจากผลการทดลอง อัลกอริทึมดีดีดีไอสามารถระบุออโตมาตาจำกัดเชิงกำหนดได้ แต่ใช้จำนวนตัวอย่างลักษณะสำหรับการเรียนรู้มากกว่าอัลกอริทึมอาร์พีเอ็นไอ

ในปี ค.ศ. 1997 นักวิจัยทางด้านการเรียนรู้ภาษา ได้จัดการแข่งขันที่ชื่อว่า แอบบาดินโก (Abbadingo competition) [39] ขึ้นมาเพื่อให้นักวิจัยเสนออัลกอริทึมการเรียนรู้ที่มีประสิทธิภาพสำหรับเรียนรู้ภาษาสม่ำเสมอ ไพรซ์ [40] เป็นผู้ชนะการแข่งขั้นครั้งนี้ โดยใช้อัลกอริทึมที่เรียกว่าอีดีเอสเอ็ม (Evidence Driven State Merging) อัลกอริทึมนี้เริ่มต้นจากการสร้างต้นไม้ยอมรับส่วนหน้าเช่นเดียวกับอัลกอริทึมอาร์พีเอ็นไอ สำหรับการเลือกคู่ของสถานะนั้นจะพิจารณาจากการให้คะแนนการผสานทุกคู่ของสถานะที่เป็นไปได้ และเลือกผสานคู่ของสถานะที่ได้คะแนนมากที่สุด อัลกอริทึมอีดีเอสเอ็มมีข้อเสียคือการไม่คงคุณสมบัติการคงสภาพต้นไม้ (tree invariant) [41] ซึ่งคุณสมบัตินี้กล่าวว่า สำหรับการพิจารณาการผสานคู่สถานะใด ๆ

ต้องมีสถานะอย่างน้อยหนึ่งสถานะเป็นรากของต้นไม้ คุณสมบัตินี้เป็นเงื่อนไขที่เพียงพอที่จะทำให้อัลกอริทึมทำงานได้ในเวลาจำกัด ทำให้อัลกอริทึมอีดีเอสเอ็มไม่สามารถรับประกันเวลาในการทำงานที่จำกัด อัลกอริทึมการเรียนรู้ที่น่าสนใจอีกอัลกอริทึมหนึ่งคือบลูฟริงก์ (Blue Fringe) [40] ซึ่งนำข้อดีของ อัลกอริทึมอีดีเอสเอ็มมาใช้และเพิ่มคุณสมบัติการคงสภาพต้นไม้เข้าไป แต่เนื่องจากปริภูมิการค้นหา (search space) ของคู่สถานะที่ดีที่สุดสำหรับการผสานน้อยกว่าอัลกอริทึมอีดีเอสเอ็ม ทำให้ความถูกต้องน้อยกว่าอัลกอริทึมอีดีเอสเอ็มเล็กน้อย แต่มีข้อดีคือรับประกันเวลาการทำงานที่จำกัด

ในปี ค.ศ. 2000 การ์เซีย และคณะ [42] ได้ทำการศึกษาเปรียบเทียบระหว่างอัลกอริทึมของโกลด์ และอัลกอริทึมอาร์พีเอ็นไอ ผลที่ได้จากการศึกษาคืออัลกอริทึมอาร์พีเอ็นไอให้ค่าความถูกต้องมากกว่า เมื่อใช้จำนวนตัวอย่างการเรียนรู้ที่เท่ากัน

ในปี ค.ศ. 2002 ซิซเชลโล และเคร์เมอร์ [43] ได้พัฒนาอัลกอริทึมอีดีเอสเอ็ม และนำไปวิเคราะห์ พบว่าค่าความถูกต้องที่เพิ่มขึ้น 27 เปอร์เซ็นท์ เกิดจากการผสานสถานะ 5 ครั้งแรก และในปีเดียวกัน อิกูเอรา [44] ได้ทำการตรวจสอบระดับของไวยากรณ์ และออโตมาตาจำกัด ที่สามารถระบุได้ในขอบเขตจำกัดจากตัวอย่างบวกและตัวอย่างลบ และเสนอเงื่อนไขเพียงพอ (sufficient conditions) สำหรับการระบุได้ในขอบเขตจำกัดด้วยเวลาและจำนวนข้อมูลเชิงพหุนาม เงื่อนไขที่เพียงพอนี้คือ การรวมอยู่ (inclusion) ของตัวอย่างลักษณะในตัวอย่างที่ใช้ในการเรียนรู้ และเสนอเงื่อนไขที่จำเป็นคือ ความกำหนด (determinism) และความเป็นเชิงเส้น (linearity)

2.4 สรุป

ในบทนี้ได้กล่าวถึงงานวิจัยต่าง ๆ ที่เกี่ยวข้องกับวิทยานิพนธ์นี้ ซึ่งประกอบด้วยงานวิจัยที่ศึกษาแบบจำลองการเรียนรู้ภาษา งานวิจัยที่ศึกษาความสามารถในการเรียนรู้จากตัวอย่างบวกเพียงอย่างเดียว และงานวิจัยที่ศึกษาความสามารถในการเรียนรู้จากตัวอย่างบวกและตัวอย่างลบ โดยการศึกษาความสามารถในการเรียนรู้ภาษายอมรับได้-เคในวิทยานิพนธ์นี้ได้นำแบบจำลองการเรียนรู้ของภาษาเชิงตัวอย่างของโกลด์มาใช้แบบจำลองหลักในการศึกษา รวมถึงได้ใช้ทฤษฎีบทเชิงบวก ทฤษฎีบทเชิงลบของโกลด์ และทฤษฎีบทความสมมูลของดูปองท์ มาใช้ในการพิสูจน์ความสามารถในการเรียนรู้ของระดับภาษายอมรับได้-เค นอกจากนั้นในวิทยานิพนธ์นี้ได้ใช้ผลการวิจัยรวมถึงเงื่อนไขเพียงพอต่าง ๆ จากงานวิจัยของอิกูเอรามาใช้ในการพิสูจน์ระดับของภาษายอมรับได้-เคเรียนรู้ได้อย่างมีประสิทธิภาพหรือไม่

บทที่ 3

ทฤษฎีพื้นฐานที่ใช้ในงานวิจัย

ในบทนี้จะกล่าวถึงทฤษฎีพื้นฐานที่ใช้ในงานวิจัยนี้ ประกอบด้วย นิยาม สัญลักษณ์ และทฤษฎีบทที่เกี่ยวข้องกับภาษารูปนัยและตัวแทนเชิงไวยากรณ์ รวมถึงวิธีการเรียนรู้ภาษาโดยการอนุมานเชิงไวยากรณ์ นอกจากนี้จะอธิบายถึงแบบจำลองทางคณิตศาสตร์ที่ใช้ศึกษาความสามารถในการเรียนรู้ภาษา

3.1 ภาษารูปนัยและตัวแทนเชิงไวยากรณ์

ในการศึกษาเรื่องภาษารูปนัย หน่วยย่อยที่สุดของภาษา คือ อักขระหรือสัญลักษณ์ (character or symbol) ซึ่งเมื่อนำตัวอักขระเหล่านี้มารวมกันในรูปแบบของเซต เราเรียกเซตนี้ว่า เซตอักขระ (alphabet) ในกรณีที่ตัวอักขระแต่ละตัวมีอันดับ สามารถแสดงอันดับก่อนหลังของตัวอักขระได้ เราเรียกเซตอักขระลักษณะนี้ว่า เซตอักขระมีอันดับ (ordered alphabet) เมื่อนำอักขระต่าง ๆ มาเรียงต่อกัน เราจะเรียกผลของการเรียงต่อกันว่า สายอักขระ (string) และเมื่อนำสายอักขระที่มีความหมายอย่างเดียวกันมารวมกลุ่มกันในรูปแบบเซต เราเรียกเซตของสายอักขระนี้ว่า ภาษา (language) โดยความหมายของสายอักขระในแต่ละภาษาเป็นตัวแสดงความแตกต่างระหว่างภาษาหนึ่งกับอีกภาษาหนึ่ง ซึ่งความหมายเกิดจากการประกอบกันของอักขระอยู่ในรูปแบบกฎที่เรียกว่า ไวยากรณ์ (grammar) สำหรับตัวแบบที่ใช้ในการรู้จำ (recognize) ภาษา เรียกว่า ออโตมาตา (automata) ทั้งไวยากรณ์และออโตมาตาเป็นตัวแทนเชิงไวยากรณ์ของภาษาที่สามารถใช้บอกความซับซ้อนของระดับของภาษาได้ เรานิยามคำศัพท์ต่าง ๆ ที่อ้างถึงในงานวิจัยดังต่อไปนี้

นิยาม 3.1 เซตอักขระ *หมายถึง เซตจำกัดของอักขระ และเรียกสมาชิกในเซตอักขระนี้ว่า* อักขระ *เราแทนเซตอักขระด้วยสัญลักษณ์* Σ

นิยาม 3.2 สายอักขระ *หมายถึง ลำดับของอักขระ และ* ความยาวของสายอักขระ *หมายถึง จำนวนอักขระที่ปรากฏในลำดับ ความยาวของสายอักขระ* w *เขียนแทนด้วยสัญลักษณ์* $|w|$ *และเราเรียกสายอักขระที่มีความยาวเท่ากับศูนย์ว่า* สายอักขระว่าง (*null string*) *แทนด้วยสัญลักษณ์* λ *และเราแทนเซตของสายอักขระทั้งหมดที่สร้างจากอักขระใน* Σ *ซึ่งรวมสายอักขระว่างด้วยสัญลักษณ์* Σ^*

ตัวอย่าง 3.1 *ให้ $\Sigma = \{a, b\}$ เป็นเซตอักขระที่ประกอบด้วยอักขระ a และ b สายอักขระที่นิยามบน Σ สร้างจากการนำอักขระ a และ b มาเรียงต่อเข้าด้วยกันเป็นลำดับ เช่น สายอักขระความยาวหนึ่ง คือ a และ b ส่วนสายอักขระความยาวสอง คือ aa, ab, ba และ bb สำหรับ Σ^* เป็นเซตของสายอักขระทุกตัวที่เป็นไปได้ที่มีความยาวตั้งแต่ศูนย์ขึ้นไป ดังนั้น $\Sigma^* = \{\lambda, a, b, aa, ab, ba, bb, aaa, aab, aba, abb, baa, bab, bba, bbb, \ldots\}$* □

นิยาม 3.3 *กำหนดให้ u, v และ w เป็นสายอักขระ และ $w = uv$ เป็นสายอักขระที่เกิดจากการเรียงต่อกันของสายอักขระ u และสายอักขระ v เราเรียกสายอักขระ u ว่า* สายอักขระส่วนหน้า (*prefix*) *ของ w และ เรียกสายอักขระ v ว่า* สายอักขระส่วนหลัง (*suffix*) *ของสายอักขระ w*

ข้อสังเกต สายอักขระว่าง เป็นสายอักขระส่วนหน้าและเป็นสายอักขระส่วนหลังของทุกสายอักขระ และสายอักขระใด ๆ จะเป็นทั้งสายอักขระส่วนหน้าและเป็นสายอักขระส่วนหลังของตัวมันเองเสมอ

นิยาม 3.4 *ให้สายอักขระ $w, u, v \in \Sigma^*$ โดยที่ $w = uv$* เซตของสายอักขระส่วนหน้า *(set of prefixes) ของ w* นิยามโดย $Pref(w) = \{u \in \Sigma^* : w = uv\}$

ตัวอย่าง 3.2 *กำหนดให้ $w = abbbba$ เป็นสายอักขระที่นิยามบนเซตอักขระ $\Sigma = \{a, b\}$ สายอักขระ ab เป็นสายอักขระส่วนหน้าตัวหนึ่งของสายอักขระ w และสายอักขระ ba เป็นสายอักขระส่วนหลังตัวหนึ่งของสายอักขระ w ส่วนเซตของสายอักขระส่วนหน้าของสายอักขระ w คือ $Pref(w) = \{\lambda, a, ab, abb, abbb, abbbb, abbbba\}$* □

นิยาม 3.5 ภาษา *หมายถึง เซตของสายอักขระที่เกิดจากการเรียงต่อกันของอักขระในเซตอักขระ Σ เราแทนภาษาด้วยสัญลักษณ์ L ภาษาที่มีสายอักขระเป็นจำนวนจำกัด เรียกว่า* ภาษาจำกัด (*finite language*) *และภาษาที่มีสายอักขระเป็นจำนวนอนันต์ เรียกว่า* ภาษาอนันต์ (*infinite language*)

ข้อสังเกต Σ^* เป็นภาษาอนันต์ภาษาหนึ่ง และภาษาทุกภาษาที่นิยามบนเซตอักขระ Σ เป็นเซตย่อยของ Σ^* ไม่ว่าจะเป็นภาษาจำกัดหรือภาษาอนันต์ก็ตาม เนื่องจากสมาชิกของภาษาทุกตัวสร้างจากการเรียงต่อกันของสัญลักษณ์ใน Σ

ในทฤษฎีภาษารูปนัย การบรรยายภาษาสามารถทำได้โดยตัวแทนเชิงไวยากรณ์ ซึ่งมีอยู่สองรูปแบบ คือ ไวยากรณ์ (grammar) และออโตมาตา (automata) ไวยากรณ์เป็นตัวแบบเชิงการสร้าง ส่วนออโตมาตาเป็นตัวแบบเชิงการรู้จำภาษา ทั้งสองตัวแบบนี้นอกจากใช้บรรยายภาษาแล้ว ยังสามารถใช้จำแนกภาษาออกเป็นระดับตามความซับซ้อนในการบรรยายได้อีกด้วย ซึ่งแต่ละตัวแบบนิยามดังนี้

นิยาม 3.6 *ไวยากรณ์ หมายถึง ตัวแบบที่ถูกกำหนดขึ้นในการบรรยายภาษา ซึ่งประกอบด้วยส่วนสำคัญสี่ส่วน* $G = (V, \Sigma, S, P)$

โดยที่ V *คือ เซตของตัวแปร*

Σ *คือ เซตอักขระ*

S *คือ ตัวแปรเริ่มต้น ซึ่งเป็นสมาชิกของเซต* V

P *คือ เซตของกฎการผลิต (production rule) ในรูปแบบ* $\alpha \rightarrow \beta$ *โดยที่* $\alpha, \beta \in (V \cup \Sigma)^*$

ความสามารถในการบรรยายภาษาของไวยากรณ์ขึ้นอยู่กับรูปแบบของกฎการผลิต นอม ชอมสกี ได้แบ่งความซับซ้อนของภาษา โดยพิจารณาความสามารถในการบรรยายภาษาของไวยากรณ์ในรูปแบบของลำดับชั้นที่เรียกว่า ลำดับชั้นของชอมสกี ไวยากรณ์แต่ละรูปแบบมีดังนี้

ไวยากรณ์ไม่มีข้อจำกัด (unrestricted grammar) มีรูปแบบคือ

$\alpha \rightarrow \beta$ *โดยที่* $\alpha, \beta \in (V \cup \Sigma)^*$

ไวยากรณ์พึ่งบริบท (context-sensitive grammar) มีรูปแบบคือ

$\alpha \rightarrow \beta$ *โดยที่* $\alpha, \beta \in (V \cup \Sigma)^*$ และ $|\alpha| \leq |\beta|$

ไวยากรณ์ไม่พึ่งบริบท (context-free grammar) มีรูปแบบคือ

$\alpha \rightarrow \beta$ *โดยที่* $\alpha \in V, \beta \in (V \cup \Sigma)^*$ และ $|\alpha| \leq |\beta|$

ไวยากรณ์สม่ำเสมอ (regular grammar) มีรูปแบบคือ

$\alpha \rightarrow \beta$ *โดยที่* $\alpha \in V, \beta \in \{a : a \in \Sigma\} \cup \{aC : a \in \Sigma, C \in V\}$ และ $|\alpha| \leq |\beta|$

ข้อสังเกต จากตัวแปร α, β ในกฎการผลิตของไวยากรณ์แต่ละแบบ สังเกตว่าไวยากรณ์ไม่มีข้อจำกัด $\alpha, \beta \in (V \cup \Sigma)^*$ นั่นคือ α, β สามารถประกอบด้วยตัวแปรหรืออักขระเป็นจำนวนเท่าใดก็ได้ ด้วยความที่ไม่มีข้อจำกัดนี้ทำให้ไวยากรณ์รูปแบบนี้จึงสามารถใช้สร้างภาษาที่มีความซับซ้อนมากได้ แต่เมื่อเทียบกับไวยากรณ์สม่ำเสมอ ซึ่ง $\alpha \in V, \beta \in \{a : a \in \Sigma\} \cup \{aC : a \in \Sigma, C \in V\}$ และ $|\alpha| \leq |\beta|$ จะเห็นว่า ตัวแปร α, β มีข้อจำกัดการใช้งานมากกว่า ทำให้ไวยากรณ์สม่ำเสมอไม่สามารถสร้างภาษาที่มีความซับซ้อนมากได้

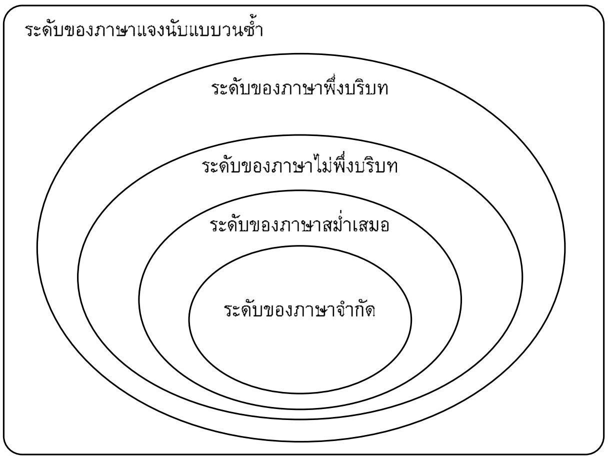
ไวยากรณ์ที่แสดงข้างต้น เรียงลำดับความสามารถในการบรรยายภาษาจากน้อยไปมาก คือ ไวยากรณ์สม่ำเสมอ ไวยากรณ์ไม่พึ่งบริบท ไวยากรณ์พึ่งบริบท และไวยากรณ์ไม่จำกัด ความสามารถในการบรรยายภาษาของแต่ละไวยากรณ์ใช้สำหรับการแบ่งระดับของภาษาตามความซับซ้อน ตามนิยามต่อไปนี้

นิยาม 3.7 ระดับของภาษา (*class of languages*) *หมายถึง เซตของภาษาที่สร้างจากไวยากรณ์ที่มีความซับซ้อนระดับเดียวกัน นั่นคือ*

- *เซตของภาษาที่สร้างจากไวยากรณ์สม่ำเสมอ เรียกว่า* ระดับของภาษาสม่ำเสมอ (*class of regular languages*)
- *เซตของภาษาที่สร้างจากไวยากรณ์ไม่พึ่งบริบท เรียกว่า* ระดับของภาษาไม่พึ่งบริบท (*class of context-free languages*)
- *เซตของภาษาที่สร้างจากไวยากรณ์พึ่งบริบท เรียกว่า* ระดับของภาษาพึ่งบริบท (*class of context-sensitive languages*)
- *เซตของภาษาที่สร้างจากไวยากรณ์ไม่มีข้อจำกัด เรียกว่า* ระดับของภาษาแจงนับแบบ-วนซ้ำ (*class of recursively enumerable languages*)

นิยาม 3.8 ระดับของภาษาจำกัด (*class of finite languages*) *หมายถึง เซตที่ประกอบด้วยภาษาจำกัด*

นิยาม 3.9 ระดับจำกัดของภาษา (*finite class of languages*) *หมายถึง เซตจำกัดที่ประกอบด้วยภาษา (เป็นภาษาจำกัดหรือภาษาอนันต์ก็ได้)* ระดับเกินจำกัดของภาษา *(super-finite class of languages) หมายถึง เซตที่ประกอบด้วยภาษาจำกัดทั้งหมด และภาษาอนันต์อย่างน้อยหนึ่งภาษา*



รูปที่ 3.1 ลำดับชั้นของภาษารูปนัย

ระดับต่าง ๆ ของภาษาสามารถแสดงความสัมพันธ์ในรูปแบบของลำดับชั้นตามรูป 3.1 ซึ่งแสดงถึงความซับซ้อนในการบรรยายภาษาของไวยากรณ์แต่ละรูปแบบ เช่น ไวยากรณ์พึ่งบริบทสามารถใช้บรรยายภาษาไม่พึ่งบริบทได้ รวมถึงสามารถใช้บรรยายภาษาสม่ำเสมอก็ได้ แต่ในทางกลับกันไวยากรณ์สม่ำเสมอใช้บรรยายภาษาสม่ำเสมอได้เท่านั้น ไม่สามารถใช้บรรยายภาษาไม่พึ่งบริบทได้

นอกจากจะใช้ไวยากรณ์เป็นตัวแบบในการแบ่งความซับซ้อนของภาษาแล้ว ตัวแบบทางคณิตศาสตร์อีกตัวแบบหนึ่งที่ใช้สำหรับบอกความซับซ้อนของภาษาได้ คือ ออโตมาตา ซึ่งสามารถตรวจสอบความเป็นสมาชิกของภาษาได้ สำหรับออโตมาตาพื้นฐานที่เกี่ยวข้องกับงานวิจัยนี้ แสดงดังนิยามต่อไปนี้

นิยาม 3.10 ออโตมาตาจำกัด (*finite automata*) *หมายถึง ตัวแบบของทางคณิตศาสตร์ ซึ่งประกอบด้วยส่วนสำคัญหกส่วน คือ* $M = (\Sigma, Q, q_0, F_A, F_R, \delta)$ *โดยที่*

Σ *แทน เซตอักขระ ซึ่งเป็นเซตจำกัด*

Q *แทน เซตของสถานะ ซึ่งเป็นเซตจำกัด*

q_0 *แทน สถานะเริ่มต้น (initial state) ซึ่ง* $q_0 \in Q$

F_A *แทน เซตของสถานะยอมรับ* (*accepting state*) *ซึ่ง* $F_A \subseteq Q$

F_R *แทน เซตของสถานะปฏิเสธ* (*rejecting state*) *ซึ่ง* $F_R \subseteq Q$

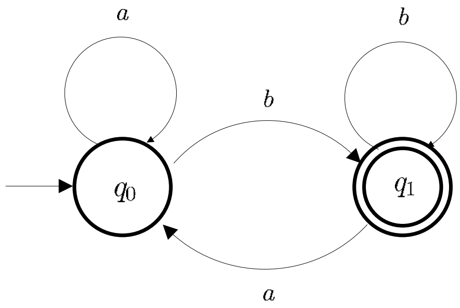
δ *แทน ฟังก์ชันการเปลี่ยน* (*transition function*) *นิยามโดย*

$$\delta : Q \times \Sigma \to Q$$

ฟังก์ชันการเปลี่ยน δ *สามารถถูกนิยามเพิ่มเติมเพื่อให้ใช้ได้กับสายอักขระดังนี้* $\delta^* : Q \times \Sigma^* \to Q$ *เราเรียกฟังก์ชันการเปลี่ยนที่นิยามเพิ่มเติมนี้ว่า* ฟังก์ชันการเปลี่ยนขยาย (*extended transition function*) และ ขนาดของออโตมาตาจำกัด *หมายถึง จำนวนสมาชิกในเซต* Q

