การสังเคราะห์เสียงพูดภาษาไทยสำหรับคำที่ไม่รู้จัก

Morphological Derivative for Unknown Words in Thai text-to-speech Synthesis

Pisit Promchan,
PsTNLP laboratory, http://geocities.com/pisitp,
Bangkok, Thailand,
E- Mail: pisitp@yahoo.com,
Mailing address: 18/57 Sukhumwid 68, Bangna,
Bkk., 10260, Tel: 5740440 ext. 216, 7449783

Wittaya Wongvachirapanich and Saanti Chinnakarn,
Department of Computer Engineering,
Prince of Songklanakharin University,
Songkla, Thailand,
E- Mail: s4010388, s4010362@thaimail.com,

บทคัดย่อ: บทความนี้นำเสนอผลการวิจัยและพัฒนาสร้างระบบการสังเคราะห์เสียงพูดภาษาไทยที่นอกจากมีความสามารถในการพูดหรืออ่านคำทั่ว ๆ ไปได้ดีแล้ว ยังมีความสามารถในการผันเสียงเพื่อการอ่านคำที่ไม่รู้จักหรือคำที่ไม่ปรากฏในพจนานุกรมได้ ที่ระดับความแม่นยำ (Precision) สูงมากพอสมควรที่เดียว ทีมงานวิจัยได้ประยุกต์ใช้หลักแนวคิดปรัชญาตรรกแห่งความคลุมเครือ (Fuzzy Logic) เป็นเครื่องมือหลักในการแก้ปัญหาในการสังเคราะห์คำที่ไม่รู้จักเหล่านั้น ในงานวิจัยนี้ได้นำเสนอแนวทางเบื้องต้นของตรรกแห่งความคลุมเครือ สถาปัตยกรรมของระบบที่ได้ออกแบบไว้ ขั้นตอนวิธีการทำวจีวิภาค การประมาณหน่วยเสียงไกล้เคียงหรือขั้นตอนของการคำณวนหาสัดส่วนของการเป็นสมาชิกในเซ็ตของความคลุมเครือ และผลการทดสอบประเมินประสิทธิภาพของระบบ พบว่าในกรณีของภาษาไทยทั่ว ๆ ระบบมีความแม่นยำเฉลี่ย 99.59% และในกรณีของข้อความที่ประกอบด้วยคำที่ไม่รู้จักทั้งหมดระบบมีความแม่นยำเฉลี่ย 96.69% ตำแหน่งที่มีความแม่นยำและในเวลาเดียวกันมีความครอบคลุม (Recall) สูงสุดคือประมาณ 98% และ 88% ในกรณีของภาษาไทยทั่ว ๆ และในกรณีของข้อความที่ประกอบด้วยคำที่ไม่รู้จักทั้งหมดตามลำดับ

Abstract: This paper presents the morphological derivative for unknown words in Thai text-to-speech synthesis. The research methodology is based on the philosophy of fuzzy logic theory. The paper contains the basic idea of the fuzzy logic, the system architecture, parsing algorithm, approximation matching algorithm for Thai or the fuzzy value calculation. The experimental and performance evaluations are also included. It is found that the system performs up-to 99.59% of precision in case of the Thai text data contain both known and unknown words. The system perform up-to 96.69% in term of precision in case of the Thai text data contains purely unknown. The highest intersection point between precision and recall are about 98% and 88% for Thai text data contain both known and unknown words and the Thai text data contains purely unknown respectively.

Keywords: Thai Text-to-Speech, Synthesis, Parsing, Unknown Word Identification, Approximate Matching, Algorithms, CTI, Fuzzy Logic

1. บทนำ

ระบบการสังเคราะห์เสียงพูดภาษาไทยเริ่มเข้ามามีบทบาทในชีวิตประจำวันของคนไทยในสังคมยุคข่าวสารสารสนเทศมากยิ่งขึ้นโดยเฉพาะอย่างยิ่งเทคโนโลยีการเชื่อมต่อกันระหว่างระบบโทรศัพท์กับคอมพิวเตอร์ (CTI: Computer Telephony Integration) ทำให้เกิดอุต

สาหกรรมใหม่ ๆ เช่น การบริการข้อมูลตามสาย การธนาคารทางโทรศัพท์ ฯลฯ งานวิจัยและพัฒนาเทคโนโลยีในสาขานี้จึงมีความสำคัญมากขึ้นเรื่อย ๆ และควรส่งเสริมให้ขยายวงกว้างออกไป

