

ห้องสมุดคณะเทคโนโลยีสารสนเทศ พระจอมเกล้าลาดกระบัง

ระบบถามตอบที่ใช้ร่วมกับโปรแกรมประยุกต์บนเว็บ

QUESTION & ANSWERING SYSTEM
ON WEB APPLICATION



H004880

โดย

ภักดี เทียนเครือ

PAK TIANKRUA

อาจารย์ที่ปรึกษา

ผศ.ดร.พรฤดี เนติโสภาค

ภก.
ภ 3135
2550

เลขหมู่.....
เลขทะเบียน..... 04880
วัน,เดือน,ปี. = 9 ต.ค. 2551

b.11978649.....
i.....

รายงานนี้เป็นส่วนหนึ่งของวิชาโครงการพัฒนาระบบงาน
หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ภาคเรียนที่ 2 ปีการศึกษา 2550

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ในชั้นเรียนเท่านั้น กรุณาอย่านำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**QUESTION & ANSWERING SYSTEM
ON WEB APPLICATION**



**A SYSTEM DEVELOPMENT PROJECT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY
FACULTY OF INFORMATION TECHNOLOGY
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2/ 2007



COPYRIGHT 2008

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์หรือการงานเพื่อการศึกษาเท่านั้น เมื่อผู้ดูแลเนื้อหาเว็บไซต์หรือเอกสารฉบับนี้มีการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อ	ระบบถามตอบที่ใช้ร่วมกับโปรแกรมประยุกต์บนเว็บ
นักศึกษา	นายภักดี เทียนเครือ
รหัสนักศึกษา	49066508
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
แขนงวิชา	วิทยาการสารสนเทศ
ปีการศึกษา	2550
อาจารย์ที่ปรึกษา	ผศ.ดร.พรฤดี เนติโสภาคกุล

บทคัดย่อ

โครงการพัฒนาระบบงานฉบับนี้นำเสนอการพัฒนาเว็บถามตอบซึ่งใช้หลักของการประมวลผลภาษาธรรมชาติ (Natural Language Processing) มาใช้งานร่วมกับ โปรแกรมประยุกต์บนเว็บ (Web Application) เพื่อช่วยให้ผู้ใช้งานเข้าถึงข้อมูลที่ต้องการจากระบบ ซึ่งเก็บอยู่บนฐานข้อมูลเชิงสัมพันธ์ (Relational Database) ได้โดยที่ไม่จำเป็นต้องเข้าใจลักษณะ โครงสร้างการจัดเก็บข้อมูลและวิธีการสืบค้นจากฐานข้อมูลเชิงสัมพันธ์ของระบบนั้นๆ โดยที่ระบบถามตอบที่พัฒนาขึ้นมาเป็นส่วนติดต่อกับผู้ใช้งานสำหรับ โปรแกรมประยุกต์บนเว็บที่มีลักษณะเป็นกล่องข้อความที่อนุญาตให้ผู้ใช้งานกรอกคำถามที่เกี่ยวข้องกับระบบเป็นภาษาอังกฤษเข้าไปในระบบ แล้วระบบจะนำคำถามที่ได้มาประมวลผลด้วยกระบวนการประมวลผลภาษาธรรมชาติด้วยวิธีการแจกแจงค่าเพื่อให้ทราบโครงสร้างของประโยคในส่วนที่เกี่ยวข้องกับการแปลงภาษาธรรมชาติไปเป็นภาษา SQL แล้วนำเอาประโยคสืบค้นภาษา SQL ที่ได้ไปสืบค้นคำตอบซึ่งเป็นข้อมูลที่เก็บอยู่ในฐานข้อมูลเชิงสัมพันธ์ของโปรแกรมประยุกต์บนเว็บ แล้วนำมาแสดงผลยังผู้ใช้งาน ทำให้โปรแกรมประยุกต์บนเว็บมีความสามารถในการโต้ตอบกับผู้ใช้งานได้ดีขึ้น

Title	Question & Answering System on Web Application
Student	Mr. Pak Tiankrua
Student ID.	49066508
Degree	Master of Science
Program	Information Technology Management
Academic Year	2007
Advisor	Assistant Prof. Dr. Pornrudee Netisopakul

ABSTRACT

This system development project propose to develop question & Answering system by principle of Natural Language Processing work together with Web Application in order to support user to approach the required data that collect in Relational Database without understandability for structure of data collection and relational database searching method. The developed Question & Answering System is application which interacts with user in "Message box" pattern. Firstly, the message box allow users to fill in English language for their question about the system. Secondly, the system will bring that question to proceed with Natural Language Processing by parsing text method to know the structure of sentence that concern with the changing Natural Language to SQL. Next, the system will bring the searching sentence in SQL to find the answer that collect in Relational Database of web application. Finally, it will show the result to user on screen. The developed system can support web application to interact with user conveniently

กิตติกรรมประกาศ

ระบบงานฉบับนี้สำเร็จได้อย่างดี ด้วยคำแนะนำ และคำปรึกษาจาก ผศ.ดร. พรฤดี เนติโสภาคกุล
ขอกราบขอบพระคุณอาจารย์เป็นอย่างสูง

ขอกราบขอบพระคุณคณาจารย์คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้า
คุณทหารลาดกระบัง ที่ได้ประสิทธิ์ประสาทวิชาให้กับข้าพเจ้า

สุดท้ายนี้ข้าพเจ้าขอกราบขอบพระคุณ บิดา มารดา และครอบครัวของข้าพเจ้าที่เป็นกำลังใจ
และให้การสนับสนุนในทุกเรื่อง ทำให้ข้าพเจ้าสามารถพัฒนาระบบงานนี้สำเร็จลุล่วงด้วยดี
คุณค่าและประโยชน์อันพึงมาจากระบบงานฉบับนี้ ข้าพเจ้าขอมอบแด่ผู้มีพระคุณทุกท่าน

ภักดี เทียนเครือ



สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VI
สารบัญรูป.....	VII
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการพัฒนาระบบงาน.....	2
1.3 ขอบเขตของการพัฒนาระบบงาน.....	2
1.4 ขั้นตอนการพัฒนาระบบงาน.....	3
1.5 รายละเอียดเนื้อหาของแต่ละบท.....	4
บทที่ 2 ทฤษฎีพื้นฐานที่ใช้ในการวิจัย และระบบ โอเอพีดีเอ็ม.....	4
2.1 ภาษา SQL (SQL: Structure Query Language).....	4
2.1.1 ประเภทของคำสั่ง SQL.....	4
2.1.2 ชนิดของข้อมูลที่ใช้ในภาษา SQL.....	5
2.1.3 การสืบค้นข้อมูลจากระบบฐานข้อมูลเชิงสัมพันธ์ด้วยภาษา SQL.....	5
2.2 การสืบค้นข้อมูลจากระบบฐานข้อมูลเชิงสัมพันธ์ด้วยภาษา SQL.....	6
2.3 การประมวลผลภาษาธรรมชาติโดยวิธีการแจกแจงค่า.....	14
2.4 นิพจน์ปรกติ.....	19
2.5 การเขียนไวยากรณ์ในการแจกแจงด้วยนิพจน์ปรกติ.....	22
บทที่ 3 การวิเคราะห์และออกแบบระบบ.....	23
3.1 การออกแบบระบบถามตอบ.....	23
3.2 แผนภาพแสดงลำดับการทำงานในระบบถามตอบ.....	26
3.2.1 แผนภาพแสดงลำดับการทำงานโดยรวม.....	27

สารบัญ (ต่อ)

	หน้า
3.2.2 แผนภาพแสดงลำดับการทำงานของกระบวนการต่างๆ.....	29
3.3 การออกแบบฐานข้อมูลระบบถามตอบ.....	36
3.3.1 ฐานข้อมูลของระบบถามตอบ.....	36
3.3.2 ฐานข้อมูล NorthWind.....	39
บทที่ 4 การพัฒนาและการทดสอบระบบงาน.....	40
4.1 การพัฒนาระบบถามตอบที่ใช้งานกับ โปรแกรมประยุกต์บนเว็บ.....	40
4.1.1 ฮาร์ดแวร์.....	40
4.1.2 ซอฟต์แวร์.....	40
4.2 ส่วนติดต่อผู้ใช้งาน.....	41
4.3 การทดสอบการทำงานของระบบ.....	45
4.3.1 การทดสอบการทำงานหน้าหลัก (Home Page).....	45
4.3.2 การทดสอบการทำงานหน้าฐานข้อมูล (Database Page).....	59
บทที่ 5 บทสรุป.....	62
5.1 สรุปผลการพัฒนาระบบงาน.....	62
5.2 ประโยชน์ที่ได้รับจากการพัฒนาระบบงาน.....	62
5.3 ข้อเสนอแนะ.....	63
บรรณานุกรม.....	64
ประวัติผู้เขียน.....	65

สารบัญตาราง

ตารางที่	หน้า
2.1 ตัวอย่างตาราง Store Information.....	7
2.2 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.2).....	7
2.3 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.3).....	8
2.4 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.4).....	8
2.5 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.6).....	9
2.6 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.12).....	10
2.7 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.14).....	10
2.8 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.17).....	11
2.9 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.20).....	12
2.10 ตัวอย่างตาราง Product.....	12
2.11 ตัวอย่างตาราง Supplier.....	12
2.12 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.24).....	13
2.13 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.25).....	13

สารบัญรูป

รูปที่	หน้า
3.1 แผนภาพแสดงลำดับการทำงานของระบบถามตอบ.....	23
3.2 แผนภาพแสดงลำดับการทำงานโดยรวมของระบบถามตอบ.....	26
3.3 แผนภาพแสดงลำดับการทำงานกระบวนการแปลงคำถามให้อยู่ในรูปแบบอย่างง่าย.....	29
3.4 แผนภาพแสดงลำดับการทำงานของกระบวนการแยกคีย์เวิร์ด.....	30
3.5 แผนภาพแสดงลำดับการทำงานของกระบวนการตรวจจับ Entity ของฐานข้อมูล.....	31
3.6 แผนภาพแสดงลำดับการทำงานของกระบวนการสร้างประโยคสืบค้น SQL.....	33
3.7 แผนภาพแสดงลำดับการทำงานของกระบวนการรวบรวมคำตอบ.....	35
3.8 แผนภาพแสดงความสัมพันธ์ระหว่าง Entity ของฐานข้อมูลระบบถามตอบ	36
3.9 แผนภาพแสดงความสัมพันธ์ระหว่าง Entity ของฐานข้อมูล NorthWind.....	39
4.1 ส่วนติดต่อผู้ใช้งานหน้าหลัก (Home page).....	41
4.2 ส่วนติดต่อผู้ใช้งานหน้าฐานข้อมูล (Database page).....	42
4.3 ส่วนติดต่อผู้ใช้งานหน้าคีย์เวิร์ด (Keyword page).....	43
4.4 ส่วนติดต่อผู้ใช้งาน หน้าไวยากรณ์ (Grammar page).....	44
4.5 ผลจากการทดสอบ 1.1).....	45
4.6 ผลจากการทดสอบ 1.2).....	46
4.7 ผลจากการทดสอบ 1.3).....	47
4.8 ผลจากการทดสอบ 1.4).....	48
4.9 ผลจากการทดสอบ 1.5).....	49
4.10 ผลจากการทดสอบ 1.6).....	50
4.11 ผลจากการทดสอบ 1.7).....	51
4.12 ผลจากการทดสอบ 1.8).....	52
4.13 ผลจากการทดสอบ 1.9).....	53
4.14 ผลจากการทดสอบ 2.1).....	54
4.15 ผลจากการทดสอบ 2.2).....	55
4.16 ผลจากการทดสอบ 2.3).....	56
4.17 ผลจากการทดสอบ 2.4).....	57
4.18 ผลจากการทดสอบ 2.5).....	58

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.19 ผลจากการทดสอบการแสดงผลรายการตาราง.....	59
4.20 ผลจากการทดสอบการแสดงผลรายการคอลัมน์.....	60
4.21 ผลจากการทดสอบการแสดงผลรายการความสัมพันธ์.....	61



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบัน โปรแกรมประยุกต์บนเว็บเข้ามามีบทบาทสำคัญต่อการดำเนินธุรกิจเป็นอย่างมาก เนื่องจาก โปรแกรมประยุกต์บนเว็บสามารถช่วยให้การบริหารจัดการต่างๆ ในการดำเนินธุรกิจทำได้รวดเร็วขึ้น ทำให้มีการพัฒนาโปรแกรมประยุกต์บนเว็บขึ้นมามากมาย ซึ่งแต่ละระบบนิยมจัดเก็บข้อมูลให้อยู่ใน ระบบฐานข้อมูลเชิงสัมพันธ์ (Relational Database) เนื่องจากสามารถจัดเก็บข้อมูลจำนวนมากได้อย่างรวดเร็วและมีประสิทธิภาพ ในการสืบค้นข้อมูลบนฐานข้อมูลเชิงสัมพันธ์นั้น ผู้ใช้งานจำเป็นจะต้องมีความรู้เกี่ยวกับภาษา SQL (Structure Query Language) ซึ่งเป็นมาตรฐานที่ใช้ในการติดต่อสื่อสารกับฐานข้อมูลเชิงสัมพันธ์ตามมาตรฐานของ ANSI (American National Standard Institute) ดังนั้น ผู้ที่ต้องการทราบข้อมูลต่างๆ ของโปรแกรมประยุกต์บนเว็บจะต้องเรียนรู้คำสั่งที่ใช้ในการสืบค้นในรูปแบบภาษา SQL นอกจากนี้ผู้ใช้งานจำเป็นที่จะต้องทราบถึงโครงสร้างของข้อมูลที่เก็บไว้ในฐานข้อมูลเชิงสัมพันธ์เพื่อที่จะสามารถสร้างคำสั่งการสืบค้นภาษา SQL ให้สอดคล้องกับโครงสร้างข้อมูลที่ทำกรเก็บไว้ได้อย่างถูกต้อง ปัญหาที่เกิดขึ้นกับโปรแกรมประยุกต์ต่างๆคือผู้ใช้งานจำเป็นจะต้องเข้าใจโครงสร้างของฐานข้อมูลและภาษา SQL เพื่อนำไปใช้สืบค้นข้อมูลที่ต้องการได้

แต่เนื่องจากในปัจจุบัน โปรแกรมประยุกต์บนเว็บถูกพัฒนาขึ้นมาเพื่อช่วยในการบริหารจัดการต่างๆ ในการดำเนินธุรกิจผู้ใช้งานส่วนใหญ่ของโปรแกรมประยุกต์บนเว็บนั้นมักจะเป็นผู้บริหารองค์กรณ์ในระดับต่างๆ ซึ่งผู้ใช้งานเหล่านี้ไม่มีความรู้พื้นฐานเกี่ยวกับฐานข้อมูลเชิงสัมพันธ์ และภาษา SQL แต่ผู้ใช้งานเหล่านี้มีความต้องการข้อมูลเพื่อใช้ในการประกอบการตัดสินใจในการดำเนินธุรกิจ เพื่อแก้ปัญหาดังกล่าว ระบบจะต้องทำให้ผู้ใช้งานเหล่านี้สามารถเข้าถึงข้อมูลที่ต้องการได้โดยไม่ต้องมีความรู้พื้นฐานเกี่ยวกับฐานข้อมูลเชิงสัมพันธ์และภาษา SQL ผู้พัฒนาจึงมีแนวคิดในการพัฒนาระบบถามตอบเพื่อเป็นส่วนติดต่อผู้ใช้งานให้ผู้ใช้งานได้ป้อนคำถามภาษาธรรมชาติ (Natural Language) ซึ่งเป็นภาษาที่ทุกคนเข้าใจกันดี เข้าไปในระบบเพื่อใช้ในการสืบค้นข้อมูลที่ต้องการแทนภาษา "SQL" โดยระบบถามตอบจะทำหน้าที่ในการประมวลผลภาษาธรรมชาติ (Natural Language Processing) จากคำถามที่ผู้ใช้งานป้อนเข้าไปแปลงเป็นรูปแบบประโยคสืบค้นภาษา SQL แล้วแสดงผลการสืบค้นกลับมายังผู้ใช้งาน

1.2 วัตถุประสงค์ของการพัฒนาระบบงาน

โครงการพัฒนาระบบงานฉบับนี้มุ่งหวังที่จะศึกษารูปแบบการพัฒนาระบบตามตอบเพื่อนำมาใช้งานร่วมกับโปรแกรมประยุกต์บนเว็บเพื่อเพิ่มความสามารถให้กับ โปรแกรมประยุกต์บนเว็บในด้านการโต้ตอบกับผู้ใช้งาน โดยระบบตามตอบจะช่วยทำให้โปรแกรมประยุกต์บนเว็บสามารถตอบสนองต่อคำถามต่างๆที่ผู้ใช้งานต้องการคำตอบจากระบบ โดยมีวัตถุประสงค์หลักคือทำให้ผู้ใช้งานทั่วไปที่ไม่มีความรู้เกี่ยวกับ โครงสร้างการจัดเก็บข้อมูลของระบบสามารถเข้าถึงข้อมูลที่ต้องการได้ ผ่านทางการป้อนคำถามที่เป็นภาษาอังกฤษผ่านทางส่วนติดต่อผู้ใช้งานของระบบตามตอบที่สร้างขึ้น โดยที่ผู้ใช้งานไม่จำเป็นต้องทราบถึงโครงสร้างการจัดเก็บ และวิธีการค้นหาข้อมูลในระบบฐานข้อมูลของโปรแกรมประยุกต์บนเว็บนั้นๆ โดยมีความคาดหวังว่าระบบตามตอบนี้ที่พัฒนาขึ้นนี้จะสามารถเป็นต้นแบบในการพัฒนาระบบสนับสนุนลูกค้าต่อไปในอนาคตได้

1.3 ขอบเขตของการพัฒนาระบบงาน

โครงการพัฒนาระบบงานฉบับนี้จะนำเสนอการพัฒนาระบบตามตอบเพื่อใช้ในการค้นหาข้อมูลในระบบฐานข้อมูลของโปรแกรมประยุกต์บนเว็บ โดยระบบตามตอบนี้จะพัฒนาขึ้นเพื่อใช้งานกับ Northwind ซึ่งเป็นฐานข้อมูลจำลองที่มาพร้อมกับระบบจัดการฐานข้อมูลของไมโครซอฟต์ SQL เซอร์ฟเวอร์ 2000 (Microsoft SQL Server 2000) นอกจากนี้รูปแบบในการประมวลผลภาษาธรรมชาติที่โครงการนี้เลือกใช้เป็นลักษณะการแจกแจงคำเพื่อคู่โครงสร้างของประโยคภาษาธรรมชาติ แล้วนำไปเปรียบเทียบกับไวยากรณ์ที่มีการเก็บไว้ เมื่อตรงกันจะสามารถแปลงเป็นประโยคสืบค้น SQL ได้ และระบบตามตอบสามารถทำการสืบค้นข้อมูลบนฐานข้อมูลเชิงสัมพันธ์ โดยแสดงรูปแบบคำตอบในลักษณะระเบียบฐานข้อมูลเท่านั้น

1.4 ขั้นตอนการพัฒนาระบบงาน

- ศึกษารูปแบบการสืบค้นข้อมูลบนฐานข้อมูลเชิงสัมพันธ์ด้วยภาษา SQL
- ศึกษารูปแบบการประมวลผลภาษาธรรมชาติด้วยวิธีการแจกแจงค่า
 1. ศึกษารูปแบบและเงื่อนไขต่างๆ ในการแปลงภาษาธรรมชาติไปเป็นภาษา SQL
 2. ศึกษาวิธีการแจกแจงค่า
 3. ศึกษาการใช้งานนิพจน์ปรกติ (Regular Expression) ในการเขียนไวยากรณ์สำหรับแจกแจงค่า
- วิเคราะห์และออกแบบระบบตามตอบที่ใช้งานร่วมกับ โปรแกรมประยุกต์บนเว็บ
- พัฒนาระบบตามตอบที่ใช้งานร่วมกับ โปรแกรมประยุกต์บนเว็บ
- ทดสอบการใช้งาน และปรับปรุงแก้ไขจุดบกพร่องของระบบ
- สรุปผลการพัฒนาระบบจากการทดสอบการใช้งาน
- จัดทำเอกสารประกอบโครงการ

1.5 รายละเอียดเนื้อหาของแต่ละบท

- บทที่ 2 นำเสนอทฤษฎีพื้นฐานที่เกี่ยวข้องกับการพัฒนาระบบตามตอบที่ใช้งานร่วมกับ โปรแกรมประยุกต์บนเว็บ
- บทที่ 3 นำเสนอแนวทางการออกแบบและพัฒนาระบบตามตอบที่ใช้งานร่วมกับ โปรแกรมประยุกต์บนเว็บ โดยวิเคราะห์ความต้องการในด้านต่างๆของระบบ เพื่อให้นำไปสู่การพัฒนา ระบบ
- บทที่ 4 นำเสนอการพัฒนาและทดสอบการใช้งานของระบบตามตอบที่ใช้งานร่วมกับ โปรแกรมประยุกต์บนเว็บ
- บทที่ 5 นำเสนอบทสรุปพร้อมข้อเสนอแนะในการพัฒนาระบบ

บทที่ 2

ทฤษฎีพื้นฐานที่ใช้ในการพัฒนาระบบถามตอบ

ในบทนี้จะกล่าวถึงทฤษฎีพื้นฐานต่างๆที่นำมาใช้ในการออกแบบระบบถามตอบ ซึ่งเนื้อหาภายในบทจะประกอบไปด้วย ภาษา SQL (SQL: Structure Query Language), การสืบค้นข้อมูลจากระบบฐานข้อมูลเชิงสัมพันธ์, การประมวลผลภาษาธรรมชาติโดยวิธีการแจกแจงค่า และนิพจน์ปรกติ (Regular Expression)

2.1 ภาษา SQL (SQL: Structure Query Language)

