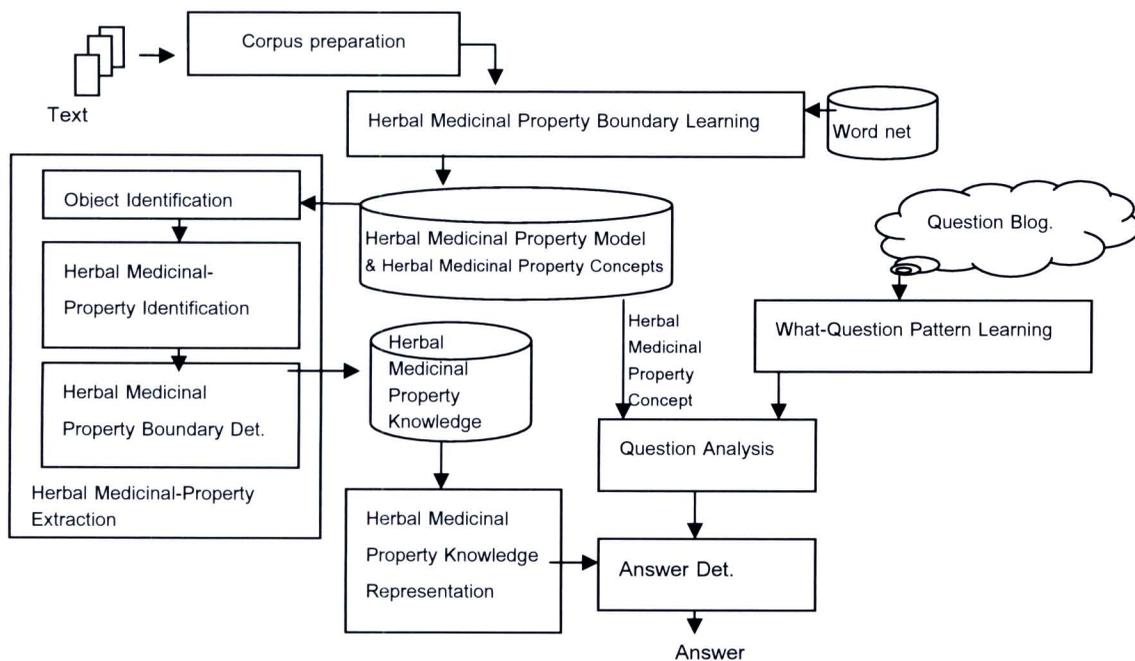




## กรรมวิธีดำเนินงาน Method

ระบบงานโดยสรุปสำหรับการสกัดความรู้เกี่ยวกับสรรพคุณทางยาของพืชสมุนไพรไทยจากเอกสารภาษาไทยเพื่อสนับสนุนระบบการตอบคำถามอัตโนมัติประกอบด้วยขั้นตอนต่างๆดังต่อไปนี้ (ดังแสดงในรูปที่3) แบ่งออกเป็นสองส่วนคือ ส่วนสกัดความรู้เกี่ยวกับสรรพคุณทางยาของพืชสมุนไพรไทยจากเอกสารภาษาไทยและส่วนการตอบคำถามชนิด “ถามอะไร”ประเภทลิสต์ และประเภทเอนติตี้เกี่ยวกับสรรพคุณทางยาของพืชสมุนไพร



รูปที่3 ระบบงานโดยสรุป

### ส่วนสกัดความรู้เกี่ยวกับสรรพคุณทางยาของพืชสมุนไพรไทยจากเอกสารภาษาไทย

การสกัดความรู้เกี่ยวกับสรรพคุณทางยาของพืชสมุนไพรไทยจากเอกสารภาษาไทย ประกอบด้วยขั้นตอนหลักๆ ดังนี้