ในระบบการสังเคราะห์เสียงพูดภาษาไทยด้วยคอมพิวเตอร์ในอดีตที่ผ่านมา มีปัญหามากในการสังเคราะห์เสียงพูดสำหรับคำที่เป็นชื่อเฉพาะ (Proper Noun) คำทับศัพท์ภาษาต่างประเทศ (Transliteration) คำศัพท์ที่เกิดจากการรวมกันของคำศัพท์อื่นใดเช่นคำสมาส คำสนธิ หรือคำที่มีการสะกดผิด ซึ่งรวมเรียกคำเหล่านี้ทั่ว ๆ ไปว่าคำที่ไม่รู้จัก (Unknown Words) หรือคือคำที่ไม่ปรากฏในพจนานุกรมนั่นเอง คำที่ไม่รู้จัก ไม่ว่าจะเป็นคำทับศัพท์หรือชื่อเฉพาะต่าง ๆ หรือแม้แต่คำที่เกิดมาใหม่ ๆ ตามยุคสมัย ยกตัวอย่างเช่น “ถนนรัชดาภิเษก” “หลวงประดิษฐ์ มนูธรรม” “ไมโครโปรเซสเซอร์” “ร้านเลมอนฟาร์ม” “นวมินทร์” “ขาโจ๋วัยจ๊าบ” ฯลฯ เหล่านี้ล้วนเป็นอุปสรรคที่สำคัญของการประมวลผลภาษาไทยด้วยคอมพิวเตอร์เป็นอย่างมาก เนื่องจากการประมวลผลภาษาไทยแทบทั้งหมดอาศัยฐานข้อมูลพจนานุกรมคำศัพท์ภาษาไทยเป็นหลักในการแบ่งคำและประมวลผลอื่น ๆ ต่อไป โดยที่นับวันคำที่ไม่รู้จักเหล่านี้จะมีเพิ่มมากขึ้นเรื่อย ๆ สาเหตุหนึ่งก็คือปัจจุบันมีความนิยมที่จะใช้คำทับศัพท์ภาษาต่างประเทศกันอย่างกว้างขวาง ในระบบการสังเคราะห์เสียงพูดภาษาไทยจากข้อความ (Thai text-to-speech) ก็เช่นเดียวกัน ปัญหาดังกล่าวนี้ทำให้ระบบนี้มีความผิดพลาดในการอ่านสูง และทำให้มีความสามารถในการอ่านข้อความทั่วไปในชีวิตประจำวันต่ำ งานวิจัยอื่น ๆ ใช้แนวทางแก้ปัญหาโดยอาศัยกฎการแปลงเสียงอักขระ (Rules Based Letter to Sound) แต่พบว่าแนวทางนั้นมีความอ่อนตัวต่ำ ในบางกรณีไม่สามารถอ่านคำที่อยู่นอกเหนือกฎได้ เช่นเคยพบว่าไม่สามารถอ่านคำสนธิกรณีของคำ “นวมินทร์” ซึ่งมาจากคำสองคำคือ “นวม” ซึ่งแปลว่าเก้าเชื่อมกับคำว่า “อินทร์” ซึ่งหมายถึงพระอินทร์

งานวิจัยนี้ได้พยายามใช้อีกแนวทางหนึ่งของการแก้ไขปัญหาดังกล่าวคือใช้วิธีการของตรรกแห่งความคลุมเครือ (Fuzzy Logic) แทนซึ่งพบว่าสามารถแก้ไขปัญหาได้ดีมากที่เดียว

2. ตรรกแห่งความคลุมเครือ

ทฤษฎีตรรกวิทยาในอดีตจะแบ่งโลกออกเป็นสองกลุ่มคือ “จริง” หรือ “เท็จ” หรืออีกนัยหนึ่งคือ “ศูนย์” หรือ “หนึ่ง” ทฤษฎีนี้ได้ถูกใช้มายาวนานในการแก้ปัญหาด้วยคอมพิวเตอร์ ซึ่งสามารถแก้ปัญหาได้ดีและเร็วสำหรับปัญหาที่มีความชัดแจ้งการเขียนโปรแกรมเพื่อแก้ปัญหาต่าง ๆ จะอาศัยหลักทางตรรกวิทยาเช่น “ถ้าเหตุการณ์นี้เป็นจริงและ/หรือเหตุการณ์โน้นเป็นเท็จให้ทำงานอย่างนั้น ถ้าไม่ ให้ทำงานอีกอย่างหนึ่ง” เป็นต้น ในกรณีการสังเคราะห์เสียงพูดจะใช้หลักตรรกวิทยาที่ว่า “ถ้าเจอหน่วยเสียงในพจนานุกรม ให้สังเคราะห์เสียงออกไป ถ้าไม่ ให้แสดงข้อความว่าไม่สามารถสังเคราะห์เสียงได้” เมื่อมนุษย์มีความต้องการที่จะแก้ปัญหาที่ซับซ้อนและคลุมเครือมากขึ้น พบว่าทฤษฎีดังกล่าวมีข้อจำกัดเป็นอันมาก ในปี ค.ศ. 1964 ศาตราจารย์ล็อตฟี ซาดาห์ (Lotfi Sadeh) แห่งมหาวิทยาลัยแคลิฟอร์เนียวิทยาเขตเบิร์กเล่ย์ได้คิดค้นและเผยแพร่ทฤษฎีใหม่คือทฤษฎีตรรกแห่งความคลุมเครือ (Fuzzy Logic) [1], [5] ซึ่งมีความเหมาะสมที่จะใช้ในการแก้ปัญหาที่ซับซ้อนขึ้นได้ดีกว่า โดยหลักปรัชญาของทฤษฎีนี้แล้วสิ่งต่าง ๆ สามารถมีสถานะได้มากกว่าสองสถานะ เช่น “เท็จ” “ไม่น่าจะจริง” “อาจจะจริง” “เป็นจริงได้” “ค่อนข้างที่จะจริง” “คงจะจริง” “จริงพอสมควร” “จริงมาก” “จริงมากกว่า” “จริงมากที่สุด” และ “จริง” หรืออาจเทียบกับปริมาณความคลุมเครือเป็นดังตารางข้างล่าง

WORDS	VALUES FUZZY
จริง	1
จริงมากที่สุด	0.90
จริงมากกว่า	0.80
จริงมาก	0.70
จริงพอสมควร	0.60
คงจะจริง	0.50
ค่อนข้างที่จะจริง	0.40
เป็นจริงได้	0.30

