

# ระบบการแปลอักษรเบรลล์ภาษาไทย - ภาษาอังกฤษ เป็นอักษรปกติ

## Thai-English Braille Translate to Text

วลัยนุช ธรรมนิตยกุล\* และ รัชฎา คงคะจันทร์\*

### บทคัดย่อ

บทความนี้ต้องการนำเสนอหลักการแปลเอกสารจากอักษรเบรลล์ชนิดสองภาษา (ไทยและอังกฤษ) เป็นอักษรปกติ โดยหลักการนี้เป็นขั้นตอนของการประมวลผลขั้นปลาย ซึ่งเป็นขั้นตอนหนึ่งของระบบการรู้จำภาพอักษรเบรลล์ และเนื่องจากรหัสอักษรเบรลล์หนึ่งๆ สามารถแปลเป็นได้ทั้งภาษาไทยและอังกฤษ ดังนั้นบทความนี้จึงนำเสนอกฎสำหรับการวิเคราะห์อักษรเบรลล์และพจนานุกรมอักษรเบรลล์เพื่อนำมาใช้สำหรับคัดกรองตัวอักษร ที่ไม่ถูกต้องออกจากนั้นจึงนำตัวอักษรที่ผ่านการคัดกรองมารวมเป็นคำ และประโยค จากนั้นใช้ N-Gram ในการคำนวณคะแนนของแต่ละประโยคร่วมกับข้อมูลจากคลังข้อมูล Orchid โดยที่ประโยคที่มีคะแนนของความเป็นไปได้สูงสุดจะถูกเลือกเป็นคำตอบ บทความนี้ประเมินประสิทธิภาพของระบบโดยทดสอบกับข้อมูลของเอกสารจำนวน 50 หน้า และทำการเปรียบเทียบกับผลที่ได้กับผู้เชี่ยวชาญ จำนวน 5 คน ซึ่งระบบให้เปอร์เซ็นต์ความถูกต้องมากกว่า 80%

**คำสำคัญ:** กฎการอ่านเบรลล์ พจนานุกรมอักษรเบรลล์ ทฤษฎี N-Gram คลังข้อมูล Orchid Corpus

### Abstract

This paper proposes an approach for translating bilingual Braille documents (Thai and English) to normal characters. The technique is used as a post-processing of Braille character recognition. Since Braille code can be translated to both Thai and English characters. Therefore, we propose the analytic rules extracted from human expert and Braille dictionary to filter the incorrect candidates. Then, the remaining candidates are combined to be a sentence. Each sentence is scored by

using n-gram with smoothing. The Orchid corpus is utilized as resource. The sentence with maximum score is selected as the results. We evaluate the system performance by testing with 50 pages of documents and compare the results with 5 human experts. The experimental results yield over 80% of accuracy.

**Keyword:** Rule base, Braille Dictionary, N-Gram, Orchid Corpus.

### 1. บทนำ

ในปัจจุบันงานวิจัยทางด้านระบบการแปลรหัสอักษรเบรลล์ ได้มีการพัฒนากันอย่างต่อเนื่อง ทั้งนี้เพื่อเป็นการลดช่องว่างในการติดต่อสื่อสารระหว่างผู้พิการทางสายตาและผู้ที่ไม่มีสายตาปกติ รวมทั้งยังช่วยเพิ่มโอกาสทางการเรียนรู้ให้ผู้พิการทางสายตา ที่จะแลกเปลี่ยนความรู้ระหว่างกันกับผู้ที่มีสายตาปกติได้

งานวิจัยทางการแปลอักษรเบรลล์ สามารถแบ่งออกเป็น 2 กลุ่ม ได้แก่ กลุ่มแรกคือกลุ่มของการแปลอักษรปกติเป็นอักษรเบรลล์ งานวิจัยในกลุ่มนี้ได้มีการพัฒนาออกมาเป็นซอฟต์แวร์สำเร็จรูปที่สามารถนำมาใช้งานได้จริง เช่น Duxbury Braille Translator [1] เป็นซอฟต์แวร์ตัวหนึ่งที่ได้รับคามนิยมอย่างแพร่หลาย เนื่องจากสามารถรองรับได้หลายภาษารวมทั้งภาษาไทยด้วยเช่นกัน และ WinBraille [2] เป็นซอฟต์แวร์ การแปลอักษรปกติเป็นอักษรเบรลล์ แต่รองรับเฉพาะภาษาอาราบิกเท่านั้น กลุ่มที่สองคือกลุ่มของการแปลอักษรเบรลล์เป็นอักษรปกติ มีงานวิจัยต่างๆ ดังนี้