นิยาม 3.11 ออโตมาตาจำกัดเชิงกำหนด (*deterministic finite automata* : DFA) *หมายถึง ออโตมาตาจำกัดที่ประกอบด้วยส่วนสำคัญหกส่วน* $M = (\Sigma, Q, q_0, F_A, F_R, \delta)$ *แต่มีการนิยามเพิ่มเติมสำหรับฟังก์ชันการเปลี่ยน คือ* $|\delta(q, x)| \leq 1$ *สำหรับทุกสถานะ* $q \in Q$ *และสำหรับทุกตัวอักขระ* $x \in \Sigma$

ตัวอย่าง 3.3 *กำหนดให้* M *เป็นออโตมาตาจำกัดขนาด 2 ซึ่งประกอบด้วย* $Q = \{q_0, q_1\}$, $\Sigma = \{a, b\}$, $F_A = \{q_1\}$, $F_R = \{q_0\}$ *และมีฟังก์ชันการเปลี่ยนซึ่งนิยามดังนี้* $\delta(q_0, a) = q_0$, $\delta(q_0, b) = q_1$, $\delta(q_1, b) = q_1$, *และ* $\delta(q_1, a) = q_0$ *จะเห็นว่า* M *เป็นออโตมาตาจำกัดเชิงกำหนด เพราะ* $|\delta(q, x)| \leq 1$ *สำหรับทุกตัวอักขระ* $x \in \Sigma$ *ออโตมาตาจำกัดเชิงกำหนด* M *สามารถแสดงด้วยแผนภาพสถานะดังรูปที่ 3.2* □



รูปที่ 3.2 แผนภาพสถานะของออโตมาตาจำกัดเชิงกำหนด

การตรวจสอบความเป็นสมาชิกของสายอักขระใด ๆ เริ่มต้นโดยกำหนดออโตมาตาให้อยู่ที่สถานะเริ่มต้น และทำงานโดยการรับข้อมูลเข้าที่อยู่ในรูปแบบสายอักขระ โดยทำการอ่านข้อมูลทีละตัวอักขระตามลำดับ จากนั้นออโตมาตาทำการเปลี่ยนจากสถานะหนึ่งไปยังอีกสถานะหนึ่งตามฟังก์ชันการเปลี่ยน และออโตมาตาหยุดการทำงานเมื่ออ่านอักขระทุกตัวในสายอักขระครบหมดแล้ว ทั้งนี้สถานะสุดท้ายจะเป็นตัวกำหนดว่าออโตมาตาจะยอมรับหรือปฏิเสธสายอักขระนั้น กล่าวคือ ถ้าสถานะสุดท้ายหลังจากการทำงานเป็นสมาชิกของเซตของสถานะยอมรับ แสดงว่าออโตมาตายอมรับสายอักขระ แต่ถ้าสถานะสุดท้ายหลังจากการทำงานเป็นสมาชิกของเซตของสถานะปฏิเสธ แสดงว่าออโตมาตาปฏิเสธสายอักขระนั้น

นิยาม 3.12 *ให้ M เป็นออโตมาตาจำกัด* ภาษาที่ถูกยอมรับโดยออโตมาตาจำกัด *เขียนแทนด้วย $\mathbb{L}(M)$ หมายถึง เซตของสายอักขระทั้งหมดที่หลังจากออโตมาตาจำกัดอ่านข้อมูลทีละตัวอักขระแล้วหยุดการทำงานที่สถานะยอมรับ นิยามโดย $\mathbb{L}(M) = \{w \in \Sigma^* : \delta^*(q_0, w) \in F_A\}$*

ตัวอย่าง 3.4 *ภาษาที่ถูกยอมรับโดยออโตมาตาจำกัด M* จากตัวอย่างที่ *3.3* คือภาษา *$\mathbb{L}(M) = \{b, ab, bb, aab, abb, bbb, \ldots\}$ หรือภาษาที่ลงท้ายด้วยตัวอักขระ b* □

นิยาม 3.13 *ให้ $S = S+ \cup S-$ เป็นเซตที่ประกอบด้วย $S+$ เซตของสายอักขระที่อยู่ในภาษา L และ $S-$ เซตของสายอักขระที่ไม่อยู่ในภาษา L* ออโตมาตายอมรับส่วนหน้าเพิ่มเติม (*augmented prefix tree acceptor* : $APTA$) *สำหรับ S หมายถึงออโตมาตาจำกัดเชิงกำหนดที่มีลักษณะคล้ายต้นไม้ ที่ประกอบด้วยส่วนสำคัญหกส่วน นิยามโดย $M_{APTA} = (\Sigma, Q, q_0, F_A, F_R, \delta)$*
โดยที่แต่ละส่วนสำคัญเป็นไปตามเงื่อนไขดังต่อไปนี้

- $Q = \{q_u : u \in Pref(S+) \cup Pref(S-)\}$
- $q_0 = q_\lambda$
- $F_A = \{q_w \in \Sigma^* : w \in S+\}$
- $F_R = \{q_w \in \Sigma^* : w \in S-\}$

– *ฟังก์ชันการเปลี่ยน δ นิยามโดย*

$\forall u \in Pref(S+) \cup Pref(S-),\ \forall a \in \Sigma,\ \forall q_u,\ q_{ua} \in Q$ *โดยที่*

$\delta(q_u, a) = q_{ua} \Leftrightarrow ua \in Pref(S+) \cup Pref(S-)$

ออโตมาตายอมรับส่วนหน้าเพิ่มเติม M_{APTA} จะยอมรับเฉพาะสายอักขระในเซต $S+$ และปฏิเสธเฉพาะสายอักขระในเซต $S-$ เท่านั้น แสดงดังตัวอย่างต่อไปนี้

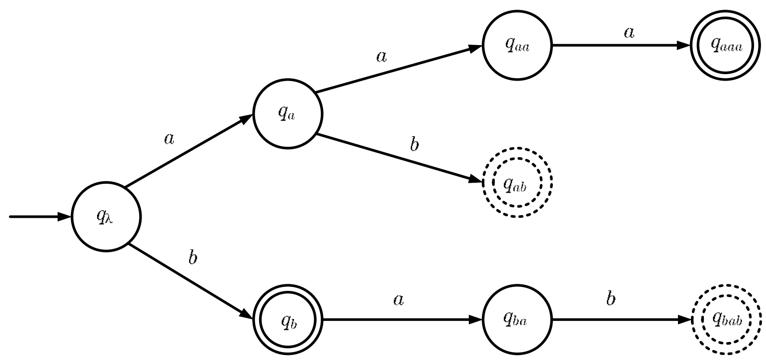
ตัวอย่าง 3.5 *กำหนดให้ $S+ = \{b, aaa\}$ และ $S- = \{ab, bab\}$ เป็นเซตของสายอักขระที่อยู่ในภาษาและเซตของสายอักขระที่ไม่อยู่ในภาษาที่นิยามบนเซตอักขระ $\Sigma = \{a, b\}$ ออโตมาตายอมรับส่วนหน้าเพิ่มเติม สำหรับตัวอย่าง $S = S+ \cup S-$ คือ $M_{APTA} = (\Sigma, Q, q_0, F_A, F_R, \delta)$ โดยที่*

– $Q = \{q_\lambda, q_a, q_b, q_{aa}, q_{ab}, q_{ba}, q_{aaa}, q_{bab}\}$
– $q_0 = q_\lambda$
– $F_A = \{q_b, q_{aaa}\}$
– $F_R = \{q_{ab}, q_{bab}\}$
– *ฟังก์ชันการเปลี่ยนสถานะ δ นิยามดังนี้*

$\delta(q_\lambda, a) = q_a,\ \delta(q_\lambda, b) = q_b,\ \delta(q_a, a) = q_{aa}, \delta(q_a, b) = q_{ab},$

$\delta(q_b, a) = q_{ba}, \delta(q_{ba}, b) = q_{bab}, \delta(q_{aa}, a) = q_{aaa}$ □

แผนภาพการผ่านของ M_{APTA} แสดงในรูปที่ 3.3 โดยแทนสถานะยอมรับด้วยรูปวงกลมซ้อนกันสองวงที่วาดด้วยเส้นทึบ และแทนสถานะปฏิเสธด้วยรูปวงกลมซ้อนกันสองวงที่วาดด้วยเส้นประ



รูปที่ 3.3 แผนภาพสถานะของ M_{APTA}

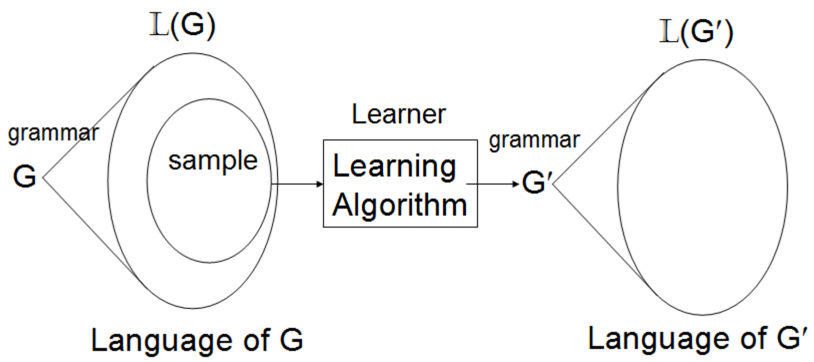
นิยาม 3.14 ภาษาสม่ำเสมอ (*regular language*) *หมายถึง ภาษาที่รู้จำได้ (recognize) โดยออโตมาตาจำกัด*

ตัวอย่าง 3.6 จากออโตมาตาในรูป 3.1 สามารถรู้จำภาษาที่ลงท้ายด้วยตัวอักขระ b หรือ $L = (a+b)^*b$ ได้ ดังนั้น L เป็นภาษาสม่ำเสมอ □

3.2 การเรียนรู้ภาษา

เนื่องจากภาษาเป็นเซต (อาจเป็นเซตจำกัด หรือเซตอนันต์) ที่สามารถบรรยายโดยใช้ตัวแทนเชิงไวยากรณ์ที่มีรูปแบบจำกัด เช่น ไวยากรณ์ หรือออโตมาตาจำกัด ดังนั้นในงานวิจัยด้านการอนุมานเชิงไวยากรณ์ จึงพิจารณาการเรียนรู้ภาษาเป็นกระบวนการหาตัวแทนเชิงไวยากรณ์ของภาษา ซึ่งส่วนใหญ่ศึกษาบนแบบจำลองการเรียนรู้เชิงตัวอย่างของโกลด์ [10]

ในกระบวนการเรียนรู้ภาษาของโกลด์ ตัวอย่าง (example) ของภาษาจะถูกส่งไปให้ผู้เรียน (learner) ผ่านวิธีการนำเสนอข้อมูลภาษา (presentation) ที่กำหนดขึ้น ผู้เรียนจะทำการอนุมาน (inference) โดยระบุ (identify) ตัวแทนเชิงไวยากรณ์ของภาษานั้น ๆ ตามตัวอย่างที่ได้รับ กระบวนการการเรียนรู้ภาษานี้สามารถแสดงดังรูป 3.4



รูปที่ 3.4 กระบวนการเรียนรู้ภาษาจากการอนุมานตัวแทนเชิงไวยากรณ์ของภาษา

จากรูปที่ 3.4 จะสังเกตเห็นว่า ในการกระบวนการเรียนรู้ภาษาจากการอนุมานไวยากรณ์ภาษาจากตัวอย่างที่ให้มานั้น ไม่จำเป็นจะต้องอนุมานให้ได้ตัวแทนเชิงไวยากรณ์ G ซึ่งเป็นไวยากรณ์ที่แท้จริงของภาษา แต่เพียงแค่อนุมานได้ตัวแทนเชิงไวยากรณ์ G' ซึ่งสามารถบรรยายภาษาที่สร้างได้จากตัวแทนเชิงไวยากรณ์ G ก็เพียงพอ หรือ $\mathbb{L}(G) = \mathbb{L}(G')$ ซึ่ง $\mathbb{L}(G)$ เป็นฟังก์ชันระบุภาษาที่นิยามดังนี้

นิยาม 3.15 *ให้ $\mathcal{L}$ เป็นระดับของภาษาบนเซตอักขระ Σ และ $\mathcal{G}$ เป็นระดับของตัวแทนเชิงไวยากรณ์ของภาษา เรานิยาม* ฟังก์ชันระบุภาษา *ดังนี้*

$$\mathbb{L} : \mathcal{G} \to \mathcal{L}$$

ข้อสังเกต ฟังก์ชันระบุภาษานี้ ใช้สำหรับระบุภาษาที่สอดคล้องกับไวยากรณ์หรือออโตมาตา ยกตัวอย่างเช่น กำหนดให้ M เป็นออโตมาตาจำกัด แล้ว $\mathbb{L}(M)$ คือภาษาที่ถูกรู้จำโดย M หรือให้ G เป็นไวยากรณ์ $\mathbb{L}(G)$ ก็คือภาษาที่สร้างได้จากไวยากรณ์ G

ตัวอย่างที่ใช้ในการเรียนรู้ภาษา สามารถพิจารณาเป็นข้อมูลในรูปแบบต่าง ๆ ที่สามารถใช้ในการเรียนรู้ภาษา และการนำตัวอย่างแต่ละตัวไปให้ผู้เรียนทำการเรียนรู้ เราเรียกว่า การนำเสนอข้อมูลภาษา ซึ่งนิยามในรูปแบบของฟังก์ชันจากเซตของจำนวนนับ $\mathbb{N}$ ไปยังเซตใด ๆ ซึ่งเป็นเซตของข้อมูลภาษาในรูปแบบต่าง ๆ นิยามของการนำเสนอข้อมูลภาษาแสดงดังนี้

นิยาม 3.16 *ให้ L เป็นภาษาที่นิยามบนเซตอักขระ Σ เราเรียกเซตย่อยของ L ว่า* เซตของตัวอย่างบวกของภาษา L *แทนด้วยสัญลักษณ์ $S+$ นิยามโดย*

$$S+ = \{w \in \Sigma^* : w \in L\}$$

และเรียกเซตของสายอักขระบางตัวที่ไม่อยู่ใน L ว่า เซตของตัวอย่างลบของภาษา L *แทนด้วยสัญลักษณ์ $S-$ นิยามโดย*

$$S- = \{w \in \Sigma^* : w \in \Sigma^* - L\}$$

ตัวอย่าง 3.7 *กำหนดให้ L เป็นภาษาที่นิยามบนเซตอักขระ $\Sigma = \{a, b\}$ และ L เป็นภาษาที่มีสายอักขระทุกตัวลงท้ายด้วย b เซตของตัวอย่างบวกของภาษา L ที่เป็นไปได้เซตหนึ่งคือ $S+ = \{b, ab, bbb\}$ และเซตของตัวอย่างลบของภาษา L ที่เป็นไปได้เซตหนึ่งคือ $S- = \{a, ba, bba, aaa\}$* □

นิยาม 3.17 *ให้ L เป็นภาษาที่นิยามบนเซตอักขระ Σ* การนำเสนอข้อมูลภาษา (presentation) *ของภาษา L คือ ฟังก์ชัน $f : \mathbb{N} \to \mathcal{X}$ โดยที่ $\mathbb{N}$ เป็นเซตของจำนวนนับและ $\mathcal{X}$ เป็นเซตใด ๆ การนำเสนอข้อมูลภาษาครั้งที่ i นิยามโดย $f_i = \{f(j) \in \mathcal{X} : 1 \le j \le i\}$*

ข้อสังเกต ฟังก์ชัน f อาจพิจารณาเป็นรูปแบบการเลือกข้อมูลภาษาของผู้นำเสนอข้อมูลก็ได้ f_i หมายถึงเซตของข้อมูลภาษาตั้งแต่ตัวแรกถึงตัวที่ i และ $f(i)$ หมายถึงข้อมูลภาษาตัวที่ i ซึ่งจะเห็นว่า $f_i \subseteq f_{i+1}$ เสมอ

ตัวอย่าง 3.8 *ให้ L เป็นภาษาที่ลงท้ายด้วยตัวอักขระ b ซึ่งภาษา L เป็นภาษาที่นิยามบนเซตอักขระ $\Sigma = \{a, b\}$ การนำเสนอข้อมูลภาษารูปแบบหนึ่งของภาษา L คือ*

$$f = \{(1, b), (2, bb), (3, abb), (4, bbb), \ldots\}$$

จะสังเกตเห็นว่า $f_2 \subseteq f_3$ เพราะว่า $f_2 = \{(1, b), (2, bb)\}$ และ $f_3 = \{(1, b), (2, bb), (3, abb)\}$ นั่นก็คือ ในการนำเสนอข้อมูลภาษาแต่ละครั้งจะรวมข้อมูลที่นำเสนอไปแล้วด้วยเสมอ เช่น f_2 นำเสนอสายอักขระ b และ bb ส่วน f_3 นำเสนอสายอักขระ b, bb ซึ่งเป็นของ f_2 และเพิ่มสายอักขระ abb เข้าไป □

สำหรับเซต $\mathcal{X}$ ซึ่งเป็นเซตใด ๆ เราสามารถพิจารณาเซต $\mathcal{X}$ เป็นเซตของข้อมูลภาษาในรูปแบบที่แตกต่างกันได้สองรูปแบบคือ รูปแบบแรกเซต $\mathcal{X}$ เป็นเซตของเซตของตัวอย่างบวกเพียงอย่างเดียว $S+$ และรูปแบบที่สอง $\mathcal{X}$ เป็นเซตของตัวอย่างบวกและตัวอย่างลบ $S+ \cup S-$ ก็ได้ สำหรับการนำเสนอที่เป็นไปได้ทั้งหมดของภาษา L นิยามในรูปแบบเซตของฟังก์ชันดังนี้

นิยาม 3.18 *ให้ L เป็นภาษาใด ๆ ในระดับภาษา $\mathcal{L}$* การนำเสนอที่เป็นไปได้ทั้งหมดของ L *นิยามดังนี้*

$$Pres(L) = \{f \ : f \text{ เป็นฟังก์ชันจาก } \mathbb{N} \text{ ไป } \mathcal{X} \text{ และ } Yield(f) = L\}$$

โดยที่ $Yield$ เป็นฟังก์ชันที่นิยามโดย $Yield : Pres(\mathcal{L}) \rightarrow \mathcal{L}$

ในส่วนของผู้เรียนภาษา เราสามารถพิจารณาเป็นอัลกอริทึมการเรียนรู้ภาษา ซึ่งนิยามในรูปแบบของฟังก์ชัน จากเซตของการนำเสนอข้อมูลภาษา ไปยังเซตของตัวแทนเชิงไวยากรณ์ของภาษา นิยามดังนี้

นิยาม 3.19 *ให้ $\mathcal{G}$ เป็นระดับของตัวแทนเชิงไวยากรณ์สำหรับระดับของภาษา $\mathcal{L}$* อัลกอริทึมการเรียนรู้ภาษา $\mathcal{A}$ *คือ โปรแกรมที่รับข้อมูลภาษาแล้วคืนค่ากลับในรูปแบบของตัวแทนเชิงไวยากรณ์ของภาษา เรานิยามอัลกอริทึมการเรียนรู้ภาษาอยู่ในรูปแบบของฟังก์ชันได้ดังนี้*

$$\mathcal{A} : \{f_i : i \in \mathbb{N}, f \in Pres(L) \ \} \rightarrow \mathcal{G}$$

3.3 ความสามารถในการเรียนรู้ภาษา

ความสามารถในการเรียนรู้ภาษาเป็นคุณสมบัติหนึ่งของระดับภาษา ระดับภาษาใดที่มีคุณสมบัติดังกล่าวแสดงว่าระดับภาษานั้นมีอัลกอริทึมการเรียนรู้สำหรับทุกภาษาในระดับนั้น แบบจำลองทางคณิตศาสตร์ที่ใช้ศึกษาความสามารถในการเรียนรู้ภาษามีอยู่หลายแบบจำลอง แต่แบบจำลองที่ยอมรับกันในกลุ่มนักวิจัยด้านนี้ มีอยู่สามแบบจำลองด้วยกัน ดังที่จะกล่าวถึงต่อไปนี้

3.3.1 การระบุภาษาได้ในขอบเขตจำกัด (Identification in the limit)

แบบจำลองที่เรียกว่า การระบุภาษาได้ในขอบเขตจำกัด เป็นแบบจำลองทางคณิตศาสตร์ที่เป็นต้นแบบของการศึกษาวิจัยด้านการเรียนรู้ภาษา โดยพิจารณาการเรียนรู้เป็นกระบวนการอนุมานไวยากรณ์ของภาษาจากตัวอย่างที่ได้รับ และเป็นกระบวนการที่ไม่สิ้นสุด (Infinite process) การพิจารณาว่าการเรียนรู้สำเร็จหรือไม่นั้น ใช้แนวคิดที่ว่า เมื่อผู้เรียนได้รับตัวอย่างสำหรับการเรียนรู้ที่เพิ่มจำนวนขึ้น ถึงจุดจำกัดจุดหนึ่ง ผู้เรียนจะอนุมานตัวแทนเชิง

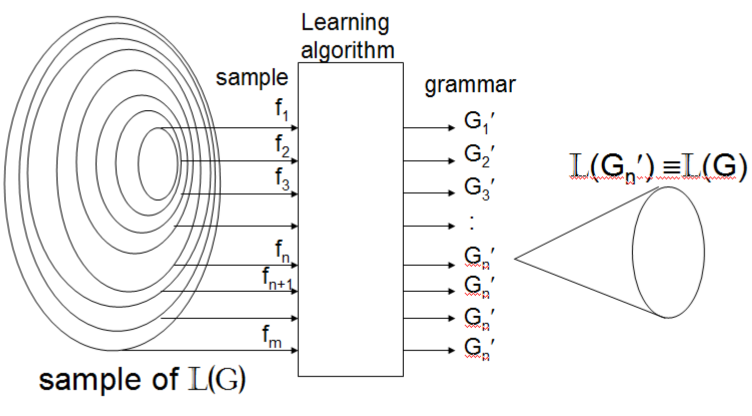
ไวยากรณ์ที่สามารถบรรยายภาษาได้ทั้งหมด และหลังจากนั้นไม่ว่าจะได้รับตัวอย่างเพิ่มขึ้นมากเท่าใดก็ตาม ผู้เรียนก็ยังอนุมานตัวแทนเชิงไวยากรณ์ได้ตัวเดิม นิยามและทฤษฎีบทที่เกี่ยวข้องต่าง ๆ แสดงดังนี้

นิยาม 3.20 *ให้* L *เป็นภาษาใด ๆ ซึ่งมี* G *เป็นตัวแทนเชิงไวยากรณ์ของ* L *และ* f *เป็นฟังก์ชันการนำเสนอตัวอย่างสำหรับ* L *เรากล่าวว่า* อัลกอริทึมการเรียนรู้ $\mathcal{A}$ สามารถระบุภาษา L ได้ในขอบเขตจำกัด *ถ้ามีจำนวนเต็ม* n *จำนวนหนึ่งที่ทำให้* $\mathbb{L}(\mathcal{A}(f_m)) = \mathbb{L}(G_n)$ *โดยที่* $\mathbb{L}(G_n) = L$ *สำหรับทุกจำนวนเต็มบวก* $m \geq n$

นิยาม 3.21 *ให้* $\mathcal{L}$ *เป็นระดับของภาษาใด ๆ ที่นิยามบนเซตอักขระ* Σ *และ* $\mathcal{G}$ *เป็นระดับของตัวแทนภาษาสำหรับระดับภาษา* $\mathcal{L}$ *เรากล่าวว่าระดับภาษา* $\mathcal{L}$ สามารถระบุภาษาได้ในขอบเขตจำกัด *ด้วยการนำเสนอตัวอย่าง* f *ถ้ามีอัลกอริทึมการเรียนรู้* $\mathcal{A}$ *ที่สามารถระบุทุกภาษา* $L \in \mathcal{L}$ *ด้วย* f *ได้ในขอบเขตจำกัด*

ข้อสังเกต จากนิยาม 3.21 จะเห็นว่า การที่จะบอกว่าระดับภาษาสามารถเรียนรู้ภาษาได้หรือไม่นั้น ไม่จำเป็นต้องหาอัลกอริทึมการเรียนรู้ที่สามารถหาไวยากรณ์ที่แท้จริงของภาษา เพียงแค่หาอัลกอริทึมการเรียนรู้ที่หาตัวแทนเชิงไวยากรณ์ที่สมมูลกับตัวแทนเชิงไวยากรณ์ที่แท้จริงของภาษาได้ ก็เพียงพอที่จะกล่าวว่าระดับภาษานั้นสามารถเรียนรู้ได้ในขอบเขตจำกัด

แนวคิดของแบบจำลองการระบุภาษาได้ในขอบเขตจำกัดนี้คือ เมื่อผู้เรียนได้รับตัวอย่างสำหรับการเรียนรู้ที่เพิ่มจำนวนขึ้นจนถึงจุดจำกัดจุดหนึ่ง ผู้เรียนจะอนุมานตัวแทนเชิงไวยากรณ์ที่สามารถบรรยายภาษาที่กำลังเรียนรู้ได้ทั้งหมด และหลังจากนั้นไม่ว่าจะได้รับตัวอย่างเพิ่มขึ้นมากเท่าใดก็ตาม ผู้เรียนก็ยังคงอนุมานไวยากรณ์ได้ตัวเดิม เราสามารถแสดงแบบจำลองการระบุภาษาได้ในขอบเขตจำกัดในรูปที่ 3.5



รูปที่ 3.5 แบบจำลองการระบุภาษาได้ในขอบเขตจำกัด

นิยาม 3.22 *ให้ $\mathcal{L}$ เป็นระดับของภาษาที่นิยามบนเซตอักขระ Σ เรากล่าว $\mathcal{L}$* สามารถเรียนรู้ได้ในขอบเขตจำกัด (*learnable in the limit*) *ถ้า $\mathcal{L}$ สามารถระบุภาษาได้ในขอบเขตจำกัด*

โกลด์ [10] ได้ใช้นิยาม 3.21 ในการพิสูจน์ความสามารถในการเรียนรู้ของระดับของภาษาในลำดับชั้นของชอมสกี ซึ่งแสดงดังทฤษฎีบท 3.1, 3.2, 3.3 และ 3.4

ทฤษฎีบท 3.1 *ระดับของภาษาจำกัดสามารถเรียนรู้ได้ในขอบเขตจำกัดด้วยตัวอย่างบวกเพียงอย่างเดียว*

พิสูจน์ การพิสูจน์ทำได้โดยการแสดงให้เห็นว่ามีอัลกอริทึมการเรียนรู้ตามนิยาม 3.21 อยู่จริง

ให้ L เป็นภาษาจำกัดใด ๆ ซึ่งมีสมาชิก n ตัว และ f เป็นการนำเสนอตัวอย่างบวกสำหรับภาษา L เนื่องจากภาษา L เป็นเซตจำกัด ดังนั้นการเรียนรู้ภาษา L นั้นทำได้โดยใช้ตัวแทนเชิงไวยากรณ์ภาษา G ที่อยู่ในรูปแบบเซตจำกัดซึ่งมีตัวอย่างบวกทั้งหมดที่ถูกนำเสนอเป็นสมาชิก เรากำหนดให้อัลกอริทึมการเรียนรู้ $\mathcal{A}$ มีขั้นตอนการทำงานดังนี้คือ $\mathcal{A}(f_i) = G_j = \cup_{k=1 \text{ to } i}\{w \in \Sigma^* : w \in f_k(k)\}$ เนื่องจากภาษา L เป็นเซตจำกัด จะเห็นว่ามีจำนวนเต็ม n ที่ทำให้ $\mathbb{L}(\mathcal{A}(f_m)) = \mathbb{L}(G_n)$ และ $\mathbb{L}(G_n) = L$ เสมอสำหรับทุกจำนวนเต็มบวก $m \geq n$ ■

ทฤษฎีบท 3.2 *ระดับจำกัดของภาษาที่ประกอบด้วยภาษาจำกัดอยู่เป็นจำนวนจำกัด และภาษาอนันต์ไม่เกินหนึ่งภาษา สามารถเรียนรู้ได้ในขอบเขตจำกัดด้วยตัวอย่างบวกเพียงอย่างเดียว*

พิสูจน์ การพิสูจน์ทำได้โดยการแสดงให้เห็นว่ามีอัลกอริทึมการเรียนรู้ตามนิยาม 3.21 อยู่จริง

ให้ $\mathcal{L}$ เป็นระดับจำกัดของภาษา ที่สมาชิกแต่ละตัวถูกเรียงลำดับ $L_0, L_1, L_2, \ldots, L_k, L_\infty$ โดยที่ $L_0, L_1, L_2, \ldots, L_k$ เป็นภาษาจำกัด และ L_∞ เป็นภาษาอนันต์ อัลกอริทึมการเรียนรู้ $\mathcal{A}$ ของระดับภาษา $\mathcal{L}$ นี้ คือ $\mathcal{A}$ จะทำการพิจารณาภาษาในระดับภาษา $\mathcal{L}$ ทีละตัว ว่าสอดคล้องกับตัวอย่างบวกที่ได้รับมาหรือไม่ ถ้าสอดคล้องจะคืนค่าภาษาตัวแรกที่เจอนั้นออกมา จากอัลกอริทึมนี้จะเห็นว่า การเรียนรู้ด้วยอัลกอริทึม $\mathcal{A}$ จะมีจำนวนเต็ม n ที่ทำให้ $\mathbb{L}(\mathcal{A}\ (f_m)) = \mathbb{L}(G_n) = \mathcal{L}$ สำหรับทุกจำนวนเต็มบวก $m \geq n$ เสมอ แสดงว่าอัลกอริทึมการเรียนรู้ $\mathcal{A}$ สามารถระบุทุกภาษาในระดับ $\mathcal{L}$ ได้ นั้นหมายความว่าระดับ $\mathcal{L}$ สามารถเรียนรู้ได้ในขอบเขตจำกัดด้วยตัวอย่างบวกเพียงอย่างเดียว ■

ทฤษฎีบท 3.3 *ระดับเกินจำกัดของภาษา ไม่สามารถเรียนรู้ได้ในขอบเขตจำกัดด้วยตัวอย่างบวกเพียงอย่างเดียว*

พิสูจน์ การพิสูจน์ทำได้ใช้เทคนิคการพิสูจน์แบบขัดแย้งดังนี้

ให้ $\mathcal{L}$ เป็นระดับภาษาเกินจำกัด ซึ่งประกอบด้วยภาษา $L_0,\ L_1,\ L_2,\ \ldots,\ L_n, \ldots$ เป็นภาษาจำกัด และภาษา L_∞ เป็นภาษาอนันต์ โดยแต่ละภาษาใน $\mathcal{L}$ นิยามดังนี้ สำหรับภาษาจำกัด $L_0 = \{w_0\}$ และ $L_{i+1} = L_i \cup \{w_{i+1}\}$ โดยที่ $L_i \subset L_\infty$ และ $L_\infty = \cup_{i\in\mathbb{N}} L_i$

สมมติให้ระดับภาษา $\mathcal{L}$ สามารถเรียนรู้ได้จากตัวอย่างบวกเพียงอย่างเดียว จากนิยาม 3.21 แสดงว่าจะต้องมีอัลกอริทึมการเรียนรู้ $\mathcal{A}$ สำหรับทุกภาษาในระดับภาษา $\mathcal{L}$ โดยที่มีจำนวนเต็ม n ค่าหนึ่งที่ทำให้ $\mathbb{L}(\mathcal{A}(f_m)) = \mathbb{L}(G)$ สำหรับทุกจำนวนเต็มบวก $m \geq n$

พิจารณาการมีอยู่ของจำนวนเต็ม n ที่ทำให้ $\mathbb{L}(\mathcal{A}(f_m)) = \mathbb{L}(G_n) = \mathcal{L}$ สำหรับทุกจำนวนเต็มบวก $m \geq n$ จะได้ดังนี้

ให้ p_i คือจำนวนเต็มสำหรับภาษา L_i ที่ทำให้ $\mathbb{L}(\mathcal{A}(f^i{}_m)) = \mathbb{L}(G_{pi}) = L_i$ โดยที่ f^i เป็นการนำเสนอตัวอย่างบวกของภาษา L_i และให้การนำเสนอข้อมูลของภาษาอื่น ๆ ใน $\mathcal{L}$ ที่สร้างดังนี้ คือ สำหรับทุก $j \leq p_{k\text{-}1}$, $f^k(j) = f^{k\text{-}1}(j)$ และ $f^k(p_{k\text{-}1}{+}1) = w_k$

พิจารณาการเรียนรู้ภาษา L_∞ ด้วยการนำเสนอตัวอย่างบวก f^∞ จะเห็นว่าเราไม่สามารถหาจำนวนเต็ม n ที่ทำให้ $\mathbb{L}(\mathcal{A}(f^\infty{}_m)) = \mathbb{L}(G_n) = L_\infty$ สำหรับ $m \geq n$ ได้ ดังนั้นแสดงว่าอัลกอริทึม $\mathcal{A}$ ไม่มีอยู่จริง ทำให้เกิดข้อขัดแย้งในการพิสูจน์ ดังนั้นเราสามารถสรุปได้ว่าไม่มีระดับเกินจำกัดของภาษาใด สามารถระบุภาษาได้ในขอบเขตจำกัดจากตัวอย่างบวกเพียงอย่างเดียว ■

ทฤษฎีบทตาม 3.1 *ระดับภาษาสม่ำเสมอ ระดับภาษาไม่พึ่งบริบท ระดับภาษาพึ่งบริบท ระดับภาษาแจงนับได้แบบวนซ้ำ ไม่สามารถเรียนรู้ได้ในขอบเขตจำกัดด้วยตัวอย่างบวกเพียงอย่างเดียว*

พิสูจน์ เนื่องจากระดับภาษาสม่ำเสมอ ระดับภาษาไม่พึ่งบริบท ระดับภาษาพึ่งบริบท ระดับภาษาแจงนับได้แบบวนซ้ำ เป็นระดับภาษาเกินจำกัด ดังนั้นจากทฤษฎีบท 3.3 เราสรุปได้ว่าระดับภาษาเหล่านี้ไม่สามารถเรียนรู้ด้วยตัวอย่างบวกเพียงอย่างเดียว ■

ทฤษฎีบท 3.4 *ระดับของภาษาวนซ้ำ สามารถเรียนรู้ได้ในขอบเขตจำกัดด้วยตัวอย่างบวกและตัวอย่างลบ*

พิสูจน์ การพิสูจน์ทำได้โดยการแสดงให้เห็นว่ามีอัลกอริทึมการเรียนรู้ตามนิยาม 3.21 อยู่จริง

ให้ $\mathcal{L}$ เป็นระดับของภาษาวนซ้ำ ซึ่งสมาชิกใน $\mathcal{L}$ สามารถเรียงลำดับได้ อัลกอริทึมการเรียนรู้ $\mathcal{A}$ ของระดับภาษา $\mathcal{L}$ นี้ คือ $\mathcal{A}$ จะทำการพิจารณาภาษาในระดับภาษา $\mathcal{L}$ ทีละตัว ว่าสอดคล้องกับตัวอย่างบวกและตัวอย่างลบที่ได้รับมาหรือไม่ ในแต่ละครั้งของการเรียนรู้จากตัวอย่าง อัลกอริทึมการเรียนรู้ $\mathcal{A}$ จะทำการกำจัดภาษาที่ไม่สอดคล้องกับตัวอย่างบวกและตัวอย่างลบที่ได้รับมา และจะทำการคืนภาษาในลำดับถัดไปออกมา จะเห็นว่าจำนวนครั้งการกำจัดภาษาเป็นจำนวนจำกัด แสดงว่ามีจำนวนเต็ม n ที่ทำให้ $\mathbb{L}(\mathcal{A}(f_m)) = \mathbb{L}(G_n) = L$ สำหรับทุกจำนวนเต็มบวก $m \geq n$ แสดงว่าอัลกอริทึม $\mathcal{A}$ สามารถเรียนรู้ทุกภาษาใน $\mathcal{L}$ ได้ นั่นหมายความว่าระดับ $\mathcal{L}$ สามารถเรียนรู้ได้ในขอบเขตจำกัดด้วยตัวอย่างบวกและตัวอย่างลบ ■

ทฤษฎีบทตาม 3.2 *ระดับของภาษาสม่ำเสมอ ระดับภาษาไม่พึ่งบริบท ระดับภาษาพึ่งบริบท สามารถเรียนรู้ได้ในขอบเขตจำกัดจากตัวอย่างบวกและตัวอย่างลบ*

พิสูจน์ เนื่องจากระดับภาษาสม่ำเสมอ ระดับภาษาไม่พึ่งบริบท ระดับภาษาพึ่งบริบท เป็นระดับย่อยแท้ของระดับภาษาวนซ้ำ แสดงว่าภาษาที่เป็นสมาชิกของระดับภาษาเหล่านี้เป็นภาษาวนซ้ำด้วย และจากทฤษฎีบท 3.4 เราจึงสรุปได้ว่า ระดับของภาษาสม่ำเสมอ ระดับภาษาไม่พึ่งบริบท ระดับภาษาพึ่งบริบท สามารถเรียนรู้ได้จากตัวอย่างบวกและตัวอย่างลบ ■

จากทฤษฎีบท 3.3 จะเห็นว่าระดับภาษาสม่ำเสมอ ระดับภาษาไม่พึ่งบริบท ระดับภาษาพึ่งบริบท และระดับภาษาแจงนับแบบวนซ้ำนั้น เป็นระดับเกินจำกัด ส่งผลให้ระดับภาษาเหล่านี้ไม่สามารถเรียนรู้ได้จากตัวอย่างบวกเพียงอย่างเดียว ซึ่งนอกจากคุณสมบัติดังกล่าวแล้วยังมีคุณสมบัติอื่น ๆ ที่สามารถใช้เป็นเงื่อนไขที่เพียงพอในการพิสูจน์ว่าไม่สามารถเรียนรู้ได้จากตัวอย่างบวกเพียงอย่างเดียว นั่นก็คือ การมีจุดจำกัด (limit point) [15] ซึ่งคุณสมบัติที่กล่าวถึงนิยามดังต่อไปนี้

นิยาม 3.23 *ระดับภาษา $\mathcal{L}$ มี* จุดจำกัด *(limit point) ก็ต่อเมื่อมีลำดับอนันต์ $<L_n>$ ซึ่ง $n \in \mathbb{N}$ เป็นลำดับของภาษาใน $\mathcal{L}$ โดยที่ $L_0 \subset L_1 \subset \ldots \subset L_n \subset \ldots$ และมีภาษา $L_p \in \mathcal{L}$ โดยที่*

$$L_p = \cup_{n \in \mathbb{N}} L_n$$

และเราเรียกภาษา L_p ว่า จุดจำกัดของระดับภาษา $\mathcal{L}$

ทฤษฎีบท 3.5 *ให้ $\mathcal{L}$ เป็นระดับของภาษาใด ๆ ถ้า $\mathcal{L}$ มีจุดจำกัดแล้ว $\mathcal{L}$ ไม่สามารถเรียนรู้ได้ในขอบเขตจำกัดจากตัวอย่างบวกเพียงอย่างเดียว*

พิสูจน์ การพิสูจน์ทำได้เช่นเดียวกับทฤษฎีบท 3.3 ■

3.3.2 การระบุภาษาได้จากข้อมูลที่กำหนดให้ (Identification from given data)

แบบจำลองการเรียนรู้ที่เรียกว่า การระบุภาษาได้จากข้อมูลที่กำหนดให้เป็นแบบจำลองที่พัฒนาขึ้นจากแบบจำลองการระบุภาษาได้ในขอบเขตจำกัดในหัวข้อ 3.3.1 โดยมีแนวคิดที่ว่าตัวอย่างสำหรับการเรียนรู้ภาษานั้น แต่ละตัวอย่างมีคุณภาพที่ทำให้การเรียนรู้สำเร็จไม่เท่ากัน การเรียนรู้จะสำเร็จหรือไม่นั้นขึ้นอยู่กับว่าผู้เรียนได้รับตัวอย่างที่มีคุณภาพดี หรือตัวอย่างที่เรียกว่า ตัวอย่างลักษณะ (characteristic examples) มาครบแล้วหรือไม่ นิยามต่าง ๆ และทฤษฎีบทที่เกี่ยวข้องกับแบบจำลองการระบุภาษาได้จากข้อมูลที่กำหนดให้มีดังนี้

นิยาม 3.24 *ให้ L เป็นภาษา เราเรียกเซตของตัวอย่าง $S = (S+, \varnothing)$ ว่า*เซตของตัวอย่างบวกของภาษา L *ถ้า $S+ \subseteq L$ และเรียกเซตของตัวอย่าง $S = (S+, S-)$ ว่า*เซตของตัวอย่างบวกและตัวอย่างลบ *ถ้า $S+ \subseteq L$ และ $S- \subseteq \Sigma^* - L$ และ* ขนาดของเซตของตัวอย่าง S *เขียนแทนด้วย $||S||$ หมายถึง ผลรวมของความยาวของสายอักขระที่เป็นสมาชิกของ S โดยนิยามดังนี้*

$$||S|| = \Sigma_{i=1}|w_i| \text{ โดยที่ } w_i \in S$$

ตัวอย่าง 3.9 *ให้ L เป็นภาษา ab^* ที่นิยามบนเซตอักขระ $\Sigma = \{a, b\}$ จงยกตัวอย่างเซตของตัวอย่างบวกที่เป็นไปได้ และเซตของตัวอย่างบวกและตัวอย่างลบที่เป็นไปได้ รวมทั้งหาขนาดของแต่ละเซต*

วิธีทำ เซตของตัวอย่างบวกที่เป็นไปได้เซตหนึ่งคือ

$S_1 = (S_1+, \varnothing)$ โดยที่ $S_1+ = \{a, ab, abb\}$

คำนวณหาขนาดของ $S_1 = (S_1+, \varnothing)$ ได้ดังนี้

$||S_1|| = ||S_1+|| = |a| + |ab| + |abb| = 1+2+3 = 6$

เซตของตัวอย่างบวกและตัวอย่างลบที่เป็นไปได้เซตหนึ่งคือ

$S_2 = (S_2+, S_2-)$ โดยที่ $S_2+ = \{ab, abbb\}$ และ $S_2- = \{\lambda, aa, baa\}$

คำนวณหาขนาดของ $S_2 = (S_2+, S_2-)$ ได้ดังนี้

$$\begin{aligned} ||S_2|| &= ||S_2+|| + ||S_2-|| \\ &= |ab| + |abbb| + |\lambda| + |aa| + |baa| = 2+4+0+2+3 = 11 \end{aligned}$$ □

นิยาม 3.25 *ให้ $S = (S+,\ S\text{-})$ และ $S' = (S'+,\ S'\text{-})$ เป็นเซตของตัวอย่างบวกและตัวอย่างลบของภาษา L เราจะกล่าวว่า S'* เป็นเซตย่อยของ *S เขียนแทนด้วย $S' \subseteq S$ ถ้า $S'+ \subseteq S+$ และ $S'\text{-} \subseteq S\text{-}$*

ข้อสังเกต ในกรณีเซตของตัวอย่าง $S = (S+, \varnothing)$ เป็นเซตของตัวอย่างบวกเพียงอย่างเดียว เราสามารถนิยามการเป็นเซตย่อยได้เช่นเดียวกัน โดยพิจารณาการเป็นเซตย่อยเฉพาะ $S+$ เท่านั้น

ตัวอย่าง 3.10 *ให้ L เป็นภาษา ab^* ที่นิยามบนเซตอักขระ $\Sigma = \{a,b\}$ และ $S = (S+,\ S-)$ โดยที่ $S+ = \{ab,\ abbb,\ abbbb\}$ และ $S- = \{\lambda,\ aa,\ baa,\ aabb\}$ เป็นเซตของตัวอย่าวบวกและตัวอย่างลบของภาษา L จงยกตัวอย่างเซตย่อยของ S*

วิธีทำ เซตย่อยของเซตของตัวอย่างบวกและตัวอย่างลบ $S = (S+,\ S-)$ ที่เป็นไปได้เซตหนึ่งก็คือ $S' = (S'+,\ S'-)$ โดยที่ $S'+ = \{ab,\ abbb\}$ และ $S'- = \{\lambda,\ aa,\ baa\}$ เนื่องจาก $S'+ \subseteq S+$ และ $S'\text{-} \subseteq S\text{-}$ $\square$

นิยาม 3.26 *ให้ S เป็นเซตของตัวอย่างของภาษา L โดยมี $\mathcal{A}$ เป็นอัลกอริทึมการเรียนรู้ เราเรียกเซต CS ซึ่งเป็นเซตของตัวอย่างว่า* เซตของตัวอย่างลักษณะ *ของภาษา L สำหรับอัลกอริทึมการเรียนรู้ $\mathcal{A}$ ถ้า $\mathbb{L}(\mathcal{A}(S)) = L$ สำหรับทุกเซตของตัวอย่าง $S \supseteq CS$*

ข้อสังเกต ในทฤษฎีการเรียนรู้ ตัวอย่างลักษณะคือ ตัวอย่างดีที่ทำให้การเรียนรู้ภาษาประสบความสำเร็จ เมื่อใดก็ตามที่ตัวอย่างลักษณะถูกรวมเข้าไปในเซตของตัวอย่างที่ใช้ในการเรียนรู้แล้วจะสามารถเรียนรู้ภาษาได้อย่างถูกต้อง จากนิยาม 3.26 จะสังเกตเห็นว่าตัวอย่างลักษณะขึ้นอยู่กับอัลกอริทึมการเรียนรู้ ซึ่งหมายความว่าอัลกอริทึมที่แตกต่างกันย่อมใช้ตัวอย่างลักษณะที่แตกต่างกัน

นิยาม 3.27 *ระดับภาษา $\mathcal{L}$* สามารถเรียนรู้ได้จากตัวอย่างที่กำหนดให้ *ก็ต่อเมื่อ มีอัลกอริทึมการเรียนรู้ $\mathcal{A}$ โดยที่ $\mathbb{L}(\mathcal{A}(S)) = L$ เมื่อ $S \supseteq CS$ สำหรับทุกภาษา L ใน $\mathcal{L}$*

ข้อสังเกต ความสามารถในการเรียนรู้ภาษาจากตัวอย่างที่กำหนดให้ตามนิยาม 3.27 แตกต่างกับความสามารถการเรียนรู้ในขอบเขตจำกัดตามนิยาม 3.21 คือ การเรียนรู้ภาษาจากตัวอย่างที่กำหนดให้เป็นพิจารณาลักษณะการเรียนรู้เป็นกระบวนการที่จำกัด (finite process) แต่การเรียนรู้ภาษาในขอบเขตจำกัดพิจารณาลักษณะการเรียนรู้เป็นกระบวนการที่ไม่จำกัด (infinite process) แต่อย่างไรก็ตาม มีทฤษฎีบทที่กล่าวว่าการเรียนรู้ทั้งสองแบบนี้สมมูลกันซึ่งได้พิสูจน์ในงานวิจัยของดูปองท์ [14] แสดงดังทฤษฎีบท 3.6

ทฤษฎีบท 3.6 *ระดับภาษาใด ๆ สามารถเรียนรู้ได้ในขอบเขตจำกัด ก็ต่อเมื่อ ระดับภาษานั้นสามารถเรียนรู้ได้จากตัวอย่างที่กำหนดให้*

