กรอบงานบนพื้นฐานของภววิทยาในฐานะที่เป็นพื้นฐานของระบบสารสนเทศสำหรับสร้างคำถามแบบปรนัย

An Ontology-based Framework as a Foundation of an Information System for Generating Multiple-choice Questions

พรภัทร์ ศิริธรรมกุล

งานวิจัยนี้ได้รับทุนสนับสนุนจากงบประมาณรายได้ ประจำปีงบประมาณ พ.ศ.2563
คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร

กรอบงานบนพื้นฐานของภววิทยาในฐานะที่เป็นพื้นฐานของระบบสารสนเทศสำหรับสร้างคำถามแบบปรนัย

An Ontology-based Framework as a Foundation of an Information System for Generating Multiple-choice Questions

พรภัทร์ ศิริธรรมกุล

งานวิจัยนี้ได้รับทุนสนับสนุนจากงบประมาณรายได้ ประจำปีงบประมาณ พ.ศ.2563
คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร

ชื่อเรื่อง : กรอบงานบนพื้นฐานของภววิทยาในฐานะที่เป็นพื้นฐานของระบบสารสนเทศสำหรับสร้างคำถามแบบปรนัย

ผู้วิจัย : ดร. พรภัทร์ ศิริธรรมกุล สาขาวิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร

พ.ศ. : 2563

บทคัดย่อ

งานวิจัยนี้นำเสนอกรอบงานบนพื้นฐานภววิทยา หรือที่เรียกว่า ออนโทโลจี สำหรับสร้างคำถามที่มีตัวเลือกแบบปรนัย ซึ่งสามารถประเมินความรู้ในระดับคิด – วิเคราะห์ โดยคำถามที่ถูกสร้างจากกรอบงานนี้มุ่งหวังที่จะสนับสนุนให้นักเรียนคิดและประยุกต์ใช้แนวคิดสัมพัทธ์เพื่อผลิตกระบวนทัศน์ที่แตกต่างกันในการแก้ปัญหา ในการทดลองของงานวิจัยนี้แสดงวิธีการใช้กรอบงานที่นำเสนอในการสร้างคำถาม โดยโครงร่างของคำถามประกอบด้วย ปัญหาและแนวทางการแก้ปัญหาเบื้องต้นในประโยคคำถาม ทางเลือกอื่น ๆ ในการแก้ปัญหาจะอยู่ในตัวเลือกแบบปรนัยซึ่งประกอบด้วยคำตอบที่ถูกต้อง และตัวลวง ผลการทดลองแสดงให้เห็นประสิทธิภาพของคำถามในแง่ของการใช้ตัวลวงเพื่อแสดงความเข้าใจผิดของนักเรียน นอกจากนี้ยังมีการใช้ดัชนีที่เรียกว่า พี-สกอร์ (P-scores) และดี-สกอร์ (D-scores) ที่แสดงช่วงความยากของคำถาม และยังสามารถเปิดเผยจุดเด่นและข้อจำกัดของนักเรียนที่มีผลการเรียนที่แตกต่างกัน ท้ายที่สุดในงานวิจัยนี้ได้เสนอวิธีการใช้กรอบงานบนพื้นฐานของภววิทยาในการสร้างคำถามแบบอัตโนมัติในระบบสารสนเทศเพื่อวัดและประเมินความรู้

Title : An Ontology-based Framework as a Foundation of an Information System for Generating Multiple-choice Questions

Researcher : Dr. Pornpat Sirithumgul Department of Computer Engineering, Faculty of Engineering, Rajamangala University of Technology Phra Nakhon

Year : 2020

ABSTRACT

This research is aimed to propose an ontology-based framework for generating multiple-choice questions that can evaluate students' knowledge in the analytical level. The questions generated from this framework could encourage students to apply relative concepts from the lesson they learn to produce different paradigms for solving the same problem. The experiment of this research demonstrates an implementation of the framework. A structure of a question set from the framework comprises a problem and an initial solution in the question phrase, and multiple choices including one correct choice, and three distractors. The experimental results show effectiveness of the questions constructed based on this framework in terms of distractors that can reveal students' misconceptions. Also, difficulty and discrimination indices – P-scores and D-scores used in the experiment show that in a range of difficulty, the generated questions can reveal strength and weaknesses of students in the groups of high and low learning performance. Eventually, this research suggests how the proposed framework could be used as a foundation for automatically generating questions in a knowledge-evaluation system.