ภาษา SQL (สามารถอ่านออกเสียงได้ 2 แบบ คือ “เอสคิวแอล” (SQL) หรือ “ซีเควล” (Squel) ย่อมาจาก Structured Query Language หรือภาษาในการสอบถามข้อมูลเป็นภาษาทางด้านฐานข้อมูลที่สามารถสร้างและปฏิบัติการกับระบบฐานข้อมูลเชิงสัมพันธ์ (Relation Database Management System) โดยเฉพาะ และเป็นภาษาที่มีลักษณะคล้ายกับภาษาอังกฤษ ภาษา SQL ถูกพัฒนาขึ้นจากแนวคิดของ Relation Calculus และ Relational Algebra เป็นหลัก ภาษา SQL เริ่มพัฒนาครั้งแรกโดย Almaden Reseach Center ของบริษัท IBM โดยที่ชื่อเริ่มแรกว่า “ซีเควล” (Sequel) ต่อมาได้เปลี่ยนชื่อเป็น “เอสคิวแอล” (SQL) หลังจากนั้นภาษา SQL ได้ถูกนำมาพัฒนาโดยผู้ผลิตซอฟต์แวร์ด้านระบบจัดการฐานข้อมูลเชิงสัมพันธ์จนเป็นที่นิยมกันอย่างแพร่หลายในปัจจุบัน โดยผู้ผลิตแต่ละรายก็พยายามที่จะพัฒนาระบบที่แตกต่างกันไปบ้าง เช่น ORACLE ACCESS SQL Base ของ Sybase INGRES หรือ SQL Server ของ Microsoft เป็นต้น ดังนั้นในปี ค.ศ. 1986 ทางด้าน American National Standards Institute (ANSI) จึงได้กำหนดมาตรฐานของ SQL ขึ้น อย่างไรก็ตามโปรแกรมฐานข้อมูลที่ขายในท้องตลาด ได้ขยายเอสคิวแอลออกไปจนเกินข้อกำหนดของ ANSI โดยเพิ่มคุณสมบัติอื่น ๆ ที่คิดว่าเป็นประโยชน์เข้าไปอีกแต่โดยหลักทั่วไปแล้วก็ยังปฏิบัติตามมาตรฐานของ ANSI ในการอธิบายคำสั่งต่าง ๆ ของภาษา SQL

2.1.1 ประเภทของคำสั่ง SQL

ภาษา SQL เป็นภาษาที่ใช้งาน ได้ตั้งแต่ระดับเครื่องคอมพิวเตอร์ส่วนบุคคลพีซีไปจนถึงระดับเมนเฟรม ประเภทของคำสั่งในภาษา SQL (The subdivision of SQL) แบ่งออกเป็น 3 ประเภท ดังนี้

- 1) ภาษาสำหรับการนิยามข้อมูล (Data Definition Language: DDL) ประกอบด้วยคำสั่งที่ใช้ในการกำหนดโครงสร้างข้อมูลว่ามีคอลัมน์อะไร แต่ละคอลัมน์เก็บข้อมูลประเภทใด รวมถึงการเพิ่มคอลัมน์ การกำหนดดัชนี การกำหนดควิวหรือตารางเสมือนของผู้ใช้ เป็นต้น
- 2) ภาษาสำหรับการจัดการข้อมูล (Data Manipulation Language: DML) ประกอบด้วยคำสั่งที่ใช้ในการเรียกใช้ข้อมูล การเปลี่ยนแปลงข้อมูล การเพิ่มหรือลบข้อมูล เป็นต้น
- 3) ภาษาควบคุม (Data Control Language: DCL) ประกอบด้วยคำสั่งที่ใช้ในการควบคุมการเกิดภาวะพร้อมกัน หรือการป้องกันการเกิดเหตุการณ์ที่ผู้ใช้หลายคนเรียกใช้ข้อมูลพร้อมกันและคำสั่งที่เกี่ยวข้องกับการควบคุมความปลอดภัยของข้อมูลด้วยการกำหนดสิทธิของผู้ใช้ที่แตกต่างกัน เป็นต้น

2.1.2 ชนิดของข้อมูลที่ใช้ในภาษา SQL

ในภาษา SQL การบรรจุข้อมูลลงในคอลัมน์ต่าง ๆ ของตารางจะต้องกำหนดชนิดของข้อมูล (Data Type) ให้แต่ละคอลัมน์ ชนิดของข้อมูลนี้จะแสดงชนิดของค่าที่อยู่ในคอลัมน์ ค่าทุกค่าในคอลัมน์ที่กำหนดจะต้องเป็นชนิดเดียวกัน เช่น ใตตารางลูกค้าคอลัมน์ที่เป็นรายชื่อลูกค้าจะต้องเป็นตัวหนังสือ ในขณะที่คอลัมน์จำนวนเงินที่ลูกค้าเป็นตัวเลข

ชนิดของข้อมูลของแต่ละคอลัมน์จะขึ้นกับลักษณะของข้อมูลแต่ละคอลัมน์ ซึ่งแบ่งได้ตามชนิดข้อมูลพื้นฐานในภาษา SQL ดังนี้

- 1) ตัวหนังสือ (Character) ในภาษา SQL จะใช้
 - ตัวหนังสือแบบความยาวคงที่ (Fixed-Length Character) จะใช้ `char (n)` หรือ `character (n)` แทนประเภทของข้อมูลที่เป็นตัวหนังสือใดๆที่มีความยาวของข้อมูลคงที่โดยมีความยาว n ตัว หนังสือประเภทนี้จะมีการจองเนื้อที่ตามความยาวที่คงที่ตามที่กำหนดไว้ ชนิดของข้อมูลประเภทนี้จะเก็บความยาวของข้อมูลได้มากที่สุดได้ 255 ตัวอักษร
 - ตัวหนังสือแบบความยาวไม่คงที่ (Variable-Length Character) จะใช้ `varchar (n)` แทนประเภทของข้อมูลที่เป็นตัวหนังสือใดๆที่มีความยาวของ

ข้อมูลไม่คงที่ โดยมีความยาว n ตัวหนังสือประเภทนี้จะมีการจองเนื้อที่ตามความยาวของข้อมูล ชนิดของข้อมูลประเภทนี้จะเก็บความยาวของข้อมูลได้มากที่สุดได้ 4000 ตัวอักษร

2) จำนวนเลข (Numeric)

- จำนวนเลขที่มีจุดทศนิยม (Decimal) ในภาษา SQL จะใช้ `dec (m,n)` หรือ `decimal (m,n)` เป็นประเภทข้อมูลที่เป็นจำนวนเลขที่มีจุดทศนิยม โดย m คือจำนวนตัวเลขทั้งหมด (รวมจุดทศนิยม) และ n คือจำนวนตัวเลขหลังจุดทศนิยม
- จำนวนเลขที่ไม่มีจุดทศนิยมในภาษา SQL จะใช้ `int` หรือ `integer` เป็นเลขจำนวนเมววกหรือลบขนาดใหญ่ เป็นตัวเลข 10 หลัก ที่มีค่าตั้งแต่ -2,147,483,648 ถึง +2,147,483,647 และในภาษา SQL จะใช้ `smallint` เป็นประเภทข้อมูลที่เป็นเลขจำนวนเต็มบวกหรือลบขนาดเล็ก เป็นตัวเลข 5 หลัก ที่มีค่าตั้งแต่ -32,768 ถึง +32,767 ตัวเลขจำนวนเต็มประเภทนี้จะมีการจองเนื้อที่น้อยกว่าแบบ `integer`
- เลขจำนวนจริง ในภาษา SQL อาจใช้ `number (n)` แทนจำนวนเลขที่ไม่มีจุดทศนิยมและจำนวนเลขที่มีจุดทศนิยม

3) ข้อมูลในลักษณะอื่น ๆ

- วันที่และเวลา (Date/Time) เป็นชนิดวันที่หรือเวลาในภาษา SQL จะใช้ `date` เป็นข้อมูลวันที่ ซึ่งจะมีหลายรูปแบบให้เลือกใช้ เช่น `yyyy-mm-dd` (1999-10-31) `dd.mm.yyyy` (31.10.1999) หรือ `dd/mm/yyyy` (31/10/1999)

2.2 การสืบค้นข้อมูลจากระบบฐานข้อมูลเชิงสัมพันธ์ด้วยภาษา SQL

การสืบค้นข้อมูลจากระบบฐานข้อมูลเชิงสัมพันธ์ด้วยภาษา SQL นั้นจะใช้ภาษาสำหรับการจัดการข้อมูล โดยในโครงการพัฒนาระบบงานฉบับนี้ จะนำเสนอเฉพาะส่วนที่เกี่ยวกับการสืบค้นข้อมูลเท่านั้น ซึ่งทั่ว ๆ ไปแล้วในการสืบค้นข้อมูลจากระบบฐานข้อมูลเชิงสัมพันธ์นั้นจะนึกถึงคำหลักๆสองคำคือ `SELECT` และ `FROM` ซึ่งจะมีการใช้คำดังกล่าวซึ่งจะเรียกรวมว่าเป็นคำสั่ง SQL ในรูปแบบ ดังนี้

`SELECT "column_name" FROM "table_name"` (2.1)

ตารางที่ 2.1 ตัวอย่างตาราง Store Information

Store_name	Sales
Los Angeles	1500
San Diego	250
Los Angeles	300
Boston	700

จากตารางที่ 2.1 จะทำการสืบค้นข้อมูล store_name จากตาราง Store Information จะใช้คำสั่ง SQL ดังนี้

```
SELECT Store_name FROM Store_Information (2.2)
```

จากคำสั่ง SQL (2.2) จะ ได้ผลการสืบค้นข้อมูลทั้งหมดจากคอลัมน์ store_name ดังตารางที่ 2.2

ตารางที่ 2.2 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.2)

Store_name
Los Angeles
San Diego
Los Angeles
Boston

การสืบค้นข้อมูลสามารถทำการสืบค้นได้มากกว่า 1 คอลัมน์ขึ้นไป โดยถ้ามีมากกว่า 1 คอลัมน์ แต่ละคอลัมน์จะต้องกันด้วยเครื่องหมายคอมม่า “,” หลังคำสั่ง SELECT ดังนี้

```
SELECT Store_name,Sales FROM Store_Information (2.3)
```

จากคำสั่ง SQL (2.3) จะ ได้ผลการสืบค้นทั้งหมดจากคอลัมน์ store_name และคอลัมน์ Sales จากตาราง Store_Information ดังตารางที่ 2.3

ตารางที่ 2.3 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.3)

Store_name	Sales
Los Angeles	1500
San Diego	250
Los Angeles	300
Boston	700

ถ้าต้องการสืบค้นข้อมูลทั้งหมดในทุกคอลัมน์ของตารางจะใช้เครื่องหมาย ดอกจัน "*" หลังคำสั่ง SELECT ดังนี้

SELECT * FROM Store_Information (2.4)

จากคำสั่ง SQL (2.4) จะ ได้ผลการสืบค้นข้อมูลทั้งหมดจากตาราง Store_Information ดังตารางที่ 2.4

ตารางที่ 2.4 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.4)

Store_name	Sales
Los Angeles	1500
San Diego	250
Los Angeles	300
Boston	700

หากต้องการสืบค้นข้อมูลทั้งหมด โดยมีการระบุเงื่อนไขด้วยจะใช้คำสั่ง WHERE ร่วมด้วย ซึ่งจะมีการใช้คำสั่งกล่าวในรูปแบบ ดังนี้

SELECT * FROM "table_name" WHERE "column_name" "condition" (2.5)

การระบุเงื่อนไขนั้น ภาษา SQL จะมีรูปแบบเงื่อนไขพื้นฐานให้ใช้ดังนี้

หากต้องการสืบค้นข้อมูลโดยต้องการระบุเกี่ยวกับการเปรียบเทียบ (=, >, <, >=, <=) จะใช้คำสั่ง SQL ในรูปแบบ ดังนี้

SELECT * FROM Store_Information WHERE Store_name = "Lost Angeles" (2.6)

SELECT * FROM Store_Information WHERE Sales = "1500" (2.7)

SELECT * FROM Store_Information WHERE Sales > "250" (2.8)

SELECT * FROM Store_Information WHERE Sales < "700" (2.9)

SELECT * FROM Store_Information WHERE Sales >= "250") (2.10)

SELECT * FROM Store_Information WHERE Sales <= "700" (2.11)

จากคำสั่ง SQL (2.6) จะ ได้ผลการสืบค้นข้อมูลทั้งหมดจากตาราง Store_Information โดยมีการระบุเงื่อนไข คือ ต้องการสืบค้นข้อมูลที่มีชื่อหรือข้อความว่า "Los Angeles" จากคอลัมน์ Store_name ดังตารางที่ 2.6

ตารางที่ 2.5 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.6)

Store_name	Sales
Los Angeles	1500
Los Angeles	300

หากต้องการสืบค้นข้อมูล โดยต้องการระบุเงื่อนไขที่มากกว่าหนึ่งเงื่อนไขจะใช้คำสั่ง AND, OR ในการเชื่อมเงื่อนไขจะใช้คำสั่ง SQL ในรูปแบบ ดังนี้

SELECT * FROM Store_Information (2.12)

WHERE Sales >= "300" AND Sales <= "1800" (2.13)

SELECT * FROM Store_Information

WHERE Sales > "1000" OR Sales < "400" (2.14)

จากคำสั่ง SQL (2.12) จำได้ผลการสืบค้นข้อมูลทั้งหมดจากตาราง Store_Information โดยมีการระบุเงื่อนไข คือ ต้องการสืบค้นข้อมูลที่คอลัมน์ Sales มีค่ามากกว่าหรือเท่ากับ "300" และน้อยกว่าหรือเท่ากับ "1800" ดังตารางที่ 2.6

ตารางที่ 2.6 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.12)

Store_name	Sales
Los Angeles	1500
Los Angeles	300
Boston	700

หากต้องการสืบค้นข้อมูลโดยต้องการระบุเงื่อนไขที่เป็นรูปแบบจะใช้คำสั่ง LIKE และตามด้วยเงื่อนไขในรูปแบบ เช่น ขึ้นต้นด้วยตัวอักษร “B” จะแทนด้วย “B%”, ลงท้ายด้วยตัวอักษร “S” จะแทนด้วย “%S”, มีตัวอักษร “T” อยู่จะแทนด้วย “%T%” ซึ่งจะใช้คำสั่ง SQL ในรูปแบบ ดังนี้

```
SELECT * FROM Store_Information
WHERE Store_name LIKE "B%" (2.14)
```

```
SELECT * FROM Store_Information
WHERE Store_name LIKE "%S" (2.15)
```

```
SELECT * FROM Store_Information
WHERE Store_name LIKE "%T%" (2.16)
```

จากคำสั่ง SQL (2.14) จะได้ผลการสืบค้นข้อมูลทั้งหมดจากตาราง Store_Information โดยมีการระบุเงื่อนไข คือ ต้องการสืบค้นข้อมูลที่คอลัมน์ Store_name โดยขึ้นต้นด้วยตัวอักษร “B” ดังตารางที่ 2.7

ตารางที่ 2.7 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.14)

Store_name	Sales
Boston	700

หากต้องการสืบค้นข้อมูลโดยต้องการระบุเงื่อนไขให้มีการเรียงลำดับของผลการสืบค้นที่ได้จะใช้คำสั่ง ORDER BY แทนคำสั่ง WHERE ตามด้วย “column_name” และตามด้วยคำสั่ง ASC, DESC ซึ่งจะใช้คำสั่ง SQL ในรูปแบบ ดังนี้

```
SELECT * FROM Store_Information
ORDER BY Sales ASC
```

 (2.17)

```
SELECT * FROM Store_Information
ORDER BY Sales DESC
```

 (2.18)

จากคำสั่ง SQL (2.17) จะได้ผลการสืบค้นข้อมูลทั้งหมดจากราย Store_Information โดยมี การระบุเงื่อนไข คือ ต้องการสืบค้นข้อมูล โดยเรียงลำดับจากน้อยไปมาก โดยดูที่คอลัมน์ Sales ดัง ตารางที่ 2.8

ตารางที่ 2.8 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.17)

Store_name	Sales
San Diego	250
Los Angeles	300
Boston	700
Los Angeles	1500

หากต้องการสืบค้นข้อมูล โดยต้องการระบุเงื่อนไขที่เป็นรูปแบบของฟังก์ชัน เช่น ต้องการหา ค่าเฉลี่ยของคอลัมน์ Sales จะใช้คำสั่ง Average ต้องการหาค่าที่มากที่สุดใช้คำสั่ง Maximum ต้องการหาค่าที่น้อยที่สุดให้ใช้คำสั่ง Minimum ต้องการหาผลรวมใช้คำสั่ง Summary ต้องการนับ จำนวนข้อมูลทั้งหมดจะใช้คำสั่ง Count ซึ่งจะใช้คำสั่ง SQL ในรูปแบบ ดังนี้

```
SELECT AVG(Sales) FROM Store_Information
```

 (2.19)

```
SELECT MAX (Sales) FROM Store_Information
```

 (2.20)

```
SELECT MIN (Sales) FROM Store_Information
```

 (2.21)

```
SELECT SUM (Sales) FROM Store_Information
```

 (2.22)

```
SELECT COUNT (*) FROM Store_Information
```

 (2.23)

จากคำสั่ง SQL (2.20) จะได้ผลการสืบค้นข้อมูลทั้งหมดจากราย Store_Information โดยมี การระบุเงื่อนไข คือ ต้องการหาค่าที่สุด โดยดูที่คอลัมน์ Sales ดังตารางที่ 2.9

ตารางที่ 2.9 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.20)

MAX(Sales)
1500

ตารางที่ 2.10 ตัวอย่างตาราง Product

Product Id	Product Name	Supplier Id
1	Toshiba	1
2	Sony	3
3	Fujitsu	3

ตารางที่ 2.11 ตัวอย่างตาราง Supplier

Supplier Id	Supplier Name
1	A
2	B
3	C

จากตารางที่ 2.10 และ 2.11 หากต้องการสืบค้นข้อมูลจากตารางที่มีความสัมพันธ์กัน เช่น ต้องการสืบค้นข้อมูล Product และ Supplier ที่มีความสัมพันธ์กัน ซึ่งจะใช้คำสั่ง SQL ในรูปแบบดังนี้

```
SELECT * FROM Products, Suppliers
```

```
WHERE Products.SupplierId = Suppliers.SupplierId
```

(2.24)

จากคำสั่ง SQL (2.24) จะได้ผลการสืบค้นข้อมูลทั้งหมดจากตาราง Products และ Suppliers ที่มีความสัมพันธ์กัน ดังตารางที่ 2.12

ตารางที่ 2.12 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.24)

Product Id	Product Name	Supplier Id	Supplier Id	Supplier Name
1	Toshiba	1	1	A
2	Sony	3	3	C
3	Fujitsu	3	3	C

จากตารางที่ 2.10 และ 2.11 หากต้องการสืบค้นข้อมูลจากตารางที่มีความสัมพันธ์กัน เช่น ต้องการสืบค้นข้อมูล Products และ Suppliers ที่มีความสัมพันธ์กัน และมีการระบุเงื่อนไขด้วย ซึ่งจะใช้คำสั่ง SQL ในรูปแบบ ดังนี้

```
SELECT * FROM Products, Suppliers
WHERE Products.SupplierId = Suppliers.SupplierId and SupplierName = "C" (2.25)
```

จากคำสั่ง SQL (2.25) จะ ได้ผลการสืบค้นข้อมูลทั้งหมดจากราย Products และ Suppliers ที่มีความสัมพันธ์กันและมีการระบุเงื่อนไขด้วย ดังตารางที่ 2.13

ตารางที่ 2.13 ผลการสืบค้นข้อมูลจากคำสั่ง SQL (2.25)

Product Id	Product Name	Supplier Id	Supplier Id	Supplier Name
2	Sony	3	3	C
3	Fujitsu	3	3	C

2.3 การประมวลผลภาษาธรรมชาติโดยวิธีการแจกแจงคำ

เนื่องจากระบบถามตอบที่ใช้ร่วมกับ โปรแกรมประยุกต์บนเว็บนี้มีลักษณะการทำงานเป็นการแปลงจากประโยคภาษาธรรมชาติให้กลายเป็นภาษา SQL หลักสำคัญสำหรับการแปลงจากภาษาใดภาษาหนึ่งไปเป็นอีกภาษาหนึ่งนั่นคือ การแปลงภาษานั้นๆจะต้องไม่ทำให้โครงสร้างทางความหมายของภาษานั้นผิดเพี้ยนไป การแจกแจงคำเป็นกระบวนการหนึ่งในการประมวลผลภาษาธรรมชาติ (Natural Language Processing) ซึ่งจะทำให้การแยกคำต่างๆที่มีในประโยคออกจากกัน แล้วระบุความหมายพร้อมทั้งหน้าที่ทางไวยากรณ์ของคำต่างๆเหล่านั้น การที่จะสามารถระบุหน้าที่และความหมายของคำต่างๆเหล่านี้ได้อย่างถูกต้องจะต้องอาศัยการป้อนข้อมูลทางไวยากรณ์เพื่อใช้สำหรับการอ้างอิงในกระบวนการระบุความหมายและหน้าที่ทางไวยากรณ์ของคำต่างๆ

เนื่องจากระบบถามตอบ จะทำหน้าที่ในการแปลงภาษาธรรมชาติไปเป็นภาษา SQL เพื่อใช้ในการสืบค้นบนฐานข้อมูลเชิงสัมพันธ์ ข้อมูลทางไวยากรณ์ที่เอาไว้สำหรับอ้างอิงในกระบวนการระบุความหมายและหน้าที่ทางไวยากรณ์ของคำต่างๆในประโยคจะต้องสามารถบ่งบอกได้ว่าคำหรือกลุ่มคำใดๆในประโยคมีความหมายในแง่การสืบค้น หรือเป็นส่วนหนึ่งส่วนใดบนฐานข้อมูลอย่างไร เพื่อจะได้สามารถเข้าใจได้ว่าประโยคคำถามที่ป้อนเข้ามาต้องการสืบค้นข้อมูลใดที่อยู่บนฐานข้อมูลตัวอย่างเช่นหากผู้ใช้งานต้องการทราบข้อมูลของผู้จำหน่ายสินค้า (Suppliers) ทั้งหมดบนฐานข้อมูล ผู้ใช้งานจะถามด้วยคำถามเช่นคำถามที่ 2.26 และหากผู้ใช้งานต้องการทราบข้อมูลของสินค้า (Products) ทั้งหมดบนฐานข้อมูล ผู้ใช้งานจะถามด้วยคำถามเช่นคำถามที่ 2.27

Show me all suppliers. (2.26)

Show me all products. (2.27)

ประโยคสืบค้นของภาษา SQL ที่ให้ความหมายเดียวกับประโยคคำถามที่ 2.26 และ 2.27 คือประโยคสืบค้นภาษา SQL 2.28 และ 2.29 ตามลำดับ

SELECT * FROM suppliers (2.28)

SELECT * FROM products (2.29)

เมื่อเรามาวិเคราะห์โครงสร้างของประโยคภาษาอังกฤษ 2.26 และ 2.27 เราจะเห็นได้ว่าทั้งสองประโยคมีโครงสร้างที่เหมือนกันแตกต่างกันเพียงข้อมูลที่ต้องการหาหรือชื่อของตารางนั่นเอง โครงสร้างดังกล่าวเราสามารถเขียนให้อยู่ในรูปแบบทั่วไปได้แก่ประโยคที่ 2.30

Show me all <table_name>. (2.30)

ประโยคสืบค้น SQL รูปแบบทั่วไปที่มีความหมายเหมือนกับประโยคที่ 2.30 สามารถเขียนได้ดังนี้

SELECT * FROM <table_name> (2.31)

คำว่า <table_name> สามารถใช้แทนด้วยชื่อตารางต่างๆที่มีในฐานข้อมูลได้เช่นตาราง suppliers และ products คึงที่กล่าวมาข้างต้น นอกจากนี้ยังมีรูปแบบประโยคภาษาอังกฤษที่สามารถอธิบายความหมายเช่นเดียวกับประโยคที่ 2.26 ดังนี้

