บทที่ 1.

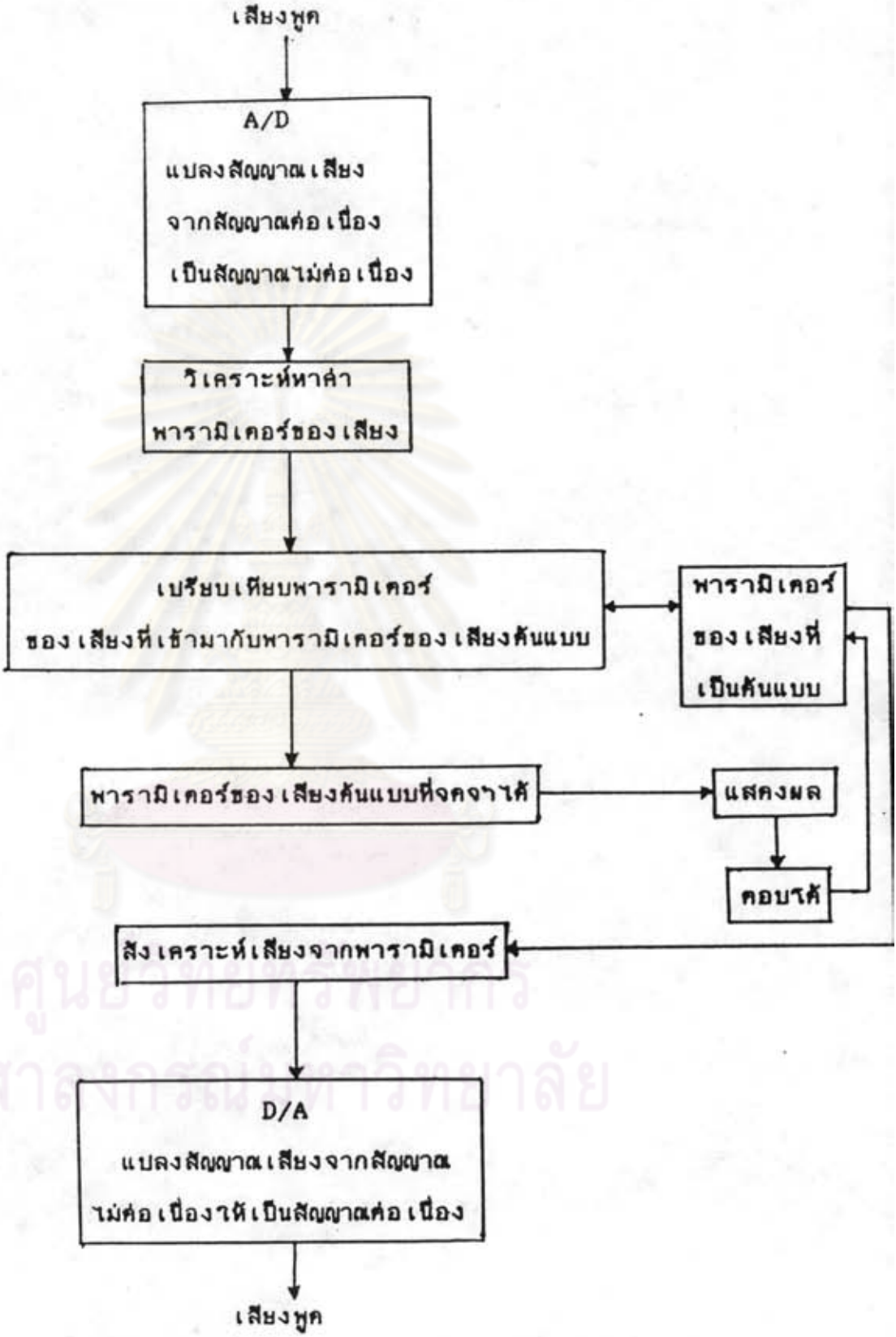
บทนำ

1.1 ความเป็นมาของปัญหา

ปัจจุบันเครื่องคอมพิวเตอร์ถูกนำมาใช้กันอย่างแพร่หลาย เพื่อช่วยในการประมวลผลให้เป็นไปอย่างรวดเร็วและถูกต้อง ซึ่งการติดต่อระหว่างเครื่องคอมพิวเตอร์กับมนุษย์โดยปกติเครื่องจะรับคำสั่งหรือข้อมูลทางแป้นพิมพ์ (key board) และจะแสดงผลทางจอภาพ (terminal) หรือเครื่องพิมพ์ (printer) แป้นพิมพ์จะทำหน้าที่เปลี่ยนข้อมูลหรือคำสั่งที่มนุษย์กดบนแป้นพิมพ์ ให้เป็นภาษาสำหรับเครื่อง (0 กับ 1) การกดแป้นพิมพ์ต้องอาศัยความชำนาญ เพราะทำงานด้วยมือ ดังนั้นจึงมีโอกาสผิดพลาดได้ง่าย และไม่สะดวกสำหรับผู้ที่ไม่คุ้นเคยกับแป้นพิมพ์ [1]

โดยปกติ คนเรามักติดต่อกันด้วยเสียงพูด เนื่องจากเป็นวิธีสื่อสารที่สะดวกและรวดเร็ว จากปัญหาในการติดต่อกับเครื่องทางแป้นพิมพ์ ดังได้กล่าวมาแล้ว ทำให้มนุษย์มีความต้องการที่จะสามารถติดต่อกับเครื่อง (man-machine communication) โดยเสียงพูด ซึ่งการติดต่อกันในระบบนี้ เครื่องจะต้องมีหน่วยรับข้อมูล (input) ที่สามารถรับเสียงของมนุษย์ และสามารถแปลเป็นสัญญาณเครื่อง (speech analyzer) และจะต้องมีส่วนผลลัพท์(output) ที่จะแปลงสัญญาณเครื่องเป็นเสียงในภาษามนุษย์ (speech sythesizer) ซึ่งสามารถเขียนเป็นโครงสร้าง (block diagram) ของระบบได้ดังรูปที่ 1.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 1 ระบบการติดต่อกับเครื่องโดยใช้เสียงพูด

เนื่องจากสัญญาณเสียงที่มนุษย์ใช้ในการติดต่อสื่อสารกันนั้น เป็นสัญญาณของคลื่นเสียง (sound wave) ซึ่งโดยปกติทั่ว ๆ ไปแล้ว จะมีรูปแบบของภาษาที่ใช้ในการติดต่อเหมือนกัน (ในภาษาเดียวกัน) และรูปแบบ (pattern) ของคลื่นเสียงที่คนเราพูดออกมานั้น จะคล้าย ๆ กัน เนื่องจากสัญญาณเสียงเป็นสัญญาณ ที่ถูกทำให้ผิดเพี้ยน(distort) ได้ง่ายจึงทำให้สัญญาณเสียงของแต่ละคนแตกต่างกัน ซึ่งขึ้นอยู่กับ [2]

1. ความเร็วในการพูด
2. เสียงรบกวน
3. การออกเสียงที่ผิดหลักเกณฑ์ทางภาษา

แต่ถ้าเราสามารถหาบลักษณะร่วมกันของสัญญานเสียงของมนุษย์ได้ เราก็จะสามารถอ้าศัยลักษณะร่วมนั้นเป็นตัวชี้ ที่จะใช้ในการแปลงจากสัญญานเสียงเป็นสัญญานเครื่องได้