กิตติกรรมประกาศ

งานวิจัยนี้สำเร็จด้วยทุนวิจัยจากงบประมาณรายได้ประจำปีงบประมาณ 2563 [รหัสทุนวิจัย 63 – 3306 – 05/3] มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร อีกทั้ง ผู้วิจัยขอขอบพระคุณคณาจารย์จากภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย คณาจารย์จากสาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร และคณาจารย์จากคณะวิทยาศาสตร์และเทคโนโลยี แขนงวิชาเทคโนโลยีสารสนเทศและการสื่อสาร ที่เอื้อเฟื้อคำแนะนำ ข้อมูล และสถานที่ที่ใช้ในงานวิจัย มา ณ ที่นี้

พรภัทร์ ศิริธรรมกุล

สารบัญ

	หน้า
บทคัดย่อ	I
ABSTRACT	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VI
สารบัญรูป	VII
บทที่ 1	8
บทนำ	8
1.1 ความเป็นมาและความสำคัญของงานวิจัย	8
1.2 วัตถุประสงค์ของการวิจัย	9
1.3 การทบทวนวรรณกรรม/สารสนเทศที่เกี่ยวข้อง	10
1.4 ขอบเขตการวิจัย	11
1.5 ประโยชน์ที่ได้รับจากงานวิจัย	11
บทที่ 2	12
ทฤษฎีที่เกี่ยวข้อง	12
2.1 การออกแบบภววิทยา	12
2.2 การออกแบบคำถามเชิงวิเคราะห์	14
2.3 การวัดคุณภาพของตัวลวงในคำถามแบบปรนัย	14
2.4 ดัชนีความยากและค่าอำนาจจำแนก	15
บทที่ 3	16
กรอบงานสำหรับสร้างคำถามแบบปรนัย	16

บทที่ 4 19

การทดลอง 19

บทที่ 5 24

ผลการทดลอง 24

บทที่ 6 28

บทสรุปงานวิจัย 28

6.1 ประเด็นถกเถียงและการพัฒนางานวิจัยต่อเนื่องในอนาคต 28

6.1 บทสรุป 31

บรรณานุกรม 32

ภาคผนวก 35

ประวัติผู้วิจัย 36

สารบัญตาราง

ตารางที่	หน้า
4.1: แนวคิดและความสัมพันธ์ระหว่างแนวคิดในโดเมนการโปรแกรมคอมพิวเตอร์	20
4.2: ข้อมูลสรุปเกี่ยวกับคำถาม Q1 - Q11 ที่สร้างขึ้น	23
5.1: จำนวนของคำถามที่ปรากฏใน 3 ระดับความยาก	25
5.2: จำนวนของคำถามที่ปรากฏใน 3 ระดับของดัชนีอำนาจจำแนก	26
6.1: ความสัมพันธ์ของข้อมูล D-scores และ P-scores	28

สารบัญภาพ

ภาพที่	หน้า
2.1: ความสัมพันธ์ของแนวคิดทั้ง 4 รูปแบบ	13
3.1: ตัวอย่างการกำหนดภววิทยาของโครงสร้างการเขียนโปรแกรมแบบลูป	17
3.2: ตัวอย่างการสร้างคำตอบเริ่มต้นจากหนึ่งแนวคิดที่กำหนดในภววิทยา	17
3.3: ตัวอย่างการสร้างคำตอบที่ถูกต้องจากหนึ่งในแนวคิดสัมพัทธ์	18
4.1: ตัวอย่างการสร้างประโยคคำถามที่ประกอบด้วยคำตอบเริ่มต้นและประโยคคำถามให้ค้นหาคำตอบโดยวิธีอื่น	21
4.2: ตัวอย่างการสร้างตัวเลือกของคำตอบที่ถูกต้อง	21
4.3: ตัวอย่างการสร้างตัวเลือกของคำตอบที่เป็นตัวลวง	22
6.1: สถาปัตยกรรมระดับบนของระบบสารสนเทศสำหรับสร้างคำถามและประเมินผลการเรียนนอัตโนมัติ	30

บทที่ 1

บทนำ

ในบทนี้เป็นการกล่าวถึงความเป็นมาและความสำคัญของงานวิจัย วัตถุประสงค์ ขอบเขตงานวิจัย และประโยชน์ที่ได้รับจากงานวิจัย ซึ่งมีรายละเอียดดังต่อไปนี้