อาจจะจริง	0.20
ไม่น่าจะจริง	0.10
เท็จ	0

ยกตัวอย่างเรื่องคนสูงกับคนเตี้ย ถ้าสมศรี สมชาย สมควร และสมศักดิ์ สูง ห้าฟุต ห้าฟุตครึ่ง หกฟุต และหกฟุตครึ่ง แล้ว ถ้าเรามองโดยใช้แนวคิดแบบตรรกวิทยา เราอาจบอกว่าสมศรีกับสมชายตัวเตี้ย และสมควรกับสมศักดิ์ตัวสูง แต่ถ้ามีสมปองเพิ่มเข้ามาอีกคน โดยที่เขามีความสูง ห้าฟุตเจ็ดนิ้ว ก็จะเกิดความคลุมเครือที่ว่าสมปองควรจะจัดอยู่ในกลุ่มพวกที่สูงหรือเตี้ยดี ลองมองปัญหาโดยใช้แนวคิดตรรกแห่งความคลุมเครือบ้าง จะเห็นว่าทุกคนเป็นสมาชิกของกลุ่มคนที่มีความสูงทั้งหมด โดยแต่ละคนมีสัดส่วนความเป็นสมาชิกที่แตกต่างกันไป เช่นสมศรีมีสัดส่วนเป็น 0.35 สมชาย 0.55 สมปอง 0.6 สมควร 0.75 และสมศักดิ์มีสัดส่วนเป็น 0.95 ศาสตราจารย์ซาดาห์เรียกกลุ่มของสมาชิกใด ๆ ว่า เซ็ตของความคลุมเครือ (Fuzzy Set) และเรียกสัดส่วนของความเป็นสมาชิกว่าค่าของความคลุมเครือ (Fuzzy Values)

ในทางตรรกวิทยาถ้าสมศรีเป็นสมาชิกของคนตัวเตี้ยแล้ว เธอก็จะไม่ใช่สมาชิกของคนตัวสูง แต่ในทางตรรกแห่งความคลุมเครือถ้าสมชายมีสัดส่วนเป็นสมาชิกของชมรมวิทยาศาสตร์ 0.75 แล้ว อีก 0.25 ที่เหลือในชีวิตเขาอาจมีสัดส่วนของความเป็นสมาชิกของชมรมประวัติศาสตร์ก็ได้ นี่คือแก่นหลักที่บ่งชี้ว่าในปัญหาที่มีความคลุมเครือ ตรรกวิทยามีความจำกัดในการแก้ปัญหาเป็นอย่างมาก และตรรกแห่งความคลุมเครือสามารถแก้ปัญาได้อย่างกว้างขวางและแม่นยำกว่า แต่มีข้อระวังว่าในปัญหาที่มีความชัดแจ้ง การแก้ไขโดยใช้ตรรกแห่งความคลุมเครืออาจใช้ต้นทุนในการแก้ไขปัญหาสูงกว่า

ในกรณีของการสังเคราะห์เสียงด้วยเทคนิคตรรกแห่งความคลุมเครือจะคำนวณหาค่าความคลุมเครือของหน่วยเสียงต่าง ๆ ในพจนานุกรมหน่วยเสียงเทียบกับคำหรือพยางค์ที่ต้องการสังเคราะห์ แล้วให้เลือกสังเคราะห์หน่วยเสียงที่มีค่าความคลุมเครือดีที่สุดออกไป เช่นคำว่า “ลินคอห์น” อาจจะพบว่าค่าความคลุมเครือของหน่วยเสียงใด ๆ เมื่อเทียบกับพยางค์ “ลิน” ไม่มีค่าที่เป็นหนึ่งเลยแต่มีหน่วยเสียงที่ได้ค่าความคลุมเครือสูงสุดสองอันดับแรกคือ “ลิ้น” มีค่าเป็น 0.99 และ “ลิง” มีค่าเป็น 0.92 จึงให้เลือกสังเคราะห์เสียงเป็น “ลิ้น-คอน” เป็นต้น

ทฤษฎีตรรกแห่งความคลุมเครือนี้เป็นต้นแบบสำหรับแนวคิดแก้ปัญหาแบบคลุมเครือวิธีใหม่ ๆ เช่นวิธีการเชิงพันธุศาสตร์ ที่อาศัยหลักสมมติฐานที่ว่า ถ้าเราคัดยีนที่มีลักษณะดีของพ่อมาผสมกับยีนที่มีลักษณะดีของแม่ ก็จะได้พันธ์ใหม่ที่มีลักษณะดีกว่าทั้งพ่อและแม่ ยกตัวอย่างกรณีของการสังเคราะห์เสียงพูดคำว่า “ลินคอห์น” เราเห็นว่าหน่วยเสียง “ลิ้น” มียีนที่ดีคือ มีอักษรนำ สระต้น และตัวสะกดเหมือนกับ “ลิน” ซึ่งเป็นหน่วยเสียงเป้าหมาย มียีนด้อยคือ เสียงโทซึ่งแตกต่างไป และ ”ลิง” มียีนที่มีลักษณะดีคือ อักษรนำ สระต้น และเสียงสามัญเหมือน “ลิน” ยีนที่ด้อยคือตัวสะกดที่ไม่เหมือนกัน ดังนั้นถ้าเราตัดยีนที่ด้อยของทั้ง “ลิ้น” และ “ลิง” ทิ้งไปแล้วนำเฉพาะยีนที่ดีทั้งหมดของทั้งสองมารวมกันเราก็จะได้หน่วยเสียงใหม่ที่มีความใกล้เคียง ”ลิน” มากกว่าทั้ง “ลิ้น” และ ”ลิง” ซึ่งเป็นพันธุพ่อและแม่เป็นต้น