โดยหลักเกณฑ์ในภาษาไทยแล้ว เสียงที่คนเปล่งออกมาแต่ละครั้ง ซึ่งประกอบไปด้วย เสียงสระ เสียงพยัญชนะ และ เสียงวรรณยุกต์ หน่วยซึ่งประกอบด้วยเสียงทั้ง 3 นี้จะมีความหมายหรือไม่มีความหมายก็ได้ เราเรียกหน่วยนี้ว่า "พยางค์" พยางค์นี้ก็คือ หน่วยที่เล็กที่สุด [3]ที่คน เราจะใช้ในการติดต่อสื่อสารกัน ดังนั้นถ้าเราสามารถสร้าง ระบบที่สามารถแยกสัญญาณเสียงพูดให้เป็นพยางค์ได้ ก็จะทำให้การพัฒนาระบบการตรวจรู้เสียงพูด [Speech Recognition System] ง่ายขึ้น

การวิจัยนี้จึงมุ่งที่จะทำการศึกษาถึงวิธีการและแนวทาง ในการที่จะตัดพยางค์ในภาษาไทยในประโยคต่อเนื่อง ให้ถูกต้องมากที่สุด แล้วจึงนำพยางค์ที่ตัดได้ไปเปรียบเทียบกับพยางค์ที่เป็นต้นแบบ เพื่อให้เครื่องสามารถตรวจรู้เสียงพูดในประโยคต่อเนื่องได้โดยใช้ต้นแบบเพียงหน่วยเสียงของพยางค์เท่านั้น

1.2 <u>วัตถุประสงค์</u>

1. เพื่อทำการศึกษาถึงวิธีการและแนวทางในการพัฒนาให้คอมพิวเตอร์สามารถตัดเสียงพูดออกเป็นพยางค์ ได้โดยอัตโนมัติ

2. เพื่อทำการศึกษาถึงวิธีการและแนวทางที่จะให้เครื่องสามารถรับฟังเสียงพูดของมนุษย์ในประโยคต่อเนื่องได้ โดยอาศัยหน่วยพยางค์เป็นต้นแบบ

3. เพื่อเป็นแนวทางในการพัฒนาระบบการตรวจรู้เสียงพูดในประโยคต่อเนื่องให้สมบูรณ์มากขึ้น

1.3 ประโยชน์ที่จะได้จากการวิจัยนี้

1. เป็นแนวทางที่จะช่วยให้มนุษย์สามารถสื่อสารกับคอมพิวเตอร์โดยเสียงพูดได้

2. เป็นแนวทางในการประยุกต์ใช้งานด้าน การใช้เสียงพูดเป็นตัวตรวจสอบผู้พูด(speaker identification)

3. เพื่อการศึกษาทางภาษาศาสตร์

1.4 ขอบเขตของการวิจัย

ในการวิจัยนี้ขั้นแรก จะทำการค้นคิดวิธีการ(Algorithm) และโปรแกรมเพื่อใช้ในการตัดพยางค์ของเสียงพูด ซึ่งพูดเป็นประโยคต่อเนื่องโดยมีขอบเขตดังนี้

1. ผู้พูดจะต้องพูด ดังพอสมควรและจะต้องไม่เร็วจนเกินไป

2. เสียงรบกวนของสภาพแวดล้อมต้องอยู่ในระดับต่ำ

3. เสียงที่พูดนั้นจะต้องเป็นเสียงพูดภาษาไทย

การตัดพยางค์ ใช้กลุ่มคำดังปรากฏในภาคผนวก ค. เป็นตัวอย่าง หลังจากนั้นจะให้ระบบตรวจรู้เสียงพูดตัวเลขศูนย์ถึงเลขเก้าที่เป็นประโยคต่อเนื่อง โดยอาศัยเสียงพูดต้นแบบเป็นหน่วยเสียงพยางค์

1.5 ขั้นตอนและวิธีดำเนินการวิจัย

1. ศึกษาและวิเคราะห์เกี่ยวกับลักษณะและองค์ประกอบของเสียงพูด
2. หาหลักเกณฑ์ที่จะมาใช้เป็นกฎเกณฑ์ในการตัดพยางค์
3. ออกแบบระบบซอฟท์แวร์ และโปรแกรมที่จะใช้ในการตัดพยางค์และตรวจรู้เสียงพูด
4. ทดสอบการทำงาน
5. สรุปผลการวิจัยและข้อเสนอแนะ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย