

รหัสโครงการ 13P33S001

## Thunder Translator

### สายฟ้า แปลภาษาอัจฉริยะ

ประเภท โครงการ BEST 2011: Thai Language Processing Software Contest

รายงานฉบับสมบูรณ์

เสนอต่อ

ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ

สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ

กระทรวงวิทยาศาสตร์และเทคโนโลยี

ประจำปีงบประมาณ 2553

โดย

นายชนพล จินดาพิทักษ์

นายพีระศักดิ์ รัตนมณี

นายวรรณชัย วัชราทักษิณ

อาจารย์ที่ปรึกษาโครงการ

นายสุรน แซ่ว่อง

สถาบันการศึกษา

ภาควิชาคอมพิวเตอร์ คณะวิศวกรรมศาสตร์

มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตหาดใหญ่

## กิตติกรรมประกาศ

การพัฒนาสายฟ้าแปลภาษาอัจฉริยะได้รับทุนอุดหนุนโครงการวิจัยพัฒนาและวิศวกรรมจากศูนย์เทคโนโลยี อิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (เนคเทค) สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติกระทรวงวิทยาศาสตร์และเทคโนโลยี ตามโครงการแข่งขันพัฒนาโปรแกรมคอมพิวเตอร์แห่งประเทศไทยครั้งที่ 13 ประจำปีงบประมาณ 2553

ขอกราบขอบคุณอาจารย์สุธน แซ่ว่อง ที่เป็นอาจารย์ที่ปรึกษาและให้คำแนะนำในการดำเนินโครงการคณะทำงานขอกราบขอบพระคุณไว้ ณ โอกาสนี้

สุดท้ายนี้คณะทำงานขอขอบพระคุณท่านผู้เกี่ยวข้องที่ไม่ ได้กล่าวนามข้างต้น ซึ่งมีส่วนช่วยเหลือในการจัดทำรายงานฉบับสมบูรณ์นี้

3 มกราคม 2554

## บทคัดย่อ

สืบต่อจาก การแข่งขันสุดยอดซอฟต์แวร์ตัดคำไทย ด้วยเทคนิคไฮบริด ซึ่งรับรางวัลชมเชยในการแข่งขัน BEST2010 ที่มีความสามารถเด่น คือความเร็วในการแบ่งคำ อีกทั้งโปรแกรมมีขนาดเล็กมาก จึงเกิดแนวคิดที่จะ พัฒนาต่อยอดให้เกิดการใช้งานจริง

โปรแกรมแปลภาษา ที่เป็น Firefox Add-on เป็นตัวเลือกแรกๆ ที่ทีมพัฒนา คิดขึ้น เนื่องจากสามารถใช้ความสามารถของการแบ่งคำไทยด้วยเทคนิคไฮบริด ได้อย่างครบครัน และยังไม่มีการใช้วิธีนี้มาก่อน

โปรแกรม Thunder Translator ที่พัฒนาเป็น Firefox Add-ons นั้น สามารถแปลความหมาย จากคำภาษาไทยเป็นภาษาอังกฤษ และจากคำภาษาอังกฤษเป็นภาษาไทย ได้โดย

## คำนำ

การแปลภาษา ทำได้โดยง่าย หากมีคำที่ต้องการ แปล ก็สามารถเปิดพจนานุกรม ซึ่งมีทั้งแบบหนังสือ หรือแบบสื่ออิเล็กทรอนิกส์ ได้ทันที แต่ถ้าหากคำที่ต้องการแปล นั้นอยู่บนเนื้อหาของเว็บไซต์ จะนำคำนั้นมาแปลภาษา ก็ค่อนข้างยุ่งยากและใช้หลายขั้นตอน

หากมีส่วนเสริมของเว็บเบราว์เซอร์ เพื่อช่วยแปล ภาษาแบบอัตโนมัติ จะทำให้ สะดวกยิ่งขึ้น และหากเพียงแค่เขียนคำ ที่ต้องการ จากประโยค แล้วสามารถแปลความหมายเฉพาะคำที่ชี้ได้ ก็ยิ่งทำให้สะดวกมากขึ้น