Jiang และคณะ [3] ได้นำเสนองานวิธีการแปลรหัสอักษรเบรลล์สำหรับภาษาแมนดาริน โดยใช้ Multi-knowledge base ในการจัดการกับประโยคที่กำกวม และสร้าง Multi-level

\* คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์

graph โดยใช้ Virtabi algorithm ในการหาลำดับของตัวอักษรภาษาจีน ด้วย Maximum likelihood และใช้ N-Best algorithm เพื่อเลือกค่าของลำดับที่เป็นไปได้มากที่สุดมาเป็นคำตอบ แต่อย่างไรก็ตามงานวิจัยนี้สามารถรองรับได้เพียงแค่ภาษาแมนดารินเท่านั้น

King และคณะ [4] ได้พยายามเปรียบเทียบการพัฒนา ระบบการแปลอักษรปกติ ในแต่ละเครื่องมือที่ใช้ในการพัฒนา ระหว่างภาษาซี จาวา และแบบยูนิโคด โดยมุ่งเน้นไปที่การเปรียบเทียบว่าเครื่องมือใดมีประสิทธิภาพในด้านของเวลาที่ใช้ในการแปลมากกว่ากัน

ศรัณย์ [5] ได้นำเสนอระบบการหาตำแหน่งของรหัส อักษรเบรลล์ที่อยู่ในภาพ รวมทั้งแก้ปัญหาในเรื่องของการประมวลผลภาพ ที่มีลักษณะโน้มน้ำเอียง หรือภาพที่ไม่ชัดเจน จากนั้นทำการรู้จำภาพอักษรเบรลล์ โดยแปลเป็นแบบตัวต่อตัว ไม่ได้มีการประสมคำของภาษาไทย ดังนั้นจึงยากต่อการอ่าน หลังจากการแปลเป็นอักษรปกติ

Murray และคณะ [6] ได้นำเสนอการพัฒนาทั้งในส่วน ของซอฟต์แวร์ และฮาร์ดแวร์ หรือ portable device สำหรับ แปลอักษรเบรลล์ เป็นอักษรปกติ โดยให้มีการรับข้อมูลเข้า โดยการสแกนภาพอักษรเบรลล์และแปลเป็นอักษรปกติ แบบทันทีทันใด แต่ความถูกต้องก็ยังต้องขึ้นอยู่กับภาพที่ ต้องชัดเจน และภาษาที่รองรับได้เพียงแค่ภาษาไทย ภาษาหนึ่ง

Al-Salman และคณะ [7] ได้นำเสนอการแปลอักษรเบรลล์ สำหรับภาษาอาราบิก เป็นอักษรปกติ โดยใช้ Rule file ซึ่ง สร้างขึ้นมาในลักษณะของ XML file เพื่อมาช่วยในการแปล งานวิจัยนี้สามารถรองรับได้เพียงภาษาอาราบิกเท่านั้น

จากงานวิจัยที่ผ่านมา สามารถพิจารณาได้ว่างานวิจัย ในกลุ่มของการแปลอักษรปกติเป็นอักษรเบรลล์นั้น ได้ถูกนำ มาพัฒนากันอย่างแพร่หลาย จนเป็นซอฟต์แวร์ที่สามารถ ใช้งานได้จริงในปัจจุบัน แต่ในขณะเดียวกันที่งานวิจัยในกลุ่ม ของการแปลอักษรเบรลล์เป็นอักษรปกติยังไม่ได้ถูกให้ความสนใจ เท่าที่ควรอีกทั้งงานวิจัยส่วนใหญ่สามารถรองรับได้ เพียงภาษาใดภาษาหนึ่งเท่านั้น ไม่สามารถแปลจากเอกสาร ที่ประกอบไปด้วยอักษรเบรลล์ ที่มีมากกว่า 1 ภาษา หรือ ประกอบไปด้วยตัวเลขได้ โดยเฉพาะอย่างยิ่งกับภาษาไทย ซึ่งยังไม่มีผู้วิจัยนำมาพัฒนาในส่วนของการแปลจากรหัส อักษรเบรลล์ เป็นตัวอักษรปกติ