Show me all our suppliers. (2.32)

Show me all suppliers. (2.33)

Show all our suppliers. (2.34)

Show all suppliers. (2.35)

Show me our suppliers. (2.36)

Show me suppliers. (2.37)

Show our suppliers. (2.38)

Show suppliers. (2.39)

ประโยคที่ 2.32 – 2.39 สามารถเขียนให้อยู่ในรูปไวยากรณ์เดียวในรูปแบบทั่วไปได้ดังนี้ โดยที่คำที่อยู่ในเครื่องหมายวงเล็บจะมีความหมายว่ามีค่าๆนั้นหรือไม่มีในประโยคก็ได้

Show (me) (all) (our) <table_names> (2.40)

นอกจากนี้เราสามารถใช้อื่นแทนคำว่า show แล้วให้มีความหมายเดียวกันกับประโยคที่ 2.40 เช่นคำว่า list และ display คึงประโยคต่อไปนี้

List me all our suppliers. (2.41)

Display me all our suppliers. (2.42)

จากประโยคที่ 2.41-2.42 เราจะสามารถเขียนไวยากรณ์ขึ้นมาใหม่เพื่อรองรับการใช้งานได้ครอบคลุมมากขึ้นดังนี้

```
[Show|List|Display] (me) (all) (our) <table_names> (2.43)
```

จากไวยากรณ์ที่ 2.43 กลุ่มคำที่เขียนคั่นด้วยเครื่องหมาย ชิดตั้ง () และอยู่ในวงเล็บสี่เหลี่ยม ([]) มีความหมายว่าจะต้องมีค่าใดค่าหนึ่งประกฏในประโยคจึงจะทำให้ประโยคสมบูรณ์ และจากไวยากรณ์ที่ 2.43 นี้จะสามารถเข้ากันได้กับประโยคภาษาอังกฤษถึง 24 ประโยค โดยที่ประโยคเหล่านี้สามารถแปลงเป็นประโยคสืบค้นภาษา SQL ประโยคเดียวดังนี้

```
SELECT * FROM <table_name> (2.44)
```

เราสามารถเพิ่มกฎไวยากรณ์ต่างๆเข้าไปเพื่อใช้สำหรับอ้างอิงการสืบค้นข้อมูลในรูปแบบต่างๆได้อีก เช่น หากต้องการเลือกสืบค้นข้อมูลเพียงคอลัมน์เดียวจากฐานข้อมูลเช่น ต้องการทราบชื่อของผู้จำหน่ายสินค้าทั้งหมด หรือต้องการทราบรายละเอียดของสินค้าทั้งหมด เราสามารถเขียนเป็นประโยคภาษาอังกฤษได้ดังนี้

```
Show the names of our suppliers. (2.45)
```

```
Show the descriptions of our products. (2.46)
```

ประโยคภาษาอังกฤษ 2.45 และ 2.46 สามารถเขียนให้อยู่ในรูปแบบไวยากรณ์ทั่วไปได้ดังนี้

```
Show (the) <column_name> of (our) <table_name> (2.47)
```

ประโยคสืบค้น SQL รูปแบบทั่วไปที่มีความหมายเหมือนกับประโยคที่ 2.47 สามารถเขียนได้ดังนี้

```
SELECT DISTINCT <column_name> FROM <table_name> (2.48)
```

จากตัวอย่างข้างต้นจะสังเกตเห็นได้ว่าในส่วนของไวยากรณ์ที่ระบุไว้ว่าเป็นชื่อของคอลัมน์เมื่อไปเปรียบเทียบกับคำอยู่ในประโยคภาษาอังกฤษของประโยคที่ 2.45 เราจะได้คำว่า names ซึ่งไม่ใช่ชื่อของคอลัมน์ที่มีอยู่ในตาราง suppliers คอลัมน์ของชื่อผู้จำหน่ายนั้นจะมีชื่อว่า name ทำให้การแปลง

ไม่ว่ากรณ์ใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประโยคสืบไม่สมบูรณ์ ปัญหานี้เราสามารถแก้ไขได้โดยใช้กลไกการแปลงคำให้อยู่ในรูปแบบพื้นฐานของ WordNet ที่ได้กล่าวไว้ในส่วนก่อนหน้า กลไกดังกล่าวจะแปลงคำว่า names ให้เป็น name ซึ่งเหมือนกับชื่อของคอลัมน์ ทำให้การแปลงประโยคสืบค้น SQL ทำได้ถูกต้อง

ในกรณีที่ต้องการเลือกสืบค้นข้อมูลสองคอลัมน์จากฐานข้อมูลเช่น ต้องการทราบชื่อและที่อยู่ของผู้จำหน่ายสินค้าทั้งหมด หรือต้องการทราบรายละเอียดของสินค้าและราคาของสินค้าทั้งหมด เราสามารถเขียนเป็นประโยคภาษาอังกฤษได้ดังนี้

Show me names and address of our suppliers. (2.49)

Show me descriptions and cost of our products. (2.50)

ประโยคภาษาอังกฤษ 2.49 และ 2.50 สามารถเขียนให้อยู่ในรูปแบบไวยากรณ์ทั่วไปได้ดังนี้

Show (me) <column_name1> and <column_name2> of (our) <table_name>. (2.51)

ประโยคสืบค้น SQL รูปแบบทั่วไปที่มีความหมายเหมือนกับประโยคที่ 2.51 สามารถเขียนได้ดังนี้

```
SELECT DISTINCT <column_name1>,<column_name2>
FROM <table_name>
```

(2.52)

ในกรณีที่ต้องการเลือกสืบค้นข้อมูลพร้อมทั้งระบุเงื่อนไขเช่น ต้องการทราบชื่อรายการสินค้าที่จำหน่ายโดยผู้จำหน่ายชื่อ ABC หรือต้องการทราบที่อยู่ของผู้จำหน่ายชื่อ XYZ เราสามารถเขียนเป็นประโยคภาษาอังกฤษได้ดังนี้

Show me the names of our products supplied by ABC. (2.53)

Show me the address of our supplier whose name is XYZ. (2.54)

ประโยคภาษาอังกฤษ 2.53 และ 2.54 สามารถเขียนให้อยู่ในรูปแบบไวยากรณ์ทั่วไปได้ดังนี้

Show (me) (the) <column_name1> of (our) <table_name>

(whose)<column_name2> [by|be] <value>

(2.55)

ประโยคสืบค้น SQL รูปแบบทั่วไปที่มีความหมายเหมือนกับประโยคที่ 2.55 สามารถเขียนได้ดังนี้

```
SELECT <column_name1> FROM <table_name>
WHERE <column_name2> = <value>
```

(2.56)

ในกรณีที่ต้องการสืบค้นข้อมูลที่มีความสัมพันธ์กันระหว่างสองตารางเช่น ต้องการทราบชื่อผู้จำหน่ายชื่อสินค้าและสินค้าที่ผู้จำหน่ายมีทั้งหมด เราสามารถเขียนเป็นประโยคภาษาอังกฤษได้ดังนี้

Show me our suppliers and their products. (2.57)

ประโยคภาษาอังกฤษ 2.57 สามารถเขียนให้อยู่ในรูปแบบไวยากรณ์ทั่วไปได้ดังนี้

Show (me) (our) <table_name1> and (their) <table_name2> (2.58)

ประโยคสืบค้น SQL รูปแบบทั่วไปที่มีความหมายเหมือนกับประโยคที่ 2.58 สามารถเขียนได้ดังนี้

```
SELECT * FROM <table_name1>, <table_name2>
WHERE <relationship_between_tables1_and_2>
```

(2.59)

จากกรณีข้างต้นตาราง table_name1 และ table_name2 จะต้องมีความสัมพันธ์ต่อกันหากทั้งสองตารางไม่มีความสัมพันธ์ต่อกันจะทำให้ข้อมูลที่สืบค้นเป็นข้อมูลที่ไม่มีความหมาย ดังนั้นในกรณีที่มีการสืบค้นข้อมูลระหว่าง2ตาราง จะต้องทำการตรวจสอบก่อนว่าทั้งสองตารางมีความสัมพันธ์ต่อกันหรือไม่ หากมี ให้ดำเนินการต่อไป

จากวิธีการที่ได้อธิบายมาข้างต้นจะเห็นได้ว่าเราสามารถใช่วิธีแจกแจงคำ โดยนำประโยคภาษาอังกฤษมาเปรียบเทียบกับไวยากรณ์ที่เขียนไว้ข้างต้น เพื่อจะได้ทราบว่าส่วนใดของประโยคทำหน้าที่ใดในประโยคสืบค้นภาษา SQL โดยความถูกต้องของการแจกแจงคำภาษาอังกฤษให้เป็นภาษา SQL ขึ้นอยู่กับความถูกต้องของกฎไวยากรณ์ที่ใช้ในการอ้างอิง

2.4 นิพจน์ปรกติ

ในวิทยาการคอมพิวเตอร์ นิพจน์ปรกติ (regular expression สามารถย่อได้เป็น regexp หรือ regex) คือสตริงที่อธิบายถึงรูปแบบของสตริงตาม โครงสร้างรูปแบบที่กำหนด นิพจน์ปรกตินั้นใช้ อยู่แพร่หลายในโปรแกรมประเภท Text editor ในการค้นหาและปรับเปลี่ยนข้อความ ภาษา โปรแกรมหลายภาษายังรองรับการใช้นิพจน์ปรกติสำหรับการจัดการและปรับเปลี่ยนสตริง

2.4.1 ภาพรวม

นิพจน์ปรกติ ใช้อธิบายรูปแบบสตริง ซึ่งการอธิบายมักจะกระชับและรัดกุมโดยที่ไม่ ต้องอธิบายเป็นรายชื่อทั้งหมด ยกตัวอย่างเช่นคำว่า ปกติ และ ปรกติ สามารถอธิบายได้ว่า "ป(ร?) กติ" ซึ่งส่วนใหญ่มีโครงสร้างการใช้ดังนี้

2.4.2 การสร้างทางเลือก

| เครื่องหมายขีดตั้ง ใช้สำหรับสร้างทางเลือกที่จะใช้ค้นหา นิพจน์ เช่น "express|expand" มีความหมายว่า express หรือ expand

2.4.3 การจัดกลุ่ม

() วงเล็บ ใช้ในการกำหนดลำดับและขอบเขตของนิพจน์ เช่น "ex(pres|pan)sion" มีความหมายว่า expression หรือ expansion

2.4.4 การบอกจำนวน

- ? ปรกติ ใช้บ่งบอกว่ามีนิพจน์ก่อนหน้านี้นี้จำนวน 0 หรือ 1 นิพจน์ (มีนิพจน์ หรือ ไม่มีก็ได้) เช่น "ปร?กติ" มีความหมายว่า ปกติ หรือ ปรกติ
- * เครื่องหมายดอกจัน ใช้บ่งบอกว่ามีนิพจน์ก่อนหน้านี้นี้จำนวน 0 นิพจน์หรือ มากกว่า เช่น "go*gle" มีความหมายว่า ggle gogle google gooogle เป็นต้น
- + เครื่องหมายบวก ใช้บ่งบอกว่ามีนิพจน์ก่อนหน้านี้นี้จำนวน 1 นิพจน์หรือ มากกว่า เช่น "go+gle" มีความหมายว่า gogle google gooogle เป็นต้น แต่ไม่ เท่ากับ ggle

- {n} วงเล็บปีกกาใช้บ่งบอกว่าต้องมีนิพจน์ก่อนหน้าจำนวน n นิพจน์ เช่น "go{5}gle" มีความหมายว่า goooogle เพียงตัวเดียว
- {m,n} วงเล็บปีกกาที่มีจุดภาคข้างใน ใช้บ่งบอกว่าต้องมีนิพจน์ก่อนหน้าจำนวน m นิพจน์ถึง n นิพจน์ เช่น "go{2,4}gle" มีความหมายว่า google gooogle และ goooogle หากไม่ระบุตัวเลข n จะหมายถึงต้องมีนิพจน์จำนวนตั้งแต่ m นิพจน์เป็นต้นไป

2.4.5 การระบุตำแหน่ง

- ^ เครื่องหมายหมวก ใช้เขียนเป็นตัวแรกสุดของนิพจน์ปกติ เพื่อบ่งบอกว่านิพจน์ที่ต้องการค้นหาอยู่ต้นบรรทัดหรือต้นสตริงเท่านั้น สามารถใช้คู่กับ \$
- \$ เครื่องหมายสกุลเงิน ดอลลาร์ ใช้เขียนเป็นตัวท้ายสุดของนิพจน์ปกติ เพื่อบ่งบอกว่านิพจน์ที่ต้องการค้นหาอยู่ท้ายบรรทัดหรือท้ายสตริงเท่านั้น สามารถใช้คู่กับ ^

2.4.6 การระบุช่วงอักขระ

- [] วงเล็บเหลี่ยม ใช้เขียนแทนอักขระหนึ่งตัวที่อยู่ในขอบเขต ซึ่งระบุไว้ภายในวงเล็บ เช่น "[bcdf]an" มีความหมายว่า ban can dan และ fan
- ^ เครื่องหมายหมวก ใช้เขียนภายในวงเล็บเหลี่ยมเป็นตัวแรก เพื่อบ่งบอกว่าเป็นอักขระหนึ่งตัวที่ไม่ได้อยู่ในขอบเขต เช่น "[^bcdf]an" มีความหมายว่า aan ean gan han ian เป็นต้น
- ยัติภังค์ ใช้เขียนภายในวงเล็บเหลี่ยมระหว่างอักขระสองตัว เพื่อบ่งบอกว่าเป็นอักขระหนึ่งตัวที่มีรหัสแอสกีหรือยูนิโคคอยู่ในช่วงนั้น เช่น "[a-z]an" มีความหมายว่า aan ban can เรื่อยไปจนถึง zan

2.4.7 เครื่องหมายอื่นๆ

- .
- มหัพภาค ใช้บ่งบอกว่าอักขระในตำแหน่งนั้นของนิพจน์จะเป็นตัวอักษรอะไรก็ได้ เช่น ".an" มีความหมายว่า aan ban can Oan lan -an เป็นต้น

แบ็กสแลช ใช้ระบุ escape character หรือสัญลักษณ์ที่ต้องการใช้ในนิพจน์ แต่ซ้ำกับสัญลักษณ์ที่กล่าวมาทั้งหมดด้านบน เช่น "[0-9]\.[0-9]\$" มีความหมายว่า 2.3\$ 9.0\$ 1.7\$ เป็นต้น

โดยเครื่องหมายต่างๆ สามารถสร้างนิพจน์ปรกติที่ซ้ำซ้อนและมีประสิทธิภาพสูงได้ ซึ่งคล้ายกับการสร้างนิพจน์คณิตศาสตร์นั่นเอง

2.5 การเขียนไวยากรณ์ในการแจกแจงด้วย นิพจน์ปรกติ

จากที่ได้ศึกษาการประมวลภาษาธรรมชาติโดยวิธีการแจกแจงค่า และการใช้นิพจน์ปรกติในการประมวลผลข้อความในส่วนข้างต้น ในส่วนนี้จะเป็นการกล่าวถึงการเขียนไวยากรณ์ที่จะนำไปใช้จริงในการทำงานของระบบถามตอบ

จากไวยากรณ์ที่ 2.43 ในส่วนที่ 2.3 ที่เขียนดังนี้

$$[\text{Show|List|Display}] (\text{me}) (\text{all}) (\text{our}) \langle \text{table_names} \rangle \quad (2.60)$$

เราสามารถเขียนให้อยู่ในรูปของนิพจน์ปรกติซึ่งมีความหมายครอบคลุม ไวยากรณ์ข้างต้นและมีส่วนขยายต่างๆเพิ่มเข้าไปเพื่อให้มีความยืดหยุ่นในการใช้งานเพิ่มขึ้น สามารถเขียนได้ดังต่อไปนี้

$$\wedge((\text{show|list|display|tell})\text{s}?((\text{to}\text{s})?(\text{me|us})?)?((\text{what|who})(\text{sbe}'(\text{s|re})|\text{sis}|\text{sare}|\text{swas}|\text{swere})\text{s}?(\text{the|every|all}(\text{s|the})?|\text{any|our})?\text{s}?(\text{our})?\text{s}?(? \langle \text{table} \rangle (\text{w}+\text{s})?\text{w}+)\text{s}?(\text{[.?'s]})?\text{s} \quad (2.61)$$

จะเห็นได้ว่าไวยากรณ์ที่เขียนด้วยนิพจน์ปรกติเพื่อนำไปใช้กับการประมวลผลจริงของโปรแกรมไม่สามารถอ่านให้เข้าใจได้โดยง่าย จึงได้มีการออกแบบ ไวยากรณ์กลางสำหรับแปลงจากภาษาที่ให้ผู้ใช้งานอ่านเข้าใจได้ง่ายไปเป็นนิพจน์ปรกติที่สามารถนำไปใช้ได้จริง เราเรียกไวยากรณ์นี้ว่า Pre-defined Grammar โดยไวยากรณ์ในรูปของนิพจน์ปรกติที่ 2.61 สามารถเขียนให้อยู่ในรูปของ Pre-defined Grammar ดังต่อไปนี้

$$\wedge(\text{ask|question}) \text{qualifier?} (\text{our})? \langle \text{table} \rangle \langle \text{eol} \rangle \quad (2.62)$$

โดย

$$\text{ask} = (\text{show|list|display|tell})\text{s}?((\text{to}\text{s})?(\text{me|us})?)?$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

question	=	(what who)(\sbe '\sire)\s \sare \swas \swere)
qualifier	=	(the every all(\sthe)? any our)
<table>	=	(?<table>(\w+\s)?\w+)
<eol>	=	([.\s])?$



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

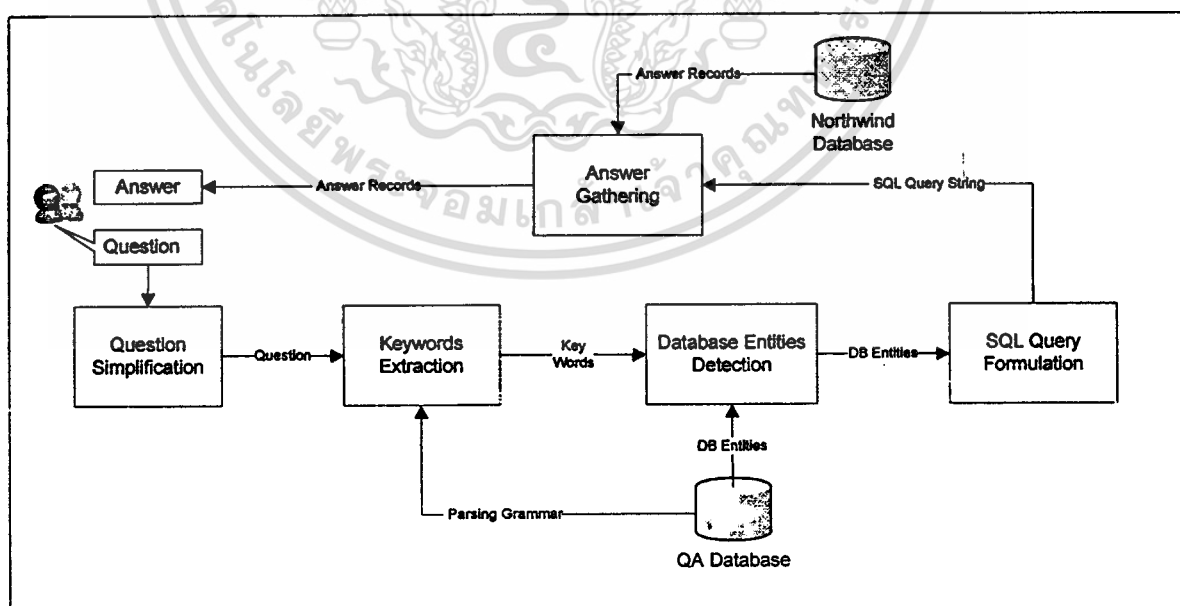
การวิเคราะห์และออกแบบระบบ

ในบทที่ 3 จะได้กล่าวถึงขั้นตอนการออกแบบระบบถามตอบที่ใช้งานร่วมกับโปรแกรมประยุกต์บนเว็บ โดยจะอธิบายถึงสถาปัตยกรรมของระบบ การทำงานโดยรวมของระบบ การทำงานแยกตามส่วนและการออกแบบโครงสร้างข้อมูลของระบบ

3.1 การออกแบบระบบถามตอบ

รูปที่ 3.1 แสดงแผนภาพแสดงลำดับการทำงานของระบบถามตอบซึ่งถูกออกแบบให้มีหน่วยย่อยต่างๆด้วยกัน 5 หน่วยดังนี้

1. หน่วยแปลงคำถามให้อยู่ในรูปอย่างง่าย (Question Simplification Module)
2. หน่วยแยกคีย์เวิร์ดคำถาม (Keywords Extraction Module)
3. หน่วยตรวจจับ Entity ต่างๆของฐานข้อมูล (Database Entities Detection Module)
4. หน่วยสร้างประโยคสืบค้น SQL (SQL Query Formulation Module)
5. หน่วยรวบรวมคำตอบ (Answer Gathering Module)



รูปที่ 3.1 แผนภาพแสดงลำดับการทำงานของระบบถามตอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.1 เราสามารถอธิบายการทำงานโดยรวมของระบบโดยการจำลองการทำงานของระบบ เริ่มต้นโดยการให้ผู้ใช้งานเริ่มต้นป้อนคำถามภาษาธรรมชาติเข้ามาในระบบ เช่น คำถามที่ต้องการทราบว่าข้อมูลของผู้จำหน่ายสินค้า (Suppliers) ทั้งหมดบนฐานข้อมูลเป็นอย่างไรบ้าง ผู้ใช้จะป้อนคำถามตามประโยคที่ 3.1 เพื่อส่งไปยังหน่วยแปลงคำถามให้อยู่ในรูปอย่างง่าย

Show me all suppliers.(3.1)

หน่วยแปลงคำถามให้อยู่ในรูปอย่างง่าย จะทำหน้าที่ในการแปลงคำถามที่ได้รับให้อยู่ในรูปแบบพื้นฐานเพื่อนำไปใช้ในการประมวลผลขั้นต่อไปด้วยการแยกคำต่างๆที่มีในประโยคคำถามออก เปลี่ยนตัวอักษรที่อยู่ในรูปตัวพิมพ์ใหญ่ทั้งหมดให้กลายเป็นตัวพิมพ์เล็ก ตัด white space ที่ปรากฏด้านหน้าและท้ายประโยคออกเปลี่ยน white space ทุกชนิดที่ปรากฏให้กลายเป็นช่องว่าง 1 ตัวอักษร และเรียบเรียงประโยคใหม่ซึ่งจะสามารถเขียนประโยคคำถามใหม่ได้ดังประโยคที่ 3.2 เพื่อส่งไปยังหน่วยแจกแจงคำถาม

show me all suppliers. (3.2)