พิสูจน์ กำหนดให้ $\mathcal{L}$ เป็นระดับภาษาใด ๆ การพิสูจน์เริ่มโดยการสมมติให้ $\mathcal{L}$ สามารถเรียนรู้ได้ในขอบเขตจำกัด จากนิยาม 3.21 แสดงว่ามีจำนวนเต็ม n ที่ทำให้ $\mathbb{L}(\mathcal{A}(f_m)) = \mathbb{L}(\mathcal{A}(f_n)) = L$ สำหรับ $m \geq n$ และสำหรับทุกภาษา L ใน $\mathcal{L}$ เนื่องจากแบบจำลองการเรียนรู้ได้ในขอบเขตจำกัดเป็นการเรียนรู้แบบไม่จำกัด (infinite learning) จะได้ว่า $f_n \subseteq f_m$ เสมอ ซึ่งจะเห็นว่า f_n คือเซตของตัวอย่างลักษณะตามนิยาม 3.26 ดังนั้นระดับภาษา $\mathcal{L}$ สามารถเรียนรู้ได้จากตัวอย่างที่กำหนดให้ด้วย

สมมติให้ระดับภาษา $\mathcal{L}$ สามารถเรียนรู้ได้จากตัวอย่างที่กำหนดให้ จากนิยาม 3.26 จะได้ว่า มีตัวอย่างลักษณะ CS ที่ทำให้ $\mathbb{L}(\mathcal{A}(S)) = L$ เมื่อ $S \supseteq CS$ สำหรับทุกภาษา L ใน $\mathcal{L}$ หรือ $\mathbb{L}(\mathcal{A}(S)) = \mathbb{L}(\mathcal{A}(CS)) = L$ พิจารณาการเรียนรู้เป็นแบบไม่จำกัด โดยให้ $f_n = CS$ และ $f_m = S$ และเนื่องจาก CS เป็นเซตจำกัด แสดงว่ามีจำนวนเต็ม n ที่ทำให้ $\mathbb{L}(\mathcal{A}(f_m)) = \mathbb{L}(\mathcal{A}(f_n))$ และ $\mathbb{L}(\mathcal{A}(f_n)) = L$ ดังนั้นระดับภาษา $\mathcal{L}$ สามารถเรียนรู้ได้ในขอบเขตจำกัด ■

ทฤษฎีบทตาม 3.3 *ระดับของภาษาสม่ำเสมอ ระดับของภาษาไม่พึ่งบริบท ระดับของภาษาพึ่งบริบท ระดับของภาษาวนซ้ำ สามารถเรียนรู้ได้จากตัวอย่างที่กำหนดให้*

พิสูจน์ เนื่องจากระดับของภาษาสม่ำเสมอ ระดับของภาษาไม่พึ่งบริบท ระดับของภาษาพึ่งบริบท ระดับของภาษาวนซ้ำ สามารถเรียนรู้ได้ในขอบเขตจำกัด (ทฤษฎีบทตาม 3.2) เราสามารถสรุปได้จากทฤษฎีบท 3.6 ว่าระดับของภาษาสม่ำเสมอ ระดับของภาษาไม่พึ่งบริบท ระดับของภาษาพึ่งบริบท ระดับของภาษาวนซ้ำ สามารถเรียนรู้ได้จากตัวอย่างที่กำหนดให้ ■

3.3.3 การระบุภาษาได้ในขอบเขตจำกัดจากเวลาและข้อมูลเชิงพหุนาม (Identification in the limit from polynomial time and data)

สำหรับแบบจำลองการเรียนรู้ภาษาที่จะกล่าวถึงในหัวข้อนี้ ได้มีการพิจารณาถึงประสิทธิภาพในการเรียนรู้ภาษาเพิ่มเข้ามาในแบบจำลอง โดยนำแบบจำลองการระบุภาษาได้จากตัวอย่างที่กำหนดให้ จากหัวข้อ 3.3.2 มาพัฒนาเพิ่มเติมในส่วนของเวลาที่ผู้เรียน (อัลกอริทึมการเรียนรู้)ใช้ในการเรียนรู้ภาษา และขนาดของตัวอย่างที่ทำให้ผู้เรียนสามารถอนุมานไวยากรณ์ที่ใช้อธิบายภาษาได้สำเร็จ แบบจำลองนี้เรียกว่า การระบุภาษาได้ในขอบเขตจำกัดจากเวลาและข้อมูลเชิงพหุนาม นิยามและทฤษฎีบทที่เกี่ยวข้องแสดงดังนี้

นิยาม 3.28 *ให้ $\mathcal{L}$ เป็นระดับของภาษา โดยมี $\mathcal{G}$ เป็นระดับของตัวแทนเชิงไวยากรณ์ของ $\mathcal{L}$ และให้ S เป็นเซตของตัวอย่างสำหรับภาษา $L \in \mathcal{L}$ ซึ่งมี CS เป็นเซตของตัวอย่างลักษณะ เราจะกล่าวว่า L* สามารถเรียนรู้ได้ในขอบเขตจำกัดจากเวลาและข้อมูลเชิงพหุนาม *ถ้ามีอัลกอริทึมการเรียนรู้ $\mathcal{A}$ สำหรับทุกภาษา $L \in \mathcal{L}$ โดยที่*

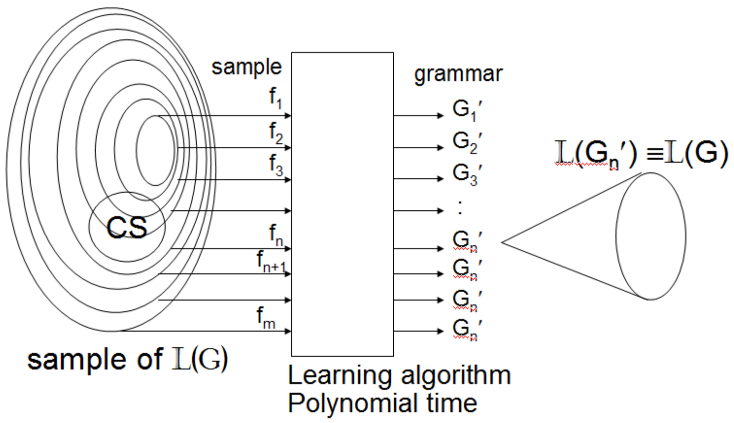
(i) $\mathbb{L}(\mathcal{A}(S)) = L$ *เมื่อ* $S \supseteq CS$

(ii) $||CS|| \leq p(||G||)$ *เมื่อ G เป็นตัวแทนเชิงไวยากรณ์ของภาษา L*

(iii) $\mathcal{A}$ *ทำงานด้วยเวลา* $\mathcal{O}(q(||S||))$

เมื่อ $p(x)$ และ $q(x)$ เป็นฟังก์ชันพหุนามของ x

จากนิยาม 3.28 จะสังเกตเห็นว่ามีการพิจารณาขนาดของเซตของตัวอย่างลักษณะ และเวลาที่อัลกอริทึมการเรียนรู้ใช้ เพิ่มเติมจากนิยาม 3.21 นั่นก็คือเวลาที่อัลกอริทึมใช้อนุมานหาตัวแทนเชิงไวยากรณ์ G จากเซตของตัวอย่าง S ต้องเป็นฟังก์ชันพหุนามของ $||S||$ และขนาดของเซตของตัวอย่างลักษณะต้องมีอัตราการเติบโตไม่เกินฟังก์ชันพหุนามของ $||G||$ แบบจำลองการระบุภาษาได้ในขอบเขตจำกัดจากเวลาและข้อมูลเชิงพหุนามแสดงดังรูปที่ 3.6



รูปที่ 3.6 แบบจำลองการระบุภาษาได้ในขอบเขตจำกัดจากเวลาและข้อมูลเชิงพหุนาม

3.4 สรุป

ในบทนี้ได้กล่าวถึงนิยาม และทฤษฎีบทที่เกี่ยวข้อง ซึ่งประกอบด้วย นิยามพื้นฐานของภาษารูปนัย ทฤษฎีบทของความสามารถในการเรียนรู้ของภาษาระดับต่าง ๆ เช่น ระดับภาษาสม่ำเสมอ ระดับภาษาไม่พึ่งบริบท ระดับภาษาพึ่งบริบท ระดับภาษาวนซ้ำ นิยามและทฤษฎีบทที่ได้กล่าวถึงในบทนี้จะถูกนำไปใช้ในการศึกษาความสามารถการเรียนรู้ภาษายอมรับได้-เค ในบทต่อไป

บทที่ 4

ระดับของภาษายอมรับได้-เค

ในบทนี้จะกล่าวถึงระดับของภาษายอมรับได้-เค (*k*-acceptable languages) โดยจะทำการศึกษาคุณสมบัติต่าง ๆ ของระดับของภาษานี้ รวมทั้งศึกษาคุณสมบัติของตัวแทนเชิงไวยากรณ์ของระดับภาษานี้ ซึ่งก็คือ ออโตมาตาจำกัดเชิงกำหนดขอบ-เค (*k*-edge deterministic finite automata) โดยเสนอผลการศึกษาในรูปแบบทฤษฎีบทและบทพิสูจน์ ซึ่งจะถูกนำไปใช้ในการศึกษาความสามารถในการเรียนรู้ของระดับภาษายอมรับได้-เคในบทต่อไป

4.1 นิยามของภาษายอมรับได้-เค

นิยาม 4.1 ให้ $\Sigma = \{a_1, a_2, \ldots, a_n\}$ *เป็นเซตอักขระ และสัญลักษณ์* $\leq$ *เป็นความสัมพันธ์ทวิภาคบน* Σ *ซึ่งมีสมบัติสะท้อน ปฏิสมมาตร และถ่ายทอดบนเซต* Σ เซตอักขระมีอันดับ *(ordered alphabet) หมายถึง เซตอักขระที่อักขระทุกตัวในเซตมีอันดับ แทนด้วยสัญลักษณ์* $\Sigma_\leq$

ตัวอย่าง 4.1 *สมมติให้* $\Sigma_\leq = \{a, b, c, d\}$ *เป็นเซตอักขระมีอันดับ เราสามารถเรียงลำดับสมาชิกในเซตอักขระ* $\Sigma_\leq$ *ได้คือ* a, b, c และ d □

นิยาม 4.2 ออโตมาตาจำกัดเชิงกำหนดขอบ-เค *(k-edge deterministic finite automata หรือ* k-DFA*) หมายถึง ตัวแบบของเครื่องจักรนามธรรม* M *ซึ่งประกอบด้วยส่วนสำคัญหกส่วน นิยามโดย*

$$M = (\Sigma_\leq, Q, q_0, F_A, F_R, \delta)$$

โดยที่

- $\Sigma_\leq$ *คือ เซตอักขระมีอันดับ ซึ่งเป็นเซตจำกัด*
- Q *คือ เซตของสถานะ ซึ่งเป็นเซตจำกัด*
- q_0 *คือ สถานะเริ่มต้น ซึ่ง* $q_0 \in Q$
- F_A *คือ เซตของสถานะยอมรับ ซึ่ง* $F_A \subseteq Q$
- F_R *คือ เซตของสถานะปฏิเสธ ซึ่ง* $F_R \subseteq Q$
- δ *คือ ฟังก์ชันการเปลี่ยน-เค (k-transition function) นิยามโดย*

$$\delta : Q \times \Sigma_\leq \times \Sigma_\leq \rightarrow Q$$

โดยที่ $\forall q \in Q$,

ถ้า $\delta_k(q, a_1, b_1) \neq \delta_k(q, a_2, b_2)$ และ $\{z : a_1 \leq z \leq b_1\} \cap \{z : a_2 \leq z \leq b_2\} = \varnothing$

แล้ว $|\{q' : \delta_k(q, x, y) = q'\}| \leq k$

ฟังก์ชันการเปลี่ยน δ สามารถถูกนิยามเพิ่มเติมเพื่อให้ใช้ได้กับสายอักขระ ดังนี้

$$\delta^* : Q \times \Sigma^*_{\leq} \rightarrow Q$$

โดยที่ $\delta^*(q, \lambda) = q$

$$\delta^*(q, aw) = \delta^*(q', w) \text{ ซึ่ง } x \leq a \leq y \text{ และ } \delta(q, x, y) = q'$$

โดยที่ $q, q' \in Q$ และ $a, x, y \in \Sigma_{\leq}, w \in \Sigma^*_{\leq}$

และเราเรียกฟังก์ชันการเปลี่ยน-เคที่นิยามเพิ่มเติมนี้ว่า ฟังก์ชันการเปลี่ยนขยาย-เค (extended k-transition function)

ข้อสังเกต ค่า k ใน k-DFA คือ จำนวนขอบที่มากที่สุดของแต่ละสถานะ ดังนั้น k-DFA สามารถพิจารณาว่าเป็น m-DFA สำหรับจำนวนเต็ม $m > k$ ได้เช่นเดียวกัน

นิยาม 4.3 *ให้ q และ p เป็นสถานะใด ๆ ในเซตของสถานะ Q ที่นิยามในออโตมาตาจำกัดเชิงกำหนดขอบ-เค $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ เราเรียก q ว่า* ตัวก่อนหน้า *$(predecessor)$ ของ p และเรียก p ว่า*ตัวตามหลัง *$(successor)$ ของ q ถ้า $\delta(q, a, b) = p$ สำหรับอักขระ $a, b \in \Sigma_{\leq}$ และแทน* เซตของตัวก่อนหน้าทั้งหมด *ของสถานะ $q \in Q$ ด้วย $Pred(q)$ และแทน* เซตของตัวตามหลังทั้งหมด *ของสถานะ $q \in Q$ ด้วย $Succ(q)$*

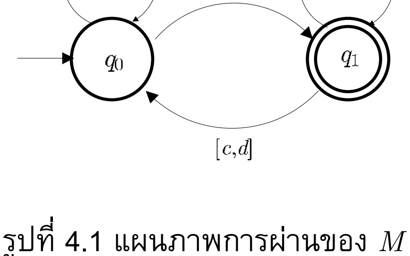
นิยาม 4.4 *ให้ δ เป็นฟังก์ชันการเปลี่ยน-เค ของออโตมาตาจำกัดเชิงกำหนดขอบ-เค $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$* จำนวนตัวตามหลังมากที่สุดของฟังก์ชันการเปลี่ยน-เค δ *แทนด้วย $\#(\delta)$ ซึ่งนิยามดังนี้*

$$\#(\delta) = max\{\ |Succ(q)|\ :\ q \in Q\ \}$$

ตัวอย่าง 4.2 *กำหนดให้ $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-2 โดยที่* $Q = \{q_0, q_1\}$, $\Sigma_{\leq} = \{a, b, c, d\}$, $F_A = \{q_0\}$ และ $F_R = \{q_1\}$ และฟังก์ชันการเปลี่ยน-*เค δ นิยามดังนี้*

$$\delta(q_0, a, c) = q_0,\ \ \delta(q_0, d, d) = q_1,\ \delta(q_1, a, b) = q_0 \text{ และ } \delta(q_1, c, d) = q_1$$

พิจารณา $\delta(q_0, d, d) = q_1$ จะ*ได้ว่าสถานะ q_0 เป็นตัวก่อนหน้าของสถานะ q_1* และ*สถานะ q_1 เป็นตัวตามหลังของสถานะ q_0 เราแสดงแผนภาพการผ่านของ M ได้ดังรูปที่ 4.1* □



รูปที่ 4.1 แผนภาพการผ่านของ M

นิยาม 4.5 *ให้* $S = S+ \cup S-$ *เป็นเซตของตัวอย่างที่ประกอบด้วยตัวอย่างบวกและตัวอย่างลบ ต้นไม้ยอมรับส่วนหน้าขยายขอบ-เคสำหรับตัวอย่าง* S *(k-edge augmented prefix tree acceptor for* S*) หมายถึง ออโตมาตาจำกัดเชิงกำหนดขอบ-เค ที่ประกอบด้วยส่วนสำคัญหกส่วน*

$$APTA = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$$

โดยที่แต่ละส่วนสำคัญเป็นไปตามเงื่อนไขดังต่อไปนี้

- $Q = \{q_u : u \in Pref(S+) \cup Pref(S-)\}$
- $q_0 = q_\lambda$
- $F_A = \{q_w \in Q : w \in S+\}$
- $F_R = \{q_w \in Q : w \in S-\}$
- *ฟังก์ชันการเปลี่ยน-เค* δ *นิยามโดย*

  $\forall u \in Pref(S+) \cup Pref(S-),\ \forall z \in \Sigma_{\leq}$

  $\delta^*(q_0, uz) = \delta(\delta^*(q_0, u), a, b) = q$

  โดยที่ $a \leq z \leq b,\ q \in Q,\ a, b, z \in \Sigma_{\leq},\ u \in \Sigma^*_{\leq}$

  ถ้า $\delta^*(q_0, u) = q_1$ *และ* $\delta^*(q_0, uz) = q_2$ *แล้ว* $\delta(q_1, a, b) = q_2$

  โดยที่ $a \leq z \leq b$ *สำหรับ* $q_0, q_1, q_2 \in Q,\ a, b, z \in \Sigma_{\leq},\ u \in \Sigma^*_{\leq}$

ตัวอย่าง 4.3 *กำหนดให้* $S+ = \{d, ad\}$ *และ* $S- = \{ac, ddb\}$ *เป็นเซตของตัวอย่างบวกและตัวอย่างลบบนเซตอักขระมีอันดับ* $\Sigma_{\leq} = \{a, b, c, d\}$ *ต้นไม้ยอมรับส่วนหน้าขยายขอบ-เคสำหรับตัวอย่าง* S *โดยที่* $S = S+ \cup S-$ *คือ* $APTA = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$

โดยที่ $Q = \{q_\lambda, q_a, q_d, q_{ac}, q_{ad}, q_{dd}, q_{ddb}\}$

$q_0 = q_\lambda$

$F_A = \{ q_d , q_{ad} \}$

$F_R = \{ q_{ac} , q_{ddb} \}$

และฟังก์ชันการเปลี่ยน-เค δ *นิยามดังนี้*

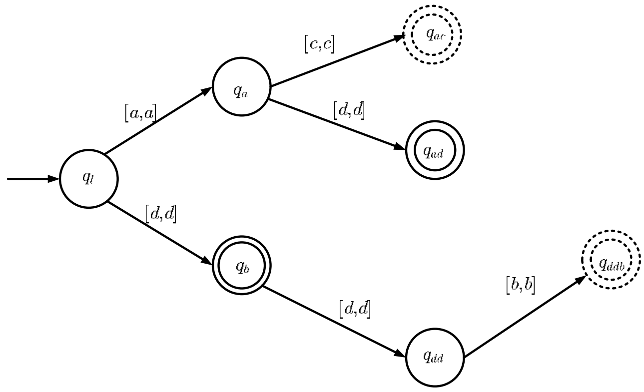
$\delta(q_\lambda, a, a) = q_a,\ \ \delta(q_\lambda, d, d) = q_d,$

$\delta(q_a, c, c) = q_{ac},\ \ \delta(q_a, d, d) = q_{ad},$

$\delta(q_d, d, d) = q_{dd},\ \ \delta(q_{dd}, b, b) = q_{ddb}$

แผนภาพการผ่านของ $APTA$ *แสดงในรูปที่ 4.2 โดยแทนสถานะยอมรับด้วยรูปวงกลมซ้อนกันสองวงที่วาดด้วยเส้นทึบ และแทนสถานะปฏิเสธด้วยรูปวงกลมซ้อนกันสองวงที่วาดด้วยเส้นประ*

□



รูปที่ 4.2 แผนภาพการผ่านของ $APTA$

นิยาม 4.6 *ให้* $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ และ $M' = (\Sigma_{\leq}, Q', q_0, F'_A, F'_R, \delta')$ *เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-เค สองตัวใด ๆ เราจะกล่าวว่า* M *เป็นออโตมาตาจำกัดเชิง*กำหนดขอบ-เคย่อย *(subautomaton) ของ* M' *ถ้าสอดคล้องกับเงื่อนไขต่อไปนี้*

(i) $Q \subseteq Q'$, $F_A \subseteq F'_A$ และ $F_R \subseteq F'_R$

(ii) สำหรับทุก $q \in Q$ และ $a, b \in \Sigma_{\leq}$ *โดยที่* $a \leq b$, $\delta(q, a, b) = \delta'(q, a, b)$

นิยาม 4.7 *ให้* $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ และ $M' = (\Sigma'_{\leq}, Q', q_0, F'_A, F'_R, \delta')$ *เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-เค สองตัวใด ๆ ให้* f *เป็นฟังก์ชันจาก* Q *ไป* Q' *เราเรียกฟังก์ชัน* f *ว่า* สาทิสสัณฐาน *(homomorphism)* ถ้า $f(\delta(q, a, b)) = \delta'(f(q), a, b)$

นิยาม 4.8 *ให้* $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ และ $M' = (\Sigma'_{\leq}, Q', q_0, F'_A, F'_R, \delta')$ *เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-เค สองตัวใด ๆ* M *เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-เค*ย่อยโดยสาทิสสัณฐาน*ของ* M' ถ้ามี สาทิสสัณฐาน $f: Q \rightarrow Q'$ *ที่สอดคล้องกับเงื่อนไขต่อไปนี้*

(i) ถ้า $q \in Q$ แล้ว $f(q) \in Q'$,

(ii) ถ้า $q \in F_A$ แล้ว $f(q) \in F'_A$,

(iii) ถ้า $q \in F_R$ แล้ว $f(q) \in F'_R$,

(iv) $\forall q \in Q$ และ $a,b \in \Sigma_{\leq}$, ถ้า $\delta(q, a, b) = p$ แล้ว $f(\delta(q, a, b)) = \delta'(f(q), a, b)$

4.2 คุณสมบัติของระดับภาษายอมรับได้-เค

ในหัวข้อนี้จะกล่าวถึงคุณสมบัติของระดับภาษายอมรับได้-เค ที่จำเป็นสำหรับการพิสูจน์ความสามารถในการเรียนรู้ในบทต่อไป

ทฤษฎีบท 4.1 *สำหรับทุกจำนวนเต็ม $k \geq 0$ ถ้า M เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-k แล้ว M เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-(k+1)*

พิสูจน์ ให้ M เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-k นิยามโดย $M= (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ จากนิยาม 4.1 จะได้ว่ามีบางสถานะ q ใน Q ที่ $|\{q' : \delta(q, x, y) = q'\}| = k$ และเนื่องจากคุณสมบัติของจำนวนเต็มที่ว่า $k < k+1$ เสมอ ทำให้ $|\{q' : \delta(q, x, y) = q'\}| < k+1$ สำหรับทุกสถานะ q ใน Q ของ M ดังนั้นจากนิยาม 4.1 เราสรุปได้ว่า M เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-$(k+1)$ ■

นิยาม 4.9 *ภาษายอมรับได้-เค (k-acceptable languages) หมายถึง ภาษาที่รู้จำได้โดยออโตมาตาจำกัดเชิงกำหนดขอบ-เค*

นิยาม 4.10 *ระดับของภาษายอมรับได้-เค (class of k-acceptable languages) หมายถึง เซตของภาษายอมรับได้-เค เขียนแทนด้วยสัญลักษณ์ $k\text{-}\mathcal{ACC}$*

ทฤษฎีบท 4.2 *สำหรับทุกจำนวนเต็ม $k \geq 0$ ระดับภาษายอมรับได้-k เป็นระดับย่อยของระดับภาษายอมรับได้-(k+1)*