ดังนั้นงานวิจัยฉบับนี้จึงขอนำเสนอระบบการแปลเอกสาร

อักษรเบรลล์ เป็นอักษรปกติ ซึ่งสามารถรองรับได้ทั้งภาษาไทย ภาษาอังกฤษ และตัวเลข โดยใช้วิธีการคัดกรองตัว อักษรที่ไม่ถูกต้องออก ก่อนที่จะนำไปเข้าสู่ขั้นตอนของการ จัดเรียงต่อคำ และประโยคเพื่อนำไปคำนวณหาค่าของคะแนน ความเป็นไปได้ของประโยคที่ต้องการ โดยใช้ N-gram ร่วม กับคลังข้อมูล (Orchid corpus) และทำการเลือกประโยคที่มีค่าคะแนนความเป็นไปได้มากที่สุดมาเป็นคำตอบ

## 2. การแปลอักษรเบรลล์เป็นอักษรปกติ

### 2.1 หลักการและลำดับอ่านอักษรเบรลล์

การอ่านอักษรเบรลล์จะอ่านจากตำแหน่งทางด้านซ้าย ไปขวาเหมือนกับอักษรปกติ กล่าวคืออ่านจากตำแหน่งที่ 1 2 3 4 5 6 และจะกำหนดแต่ละรหัสตามตำแหน่งของจุด อักษรเบรลล์

โดยการอ่านอักษรเบรลล์สำหรับตัวอักษรตัวพิมพ์ใหญ่ ในภาษาอังกฤษ (Capital letter) จะต้องใช้รหัสอักษรเบรลล์ ที่เป็นสัญลักษณ์ตัวพิมพ์ใหญ่ (Capital sign) นำหน้าตัวอักษร ภาษาอังกฤษที่ต้องการแสดงเป็นตัวพิมพ์ใหญ่ 1 ตัว เช่น “Capital sign | C | a | t” แต่ถ้าต้องการจะให้กลุ่มคำทั้งคำ เป็นตัวอักษรภาษาอังกฤษตัวใหญ่ทั้งคำ จะต้องใช้รหัสอักษร เบรลล์ที่เป็นสัญลักษณ์ตัวพิมพ์ใหญ่ (Capital sign) นำหน้า ตัวอักษรภาษาอังกฤษ 2 ตำแหน่ง เช่น “Capital sign | Capital sign | C | A | T”

สำหรับการอ่านอักษรเบรลล์ของตัวเลข จะต้องใช้รหัส อักษรเบรลล์ที่เป็นสัญลักษณ์ตัวเลข (Number sign) นำหน้า กลุ่มของตัวเลข เช่น “Number sign | 1 | 2 | 3”

ลำดับการอ่านหรือประสมตัวอักษร สำหรับอักษรเบรลล์ ภาษาไทย จะเรียงในลักษณะ พยัญชนะ สระ วรรณยุกต์ โดยในส่วนของสระจะแบ่งเป็น สระที่ต้องอยู่หน้าพยัญชนะ ได้แก่ “โ | ค | ร | ง | ก | สระอา | ร” อ่านว่า “โครงการ”, สระที่ต้องอยู่ตามหลังพยัญชนะ ได้แก่ “ก | สระอา | ร” อ่านว่า “การ” หรือ “บ | สระอา | ไม่โท | น” อ่านว่า “บ้าน”

### 2.2 ลักษณะของอักษรเบรลล์สำหรับภาษาไทย และ ภาษาอังกฤษ

เนื่องจากอักษรเบรลล์มีลักษณะเป็นจุดหกจุด เรียงต่อกัน อย่างเป็นรหัส จะได้ออกมา 63 กลุ่ม และในแต่ละภาษาก็ จะมีหลักการกำหนดรหัสอักษรเบรลล์ที่แตกต่างกันไป แต่ อย่างไรก็ตามจะพบว่ามีตัวอักษรที่มีรหัสอักษรเบรลล์ซ้ำกัน

ทั้งในภาษาที่แตกต่างกันเช่น ภาษาไทยและภาษาอังกฤษ ดังนั้นการอ่านเอกสารอักษรเบรลล์ที่ประกอบด้วยภาษามากกว่าหนึ่งภาษา จึงต้องอาศัยความชำนาญเฉพาะด้าน โดยพิจารณาจากตัวอักษรที่อยู่รอบข้าง ว่าเป็นภาษาใด หรือเป็นคำอ่านใด สามารถดูตัวอย่างของตัวอักษรที่มีรหัสซ้ำกัน ตามตารางที่ 1