การแปลภาษาบนเนื้อหาของเว็บไซต์ จะไม่เป็นปัญหาอีกต่อไป โดยการมีส่วนเสริมของ Firefox ที่มีชื่อว่า Thunder Translator

## สารบัญ

บทนำ.....	4
ที่มาของโครงการ.....	4
หลักการและเหตุผล .....	5
วัตถุประสงค์และเป้าหมาย .....	5
รายละเอียดของการพัฒนา .....	6
ทฤษฎี หลักการและเทคนิคที่ใช้ .....	6
เครื่องมือที่ใช้พัฒนา .....	6
รายละเอียดโปรแกรมที่ได้พัฒนาในเชิงเทคนิค (software specification)....	6
ขอบเขตและข้อจำกัดของโปรแกรมที่พัฒนา .....	7
ผลการทดสอบโปรแกรม .....	8
ผลการทดสอบโปรแกรม .....	8
ตัวอย่างการทดสอบโปรแกรม .....	8
สรุปผลและข้อเสนอแนะ .....	9
ปัญหาและอุปสรรค .....	9
แนวทางการพัฒนาต่อ .....	9
วิธีการติดตั้ง .....	10
คำแนะนำการใช้งาน .....	11
เอกสารอ้างอิง .....	12

## บทนำ

### 1.1 ที่มาของโครงการ

Thunder Translator เป็น Firefox Add-ons เพื่อช่วยแปลคำศัพท์จากภาษาไทยเป็นภาษาไทย หรือจากภาษาไทยเป็นภาษาอังกฤษ และจากภาษาอังกฤษเป็นภาษาไทย ได้รวดเร็ว โดยนำ Mouse Pointer ไปชี้ตำแหน่งคำภาษาไทยที่ต้องการแปลแล้วจะมี Pop-Up ที่เป็นความหมายของคำนั้นขึ้นมาทันที

การทำงานมีองค์ประกอบที่สำคัญ 2 ระบบ คือ

#### 1. การแบ่งคำไทย ด้วยเทคนิค Hybrid

พัฒนาต่อยอดจากการแบ่งคำด้วยเทคนิค Hybrid ในการแข่งขัน BEST2010 เนื่องจากความเร็วของการแบ่งคำไทยของเทคนิคนี้ ทำให้ Thunder Translator ทำงานได้รวดเร็ว

#### 2. การแปลความหมายของคำที่แบ่งแล้ว

เทียบคำจากฐานข้อมูลเพื่อดึงความหมายที่ถูกต้องหรือใกล้เคียง และแสดงผล  
ในทันที

เนื่องด้วย Firefox Web Browser เปิดให้นักพัฒนาสามารถพัฒนา Add-ons ด้วยภาษา Java ซึ่งทั้งหมดนี้เป็น Open Source ทำให้เป็นเหตุผลหลักที่เลือกพัฒนา Add-ons โดยใช้ Eclipse เป็นเครื่องมือช่วยในการเขียนและพัฒนา ร่วมกับ Java Development Kit (JDK).

## 1.2 หลักการและเหตุผล

Internet มีบทบาทกับการเรียนรู้มากขึ้น โดยเฉพาะอย่างยิ่งทางด้านภาษา ซึ่งเป็นหน้าที่ของ Web Browser ที่จะจัดการกับเนื้อหาเหล่านั้น เพื่อตอบสนองต่อผู้ใช้อย่างมีประสิทธิภาพ

ในด้านการแปลภาษา มีผู้พัฒนา Add-ons ทางด้านภาษาขึ้นมามากมาย เช่นเพื่อช่วยแปลความหมายจากคำภาษาอังกฤษเป็นคำภาษาไทย แต่ยังไม่มี Add-ons ใดที่สามารถแปลคำจากภาษาไทยเป็นภาษาอังกฤษที่เร็วและแม่นยำพอ เนื่องจากภาษาไทยมีลักษณะการเชื่อมคำที่ยากต่อการระบุคำให้แน่นอน และความเร็วในการแบ่งคำไทยออกจากประโยคยังช้ามาก [1]

ด้วยเหตุผลดังกล่าว ทำให้เกิดแนวคิดที่จะพัฒนา Add-ons ที่สามารถแปลความหมายคำภาษาไทยจากข้อความที่ได้อย่างรวดเร็วและมีประสิทธิภาพ ภายใต้ชื่อ “THUNDER TRANSLATOR” สื่อถึงความรวดเร็ว ความแม่นยำ และประสิทธิภาพในการแปล

## 1.3 วัตถุประสงค์และเป้าหมาย

1. Thunder Translator Add-ons สามารถแปลความหมายคำภาษาไทยเป็นภาษาอังกฤษ และจากภาษาอังกฤษเป็นภาษาไทยได้อย่างรวดเร็ว
2. สามารถแยกคำภาษาไทยจากประโยคได้ โดยการชี้ Mouse Pointer ไปยังคำในเว็บเพจ

## รายละเอียดของการพัฒนา

### 2.1 ทฤษฎี หลักการและเทคนิคที่ใช้

ตัวโปรแกรมใช้ภาษา Java ในการพัฒนาเป็น Add-ons บน Web Browser ยอดนิยมอย่าง Firefox เนื่องจากภาษา Java และ Firefox Web Browser เป็น Freeware ที่มีผู้ใช้งานเป็นจำนวนมาก อีกทั้งยังสามารถทำงานได้บนทุกระบบปฏิบัติการ

ด้วยเหตุผลดังกล่าวทำให้เลือกใช้ เทคโนโลยีในการพัฒนาด้วยภาษา Java เพื่อพัฒนา Add-ons ของ Mozilla Firefox Web Browser [2]

### 2.2 เครื่องมือที่ใช้พัฒนา

1. Eclipse
2. Java Development Kit (JDK)

### 2.3 รายละเอียดโปรแกรมที่ได้พัฒนาในเชิงเทคนิค (software specification)

Thunder Translate Add-ons เป็น Add-ons ที่ทำงานบน Mozilla Firefox Web Browser เพื่อแปลความหมายคำศัพท์ภาษาไทยเป็นภาษาอังกฤษ และจากภาษาอังกฤษเป็นภาษาไทย ที่ Mouse Pointer ซ้ำอยู่ โดยมุ่งเน้นความรวดเร็ว ซึ่งแบ่งการทำงานออกเป็น 2 รูปแบบการใช้งานคือ

#### 1. Automatic Translation Mode

โปรแกรมจะคำนวณหาคำที่ Mouse Pointer ซ้ำอยู่ในประโยคใดบนหน้าเว็บเพจ โดยใช้การแบ่งคำไทยด้วยเทคนิค Hybrid จาก BEST2010 เพื่อแบ่งคำภาษาไทย แล้วเทียบกับฐานข้อมูลคำศัพท์และดึงความหมายออกมาแสดงผล ในรูปแบบของ Pop-Up ส่วนการแปลคำจากภาษาอังกฤษนั้น

#### 2. Manual Translation Mode

ผู้ใช้งานเป็นผู้เลือกคำที่ต้องการแปล โดยเลือก Highlight คำที่ต้องการและกดคีย์ลัดเพื่อสั่งงานให้ Thunder Translator แปลคำศัพท์ให้



- **Input/Output Specification**

**Input :**

- ข้อความใดๆ บนเว็บเพจที่เปิดด้วย Mozilla Firefox ที่ Mouse Pointer ชี้อยู่ หรือข้อความที่ทำการ Highlight ไว้

**Output :**

- หน้าต่าง Pop-Up แสดงความหมายของคำศัพท์ที่ต้องการ

- **Functional Specification**

- โปรแกรมแบ่งและแปลความหมายคำได้รวดเร็ว และถูกต้อง
- สามารถเลือกโหมดการแปลได้ตามความต้องการของผู้ใช้ (Automatically Mode, Manual Mode)
- สามารถเลือกโหมดตามความต้องการของผู้ใช้(Precision, Performance)

## 2.4 ขอบเขตและข้อจำกัดของโปรแกรมที่พัฒนา

- ใช้ได้กับข้อความภาษาไทยหรือภาษาอังกฤษเท่านั้น
- ไม่สามารถแปลคำศัพท์จากข้อมูลที่เป็นรูปภาพได้
- ไม่สามารถแบ่งคำที่เขียนผิด หรือไม่มีความหมายได้
- ไม่สามารถแปลคำหรือตัดคำจากคำแสดงได้