1.1 ความเป็นมาและความสำคัญของงานวิจัย

งานวิจัยนี้ได้รับแรงบันดาลใจจากความต้องการในการสร้างหน่วยประมวลผลความรู้อัตโนมัติสำหรับการเรียนการสอนออนไลน์ เช่น การเรียนผ่านระบบ MOOCs (Massive Open Online Courses) (“Massive open online course,” n.d.) และ Coursera (“Coursera,” n.d.) และการเรียนการสอนทางไกลในประเทศไทย ได้แก่ การเรียนการสอนของมูลนิธิการศึกษาทางไกลผ่านดาวเทียมในพระบรมราชูปถัมภ์ และการเรียนการสอนของมหาวิทยาลัยสุโขทัยธรรมาธิราชซึ่งเป็นมหาวิทยาลัยเปิดแห่งเดียวในประเทศไทยที่มีจำนวนนักศึกษาที่ลงทะเบียนเรียนต่อหนึ่งวิชาเป็นจำนวนมหาศาล อย่างไรก็ดี การเรียนการสอนทางไกลยังมีข้อจำกัดอยู่คือ ผู้เรียนและผู้สอนยังสื่อสารถึงกันในลักษณะของอะซิงโครนัส (asynchronous)[1] ทำให้เป็นไปได้ยากสำหรับผู้สอนที่จะสามารถประเมินความเข้าใจของนักเรียน และเตรียมการสอนในครั้งถัดไปได้อย่างมีประสิทธิภาพ

การใช้แบบฝึกหัดท้ายบท หรือข้อสอบวัดการเรียนรู้จึงมีความจำเป็นเพื่อประเมินความรู้ของผู้เรียนก่อนที่จะมีการเรียนการสอนในครั้งถัดไป แต่ข้อจำกัดของการใช้แบบฝึกหัดท้ายบทและข้อสอบวัดประเมิน คือความต้องการแรงงาน (effort) และเวลาจำนวนมากเพื่อสร้างชุดคำถาม และเพื่อประเมินความรู้ของนักเรียนจากคำตอบของคำถามที่สร้างขึ้น อีกทั้งการเรียนออนไลน์ และการเรียนในมหาวิทยาลัยเปิด โดยปกติแล้วจะมีผู้เรียนเป็นจำนวนมหาศาล จึงเป็นไปได้ยากที่จะมีบุคลากรที่เพียงพอสำหรับการสร้างชุดคำถามที่หลากหลายสำหรับการประเมินผลความรู้ และแทบจะเป็นไปไม่ได้เลยที่ผู้สอนจะสามารถตอบกลับ (feedback) ผู้เรียนได้อย่างทันท่วงที

ปัญหาในข้างต้นจึงเป็นที่มาของงานวิจัยนี้ที่มุ่งเน้นค้นหาวิธี กระบวนการการทำงานสำหรับระบบสารสนเทศที่อยู่ในลักษณะของระบบผู้เชี่ยวชาญ (expert system) ที่สามารถสร้างคำถามและวัดการเรียนรู้ได้อย่างอัตโนมัติ โดยคำถามที่สร้างจะเป็นแบบปรนัยซึ่งถูกมองว่าเหมาะสมกว่าคำถามแบบอัตนัยในแง่ที่ว่า คำถามแบบปรนัย (ในที่นี้หมายถึง คำถามที่มีการเลือกคำตอบที่ถูกต้องที่สุดจากเซตของตัวเลือก หรือที่เรียกว่า multiple-choice question) สามารถนำไปประมวลผลต่อได้ง่าย

[1] หมายถึงการสื่อสารที่ผู้รับสารและผู้ส่งสาร ในบริบทนี้หมายถึงการสื่อสารระหว่างผู้เรียนและผู้สอน ที่ไม่ได้อยู่ในช่วงเวลาเดียวกัน มักพบในการสอนออนไลน์โดยใช้วิดิโอ หรือกระทู้สนทนาออนไลน์ (online discussions)

ในขณะที่คำถามแบบอัตนัย เช่น การเติมคำในช่องว่าง (short-essay question) และการตอบแบบพรรณา (long-essay question) ยังต้องอาศัยเทคนิคการวิเคราะห์ภาษาเขียน เช่น การประมวลผลภาษาธรรมชาติ (Natural Language Processing: NLP) ซึ่งทำให้การประมวลผลความรู้จากงานเขียนเป็นไปได้ยาก

อย่างไรก็ดี การใช้คำถามแบบปรนัยยังเป็นที่กังขาของนักวิชาการถึงประสิทธิภาพในการวัดความรู้ในระดับการวิเคราะห์ (analytical level) ของนักเรียน อีกทั้งในงานวิจัยก่อนหน้า (Kunichika Agarwal and Mannem 2011; Smith et al. 2010; Alsubait et al. 2016; Kwankajornkiet et al. 2016) จะเป็นคำถามประเภท *'อะไร'* (what question) ซึ่งเป็นคำถามประเภทวัดความจำ (remembering-level question) มากกว่าคำถามประเภท *'อย่างไร'* และ *'ทำไม'* (how and why) ซึ่งอยู่ในระดับความเข้าใจ (understanding-level question) และระดับการคิดวิเคราะห์ (analysis-level question) ดังที่ได้ระบุไว้ใน Bloom's taxonomy (Forehand, 2005)