#### 1. การเตรียมคลังข้อมูล (Corpus Preparation)

ก่อนที่จะมีการเตรียมคลังข้อมูลจะต้องมีการศึกษาพฤติกรรมทางภาษาของชุดข้อมูลที่ได้จากเอกสารสมุนไพรไทยดาวโหลดจากเว็บไซต์กรมส่งเสริมการเกษตร [www.doae.go.th](http://www.doae.go.th) และโครงการอนุรักษ์พันธุพืช <http://www.rspg.or.th/index.html> ว่ามีความเหมาะสมที่จะใช้เป็นคลังข้อมูลสำหรับงานวิจัยนี้ ประโยคต่างๆที่ปรากฏอยู่ในชุดข้อมูลดังกล่าวเราจะสนใจเฉพาะประโยคธรรมดา (Simple Sentence หรือ EDU) ซึ่งสามารถแสดงได้ด้วย Regular Express ดังนี้  $NP1 V_{mp} NP2$  เมื่อ  $NP1$  คือ เซตของนามวลีที่เป็นเอนติตี้สมุนไพรหรือส่วนประกอบของเอนติตี้สมุนไพร  $V_{mp}$  คือเซตความคิดกริยาที่แสดงให้เห็นสรรพคุณทางยา(Medicinal-Property Verb Concept)

และ NP2 คือเซตความคิดนามวลีเกี่ยวกับอาการ, โรค, และ เชื้อโรค (ตามที่ได้กล่าวไว้ในบทนำ) อย่างไรก็ตามงานวิจัยนี้สนใจสรรพคุณทางยาที่แสดงให้เห็นหรือเข้าใจได้ด้วยเหตุการณ์ (Event) ซึ่งสามารถแทนได้ด้วยกริยา  $v_{mp}$  เมื่อ  $v_{mp} \in V_{mp}$

$V_{mp} = \{\text{“รักษา/cure” “บรรเทา/relieve” “แก้/stop,prevent” “ขับ/release” “ลด/reduce” “เพิ่ม/increase” ...}\}$

ดังแสดงสรุปได้ในตารางที่1 (จากข้อมูลตัวอย่างสุ่มจำนวน500 EDUsจากคลังข้อมูลที่ดาวโหลดมา)

**ตารางที่1** แสดงพฤติกรรมทางภาษาของคำกริยาที่มีความคิดเป็นสรรพคุณทางยา(Medicinal-Property Verb Concept) และไม่มีความคิดเป็นสรรพคุณทางยา(Non Medicinal-Property Verb Concept) จาก Surface Form เดียวกัน จาก คลังข้อมูล500 EDUs

| Medicinal-Property Verb Concept<br>$v_{mp}$ | Number of Occurrences<br>(Surface Form) | Number of Medicinal-Property-Concept Occurrences | Number of Non-Medicinal-Property-Concept Occurrences |
|---|---|--|--|
| แก้(Stop, Relief)                           | 68                                      | 60   | 8  |
| ขับ (Discharge, Release)                    | 25                                      | 23   | 2  |
| บรรเทา (Relief)                             | 5                                       | 5  | -  |
| รักษา (Treat, Cure)                         | 11                                      | 8  | 3  |
| ลด (Decrease)                               | 15                                      | 9  | 6  |
| ถ่าย (Excrete)                              | 5                                       | 3  | 2  |
| บำรุง (Norish)                              | 10                                      | 6  | 4  |