# บทที่ 3 ผลการทดสอบโปรแกรม

## 3.1 ผลการทดสอบ โปรแกรม

โปรแกรมสามารถแปลภาษาได้อย่างถูกต้อง จากภาษาไทยเป็นภาษาอังกฤษ และจากภาษาอังกฤษเป็นภาษาไทย

## 3.2 ตัวอย่างการทดสอบโปรแกรม

โรงเรียนของเราอยู่ คุณครูของเราใจดีทุกคน เด็กก็ไม่ซุกซน เพราะพวกเราทุกคนชอบไปโรงเรียน  
*My school is lively. My Teacher is very kindly. We are not be naughty. Because we like to go school.*

**kindly**  
[ADJ] : อ่อนโยน,  
[ADV] : ด้วยจิตใจที่ดี,

โรงเรียนของเราอยู่ คุณครูของเราใจดีทุกคน เด็กก็ไม่ซุกซน เพราะพวกเราทุกคนชอบไปโรงเรียน  
*My school is lively. My Teacher is very kindly. We are not be naughty. Because we like to go school.*

**ซุกซน**  
[ADJ] : naughty,  
[ADV] : naughtily,  
[V] : be naughty,

ในการทดสอบนั้น pointer ของเรานั้นจะต้องทำการชี้ที่ข้อความ และทำการกด Ctrl ใน โหมด Manual หรือตั้งค้างไว้ในบริเวณข้อความจนครบเวลาในโหมด Auto

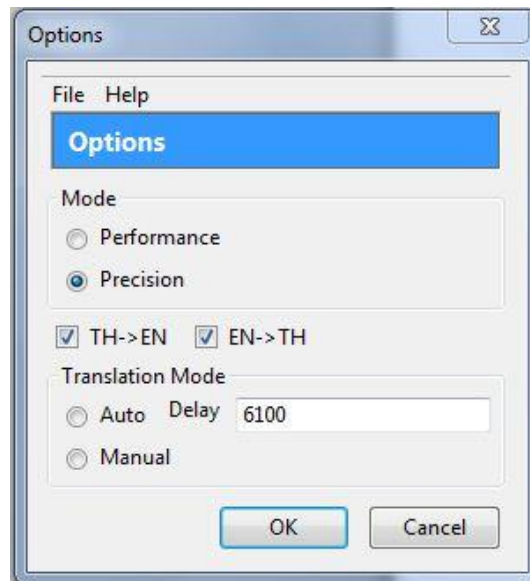
โรงเรียนของเราอยู่ คุณครูของเราใจดีทุกคน เด็กก็ไม่ซุกซน เพราะพวกเราทุกคนชอบไปโรงเรียน  
*My school is lively. My Teacher is very kindly. We are not be naughty. Because we like to go school.*

**naught**  
[N] : ความล้มเหลว, ศูนย์,

โรงเรียนของเราอยู่ คุณครูของเราใจดีทุกคน เด็กก็ไม่ซุกซน เพราะพวกเราทุกคนชอบไปโรงเรียน  
*My school is lively. My Teacher is very kindly. We are not be naughty. Because we like to go school.*

**ทุกคน**  
[IDM] : -, all and sundry, every Tom, Dick, and Harry, every last one, every living soul, every man jack of, one and all,  
[PRON] : all, everybody, everyone,  
[SL] : Uncle Tom Cobby and all,

ในการทดสอบนั้น จะทำการ Highlight ที่ข้อความที่ต้องการทราบความหมาย และทำการกด Ctrl ในโหมด Manual หรือตั้งค้างไว้ในบริเวณข้อความไว้จนครบเวลาในโหมด Auto