หน่วยแยกคีย์เวิร์ดจะนำเอาคำถามที่ 3.2 มาเพื่อทำการเปรียบเทียบกับ ไวยากรณ์ที่ใช้ในการแจกแจงคำถามที่เก็บไว้ในฐานข้อมูล ซึ่งอธิบายไว้ในบทที่ 2 โดยประโยคที่ 3.2 สามารถเปรียบเทียบได้ตรงกับไวยากรณ์ดังต่อไปนี้

[show|list|display] (me) (all) (our) <table_name> (3.3)

ไวยากรณ์ดังกล่าวจะมีการระบุรูปแบบภาษาสืบค้น SQL ที่ใช้สำหรับการสร้างประโยคสืบค้น SQL จะเป็นดังคำสั่งที่ 3.4

SELECT * FROM <table_name> (3.4)

ส่วนที่เป็นคีย์เวิร์ดจะเป็นส่วนที่เขียนอยู่ในเครื่องหมาย <> โดยจะมีการกำกับว่าคีย์เวิร์ดนี้มีหน้าที่อะไรในประโยคสืบค้น SQL ที่จะทำการสร้างขึ้น จากไวยากรณ์ที่ 3.3 เมื่อเทียบกับประโยค 3.2 เราจะสามารถดึงคีย์เวิร์ดคำว่า suppliers ออกมาได้โดยคีย์เวิร์ดที่ได้จะมีหน้าที่เป็นชื่อของตารางที่จะทำการสืบค้นในประโยคสืบค้น SQL 3.4 คีย์เวิร์ด suppliers ที่แยกออกมาได้จะถูกส่งไปยัง

หน่วยตรวจจับ Entity ต่างๆของฐานข้อมูล และรูปแบบประโยคสืบค้น 3.4 จะถูกส่งไปยังหน่วยสร้างประโยคสืบค้น SQL

หน่วยตรวจจับ Entity ต่างๆของฐานข้อมูลจะนำเอาคีย์เวิร์ดที่ได้พร้อมกับหน้าที่ในประโยคสืบค้น ไปค้นหาในข้อมูลโครงสร้างฐานข้อมูลว่าคีย์เวิร์ดที่ได้หมายความว่าถึง Entity ใดของฐานข้อมูล ในที่นี้คีย์เวิร์ดที่ได้คือ suppliers มีหน้าที่เป็นชื่อของตาราง ซึ่งเมื่อเปรียบเทียบกับข้อมูลโครงสร้างฐานข้อมูล Northwind จะหมายถึงตารางที่มีชื่อว่า Suppliers ชื่อตารางดังกล่าวจะถูกส่งไปยังหน่วยสร้างประโยคสืบค้น SQL

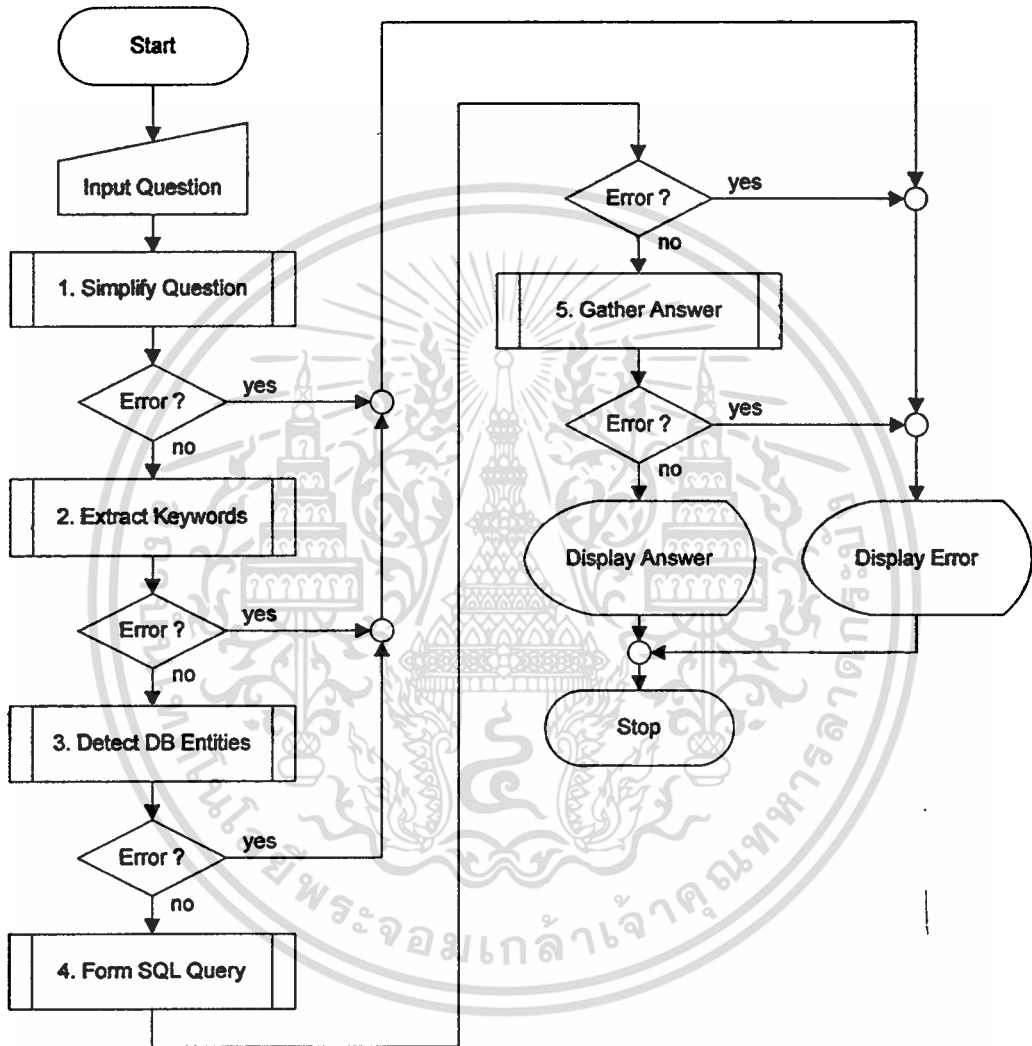
หน่วยสร้างประโยคสืบค้นจะนำเอารูปแบบประโยคสืบค้น SQL และ Entity ต่างๆของฐานข้อมูลมาสร้างประโยคสืบค้น จากรูปแบบประโยคสืบค้นที่ 3.4 และชื่อตาราง Suppliers ที่ได้จะทำให้สามารถสร้างประโยคสืบค้น SQL ได้ดังนี้

SELECT * FROM Suppliers (3.5)

ประโยคสืบค้นภาษา SQL ที่ 3.5 จะถูกส่งไปยังหน่วยรวบรวมคำตอบโดยจะติดต่อกับระบบจัดการฐานข้อมูลของฐานข้อมูล Northwind เพื่อทำการสืบค้นคำตอบตามที่ผู้ใช้งานต้องการ โดยการส่งประโยคสืบค้นภาษา SQL ที่ 3.5 ให้กับระบบจัดการฐานข้อมูล จากนั้นหน่วยรวบรวมคำตอบจะรวบรวมระเบียบฐานข้อมูลที่สืบค้นได้เพื่อนำไปแสดงผลให้กับผู้ใช้งานเพื่อเป็นคำตอบของคำถามที่ 3.1

3.2 แผนภาพแสดงลำดับการทำงานในระบบถามตอบ

เพื่อให้เข้าใจการทำงานของระบบถามตอบที่ใช้งานร่วมกับโปรแกรมประยุกต์บนเว็บได้อย่างดี ขึ้น เราสามารถอธิบายการทำงานของระบบด้วยแผนภาพแสดงลำดับการทำงานซึ่งเขียนอธิบายลำดับการทำงานของระบบ และกระบวนการต่างๆภายในระบบถามตอบ



รูปที่ 3.2 แผนภาพแสดงลำดับการทำงานโดยรวมของระบบถามตอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.1 แผนภาพแสดงลำดับการทำงานโดยรวม

แผนภาพแสดงลำดับการทำงานโดยรวมแสดงตามรูปที่ 3.2 เป็นการแสดงลำดับการทำงานของระบบถามตอบ เพื่อให้เห็นภาพรวมของการทำงานของระบบว่ามีเริ่มทำงานที่กระบวนการใดและแต่ละกระบวนการหลักของระบบทำงานตามลำดับอย่างไร จากแผนภาพที่ 3.2 เราสามารถอธิบายลำดับการทำงานของระบบได้ดังต่อไปนี้

เริ่มต้นระบบด้วยการป้อนคำถามเข้าไปในระบบและจะถูกส่งไปยังกระบวนการแปลงคำถามให้อยู่ในรูปอย่างง่าย (Simplify Question Process) หากเกิดข้อผิดพลาดจะแสดงข้อผิดพลาดรายงานให้ผู้ใช้งานทราบหากไม่มีข้อผิดพลาดจะส่งไปประมวลผลต่อยังกระบวนการแยกคีย์เวิร์ด (Extract Keywords Process) หากเกิดข้อผิดพลาดจะแสดงข้อผิดพลาดรายงานให้ผู้ใช้งานทราบหากไม่มีข้อผิดพลาดจะส่งไปประมวลผลต่อยังกระบวนการตรวจจับ Entity ต่างๆของฐานข้อมูล (Detect Database Entities Process) หากเกิดข้อผิดพลาดจะแสดงข้อผิดพลาดรายงานให้ผู้ใช้งานทราบหากไม่มีข้อผิดพลาดจะส่งไปประมวลผลต่อยังกระบวนการสร้างสร้างประโยคสืบค้น (Form SQL Query Process) หากเกิดข้อผิดพลาดจะแสดงข้อผิดพลาดรายงานให้ผู้ใช้งานทราบหากไม่มีข้อผิดพลาดจะส่งไปประมวลผลต่อยังกระบวนการรวบรวมคำตอบ (Gather Answer Process) หากเกิดข้อผิดพลาดจะแสดงข้อผิดพลาดรายงานให้ผู้ใช้งานทราบหากไม่มีข้อผิดพลาดจะแสดงผลคำตอบให้ผู้ใช้งานทราบในรูปแบบ ระเบียบฐานข้อมูล

จากแผนภาพแสดงลำดับการทำงาน โดยรวมจะมีกระบวนการหลักอยู่ด้วยกัน 5 กระบวนการ ดังต่อไปนี้

- 1) กระบวนการแปลงคำถามให้อยู่ในรูปอย่างง่าย (Simplify Question Process) เป็นกระบวนการประมวลคำถามเบื้องต้นเพื่อให้พร้อมในการทำงานในระบบถามตอบ กระบวนการนี้จะทำหน้าที่ในการเรียบเรียงประโยคคำถามให้อยู่ในรูปแบบที่พร้อมสำหรับการประมวลผลในระบบถามตอบ หากขั้นตอนในการแปลงคำถามมีปัญหาจะส่งข้อผิดพลาดไปเพื่อรายงานความผิดพลาดให้กับผู้ใช้งานทราบ
- 2) กระบวนการแยกคีย์เวิร์ด (Extract Keyword Process) เป็นกระบวนการแยกคีย์เวิร์ดต่างๆ ที่มีในประโยคคำถามออกมา ในกระบวนการนี้จะนำเอาคำถามที่ได้รับมาทำการประมวลผลภาษาธรรมชาติด้วยวิธีการแจกแจงคำ โดยจะอาศัยไวยากรณ์ที่ได้มีการเก็บไว้ในฐานข้อมูล ซึ่งจะทำได้ทำให้สามารถแยกคีย์เวิร์ดที่เกี่ยวข้อง และรูปแบบที่ใช้ในการสร้างประโยคสืบค้น SQL (SQL Query Pattern) โดยที่คีย์เวิร์ดต่างๆที่แยกออกมาได้จะถูกนำไปใช้ในกระบวนการตรวจจับ Entity ของฐานข้อมูล (Detect DB Entity Process) และรูปแบบที่ใช้ในการสร้างประโยคสืบค้น SQL จะถูกส่งไปยังกระบวนการสร้างประโยค

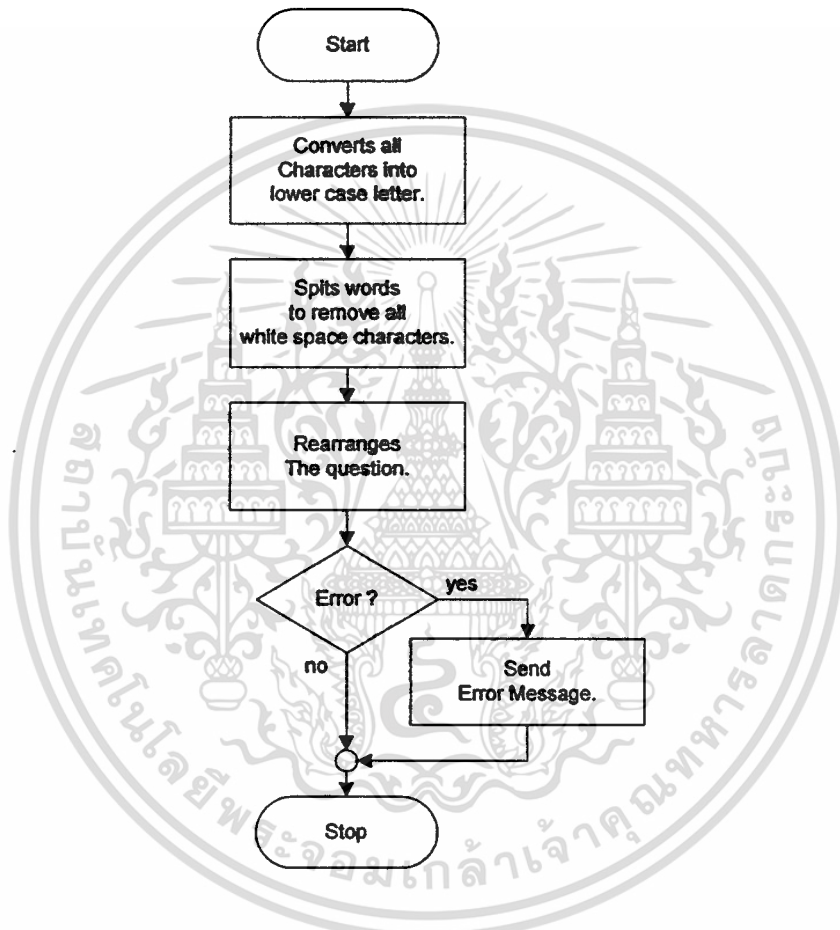
สืบค้น SQL (Form SQL Query Process) หากขั้นตอนในการแยกคีย์เวิร์คมีปัญหาจะส่งข้อผิดพลาดไปเพื่อรายงานความผิดพลาดให้กับผู้ใช้งานทราบ

- 3) กระบวนการตรวจจับ Entity ของฐานข้อมูล (Detect DB Entity Process) เป็นกระบวนการที่คอยตรวจสอบว่าในคำถามควรจะหาข้อมูลจาก Entity ใดบ้างของฐานข้อมูล โดยกระบวนการนี้จะอาศัยข้อมูลคีย์เวิร์คต่างๆที่ได้จากกระบวนการแยกคีย์เวิร์คเพื่อนำไปเปรียบเทียบกับข้อมูลการจับคู่คีย์เวิร์คในฐานข้อมูลจากข้อมูลโครงสร้างของฐานข้อมูล (Database Schema) โดยจะใช้ประเภทของฐานข้อมูลมาช่วยในการกรองข้อมูลอีกระดับหนึ่ง Entity ต่างๆที่ตรวจจับได้จะถูกส่งไปยังกระบวนการสร้างประโยคสืบค้น SQL (Form SQL Query Process) คีย์เวิร์คใดที่ไม่ตรงตามค่าที่ระบุไว้ในโครงสร้างฐานข้อมูลจะถูกส่งต่อไปยังกระบวนการจับคู่ Entity ของฐานข้อมูล (Match Database Entity Process) หากขั้นตอนในการตรวจจับ Entity ของฐานข้อมูลมีปัญหาจะส่งข้อผิดพลาดไปรายงานความผิดพลาดให้กับผู้ใช้งานทราบ
- 4) กระบวนการสร้างประโยคสืบค้น SQL (Form SQL Query Process) เป็นกระบวนการที่จะสร้างประโยคสืบค้น SQL จาก Entity ของฐานข้อมูลและรูปแบบประโยคสืบค้นเพื่อนำไปดึงคำตอบออกมาจากฐานข้อมูลหลักของระบบเมื่อได้คำตอบจะส่งข้อมูลคำตอบให้ผู้ใช้งานทราบ หากขั้นตอนในการสร้างประโยคสืบค้น SQL มีปัญหาจะส่งข้อผิดพลาดไปรายงานความผิดพลาดให้กับผู้ใช้งานทราบ
- 5) กระบวนการรวบรวมคำตอบ (Gather Answer Process) เป็นกระบวนการที่จะนำเอาประโยคสืบค้นที่สร้างได้ส่งไปยังระบบจัดการฐานข้อมูลของฐานข้อมูล NorthWind เพื่อสืบค้นระเบียบข้อมูลที่เป็นคำตอบ เมื่อได้รับคำตอบจากระบบจัดการฐานข้อมูลแล้วจะแสดงผลให้กับผู้ใช้งานได้ทราบ หากขั้นตอนในการจับคู่ Entity ของฐานข้อมูล มีปัญหาจะส่งข้อผิดพลาดไปรายงานความผิดพลาดให้กับผู้ใช้งานทราบ

3.2.2 แผนภาพแสดงลำดับการทำงานของกระบวนการต่างๆ

แผนภาพแสดงลำดับการทำงานของกระบวนการหลักนี้ จะเป็นการนำเอากระบวนการหลักที่แสดงไว้ในแผนภาพแสดงลำดับการทำงานของกระบวนการ โดยรวมรูปที่ 3.2 มาแจกแจงลำดับการทำงานที่กระบวนการดังกล่าวต่อไปนี้

1) กระบวนการแปลงคำถามให้อยู่ในรูปแบบอย่างง่าย (Simplify Question Process)



รูปที่ 3.3 แผนภาพแสดงลำดับการทำงานของกระบวนการแปลงคำถามให้อยู่ในรูปแบบอย่างง่าย

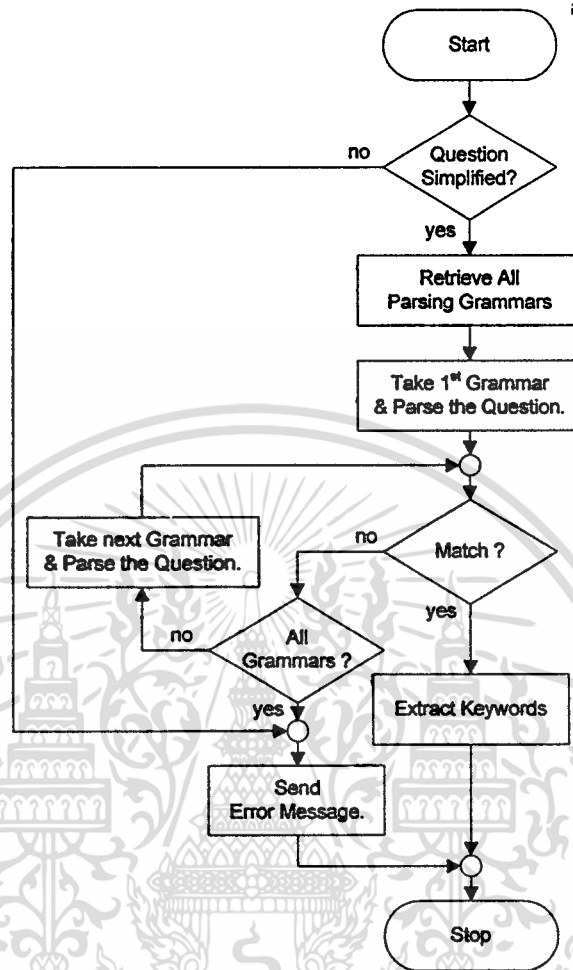
จากรูปที่ 3.3 แสดงแผนภาพแสดงลำดับการทำงานของกระบวนการแปลงคำถามให้อยู่ในรูปแบบอย่างง่ายซึ่งมีลำดับการทำงานดังนี้

แปลงตัวอักษรต่างๆ ในคำถามที่ได้รับให้เป็นตัวพิมพ์เล็กทั้งหมด

1. แยกคำต่างๆ ที่มีออกจากกันเพื่อกำจัด white space character ต่างๆ ที่มี
2. เรียบเรียงประโยคใหม่ โดยให้มีช่องว่างระหว่างคำเท่ากับ 1 white space character
3. หากมีข้อผิดพลาดจะสร้างข้อความรายงานความผิดพลาด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) กระบวนการแยกคีย์เวิร์ด (Extract Keywords Process)



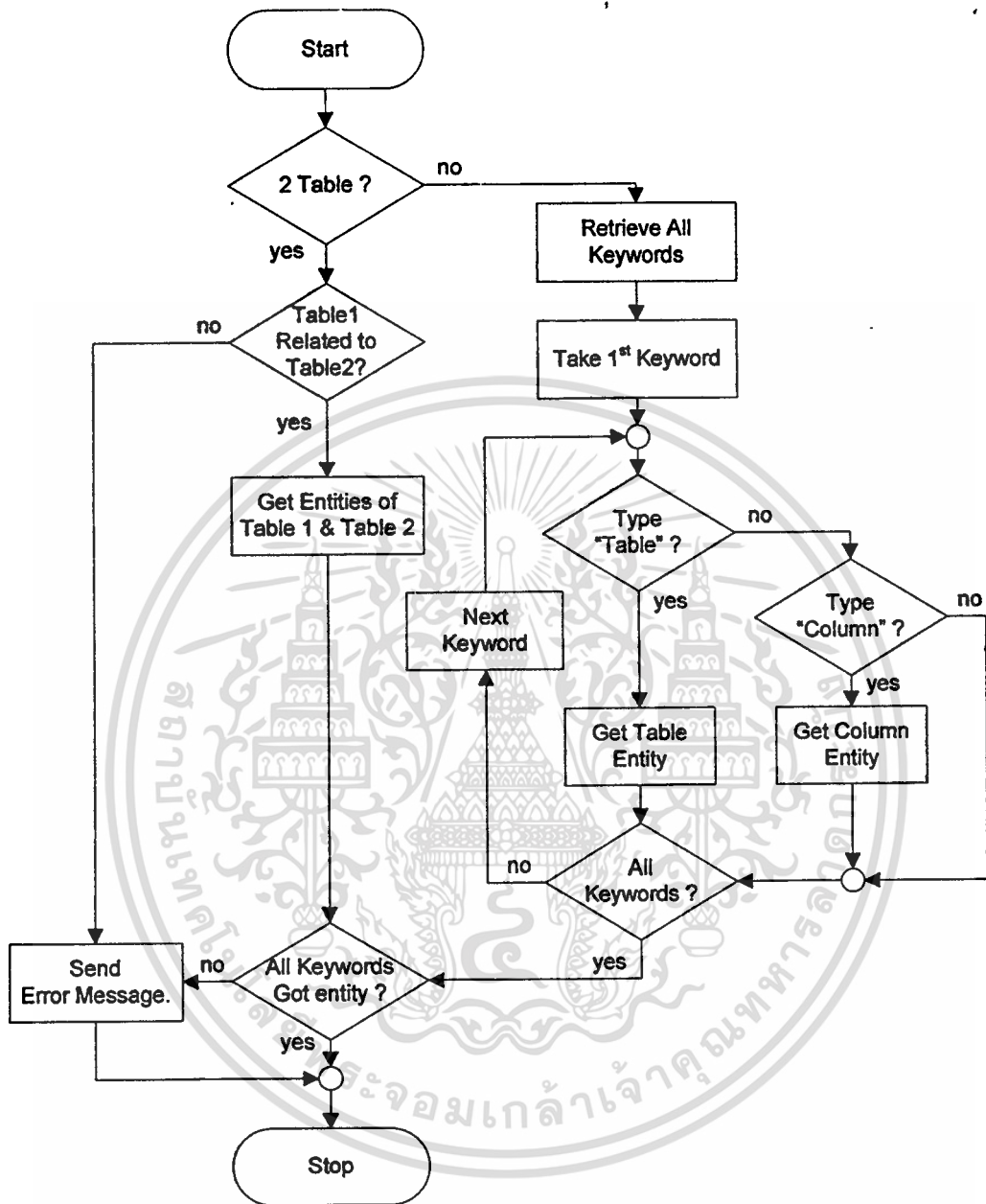
รูปที่ 3.4 แผนภาพแสดงลำดับการทำงานของกระบวนการแยกคีย์เวิร์ด