พิสูจน์ ให้ $k\text{-}\mathcal{ACC}$ แทนระดับภาษายอมรับได้-k และ $(k+1)\text{-}\mathcal{ACC}$ แทนระดับภาษายอมรับได้-$(k+1)$ สำหรับการพิสูจน์ทฤษฎีบทนี้ เราจะแสดงให้เห็นว่า มีออโตมาตาจำกัดเชิงกำหนดขอบ-$(k+1)$ ที่สามารถรู้จำทุกภาษายอมรับได้-k

ให้ $\Sigma_{\leq} = \{a_1, a_2, \ldots, a_n\}$ โดยที่ $n \geq k$ และให้ $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-k ที่รู้จำภาษายอมรับได้-k ซึ่งแทนภาษานี้ด้วย L จากทฤษฎีบท 4.1 แสดงว่ามีออโตมาตาจำกัดเชิงกำหนดขอบ-$(k+1)$ ที่สามารถรู้จำภาษาภาษายอมรับได้-k L ได้ ซึ่งจะเห็นได้ชัดว่าทุก ๆ ภาษายอมรับได้-k เป็นภาษายอมรับได้-$(k+1)$ ด้วย ดังนั้นเราสรุปได้ว่า $k\text{-}\mathcal{ACC} \subseteq (k+1)\text{-}\mathcal{ACC}$ ■

ทฤษฎีบท 4.3 *สำหรับจำนวนเต็ม $n \geq 0$, $\cup_{i = 0 \text{ to } n} \Sigma^{i}_{\leq}$ เป็นภาษาในระดับ $1\text{-}\mathcal{ACC}$*

พิสูจน์ การพิสูจน์ทำได้โดยใช้วิธีอุปนัยทางคณิตศาสตร์ดังนี้

ขั้นตอนพื้นฐานอุปนัย : เราจะแสดงให้เห็นว่าในกรณี $n = 0$ ทฤษฎีบทนี้เป็นจริง

พิจารณาภาษา $L_{(0)} = \cup_{i = 0 \text{ to } 0}\Sigma^{i}_{\leq} = \{\lambda\}$ จะเห็นว่าภาษา $L_{(0)}$ สามารถรู้จำได้โดยออโตมาตาเชิงกำหนดขอบ-1 แทนด้วย $M_{(0)} = (\Sigma_{\leq}, Q_{(0)}, q_0, F_{A(0)}, F_{R(0)}, \delta_{(0)})$ โดยที่แต่ละส่วนสำคัญนิยามดังนี้ $\Sigma_{\leq} = \{a_1, a_2, \ldots , a_n\}$, $Q_{(0)} = \{q_0, q_1\}$, $F_{A(0)} = \{q_0\}$, $F_{R(0)} = \{q_1\}$ และ $\delta_{(0)} = \{(q_0, a_1, a_n, q_1), (q_1, a_1, a_n, q_1)\}$ จะเห็นว่า $\#(\delta_{(0)}) = 1$ ทำให้ $M_{(0)}$ เป็นออโตมาตาเชิงกำหนดขอบ-1 ดังนั้นจากนิยาม 4.9 แสดงว่า $L_{(0)} \in 1\text{-}\mathcal{ACC}$

ขั้นตอนการอุปนัย : สมมติให้กรณี $n = k$ เป็นจริง และจะแสดงให้เห็นว่าในกรณี $n = k+1$ ทฤษฎีบทนี้เป็นจริง

ให้ $L_{(k)} = \cup_{i=0 \text{ to } k} \Sigma^i_{\leq} \in$ 1-$\mathcal{ACC}$ เป็นจริง ซึ่งจะได้ว่ามีออโตมาตาจำกัดเชิงกำหนดขอบ-1 $M_{(k)} = (\Sigma_{\leq}, Q_{(k)}, q_0, F_{A(k)}, F_{R(k)}, \delta_{(k)})$ โดยที่ $\Sigma_{\leq} = \{a_1, a_2, \ldots , a_n\}$, $Q_{(k)} = \{q_0, q_1, \ldots q_{k+1}\}$, $F_{A(k)} = \{q_0, q_1, \ldots q_k\}$, $F_{R(k)} = \{q_{k+1}\}$ และ $\delta_{(k)} = \{(q_0, a_1, a_n, q_1), (q_1, a_1, a_n, q_2), \ldots ,(q_k, a_1, a_n, q_{k+1}), (q_{k+1}, a_1, a_n, q_{k+1})$ } ซึ่ง $\#(\delta_{(k)}) = 1$ สำหรับภาษา $L_{(k)}$

พิจารณา $L_{(k+1)} = \cup_{i=0 \text{ to } k+1}\Sigma^i_{\leq} = \cup_{i=0 \text{ to } k} \Sigma^i_{\leq} \cup \{w \in \Sigma^*_{\leq} : |w| = k+1 \}$ จะเห็นว่าออโตมาตาที่รู้จำภาษา $L_{(k+1)}$ สามารถสร้างได้จากออโตมาตา $M_{(k)}$ ดังนี้ $M_{(k+1)} = (\Sigma_{\leq}, Q_{(k+1)}, q_0, F_{A(k+1)}, F_{R(k+1)}, \delta_{(k+1)})$ โดยที่ $Q_{(k+1)} = Q_{(k)} \cup \{q_{k+2}\}$, $F_{A(k+1)} = F_{A(k)} \cup \{q_{k+1}\}$, $F_{R(k+1)} = \{q_{k+2}\}$ และ $\delta_{(k+1)} = (\delta_{(k)} - \{(q_{k+1}, a_1, a_n, q_{k+1})\}) \cup \{(q_{k+1}, a_1, a_n, q_{k+2}), \{(q_{k+2}, a_1, a_n, q_{k+2})\}\}$ ซึ่งเห็นได้ชัดว่า $\#(\delta_{(k+1)}) = 1$ ทำให้ $M_{(k+1)}$ เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-1 ดังนั้น $L_{(k+1)} \in$ 1-$\mathcal{ACC}$ ดังนั้นจากขั้นตอนพื้นฐานอุปนัยและขั้นตอนการอุปนัย เราสรุปได้ว่าสำหรับจำนวนเต็ม $n \geq 0$, $\cup_{i=0 \text{ to } n}\Sigma^i_{\leq}$ เป็นภาษาในระดับ 1-$\mathcal{ACC}$ ■

ทฤษฎีบท 4.4 *ระดับภาษา 1-$\mathcal{ACC}$ มีจุดจำกัด*

พิสูจน์ การพิสูจน์ทำได้โดยแสดงให้เห็นว่ามีลำดับอนันต์และภาษาที่เป็นจุดจำกัดซึ่งเป็นเงื่อนไขเพียงพอของการมีอยู่ของจุดจำกัดตามนิยามที่ 3.23

จากทฤษฎีบท 4.3 จะได้ว่า สำหรับทุกจำนวนเต็ม $n \geq 0$, $L_{(n)} = \cup_{i=0 \text{ to } n}\Sigma^i_{\leq}$ เป็นภาษาในระดับภาษา 1-$\mathcal{ACC}$ และจากภาษา $L_{(n+1)} = \cup_{i=0 \text{ to } n}\Sigma^i_{\leq} \cup \{w \in \Sigma^*_{\leq} : |w| = n+1 \}$ หรือ $L_{(n+1)} = L_{(n)} \cup \{w \in \Sigma^*_{\leq} : |w| = n+1 \}$ ซึ่งแสดงว่า $L_{(n)} \subset L_{(n+1)}$ สำหรับทุกจำนวนเต็ม $n \geq 0$ ดังนั้นเห็นได้ชัดว่า มีลำดับอนันต์ $L_{(0)} \subset L_{(1)} \subset L_{(2)} \subset \ldots \subset L_{(n)} \subset \ldots$ ในระดับภาษา 1-$\mathcal{ACC}$

พิจารณาภาษา $L = \Sigma^*_{\leq}$ จะเห็นว่าเราสามารถสร้างออโตมาตาจำกัดเชิงกำหนดขอบ-1 ที่รู้จำภาษา L ได้ นั่นคือ $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ โดยที่ $\Sigma_{\leq} = \{a_1, a_2, \ldots , a_n\}$, $Q = \{q_0\}$, $F_A = \{q_0\}$, $F_R = \varnothing$ และ $\delta = \{(q_0, a_1, a_n, q_0)\}$ ซึ่งจะเห็นว่า $\#(\delta) = 1$ สำหรับภาษา $L = \Sigma^*_{\leq}$ จากนิยาม 4.9 แสดงว่า $L \in$ 1-$\mathcal{ACC}$

พิจารณาจากนิยาม 3.23 จะเห็นว่าภาษา $L = \Sigma^*_{\leq}$ สอดคล้องกับเงื่อนไขที่ว่า $L = \cup_{n\in\mathbb{N}} L_{(n)}$ ดังนั้นเราสามารถสรุปได้ว่า ระดับภาษา 1-$\mathcal{ACC}$ มีจุดจำกัด โดยที่ภาษา $\Sigma^*_{\leq}$ คือจุดจำกัด ■

ทฤษฎีบท 4.5 *สำหรับจำนวนเต็ม $k \geq 1$ ระดับภาษา k-$\mathcal{ACC}$ มีจุดจำกัด*

พิสูจน์ เนื่องจากระดับภาษา 1-$\mathcal{ACC}$ มีจุดจำกัด และเห็นได้ชัดว่าระดับภาษา 1-$\mathcal{ACC}$ เป็นระดับย่อยของทุกระดับภาษา k-$\mathcal{ACC}$ สำหรับจำนวนเต็ม $k \geq 1$ (ทฤษฎีบท 4.2) ดังนั้นเราสามารถสรุปโดยใช้นิยาม 3.23 ว่าระดับภาษา k-$\mathcal{ACC}$ มีจุดจำกัด เมื่อ $k \geq 1$ ■

ผลที่ได้จากการพิสูจน์ทฤษฎีบท 4.5 จะถูกนำไปใช้ในการพิสูจน์ความสามารถในการเรียนรู้ของระดับภาษา k-$\mathcal{ACC}$ ในบทถัดไป

ในงานวิจัยนี้ นอกจากจะทำการศึกษาคุณสมบัติพื้นฐานต่าง ๆ ของระดับภาษา k-$\mathcal{ACC}$ แล้ว ยังได้ทำการศึกษาความสัมพันธ์ระหว่างระดับภาษา k-$\mathcal{ACC}$ และระดับภาษาสม่ำเสมอ โดยจะแสดงในทฤษฎีบทต่อไปนี้

ทฤษฎีบท 4.6 *สำหรับจำนวนเต็ม $k \geq 0$ ระดับภาษายอมรับได้-เค เป็นเซตย่อยของระดับภาษาสม่ำเสมอ*

พิสูจน์ ให้ k-$\mathcal{ACC}$ แทนระดับภาษายอมรับได้-เค และ $\mathcal{REG}$ แทนระดับภาษาสม่ำเสมอ

การพิสูจน์ทำได้โดยแสดงให้เห็นว่า ทุกภาษาในระดับภาษา k-$\mathcal{ACC}$ เป็นภาษาในระดับภาษา $\mathcal{REG}$ นั่นก็คือ ถ้า L เป็นภาษาใดที่รู้จำด้วยออโตมาตาจำกัดเชิงกำหนดขอบ-เค แล้ว L สามารถรู้จำได้ด้วยออโตมาตาจำกัดเชิงกำหนด

ให้ $M_k = (\Sigma_\leq, Q, q_0, F_A, F_R, \delta_k)$ นิยามโดย $\Sigma_\leq = \{a_1, a_2, \ldots , a_n\}$, $\delta_k = \{(p, a, b, q) : p, q \in Q$ และ $a, b \in \Sigma_\leq\}$ ซึ่ง $\#(\delta_k) = k$ เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-เค ที่สามารถรู้จำภาษา L ได้ เราสามารถสร้างออโตมาตาจำกัดเชิงกำหนดที่สามารถรู้จำภาษา L ได้เสมอ โดยนิยามออโตมาตาจำกัดเชิงกำหนดดังนี้ $M = (\Sigma, Q, q_0, F_A, F_R, \delta)$ โดยที่ $\Sigma = \{a : a \in \Sigma_\leq\}$ และ $\delta = \{(p, z ,q) : a \leq z \leq b$ สำหรับทุก $(p, a, b, q) \in \delta_k\}$ จะเห็นว่าออโตมาตาจำกัดเชิงกำหนด M สามารถรู้จำภาษา L ได้ ดังนั้น เราสามารถสรุปได้ว่า k-$\mathcal{ACC} \subseteq \mathcal{REG}$ สำหรับจำนวนเต็ม $k \geq 0$ ■

ทฤษฎีบท 4.7 $\cup_{k=0}$ k-$\mathcal{ACC} = \mathcal{REG}$ *โดยที่* k-$\mathcal{ACC}$ *แทนระดับภาษายอมรับได้-เค และ* $\mathcal{REG}$ *แทนระดับภาษาสม่ำเสมอ*

พิสูจน์ การพิสูจน์ทำได้แสดงให้เห็นว่า $\cup_{k=0}$ k-$\mathcal{ACC} \subseteq \mathcal{REG}$ และ $\mathcal{REG} \subseteq \cup_{k=0}$ k-$\mathcal{ACC}$

กำหนดให้ $M = (\Sigma, Q, q_0, F_A, F_R, \delta)$ ออโตมาตาจำกัดเชิงกำหนดตัวหนึ่งที่รู้จำภาษา L แต่ละส่วนสำคัญของ M นิยามโดย $\Sigma = \{a_1, a_2, \ldots , a_n\}$, $\delta = \{(p, a, q) : p, q \in Q$ และ $a \in \Sigma\}$ เราสามารถสร้างออโตมาตาจำกัดเชิงกำหนดขอบ-k เมื่อ k เป็นจำนวนเต็มใด ๆ แทนด้วย $M_k = (\Sigma_\leq, Q, q_0, F_A, F_R, \delta_k)$ ซึ่งแต่ละส่วนสำคัญของ M_k นิยามดังนี้ $\Sigma_\leq = \{a_i : a_i \in \Sigma\}$ และ $\delta_k = \{(p, a, a, q) :$ สำหรับทุก $(p, a, q) \in \delta\}$ จะเห็นว่าออโตมาตาจำกัดเชิงกำหนด M_k

สามารถรู้จำภาษา L ได้ ดังนั้น เราสามารถสรุปได้ว่า $\mathcal{REG} \subseteq \cup_{k=0}$ k-$\mathcal{ACC}$ สำหรับจำนวนเต็ม $k \geq 0$ และจากทฤษฎีบท 4.6 ที่ได้แสดงว่าสำหรับจำนวนเต็ม $k \geq 0$, k-$\mathcal{ACC} \subseteq \mathcal{REG}$ ดังนั้นเราสามารถสรุปได้ว่า $\cup_{k=0}$ k-$\mathcal{ACC} = \mathcal{REG}$ ■

4.3 สรุป

ในบทนี้ เราได้กล่าวถึงนิยามของภาษายอมรับได้-เค ออโตมาตาเชิงกำหนดขอบ-เค และสาทิสสัณฐานของออโตมาตาเชิงกำหนดขอบ-เค และนิยามอื่น ๆ ที่จำเป็นต้องใช้ในการพิสูจน์เรื่องความสามารถในการเรียนรู้ของระดับภาษายอมรับได้-เคในบทที่ 5 นอกจากนั้นเราได้ทำการพิสูจน์ทฤษฎีบทเกี่ยวกับคุณสมบัติของระดับภาษายอมรับได้-เค ซึ่งได้แก่ ทฤษฎีบทที่แสดงความสัมพันธ์ระหว่างระดับภาษายอมรับได้-เค เมื่อกำหนดค่าเคแตกต่างกัน (ทฤษฎีบท 4.1 และ 4.2) และได้พิสูจน์ให้เห็นว่าระดับภาษายอมรับได้-เคมีจุดจำกัด (ทฤษฎีบท 4.3, 4.4 และ 4.5) อันเป็นเงื่อนไขที่เพียงพอในการพิสูจน์ว่าระดับภาษายอมรับได้-เคไม่สามารถเรียนรู้ได้จากการนำเสนอด้วยตัวอย่างบวกเพียงอย่างเดียว รวมถึงได้พิสูจน์ให้เห็นว่าระดับภาษายอมรับได้-เคเป็นระดับย่อยของภาษาสม่ำเสมอ (ทฤษฎีบท 4.6 และ 4.7)

บทที่ 5

ความสามารถในการเรียนรู้ของระดับภาษายอมรับได้-เค

ในบทนี้จะศึกษาคุณสมบัติ ความสามารถในการเรียนรู้ของระดับภาษายอมรับได้-เค โดยแบ่งการศึกษาออกเป็นสองส่วน คือ ส่วนแรกจะศึกษาความสามารถในการเรียนรู้ของระดับภาษายอมรับได้-เค ในกรณีที่มีตัวอย่างบวกเพียงอย่างเดียว และส่วนที่สองจะศึกษาความสามารถในการเรียนรู้ของระดับภาษายอมรับได้-เค ในกรณีที่มีทั้งตัวอย่างบวกและตัวอย่างลบ ผลที่ได้จากการศึกษาจะแสดงในรูปแบบของทฤษฎีบทและบทพิสูจน์

5.1 การเรียนรู้จากตัวอย่างบวกเพียงอย่างเดียว

การพิสูจน์ว่าระดับของภาษาใด ๆ สามารถเรียนรู้ได้จากตัวอย่างบวกเพียงอย่างเดียวหรือไม่นั้น ในกรณีที่จะพิสูจน์ว่าระดับของภาษานั้นไม่สามารถเรียนรู้ได้จากตัวอย่างบวกเพียงอย่างเดียวทำได้หลายวิธี อาทิเช่น การแสดงให้เห็นว่าไม่มีอัลกอริทึมการเรียนรู้ใดที่สามารถเรียนรู้ทุกภาษาในระดับนั้นได้ ซึ่งเป็นวิธีที่ทำได้ยาก วิธีหนึ่งที่นิยมใช้กันก็คือ การแสดงว่าระดับของภาษาเป็นระดับเกินจำกัด (ตามนิยาม 3.9) หรือ การแสดงว่าระดับของภาษามีจุดจำกัด (ตามนิยาม 3.23) ซึ่งเป็นเงื่อนไขเพียงพอในการสรุปว่าระดับภาษานั้นไม่สามารถเรียนรู้ได้จากตัวอย่างบวกเพียงอย่างเดียว แต่ในทางกลับกันถ้าต้องการแสดงว่าระดับของภาษาสามารถเรียนรู้ได้นั้น ต้องแสดงให้เห็นว่ามีอย่างน้อยหนึ่งอัลกอริทึมที่สามารถเรียนรู้ทุกภาษาที่อยู่ในระดับของภาษานั้นได้

จากการศึกษาคุณสมบัติต่าง ๆ ของระดับภาษายอมรับได้-เคในบทที่ 4 ซึ่งได้พิสูจน์ว่าระดับภาษายอมรับได้-เคมีจุดจำกัด ดังนั้นระดับภาษายอมรับได้-เคจึงไม่สามารถเรียนรู้ได้จากตัวอย่างบวกเพียงอย่างเดียว ซึ่งแสดงได้ตามทฤษฎีบทต่อไปนี้

ทฤษฎีบท 5.1 *สำหรับจำนวนเต็ม $k \geq 1$ ระดับภาษายอมรับได้-เคไม่สามารถเรียนรู้ได้ในขอบเขตจำกัดจากตัวอย่างบวกเพียงอย่างเดียว*

พิสูจน์ การพิสูจน์ทำได้โดยแสดงให้เห็นว่าระดับภาษายอมรับได้-เคมีจุดจำกัด ซึ่งเป็นเงื่อนไขที่เพียงพอในการแสดงให้เห็นว่าระดับภาษานั้นไม่สามารถเรียนรู้ได้ในขอบเขตจำกัดจากตัวอย่างบวกเพียงอย่างเดียว

ให้ k-ACC แทนระดับภาษายอมรับได้-เค เนื่องจากทฤษฎีบท 4.5 ที่แสดงว่าสำหรับจำนวนเต็ม $k \geq 1$ ระดับภาษา k-ACC มีจุดจำกัด และจากทฤษฎีบท 3.5 ที่กล่าวว่า “ถ้าระดับภาษาใดมีจุดจำกัด แล้วระดับภาษานั้นไม่สามารถเรียนรู้ได้ในขอบเขตจำกัดจากตัวอย่างบวกเพียง

อย่างเดียว” ดังนั้นเราสามารถสรุปได้ว่า สำหรับจำนวนเต็ม $k \geq 1$ ระดับภาษา k-$\mathcal{ACC}$ ไม่สามารถเรียนรู้ได้ในขอบเขตจำกัดจากตัวอย่างบวกเพียงอย่างเดียว ■

5.2 การเรียนรู้จากตัวอย่างบวกและตัวอย่างลบ

ในหัวข้อนี้เราจะพิสูจน์ให้เห็นว่าระดับของภาษายอมรับได้-เค สามารถเรียนรู้ได้ในขอบเขตจำกัดจากตัวอย่างบวกและตัวอย่างลบ และนอกเหนือไปกว่านั้นระดับของภาษายอมรับได้-เคสามารถเรียนรู้ได้ในขอบเขตจำกัดอย่างมีประสิทธิภาพจากตัวอย่างบวกและตัวอย่างลบอีกด้วย ซึ่งในการพิสูจน์จำเป็นต้องแสดงให้เห็นว่ามีอัลกอริทึมการเรียนรู้ตามนิยาม 3.21 ดังนั้นในหัวข้อนี้จะเสนออัลกอริทึมการเรียนรู้สำหรับระดับภาษายอมรับได้-เค และได้แสดงให้เห็นว่ามีตัวอย่างลักษณะ ซึ่งเป็นตัวอย่างที่ทำให้การเรียนรู้สำเร็จสำหรับการเรียนรู้ภาษายอมรับได้-เค

5.2.1 อัลกอริทึมการเรียนรู้ภาษายอมรับได้-เค

อัลกอริทึมสำหรับเรียนรู้ภาษายอมรับได้-เค ที่เสนอในงานวิจัยนี้เรียกว่า อัลกอริทึมเคเอซีแอลไอ ($KACLI$ หรือ *k*-acceptable language inference) ซึ่งเป็นอัลกอริทึมการเรียนรู้โดยการอนุมานออโตมาตาจำกัดเชิงกำหนดขอบ-เคจากตัวอย่างบวกและตัวอย่างลบ อัลกอริทึมเคเอซีแอลไอ พัฒนามาจากแนวคิดของอัลกอริทึมการเรียนรู้อาร์พีเอ็นไอ อีดีเอสเอ็ม และบลูฟริงก์ ซึ่งเป็นอัลกอริทึมการเรียนรู้ของระดับภาษาสม่ำเสมอที่รู้จักกันดีในงานวิจัยทางด้านการอนุมานเชิงไวยากรณ์ การทำงานของอัลกอริทึมเคเอซีแอลไอประกอบด้วยขั้นตอนต่าง ๆ คือ การสร้างตัวยอมรับต้นไม้ส่วนหน้าขยายขอบ-เค การเลือกสถานะ การลดจำนวนตัวตามหลัง การผสานสถานะ การผสานสถานะแบบวนซ้ำ และการตรวจสอบความสอดคล้อง ขั้นตอนเหล่านี้แสดงด้วยอัลกอริทึมดังต่อไปนี้

อัลกอริทึม 5.1 *อัลกอริทึมบิวเอพีทีเอ สำหรับการสร้างตัวยอมรับต้นไม้ส่วนหน้าขยายขอบ-เค*
Algorithm : *BUILD_APTA*