งานวิจัยนี้จึงมุ่งหวังที่จะนำเสนอกรอบแนวคิด (framework) บนพื้นฐานของหลักออนโทโลจี (ontology) หรือที่เรียกว่า หลักภววิทยา อันจะเป็นแกนหลักสำคัญของการสร้างซอฟต์แวร์ที่ฝังตัวในระบบผู้เชี่ยวชาญที่ใช้สำหรับสร้างคำถามแบบปรนัยที่สนับสนุนให้นักเรียนได้ใช้ทักษะการคิดวิเคราะห์ในการตอบคำถาม อีกทั้งกรอบแนวคิดนี้จะยังสามารถนำไปสู่ การออกแบบหน่วยประมวลผลการเรียนรู้สำหรับแนะนำผู้สอนถึงประเด็นที่มีแนวโน้มที่จะเป็นจุดอ่อนของการเรียนรู้ของผู้เรียนอีกด้วย

1.2 วัตถุประสงค์ของการวิจัย

1) เพื่อสร้างองค์ความรู้ใหม่ทางสาขาภววิทยา โดยองค์ความรู้ใหม่นี้อยู่ในรูปแบบของกรอบงาน (framework) สำหรับสร้างชุดคำถามเพื่อใช้ทดสอบความรู้ในระดับคิดวิเคราะห์
2) เพื่อนำเสนอวิธีการประยุกต์ใช้กรอบงานเพื่อเป็นกลไกของซอฟต์แวร์และระบบสารสนเทศในรูปแบบของผู้เชี่ยวชาญ (expert system) ที่สามารถสร้างคำถามแบบปรนัยและวิเคราะห์คำตอบของคำถามได้ในเชิงลึก
3) เพื่อสร้างนวัตกรรมยกระดับการศึกษา โดยมุ่งหวังเพื่อช่วยลดแรงงาน ในการสร้างคำถามที่มีจำนวนมากและหลากหลาย อีกทั้ง นวัตกรรมนี้จะช่วยยกระดับคุณภาพการเรียนของผู้เรียนได้อีกด้วย

1.3 การทบทวนวรรณกรรม/สารสนเทศที่เกี่ยวข้อง

งานวิจัยนี้จัดอยู่ในกลุ่มเดียวกับงานวิจัยที่เสนอวิธีการสร้างคำถามแบบปรนัยโดยอัตโนมัติ (Kunichika et al. 2004; Heilman and Noah 2010; Brown et al. 2005; Mitkov et al. 2006; Goto et al. 2010; Kumar et al. 2015; Agarwal and Mannem 2011; Smith et al. 2010; Alsubait et al. 2016; Kwankajornkiet et al. 2016) โดยวัตถุประสงค์เบื้องต้นของงานวิจัยในกลุ่มนี้คือ ต้องการลดต้นทุน เวลา และแรงงานของผู้สอนในการสร้างคำถามเพื่อประเมินผลความรู้ของนักเรียน จุดเริ่มต้นของงานวิจัยในกลุ่มนี้จะเน้นไปที่การสร้างระบบอัตโนมัติสำหรับผลิตชุดคำถามเพื่อวัดความรู้ทางภาษา (Kunichika et al. 2004; Heilman and Noah 2010; Brown et al. 2005; Mitkov et al. 2006) เช่น งานวิจัยของ Kunichika et al. (2004) ที่เน้นการสร้างคำถามเพื่อทดสอบไวยากรณ์ภาษาอังกฤษ และงานวิจัยของ Brown et al. (2005) ที่เน้นการสร้างคำถามเพื่อทดสอบคำศัพท์ภาษาอังกฤษ งานวิจัยยุคถัดมาในกลุ่มนี้ใช้การสร้างคำถามแบบอัตโนมัติเพื่อทดสอบความเข้าใจในเนื้อหาของบทเรียน (content) มากกว่าเพื่อการทดสอบความสามารถทางภาษา เช่น งานวิจัยของ Mitkov et al. (2006) ใช้ระบบอัตโนมัติเพื่อสร้างคำถามประเภท wh-questions เช่น *'what'* และ *'which'* เพื่อเป็นการทดสอบความเข้าใจของผู้เรียนเกี่ยวกับบทความที่อ่าน แต่ข้อจำกัดของวิธีการสร้างคำถามประเภทนี้ คือคำถามจะถูกสร้างจากกลุ่มคำนามที่ทำหน้าที่เป็นกรรม (object) ในประโยคเท่านั้น และคำถามที่สร้างขึ้นก็ยังถูกจำกัดการวัดความรู้ที่อยู่ในระดับการวัดความจำ (remembering-level question) เท่านั้น