จากรูปที่ 3.4 แสดงแผนภาพแสดงลำดับการทำงานของกระบวนการแยกคีย์เวิร์ดซึ่งมีลำดับการทำงานดังนี้

1. หากคำถามไม่อยู่ในรูปร่างง่ายให้สร้างข้อความแสดงความผิดพลาดแล้วหยุดการทำงาน หากอยู่ในรูปร่างง่ายแล้วให้ทำงานต่อไป
2. ดึงข้อมูลไวยากรณ์ที่ใช้สำหรับแจกแจงคำ
3. ทำการตรวจสอบโดยการแจกแจงคำถาม โดยตรวจทีละไวยากรณ์ หากไม่สามารถตรวจพบไวยากรณ์ที่สามารถใช้ในการแจกแจงคำถามได้ให้แสดงข้อความรายงานความผิดพลาด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) กระบวนการตรวจจับ Entity ของฐานข้อมูล (Detect DB Entity Process)



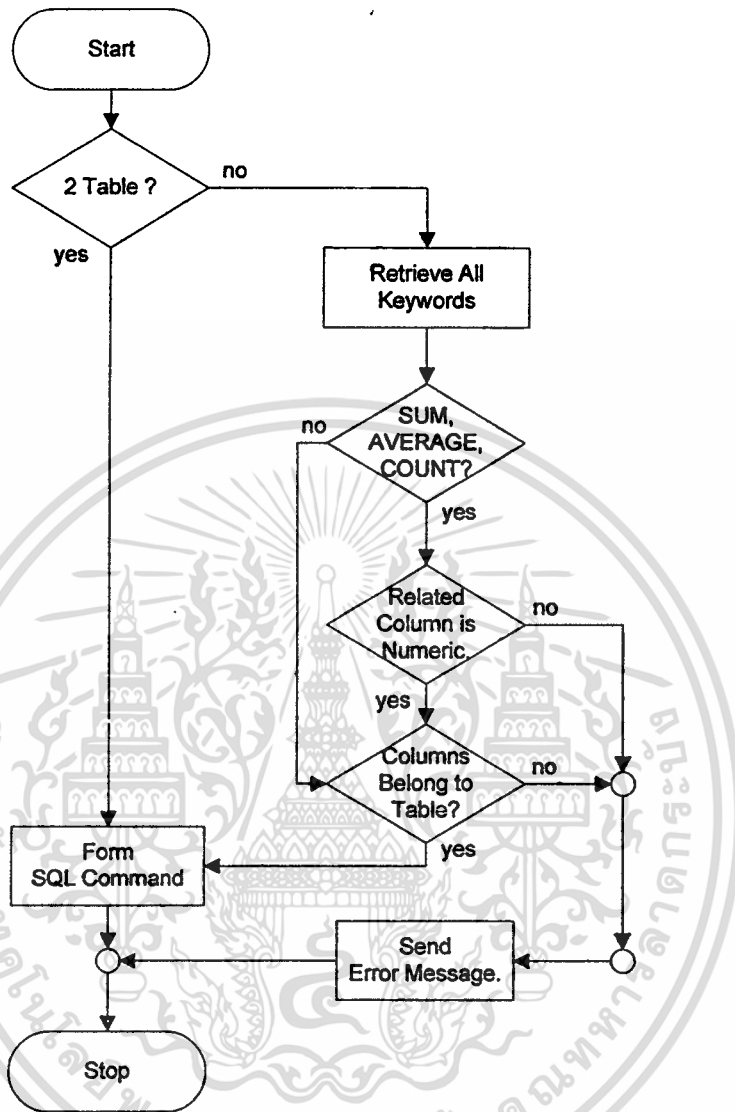
รูปที่ 3.5 แผนภาพแสดงลำดับการทำงานของกระบวนการตรวจจับ Entity ของฐานข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.5 แสดงแผนภาพแสดงลำดับการทำงานของกระบวนการแยกคีย์เวิร์คซึ่งมีลำดับการทำงานดังนี้

1. หากพบคีย์เวิร์คชนิด table2 ตัว ให้ทำการตรวจสอบความสัมพันธ์ของทั้งสองตารางหากทั้งสองตารางไม่มีความสัมพันธ์กันให้สร้างข้อความแสดงความผิดพลาดแล้วหยุดการทำงาน หากพบความสัมพันธ์ของทั้งสองตารางให้ตรวจหา Entity ชนิด table จากข้อมูลโครงสร้างฐานข้อมูล แล้วนำมาเพิ่มเป็นข้อมูลในคีย์เวิร์คนั้นๆ
2. หากพบคีย์เวิร์คชนิด table ตัวเดียวให้ทำการตรวจหา Entity ของคีย์เวิร์คแต่ละชนิดจากข้อมูล โครงสร้างฐานข้อมูล แล้วนำมาเพิ่มเป็นข้อมูลในคีย์เวิร์คนั้นๆ จนกระทั่งครบทุกๆคีย์เวิร์ค
3. เมื่อทำการประมวลผลครบทุกๆคีย์เวิร์คแล้วให้ทำการตรวจดูว่าคีย์เวิร์คแต่ละตัวมีข้อมูลในส่วนของ Database Entity หรือไม่ หากมีตัวใดตัวหนึ่งหรือทั้งหมดไม่มีข้อมูลของ DB Entity ให้สร้างข้อความแสดงความผิดพลาดแล้วหยุดการทำงาน หากทุกๆคีย์เวิร์คพบข้อมูล Database Entity แล้วไม่ต้องรายงานความผิดพลาด

4) กระบวนการสร้างประโยคสืบค้น SQL (Form SQL Query Process)



รูปที่ 3.6 แผนภาพแสดงลำดับการทำงานของกระบวนการสร้างประโยคสืบค้น SQL

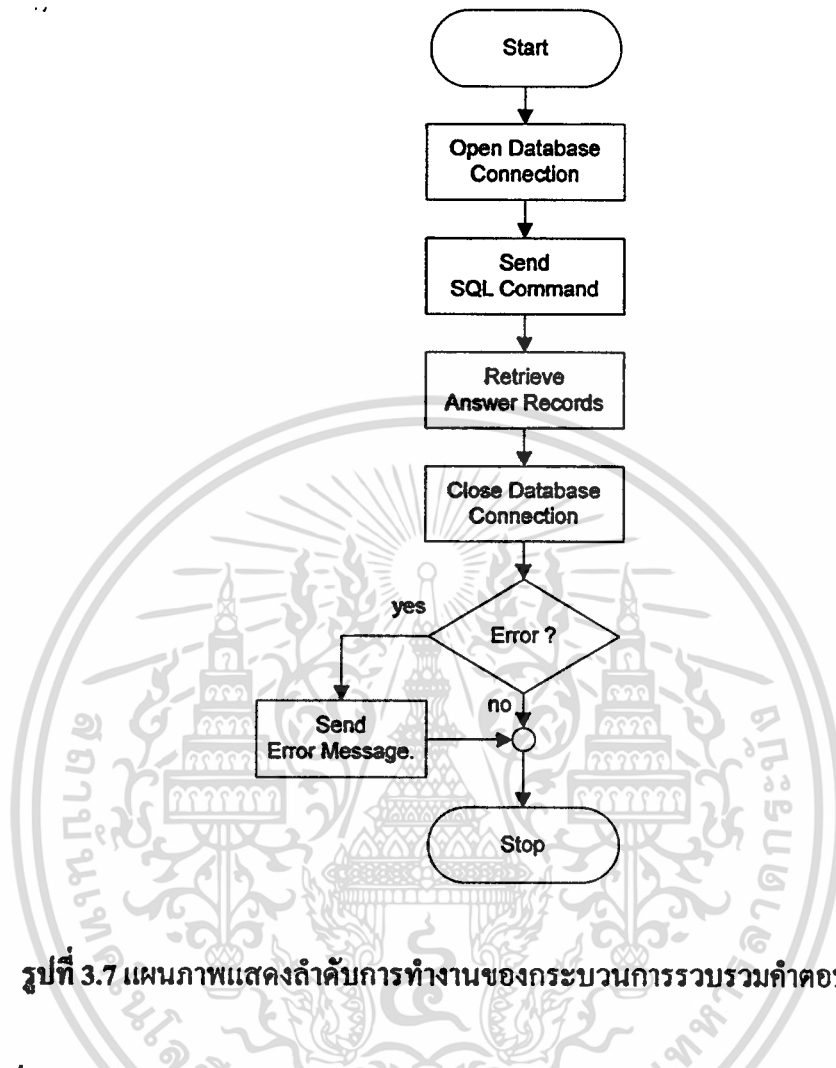
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.6 แสดงแผนภาพแสดงลำดับการทำงานของกระบวนการสร้างประโยคสืบค้น SQL ซึ่งมีลำดับการทำงานดังนี้

1. หากพบคีย์เวิร์ดชนิด table2 ตัวให้ทำการสร้างประโยคสืบค้นจากข้อมูลคีย์เวิร์ดที่ได้
2. หากชนิดของคำสั่ง SQL เป็นชนิด COUNT ,SUM หรือ AVERAGE ให้ทำการตรวจสอบว่าคีย์เวิร์ดที่ใช้มีชนิดข้อมูลที่สามารถประมวลผลด้วยคำสั่งดังกล่าวได้หรือไม่ หากไม่สามารถประมวลผลได้ให้สร้างข้อความแสดงความคิดเห็นแล้วหยุดการทำงาน
3. ตรวจสอบคีย์เวิร์ดชนิด Column ว่ามีความสัมพันธ์กับ คีย์เวิร์ดชนิด Table หรือไม่ หากมีตัวหนึ่งตัวใด ไม่สัมพันธ์กันให้สร้างข้อความแสดงความคิดเห็นแล้วหยุดการทำงาน
4. หากไม่พบความผิดพลาดใดๆให้ทำการสร้างประโยคสืบค้น SQL จากข้อมูลคีย์เวิร์ดที่ได้



5) กระบวนการรวบรวมคำตอบ (Gather Answer Process)



รูปที่ 3.7 แผนภาพแสดงลำดับการทำงานของกระบวนการรวบรวมคำตอบ

จากรูปที่ 3.7 แสดงแผนภาพแสดงลำดับการทำงานของกระบวนการรวบรวมคำตอบซึ่งมีลำดับการทำงานดังนี้

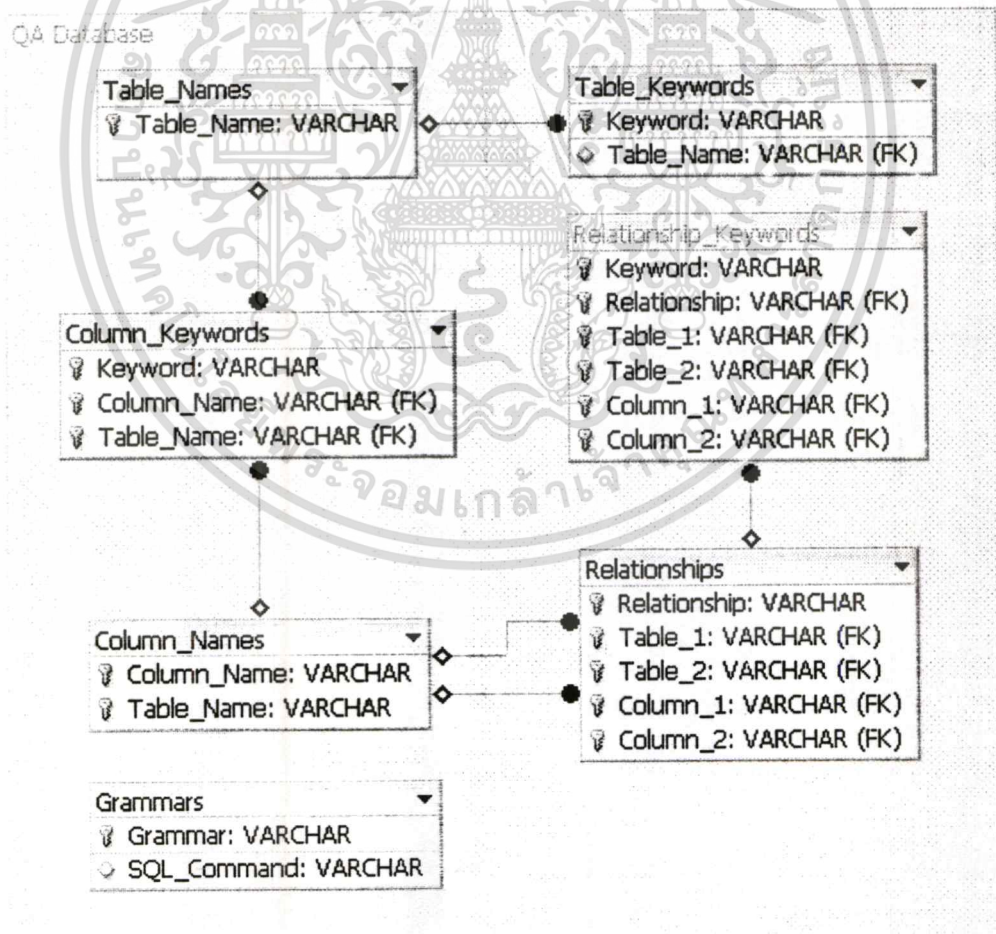
- 1 สร้างการเชื่อมต่อพร้อมทั้งเปิดการเชื่อมต่อเข้ากับฐานข้อมูล NorthWind
- 2 ส่งคำสั่ง SQL ไปยังฐานข้อมูล NorthWind
- 3 รวบรวมคำตอบที่ได้ในรูปแบบระเบียบฐานข้อมูล
- 4 ปิดการเชื่อมต่อกับฐานข้อมูล

3.3 การออกแบบฐานข้อมูลระบบถามตอบ

ระบบถามตอบนี้จะอาศัยข้อมูลต่างๆในการประมวลผลของระบบซึ่งถูกเก็บไว้ในรูปแบบฐานข้อมูล โดยฐานข้อมูลที่ใช้ในระบบถามตอบที่ใช้งานกับโปรแกรมประยุกต์บนเว็บนี้มีด้วยกัน 2 ฐานข้อมูลคือ ฐานข้อมูลของระบบถามตอบ ซึ่งจะเก็บข้อมูลต่างๆที่ใช้สำหรับการประมวลผล และฐานข้อมูล NorthWind เป็นฐานข้อมูลตัวอย่าง ที่มาพร้อมกับระบบจัดการฐานข้อมูลไมโครซอฟต์ SQL Server 2000 ซึ่งระบบถามตอบใช้เป็นฐานข้อมูลเป้าหมาย

3.3.1 ฐานข้อมูลของระบบถามตอบ

ฐานข้อมูลของระบบถามตอบนี้ถูกสร้างขึ้นบนระบบจัดการฐานข้อมูลเชิงสัมพันธ์ของไมโครซอฟต์ SQL Server 2000 โดยฐานข้อมูลนี้จะเก็บข้อมูลต่างๆที่จำเป็นสำหรับการประมวลผลของระบบถามตอบ เช่น ข้อมูลโครงสร้างของฐานข้อมูลเป้าหมาย, คีย์เวิร์ดต่างๆ, และไวยากรณ์ที่ใช้ในการประมวลผลภาษาธรรมชาติ ฐานข้อมูลของระบบถามตอบมีโครงสร้างดังรูปที่ 3.8



รูปที่ 3.8 แผนภาพแสดงความสัมพันธ์ระหว่าง Entity ของฐานข้อมูลระบบถามตอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) ตาราง Table_Names

เป็นตารางที่ใช้เก็บชื่อของตารางต่างๆที่มีในฐานข้อมูล Northwind มีคอลัมน์ต่างๆดังต่อไปนี้

- Table_Name เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของตาราง

2) ตาราง Table_Keywords

เป็นตารางที่ใช้เก็บข้อมูลคีย์เวิร์ดชนิด Table ของระบบถามตอบมีคอลัมน์ต่างๆดังต่อไปนี้

- Keyword เป็นคอลัมน์ที่ใช้สำหรับเก็บคำที่ใช้เป็นคีย์เวิร์ด
- Table_Name เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของตาราง

3) ตาราง Column_Names

เป็นตารางที่ใช้เก็บชื่อของคอลัมน์ต่างๆที่มีในฐานข้อมูล Northwind มีคอลัมน์ต่างๆดังต่อไปนี้

- Column_Name เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของตาราง
- Table_Name เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของตาราง

4) ตาราง Column_Keywords

เป็นตารางที่ใช้เก็บข้อมูลคีย์เวิร์ดชนิด Column มีคอลัมน์ต่างๆดังต่อไปนี้

- Keyword เป็นคอลัมน์ที่ใช้สำหรับเก็บคำที่ใช้เป็นคีย์เวิร์ด
- Column_Name เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของตาราง
- Table_Name เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของตาราง

5) ตาราง Relationships

เป็นตารางที่ใช้เก็บความสัมพันธ์ต่างๆที่มีในฐานข้อมูล Northwind มีคอลัมน์ต่างๆดังต่อไปนี้

- Relationship เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของความสัมพันธ์
- Table_1 เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของตาราง
- Column_1 เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของตาราง
- Table_2 เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของตาราง
- Column_2 เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของตาราง

6) ตาราง Relationship_Keywords

เป็นตารางที่ใช้เก็บข้อมูลคีย์เวิร์ดชนิด Relationship มีคอลัมน์ต่างๆดังต่อไปนี้

- Keyword เป็นคอลัมน์ที่ใช้สำหรับเก็บค่าที่ใช้เป็นคีย์เวิร์ด
- Relationship เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของความสัมพันธ์
- Table_1 เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของตาราง
- Column_1 เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของตาราง
- Table_2 เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของตาราง
- Column_2 เป็นคอลัมน์ที่ใช้สำหรับเก็บชื่อของตาราง

7) ตาราง Grammars

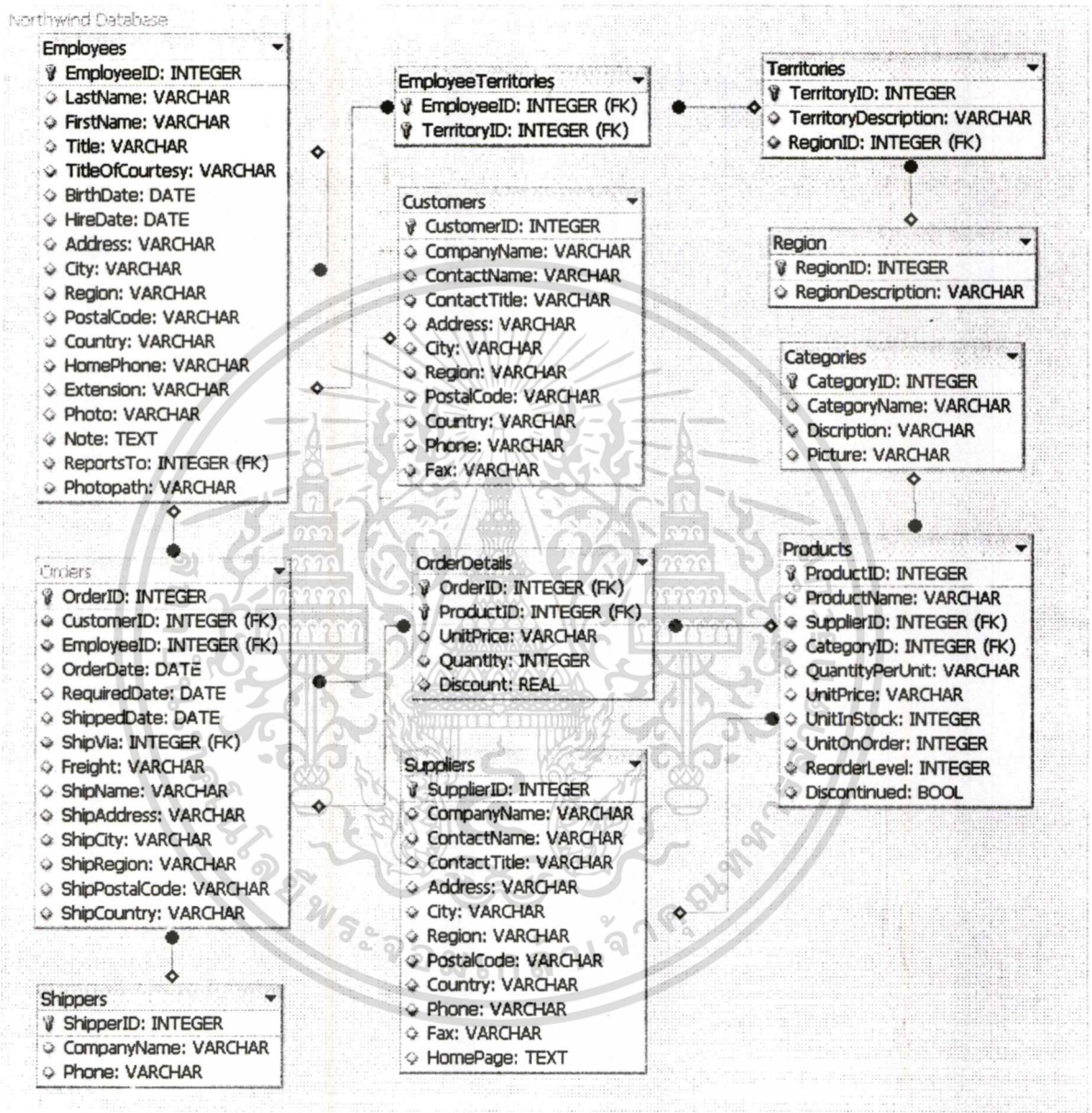
เป็นตารางที่ใช้เก็บข้อมูลไวยากรณ์ที่ใช้สำหรับการประมวลผลภาษาธรรมชาติของระบบถามตอบซึ่งมีคอลัมน์ต่างๆดังต่อไปนี้