จากตารางที่1จะเห็นว่า  $v_{mp}$  ที่ปรากฏส่วนใหญ่มีความหมายหรือแนวความคิดเป็นความคิดกริยาที่แสดงให้เห็นสรรพคุณทางยา (Medicinal-Property Concept) ดังนั้นคลังข้อมูลที่ดาวโหลดมานี้สามารถใช้ในการงานวิจัยนี้ โดยเริ่มจากการเตรียมคลังข้อมูลประมาณ 3000 EDUs (2000 EDUs สำหรับการเรียนรู้ EDUs สรรพคุณทางยาของพืชสมุนไพรไทย 1000สำหรับการทดสอบและประเมินผล) ข้อมูลที่เป็นเอกสารเหล่านี้ก่อนที่จะนำมาประมวลผลทางภาษารวมชาติจะต้องผ่านขั้นตอนการตัดคำโดยใช้ซอฟต์แวร์ตัดคำภาษาไทยที่สามารถแก้ปัญหาขอบเขตคำ และขณะเดียวกันสามารถกำกับหน้าที่ของคำ (Part of Speech) ได้ (Sudprasert and Kawtrakul, 2003) ซึ่งรวมถึงการทำ Name Entity (Chanlekha and Kawtrakul, 2004), และการรับรู้คำ (Word-Formation Recognition) (Pengphon et al., 2002) เพื่อที่จะแก้ปัญหาขอบเขตของ Thai Name Entity และนามวลี หลังจากนั้นต้องทำการตัดประโยคในระดับ EDU ด้วยวิธีการของ (Chareonsuk et al., 2005) และสุดท้ายทำการกำกับ EDUs ที่เป็นความรู้สรรพคุณทางยาของพืชสมุนไพร (Medicinal Property Tag) ที่ได้ออกแบบดังแสดงในรูปที่4

“พริกไทยดำ.....

พริกไทยดำมีรสเผ็ดอุ่น EDU1เมื่อรับประทานเข้าไป EDU2จะรู้สึกอุ่นวามที่ท้อง EDU3 ช่วยขับลม EDU4ขับ  
ปัสสาวะ EDU5แก้อืดท้องเฟ้อ EDU6แก้ไข้มาลาเรีย EDU7แก้อหิวาตกโรค, EDU8ใช้ก้านพริกไทย 10  
ก้าน”

<Topic\_name Entity-concept=black pepper/herb>พริกไทยดำ</Topic\_name>.....

<EDU1> เมื่อ <NP1 concept=person>φ</NP1>

<VP><Vmp concept=consume>รับประทาน </Vmp>เข้าไป</EDU>

<id =1 class=Medicinal Property>

<EDU2><NP1 concept= black pepper/herb>φ</NP1>

<VP>จะ<Vmp concept=be warm>รู้สึกอุ่นวาม </Vmp>ที่ท้อง</EDU>

<EDU3><NP1 concept= black pepper/herb>φ</NP1>

<VP><Vmp concept=release>ช่วยขับ </Vmp>

<NP2 concept=gas>ลม</NP2></VP></EDU>

<EDU4><NP1 concept= black pepper/herb>φ</NP>

<VP>< Vmp concept=discharge/release >ขับ </ Vmp >

<NP2 concept=urine>ปัสสาวะ</NP2></VP></EDU>

<EDU5><NP1 concept= black pepper/herb>φ</NP1>

<VP>< Vmp concept=stop>แก้ </ Vmp >

<NP2 concept=flatulence/symptom >ท้องอืดท้องเฟ้อ</NP2></VP></EDU>

<EDU6><NP1 concept= black pepper/herb>φ</NP1>

<VP>< Vmp concept=cure>แก้ </ Vmp >

<NP2 concept= malaria >ไข้มาลาเรีย</NP2></VP></EDU>

<EDU7><NP1 concept= black pepper/herb>φ</NP1>

<VP>< Vmp concept=cure>แก้ </ Vmp >

<NP2 concept= cholera >อหิวาตกโรค</NP2></VP></EDU>

</id>

Vmp is the medicinal property verb tag

**รูปที่4** ตัวอย่างการกำกับ EDUที่เป็นความรู้สรรพคุณทางยาของพืชสมุนไพร

## 2. การเรียนรู้ขอบเขตของสรรพคุณของพืชสมุนไพร (Herbal Medicinal Property Boundary Learning)

จากคลังข้อมูลที่ได้กำกับกับความหมายสรรพคุณทางยาของพืชสมุนไพรในระดับ EDU ของขั้นตอนก่อนหน้านี้ ทำการสกัดลักษณะเฉพาะหรือฟีเจอร์ต่างๆ (Features) ที่เป็นทั้งแนวความคิด (Concept) และเซอร์เฟซฟอร์ม (Surface Form) ของกริยาที่มีความคิดเป็นสรรพคุณทางยาของพืชสมุนไพร รวมไปถึงนามวลีที่อยู่ก่อนหน้าและหลังกริยา เก็บเป็นฐานข้อมูลความคิดทางยาของพืชสมุนไพรของ  $v_{mp}$  (เมื่อ  $v_{mp} \in V_{mp}$ ) , NP1, และ NP2 ดังแสดงในตารางที่2