Input : a set of positive and negative examples $S = (S+,\ S-)$

Output : a *k*-APTA : $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$

$q_0 \leftarrow q_\lambda$

$Q \leftarrow \{q_u : u \in Pref(S)\}$

FOR $q_u \in Q$ **DO**

  FOR $a \in \Sigma_{\leq}$ **DO**

    $\delta(q_u, a, a) \leftarrow q_{ua}$

  ENDFOR

ENDFOR

FOR $q_u \in Q$ **DO**

IF $u \in S+$ **THEN**

$F_A \leftarrow F_A \cup \{q_u\}$

ENDIF

IF $u \in S-$ **THEN**

$F_R \leftarrow F_R \cup \{q_u\}$

ENDIF

ENDFOR

คำอธิบายอัลกอริทึม 5.1 อัลกอริทึมบิวเอพีทีเอมีข้อมูลรับเข้า คือเซตของตัวอย่างบวกและตัวอย่างลบ และมีข้อมูลส่งออกคือออโตมาตาที่เรียกว่าต้นไม้ยอมรับส่วนหน้าขยายขอบ-เค ซึ่งลักษณะของออโตมาตาแบบนี้คือ ออโตมาตาจะทำการยอมรับทุกสายอักขระในเซตของตัวอย่างบวก และปฏิเสธทุกสายอักขระในเซตของตัวอย่างลบเท่านั้น

อัลกอริทึม 5.2 *อัลกอริทึมเมคซีเคียว สำหรับการลดจำนวนตัวตามหลังของสถานะที่กำหนด*

Algorithm : *MAKE_SECURE*

Input : a k-DFA : $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ and a state q

Output : a k-DFA : $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ that the state q is secured

$q_\alpha \leftarrow q_u : u = min_{\leq}\{SP(q') : \delta(q, a, a) = q'\}$

FOR $q_\beta \in \delta(q_\alpha, a, a) : \alpha \leq \beta$ **DO**

IF $(q_\alpha \in F_A \wedge q_\beta \in F_R) \vee (q_\beta \in F_A \wedge q_\alpha \in F_R)$ **THEN**

$q_\alpha \leftarrow q_\beta$

ELSE

$M \leftarrow REC\_MERGE(M, q_\alpha, q_\beta)$

ENDIF

ENDFOR

คำอธิบายอัลกอริทึม 5.2 อัลกอริทึมเมคซีเคียวมีข้อมูลรับเข้าสองตัวคือ คือออโตมาตาจำกัดเชิงกำหนดขอบ-เค $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ และสถานะ $q \in Q$ ส่วนข้อมูลส่งออกของอัลกอริทึมนี้คือออโตมาตาจำกัดเชิงกำหนดขอบ-เค โดยที่สถานะ q มีจำนวนตัวตามหลังไม่เกิน k ตัว อัลกอริทึมนี้ใช้สำหรับลดจำนวนขอบของสถานะ q ที่กำหนดให้มากที่สุด ก่อนที่จะทำการผสานสถานะ โดยทำการผสานตัวตามหลังของสถานะ q ทุกตัวที่ไม่ขัดแย้งเข้าด้วยกัน (สถานะยอมรับทำการผสานกับสถานะปฏิเสธไม่ได้) โดยพิจารณาจากอันดับเชิงพจนานุกรมและความยาวจากน้อยไปมาก

อัลกอริทึม 5.3 *อัลกอริทึมชูส สำหรับการเลือกสถานะ*

Algorithm : *CHOOSE*

Input : a k-DFA : $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$

Output : the state q_α that the shortest string is reached

$q_\alpha \leftarrow q_\lambda$
FOR $q_u \in Q$ **DO**
  IF $SP(q_u) \leq SP(q_\alpha)$ **THEN**
    $q_\alpha \leftarrow q_u$
  ENDIF
ENDFOR

คำอธิบายอัลกอริทึม 5.3 อัลกอริทึมชูส มีข้อมูลรับเข้าคือ คือ ออโตมาตาจำกัดเชิงกำหนดขอบ-เค $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ และข้อมูลส่งออกคือ สถานะ $q \in Q$ อัลกอริทึมนี้ใช้สำหรับเลือกสถานะในออโตมาตาจำกัดเชิงกำหนดขอบ-เคที่มีสายอักขระสั้นที่สุดเข้าถึง โดยการพิจารณาอันดับเชิงพจนานุกรมและความยาว

อัลกอริทึม 5.4 อัลกอริทึมเมิร์จ สำหรับการผสานสถานะ

Algorithm : *MERGE*

Input : a k-DFA : $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ and two states : q_u, q_v

Output : a k-DFA : $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$

FOR $q \in Q$ **DO**
  FOR $a, b \in \Sigma_{\leq} : a \leq b$ DO
    IF $q_v \in \delta(q, a, b)$ **THEN**
      $\delta(q, a, b) \leftarrow \delta(q, a, b) \cup \{q_u\}$
    ENDIF
  ENDFOR
ENDFOR
FOR $a, b \in \Sigma_{\leq} : \delta(q_u, a, b) \neq \varnothing$ **DO**
  FOR $a \leq z \leq b$ **DO**
    $\delta(q_u, z, z) \leftarrow \delta(q_u, a, b)$
  ENDFOR
  $\delta(q_u, a, b) \leftarrow \varnothing$
ENDFOR
FOR $a, b \in \Sigma_{\leq} : \delta_k (q_v, a, b) \neq \varnothing$ **DO**

```
        FOR a ≤ z ≤ b DO
            δ(q_v, z, z) ← δ(q_v, a, b)
        ENDFOR
        δ(q_v, a, b) ← ∅
ENDFOR
FOR q∈Q  DO
        FOR z ∈ Σ≤ DO
            IF q ∈ δ(q_v, z, z) THEN
                δ(q_u, z, z) ← δ(q_u, z, z) ∪ {q}
            ENDIF
        ENDFOR
ENDFOR
FOR  a, b ∈ Σ≤ : δ(q_u, a, b) ≠ ∅ DO
        FOR c, d ∈ Σ≤ : δ(q_u, c, d) ≠ ∅ DO
            IF  δ(q_u, a, b) = δ(q_u, c, d) THEN
                δ(q_u, a, d) ← δ(q_u, a, b)
                δ(q_u, a, b) ← ∅
                δ(q_u, c, d) ← ∅
            ENDIF
        ENDFOR
ENDFOR
IF  q_v ∈ F_A THEN
        F_A ← F_A ∪ { q_u }
ENDIF
IF  q_v ∈ F_R THEN
        F_R ← F_R ∪ { q_u }
ENDIF
Q ← Q – { q_v }
```

คำอธิบายอัลกอริทึม 5.4 อัลกอริทึมเมิร์จ มีข้อมูลรับเข้าคือ คือออโตมาตาจำกัดเชิงกำหนดขอบ-เค $M = (\Sigma_{\le}, Q, q_0, F_A, F_R, \delta)$ และสถานะ $q_u, q_v \in Q$ ข้อมูลส่งออกคือออโตมาตาจำกัด

ขอบ-เค $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ ซึ่งได้ทำการผสานสถานะ q_v เข้ากับสถานะ q_u แล้ว อัลกอริทึมนี้ใช้สำหรับผสานสถานะสองสถานะ q_u และ q_v เข้าด้วยกัน โดยนำสถานะ q_v เข้ามารวมกับสถานะ q_u สถานะใดในออโตมาตาจำกัดเชิงกำหนดขอบ-เค ที่มีสถานะ q_v เป็นตัวตามหลังให้นำสถานะ q_u มาเป็นตัวตามหลังแทน ส่วนตัวตามหลังทั้งหมดของสถานะ q_v ให้มารวมเป็นตัวตามหลังของสถานะ q_u ออโตมาตาจำกัดขอบ-เคที่ได้จากอัลกอริทึมนี้อาจเป็นออโตมาตาจำกัดเชิงกำหนด หรือเป็นออโตมาตาจำกัดเชิงไม่กำหนดก็ได้

อัลกอริทึม 5.5 อัลกอริทึมเรคเมิร์จ สำหรับการผสานสถานะแบบวนซ้ำ

Algorithm : *REC_MERGE*

Input : a k-DFA : $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ and two states : q_u, q_v

Output : a k-DFA : $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$

1: $M \leftarrow MERGE(M, q_u, q_v)$
2: **FOR** $a, b \in \Sigma_{\leq}$ **DO**
3: **IF** $|\delta(q_u, a, b)| > 0$ **THEN**
4: $M \leftarrow REC\_MERGE(M, q, q') : q, q' \in \delta(q_u, a, b)$
5: **ENDIF**
6: **ENDFOR**

คำอธิบายอัลกอริทึม 5.5 อัลกอริทึมเรคเมิร์จ มีข้อมูลรับเข้าคือ คือออโตมาตาจำกัดเชิงกำหนดขอบ-เค $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ และสถานะ $q_u, q_v \in Q$ ข้อมูลส่งออกคือออโตมาตาจำกัดเชิงกำหนดขอบ-เค $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ ซึ่งผสานสถานะ q_v เข้ากับสถานะ q_u แล้ว อัลกอริทึมนี้ใช้สำหรับผสานสถานะ 2 สถานะใด ๆ ในออโตมาตาจำกัดเชิงกำหนดขอบ-เคที่ถูกเลือกมาทำการผสานเข้าด้วยกันเป็นสถานะเดียว โดยทำการผสานแบบวนซ้ำ (recursively merging) นั่นก็คือ ถ้าออโตมาตาที่ได้จากการผสานสถานะเกิดความไม่กำหนดแน่ (nondeterminism) ต้องทำการผสานคู่ของสถานะที่ทำให้ออโตมาตาที่เป็นสาเหตุของความไม่กำหนดแน่นี้ ฟังก์ชันนี้จะถูกเรียกแบบวนซ้ำไปเรื่อยจนกระทั่งออโตมาตาที่ได้มีคุณสมบัติความกำหนดแน่ (determinism property)

อัลกอริทึม 5.6 อัลกอริทึมคอมแพททิเบิล สำหรับตรวจสอบความสอดคล้อง

Algorithm : *COMPATIBLE*

Input : a k-DFA : $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ and
a set of positive and negative examples $S = (S+, S-)$

Output : *TRUE* or *FALSE*

1: **FOR** $w \in S+$ **DO**
2: **IF** $\delta^*(q_0, w) \cap F_R \neq \varnothing$ **THEN**
3: **RETURN** *FALSE*

ENDIF

ENDFOR

FOR $w \in S-$ **DO**

IF $\delta^*(q_0, w) \cap F_A \neq \varnothing$ **THEN**

RETURN *FALSE*

ENDIF

ENDFOR

RETURN *TRUE*

คำอธิบายอัลกอริทึม 5.6 อัลกอริทึมคอมแพททิเบิล มีข้อมูลรับเข้าคือ คือออโตมาตาจำกัดเชิงกำหนดขอบ-เค $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ และเซตของตัวอย่างบวกและตัวอย่างลบ ข้อมูลส่งออกคือค่าความจริง (จริงหรือเท็จ) อัลกอริทึมนี้ใช้สำหรับตรวจสอบความสอดคล้องของออโตมาตาจำกัดเชิงกำหนดขอบ-เคที่ได้รับจากการผสานสถานะ กับเซตของตัวอย่างการเรียนรู้ที่กำหนดให้ ถ้าออโตมาตายอมรับสายอักขระทุกตัวในเซตของตัวอย่างบวก และปฏิเสธสายอักขระทุกตัวในตัวอย่างลบ แล้วจะถือว่าออโตมาตานั้นสอดคล้องกับตัวอย่างก็จะส่งค่าจริงออกมา แต่ในทางกลับกันจะส่งค่าเท็จออกมา

อัลกอริทึม 5.7 *อัลกอริทึมเคเอซีแอลไอ สำหรับการเรียนรู้ภาษายอมรับได้-เค*

Algorithm : *KACLI*

Input : a set of positive and negative examples $S = (S+, S-)$

Output : a k-DFA : $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$

$M \leftarrow BUILD\_APTA(S)$

$K \leftarrow \varnothing$

WHILE $Q - K \neq \varnothing$ **DO**

$q_\alpha \leftarrow CHOOSE(Q - K)$

$K \leftarrow K \cup \{q_\alpha\}$

$M \leftarrow MAKE\_SECURE(M, q_\alpha)$

FOR $q_\beta \in \delta(q_\alpha, a, b) : a, b \in \Sigma_{\leq}$ **DO**

FOR $q_\omega \in K$ **DO**

IF $COMPATIBLE(REC\_MERGE(M, q_\omega, q_\beta), S)$ = TRUE **THEN**

$M \leftarrow REC\_MERGE(M, q_\omega, q_\beta)$

ENDIF

ENDFOR

$K \leftarrow K \cup \{q_\beta\}$

ENDFOR

ENDWHILE

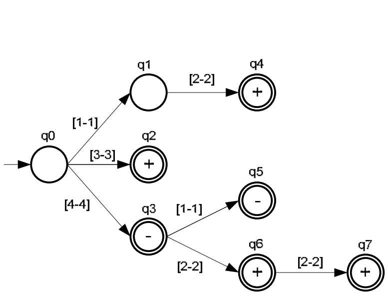
คำอธิบายอัลกอริทึม 5.7 อัลกอริทึมเคเอซีแอลไอมีข้อมูลรับเข้าคือเซตของตัวอย่างบวกและตัวอย่างลบ และมีข้อมูลส่งออกคือออโตมาตาจำกัดเชิงกำหนดขอบ-เค การทำงานของอัลกอริทึมนี้เริ่มต้นด้วยการสร้างต้นไม้ยอมรับส่วนหน้าขยายขอบ-เค (อัลกอริทึม 5.1) และใช้หลักการของการทำให้ออโตมาตาที่ได้นั้นเป็นทั่วไปมากขึ้น (generalization) จากต้นไม้ยอมรับส่วนหน้าขอบ-เค ที่มีความจำเพาะ (specification) กับตัวอย่างที่ได้รับมาเท่านั้น ขั้นตอนการทำให้เป็นทั่วไปนี้สามารถทำได้โดยวิธีการผสานสถานะให้เหลือจำนวนน้อยที่สุด (อัลกอริทึม 5.4 และ 5.5) ซึ่งออโตมาตาที่ได้จากการผสานยังคงสอดคล้องกับตัวอย่างที่กำหนดให้ แต่เนื่องจากออโตมาตาที่เราต้องเรียนรู้นั้นจะต้องมีจำนวนขอบของทุกสถานะไม่เกินค่า k ที่กำหนดไว้ ดังนั้นแนวคิดที่ใช้ในการแก้ปัญหาของอัลกอริทึมนี้คือ การลดจำนวนขอบของสถานะให้มากที่สุดก่อนที่จะทำการผสานสถานะ (อัลกอริทึม 5.3) สำหรับการเลือกคู่ของสถานะสำหรับการผสานนั้นเราจะทำการเลือกโดยพิจารณาอันดับเชิงพจนานุกรมและความยาว (อัลกอริทึม 5.2) และส่วนการกำหนดเงื่อนไขสำหรับการผสานสถานะเราจะใช้การพิจารณาว่าออโตมาตาที่ได้นั้นสอดคล้องกับตัวอย่างที่ได้รับหรือไม่ (อัลกอริทึม 5.6)

ตัวอย่าง 5.1 *กำหนดให้ภาษา* $L = ((1+2+3)+(4+5)1^*(2+3+4+5))^*$ *และ* $S+ = \{3, 12, 42, 422\}$ *เป็นเซตของตัวอย่างบวก และให้* $S- = \{4, 41\}$ *เป็นเซตของตัวอย่างลบของภาษา* L *ที่นิยามบนเซตอักขระที่มีอันดับ* $\Sigma_{\leq} = \{1, 2, 3, 4, 5\}$ *จงสร้างออโตมาตาจำกัดเชิงกำหนดขอบ-เค*

วิธีทำ ให้ $S = (S+, S-)$ แทนเซตของตัวอย่างบวกและตัวอย่างลบเป็นข้อมูลรับเข้าสำหรับอัลกอริทึมเคเอซีแอลไอ การหาออโตมาตาจำกัดเชิงกำหนดขอบ-เค จากตัวอย่าง S ทำได้ดังนี้

ขั้นตอนเริ่มต้นของเคเอซีแอลไอ คือ การสร้างออโตมาตาจำกัดเชิงกำหนดขอบ-เค $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ ตามอัลกอริทึมบิวเอพีทีเอ โดยที่ M เป็นต้นไม้ยอมรับส่วนหน้าขยายขอบ-เค สำหรับตัวอย่าง S ซึ่ง M ต้องยอมรับทุกสายอักขระในเซตของตัวอย่างบวก $S+$ และปฏิเสธทุกสายอักขระในเซตของตัวอย่างลบ $S-$ เท่านั้น โดยที่แต่ละสถานะ $q \in Q$ ของออโตมาตา M สร้างได้จากเซตของสายอักขระส่วนหน้าของ $S+$ และ $S-$ นั่นก็คือ $Pref(S+) \cup Pref(S-)$ โดยที่ $Pref(S+) = \{\lambda, 1, 3, 4, 12, 41, 42, 422\}$ และ $Pref(S) = \{\lambda, 4, 41\}$ ออโตมาตายอมรับส่วนหน้าขยายขอบ-เค M สำหรับเซตของตัวอย่าง S ที่ได้แสดงดังรูปที่ 5.1

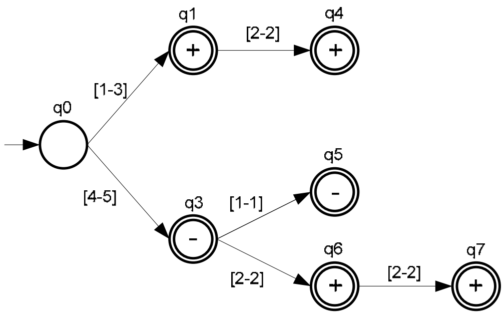
□



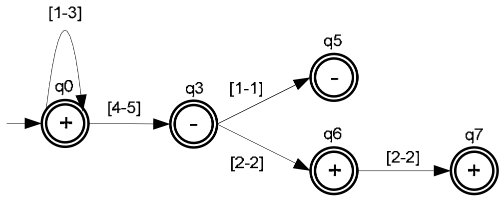
รูปที่ 5.1 ต้นไม้ยอมรับส่วนหน้าขยายขอบ-เค

ขั้นตอนต่อไปเป็นขั้นตอนการทำออโตมาตาจำกัดเชิงกำหนดขอบ-เค ให้มีความเป็นทั่วไปมากขึ้น (generalization) จากต้นไม้ยอมรับส่วนหน้าขยายขอบ-เคที่มีความจำเพาะ (specific) กับตัวอย่างเท่านั้น ขั้นตอนนี้ทำได้โดยวิธีการผสานสถานะให้เหลือจำนวนน้อยที่สุด โดยที่ยังคงสอดคล้องกับตัวอย่างที่กำหนดให้

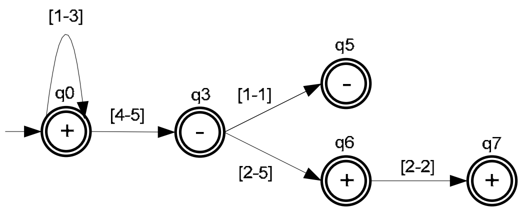
จากออโตมาตาเริ่มต้น M ทำการเลือกสถานะโดยฟังก์ชันชูส ได้สถานะ $q_\alpha = q_0$ จากนั้นทำการจำนวนลดสถานะโดยฟังก์ชันเมคซีเคียว ซึ่งทำการผสานสถานะที่เป็นตัวตามหลังของสถานะ q_0 ซึ่งก็คือ สถานะ q_1, q_2 และ q_3 ผลที่ได้จากการผสานแสดงดังรูปที่ 5.2 (ก) ขั้นตอนต่อไปคือการผสานตัวตามหลังของสถานะที่ถูกเลือกกับสถานะก่อนหน้า ถ้าไม่สามารถผสานได้ให้ทำการเลือกสถานะต่อไป การตัดสินใจว่าจะผสานได้หรือไม่นั้นพิจารณาจากฟังก์ชันคอมแพททิเบิล สำหรับในตัวอย่างนี้ สถานะ q_1 สามารถผสานเข้ากับสถานะ q_0 ออโตมาตาที่ได้จากการผสานสอดคล้องกับตัวอย่าง S แสดงดังรูปที่ 5.2 (ข) จากนั่นเลือกสถานะ $q_\alpha = q_1$ และทำการลดสถานะของตัวตามหลังของ q_1 แสดงดังรูปที่ 5.2 (ค) ทำแบบนี้ต่อไปจนกระทั่งทุกสถานะถูกเลือก จะได้ออโตมาตาจำกัดเชิงกำหนดขอบ-เค ดังรูปที่ 5.2 (ง)



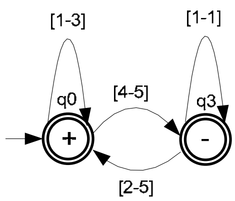
(ก)



(ข)



(ค)



(ง)

รูปที่ 5.2 การทำงานของอัลกอริทึมเคเอซีแอลไอ

ทฤษฎีบท 5.2 *อัลกอริทึมเคเอซีแอลไอ ทำงานด้วยเวลา $\mathcal{O}(p(||S||))$ โดยที่ $S = (S+, S-)$ เป็นเซตของตัวอย่างบวกและตัวอย่างลบ และ $p(x)$ เป็นฟังก์ชันพหุนามของ x*

พิสูจน์ พิจารณาจากอัลกอริทึมเคเอซีแอลไอจะเห็นว่าเวลาทำงานของอัลกอริทึมนี้ประกอบด้วย

(i) เวลาสำหรับการสร้างต้นไม้ยอมรับส่วนหน้าขยายขอบ-เค จากตัวอย่าง S ด้วยฟังก์ชัน $BUILD\_APTA$ การทำงานของฟังก์ชันนี้อธิบายด้วยอัลกอริทึมบิวเอพีทีเอ (อัลกอริทึม 5.1) ซึ่งเวลาทำงานขึ้นอยู่กับการเชื่อมโยงแต่ละสถานะ $q \in Q$ เข้าด้วยกันเพื่อสร้างฟังก์ชันการเปลี่ยน-เค δ ซึ่ง $Q = \{q_u : u \in Pref(S)\}$ จะเห็นว่าจำนวนสมาชิกของเซต Q มีอย่างมากเท่ากับจำนวนสายอักขระส่วนหน้าในเซต S นั่นก็คือ $\mathcal{O}(||S||)$ ดังนั้นเวลาการทำงานของฟังก์ชันนี้คือ $\mathcal{O}(||S||^2)$

(ii) เวลาการหาออโตมาตาจำกัดเชิงกำหนดขอบ-เค ในลูป WHILE ซึ่งจำนวนรอบการวนซ้ำทั้งหมดไม่เกินจำนวนสถานะของ Q นั่นก็คือ $\mathcal{O}(||S||)$ ครั้ง โดยที่แต่ละครั้งของการทำซ้ำประกอบด้วยการทำงานของฟังก์ชันต่าง ๆ ซึ่งใช้เวลาดังนี้

- ฟังก์ชัน $CHOOSE$: ใช้เวลา $\mathcal{O}(||S||^2)$
- ฟังก์ชัน $MAKE\_SECURE$: ใช้เวลา $\mathcal{O}(|\Sigma_{\leq}|)$
- ฟังก์ชัน REC_MERGE : ใช้เวลา $\mathcal{O}(||S||^2)$
- ฟังก์ชัน $COMPATIBLE$: ใช้เวลา $\mathcal{O}(||S||^2)$

ดังนั้น เวลาในการทำงานทั้งหมดของอัลกอริทึมเคเอซีแอลไอ คือ $\mathcal{O}(||S||^4)$ ■