- Grammar เป็นคอลัมน์ที่ใช้สำหรับเก็บคำไวยากรณ์ที่ใช้ในการแจกแจงคำ
- SQL_Command เป็นคอลัมน์ที่ใช้สำหรับเก็บประเภทของคำสั่ง SQL ที่ใช้ในการสร้างประโยคสืบค้น



3.3.2 ฐานข้อมูล NorthWind

ฐานข้อมูล Northwind เป็นฐานข้อมูลเป้าหมายที่ใช้สำหรับการสืบค้นข้อมูลของระบบถามตอบ มีโครงสร้างดังรูปที่ 3.9



รูปที่ 3.9 แผนภาพแสดงความสัมพันธ์ระหว่าง Entity ของฐานข้อมูล NorthWind

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

การพัฒนาและการทดสอบระบบงาน

ในบทนี้จะกล่าวถึงการพัฒนาเว็บที่ใช้งานกับ โปรแกรมประยุกต์บนเว็บ ซึ่งจะอธิบายถึง สภาพแวดล้อมต่างๆที่ใช้ในการพัฒนาระบบ เช่นระบบปฏิบัติการ ภาษาที่ใช้ในการเขียนโปรแกรมของระบบ และเครื่องมือต่างๆที่ใช้ในการพัฒนาระบบ ส่วนการทดสอบการทำงานของระบบจะได้ทำการทดสอบการใช้งานระบบในเงื่อนไขต่างๆเพื่อจำลองการใช้งานจริงของผู้ใช้งาน

4.1 การพัฒนาระบบตามตอบที่ใช้งานกับโปรแกรมประยุกต์บนเว็บ

ระบบตามตอบที่ใช้งานกับ โปรแกรมประยุกต์บนเว็บนี้อาศัยสภาพแวดล้อมในการพัฒนาระบบต่างๆดังต่อไปนี้

4.1.1 ฮาร์ดแวร์

เครื่องคอมพิวเตอร์ที่ใช้ในการพัฒนาและทดสอบระบบ โดยมีคุณสมบัติดังนี้

- Notebook IBM R40 : Centrino 1.3
- RAM : 512 MB
- Hard Disk : 40 GB
- ระบบปฏิบัติการ Windows XP Professional

4.1.2 ซอฟต์แวร์

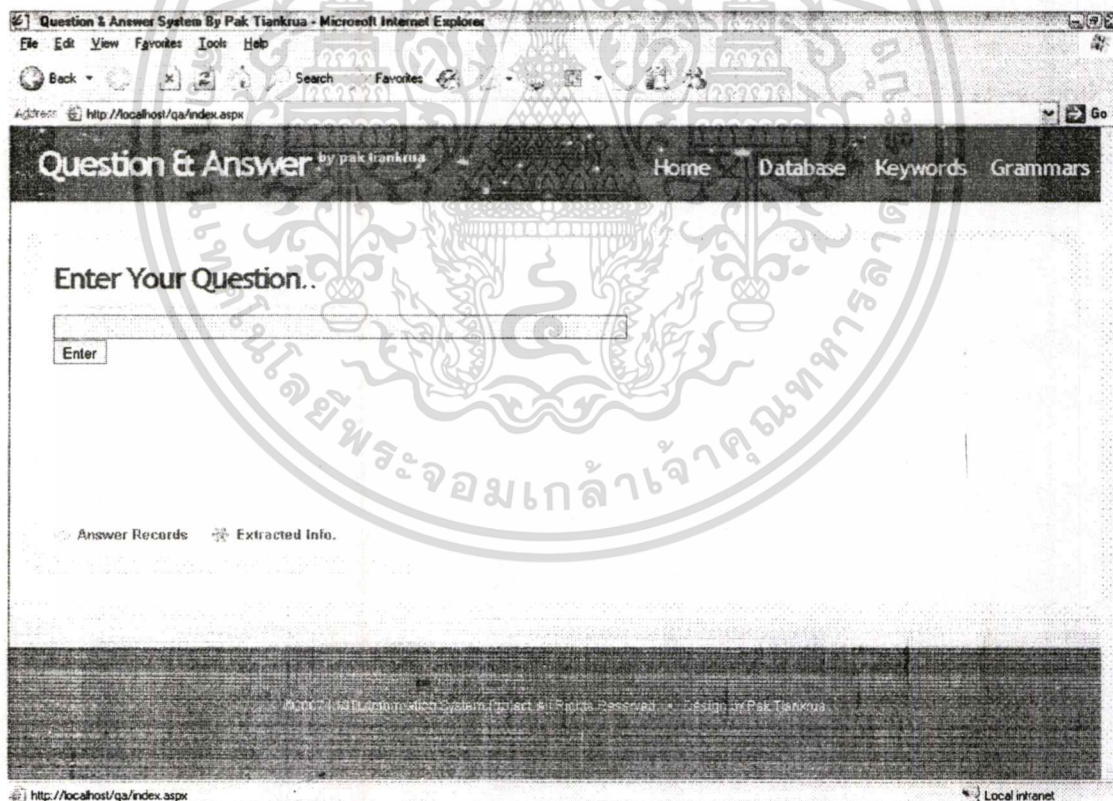
ซอฟต์แวร์ที่ใช้การพัฒนาและทดสอบระบบ มีดังนี้

- Microsoft Visual Web Developer 2005
- Microsoft SQL Server 2000
- Expresso
- Microsoft Office
- Microsoft Internet Explorer version 6.0
- Microsoft Internet Information System (IIS)

4.2 ส่วนติดต่อผู้ใช้งาน

ส่วนติดต่อผู้ใช้งานระบบถามตอบที่ใช้งานกับโปรแกรมประยุกต์บนเว็บ จะเป็นลักษณะเว็บเพจ ซึ่งประกอบด้วยหน้าหลัก 4 หน้า ดังนี้

- 1) หน้าหลัก (Home page) เป็นหน้าที่เป็นหน้าแรกเมื่อเริ่มการทำงานของระบบมีชื่อไฟล์ว่า index.aspx ซึ่งทำงานร่วมกับไฟล์โค้ด index.aspx.cs จะมีกล่องสำหรับป้อนคำถาม และมีเมนูสำหรับการเรียกหน้าอื่นๆของระบบอยู่ด้านบนดังแสดงในรูปที่ 4.1 และมีฟังก์ชันการทำงานดังนี้
 - ป้อนคำถาม
 - แสดงผลลัพธ์ที่ได้จากการสืบค้นฐานข้อมูล
 - แสดงข้อผิดพลาดต่างๆที่เกิดขึ้นในระบบ
 - แสดงข้อมูลต่างๆที่ได้จากการประมวลผลของระบบ

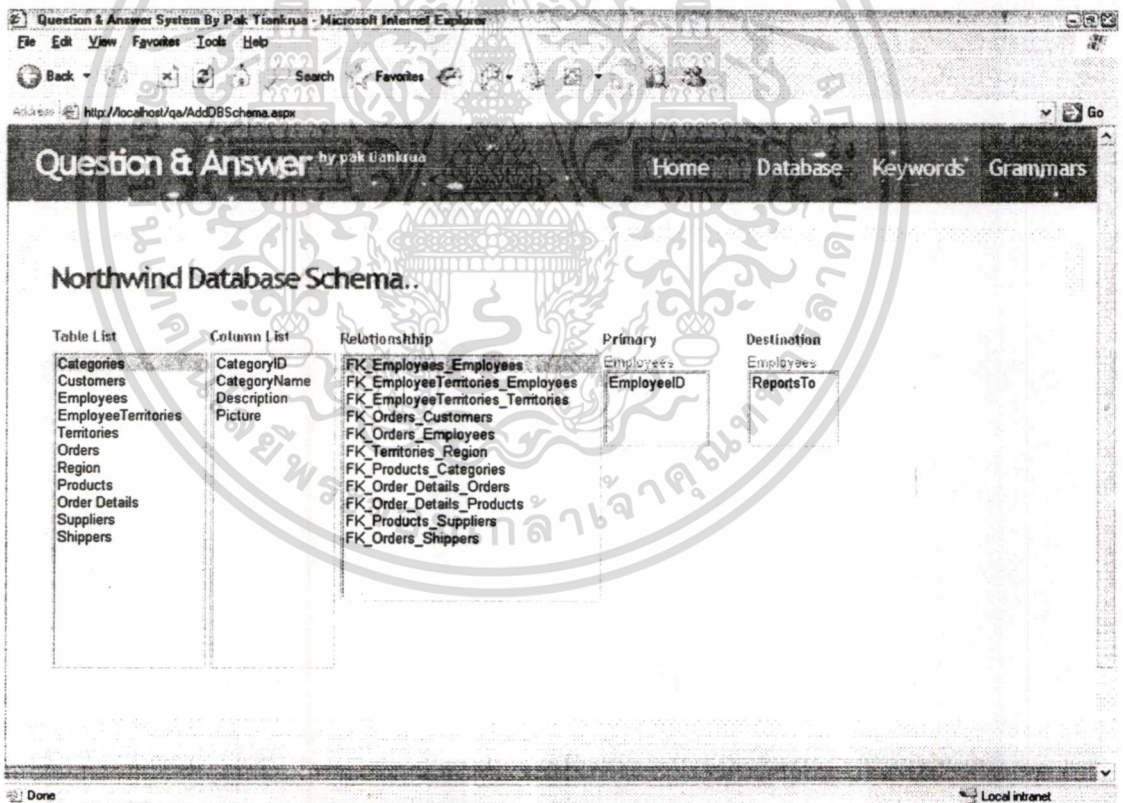


รูปที่ 4.1 ส่วนติดต่อผู้ใช้งานหน้าหลัก (Home page)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) หน้าโครงสร้างฐานข้อมูล (Database page) เป็นหน้าที่แสดงโครงสร้างของฐานข้อมูล NorthWind มีชื่อไฟล์ว่า AddDBSchema.aspx ซึ่งทำงานร่วมกับไฟล์โค้ด AddDBSchema.aspx.cs จะมีกล่องข้อความต่างๆ กล่องแรกด้านซ้ายมีรายชื่อของตารางต่างๆที่มีอยู่ในฐานข้อมูล NorthWind กล่องถัดมาจะแสดงรายชื่อของคอลัมน์ต่างๆของตารางด้านซ้ายกล่องที่ 4 แสดงรายการความสัมพันธ์ต่างๆของตารางในฐานข้อมูล NorthWind เมื่อเลือกรายการใดในกล่องข้อความนี้จะปรากฏรายชื่อของตารางที่มีความสัมพันธ์กันของความสัมพันธ์ที่เลือกในกล่องข้อความถัดไปด้านขวามือ และมีเมนูสำหรับการเรียกหน้าอื่นๆของระบบอยู่ด้านบนดังแสดงในรูปที่ 4.2 และมีฟังก์ชันการทำงานดังนี้

- แสดงรายการตารางต่างๆในฐานข้อมูล NorthWind
- แสดงรายการคอลัมน์ของตารางต่างๆในฐานข้อมูล NorthWind
- แสดงรายการความสัมพันธ์ของตารางต่างๆในฐานข้อมูล NorthWind

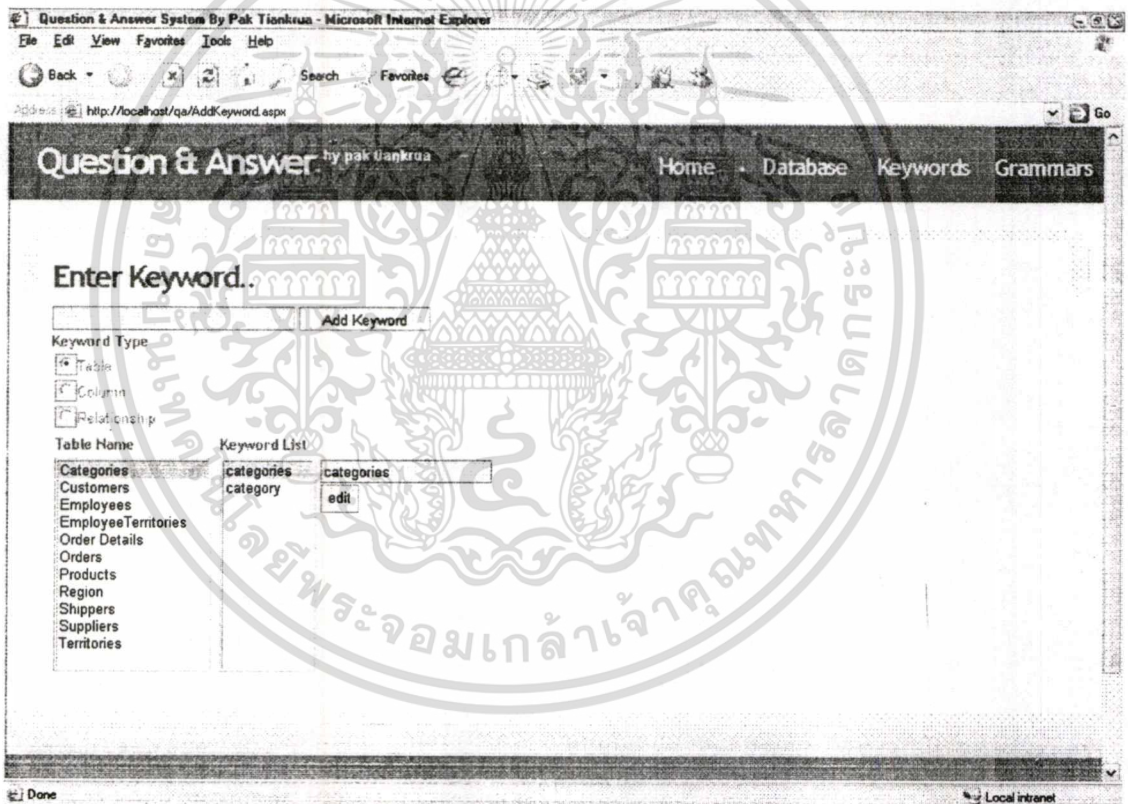


รูปที่ 4.2 ส่วนติดต่อผู้ใช้งานหน้าฐานข้อมูล (Database page)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) **หน้าคีย์เวิร์ด (Keywords page)** เป็นหน้าที่แสดงข้อมูลของคีย์เวิร์ดต่างๆที่ใช้ในการประมวลผลภาษาธรรมชาติแบบการแจกแจงของระบบถามตอบมีชื่อไฟล์ว่า AddKeyword.aspx ซึ่งทำงานร่วมกับไฟล์โค้ด AddKeyword.aspx.cs จะมีกล่องข้อความสำหรับเพิ่มคีย์เวิร์ดต่างๆเข้าไปในระบบถามตอบ รายการให้เลือกแสดงคีย์เวิร์ดประเภทต่างๆ กล่องข้อความต่างๆด้านล่างแสดงรายการคีย์เวิร์ดต่างๆที่มีอยู่ในระบบถามตอบ และมีเมนูสำหรับการเรียกหน้าอื่นๆของระบบอยู่ด้านบนดังแสดงในรูปที่ 4.3 ซึ่งมีฟังก์ชันการทำงานดังนี้

- แสดงรายการคีย์เวิร์ดต่างๆที่มีในฐานข้อมูลระบบถามตอบ
- เพิ่มคีย์เวิร์ดเข้าไปในระบบถามตอบ
- แก้ไขคีย์เวิร์ดต่างๆที่มีในระบบถามตอบ

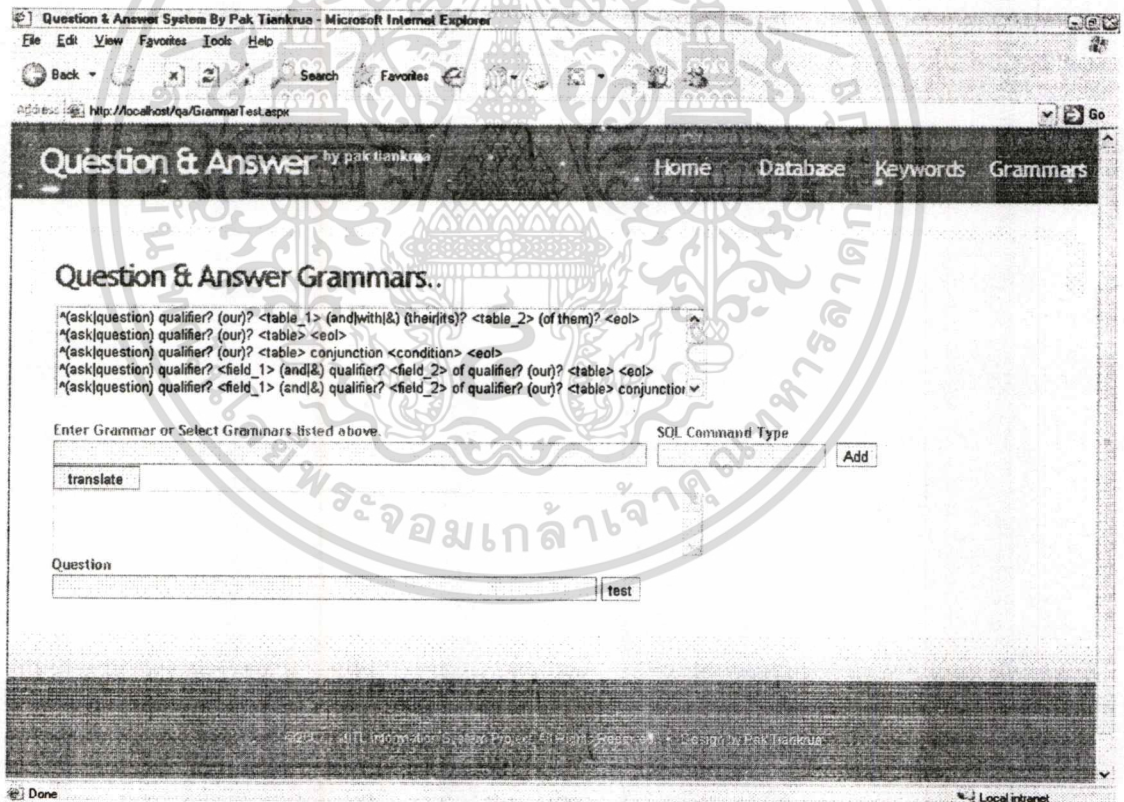


รูปที่ 4.3 ส่วนติดต่อผู้ใช้งานหน้าคีย์เวิร์ด (Keyword page)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4) หน้าไวยากรณ์ (Grammar page) เป็นหน้าที่แสดงข้อมูลเกี่ยวกับไวยากรณ์ต่างๆที่ระบบถามตอบใช้ในการประมวลผลภาษาธรรมชาติแบบวิธีแจกแจงคำมีชื่อไฟล์ว่า GrammarTest.aspx ซึ่งทำงานร่วมกับไฟล์โค้ด GrammarTest.aspx.cs จะมีกล่องข้อความแสดงรายการของไวยากรณ์ที่เก็บไว้ในฐานข้อมูล กล่องข้อความสำหรับเพิ่มไวยากรณ์เข้าไปในระบบหรือแก้ไขไวยากรณ์ที่มีอยู่เดิม กล่องข้อความด้านล่างสำหรับป้อนคำถามเพื่อทดสอบไวยากรณ์ และมีเมนูสำหรับการเรียกหน้าอื่นๆของระบบอยู่ด้านบนดังแสดงในรูปที่ 4.4 ซึ่งมีฟังก์ชันการทำงานดังนี้

- เพิ่มไวยากรณ์ในการแจกแจงคำถามเข้าไปในระบบ
- แก้ไขไวยากรณ์ที่มีอยู่
- ทดสอบการทำงานของไวยากรณ์
- แสดงรายการไวยากรณ์ต่างๆที่มีอยู่ในระบบ
- แสดงการแปลงไวยากรณ์ให้อยู่ในรูปของนิพจน์ปรกติ (Regular Expression)



รูปที่ 4.4 ส่วนติดต่อผู้ใช้งานหน้าไวยากรณ์ (Grammar page)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3) การสืบค้นจากตารางเดียวระบุสองคอลัมน์ไม่ระบุเงื่อนไข

- ป้อนคำถาม “Display me all the product id and product name of products.”
- กดปุ่ม Enter เพื่อประมวลผล
- ผลลัพธ์ออกมาดังรูปที่ 4.7

Question & Answer System By Pak Tiankrua - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Stop Search Favorites

Address: http://localhost/qa/index.aspx

Question & Answer by pak.tiankrua Home Database Keywords Grammars

Enter Your Question..

Display me all the product id and product name of products.

Enter

Your Answer Records..