**ตารางที่2** แสดงฟีเจอร์ที่เป็นกริยาแสดงสรรพคุณทางยาของพืชสมุนไพร  $V_{mp}$  รวมทั้งสารสนเทศ(นามวลี)ที่อยู่รอบๆกริยา

| no | ID   | NP1/<br>concept                                   | $V_{mp}$ / concept | NP2/ concept                           | class |
|----|------|---|--------------------|--|-------|
| 1  | 001  | พริกไทยดำ/black<br>pepper                         | รู้สึกอุ่น/be warm | -                                      | yes   |
| 2  | 001  | พริกไทยดำ/black<br>pepper                         | ช่วยขับ/ release   | ลม/gas                                 | yes   |
| 3  | 001  | พริกไทยดำ/black<br>pepper                         | ขับ/ release       | ปัสสาวะ /urine                         | yes   |
| 4  | 002  | พริกไทยดำ/black<br>pepper                         | แก้ /stop          | ท้องอืดท้องเฟ้อ<br>/flatulence,symptom | yes   |
| 5  | 002  | พริกไทยดำ/black<br>pepper                         | แก้ /cure          | ไข้มาลาเรีย /malaria<br>disease        | yes   |
| 6  | 003  | พริกไทยดำ/black<br>pepper                         | แก้ /cure          | อหิวาตกโรค /cholera<br>disease         | No    |
| 7  | 004  | ฟ้าทะลายโจร/<br>Andrographis<br>paniculata        | บรรเทา/ relief     | อาการเจ็บคอ/sore<br>throat Symptom     | yes   |
| 8  | 004  | ฟ้าทะลายโจร/<br>Andrographis<br>paniculata        | บรรเทา/ relief     | อาการหวัด<br>/flu symptom              | yes   |
| 9  | 004  | ใบฟ้าทะลายโจร/<br>Andrographis<br>paniculata Leaf | แก้ /stop          | ท้องเสีย<br>/diarrhea                  | No    |
| .. | .... | ....  | ...                | ...                                    | ...   |

ซึ่งถูกนำไปใช้สำหรับการเรียนรู้ขอบเขตของ EDUs สรรพคุณทางยาของพืชสมุนไพรไทย ด้วยการหาความถี่ของ  $v_{mp}$  pair ของคู่ EDU ที่อยู่ติดกันในหนึ่งกรอบหน้าต่างพร้อมทั้งเลื่อนกรอบหน้าต่างไปด้วยระยะทางหนึ่ง EDU ว่ามีความหมายเป็นสรรพคุณทางยาของสมุนไพรไทยหรือไม่จากคลังข้อมูลที่ได้กำกับกับความหมายสรรพคุณทางยาของพืชสมุนไพรในระดับ EDU ของขั้นตอนก่อนหน้านี้นี้ดังที่ได้แสดงในตารางที่3