5.2.2 ตัวอย่างลักษณะสำหรับการเรียนรู้ภาษายอมรับได้-เค

ในทฤษฎีการเรียนรู้ภาษา ตัวอย่างลักษณะเป็นตัวอย่างที่ทำให้การเรียนรู้ประสบความสำเร็จ หรือกล่าวอีกอย่างหนึ่งได้ว่า เป็นตัวอย่างที่ทำให้ผู้เรียนสามารถระบุตัวแทนเชิงไวยากรณ์ที่สมมูลกับตัวแทนเชิงไวยากรณ์ที่แท้จริงของภาษาได้ การมีอยู่จริง (existence) ของตัวอย่างลักษณะ จึงถูกใช้เป็นเงื่อนไขที่เพียงพอในการพิสูจน์ว่าระดับภาษาสามารถเรียนรู้ได้จากตัวอย่างที่กำหนดให้ ในหัวข้อนี้เราจะแสดงให้เห็นว่ามีตัวอย่างลักษณะสำหรับระดับภาษายอมรับได้-เค เพื่อใช้ในการพิสูจน์ว่าระดับภาษายอมรับได้-เคสามารถเรียนรู้ได้จากตัวอย่างที่กำหนดให้ ซึ่งนำไปสู่การพิสูจน์ว่าระดับภาษายอมรับได้-เคสามารถเรียนรู้ได้ในขอบเขตจำกัดโดยตัวอย่างลักษณะอยู่ในรูปแบบของตัวอย่างบวกและตัวอย่างลบ นอกจากนี้จะแสดงให้เห็นว่าขนาดของตัวอย่างลักษณะแปรผันกับขนาดของออโตมาตาจำกัดเชิงกำหนดขอบ-เค ซึ่งอธิบายได้ด้วยฟังก์ชันพหุนาม การแสดงให้เห็นว่ามีตัวอย่างลักษณะสำหรับระดับภาษายอมรับได้-เค เรานิยามเพิ่มเติมดังนี้

นิยาม 5.1 *ให้ $M = (\Sigma_{\le}, Q, q_0, F_A, F_R, \delta)$ เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-เค สายอักขระส่วนหน้าที่สั้นที่สุด (shortest prefix) ของสถานะ $q \in Q$ นิยามดังนี้*

$$Short(q) = min\{u \in \Sigma^*_{\le} : \delta^* (q_\lambda, u) = q\}$$

และเซตของอักขระส่วนหน้าที่สั้นที่สุดของ M นิยามดังนี้

$$SP(M) = \{u \in \Sigma^*_{\le} : \forall q \in Q, u = Short(q)\}$$

นิยาม 5.2 *ให้ $M = (\Sigma_{\le}, Q, q_0, F_A, F_R, \delta)$ เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-เค เซตแก่นกลาง (Kernel set) ของ M นิยามดังนี้*

$$N(M) = \{\lambda\} \cup \{ ua : \forall q \in Q, u = Short(q) \text{ และ } \delta(q, a, b) \neq \varnothing \}$$
$$\cup \{ ub : \forall q \in Q, u = Short(q) \text{ และ } \delta(q, a, b) \neq \varnothing \}$$

จากนิยาม 5.1 และนิยาม 5.2 เราใช้เซตของสายอักขระส่วนหน้าที่สั้นที่สุด และเซตแก่นกลาง ในการสร้างตัวอย่างลักษณะในรูปแบบตัวอย่างบวกและตัวอย่างลบสำหรับการเรียนรู้ภาษายอมรับได้-เคโดยใช้อัลกอริทึมเคเอซีแอลไอ เราเรียกอัลกอริทึมการสร้างตัวอย่างลักษณะนี้ว่าอัลกอริทึมคอนสตรัคท์ซีเอส (ConstructCS) แสดงดังนี้

อัลกอริทึม 5.8 อัลกอริทึมคอนสตรัคซีเอส สำหรับการสร้างตัวอย่างลักษณะ

Algorithm : *ConstructCS*
Input : a k-DFA $M = (\Sigma_{\le}, Q, q_0, F_A, F_R, \delta)$
Output : a characteristic set $CS = (CS+, CS-)$

$CS+ \leftarrow \varnothing$
$CS- \leftarrow \varnothing$
FOR $u \in N(M)$ **DO**
  IF $\delta^* (q_0, u) \in F_A$ **THEN**
    $CS+ \leftarrow CS+ \cup \{u\}$
  ELSE IF $\delta^*(q_0, u) \in F_R$ **THEN**
      $CS- \leftarrow CS- \cup \{u\}$
    ENDIF
  ENDIF
ENDFOR
FOR $v \in SP(M)$ **DO**
  FOR $u \in N(M)$ **DO**
    IF $\delta^*(q_0, u) \neq \delta^*(q_0, v)$ **THEN**
      $w \leftarrow \min_{\le}\{w \in \Sigma^*_{\le} : (\delta^*(q_u, w) \in F_A \wedge \delta^*(q_v, w) \in F_R)$
        $\vee (\delta^*(q_u, w) \in F_R \wedge \delta^*(q_v, w) \in F_A) \}$
      IF $\delta^* (q_0, uw) \in F_A$ **THEN**
        $CS+ \leftarrow CS+ \cup \{uw\}$

```
            CS– ← CS– ∪ {vw}
        ELSE
            CS+ ← CS+ ∪ {vw}
            CS– ← CS– ∪ {uw}
        ENDIF
    ENDIF
  ENDFOR
ENDFOR
```

ตัวอย่าง 5.2 *กำหนดให้ $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-2 ซึ่งนิยามโดย $\Sigma_{\leq} = \{1, 2, 3, 4, 5\}$, $Q = \{q_0, q_1\}$, $F_A = \{q_0\}$ และ $F_R = \{q_1\}$ และฟังก์ชันการเปลี่ยน δ นิยามดังนี้ $\delta(q_0, 1, 3) = q_0$, $\delta(q_0, 4, 5) = q_1$, $\delta(q_1, 1, 1) = q_1$ และ $\delta(q_1, 2, 5) = q_0$ จงสร้างเซตของตัวอย่างลักษณะตามอัลกอริทึม 5.8*

วิธีทำ อัลกอริทึมคอนสตรัคท์ซีเอส สร้างเซตของตัวอย่างลักษณะของออโตมาตา M ได้ดังนี้ เริ่มต้นด้วยการสร้างเซตของอักขระส่วนหน้าที่สั้นที่สุดและเซตแก่นกลางของ M

$SP(M) = \{\lambda, 4\}$ และ $N(M) = \{\lambda, 1, 3, 4, 5, 41, 42, 45\}$

จากนั้นนำสายอักขระจากเซตทั้งสองมาเปรียบเทียบกันเพื่อหาเซตของตัวอย่างลักษณะ ดังนั้นจะได้เซตของตัวอย่างลักษณะของออโตมาตา M คือ

$CS+ = \{\lambda, 1, 3, 42, 45\}$ และ $CS\text{-} = \{4, 5, 41\}$ □

ทฤษฎีบท 5.3 *ขนาดของตัวอย่างลักษณะ ($||CS||$) ของภาษายอมรับได้-เค L สำหรับอัลกอริทึมเคเอซีแอลไอ คือ $\mathcal{O}(n^3k)$ โดยที่ n คือขนาดของออโตมาตาจำกัดเชิงกำหนดขอบ-เคของ L*

พิสูจน์ ให้ L เป็นภาษายอมรับได้-เค และ M เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-เคของ L โดย $M = (\Sigma_{\leq}, Q, q_0, F_A, F_R, \delta)$ ซึ่งมีจำนวนสถานะเท่ากับ n สถานะ

การพิสูจน์ทำได้โดยพิจารณาจากอัลกอริทึมคอนสตรัคท์ซีเอส ซึ่งจะเห็นว่าเซตของตัวอย่างลักษณะ CS สร้างจากเซตของสายอักขระส่วนหน้าที่สั้นที่สุด $SP(M)$ และเซตแก่นกลาง $N(M)$ ส่งผลให้ ขนาดของเซต CS (แทนด้วย $||CS||$) ขึ้นอยู่กับ $|SP(M)|$ และ $|N(M)|$

พิจารณาจากนิยาม 5.2 จะได้ว่า $|SP(M)| = n$ และ $|N(M)| = 2nk+1$ และพิจารณาการสร้างตัวอย่างลักษณะจากอัลกอริทึมคอนสตรัคท์ซีเอส จะได้ว่า

$|CS+| = \mathcal{O}(n^2k)$ และ $|CS\text{-}| = \mathcal{O}(n^2k)$

ต่อไปพิจารณาสายอักขระที่สั้นที่สุดที่ทำให้สายอักขระที่อยู่ในเซต $SP(M)$ และเซต $N(M)$ แตกต่างกัน จะเห็นว่าสายอักขระดังกล่าวมีความยาวได้ไม่เกิน n ดังนั้นขนาดของตัวอย่างลักษณะคือ $||CS|| = \mathcal{O}(n^3k)$ ■

สำหรับการแสดงว่าตัวอย่างลักษณะที่สร้างจากอัลกอริทึมคอนสตรัคท์ซีเอสเป็นตัวอย่างลักษณะของภาษายอมรับได้-เค เราจะแสดงด้วยบทตั้ง และทฤษฎีบทต่อไปนี้ โดยกำหนดให้ $\mathcal{A}$ แทนอัลกอริทึมเคเอซีแอลไอ และให้ M แทนออโตมาตาจำกัดเชิงกำหนดขอบ-เคของภาษายอมรับได้ L และ $S = (S+,\ S\text{-})$ เป็นเซตของตัวอย่าง โดยมี $CS = (CS+,\ CS\text{-})$ เป็นเซตของตัวอย่างลักษณะที่สร้างจากอัลกอริทึมคอนสตรัคท์ซีเอสของภาษา L

บทตั้ง 5.1 *ให้ L เป็นภาษายอมรับได้-เค ถ้า $CS \subseteq S \subseteq L$ แล้ว $\mathbb{L}(\mathcal{A}(S)) \subseteq \mathbb{L}(M)$*

พิสูจน์ การพิสูจน์ทำได้โดยแสดงว่าออโตมาตาจำกัดเชิงกำหนดขอบ-เคที่ได้จากอัลกอริทึม $\mathcal{A}$ เป็นออโตมาตาย่อยของออโตมาตา M เสมอ โดยใช้วิธีอุปนัยเชิงคณิตศาสตร์

ให้ $M = (\Sigma_{\leq},\ Q,\ q_0,\ F_A,\ F_R,\ \delta)$ เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-เค ของภาษายอมรับได้-เค L และให้ $M_{(i)} = (\Sigma_{\leq},\ Q_{(i)},\ q_0,\ F_{A(i)},\ F_{R(i)},\ \delta_{(i)})$ เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-เค ที่ได้จากอัลกอริทึม $\mathcal{A}$ รอบที่ i ของการทำซ้ำของลูป WHILE ในอัลกอริทึม 5.8

จากนิยาม 4.8 สำหรับการแสดงว่า $M_{(i)}$ เป็นออโตมาตาย่อยของ M นั้นทำได้โดยกำหนดฟังก์ชันสาทิสสัณฐาน $f:\ Q_{(i)} \rightarrow Q$ นิยามดังนี้

$$f(q) = \delta^*(q_0,\ u_q) \text{ สำหรับทุก } q \in Q_{(i)} \qquad (1)$$

โดยที่ u_q เป็นสายอักขระส่วนหน้าของ S ที่ทำให้จำนวนครั้งการผ่านของ $\delta^*_{(i)}(q_0,\ u_q) = q$ น้อยที่สุด โดยต้องทำการแสดงให้เห็นว่าฟังก์ชัน f จะต้องสอดคล้องกับเงื่อนไขต่อไปนี้

(i) ถ้า $q \in Q_{(i)}$ แล้ว $f(q) \in Q$,

(ii) ถ้า $q \in F_{A(i)}$ แล้ว $f(q) \in F_A$,

(iii) ถ้า $q \in F_{R(i)}$ แล้ว $f(q) \in F_R$,

(iv) สำหรับ $q \in Q_{(i)}$ และ $a \in \Sigma_{\leq}$, ถ้า $\delta_{(i)}(q,\ a,\ b) = p$ แล้ว $f(\delta_{(i)}(q,\ a,\ b)) = \delta(f(q),\ a,\ b)$

ขั้นตอนพื้นฐาน : จะแสดงให้เห็นว่า $\mathbb{L}(M_{(0)}) \subseteq \mathbb{L}(M)$

เนื่องจากอัลกอริทึมเคเอซีแอลไอ เริ่มต้นด้วยการสร้างเป็นต้นไม้ยอมรับส่วนหน้าขอบ-เคจากตัวอย่างที่กำหนดให้ (ในกรณีนี้คือ $S \supseteq CS$) จะได้ว่า $M_{(0)}$ เป็นต้นไม้ยอมรับส่วนหน้าที่ได้จาก $\mathcal{A}(S)$ ซึ่งนิยามดังนี้ $M_{(0)} = (\Sigma_{\leq},\ Q_{(0)},\ q_0,\ F_{A(0)},\ F_{R(0)},\ \delta_{(0)})$ โดยที่ $q_0 = q_\lambda$, $Q_{(0)} = \{q_u :\ u \in Pref(S)\}$, $F_{A(0)} = \{q_u \in Q_{(0)}\text{: } u \in S+\}$, $F_{R(0)} = \{q_u \in Q_{(0)} :\ u \in S\text{-}\}$, $\delta_{(0)}(q_u,\ a,\ a) = q_{ua}$ โดยที่ $u,\ ua \in Pref(S)$

สำหรับการแสดงให้เห็นว่า $\mathbb{L}(M_{(0)}) \subseteq \mathbb{L}(M)$ เราจะต้องแสดงให้เห็นว่าฟังก์ชันสาทิสสัณฐาน f ที่กำหนดให้ใน (1) สอดคล้องกับเงื่อนไข (i), (ii), (iii) และ (iv) โดยพิจารณาแต่ละเงื่อนไขได้ดังนี้

(i) สมมติให้ $q_u \in Q_{(0)}$ เนื่องจาก M เป็นต้นไม้ยอมรับส่วนหน้าขยายขอบ-เค จะได้ว่า $u \in Pref(S) \subseteq Pref(\mathbb{L}(M))$ ซึ่งจะได้ว่า $\delta^*(q_0, u) \in Q$ และจาก (1) จะได้ $f(q_u) \in Q$ ดังนั้นเงื่อนไข (i) เป็นจริง

(ii) สมมติให้ $q_u \in F_{A(0)}$ เนื่องจาก $u \in S+ \subseteq \mathbb{L}(M)$ จะได้ว่า $\delta^*(q_0, u) \in F_A$ จาก (1) จะได้ $f(q_u) \in F_A$ ดังนั้นเงื่อนไข (ii) เป็นจริง

(iii) สมมติให้ $q_u \in F_{R(0)}$ เนื่องจาก $u \in S- \subseteq \Sigma^*_{\leq} - \mathbb{L}(M)$ จะได้ว่า $\delta^*(q_0, u) \in F_R$ จาก (1) จะได้ $f(q_u) \in F_R$ ดังนั้นเงื่อนไข (iii) เป็นจริง

(iv) พิจารณา $\delta_{(0)}(q_0, a, b) = p$ สำหรับทุก q, p ใน $Q_{(0)}$ และให้ $u_p = \{u_q z : a \leq z \leq b\}$ จะเห็นว่าสายอักขระ u_p เป็นสายอักขระที่สั้นที่สุดที่เข้าถึงสถานะ p

$$\begin{aligned}
\delta(f(q), a, b) &= \delta(\delta^*(q_0, u_q), a, b) && ;\ \text{จาก (1)}\\
&= \delta^*(q_0, u_q z) && ;\ a \leq z \leq b\\
&= \delta^*(q_0, u_p) && ;\ u_q z = u_p\\
&= f(p) && ;\ \text{จาก (1)}\\
&= f(\delta_{(0)}(q, a, b)) && ;\ \delta_{(0)}(q, a, b) = p
\end{aligned}$$

จะเห็นได้ว่า $f(\delta_{(0)}(q, a, b)) = \delta(f(q), a, b)$ ทำให้เงื่อนไข (iv) เป็นจริง

ดังนั้นจาก (i), (ii), (iii) และ (iv) เราสามารถสรุปได้ว่า $\mathbb{L}(M_{(0)}) \subseteq \mathbb{L}(M)$ เป็นจริง

ขั้นตอนการอุปนัย : สมมติให้ $\mathbb{L}(M_{(t)}) \subseteq \mathbb{L}(M)$ เป็นจริง และแสดงให้เห็นว่า $\mathbb{L}(M_{(t+1)}) \subseteq \mathbb{L}(M)$ เป็นจริง

จากการสมมติให้ $\mathbb{L}(M_{(t)}) \subseteq \mathbb{L}(M)$ เป็นจริง จะได้ว่า เงื่อนไข (i), (ii), (iii) และ (iv) เป็นจริงสำหรับออโตมาตา $M_{(t)}$ และจากอัลกอริทึมคอนสตรัคท์ซีเอสจะเห็นว่าออโตมาตาจำกัดเชิงกำหนดขอบ-เค $M_{(t+1)}$ สร้างได้จากการผสานสถานะในออโตมาตาจำกัดเชิงกำหนดขอบ-เค $M_{(t)}$ ด้วยรูปแบบการผสานสถานะที่แตกต่างกัน 3 รูปแบบ คือ

รูปแบบที่ 1: $[M_{(t+1)} = REC\_MERGE(M_{(t)}, q_\omega, q_\beta)$ โดยที่ $\delta_{(t)}(q_u, a_1, b_1) = q_\omega$ และ $\delta_{(t)}(q_u, a_2, b_2) = q_\beta]$

พิจารณาการผสานสถานะรูปแบบนี้ จะได้ว่า

$Q_{(t+1)} = Q_{(t)} - \{q_\beta\} \subseteq Q_{(t)}$,

$F_{A(t+1)} = F_{A(t)} - \{q_\beta\} \subseteq F_{A(t)}$,

$F_{R(t+1)} = F_{R(t)} - \{q_\beta\} \subseteq F_{R(t)}$,

และ $\delta_{(t+1)} = \delta_{(t)} - (\{(q_u, a, b, q_\beta)\} \cup \{(q_\beta, a, b, q_v) : q_v \in Q_{(t)}\}) \subseteq \delta_{(t)}$

และจากการสมมติว่า $M_{(t)}$ สอดคล้องกับเงื่อนไข (i), (ii), (iii) และ (iv) ดังนั้น $M_{(t+1)}$ที่ได้จากการผสานสถานะรูปแบบนี้ สอดคล้องกับเงื่อนไข (i), (ii), (iii) และ (iv) ด้วย

รูปแบบที่ 2: $[M_{(t+1)} = Merge(M_{(t)}, q_\omega, q_\beta)$ โดยที่ $\delta_{(t)}(q_u, a_1, b_1) = q_\omega$ และ

$\delta_{(t)}(q_v, a_2, b_2) = q_\beta, q_u \neq q_v]$

พิจารณาการผสานสถานะรูปแบบนี้ จะได้ว่า

$$Q_{(t+1)} = Q_{(t)} - \{q_\beta\} \subseteq Q_{(t)},$$
$$F_{A(t+1)} = F_{A(t)} - \{q_\beta\} \subseteq F_{A(t)},$$
$$F_{R(t+1)} = F_{R(t)} - \{q_\beta\} \subseteq F_{R(t)},$$

และ $\delta_{(t+1)} = \delta_{(t)} - (\{(q_v, a, b, q_\beta)\} \cup \{(q_\beta, a, b, q_v) : q_v \in Q_{(t)}\}) \cup \{ (q_v, a, b, q_\omega)\} \subseteq \delta_{(t)}$ และจากการสมมติว่า $M_{(t)}$ สอดคล้องกับเงื่อนไข (i), (ii), (iii) และ (iv) ดังนั้น $M_{(t+1)}$ ที่ได้จากการผสานสถานะรูปแบบนี้ สอดคล้องกับเงื่อนไข (i), (ii), (iii) และ (iv) ด้วย

รูปแบบที่ 3: $[M_{(t+1)} = M_{(t)}$ เนื่องจากไม่สามารถผสานสถานะได้].