Product ID	Product Name
07	Milk - 300ml
08	Apple - 1kg
09	Banana - 1kg
10	Orange - 1kg
11	Watermelon - 1kg
12	Tomato - 1kg
13	Carrot - 1kg
14	Broccoli - 1kg
15	Spinach - 1kg
16	Chicken - 1kg
17	Beef - 1kg
18	Pork - 1kg
19	Egg - 1kg
20	Rice - 1kg
21	Instant Noodle - 1kg
22	Instant Soup - 1kg
23	Instant Tea - 1kg
24	Instant Coffee - 1kg
25	Instant Milk - 1kg
26	Instant Fruit - 1kg
27	Instant Vegetable - 1kg
28	Instant Meat - 1kg
29	Instant Fish - 1kg
30	Instant Seafood - 1kg
31	Instant Dessert - 1kg

Done Local intranet

รูปที่ 4.7 ผลจากการทดสอบ 1.3)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.4) การสืบค้นจากตารางเดียวโดยระบุเงื่อนไข

- ป้อนคำถาม “Who are employees who first name start with c?”
- กดปุ่ม Enter เพื่อประมวลผล
- ผลลัพธ์ออกมาดังรูปที่ 4.8

The screenshot shows a web browser window titled "Question & Answer System By Pak Tiankrua - Microsoft Internet Explorer". The address bar shows "http://localhost/qa/index.aspx". The page content includes a search form with the question "Who are employees who first name start with a?" and an "Enter" button. Below the form, the "Your Answer Records.." section displays a table with the following data:

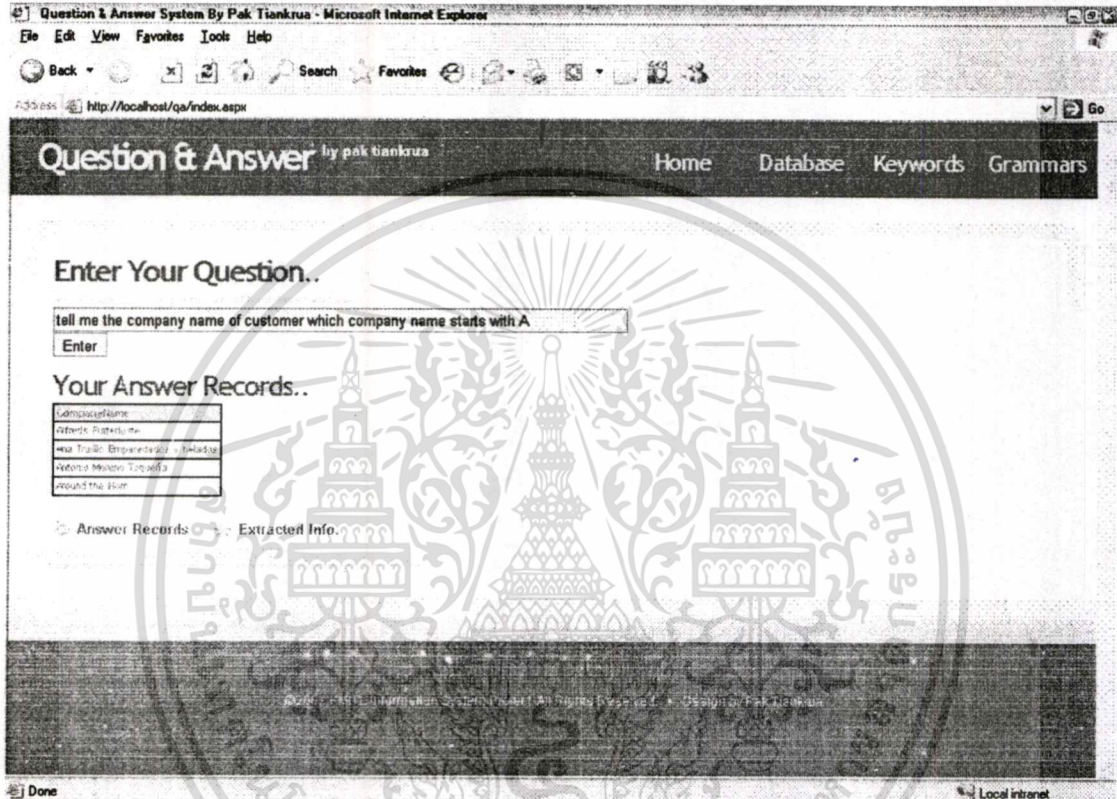
Employee ID	Last Name	First Name	Title	Title of Course	Birth Date	Age	Address	City	Region	Postal Code	Country	Phone Number	Extension	Photo	Notes	Reports To	Photo
1	Fisher	Andrew	Chief Executive Officer	Ph.D.	29/04/1952	52	17 001 10th Ave	Tampa	FL	33601	USA	(813) 555-1142	5467	Andrew.Fisher.jpg	Andrew was not and but DT2 commenced in 1974 and a Ph.D. in International Management from the University of Colorado in 1981. He is fluent in French and Italian and reads German. He joined the company as a sales manager in January 1992 and has been promoted to Chief Executive Officer on 15 Aug 2001.		Emp101

รูปที่ 4.8 ผลจากการทดสอบ 1.4)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.5) การสืบค้นจากตารางเดียวระบุหนึ่งคอลัมน์โดยระบุเงื่อนไข

- ป้อนคำถาม “Tell me the company name of customer which company name start with A?”
- กดปุ่ม Enter เพื่อประมวลผล
- ผลลัพธ์ออกมาดังรูปที่ 4.9

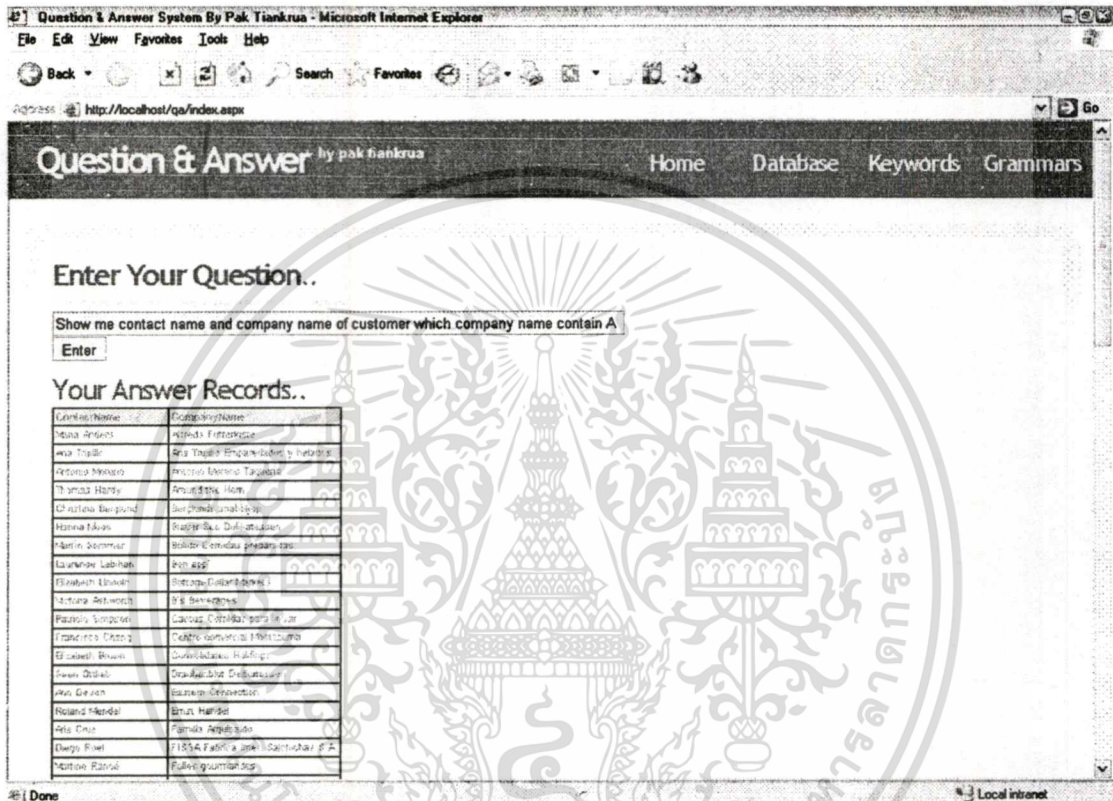


รูปที่ 4.9 ผลจากการทดสอบ 1.5)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.6) การสืบค้นจากตารางเดียวระบุสองคอลัมน์ โดยระบุเงื่อนไข

- ป้อนคำถาม “Show me contact name and company name of customer which company name contain A”
- กดปุ่ม Enter เพื่อประมวลผล
- ผลลัพธ์ออกมาดังรูปที่ 4.10



Question & Answer System By Pak Tiangkrua - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Search Favorites

Address http://localhost/qa/index.aspx

Question & Answer by pak tiangkrua Home Database Keywords Grammars

Enter Your Question..

Show me contact name and company name of customer which company name contain A

Enter

Your Answer Records..

Contact Name	Company Name
Maria Anders	Alfreds Futterkiste
Anna Berglund	Bra Turistbyrå
Antonio Moreno	Antonio Moreno Taquería
Tina Turner	Turner & Sons
Charlotte Bacard	Bacard
Hanna Moen	Richter Supermarketer
Martín Sommer	Comercial del Norte
Elizabeth Labish	Labish
Elizabeth Linde	Linde
Victoria DeBorja	DeBorja
Pablo Pichon	Pichon
Francisco Chang	Chang
Elizabeth Brown	Brown
Samuel Niles	Niles
Andrés Deven	Deven
Roland Mendel	Mendel
Arti Cruz	Cruz
Diego Roel	Roel
Martín Ramírez	Ramírez

Done Local intranet

รูปที่ 4.10 ผลจากการทดสอบ 1.6)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.7) การสืบค้นจากสองตารางไม่ระบุเงื่อนไข

- ป้อนคำถาม “Display to me all the suppliers and their products.”
- กดปุ่ม Enter เพื่อประมวลผล
- ผลลัพธ์ออกมาดังรูปที่ 4.11

Question & Answer System By Pak Tiankrua - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address http://localhost/qa/index.aspx

Question & Answer by pak.tiankrua Home Database Keywords Grammars

Enter Your Question..

Display to me all the suppliers and their products.

Enter

Your Answer Records..

SupplierID	GoodsName	GoodsName	ContactURL	Package	City	Region	PostalCode	Country	Phone	Fax	HomePage
1	Besse Liquid	Charlotte Cooper	Purchasing Manager	40 Gildart St	London		EC1A 4EP	UK	1 21 555 2222		
1	Evric Liquid	Charlotte Cooper	Purchasing Manager	40 Gildart St	London		EC1A 4EP	UK	1 21 555 2222		
1	Evric Liquid	Charlotte Cooper	Purchasing Manager	40 Gildart St	London		EC1A 4EP	UK	1 21 555 2222		
2	New Orleans Cabin Light	Shirley Barks	Order Administrator	P.O. Box 70209	New Orleans, LA		70117	USA	1 504 555 4222		PC.COM/HTML
2	New Orleans Cabin Light	Shirley Barks	Order Administrator	P.O. Box 70209	New Orleans, LA		70117	USA	1 504 555 4222		PC.COM/HTML
2	Granola Light Household	Regina Murphy	Order Administrator	707 Orleans Rd	New York, NY		40134	USA	2 12 555 5725	1 212 555 2000	
2	Granola Light Household	Regina Murphy	Order Administrator	707 Orleans Rd	New York, NY		40134	USA	2 12 555 5725	1 212 555 2000	
3	Granola Light Household	Regina Murphy	Order Administrator	707 Orleans Rd	New York, NY		40134	USA	2 12 555 5725	1 212 555 2000	
3	Tea's Traces	Scott Wagner	Marketing Manager	101 National	Atlanta, GA		303	USA	404 355 4011		
3	Tea's Traces	Scott Wagner	Marketing Manager	101 National	Atlanta, GA		303	USA	404 355 4011		
3	Tea's Traces	Scott Wagner	Marketing Manager	101 National	Atlanta, GA		303	USA	404 355 4011		

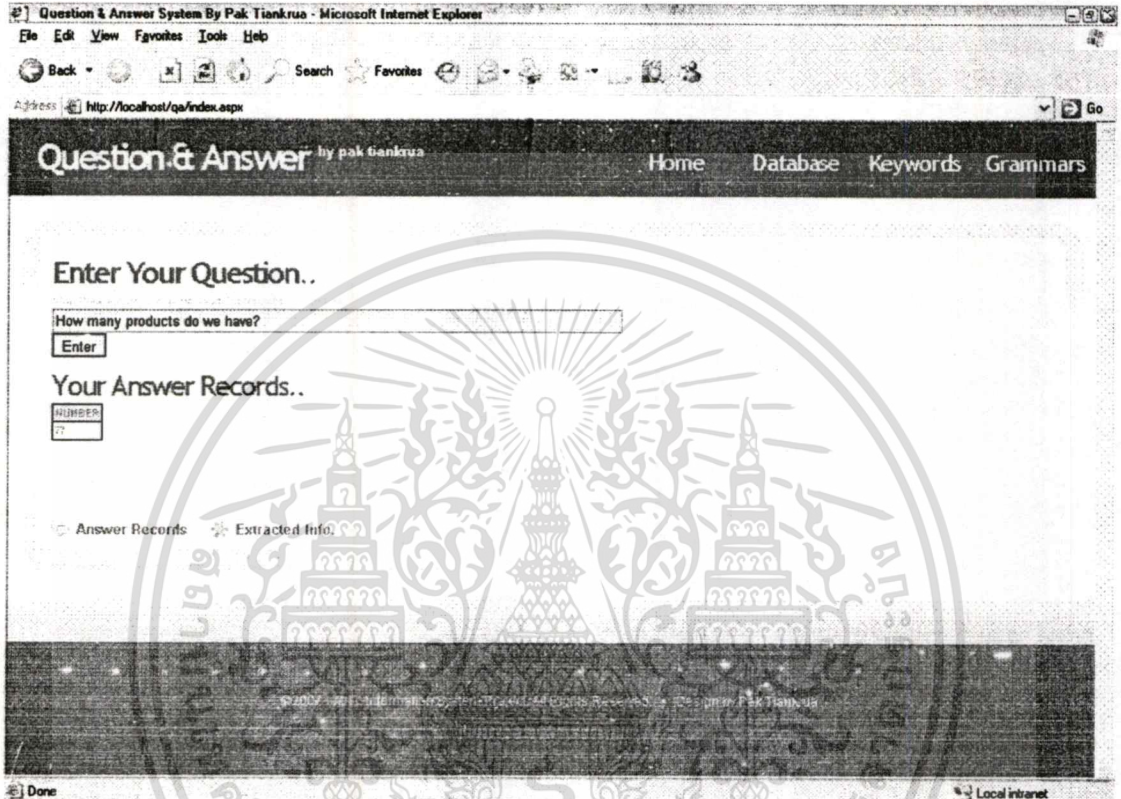
Done Local intranet

รูปที่ 4.11 ผลจากการทดสอบ 1.7)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.8) การนับจำนวนไม่ระบุเงื่อนไข

- ป้อนคำถาม “How many products do we have?”
- กดปุ่ม Enter เพื่อประมวลผล
- ผลลัพธ์ออกมาดังรูปที่ 4.12

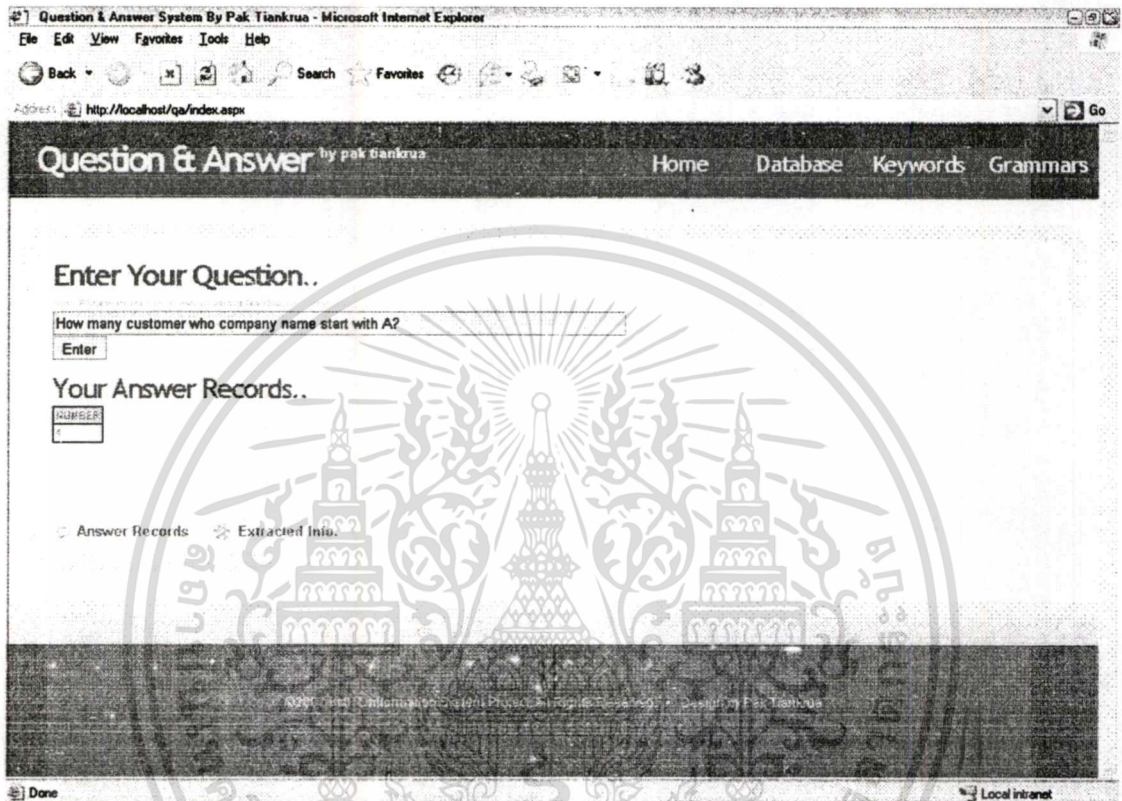


รูปที่ 4.12 ผลจากการทดสอบ 1.8)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.9) การนับแบบระบุเงื่อนไข

- ป้อนคำถาม “How many customer who company name start with A?”
- กดปุ่ม Enter เพื่อประมวลผล
- ผลลัพธ์ออกมาดังรูปที่ 4.13



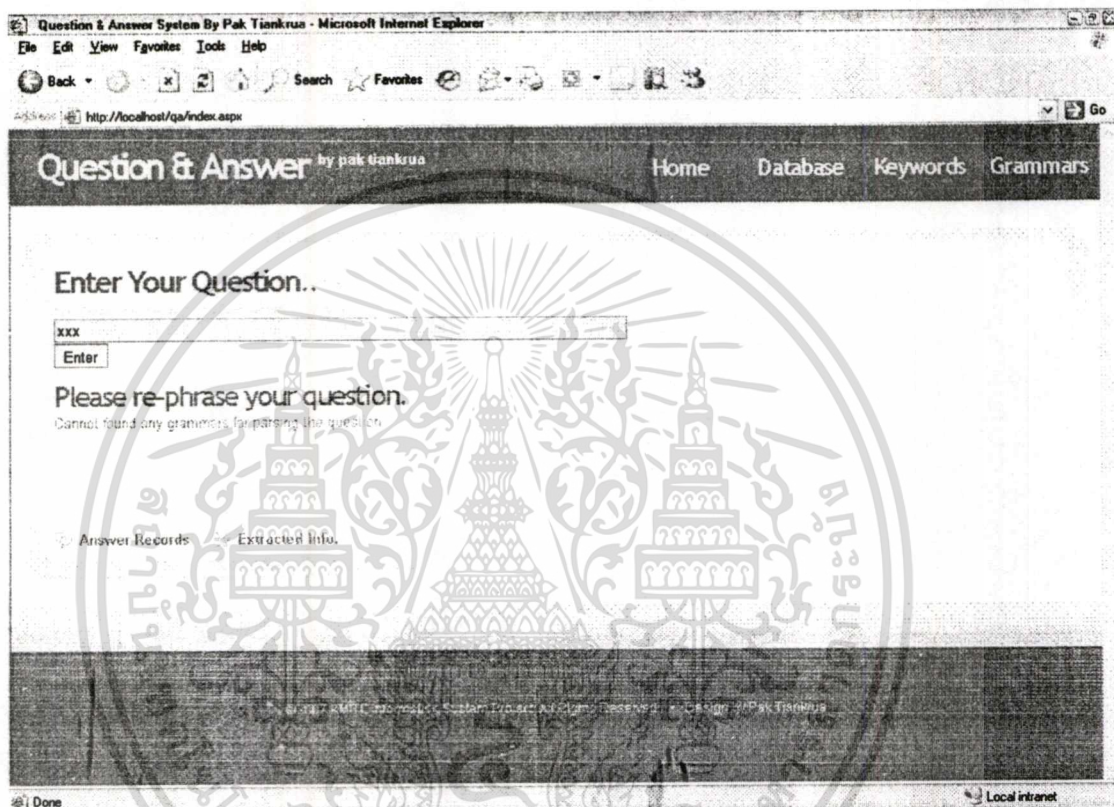
รูปที่ 4.13 ผลจากการทดสอบ 1.9)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) ทดสอบการรายงานความผิดพลาด

2.1) เมื่อไม่สามารถประมวลผลคำถามได้

- ป้อนคำถาม “xxx”
- กดปุ่ม Enter เพื่อประมวลผล
- ผลลัพธ์ออกมาดังรูปที่ 4.14

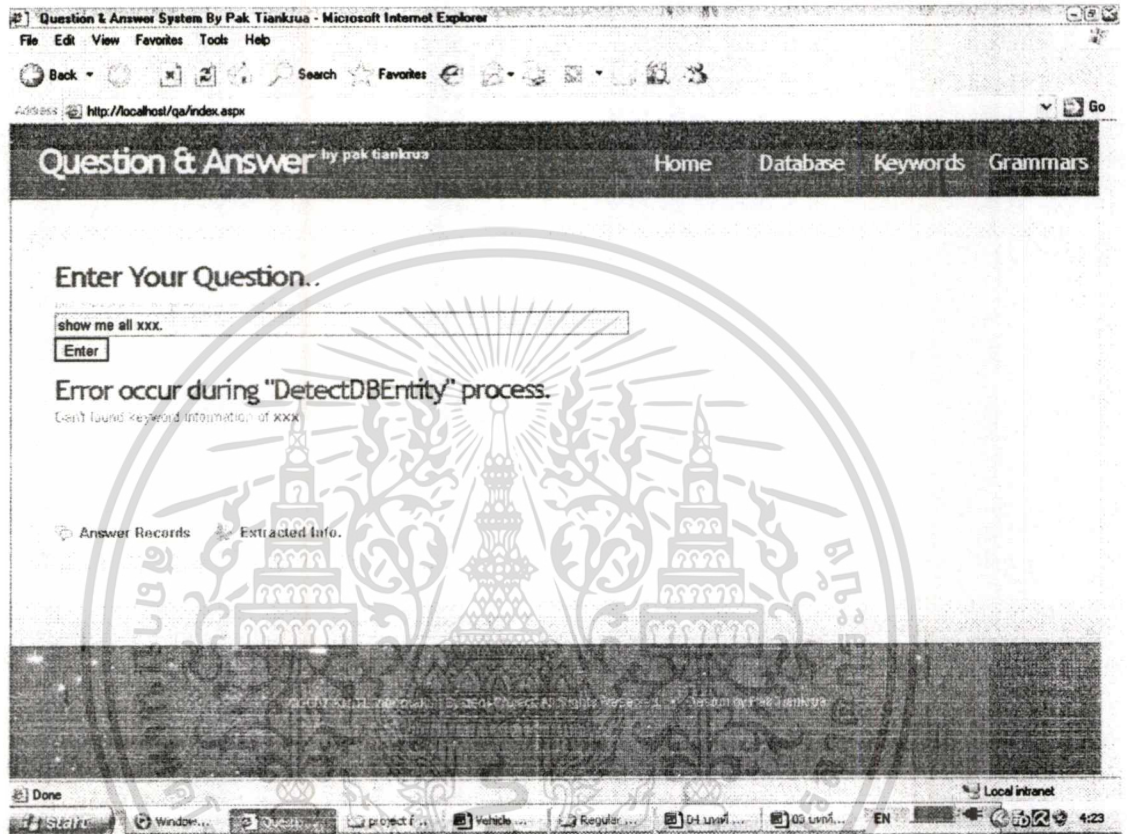


รูปที่ 4.14 ผลจากการทดสอบ 2.1)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2) เมื่อไม่สามารถพบข้อมูลศัพท์เว็บ