ตารางที่ 3 แสดงค่าความน่าจะเป็นของ  $v_{mp}$  จาก  $v_{mp}$  pair คือ  $v_{mp\_at\_ij}$  และ  $v_{mp\_at\_ij+1}$  ที่มีความคิดเป็นสรรพคุณทางยาของพืชสมุนไพร และไม่มีความคิดเป็นสรรพคุณทางยาของพืชสมุนไพร

| $v_{mp\_at\_ij}$   | Medicinal Property Verb<br>Concept | Non Medicinal Property Verb<br>Concept |
|--------------------|------------------------------------|--|
| stop               | 0.4110                             | 0.1731                                 |
| release            | 0.1507                             | 0.1346                                 |
| relief             | 0.0069                             | 0.0385                                 |
| treat              | 0.1367                             | 0.0096                                 |
| discharge          | 0.0137                             | 0.0385                                 |
| be-drug            | 0.0205                             | 0.0096                                 |
| ...                | ...                                | ...                                    |
| $v_{mp\_at\_ij+1}$ | Medicinal Property Verb            | Non Medicinal Property Verb.           |
| stop               | 0.375                              | 0.1091                                 |
| release            | 0.1447                             | 0.0273                                 |
| reduce             | 0.0197                             | 0.0091                                 |
| treat              | 0.0132                             | 0.0182                                 |
| discharge          | 0.0263                             | 0.0091                                 |
| be-drug            | 0.0132                             | 0.0182                                 |
|                    | ....                               | .....                                  |

### 3. การสกัดความรู้สรรพคุณทางยาของพืชสมุนไพร (Herbal Medicinal Property Extraction)

ขั้นตอนนี้แบ่งออกเป็นสองส่วนคือ ส่วนการระบุวัตถุหรือเอนติตี้สมุนไพรและการระบุสรรพคุณทางยาของพืชสมุนไพร และส่วนขั้นตอนการหาขอบเขตของสรรพคุณทางยาของพืชสมุนไพร

#### 3.1 การระบุวัตถุและการระบุสรรพคุณทางยาของพืชสมุนไพร (Object Identification and Herbal Medicinal-Property Identification)

ขั้นตอนนี้เป็นการระบุวัตถุหรือเอนติตี้สมุนไพร โดยใช้ชื่อเอกสารเป็น TopicName (จากบทนำ)  $np1_i = part + TopicName$

$$Np1_i \in NP1 \quad (i=1,2,\dots,m)$$

ต่อมาระบุ EDU ที่แสดงสรรพคุณทางยา โดยสแกนไปตามลำดับของ EDU ที่ปรากฏบนเอกสารพืชสมุนไพรไทยใช้เพื่อค้นหา Lexico Syntactic Pattern (NP1 V<sub>mp</sub> NP2) หากพบให้ดำเนินการในขั้นตอนถัดไป หากไม่พบให้ระบุวัตถุและระบุ EDU ที่แสดงสรรพคุณทางยากับเอกสารใหม่เช่นนี้ไปเรื่อยๆ

#### 3.2 การหาขอบเขตของสรรพคุณทางยาของพืชสมุนไพร

ขั้นตอนนี้สามารถทำได้สองวิธีที่แตกต่างกัน

ก. การหาขอบเขตของสรรพคุณทางยาของพืชสมุนไพรสามารถหาได้โดยการหาค่า  $argmax$  จากสมการ Naïve Bayes ต่อไปนี้ กับค่าความน่าจะเป็น ของ  $v_{mp\_at\_ij}$  และ  $v_{mp\_at\_ij+1}$  ในตารางที่ 2 ดังแสดงในอัลกอริทึมของรูปที่ 5

$$\begin{aligned}
 \text{HerbalMedicinalPropertyBoundaryClass} &= \arg \max_{\text{class} \in \text{Class}} P(\text{class} | v_{mp\_at\_ij}, v_{mp\_at\_ij+1}) \\
 &= \arg \max_{\text{class} \in \text{Class}} P(v_{mp\_at\_ij} | \text{class}) P(v_{mp\_at\_ij+1} | \text{class}) P(\text{class}) \quad (5)
 \end{aligned}$$

where  $v_{mp\_at\_ij} \in V_{mp}$  and  $v_{mp\_at\_ij+1} \in V_{mp}$  ( $V_{mp}$  is a medicinal\_property\_verb\_concept set)  
 $i = \{1, 2, \dots, n\}$      $j = \{1, 2, \dots, k\}$

Assume that each EDU is represented by (NP1  $V_{mp}$  NP2). L is a list of EDU.  $V_{mp}$  is the medicinal-property verb concept set. NP1 is the herbal noun phrase concept set. NP2 is the symptom/disease noun phrase concept set.