แนวความคิดตรรกแห่งความคลุมเครือ (Fuzzy Logic) ได้ถูกคิดค้นขึ้นในอเมริกา และได้รับการศึกษาค้นคว้าและพัฒนาเป็นผลิตภัณฑ์ในชีวิตประจำวันมากมายโดยประเทศญี่ปุ่น ไม่ว่าจะเป็นรถไฟฟ้าที่มีการออกตัวและเข้าเทียบจอดชานชาลาอย่างนิ่มนวลโดยที่ผู้โดยสารไม่มีความจำเป็นที่จะต้องยึดเกาะราวทรงตัวแต่อย่างไร เครื่องควบคุมอุณหภูมิของเครื่องปรับอากาศที่มีความสบายดังควบคุมด้วยความรู้สึก ระบบคอมพิวเตอร์ที่ใช้ควบคุมระบบเกียร์อัตโนมัติของรถยนต์สมัยใหม่ที่มีความเรียบนุ่ม ตู้เย็นที่รู้ว่าในเวลากลางคืนอัตราการเปิดปิดประตูใช้งานต่ำจึงควบคุมการทำงานให้ประหยัดพลังงานมากขึ้น ระบบคอมพิวเตอร์ที่รู้จำเสียงพูด ตัวเขียนรวมไปถึงระบบคอมพิวเตอร์ที่สามารถสังเคราะห์ อ่านพูด ออกเสียงคำที่ไม่รู้จัก ผลิตภัณฑ์เหล่านี้ล้วนได้รับการพัฒนาออกแบบมาโดยอาศัยหลักปรัชญาของตรรกแห่งความคลุมเครือทั้งสิ้นเราจะพบว่าความละเอียดแม่นยำหรือแม้แต่การประหยัดพลังงานของระบบที่อาศัยแนวคิดดังกล่าวนี้มีสูงกว่าระบบธรรมดาเป็นอันมาก

3. สถาปัตยกรรมของระบบ

ทีมงานวิจัยได้ทำการวิจัยและพัฒนาระบบการสังเคราะห์เสียงพูดภาษาไทยทั้งแต่ปี พ.ศ. 2541 จนถึงปี พ.ศ. 2542 ก็สามารถพัฒนาระบบการสังเคราะห์เสียงพูดแบบทั่ว ๆ ไปได้ และได้พัฒนาต่อเนื่องจนถึงปี พ.ศ. 2543 ก็สามารถพัฒนาระบบที่สามารถสังเคราะห์คำที่ไม่ปรากฏในพจนานุกรมหรือคำที่ไม่รู้จักได้ในระดับความละเอียดแม่นยำสูงประมาณ 96.69% สำหรับเอกสารที่ประกอบด้วยคำที่ไม่รู้จักทั้งหมด และมีความละเอียดแม่นยำสูงถึง 99.59% สำหรับเอกสารทั่ว ๆ ไปที่ประกอบด้วยคำที่รู้จักและคำที่ไม่รู้จักผสมปนเปกันไป สถาปัตยกรรมของระบบการสังเคราะห์เสียงที่ได้ออกแบบไว้ซึ่งประกอบไปด้วยหกส่วนหลักคือส่วนแรกเป็นการตัดคำ (Thai word segmentation algorithm) [2], [3] ทำหน้าที่แบ่งประโยคภาษาไทยซึ่งเขียนติดกันเป็นพืดให้เป็นคำ ส่วนที่สองคือส่วนวจีวิภาค (Parsing Algorithm) คือส่วนที่ทำหน้าที่แบ่งคำที่ไม่รู้จักเป็นพยางค์ย่อย ๆ ส่วนที่สามคือส่วนตามหาคำที่ไม่รู้จัก (Unknown Words Identification) ส่วนที่สี่คือส่วนพจนานุกรมการอ่านออกเสียง (Dictionary Based Pronunciation) ส่วนที่ห้าคือส่วนการประมาณหน่วยเสียง (Approximation Matching Algorithm) หรือคือส่วนที่ทำการคำนวณหาค่าความคลุมเครือ (Fuzzy Values) เพื่อจัดกลุ่มความคลุมเครือ (Fuzzy Set) ส่วนที่หกคือส่วนประมวลผลสัญญาณและสังเคราะห์เสียง (Digital Signal Processing and Synthesis) ทีมงานวิจัยได้สร้างระบบต้นแบบเพื่อให้บุคคลทั่วไปสามารถทดลองใช้งานได้ทางโทรศัพท์โดยได้ทำการออกแบบส่วนเชื่อมต่อกับโทรศัพท์ (CTI : Computer Telephony Integration) ดังรูป

โดยผู้ใช้งานสามารถดูรายละเอียดการให้บริการและเบอร์โทรศัพท์ที่จะหมุนมาทดลองใช้งานได้จากแม่ข่ายอินเทอร์เน็ต

"www.geocities.com/pisitp"

4. วจีวิภาค

วิจีวิภาค (Parsing Algorithm) หรือขั้นตอนและวิธีการแบ่งพยางค์จากคำใด ๆ ทั้งที่รู้จักและไม่รู้จักในงานวิจัยนี้ทีมงานได้พัฒนากฎ 16 กฎจากหลักภาษาไทยของอาจารย์กำชัย ทองหล่อเป็นหลักโดยกฎดังกล่าวสามารถอธิบายได้ดังนี้

