

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ	IV
สารบัญตาราง	VI
สารบัญภาพ	VII
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา	2
1.3 ขอบเขตของการศึกษา	3
1.4 ขั้นตอนของการศึกษา.....	3
1.5 รายละเอียดในแต่ละบท.....	4
บทที่ 2 การแปลภาษาด้วยเครื่องคอมพิวเตอร์.....	5
2.1 ระบบการแปลโดยตรง (Direct Machine Translation Strategy)	5
2.2 ระบบการแปลแบบเปลี่ยน (Transfer Machine Translation Strategy).....	6
2.2.1 การวิเคราะห์ภาษาต้นทาง (Source Language Analysis)	6
2.2.2 การเปลี่ยน (Transfer)	6
2.2.3 การสร้างภาษาปลายทาง (Target Language Generation)	6
2.3 ระบบการแปลภาษาแบบการใช้ภาษากลาง (Interlingual Machine Translation strategy) 6	
2.4 การใช้กฎไวยากรณ์ช่วยในการแปล (Rule-Based MT)	7
2.4.1 Transfer-Based MT	7
2.4.2 Interlingua-Based MT	7
2.5 การใช้ฐานบทความช่วยในการแปล (Corpus-Based MT).....	8
2.5.1 ระบบแปลภาษาด้วยเครื่องแบบอิงสถิติ (Statistical-Based MT)	8
2.5.2 ระบบแปลภาษาด้วยเครื่องแบบอิงตัวอย่าง (Example-Based MT).....	9
2.6 ระบบแปลภาษาอังกฤษ-ไทยภาค.....	11

สารบัญ (ต่อ)

	หน้า
บทที่ 3 ระบบแปลภาษาอังกฤษ-ไทยด้วยเครื่องแบบอิงตัวอย่าง โดยใช้ตัวแบบเอ็นแกรม.....	16
3.1 สถาปัตยกรรมระบบ (System Architecture)	16
3.2 แนวคิดของเอ็นแกรม (The n -gram Approach)	17
3.3 ส่วนโปรแกรมการวิเคราะห์แบบเอ็นแกรม (n -gram analysis Component)	18
3.4 ส่วนโปรแกรมการก่อกำเนิดแบบเอ็นแกรม (n -gram generation Component).....	19
3.5 คลังข้อความ (Corpus)	20
3.5.1 คลังข้อความแบบเดี่ยว (Monolingual Corpus)	21
3.5.2 คลังข้อความแบบคู่ (Bilingual Paralleled Corpus).....	23
บทที่ 4 การประเมินค่าความถูกต้องของผลการแปล	26
4.1 การประเมินค่าความถูกต้องของผลการแปลสำหรับระบบแปลภาษาด้วยเครื่อง	26
4.2 วิธีการทดสอบความถูกต้องของการเรียนรู้.....	27
4.2.1 การตรวจสอบความสมเหตุสมผลแบบไขว้ (Cross-Validation)	27
4.2.2 การประเมินค่าแบบอัตโนมัติของ BLEU-4	27
4.3 การเตรียมข้อมูลสำหรับการประเมินค่าการแปล	29
4.4 ผลการทดลอง.....	29
4.4.1 ประเมินค่าความถูกต้องของการแปลแบบอัตโนมัติ	29
4.4.2 ประเมินค่าความถูกต้องของการแปลโดยมนุษย์.....	31
4.5 สรุปผลการประเมินค่าการแปล	34
บทที่ 5 สรุปการวิจัยและข้อเสนอแนะ.....	39
5.1 สรุปผลการวิจัย.....	39
5.2 ข้อเสนอแนะ.....	39
เอกสารอ้างอิง	41
ภาคผนวก	43
ประวัติผู้เขียน	50

สารบัญตาราง

ตารางที่	หน้า
4.1 ตารางสรุปการเข้าคู่แบบแม่นยำตรง.....	30
4.2 ตารางสรุปผลการประเมินค่าแบบอัตโนมัติของ BLEU-4.....	31
4.3 ตารางสรุปผลการทดลองความถูกต้องของการเรียงลำดับคำ.....	32
4.4 ตารางสรุปผลการทดลองความถูกต้องของการเลือกคำให้เหมาะสมตามบริบท.....	33
4.5 ตารางสรุปผลการทดลองความถูกต้องของการแปลวลี	34

สารบัญญภาพ

ภาพที่	หน้า
2.1 ระบบการแปลภาษาแบบการใช้ภาษากลาง	7
2.2 คู่ประโยคแบบขนานที่มีการวางแนว (Alignment Paralleled Sentence).....	8
2.3 ระบบ ATR.....	9
2.4 แนวความคิดระบบแปลภาษาด้วยเครื่องแบบอิงตัวอย่าง	10
2.5 แสดงขั้นตอนการแปลภาษาของระบบภาษิต.....	12
2.6 สถาปัตยกรรมระบบแปลภาษาอังกฤษ-ไทย “ภาษิต”	12
2.7 ตัวอย่างความสัมพันธ์ระหว่างมโนทัศน์เชิงอรรถศาสตร์	13
2.8 ลำดับชั้นของมโนทัศน์เชิงอรรถศาสตร์ (Semantic Conceptual Hierarchy).....	14
2.9 ตัวอย่างข้อมูลในพจนานุกรมของคำว่า take	15
3.1 ระบบแปลภาษาอังกฤษ-ไทยด้วยเครื่องแบบอิงตัวอย่าง โดยใช้ตัวแบบเอ็นแกรม	17
3.2 ตัวอย่างการทำงานส่วนโปรแกรมการก่อกำเนิดแบบเอ็นแกรม.....	20
3.3 การแก้ปัญหาความกำกวม (1/2).....	21
3.4 การแก้ปัญหาความกำกวม (2/2).....	22
3.5 ตัวอย่างคลังข้อความแบบเดี่ยวสำหรับภาษาต้นทาง	22
3.6 ตัวอย่างคลังข้อความแบบเดี่ยวสำหรับภาษาปลายทาง	23
3.7 ตัวอย่างคลังข้อความแบบคู่.....	24
4.1 แสดงการเตรียมชุดข้อมูลเพื่อสอนและทดสอบ.....	27