$v_{ij}$  or  $v_{mp\_at\_ij}$ ,  $v_{ij+1}$  or  $v_{mp\_at\_ij+1}$  are learned verbs as elements of the  $V_{mp}$  set  
 HERBAL\_MEDICINAL\_PROPERTY\_EXTRACTION1 ( L,  $V_{mp}$ , NP1, NP2 )

```

1  i ← 1, j ← 1 R ← ∅    MEDPROPi ← ∅
2  while i ≤ length[L] do
3  begin_while1
4  If np1i ∈ NP1 ^ vi ∈ Vmp ^ np2i ∈ NP2 /*find the medicinal property EDU
5  bd=yes ; MEDPROPi ← MEDPROPi ∪ {j}
7  while ( vij ∈ Vmp ) ^ ( vij+1 ∈ Vmp ) ^ bd=yes do
8  begin_while2    /* Boundary determination
9  bd = arg maxc ∈ {yes, no} P(vij | c) P(vij+1 | c) P(c)
10    if bd = yes then
11        MEDPROPi ← MEDPROPi ∪ {j+1}; j = j+1
13    end_while2
14    R = R ∪ {MEDPROPi }
15    j = 1; i = i+1
16    end_while1
17 : Return
  
```

รูปที่ 5. อัลกอริทึม Medicinal Property Boundary Extraction โดย Naive Bayes

ข. การหาขอบเขตของสรรพคุณทางยาของพืชสมุนไพรโดยใช้ทฤษฎีเซนต์อริริง กล่าวคือให้สแกน np1<sub>i</sub> ของคู่ EDU

Assume that each EDU is represented by (NP1  $V_{mp}$  NP2). L is a list of EDU.  $V_{mp}$  is the medicinal-property verb concept set. NP1 is the herbal noun phrase concept set. NP2 is the symptom/disease noun phrase concept set.

$v_{ij}$  or  $v_{mp\_at\_ij}$ ,  $v_{ij+1}$  or  $v_{mp\_at\_ij+1}$  are learned verbs as elements of the  $V_{mp}$  set  
 HERBAL\_MEDICINAL\_PROPERTY\_EXTRACTION2 ( L,  $V_{mp}$ , NP1, NP2 )

```

1  i ← 1, j ← 1 R ← ∅    MEDPROPi ← ∅
2  while i ≤ length[L] do
3  begin_while1
4  If np1i ∈ NP1 ^ vi ∈ Vmp ^ np2i ∈ NP2 /*find the medicinal property EDU
5  bd=yes ; MEDPROPi ← MEDPROPi ∪ {j}
7  while ( vij ∈ Vmp ) ^ ( vij+1 ∈ Vmp ) ^ ( np1ij = np1ij+1 ) do
8  begin_while2    /* Boundary determination
9
10        MEDPROPi ← MEDPROPi ∪ {j+1}; j = j+1
11    end_while2
12    R = R ∪ {MEDPROPi }
13    j = 1; i = i+1
14    end_while1
15 : Return
  
```

รูปที่ 6. อัลกอริทึม Medicinal Property Boundary Extraction โดย Centering Theory

#### 4. การแทนความรู้สรรพคุณทางยาของพืชสมุนไพร (Herbal Medicinal Property Knowledge Representation)

นำความรู้สรรพคุณทางยาของพืชสมุนไพรที่สกัดได้จากขั้นตอนก่อนหน้านี้มาเก็บในฐานข้อมูลในรูปแบบของฐานข้อมูลเชิงวัตถุที่มี แททริบิวต์ดังนี้