มีงานวิจัยอีกกลุ่มหนึ่ง (Goto et al. 2010; Kumar et al. 2015; Agarwal and Mannem 2011; Smith et al. 2010; Kwankajornkiet et al. 2016) เลือกที่จะสร้างประโยคคำถาม (question phase) ประเภทเติมคำในช่องว่าง แทนที่การสร้างคำถามแบบ wh-questions เพื่อเป็นการขยายโอกาสให้เกิดความหลากหลายในการสร้างคำถามเพื่อถามเนื้อหาในบทเรียน แต่คำถามแบบเติมคำในช่องว่างก็ยังมีอุปสรรคที่พบอย่างน้อย 3 อย่างได้แก่ (1) ประโยคที่ถูกสร้าง หรือเลือกใช้เพื่อเป็นประโยคคำถามยังไม่มีเนื้อหาสาระเพียงพอที่จะใช้เป็นคำถามแบบเติมคำในช่องว่าง รวมถึงยังไม่เหมาะสมที่จะใช้เป็นประโยคคำถามเพื่อวัดความรู้ (2) ปัญหาในการเลือกคำสำคัญ (keyword) ในประโยคคำถาม กล่าวคือ เมื่อคำสำคัญถูกดึงออกจากประโยคคำถาม พบว่าผู้อ่านไม่สามารถเข้าใจเนื้อหาโดยรวมของประโยคจากคำที่เหลือในประโยคได้และ (3) ปัญหาในการสร้างตัวลวง (distractors) ที่เหมาะสมในเซตของตัวเลือกแบบปรนัย โดยตัวลวงที่เหมาะสมในที่นี้หมายถึงตัวเลือกแบบปรนัยที่ไม่มีคุณลักษณะของ *"ตัวเลือกที่ไม่ทำงาน"* (non-functioning option) (Haladyna and Downing 1993; Cizek and O'Day 1994; Sidick et al. 1994; Delgado and Prieto 1998; Roger and Harley 1999; Woodford and Bancroft 2005; Shizuka et al. 2006; Tarrant et al. 2009)

ถึงแม้ทิศทางของงานวิจัยในกลุ่มนี้จะดีขึ้นในแง่ของการสร้างคำถามเพื่อวัดเนื้อหาความรู้ การมุ่งประเด็นของการพัฒนาก็ยังอยู่ในระดับการพัฒนาภาษาของประโยคคำถามให้ใกล้เคียงการแต่งประโยคของมนุษย์มากกว่าการมุ่งประเด็นสร้างประสิทธิภาพของคำถามเพื่อตรวจวัดระดับความรู้ความเข้าใจ (cognition levels) รวมถึงเทคนิคที่นำมาใช้ก็ยังอยู่ในระดับของการใช้เทคนิคการประมวลผลภาษาธรรมชาติ (Natural Language Processing: NLP) มากกว่าการใช้เทคนิคเพื่อค้นหาความรู้ความเข้าใจของมนุษย์ ดังแสดงในงานวิจัยของ Goto et al. (2010) และ Kumar et al. (2015)

งานวิจัยนี้จึงมุ่งเน้นที่จะลดข้อจำกัดของงานวิจัยก่อนหน้า ด้วยการใช้เทคนิคทางภววิทยา (ontology) เพื่อช่วยสร้างคำถามแบบปรนัยที่เน้นให้ประโยคคำถามที่ถูกสร้างขึ้นสามารถเป็นปัจจัยในการกระตุ้นให้ผู้เรียนคิดมากขึ้นถึงวิธีการ (paradigm) ที่แตกต่างแต่ยังอยู่โดเมนความรู้เดียวกัน ซึ่งประโยชน์ของคำถามในลักษณะนี้จะช่วยให้นักเรียนสามารถพัฒนาแนวทางการแก้ปัญหาซึ่งมีจุดเด่นและข้อจำกัดที่แตกต่างกันออกไป อีกทั้งตัวเลือกแบบปรนัยรวมถึงตัวลวงที่ถูกสร้างขึ้นจากวิธีการที่งานวิจัยนี้นำเสนอยังจะสามารถนำไปใช้เพื่อวิเคราะห์หาความรู้ความเข้าใจ (cognition level) รวมถึงความเข้าใจที่ผิดพลาด (misconception) หรือความสับสนในแนวคิด (confusion) ของนักเรียนได้อีกด้วย