คำนิยาม

พ หมายถึง พยัญชนะไทย

(พ) หมายถึง พยัญชนะตัวควบกล้ำ ซึ่งอาจจะมีหรือไม่มีก็ได้

รร หมายถึง ตัว ร หัน

สนำ หมายถึง สระนำในภาษาไทย

สตาม หมายถึง สระตามในภาษาไทย

ด หมายถึง ตัวสะกดที่ใช้ในภาษาไทย

[ย,อ] หมายถึง พยัญชนะที่ใช้เป็นสระ ซึ่งจะขาดไม่ได้ เป็น ตัว ย หรือ ตัว อ เท่านั้น

<u>กฎข้อ 1.</u> กฎของการใช้ อักษร อ นำ ย ซึ่งกรณีนี้เป็นกรณีพิเศษ ได้แก่คำว่า อย่า อยู่ อย่าง

<u>กฎข้อ 2.</u> กฎของการใช้ อักษร อ นำ ย ซึ่งกรณีนี้เป็นกรณีพิเศษ ได้แก่คำว่า อยาก

<u>กฎข้อ 3</u>. กฎของการใช้ตัว ร. หัน ในกรณีที่มีตัวสะกดตามหลังสามารถแทนด้วยสัญลักษณ์คือ พ รร ด ตัวอย่างของคำเช่น พรรค วรรณ

<u>กฎข้อ 4.</u> กฎของการใช้ตัว ร. หัน ในกรณีที่ไม่มีตัวสะกดตามหลังสามารถแทนด้วยสัญลักษณ์คือ พ รร ตัวอย่างของคำเช่น พรรษา วรรณา

<u>กฎข้อ 5.</u> กฎของการใช้ สระ เอา เอาะ
สามารถแทนด้วยสัญลักษณ์คือ เ พ พ า (ะ)
ตัวอย่างของคำเช่น เพราะ เขลา เค้า

<u>กฎข้อ 6.</u> กฎของการใช้ คำที่อ่านออกเสียงเป็น เอิง สามารถแทนด้วยสัญลักษณ์คือ เ พ (พ) ด ตัวอย่างของคำเช่น เพลิง เพิง

<u>กฎข้อ 7.</u> กฎของการใช้ สระ เออะ เออ
สามารถแทนด้วยสัญลักษณ์คือ เ พ อ (ะ)
ตัวอย่างของคำเช่น เลอะเทอะ เผลอ เจอ

<u>กฎข้อ 8.</u> กฎของการใช้ สระ เอียะ เอีย เอือะ เอือ ตามด้วยตัวสะกด สามารถแทนด้วยสัญลักษณ์คือ เ พ (พ,พ) [ย,อ] ด ตัวอย่างของคำเช่น เพื่อน เรียน เกลียด

<u>กฎข้อ 9.</u> กฎของการใช้ สระ เอียะ เอีย เอือะ เอือ สามารถแทนด้วยสัญลักษณ์คือ เ พ (พ,พ) [ย,อ] (ะ) ตัวอย่างของคำเช่น เสีย เผียะ เพื่อ เกือะ

<u>กฎข้อ 10.</u> กฎของการใช้ สระนำ โอะ โอ ไอ ใอ เอะ เอ สามารถแทนด้วยสัญลักษณ์คือ ส นำ พ (พ) (ะ) ตัวอย่างของคำเช่น ไกล โกน โต๊ะ ใจ เกะ เห

<u>กฎข้อ 11.</u> กฎของการใช้ สระโอ ไอ ใอ เอ ที่มีตัวสะกดสามารถแทนด้วยสัญลักษณ์คือ ส นำ พ (พ) ด ตัวอย่างของคำเช่น โอ่ง โล่ง โกน เพลง

<u>กฎข้อ 12.</u> กฎของการใช้ พยัญชนะแล้วตามด้วยสระ ตัวสะกด สามารถแทนด้วยสัญลักษณ์คือ พ (พ) ส ตาม ด ตัวอย่างของคำเช่น กราม พราหม กัด

<u>กฎข้อ 13.</u> กฎของการใช้ พยัญชนะ สระสามารถแทนด้วยสัญลักษณ์คือ พ (พ) ส ตาม
ตัวอย่างของคำเช่น พระ กะทิ

<u>กฎข้อ 14.</u> กฎของการใช้ พยัญชนะ ตัวสะกด สามารถแทนด้วยสัญลักษณ์คือ พ (พ) ด ตัวอย่างของคำเช่น กด ห่ม

<u>กฎข้อ 15.</u> กฎของการใช้ พยัญชนะที่ตามด้วยตัวสะกด ร ล สามารถแทนด้วยสัญลักษณ์คือ พ [ร,ล]ตัวอย่างของคำเช่น พร พล กร กล

<u>กฎข้อ 16.</u> กฎของการใช้ พยัญชนะที่ออกเสียง อะ กึ่งเสียง สามารถแทนด้วยสัญลักษณ์คือ พ ตัวอย่างของคำเช่น ขจี อนัตตา สนิท