พิจารณาการผสานสถานะรูปแบบนี้ จะได้ว่า

$$Q_{(t+1)} = Q_{(t)},$$
$$F_{A(t+1)} = F_{A(t)},$$
$$F_{R(t+1)} = F_{R(t)},$$

และ $\delta_{(t+1)} = \delta_{(t)}$

และจากการสมมติว่า $M_{(t)}$ สอดคล้องกับเงื่อนไข (i), (ii), (iii) และ (iv) ดังนั้น $M_{(t+1)}$ที่ได้จากการผสานสถานะรูปแบบนี้ สอดคล้องกับเงื่อนไข (i), (ii), (iii) และ (iv) ด้วย

จากการพิจาณาการผสานสถานะทั้ง 3 รูปแบบแล้ว เราสามารถสรุปได้ว่า $M_{(t+1)}$สอดคคล้องกับเงื่อนไข (i), (ii), (iii) และ (iv) และจากขั้นตอนพื้นฐานและขั้นตอนอุปนัยเราสามารถสรุปได้ว่า ถ้า $CS \subseteq S \subseteq L$ แล้ว $\mathbb{L}(\mathcal{A}(S)) \subseteq \mathbb{L}(M)$ ■

บทตั้ง 5.2 *ให้ L เป็นภาษายอมรับได้-เค ถ้า $CS \subseteq S \subseteq L$ แล้ว $\mathbb{L}(M) \subseteq \mathbb{L}(\mathcal{A}(S))$*

พิสูจน์ ให้ $M = (\Sigma_\leq, Q, q_0, F_A, F_R, \delta)$ เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-เค ของภาษายอมรับได้-เค L และ $M_{(n)} =(\Sigma_\leq, Q_{(n)}, q_0, F_{A(n)}, F_{R(n)}, \delta_{(n)})$ เป็นออโตมาตาจำกัดเชิงกำหนดขอบ-เคที่ได้จากอัลกอริทึม $\mathcal{A}$ เราจะพิสูจน์โดยแสดงให้เห็นว่า M เป็นออโตมาตาเชิงกำหนดขอบ-เคย่อยโดยสาทิสสัณฐานของ $M_{(n)}$ ตามนิยาม 4.8 โดยกำหนดฟังก์ชันสาทิสสัณฐาน $f : Q \rightarrow Q_{(n)}$ ซึ่งจะทำให้เราสรุปได้ว่า $\mathbb{L}(M) \subseteq \mathbb{L}(\mathcal{A}(S))$ เมื่อ $CS \subseteq S$ ฟังก์ชัน f นิยามดังนี้

$$f(q) = \delta^*_{(n)}(q_0, u_q) \text{ สำหรับทุก } q \in Q \quad (2)$$

โดยที่ u_q เป็นสายอักขระส่วนหน้าของ L ที่ทำให้จำนวนครั้งการผ่านของ $\delta^*(q_0, u_q) = q$ น้อยที่สุด โดยต้องทำการแสดงให้เห็นว่าฟังก์ชัน f จะต้องสอดคล้องกับเงื่อนไขต่อไปนี้

(i) ถ้า $q \in Q$ แล้ว $f(q) \in Q_{(n)}$,

(ii) ถ้า $q \in F_A$ แล้ว $f(q) \in F_{A(n)}$,

(iii) ถ้า $q \in F_R$ แล้ว $f(q) \in F_{R(n)}$,

(iv) สำหรับ $q \in Q$ และ $a \in \Sigma_{\leq}$, ถ้า $\delta(q, a, b) = p$ แล้ว $f(\delta(q, a, b)) = \delta_{(n)}(f(q), a, b)$

โดยแสดงดังต่อไปนี้

(i) สมมติให้ $q \in Q$ ในออโตมาตาจำกัดเชิงกำหนดขอบ-เค M จะได้ว่าสถานะ q เป็นไปได้ 2 กรณี คือ $q \in F_A$ หรือ $q \in F_R$ พิจารณาแต่ละรูปแบบดังนี้
กรณีแรก: $q \in F_A$: ถ้า $q \in F_A$ แล้วจากอัลกอริทึมคอนสตรัคท์ซีเอส จะได้ว่า $u_q \in CS+$ ซึ่ง $\delta^*(q_0, u_q) = q$ ทำให้ได้ว่า $u_q \in \mathbb{L}(M_{(n)})$ ดังนั้น $\delta^*_{(n)}(q_0, u_q) \in Q_{(n)}$ ซึ่งทำให้ $f(q) \in Q_{(n)}$ เนื่องจาก $f(q) = \delta^*_{(n)}(q_0, u)$ ตามนิยาม
กรณีที่สอง: $q \in F_R$: ถ้า $q \in F_R$ แล้วจากอัลกอริทึมคอนสตรัคท์ซีเอส จะได้ว่า $u_q \in CS-$ ซึ่ง $\delta^*(q_0, u_q) = q$ ทำให้ได้ว่า $u_q \in \Sigma^* - \mathbb{L}(M_{(n)})$ ดังนั้น $\delta^*_{(n)}(q_0, u_q) \in Q_{(n)}$ซึ่งทำให้ $f(q) \in Q_{(n)}$ เนื่องจาก $f(q) = \delta^*_{(n)}(q_0, u)$ ตามนิยาม

(ii) สมมติให้ $q \in F_A$ ในออโตมาตาจำกัดเชิงกำหนดขอบ-เค M จะได้ว่า $u_q \in CS+$ ซึ่ง $\delta^*(q_0, u_q) = q$ ดังนั้นจะเห็นว่า $\delta^*_{(n)}(q_0, u_q) \in F_{A(n)}$ ซึ่งทำให้ $f(q) \in F_{A(n)}$ เนื่องจาก $f(q) = \delta^*(q_0, u)$ ตามนิยาม

(iii) สมมติให้ $q \in F_R$ ในออโตมาตาจำกัดเชิงกำหนดขอบ-เค M จะได้ว่า $u_q \in CS-$ ซึ่ง $\delta^*(q_0, u_q) = q$ ดังนั้นจะเห็นว่า $\delta^*_{(n)}(q_0, u_q) \in F_{R(n)}$ ซึ่งทำให้ $f(q) \in F_{R(n)}$ เนื่องจาก $f(q) = \delta^*(q_0, u)$ ตามนิยาม

(iv) จะแสดงให้เห็นว่า $\delta_{(n)}(f(q), a, b)) = f(\delta(q, a, b)\ )$ โดยพิจารณา $\delta(q, a, b) = p$ สำหรับทุก q, p ใน Q และให้ $u_p = \{u_q z : a \leq z \leq b\}$ จะเห็นว่าสายอักขระ u_p เป็นสายอักขระที่สั้นที่สุด

$$
\begin{aligned}
\delta_{(n)}(f(q), a, b)) &= \delta_{(n)}(\delta^*_{(n)}(q_0, u_q), a, b)) && ;\ \text{จาก } (2) \\
&= \delta^*_{(n)}(q_0, u_q z) && ;\ a \leq z \leq b \\
&= \delta^*_{(n)}(q_0, u_p) && ;\ u_q z = u_p \\
&= f(p) && ;\ \text{จาก } (2) \\
&= f(\delta(q, a, b)) && ;\ \delta(q, a, b) = p
\end{aligned}
$$

ดังนั้น จาก (i), (ii), (iii) และ (iv) เราสามารถสรุปว่า $\mathbb{L}(M) \subseteq \mathbb{L}(\mathcal{A}(S))$ เมื่อ $CS \subseteq S$ ■

ทฤษฎีบท 5.4 *มีตัวอย่างลักษณะของภาษายอมรับได้-เค สำหรับอัลกอริทึมเคเอซีแอลไอ*

พิสูจน์ จากบทตั้ง 5.1 และ 5.2 เราสามารถสรุปได้ว่ามีตัวอย่างลักษณะของภาษายอมรับได้-เค L สำหรับอัลกอริทึมเคเอซีแอลไอ ■

ทฤษฎีบท 5.5 *ระดับภาษายอมรับได้-เค สามารถเรียนรู้ได้จากข้อมูลที่กำหนดให้จากตัวอย่างบวกและตัวอย่างลบ*

พิสูจน์ จากนิยาม 3.27 และทฤษฎีบท 5.4 เราสามารถสรุปได้ว่า ระดับภาษายอมรับได้-เค สามารถเรียนรู้ได้จากข้อมูลที่กำหนดให้จากตัวอย่างบวกและตัวอย่างลบ ■

ทฤษฎีบท 5.6 *ระดับภาษายอมรับได้-เค สามารถเรียนรู้ได้ในขอบเขตจำกัดจากตัวอย่างบวกและตัวอย่างลบ*

พิสูจน์ จากทฤษฎีบท 3.6 และทฤษฎีบท 5.5 เราสามารถสรุปได้ว่า ระดับภาษายอมรับได้-เค สามารถเรียนรู้ได้ในขอบเขตจำกัดจากตัวอย่างบวกและตัวอย่างลบ ■

ทฤษฎีบท 5.7 *สำหรับจำนวนเต็ม $k \geq 0$ ระดับภาษายอมรับได้-เค สามารถเรียนรู้ได้ในขอบเขตจำกัดโดยใช้เวลาและจำนวนข้อมูลเชิงพหุนามจากตัวอย่างบวกและตัวอย่างลบ*

พิสูจน์ จากทฤษฎีบท 5.2, 5.3 และ 5.4 เราสามารถสรุปได้ว่า สำหรับจำนวนเต็ม $k \geq 0$ ระดับภาษายอมรับได้-เค สามารถเรียนรู้ได้ในขอบเขตจำกัดโดยใช้เวลาและจำนวนข้อมูลเชิงพหุนามจากตัวอย่างบวกและตัวอย่างลบ ■

5.3 สรุป

ในบทนี้ เราได้พิสูจน์ให้เห็นว่าระดับภาษายอมรับได้-เค ไม่สามารถเรียนรู้ได้ในขอบเขตจำกัดจากตัวอย่างบวกเพียงอย่างเดียว (ทฤษฎีบท 5.1) โดยใช้คุณสมบัติของการมีอยู่ของจุดจำกัด และได้ทำการพิสูจน์ว่าระดับภาษายอมรับได้-สามารถเรียนรู้ได้ในขอบเขตจำกัด (ทฤษฎีบท 5.6) และสามารถเรียนรู้ได้จากตัวอย่างที่กำหนดให้ (ทฤษฎีบท 5.5) นอกจากนั้นยังได้พิสูจน์ให้เห็นว่า ระดับภาษายอมรับได้-สามารถเรียนรู้ได้ในจำกัดโดยใช้เวลาและจำนวนข้อมูลเชิงพหุนามจากตัวอย่างบวกและตัวอย่างลบ (ทฤษฎีบท 5.7) โดยแสดงให้เห็นถึงการมีอยู่ของอัลกอริทึมการเรียนรู้ (อัลกอริทึมเคเอซีแอลไอ) ซึ่งทำงานด้วยเวลาเชิงพหุนาม (ทฤษฎีบท 5.2) และแสดงให้เห็นถึงการมีอยู่ของตัวอย่างลักษณะ (ทฤษฎีบท 5.4) ซึ่งขนาดของตัวอย่างลักษณะนี้มีอัตราการเติบโตไม่เกินฟังก์ชันพหุนามของขนาดของออโตมาตาจำกัดเชิงกำหนดขอบ-เค (ทฤษฎีบท 5.3)

บทที่ 6

สรุปผลการวิจัยและข้อเสนอแนะ

วิทยานิพนธ์นี้ศึกษาความสามารถในการเรียนรู้ของระดับภาษายอมรับได้-เค ซึ่งมีออโตมาตาจำกัดเชิงกำหนดขอบ-เคเป็นตัวแทนเชิงไวยากรณ์ โดยใช้แบบจำลองการเรียนรู้เชิงตัวอย่างที่เรียกว่า การระบุภาษาได้ในขอบเขตจำกัด แบบจำลองนี้ถูกนำมาศึกษาความสามารถในการเรียนรู้ของระดับของภาษายอมรับได้-เคใน 2 กรณี คือ ความสามารถในการเรียนรู้จากตัวอย่างบวกเพียงอย่างเดียว และความสามารถในการเรียนรู้จากตัวอย่างบวกและตัวอย่างลบ

คุณสมบัติต่าง ๆ ที่จำเป็นของระดับของภาษายอมรับได้-เค ได้ถูกตรวจสอบเชิงทฤษฎีเพื่อนำไปใช้ในการพิสูจน์ว่าระดับของภาษายอมรับได้-เคนี้มีความสามารถในการเรียนรู้หรือไม่ สำหรับการเรียนรู้จากตัวอย่างบวกเพียงอย่างเดียว ระดับของภาษายอมรับได้-เคถูกพิสูจน์ให้เห็นในวิทยานิพนธ์นี้ว่าไม่สามารถเรียนรู้ได้ในขอบเขตจำกัดจากตัวอย่างบวกเพียงอย่างเดียว (ทฤษฎีบท 5.1) โดยใช้คุณสมบัติของการมีอยู่ของจุดจำกัด และจากผลการศึกษานี้เราสามารถสรุปได้ว่าการใช้ตัวอย่างบวกอย่างเดียวเป็นข้อมูลในการเรียนรู้ ไม่เพียงพอที่จะทำให้การเรียนรู้ภาษาระดับนี้สำเร็จ ซึ่งหมายความว่าจำเป็นต้องใช้ข้อมูลชนิดอื่นเพิ่มเติมสำหรับระดับภาษายอมรับได้-เค อาทิเช่น ตัวอย่างลบ เป็นต้น

จากการศึกษาความสามารถในการเรียนรู้ระดับของภาษายอมรับได้-เคจากตัวอย่างบวกและตัวอย่างลบ เราได้แสดงให้เห็นว่ามีอัลกอริทึมการเรียนรู้ และมีตัวอย่างลักษณะที่ทำให้การเรียนรู้ภาษาสำเร็จ ซึ่งการมีตัวอย่างลักษณะนี้ถูกใช้เป็นเงื่อนไขที่เพียงพอในการสรุปว่าระดับของภาษายอมรับได้-เคสามารถเรียนรู้ได้จากข้อมูลที่กำหนดให้จากตัวอย่างบวกและตัวอย่างลบ (ทฤษฎีบท 5.5) และนำไปสู่การสรุปว่าระดับของภาษายอมรับได้-เคสามารถเรียนรู้ได้ในขอบเขตจำกัดจากตัวอย่างบวกและตัวอย่างลบ (ทฤษฎีบท 5.6) นอกจากนี้ในงานวิจัยนี้ยังได้ศึกษาถึงเรื่องประสิทธิภาพของการเรียนรู้ของระดับภาษายอมรับได้-เค โดยได้แสดงให้เห็นว่าอัลกอริทึมการเรียนรู้เคเอซีแอลไอที่ได้เสนอนั้นสามารถทำงานด้วยเวลาเชิงพหุนาม และขนาดของตัวอย่างลักษณะที่อัลกอริทึมเคเอซีแอลไอใช้นั้นอธิบายด้วยฟังก์ชันพหุนาม ซึ่งนำไปสู่การสรุปว่าระดับของภาษายอมรับได้-เคสามารถเรียนรู้ได้ในขอบเขตจำกัดโดยใช้เวลาและจำนวนข้อมูลเชิงพหุนามจากตัวอย่างบวกและตัวอย่างลบ (ทฤษฎีบท 5.7)

ข้อเสนอแนะสำหรับงานวิจัยที่น่าจะนำไปศึกษาต่อ ได้แก่ การศึกษาความสามารถในการเรียนรู้ของระดับภาษายอมรับได้-เค โดยใช้แบบจำลองการเรียนรู้อื่น ๆ เช่น แบบจำลองการเรียนรู้แบบโต้ตอบได้ (Active learning model) แบบจำลองการเรียนรู้แบบถูกต้องโดยประมาณที่น่าจะเป็นไปได้ (Probably approximation correct learning model) และ

เนื่องจากการศึกษาในวิทยานิพนธ์นี้เป็นการศึกษาเชิงทฤษฎีเท่านั้น การพัฒนาอัลกอริทึมเรียนรู้ภาษายอมรับได้-เคให้มีประสิทธิภาพมากขึ้น และเหมาะสมที่จะนำไปประยุกต์กับปัญหาจริงก็เป็นอีกประเด็นหนึ่งที่น่าสนใจนำไปศึกษาวิจัย

ผู้ที่สนใจศึกษาเรื่องความสามารถในการเรียนรู้ภาษาในแบบจำลองการเรียนรู้ภาษาแบบโต้ตอบได้ และแบบจำลองการเรียนรู้ภาษาแบบถูกต้องโดยประมาณที่น่าจะเป็นไปได้ ควรเริ่มจากการศึกษานิยาม และทฤษฎีบท และทำความเข้าใจบทพิสูจน์ต่าง ๆ รวมถึงเงื่อนไขเพียงพอและเงื่อนไขจำเป็นที่ใช้ในการพิสูจน์ ซึ่งถูกเสนอไว้ในบทความวิชาการด้านการเรียนรู้ภาษาที่ได้ศึกษาความสามารถในการเรียนรู้ของระดับภาษาต่าง ๆ เช่น ระดับภาษาสม่ำเสมอ ระดับภาษาไม่พึ่งบริบท บนแบบจำลองการเรียนรู้ทั้งสองแบบนี้

- ในกรณีที่ศึกษาบนแบบจำลองการเรียนรู้แบบโต้ตอบได้ ระดับภาษาสม่ำเสมอได้ถูกพิสูจน์แล้วว่าไม่สามารถเรียนรู้ได้จากการใช้ข้อซักถามความเป็นสมาชิก (membership query) หรือข้อซักถามความสมมูล (equivalence query) เพียงอย่างเดียว แต่จำเป็นต้องใช้ข้อซักถามทั้งสองแบบนี้ถึงจะเรียนรู้ได้
- ในกรณีที่ศึกษาบนแบบจำลองการเรียนรู้แบบถูกต้องโดยประมาณที่น่าจะเป็นไปได้ ระดับภาษาสม่ำเสมอได้ถูกพิสูจน์แล้วว่าไม่สามารถเรียนรู้ได้ด้วยความถูกต้องโดยประมาณที่น่าจะเป็นไปได้จากการใช้ตัวอย่างบวกเพียงอย่างเดียว

เนื่องด้วยภาษายอมรับได้-เค เป็นภาษาที่นิยามบนเซตอักขระมีอันดับ และเป็นระดับย่อยของระดับภาษาสม่ำเสมอ จึงเป็นที่น่าสนใจว่าความสามารถของระดับของภาษายอมรับได้-เคจะเป็นอย่างไรบนแบบจำลองการเรียนรู้ทั้งสองแบบจำลองนี้

การพัฒนาอัลกอริทึมการเรียนรู้ของระดับภาษายอมรับได้-เค ให้มีประสิทธิภาพมากขึ้น และสามารถนำไปประยุกต์ใช้กับปัญหาต่าง ๆ ได้นั้น เป็นอีกประเด็นหนึ่งที่น่าสนใจ โดยมีแนวทางในการพัฒนาคือ การเพิ่มฟังก์ชันช่วยค้นหา (heuristic function) เข้าไปในอัลกอริทึมการเรียนรู้เพื่อเลือกคู่ของสถานะในการผสาน โดยอาจเป็นพิจารณาค่าความถี่ของข้อมูลที่ผ่านสถานะ หรือพิจารณาในลักษณะของการให้คะแนนการผสานทุกคู่ของสถานะที่เป็นไปได้ แล้วเลือกคู่ของสถานะที่ให้คะแนนมากที่สุด ก็เป็นวิธีการหนึ่งที่น่าสนใจในการพัฒนาอัลกอริทึมการเรียนรู้

รายการอ้างอิง

[1] Chomsky, N. Syntactic Structures. The Hague, Mouton & Co, The Netherlands, 1957.

[2] Chomsky, N. Tree models for the description of languages. In Proceedings of the Symposium on Information Theory, 1956.

[3] Shieber, S. M. Evidence against the context-freeness of natural languages. Linguistics and Philosophy, 8 (1958): 333-343.

[4] Fu, K. S., and Booth, T. L. Grammatical inference: Introduction and survey. Part I, IEEE Transaction on Systems Man and Cybernetics 5 (1975): 59–72.

[5] Fu, K. S., and Booth, T. L. Grammatical inference: Introduction and survey. Part II, IEEE Transaction on System Man and Cybernetics 5 (1975): 409–423.

[6] Vidal, E. Grammatical inference: an introductory survey. In Proceedings. of the second International Colloquium Grammatical Inference (ICGI'94), 1994.

[7] Sakakibara, Y. Recent advances of grammatical inference. Theoretical Computer Science 185 (1997): 15–45.

[8] Higuera, C. D. Current trends in grammatical inference, in Advances in Pattern Recognition. Joint IAPR International Workshops SSPR+SPR 2000, 2000,

[9] Higuera, C. D. A bibliographical study of grammatical inference. Pattern recognition, 38 (2005): 1332-1348.

[10] Gold, E. M. Language identification in the limit. Information and Control, 10 (1967): 447-474.

[11] Angluin, D. Queries and concept learning. Machine Learning Journal, 2 (1987): 319-342.

[12] Valiant, L. G. A theory of the learnable. Communications of the ACM 27(1984): 1134-1142.

[13] Gold, E. M. Complexity of automaton identification from given data. Information and Control, 37 (1978): 302-320.

[14] Dupont, P. Utilisation et apprentissage de modeles de langage pour la reconnaissance de la parole continue. Doctoral dissertation, l' ENST Paris, 1996.

[15] Higuera, C. D. Grammatical inference: Learning Automata and Grammars. Cambridge University Press, Cambridge, 2010.

[16] Higuera, C. D. Characteristic sets for polynomial grammatical inference. Machine Learning Journal, 27 (1997): 125-138.

[17] Floriencio, C. Learning categorical grammars. Doctoral dissertation, University of Utrecht, 2003.

[18] Angluin, D. Inference of reversible languages. Journal of Association for Computing Machinery, 29 (1982): 741-765.

[19] Garcia, P., and Vidal, E. Inference of k-testable languages in the strict sense and applications to syntactic pattern recognition, IEEE Transaction on Pattern Analysis and Machine Intelligence, 12 (1990): 920-925.

[20] Emerald, J. D., et al. Learning Code Regular and Code Linear Languages, In Proceedings of Third International Colloquium on Grammatical Inference, 1996.

[21] Lange, S. A Note on Polynominal-Time Inference of *k*-Variable Pattern Languages, In Proceedings of the 1st International Workshop on Nonmonotonic and Inductive Logic, 1990.

[22] Mäkinen, E. A family of languages which is polynomialtime learnable from positive data in Pitt's sense, International Journal of computer mathematics, 61 (1996): 175-179.

[23] Yokomori, T. On polynomial-time learnability in the limit of strictly deterministic automata, Machine Learning 19 (1995) 153-179.

[24] Ruiz, J., and García, P. Learning k-piecewise testable languages from positive data. In Proceedings of the 3th International Colloquium on Grammatical Inference, 1996.

[25] Koshiba, T., et al. Learning Deterministic even Linear Languages From Positive Examples. Theoretical Computer Science 185(1997): 63-79.

[26] Mäkinen, E. Inferring uniquely terminating regular languages from positive data, Information Processing Letters, 2(1997).

[27] Fernau, H. Identification of function distinguishable languages, Theoretical Computer Science, 290 (2003): 1679-1711.

[28] Denis, F., et al. Learning regular languages using RFSA, In Proceedings of Algorithmic Learning Theory 2001, 2001.

[29] Yokomori, T. Polynomial-time identification of very simple grammars from positive data. Theoretical Computer Science 1 (2003): 179–206.

[30] Latteux, M., et al. Identification of biRFSA languages. Theoretical Computer Science 356 (2006): 212-223.

[31] Clark, A., and Eyraud, R. Polynomial Identification in the Limit of Substitutable Context-free Languages. The Journal of Machine Learning Research 8(2007): 1725-1745.

[32] Yoshinaka, R. Identification in the Limit of k, l-Substitutable Context-Free Languages. In Proceedings of the 9th International Colloquium on Grammatical Inference(ICGI'08). 2008.

[33] Yoshinaka, R. Efficient Learning of Multiple Context-Free Languages with Multidimensional Substitutability from Positive Data.Theoretical Computer Science 356 (2006): 212-223.

[34] Wakatsuki, M., and Tomita, E. Polynomial time identification of strict prefix deterministic finite state transducers. In Proceedings of the tenth international colloquium on grammatical inference, 2010.

[35] Trakhtenbrot, B., and Barzdin, Y. Finite Automata: Behavior and Synthesis. North-Holland, Amsterdam, 1973.

[36] Oncina, J., and Garcia, P. Identifying regular languages in polynomial time. Advances in structural and syntactic pattern recognition, 5 (1992): 99-108.

[37] Lang, K. Random DFA's can be approximately learned from sparse uniform examples. In Proceedings of COLT 1992, 1992.

[38] Higuera, C. D., et al. Identification of DFA: data-dependent versus data-independent algorithm. In Proceedings of international colloquium grammatical inference 1996, 1996.

[39] Lang, K. J., and Pearlmutter B. A. The Abbadingo one DFA learning competition, 1997.

[40] Lang K. J. et al. Results of the Abbadingo one DFA learning competition and a new evidence-driven state merging algorithm. In Proceedings of international Colloquium on Grammatical inference 1998, 1998.

[41] Lambeau, B., et al. State-Merging DFA Induction Algorithms with Mandatory Merge Constraints. In Proceedings of the tenth international Colloquium on Grammatical inference, 2008.

[42] Garcia, P., et al. A Comparative Study of Two Algorithms for Automata Identification. In Proceedings of international Colloquium on Grammatical inference 2000, 2000.

[43] Cicchello, O., and Kremer, S.C. Inducing grammars from sparse data sets: A survey of algorithms and results, Journal of Machine Learning Research, 4(2003): 603-632.

[44] Higuera, C. D., and Oncina, J. On Sufficient Conditions to Identify in the Limit Classes of Grammars from Polynomial Time and Data, In Proceedings of international Colloquium on Grammatical inference 2002, 2002.

ประวัติผู้เขียนวิทยานิพนธ์

นายอนุชิต จิตพัฒนกุล เกิดวันที่ 20 พฤศจิกายน พ.ศ. 2520 ที่จังหวัดกรุงเทพมหานคร สำเร็จการศึกษาปริญญาวิทยาศาสตรบัณฑิต สาขาวิชาคณิตศาสตร์ประยุกต์ จากภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ประยุกต์ จากสถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ ในปีการศึกษา 2542 และสำเร็จการศึกษาในหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคณนา จากภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2546 หลังจากนั้นได้เข้าศึกษาในหลักสูตรวิศวกรรมศาสตรดุษฎีบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์ ที่ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2547 โดยได้รับทุนอุดหนุนการศึกษา โครงการพัฒนาอาจารย์สาขาขาดแคลน(คณิตศาสตร์)เพื่อศึกษาในประเทศ ตามความต้องการของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