ID= รหัสสมุนไพร

X=ชื่อสมุนไพร หรือ TopicName

Y=ส่วนของพืช หรือ part

$np1_i = \text{part} + \text{TopicName}$  เมื่อ  $i=1,2,\dots,m$

$mp_j = \text{สรรพคุณ หรือ } V_{mp} \text{ NP2}$  เมื่อ  $j=1,2,\dots,n$

P=วิธีการเตรียม (แถมทริบิวท์“วิธีการเตรียม” สำหรับงานวิจัยครั้งต่อไป)

ฉะนั้นสามารถแทนความรู้สรรพคุณทางยาของพืชสมุนไพรที่เก็บอยู่ในฐานข้อมูลดังกล่าวออกมาในรูปแบบของ Predicate Representation ดังนี้

$\text{Herb\_ID}(\text{ID}) \wedge \text{Herb}(X) \wedge \text{Part\_of\_Herb}(Y) \wedge \text{Medcinal\_Property\_List}(mp_1, mp_2, \dots, mp_n)$

รหัสสมุนไพร(ID)  $\wedge$  ชื่อสมุนไพร(X)  $\wedge$  ส่วนประกอบพืชสมุนไพร (Y)  $\wedge$  รายชื่อสรรพคุณ( $mp_1, mp_2, \dots, mp_n$ )

## ส่วนการตอบคำถามชนิด “ถามอะไร” ประเภทลิสต์ และประเภทเอนติตี้เกี่ยวกับสรรพคุณทางยาของพืชสมุนไพร

การตอบคำถามชนิด “ถามอะไร” ประเภทลิสต์ และประเภทเอนติตี้เกี่ยวกับสรรพคุณทางยาของพืชสมุนไพร ประกอบด้วยขั้นตอนต่าง ๆ ดังต่อไปนี้

### 1. การเรียนรู้แพทเทิร์นของคำถามอะไร (what-Question Pattern Learning)

ขั้นตอนนี้จะต้องมีการรวบรวมคำถามจากบล็อก (Blog) ต่าง ๆ บนเว็บไซต์ต่าง ๆ ที่ถามเกี่ยวกับสรรพคุณของพืชสมุนไพรจำนวน 150 คำถามเพื่อใช้เรียนรู้หาแพทเทิร์นของคำถามประเภท “อะไร” และ “อะไรบ้าง” และรวบรวมคำถามอีก 50 คำถามเพื่อใช้ในการทดสอบระบบการตอบคำถาม โดยคำถามเหล่านี้ทั้งหมดจะต้องมีขนาดความยาวไม่เกิน 10 คำ และผ่านกระบวนการตัดคำโดยใช้ซอฟต์แวร์ตัดคำภาษาไทยที่สามารถแก้ปัญหาขอบเขตคำและขณะเดียวกันสามารถกำกับหน้าที่ของคำ (Part of Speech) ได้ (Sudprasert and Kawtrakul, 2003) ซึ่งสามารถเข้าใช้ระบบได้ถูกต้องพอสมควรโดยผ่านระบบอินเทอร์เน็ต หลังจากนั้นนำมาศึกษาหาแพทเทิร์นของคำถามชนิด “ถามอะไร” ประเภทลิสต์ และประเภทเอนติตี้เกี่ยวกับสรรพคุณทางยาของพืชสมุนไพรได้แพทเทิร์นทั้งหมดได้ แพทเทิร์นประเภทลิสต์ 6 แพทเทิร์น ประเภทเอนติตี้ 3 แพทเทิร์นดังนี้ (จากบทนำ กำหนดให้  $np1_i = \text{part} + \text{TopicName}$ ,

$\text{TopicName} \in \text{Herb}$

$\text{part} \in \{\text{null}, \text{“ใบ”}, \text{“ดอก”}, \text{“ราก”}, \text{“ต้น”}, \text{“เมล็ด”}, \text{“ผล”}, \dots\}$

$np1_i \in \text{NP1}$  ( $i=1,2,\dots,m$ ),  $np2_l \in \text{NP2}$  ( $l=1,2,\dots,h$ ), และ  $v_{mp} \in V_{mp}$ )