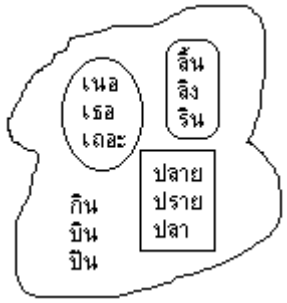
5. การติดตามหาคำที่ไม่รู้จัก

กระบวนการติดตามหาร่องรอยของคำที่ไม่รู้จักเพื่อแยกออกมาจากคำที่รู้จักใด ๆ ทำโดยเริ่มจากเมื่อเกิดการตัดคำผิดพลาดเกิดขึ้นให้นำเศษของคำไปรวมกับคำที่เพิ่งตัดแบ่งแยกมาได้ก่อนหน้านี้แล้วผ่านกระบวนการวจีวิภาคถ้าทำได้สำเร็จให้ถือว่าผลของการรวมกันคือคำที่ไม่รู้จักหนึ่งคำ แต่ถ้าทำไม่สำเร็จคือเหลือเศษของพยางค์ให้นำเศษที่เหลือไปรวมกับคำที่สามารถตัดได้หลังจากเศษของพยางค์นี้แล้วลองทำวจีวิภาคดู ถ้าทำสำเร็จให้ถือว่าผลของการรวมทั้งสองได้คำที่ไม่รู้จักสองคำ ถ้าทำไม่สำเร็จให้นำทั้งสามส่วนคือคำที่ตัดได้ก่อนหน้า เศษของคำ และคำที่ตัดได้หลังเศษคำมารวมกันเป็นคำที่ไม่รู้จักหนึ่งคำ หลังจากนี้ทำกระบวนการต่อไปคือวจีวิภาคกับคำที่ไม่รู้จักทั้งหมดเพื่อแยกพยางค์และประมาณหน่วยเสียงต่อไป

6. การประมาณหน่วยเสียง

การประมาณหน่วยเสียงภาษาไทย (Approximation Matching Algorithm for Thai) ในบางกรณีอาจเรียกว่าการหาเหตุผลเชิงประมาณการ (Approximate Reasoning) [5] ทำโดยการคำนวณหาค่าความคลุมเครือ (Fuzzy Values) หรือสัดส่วนของความเป็นสมาชิกของเซ็ตของความคลุมเครือ (Fuzzy Set) ของพยางค์ที่ตัดแบ่งออกมาโดยส่วนวจีวิภาคที่กล่าวมาแล้วจากพจนานุกรมหน่วยเสียง

โดยให้พจนานุกรมหน่วยเสียงที่มีอยู่ 3,996 หน่วยคือเซ็ตของความคลุมเครือ (Fuzzy Set) ที่มีเซ็ตย่อยคือหน่วยเสียงต่าง ๆ ที่มีความคล้ายกัน ดังรูป



หน่วยเสียงต่าง ๆ จะมีสัดส่วนของความเป็นสมาชิกของเซ็ตย่อยใด ๆ ไม่เท่ากัน สัดส่วนของความเป็นสมาชิกต่าง ๆ นี้คือค่าของความคลุมเครือ (Fuzzy Values)

การคำนวณหาค่าความคลุมเครือทำโดยการใช้มาตรวัด (Metrics) 6 ตัวคือความใกล้เคียงกันของสระหน้า (M1) ความใกล้เคียงกันของพยัญชนะต้น (M2) ความใกล้เคียงกันของพยัญชนะควบกล้ำ (M3) ความใกล้เคียงกันของสระตาม (M4) ความใกล้เคียงกันของตัวสะกด (M5) และความใกล้เคียงกันของวรรณยุกต์ (M6) ทำการวัดเปรียบเทียบระหว่างพยางค์ต้นทางกับหน่วยเสียงที่มีอยู่ในพจนานุกรมหน่วยเสียงที่ต้องการสังเคราะห์ จะได้เวกเตอร์ของคลามคลุมเครือของหน่วยเสียงใด ๆ เป็น

Vf = [M1, M2, M3, M4, M5, M6]

เมื่อ Vf คือเวกเตอร์ของความคลุมเครือของหน่วยเสียงใด ๆ

Mi คือมาตรวัดต่าง ๆ ของความใกล้เคียงทางเสียงใด ๆ

โดยให้ค่าน้ำหนักที่แตกต่างกันไปของแต่ละมาตรวัด กรณีของพยัญชนะต้นมีการแบ่งออกเป็นสอง

กลุ่มคือกลุ่มที่มีรูปเสียงเดียวกัน [4] ดังแสดงในตารางข้างล่าง

ก	ฏ ต	ฝ พ
ช ซ ฌ ญ ฆ	ฏ ต	ม
ง	ฐ ท ฒ ณ ธ ฑ	ร
จ	ฌ น	ล ฟ
ฉ ช ฌ	บ	ว
ช ศ ษ ส	ป	ห ฮ
ญ ย	ผ พ ภ	อ

อีกกลุ่มคือกลุ่มที่มีฐานกำเนิดของเสียงเดียวกัน ตามหลักภาษาไทยของอาจารย์กำชัย ทองหล่อ คือ

เกิดจากฐานคอ	ก ข ค ฆ ฃ ง	เรียกว่า วรรค ก
เกิดจากฐานเพดาน	จ ฉ ช ซ ฌ ญ	เรียกว่า วรรค จ
เกิดจากฐานปุ่มเหงือก	ฏ ฏ ฐ ฑ ฒ ณ	เรียกว่า วรรค ฏ
เกิดจากฐานฟัน	ด ต ถ ท ธ น	เรียกว่า วรรค ด
เกิดจากฐานริมฝีปาก	บ ป ผ ฝ พ ฟ ภ ม	เรียกว่า วรรค บ
เกิดจากฐานต่างๆ กัน	ย ร ล ว ศ ษ ส ห ฬ อ ฮ	เรียกว่า เศษวรรค