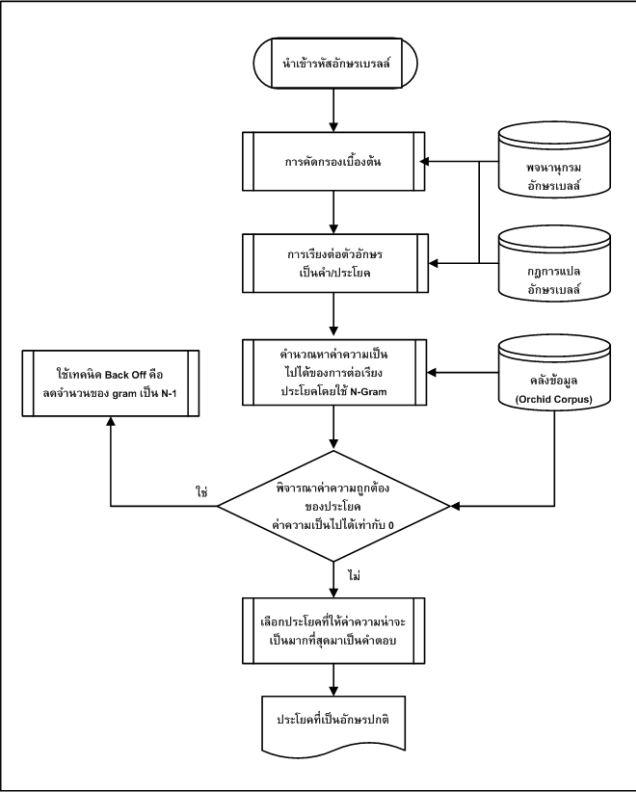
ตารางที่ 1 ตัวอย่างของตัวอักษรที่มีรหัสอักษรเบรลล์ซ้ำกัน

พยัญชนะภาษาไทย	พยัญชนะอังกฤษ	รหัสเบรลล์
ก	G	1245
ข	K	13
ด	D	45

ตัวอย่างของความกำกวมที่เกิดจากการที่มีตัวอักษรที่มีรหัสอักษรเบรลล์ซ้ำกัน เช่น รหัส {24, 145, 13456} สามารถแปลงเป็นอักษรปกติภาษาไทยได้ดังนี้ {โ, ด, ย} หรือภาษาอังกฤษได้ดังนี้ {i, d, y}

3. วิธีการดำเนินการวิจัย

ภาพรวมของระบบการแปลรหัสอักษรเบรลล์เป็นอักษรปกติ สามารถดูภาพรวมได้ ดังภาพที่ 1



ภาพที่ 1 ภาพรวมของระบบการแปลอักษรเบรลล์เป็นอักษรปกติ

3.1 การสร้างพจนานุกรมอักษรเบรลล์

การสร้างพจนานุกรมของรหัสอักษรเบรลล์ ได้ทำการเก็บรวบรวมรหัสอักษรเบรลล์ และตัวอักษรปกติ ทั้งภาษาไทย ภาษาอังกฤษ และตัวเลข มาสร้างเป็นตารางข้อมูล ซึ่งโครงสร้างของตารางข้อมูลจะประกอบไปด้วยคอลัมน์ต่างๆ ดังนี้

- คอลัมน์ที่ 1 กลุ่มของรหัสอักษรเบรลล์
  - คอลัมน์ที่ 2 กลุ่มของตัวอักษรภาษาไทย
  - คอลัมน์ที่ 3 กลุ่มของตัวอักษรภาษาอังกฤษ
  - คอลัมน์ที่ 4 กลุ่มของตัวเลข
  - คอลัมน์ที่ 5 กลุ่มของสระภาษาไทยที่จะต้องอยู่ตำแหน่งข้างหน้าตัวอักษรภาษาไทย
  - คอลัมน์ที่ 6 กลุ่มของสระหรือวรรณยุกต์ภาษาไทยที่จะต้องอยู่ตำแหน่งตามหลังตัวอักษรภาษาไทย
- ทั้งนี้การออกแบบตารางข้อมูล หรือพจนานุกรมจะต้องสอดคล้องกับกฎที่จะใช้ในส่วนของการคัดกรองเบื้องต้นและการเรียงตัวอักษรเป็นคำ /ประโยค ดูตัวอย่างตามตารางที่ 2