1.4 ขอบเขตการวิจัย

1) ออกแบบและสร้างกรอบงาน (framework) สำหรับสร้างคำถามแบบปรนัย ที่สามารถวัดความรู้ในระดับการคิดวิเคราะห์ได้ (analytical level)
2) ทดสอบประสิทธิภาพและความน่าเชื่อถือของกรอบงานที่สร้างขึ้น โดยการทดสอบจะเน้นแสดงให้เห็นถึงประสิทธิภาพของการนำกรอบงานไปใช้สร้างคำถามเพื่อใช้วัดความรู้ในโดเมนความรู้ตัวอย่าง ตลอดจนแสดงให้เห็นว่าคำถามที่สร้างจากกรอบงานสามารถนำไปใช้เพื่อวัดความเชี่ยวชาญ และข้อจำกัดของนักเรียนได้จริง
3) แสดงแนวโน้มของการนำกรอบงานไปใช้เป็นกลไกสำคัญสำหรับสร้างระบบสารสนเทศในรูปแบบของระบบผู้เชี่ยวชาญ (expert system) ที่มีกระบวนการทำงานในการสร้างคำถามและประเมินผลความรู้อัตโนมัติ

1.5 ประโยชน์ที่ได้รับจากงานวิจัย

กรอบงาน (framework) และสถาปัตยกรรมของระบบสารสนเทศ (information system architecture) ซึ่งเปรียบเสมือนเป็นพิมพ์เขียว (blue print) สำหรับสร้างระบบผู้เชี่ยวชาญ

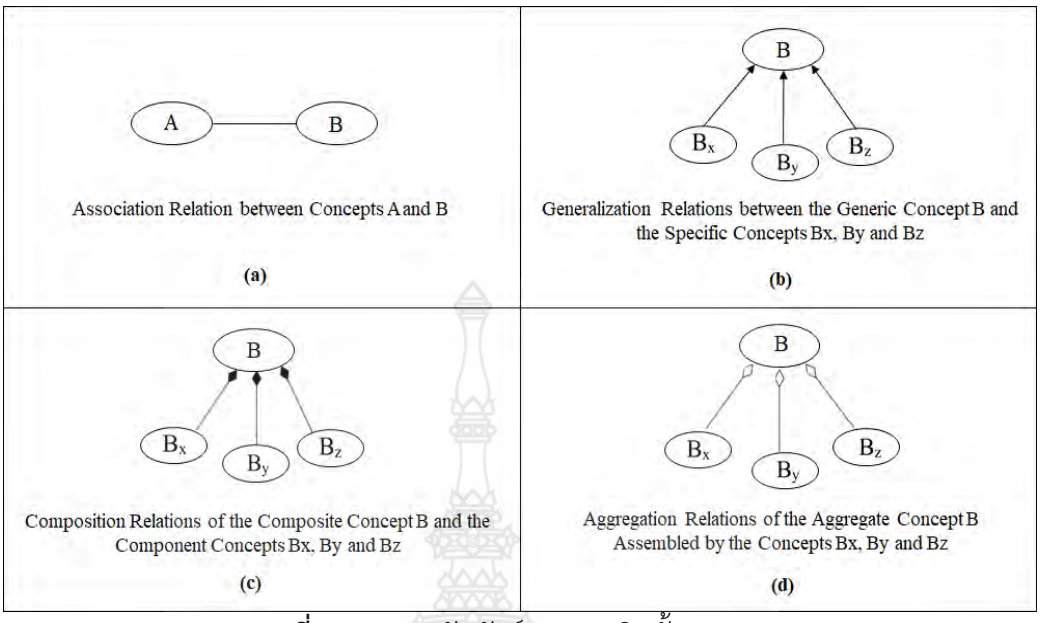
บทที่ 2

ทฤษฎีที่เกี่ยวข้อง

ทฤษฎีพื้นฐานที่นักวิจัยนำมาใช้ในโครงการวิจัยนี้ ได้แก่ การออกแบบภววิทยา (ontology design) ซึ่งถูกนำมาใช้เพื่อออกแบบกรอบงาน (framework) สำหรับสร้างคำถามสำหรับวัดความรู้แบบปรนัย โดยกรอบงานที่สร้างจะอ้างอิงมาจากงานวิจัยของ Ragonis (2012) ซึ่งแนะนำประเภทของคำถามที่เหมาะสมจะใช้ถามเพื่อวัดความรู้ที่อยู่ในกลุ่มวิชาวิทยาศาสตร์คอมพิวเตอร์ (computer science) การอ้างอิงงานวิจัยนี้จึงทำให้เกิดแนวคิดของการสร้างคำถามในเชิงวิเคราะห์ (analytical-level question) ได้เป็นอย่างดี นอกจากนี้ งานวิจัยนี้ยังอ้างถึงหลักการวิเคราะห์คุณภาพของคำถามที่ใช้เป็นเครื่องมือวัดความรู้ซึ่งแบ่งได้เป็น 3 ลักษณะ ได้แก่ (1) การวัดประสิทธิภาพของตัวลวง (distractor) ที่ปรากฏอยู่ในตัวเลือก (2) การวัดระดับความยากของคำถาม โดยเฉพาะอย่างยิ่งสำหรับคำถามในข้อสอบมาตรฐานที่เชื่อถือได้ งานวิจัยก่อนหน้า Pande et al. (2013) แนะนำว่าคำถามที่ดีไม่ควรยากหรือง่ายจนเกินไป เพราะคำถามที่ง่าย หรือยากจนเกินไปจะไม่สามารถวัดความรู้ได้เลย และ (3) การวัดดัชนีอำนาจจำแนก ซึ่งก็คือ การวิเคราะห์คุณภาพของคำถามจากการบ่งชี้ประสิทธิภาพในการตอบคำถามได้ถูกต้องเมื่อเปรียบเทียบกับประสิทธิภาพในการเรียนโดยทั่วไปของนักเรียน กล่าวคือ คำถามที่ได้มาตรฐานจะสามารถแบ่งแยกกลุ่มของนักเรียนเรียนเก่งออกจากกลุ่มนักเรียนเรียนอ่อนได้ (Pande et al. 2013) และสอดคล้องกับผลการเรียนโดยทั่วไปของนักเรียน รายละเอียดของทฤษฎีที่กล่าวข้างต้น เป็นดังต่อไปนี้

2.1 การออกแบบภววิทยา

การออกแบบภววิทยาเป็นการกำหนดกรอบ (scope) ของโดเมนความรู้ (knowledge domain) ที่สนใจ โดยทั่วไปแล้วในกรอบภววิทยาหนึ่งจะประกอบด้วยส่วนสำคัญ 2 ส่วน ได้แก่ ส่วนของการอธิบายความรู้ด้วยแนวคิดสัมพัทธ์ (relative concepts) และส่วนของการอธิบายความสัมพันธ์ของแนวคิด (concepts' relations) ภาพที่ 2.1 แสดงความสัมพันธ์ของแนวคิด 4 รูปแบบ ดังนี้



ภาพที่ 2.1: ความสัมพันธ์ของแนวคิดทั้ง 4 รูปแบบ

ภาพที่ 2.1 (a) แสดงความสัมพันธ์แบบแอสโซชิเอชัน (association) ระหว่างแนวคิด (concept) A และ B ความสัมพันธ์นี้ใช้เพื่ออธิบายสองแนวคิดใด ๆ ที่โดยปกติแล้วเกิดขึ้นพร้อมกันในความเป็นจริง เช่น เทียนกับเค้กวันเกิด (Budanitsky & Hirst, 2001) ความสัมพันธ์นี้ยังสามารถอธิบายแนวคิดตรงกันข้ามได้ด้วย เช่น Black-box กับ White-box testing (Sirithumgul & Olfman, 2013)

ภาพที่ 2.1 (b) แสดงความสัมพันธ์แบบเจเนอรัลไรเซชัน (generalization) ซึ่งเป็นความสัมพันธ์ระหว่างแนวคิดทั่วไป (generic concept) B และแนวคิดจำเพาะ (specific concepts) B_x, B_y และ B_z โดยความคิดจำเพาะเหล่านี้จะได้รับการถ่ายทอดคุณสมบัติมาจากแนวคิด B และอีกนัยหนึ่ง แนวคิดจำเพาะเหล่านี้ ในแต่ละตัวก็มีคุณสมบัติบางอย่างเฉพาะตัวของตัวเอง (Bollegala et al., 2007)

ภาพที่ 2.1 (c) และ (d) แสดงความสัมพันธ์แบบพาร์ทและโฮล (parts and whole) โดยในภาพที่ 2.1 (c) เป็นความสัมพันธ์ที่เรียกว่าคอมโพสิชัน (composition) และภาพที่ 2.1 (d) เป็นความสัมพันธ์ที่เรียกว่า แอกกริเกชัน (aggregation) ความแตกต่างของความสัมพันธ์ทั้งสองแบบคือ ในขณะที่ความสัมพันธ์แบบ composition จะถูกมองว่า B_x, B_y และ B_z ถูกหลอมรวมกันเป็นชิ้น B เดียวกัน เช่น ล้อรถยนต์ถูกมองเป็นส่วนหนึ่งของรถยนต์ ส่วนความสัมพันธ์แบบ aggregation มอง Bx, By และ Bz แยกจาก B เช่นห้องเรียน กับนักเรียนในห้องเรียน เป็นต้น