พยัญชนะ 5 วรรคข้างต้น คือ ตั้งแต่ตัว ก ถึง ม มี 33 ตัวรวมเรียกว่า พยัญชนะวรรค เพราะจัดเข้าเป็นหมู่กันได้ตามฐานที่เกิด และจะใช้พยัญชนะควบกล้ำในพวกเดียวกันไม่ได้ พยัญชนะวรรค ที่ 6 คือ ตัว ย ถึง ตัว ฮ มี 11 ตัว เรียกว่าพยัญชนะ เศษวรรค เพราะมีเสียงและฐานที่เกิดแตกต่างกันซึ่งจะจัดเข้าพวกกับพยัญชนะวรรคไม่ได้ จึงต้องจัดแยกเป็นวรรคพิเศษ ถ้าอยู่ในสองกลุ่มนี้จะได้คะแนนสองสุดเป็นสองอันดับแรก และถ้าอยู่นอกสองกลุ่มนี้จะได้คะแนนต่ำมาก

และมีการให้คะแนนในกรณีที่เป็นคำควบกล้ำ สระตาม คำที่เป็นตัวสะกดแม่ต่าง ๆ และเสียงวรรณยุกต์ที่ต่างกัน นำคะแนนที่ได้มาหาค่าคะแนนเฉลี่ยดังสูตรด้านล่างนี้

$$Vf_{avg} = 1/n * [\sum_{i=1}^{6} M[i] * W_i / (W_i * freq_i)]$$

เมื่อ Vf คือค่าความคลุมเครือ (Fuzzy Values)

M[i] คือ มาตรวัดต่าง ๆ ของความใกล้เคียงทางเสียง

W คือ น้ำหนักคะแนนที่ให้ในแต่ละมาตรวัด

freq คือ ความถี่ของการใช้มาตรวัดใด ๆ

n คือ ค่าคะแนนสูงสุดที่เป็นไปได้

i = 1 คือ การวัดความใกล้เคียงทางเสียงของสระหน้า

i = 2 คือ การวัดความใกล้เคียงทางเสียงของพยัญชนะต้น

i = 3 คือ การวัดความใกล้เคียงทางเสียงของพยัญชนะควบกล้ำ

i = 4 คือ การวัดความใกล้เคียงทางเสียงของสระตาม

i = 5 คือ การวัดความใกล้เคียงทางเสียงของตัวสะกด

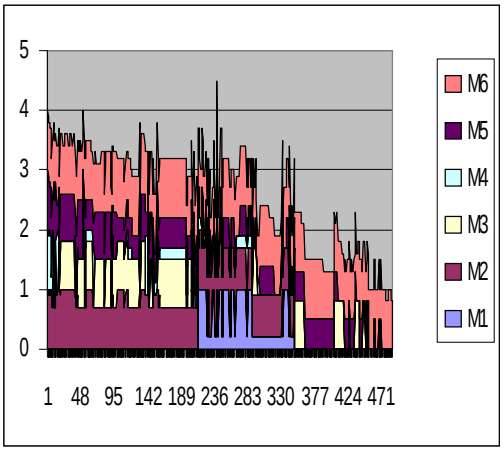
i = 6 คือ การวัดความใกล้เคียงทางเสียงของวรรณยุกต์

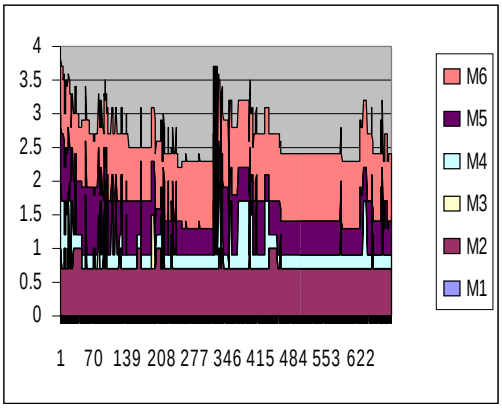
ยกตัวอย่างสมมุติหน่วยเสียงเป้าหมายได้แก่ “กราบ “ คำที่ต้องการพิจารณาได้แก่ “กาด” ขั้นแรกจะไม่มีสระหน้าทำให้คะแนน M[1] เป็น 0 ต่อไปจะพิจารณาพยัญชนะต้น ได้แก่ ตัว ก ซึ่งคะแนน M[2] ที่ได้จะเป็น 1.0 คูณกับค่า w2 ทำให้ได้คะแนนเป็น 11 ขั้นตอนที่สองจะเปรียบเทียบ พยัญชนะควบกล้ำ ได้แก่ ตัว ร ซึ่งปรากฏว่าคะแนน M[3] ที่ได้จะเป็น 0.2 คูณกับค่า w3 ทำให้ได้คะแนนเป็น 1.4 ขั้นตอนที่สามจะเปรียบเทียบ สระตาม ได้แก่ สระ อา คะแนน M[4] ที่ได้จะเป็น 1.0 คูณกับค่า w4 ทำให้ได้คะแนนเป็น 9.0 ขั้นตอนที่สี่จะเปรียบเทียบตัวสะกด ได้แก่ แม่กบ กับ แม่ กด คะแนน M[5] ที่ได้เป็น 0.8 คูณกับค่า w5 ทำให้ได้คะแนนเป็น 8.8 นำคะแนนทั้งหมดมารวมกันได้ 30.2 คะแนน + 1 คะแนนในกรณีที่เป็นวรรณยุกต์เดียวกัน M[6] รวมเป็น 31.2 นำค่าที่ได้หารด้วย weight ทั้งหมด นั่นคือคะแนน = 31.2/39 = 0.80 คือค่าเฉลี่ยของความคลุมเครือที่ได้ ตารางข้างล่างแสดงตัวอย่างค่าความคลุมเครือของพยางค์เป้าหมายต่าง ๆ ที่เกิดขึ้นเมื่อพยางค์ต้นทางคือ “ลิน” และ “คอห์น” ตามลำดับ