ตารางที่ 2 ตัวอย่างโครงสร้างตารางพจนานุกรมอักษรเบรลล์

Code	TH	EN	Number	Vowel (Front)	Vowel (behide)
135	อ	O			
124		f	6	*อ	
6	Capital sign				
3456	Number sign				
1		a	1		*อะ

3.2 การสร้างกฎการอ่านอักษรเบรลล์

การสร้างกฎการอ่านอักษรเบรลล์ จะแบ่งเป็นการกำหนดกฎที่ใช้ในการคัดกรองเบื้องต้น และกฎที่ใช้ในการเรียงตัวอักษรเป็นคำ ดูตามตารางที่ 3

3.3 การคัดกรองเบื้องต้น

การคัดกรองเบื้องต้นคือการพิจารณารหัสอักษรเบรลล์ที่นำเข้าสู่ระบบ โดยทำการนำไปค้นหาจากตารางข้อมูลพจนานุกรมอักษรเบรลล์ และใช้กฎการคัดกรองเบื้องต้นที่สร้างไว้ เพื่อนำมาช่วยในการตัดตัวเลือกที่ไม่ถูกต้องออก เนื่องจากหนึ่งรหัสอักษรเบรลล์ สามารถเป็นตัวอักษรภาษาไทย ภาษาอังกฤษ หรือตัวเลขก็ได้ ตัวอย่างเช่น คำแปลที่ถูกต้องคือคำว่า “การ”



### ตารางที่ 3 การกำหนดกฎที่ใช้ในการคัดกรองเบื้องต้นและกฎที่ใช้ในการเรียงต่อตัวอักษรเป็นคำ

กฎที่ใช้ในการคัดกรองเบื้องต้น	
1	ตำแหน่งแรกของชุดข้อมูลรหัสอักษรเบรลล์จะต้องเป็นตัวอักษรภาษาอังกฤษ (Col 3) หรือภาษาไทย (Col 2) หรือสระภาษาไทยที่จะต้องอยู่ตำแหน่งข้างหน้าตัวอักษร (Col 5)
2	ถ้าตำแหน่งแรกเป็นรหัส 3456 แล้วตำแหน่งต่อจากนั้นจะต้องเป็นตัวเลข (Col4)
3	ถ้าตำแหน่งแรกเป็นรหัส 6 แล้ว ตำแหน่งต่อจากนั้นจะเป็นตัวอักษรภาษาอังกฤษตัวใหญ่ (Col 3) เฉพาะตำแหน่งที่สอง
4	ถ้าตำแหน่งแรกและตำแหน่งที่สองเป็นรหัส 6 แล้ว ตำแหน่งต่อจากนั้นจะเป็นตัวอักษรภาษาอังกฤษตัวใหญ่ (Col 3) จนครบทุกตำแหน่งหรือทั้งคำ
กฎที่ใช้ในการเรียงต่อตัวอักษรเป็นคำ	
1	ถ้าพบตำแหน่งที่เป็นสัญลักษณ์พิเศษต่างๆ เช่น Capital sign, Number sign, Space sign ให้ลบออกจากประโยคนั้นๆ
2	ถ้าพบตำแหน่งใดที่เป็น “*อ” และตำแหน่งนั้นเป็นกลุ่มของสระที่อยู่ด้านหน้าตัวอักษร (Col 5) จะต้องนำค่าตัวอักษรที่ตำแหน่งต่อไปมาแทนที่ “*อ”
3	ถ้าพบตำแหน่งใดที่เป็น “*อ” และตำแหน่งนั้นเป็นกลุ่มของสระที่อยู่ด้านหลังตัวอักษร (Col 6) จะต้องนำค่าตัวอักษรที่ตำแหน่งก่อนหน้ามาแทนที่ “*อ”
4	ข้อมูลที่นำเข้าไปในตารางการจัดเรียงมีลักษณะแบ่งเป็น 3 คอลัมน์ ได้แก่ คอลัมน์ที่เก็บตัวอักษรภาษาไทย (รวมทั้ง สระ วรรณยุกต์ของภาษาไทย) ภาษาอังกฤษ และตัวเลข มีคอลัมน์ใดที่มีตัวอักษรครบจำนวนของคำที่นำเข้า ระบบจะดึงคำนั้นออกมา