- ป้อนคำถาม “show me all xxx”
- กดปุ่ม Enter เพื่อประมวลผล
- ผลลัพธ์ออกมาดังรูปที่ 4.15

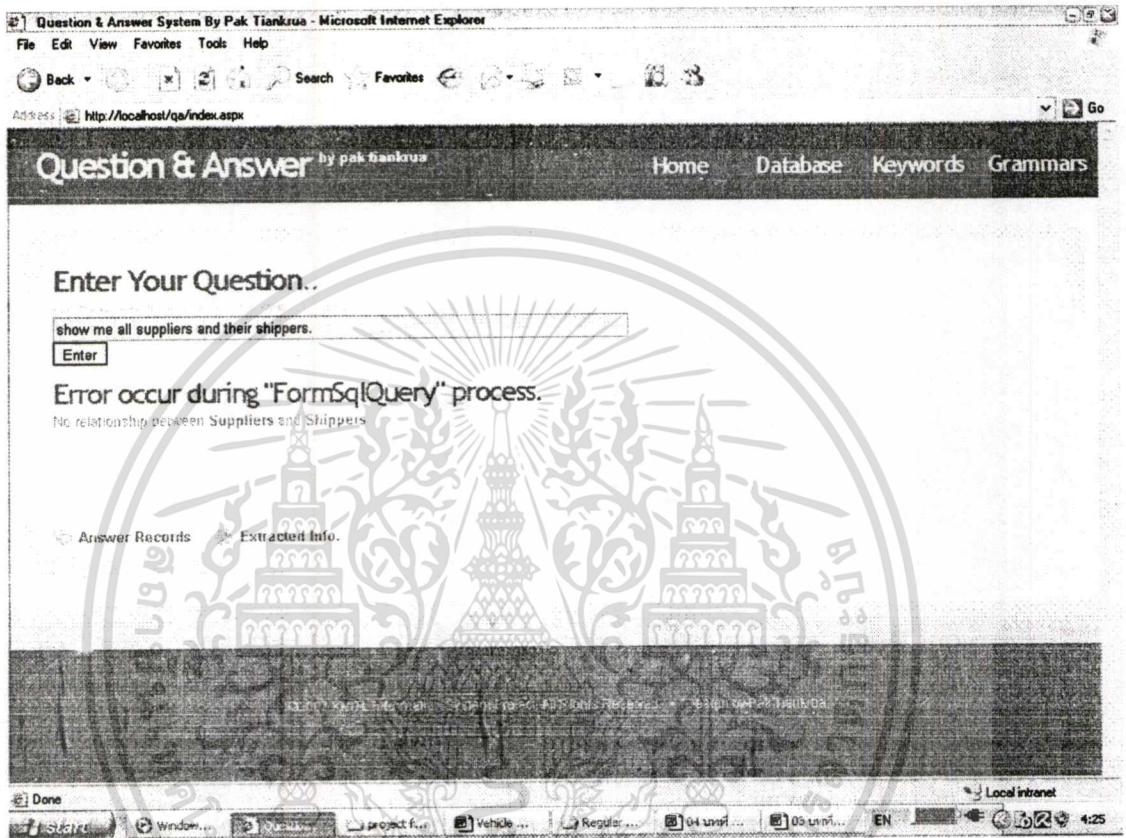


รูปที่ 4.15 ผลจากการทดสอบ 2.2)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3) เมื่อตารางทั้งสองไม่มีความสัมพันธ์กัน

- ป้อนคำถาม “show me all suppliers and their shippers.”
- กดปุ่ม Enter เพื่อประมวลผล
- ผลลัพธ์ออกมาดังรูปที่ 4.16

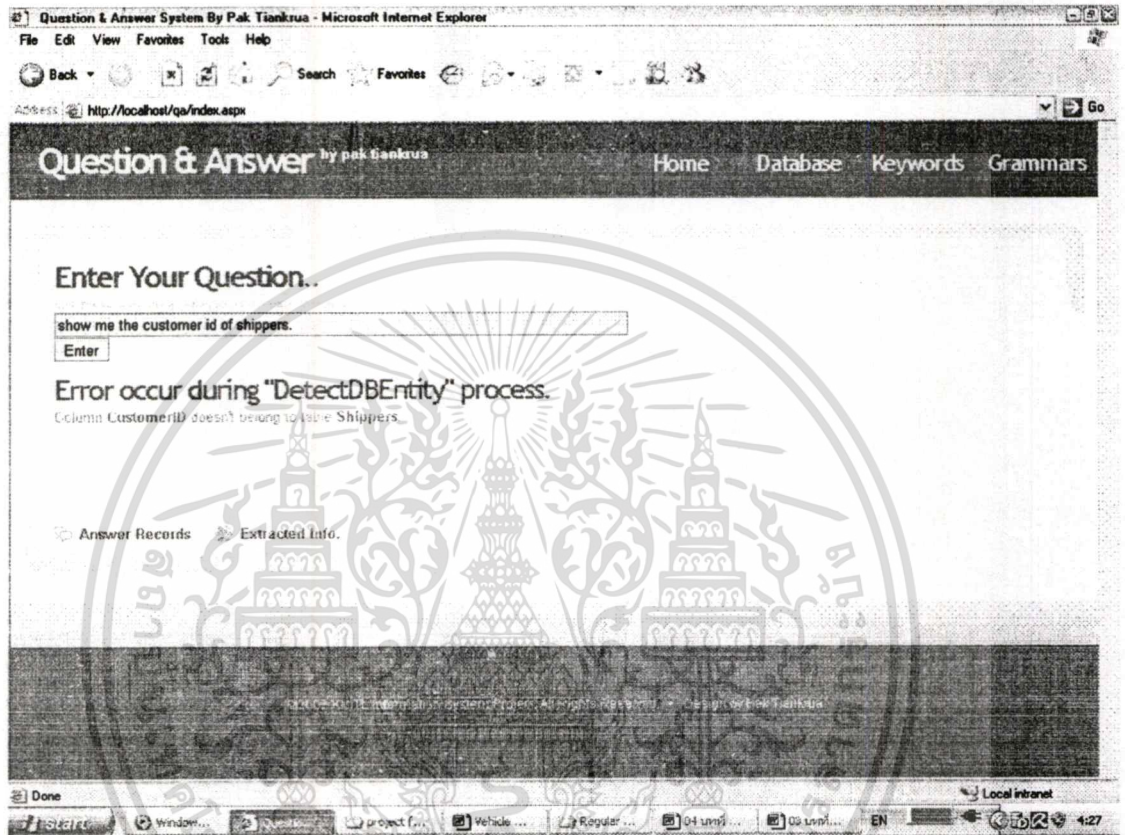


รูปที่ 4.16 ผลจากการทดสอบ 2.3)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4) เมื่อคอลัมน์และตาราง ไม่มีความสัมพันธ์กัน

- ป้อนคำถาม “Show me the customer id of shippers.”
- กดปุ่ม Enter เพื่อประมวลผล
- ผลลัพธ์ออกมาดังรูปที่ 4.17

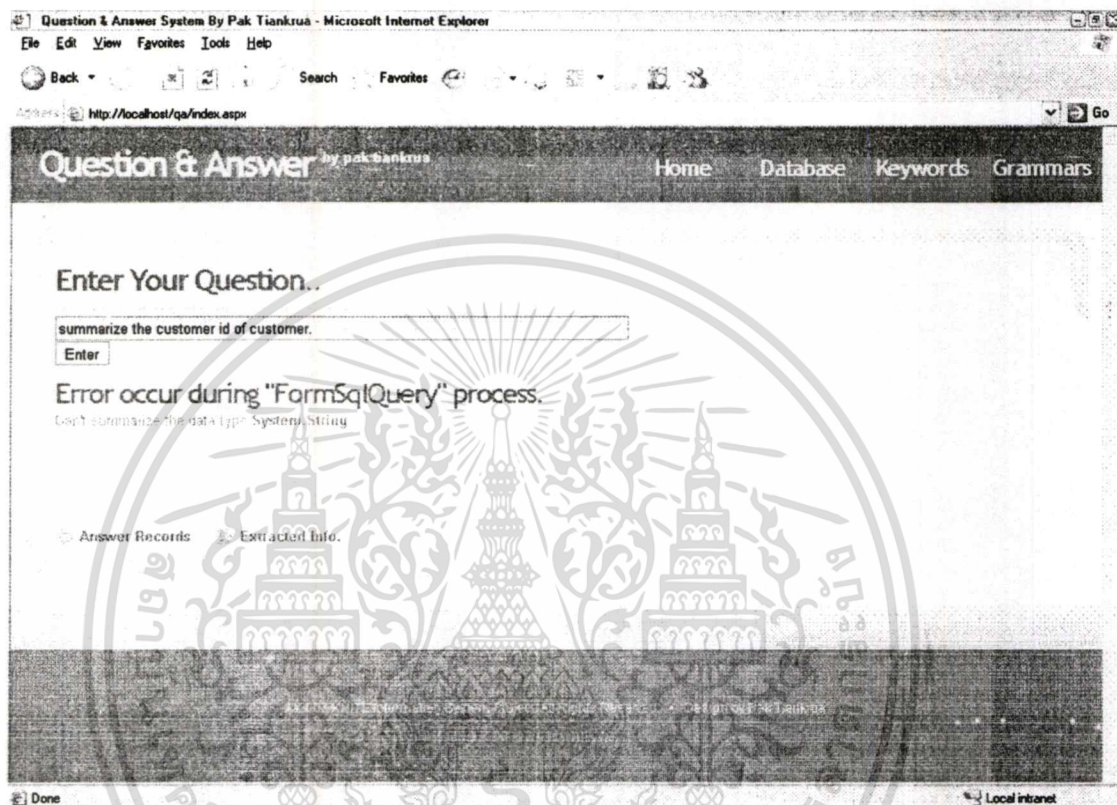


รูปที่ 4.17 ผลจากการทดสอบ 2.4)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5) เมื่อไม่สามารถนับได้

- ป้อนคำถาม “Summarize the customer id of customer.”
- กดปุ่ม Enter เพื่อประมวลผล
- ผลลัพธ์ออกมาดังรูปที่ 4.18



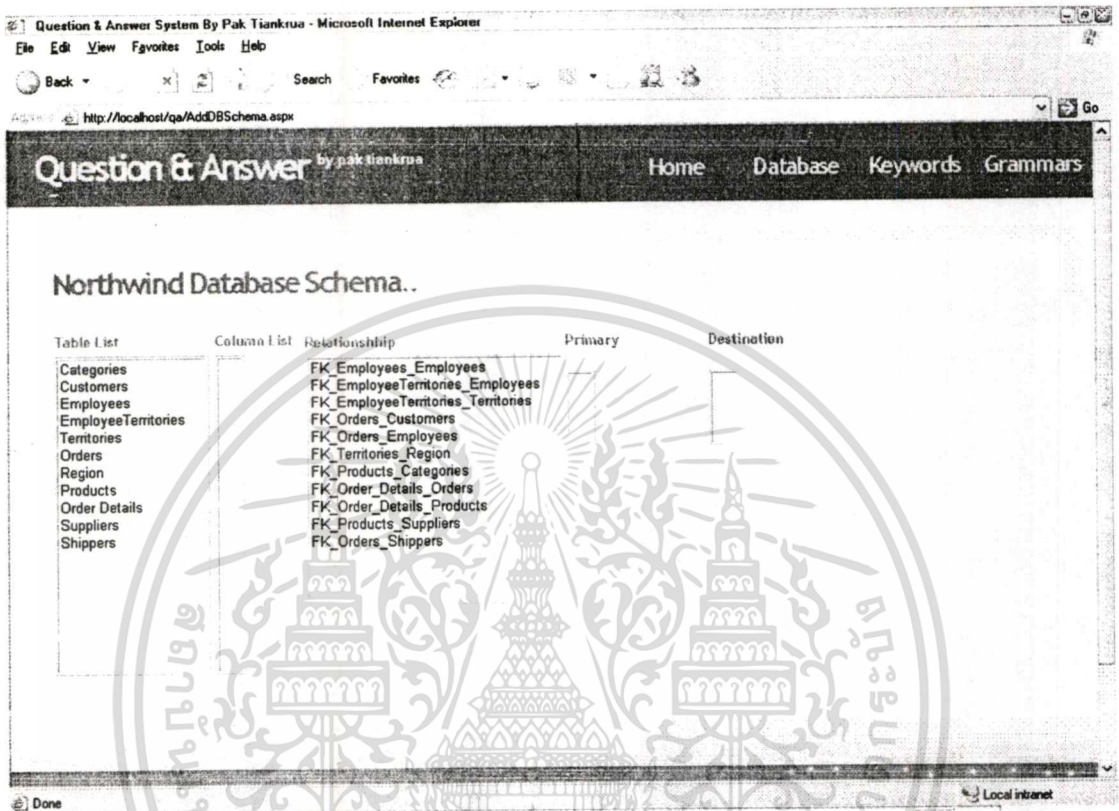
รูปที่ 4.18 ผลจากการทดสอบ 2.5)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.2 การทดสอบการทำงานหน้าฐานข้อมูล (Database Page)

1) การแสดงผลรายการตาราง

- แสดงผลตามรูปที่ 4.19



รูปที่ 4.19 ผลจากการทดสอบการแสดงผลรายการตาราง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) การแสดงผลรายการคอลัมน์

- เลือกรายชื่อตาราง Customers
- แสดงผลตามรูปที่ 4.20

Question & Answer System By Pak Tiankrua - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address http://localhost/qa/AddDBSchema.aspx

Question & Answer by pak.tiankrua Home Database Keywords Grammars

Northwind Database Schema..

Table List	Column List	Relationship	Primary	Destination
Categories	CustomerID	FK_Employees_Employees		
Customers	CompanyName	FK_EmployeeTerritories_Employees		
Employees	ContactName	FK_EmployeeTerritories_Territories		
EmployeeTerritories	ContactTitle	FK_Orders_Customers		
Territories	Address	FK_Orders_Employees		
Orders	City	FK_Territories_Region		
Region	Region	FK_Products_Categories		
Products	PostalCode	FK_Order_Details_Orders		
Order Details	Country	FK_Order_Details_Products		
Suppliers	Phone	FK_Products_Suppliers		
Shippers	Fax	FK_Orders_Shippers		

Done Local internet

รูปที่ 4.20 ผลจากการทดสอบการแสดงผลรายการคอลัมน์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) การแสดงผลรายการความสัมพันธ์

- เลือกรายชื่อความสัมพันธ์ FK_Employees_Employees
- แสดงผลตามรูปที่ 4.21

Question & Answer System By Pak Tiankrua - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Go

http://localhost/qa/AddDBSchema.aspx

Question & Answer by pak tiankrua Home Database Keywords Grammars

Northwind Database Schema..

Table List	Column List	Relationship	Primary	Destination
Categories	CustomerID	FK_Employees_Employees	Employees	Employees
Customers	CompanyName	FK_EmployeesTerritories_Employees	EmployeeID	ReportsTo
Employees	ContactName	FK_EmployeeTerritories_Territories		
EmployeeTerritories	ContactTitle	FK_Orders_Customers		
Territories	Address	FK_Orders_Employees		
Orders	City	FK_Territories_Region		
Region	Region	FK_Products_Categories		
Products	PostalCode	FK_Order_Details_Orders		
Order Details	Country	FK_Order_Details_Products		
Suppliers	Phone	FK_Products_Suppliers		
Shippers	Fax	FK_Orders_Shippers		

Done Local intranet

รูปที่ 4.21 ผลจากการทดสอบการแสดงผลรายการความสัมพันธ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

บทสรุป

5.1 สรุปผลการพัฒนาระบบงาน

การพัฒนาระบบถามตอบที่ใช้ร่วมกับ โปรแกรมประยุกต์บนเว็บ ได้พัฒนาขึ้นเพื่อเป็นเครื่องมือ เพื่อช่วยให้การเข้าถึงข้อมูลที่ต้องการบน โปรแกรมประยุกต์บนเว็บทำได้สะดวกขึ้น กล่าวคือ ผู้ใช้งานสามารถป้อนคำถามที่อยากรู้เกี่ยวกับระบบ โปรแกรมประยุกต์บนเว็บนั้น แล้วระบบถามตอบจะทำหน้าที่ในการแปลงประโยคคำถามที่ผู้ใช้เป็นคนป้อนเข้ามา ให้กลายเป็นภาษาสืบค้น SQL ซึ่งใช้สำหรับการสืบค้นข้อมูลต่างๆบนระบบฐานข้อมูลเชิงสัมพันธ์ โดยวิธีการประมวลผล ภาษาธรรมชาติแบบแจกแจงค่า โดยได้มีการระบุไวยากรณ์ต่างๆที่ใช้สำหรับการแจกแจงประโยคคำถามไว้ในฐานข้อมูล

หลังจากผู้พัฒนาได้พัฒนาระบบถามตอบแล้ว พบว่าระบบถามตอบสามารถช่วยให้ผู้ใช้งานที่ไม่มีความรู้เกี่ยวกับภาษาสืบค้น SQL สามารถสืบค้นข้อมูลต่างๆที่อยู่บนฐานข้อมูลได้อย่างไรก็ตาม แม้ว่าผู้พัฒนาสามารถพัฒนาระบบถามตอบเพื่อใช้งานกับ โปรแกรมประยุกต์บนเว็บขึ้นมาสำเร็จ แต่ระบบถามตอบนี้ยังไม่สามารถตอบสนองต่อภาษาธรรมชาติได้อย่างครอบคลุมทุกๆกรณี และยัง สามารถประมวลผลภาษาธรรมชาติที่เป็นเพียงภาษาอังกฤษเท่านั้น เนื่องจากไวยากรณ์ที่ใช้ในการประมวลผลภาษาธรรมชาติมีอย่างจำกัด แต่ระบบนี้สามารถขยายความสามารถในการประมวลผลได้ เนื่องจากผู้พัฒนาออกแบบให้ระบบสามารถเพิ่มเติมหรือแก้ไขไวยากรณ์ที่มีอยู่ได้

5.2 ประโยชน์ที่ได้รับจากการพัฒนาระบบงาน

ประโยชน์ที่ได้รับจากการพัฒนาระบบในโครงการนี้ สรุปได้ดังนี้

1. ได้นำความรู้ที่ได้ศึกษามาประยุกต์ใช้ในการวิเคราะห์ ออกแบบ และพัฒนาระบบงาน รวมถึงได้เรียนรู้ภาษาและเครื่องมือที่ใช้ในการพัฒนาระบบ เพื่อที่จะสามารถนำไปประยุกต์ใช้กับการพัฒนาระบบงานอื่นๆ ได้ในภายหลัง
2. ได้พัฒนาระบบงานที่นำไปช่วยปรับปรุงประสิทธิภาพการทำงานขององค์กร ทำให้สามารถใช้ทรัพยากรที่มีได้อย่างมีประสิทธิภาพ

5.3 ข้อเสนอแนะ

หลังจากผู้พัฒนาสามารถพัฒนาระบบถามตอบเพื่อใช้งานกับโปรแกรมประยุกต์บนเว็บขึ้นมาสำเร็จแล้ว ผู้พัฒนาเล็งเห็นว่าระบบถามตอบที่พัฒนาขึ้นมานั้นสามารถทำให้มีประโยชน์ในการนำไปใช้งานได้มากขึ้นโดยเพิ่มขีดความสามารถดังต่อไปนี้

1. เพิ่มความสามารถในการประมวลผลภาษาอื่นๆเช่น ภาษาไทย เพื่อให้สามารถตอบสนองต่อผู้ใช้งานที่เป็นคนไทยได้ดีขึ้น
2. พัฒนาเป็นโปรแกรมที่สามารถติดตั้งได้โดยง่าย เพื่อสามารถนำไปประยุกต์ใช้ได้กับโปรแกรมประยุกต์ต่างๆได้
3. พัฒนาให้โปรแกรมสามารถเพิ่มเติมชุดของไวยากรณ์จากแฟ้มไวยากรณ์นอกระบบได้ เพื่อให้ระบบมีความยืดหยุ่นเพิ่มขึ้น
4. พัฒนาให้โปรแกรมสามารถใช้งานร่วมกับฐานข้อมูลอื่นๆได้โดยง่าย
5. ในกระบวนการจับคู่คีย์เวิร์ดเข้ากับ entity ของฐานข้อมูลให้ใช้อัลกอริทึมทางด้านภาษาในการเดาความหมายของคำนั้นๆ เช่น Heuristic แทนที่การใช้การระบุคำ
6. สร้างไลบรารีในการทำงานของโปรแกรมหรือทำชุด Application Programming Interface ของโปรแกรม เช่น web service เพื่อให้ระบบอื่นสามารถเรียกใช้การทำงานของระบบถามตอบได้
7. สร้างให้ระบบสามารถตอบสนองต่อคำสั่ง SQL ของระบบจัดการฐานข้อมูลที่หลากหลายในส่วน of หน่วย SQL Formulation Module

บรรณานุกรม

- Chun-Chia Wang et al. 2006. "An Application of Question Answering System for Collaborative Learning." In **Proceedings of the 26th IEEE International Conference on Distributed Computing Systems Workshop (ICDCS 2006)**. Lisbon, Portugal, July 4-6.
- Cody Kwok et al. 2001. "Scaling Question Answering to the Web." In **Proceedings of the tenth international conference on WWW2001**. Hong Kong, May 1-5.
- David Tomas et al. 2005. "Building an XML framework for Question Answering." In **Proceedings of the Cross Language Evaluation Forum 2005**. Vienna, Austria, Sep 21-23.
- Luiz Augusto et al. 2005. "Extracting Exact Answers using a Meta Question Answering System." In **Proceedings of the Australasian Language Technology Workshop 2005**. Sydney, Australia, Dec 10-11.
- Xin Li and Dan Roth 2002. "Learning Question Classifiers." In **Proceedings of the 19th International Conference on Computational Linguistics 2002**. Taipei, Taiwan, Aug 24–Sep 1.
- Zhiguo Gong and Mei Pao Chan. 2005. "iQA: An Intelligent Question Answering System." In **Proceedings of the International Conference on Asian Digital Libraries**. Bangkok, Thailand, Dec 12-15. pp. 332-341

ประวัติผู้เขียน

ชื่อผู้เขียน นายศักดิ์ เทียนเครือ
 วัน เดือน ปี เกิด 6 มกราคม 2524
 สถานที่เกิด จังหวัดราชบุรี

ประวัติการศึกษา

ระดับปริญญาบัตรวิชาชีพ โรงเรียนเทคนิคคอนบอสโกบ้านโป่ง สาขาช่างไฟฟ้า
 จังหวัด ราชบุรี
 ระดับอุดมศึกษา วิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมไฟฟ้า
 คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้า
 พระนครเหนือ ปี 2544

ประวัติการทำงาน

พ.ศ. 2544 – ปัจจุบัน บริษัท โซนี่ ดีไวซ์ เทคโนโลยี (ประเทศไทย) จำกัด