จากโปรแกรม ผู้ใช้จะสามารถเลือกและโหมดการทำงานได้ตามต้องการ

## บทที่ 4 สรุปผลและข้อเสนอแนะ

### 4.1 ปัญหาและอุปสรรค

- การทำงานของ Add-ons Firefox นั้นมีข้อจำกัดในการทำงาน

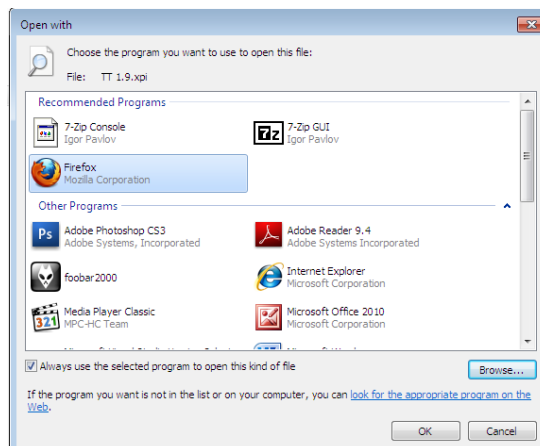
### 4.2 แนวทางการพัฒนาต่อ

- สามารถทำงานในรูปแบบของ ไทย -ไทย
- เพิ่มความสามารถในการตัดคำไทยในโหมด Precision

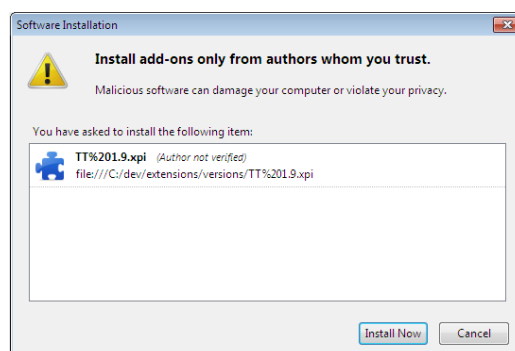
## การติดตั้งโปรแกรม

\* การติดตั้งโปรแกรมจะต้องมี Mozilla Firefox และ Java ติดตั้งภายในเครื่องก่อน

1. รันไฟล์ Thunder Translator firefox add-on installer TT1.9.xpi เลือก open with...
2. เลือก application firefox.exe



### 3. กด Install



## คำแนะนำการใช้งาน

หลังจากติดตั้งโปรแกรมเรียบร้อยแล้ว เมื่อเปิดใช้งาน Firefox Web browser ให้คลิกไปที่ Tools > Add-ons เลือก Thunder Translator แล้วคลิกที่ปุ่ม Option เพื่อตั้งค่าการทำงานของโปรแกรม โดยมีรายละเอียดของตัวเลือกดังนี้

### 1. เลือกโหมดการแปล Automatic หรือ Manual

- Automatic : เมื่อเลื่อน Mouse Pointer ไปยังคำที่ต้องการแปล โปรแกรมจะแบ่งคำในตำแหน่งนั้นโดยอัตโนมัติ และแปลความหมายออกมา หลังจากเวลาที่กำหนดไว้)delay(
- Manual : เลื่อน Mouse Pointer ไปยังคำที่ต้องการแปล และเมื่อต้องการแปล ให้กดปุ่ม Ctrl หนึ่งครั้ง โปรแกรมจึงจะเริ่มแบ่งคำ และแปลความความหมายคำนั้นออกมา

### 2. เลือกระดับการแบ่งคำ Precision หรือ Performance

- Precision : โปรแกรมจะใช้เทคนิคการตัดคำที่เพิ่มความถูกต้อง และแม่นยำในการแบ่งคำมากขึ้น แต่จะใช้เวลามากขึ้นด้วย
- Performance : โปรแกรมจะใช้การแบ่งคำแบบมาตรฐาน เพื่อให้โปรแกรมทำงานได้เร็วที่สุด

## เอกสารอ้างอิง

[https://developer.mozilla.org/en/XUL\\_Reference](https://developer.mozilla.org/en/XUL_Reference)

[https://developer.mozilla.org/en/XUL\\_Tutorial](https://developer.mozilla.org/en/XUL_Tutorial)

<http://www.w3schools.com/js/default.asp>

<http://www.w3schools.com/sql/default.asp>

[http://www.w3schools.com/css/css\\_reference.asp](http://www.w3schools.com/css/css_reference.asp)