- ระบบนำเข้ารหัสอักษรเบรลล์ “1245, 16, 1235”
- เมื่อนำรหัสอักษรเบรลล์ ไปหาในตารางข้อมูลพจนานุกรม จะได้ผลลัพธ์ดังนี้

1245 = {ก}, {ง}, {เลข 7}

16 = {\*อา}

1235 = {ร}, {ร}

- เมื่อผ่านการคัดกรองเบื้องต้นโดยการนำกฎการอ่านรหัสอักษรเบรลล์ จะได้ผลลัพธ์ดังนี้

1245 = {ก}, {ง}, {เลข 7}

16 = {\*อา}

1235 = {ร}, {ร}

ในกรณีนี้จะสามารถตัด เลข 7 ที่ตำแหน่งแรก โดยอ้างอิงกฎที่ใช้ในการคัดกรองเบื้องต้น ข้อที่ 2

### 3.4 การเรียงต่อตัวอักษรเป็นคำ/ประโยค

1) ขั้นตอนนี้เป็นการจัดเรียงแต่ละตัวอักษรที่ได้จากขั้นตอนแรก โดยมาเรียงประกอบกันเป็นคำ (Character Ordering) ซึ่งวิธีการก็คือ จะทำการสร้างตารางการเรียงต่อตัวอักษรเป็นคำ โดยแบ่งเป็น 3 คอลัมน์ ได้แก่ ภาษาไทย ภาษาอังกฤษ และตัวเลข สามารถดูตัวอย่างตามตารางที่ 4

### ตารางที่ 4 ตัวอย่างการเรียงต่อตัวอักษรเป็นคำ

ตัวอักษรภาษาไทย	ตัวอักษรภาษาอังกฤษ	ตัวเลข
{ โ*อ }	{ i }	
{ น }	{ n }	
{ ถ }	{ t }	
{ *อัว }	{ e }	
{ ร }	{ r }	

จากนั้นจะพิจารณาว่า คอลัมน์ใดที่มีจำนวนของตัวอักษรมากที่สุด ตามกฎของการเรียงต่อคำข้อที่ 4 ซึ่งจากตัวอย่างจะได้คำ ออกมา 2 คำเนื่องจาก คอลัมน์ภาษาไทย และภาษาอังกฤษมีจำนวนของตัวอักษรเท่ากัน สามารถแสดงได้ดังนี้

คำที่ 1 ได้แก่ “โน้ตวร” (การจัดเรียงคำนี้ใช้กฎการจัดเรียงตัวอักษรเป็นคำในข้อที่ 1 และ 2)

คำที่ 2 ได้แก่ “inter”

2) ในส่วนของขั้นตอนการเรียงต่อคำ เป็นประโยค จะทำการนำคำที่เป็นไปได้มาเรียงต่อกันในระดับของประโยค เพื่อนำไปใช้ในขั้นตอนต่อไป ซึ่งในขั้นตอนนี้ถ้าจะมีความเป็นไปได้ที่จะได้ผลลัพธ์มากกว่า 1 ประโยค ได้แก่ กรณีที่มีตัวอักษรที่เรียงต่อกันเป็นคำ และไม่เข้าเงื่อนไขของกฎที่ตั้งไว้ ตัวอย่างเช่น

P1{สนับสนุน | โดย | BRAILLE}

P2{สนับสนุน | idy | BRAILLE}

จากตัวอย่างการเรียงต่อคำเป็นประโยค มีความเป็นไปได้สองประโยค เนื่องจาก คำในตำแหน่งที่ 2 สามารถมีความเป็นไปได้ทั้ง {โดย}, {idy} ดังนั้น จึงสร้างรูปประโยคออกมาได้สองประโยค

### 3.5 การวิเคราะห์ความเป็นไปได้ของคำที่เกิดขึ้นเป็นประโยค

ขั้นตอนนี้จะนำ ประโยคที่ได้จากขั้นตอนก่อนหน้า มาพิจารณาหาคำที่เป็นไปได้มากที่สุดมาเป็นคำตอบ โดยงาน