UNITS	VALUES Fuzzy	M[I]
ลิ้น	0.99	[0 1 0 1 1 0.8]
ลิง	0.92	[0 1 0 1 0.8 1]
อิน	0.88	[0 0.7 0 1 1 1]
สิน	0.88	[0 0.7 0 1 1 1]

UNITS	VALUES Fuzzy	M[I]
คอน	1.00	[0 1 0 1 1 1.1]
ขอน	0.96	[0 0.9 0 1 1 1]
คอย	0.92	[0 1 0 1 0.8 1]
คอม	0.92	[0 1 0 1 0.8 1]

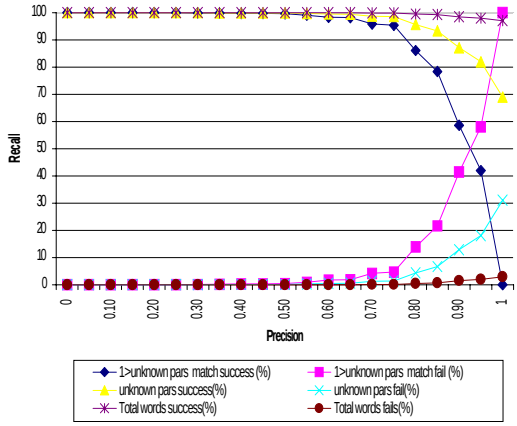
รูปข้างล่างแสดงผลการนำเวกเตอร์ของความคลุมเครือของ“ลิน” และ “คอห์น” ตามลำดับไปพลอตกราฟ





7. ผลการทดลองและประเมินประสิทธิภาพ

ได้ทำการทดลองหลายครั้งกับข้อความภาษาไทยคือ “แผนพัฒนาเศรษฐกิจและสังคมแห่งชาติฉบับที่ 8” ซึ่งประกอบด้วยข้อความภาษาไทยที่ผสมปนเประหว่างคำที่รู้จักจำนวน 55,573 คำและไม่รู้จักจำนวน 3,249 คำ โดยทำการค่อย ๆ เปลี่ยนแปลงค่าความยอมรับได้ของค่าความคลุมเครือน้อยที่สุดที่จะยอมสังเคราะห์หรือมองอีกนัยหนึ่งคือความแม่นยำชัดเจน (Precision)



ของเสียงที่สังเคราะห์ได้หรือความสามารถในการอ่านคำได้ใกล้เคียงตั้งแต่ 0 ถึง 1 พบว่าได้ค่าเฉลี่ยของความแม่นยำ เป็น 99.59% และถ้าแยกพิจารณาเฉพาะคำที่ไม่รู้จักแล้วพบว่าได้ค่าเฉลี่ยดังกล่าวเป็น 96.69% ตำแหน่งที่มีความแม่นยำและในเวลาเดียวกันมีความครอบคลุม (Recall) สูงสุดคือ 98% และ 88% ในกรณีของภาษาไทยทั่ว ๆ และในกรณีของข้อความที่ประกอบด้วยคำที่ไม่รู้จักทั้งหมดตามลำดับ รูปข้างบนคือผลการทดลองดังกล่าว

8. สรุป

จากผลการทดลองที่ได้สามารถสรุปได้ว่าแนวคิดแบบตรรกแห่งความคลุมเครือสามารถช่วยแก้ปัญหาการสังเคราะห์เสียงพูดคำที่ไม่รู้จักภาษาไทยได้ดีมากทีเดียวและมีความยืดหยุ่นในการทำงานสูงกล่าวคือถึงแม้ว่ามีคำใหม่ ๆ ที่กฎเกณฑ์ต่าง ๆ อาจยังไม่ครอบคลุมไปถึงในอนาคตก็สามารถอ่านออกเสียงไปได้และมีความสามารถที่จะเลือกความแม่นยำที่อาจมีค่าแตกต่างกันที่เหมาะสมกับการประยุกต์ใช้งานแตกต่างกันได้

เอกสารอ้างอิง

[1]. Daniel Mcneill, Paul Freiberger “Fuzzy Logic, The Discovery of a Revolutionary Computer Technology and how it is changing our World”, Simon & Schuster, USA, 1993

[2]. Pisit Promchan, Dr. Yunyong Tengamnuay, “Performance Comparison of Thai Words Segmentation Algorithms”, NCSEC98 conference, KU, BKK, Thailand, 1998

[3]. Pisit Promchan, “Analysis of Guidelines for Performance Comparison of Thai Words Separation Programs”, Master Thesis, dept. of Computer engineering, Chulalongkorn University, BKK, Thailand, 1997

[4]. P. Suwanvisat and S. Prasitjutrakul, "Transliterated Word Encoding and Retrieval Algorithms for Thai-English Cross-Language Retrieval" the National Computer Science and Engineering Conference 1999, Bangkok, Thailand

[5] Hans Bandemer, Siegfried Gottwald, “Fuzzy Sets Fuzzy Logic Fuzzy Methods with Applications”, John Wiley & Sons, England 1996

[6] Dr. Ampai P., Kreangchai C., Pisit P., Pakin J., “Thai Text-to-Speech”, Research report #1 submitted to national research council, May, 1999.