### ประเภทลิสต์

- $np1_i + \text{“มี”} + \text{Focus}(\text{“สรรพคุณ”} + \text{“อะไรบ้าง”})$
- $\text{Focus}(\text{“ลิสต์”} + \text{“สรรพคุณ”}) + \text{[“ของ”]} + np1_i$
- $\text{Focus}(\text{“ลิสต์”} + \text{TopicName}) + \text{[“มีสรรพคุณ”]} + v_{mp} + np2_l$
- $\text{Focus}(\text{[“แสดง | บอก”]} + \text{“สรรพคุณ”}) + \text{[“ของ”]} + np1_i + \text{“มี”} + \text{Focus}(\text{“อะไรบ้าง”})$
- $\text{Focus}(\text{[“แสดง | บอกชื่อ”]} + \text{“สมุนไพร”} + \text{“อะไรบ้าง”}) + \text{[“มีสรรพคุณ”]} + v_{mp} + np2_l$

### ประเภทเอนติตี้ “x อะไร”

- $np1_i + \text{“มี”} + \text{Focus}(\text{“สรรพคุณ”} + \text{“อย่างไร”})$
- $\text{Focus}(\text{[“แสดง | บอกชื่อ”]} + \text{“สมุนไพร”} + \text{“อะไร”}) + \text{[“มีสรรพคุณ”]} + v_{mp} + np2_l$

## 2. การวิเคราะห์คำถาม(Question Analysis)

นำแพทเทิร์นคำถามชนิด “ถามอะไร” ประเภทลิสต์ และประเภทเอนดีตีเกี่ยวกับสรรพคุณทางยาของพืชสมุนไพรพร้อม Focusของคำถาม จากขั้นตอนก่อนหน้ามาทำการวิเคราะห์คำถามที่ต้องการทดสอบว่าตรงกับแพทเทิร์นไหน

## 3. การหาคำตอบ(Answer Determination)

นำแพทเทิร์นที่ตรงกับคำถามที่ต้องการทดสอบจากขั้นตอนก่อนหน้ามาทำการหาคำตอบโดยแมชชีน (Matching) ระหว่าง TopicName จากแพทเทิร์น กับ X ใน Predicate Representation

np1 จากแพทเทิร์น กับ Y ใน Predicate Representation

$v_{mp} + np2_i$  จากแพทเทิร์น กับ  $mp_j$  ใน Predicate Representation

และส่วนของ Predicate Representation (ที่ตรงกับส่วน Focus ของแพทเทิร์น) คือส่วนที่แสดงคำตอบ ตัวอย่างเช่น

คำถาม: ไปโหระพามีสรรพคุณอะไรบ้าง

ไป/ncn โหระพา/ncp มี/vt สรรพคุณ/ncn อะไรบ้าง/qw  
 └──────────┬──────────┘ └──────────┬──────────┘  
 part + TopicName + “มี” + Focus(“สรรพคุณ” + “อะไรบ้าง”)  
 └──────────┬──────────┘

ตรงแพทเทิร์นประเภทลิสต์ที่3:                      np1<sub>i</sub> +                      “มี” + Focus(“สรรพคุณ” + “อะไรบ้าง”)

**Predicate Representation:**                      └──────────┬──────────┘                      └──────────┬──────────┘

Herb\_ID(ID) ^ Herb(X) ^ Part\_of\_Herb(Y) ^ Medcinal\_Property\_List( $mp_1, mp_2, \dots, mp_n$ )

รหัสสมุนไพร(ID) ^ ชื่อสมุนไพร(X) ^ ส่วนประกอบพืชสมุนไพร (Y) ^ รายชื่อสรรพคุณ( $mp_1, mp_2, \dots, mp_n$ )

**Answer:** รายชื่อสรรพคุณ( $mp_1, mp_2, \dots, mp_n$ )

รายชื่อสรรพคุณ(ขับลม, แก้ปวดฟัน, แก้ไอ, หลอดลมอักเสบ)