วิจัยนี้ได้เลือกใช้หลักการของ 3-Gram มาช่วยพิจารณาหา ค่าความถี่ของคำที่เกิดขึ้นในคลังข้อมูล โดยที่ 3-Gram คือ การให้ค่าความน่าจะเป็นของ คำใดๆ โดยพิจารณาที่ละ 3 คำ ซึ่งสามารถเขียนค่าความน่าจะเป็น ในรูปของ  $P(w_3|w_2 w_1)$  ซึ่งหมายถึงความน่าจะเป็นที่คำ  $w_3$  จะเกิดตามหลังคำ  $w_1$  และ  $w_2$  ตัวอย่างเช่น  $S1 \{ \text{สนับสนุน โดย ประเทศ THAILAND} \}$   
 $= (\text{ประเทศ | โดย สนับสนุน}) \cdot (\text{THAILAND | ประเทศ โดย})$   
 $= (0.7)(0.8)$   
 $= 0.56$   
 $S2 \{ \text{สนับสนุน idy ประเทศ THAILAND} \}$   
 $= (\text{ประเทศ | idy สนับสนุน}) \cdot (\text{THAILAND | ประเทศ idy})$   
 $= (0)(0)$   
 $= 0$

จากตัวอย่างจะเห็นได้ว่าค่าความเป็นไปได้ของคำที่เรียงต่อกันเป็นประโยคที่ 1 มีค่าความเป็นไปได้มากกว่าดังนั้นจึงเลือกประโยคที่ 1 มาเป็นคำตอบ และจากกรณีตัวอย่างเราจะพบว่าประโยคที่ 2 มีค่าเป็น 0 ซึ่งในกรณีนี้เป็นเพราะว่าไม่พบคำที่ต้องการพิจารณาในคลังข้อมูล แต่ก็อาจจะไม่ได้หมายความว่าประโยคนั้นจะเป็นประโยคที่ผิด ดังนั้นงานวิจัยนี้จึงมีการแก้ไขปัญหาดังกล่าวโดยเมื่อพิจารณาลำดับการเกิดขึ้นของคำ ด้วย 3-Gram ไม่พบ จะใช้วิธีการของ Back-off ซึ่งก็คือการลดจำนวนแกรมเพื่อพิจารณา เช่น หากพิจารณาด้วย 3-Gram ไม่พบ ให้ลองลดเหลือ 2-Gram และพิจารณา หากไม่พบอีกให้ลดลง 1-Gram ดังสมการที่ 1

$$P(w_i | w_{i-2}^{i-1}) = \begin{cases} \beta(w_{i-2}^i) P_m(w_i | w_{i-2}^{i-1}) & \text{If } N(w_{i-2}^i) > 0 \\ \alpha(w_{i-2}^{i-1}) P(w_i | w_{i-1}) & \text{If } N(w_{i-2}^{i-1}) > 0 \\ P(w_i | w_{i-1}) & \text{Otherwise} \end{cases} \quad (1)$$

#### 4. ผลการทดลอง

ในส่วนของการทดลองทางผู้วิจัยได้มุ่งเน้นไปที่ประสิทธิภาพในแง่ของความถูกต้องแม่นยำในการแปลรหัสเบรลล์ เป็นหลัก โดยที่จะมีการประเมินเปรียบเทียบความถูกต้องที่ได้จากระบบ และการอ่านอักษรเบรลล์ของผู้เชี่ยวชาญทางด้านอักษรเบรลล์

##### 4.1 การเตรียมข้อมูลเพื่อใช้ในการทดลอง

การเตรียมข้อมูลเพื่อใช้ในการทดลอง จะต้องทำการ

สร้างชุดรหัสอักษรเบรลล์เป็นประโยค เพื่อใช้ในการทดสอบ โดยนำเข้าในลักษณะของ text file โดยใช้สัญลักษณ์ “,” ในการแบ่งตัวอักษร และใช้สัญลักษณ์ “|” ในการแบ่งคำ

##### 4.2 การประเมินผลการทดลอง

การประเมินผลการทดลองได้มีการทดสอบกับข้อมูลของเอกสารจำนวน 50 หน้าและทำการเปรียบเทียบกับผลที่ได้จากการแปลจากผู้เชี่ยวชาญ 5 คน ซึ่งผลลัพธ์พบว่าระบบให้เปอร์เซ็นต์ความถูกต้องมากกว่า 80% ซึ่งหลักการพิจารณาความถูกต้องของการแปลเอกสาร จะพิจารณาเป็นประโยค กล่าวคือ ภายในหนึ่งประโยค หากแปลถูกหนึ่งคำก็จะให้ 1 คะแนน ตัวอย่างเช่น หากในหนึ่งประโยคประกอบด้วยคำ 10 คำ และระบบสามารถแปลได้ถูกต้อง 8 คำ สามารถคิดเป็นเปอร์เซ็นต์ความถูกต้องเท่ากับ 80% อ้างอิงสูตรการคำนวณที่ (2)

$$Humsent = \frac{\sum_{i=1}^n Ac_i}{n}, t > 0 \quad (2)$$

ตารางที่ 5 การเปรียบเทียบประสิทธิภาพในกรณีของความถูกต้องแม่นยำระหว่างระบบ กับผู้เชี่ยวชาญ

ผู้ทดลอง	% ความถูกต้องในการแปลรหัสอักษรเบรลล์เป็นอักษรปกติ		
	Thai	English	Thai+Eng
ระบบ	85%	90%	80%
ผู้ฝึกทางสายตา	100%	100%	100%

#### 5. บทสรุป

งานวิจัยนี้มีจุดมุ่งหมายที่ต้องการจะพัฒนาสื่อที่จะสามารถเพิ่มช่องทางการติดต่อสื่อสารระหว่างผู้พิการทางสายตา และผู้ที่มีสายตาเป็นปกติได้ และงานวิจัยนี้ก็ได้เลือกทำการวิจัยในกลุ่มของการแปลรหัสเบรลล์เป็นอักษรปกติ เนื่องจากงานวิจัยส่วนใหญ่ในกลุ่มนี้ยังมีข้อจำกัดเกี่ยวกับภาษา ซึ่งสามารถรองรับได้เพียงภาษาไทย ภาษาหนึ่ง

งานวิจัยนี้สามารถนำไปประยุกต์ใช้กับงานทางด้านความรู้จำภาพอักษรเบรลล์ หรือนำไปใช้สำหรับผู้ที่สนใจที่จะศึกษาเกี่ยวกับอักษรเบรลล์โดยที่ไม่ต้องเสียเวลาไปศึกษา

งานวิจัยนี้จะมีประสิทธิภาพสำหรับการรู้จำภาพรหัสอักษรเบรลล์ ในกรณีที่ มีทั้งภาษาไทย และภาษาอังกฤษ



รวมทั้งตัวเลข ประกอบอยู่ในเอกสารเดียวกันได้ รวมทั้งผลที่ได้จากการรู้จำภาพอักษรเบรลล์จะต้องได้รหัสอักษรเบรลล์ที่ถูกต้อง ระบบจึงจะสามารถทำงานได้ดี และให้ผลลัพธ์ได้มากที่สุดถึง 90% แต่อย่างไรก็ดีเราอาจจะพบข้อผิดพลาดที่เกิดจากรู้จำภาพรหัสอักษรเบรลล์ ซึ่งทำให้ระบบการรู้จำรหัสอักษรเบรลล์ผิด ทำให้ระบบการแปลอักษรเบรลล์อาจจะเกิดความผิดพลาดตามไปด้วย ในส่วนของการแก้คำผิดที่เกิดขึ้นนั้น ได้อยู่ในขั้นตอนของการพัฒนาในงานวิจัยต่อไป

## 6. เอกสารอ้างอิง

- [1] Duxbury System, <http://www.duxburysystem.com>
- [2] Index Braille, <http://www.indexbraille.com>
- [3] J. Minghu, Z. Xiaoyan, X.Ying, T.Gang, and Y. Bazong, "Segmentation of Mandarin Braille Word and Braille Translation Based on Multi-Knowledge", *Proceeding of ICSP 2000*, China, 2000.
- [4] A. King, G. Evens, and P. Blenkhorn, "A platform-independent Braille translator" *Department of Computation*, 2002.
- [5] ศรัณย์ เกตุศรีเมฆ, การรู้จำภาพอักษรเบรลล์ จุฬาลงกรณ์มหาวิทยาลัย, 2005.
- [6] I. Murray and A. Pasquale, "A Portable Device for the Translation of Braille to Literary Text" Australia, 2007.
- [7] A. Al-Salman, M. Alkanhal, Y. AlOuali, and H. Al- Rashed "Translating Arabic Braille into Text", *Proceeding of iiWAS 2007*, 2007.