2.2 การออกแบบคำถามเชิงวิเคราะห์

การสร้างคำถามเชิงวิเคราะห์ ดังที่ได้แนะนำในเอกสารวิจัยของ Ragonis (2012) มีอยู่หลากหลายรูปแบบ สำหรับงานวิจัยชิ้นนี้ผู้วิจัยเลือกใช้งานแบบที่ 12 (type-12) ที่เรียกว่า การแปลงการแก้ปัญหาจากรูปแบบหนึ่งไปเป็นอีกรูปแบบหนึ่ง (transforming a solution from one representation to another) ซึ่งเป็นมุมมองของการใช้กระบวนทัศน์ (paradigm) ที่หลากหลายในการแก้ปัญหาเดียวกันมาใช้ในสร้างกรอบงาน (framework) ที่นำเสนอในงานวิจัยนี้ ตัวอย่างของคำถามในลักษณะนี้ สามารถยกตัวอย่างได้จากวิชาการโปรแกรมคอมพิวเตอร์ เช่น (a) การใช้เมธอด (method) ที่แตกต่างกันในการเรียงตัวเลขที่อยู่ใน array, (b) การสร้างเมธอดเดียวกันด้วยภาษาสำหรับการเขียนโปรแกรมคอมพิวเตอร์ที่ต่างกัน เช่น Java, C และ Python เป็นต้น, (c) การใช้โครงสร้างของการโปรแกรมคอมพิวเตอร์ที่ใกล้เคียงกัน (หรือมีลักษณะ/คุณสมบัติบางอย่างร่วมกัน) เช่น if-statement และ switch-case statement ในการแก้ปัญหาเดียวกัน และ (d) การใช้อัลกอริทึมที่แตกต่างกัน เช่น loop structure และ recursive method ในการแก้ปัญหางานอย่างเดียวกัน

2.3 การวัดคุณภาพของตัวลวงในคำถามแบบปรนัย

นอกจากนี้ทฤษฎีที่นำมาใช้ในการสร้างคำถามแล้ว งานวิจัยนี้ยังมีการใช้ดัชนีที่ใช้ชี้วัดคุณภาพของคำถามด้วย ได้แก่ การวัดคุณภาพคำตอบที่เป็นตัวลวง (distractor) จากการพิจารณาตัวลวงที่ไม่ได้ทำงาน (non-functioning option) กล่าวคือ หากตัวลวงในคำถามหนึ่งถูกเลือกน้อยกว่า 5 เปอร์เซ็นต์ของจำนวนผู้เข้าสอบทั้งหมด ตัวลวงนี้ถือว่าเป็น non-functioning option

ในการพิจารณาคุณภาพของคำถามแบบปรนัย จะทำได้จากจำนวน non-functioning option ที่ปรากฏในชุดคำถาม หากพบว่าในชุดคำถามมีตัวลวงอย่างน้อยหนึ่งตัวที่ไม่เป็น non-functioning option ชุดคำถามนั้นถือว่าเชื่อถือได้ (Haladyna and Downing 1993; Cizek and O'Day 1994; Sidick et al. 1994; Delgado and Prieto 1998; Roger and Harley 1999; Woodford and Bancroft 2005; Shizuka et al. 2006; Tarrant et al. 2009)

2.4 ดัชนีความยากและค่าอำนาจจำแนก

ดัชนีวัดความยาก (Difficulty Index หรือที่เรียกว่า P-scores) และดัชนีอำนาจจำแนก (Discrimination Index หรือที่เรียกว่า D-scores) (Pande et al., 2013; Rao et al., 2016) แสดงในสมการ (eq.1) และ (eq.2) ตามลำดับ

$$P - scores = \frac{h+l}{n} \times 100 \qquad \textbf{(eq.1)}$$

$$D - scores = \frac{h-l}{n} \times 2 \qquad \textbf{(eq.2)}$$

เมื่อ h คือจำนวนของนักเรียนที่ตอบคำถามได้ถูกต้อง และอยู่ในกลุ่มนักเรียนเรียนดี
l คือจำนวนของนักเรียนที่ตอบคำถามได้ถูกต้อง และอยู่ในกลุ่มนักเรียนเรียนอ่อน
n คือจำนวนของนักเรียนทั้งสองกลุ่ม (ทั้งเรียนดี และเรียนอ่อน) รวมถึงนักเรียนที่ไม่ได้ตอบคำถามด้วย

ดัชนีวัดความยาก (P-scores) และดัชนีอำนาจจำแนก (D-scores) ถูกนำมาใช้เพื่อแสดงคุณภาพของคำถามที่ถูกสร้างตามกรอบงาน (framework) ที่นำเสนอ โดยคุณภาพจะสามารถบอกได้จากค่า P-scores ที่อยู่ในช่วง 30 ถึง 70 หมายถึง ค่าดัชนีที่ต้องการ ซึ่งสะท้อนว่าคำถามที่สร้างขึ้นไม่ยากหรือง่ายจนเกินไป และค่า D-scores ที่อยู่ในช่วง 0.20 ถึง 0.29 หมายถึง ค่าดัชนีที่ต้องการ ซึ่งสะท้อนว่าคำถามสามารถใช้แยกแยะความสามารถของนักเรียนเรียนเก่งและนักเรียนเรียนอ่อนได้อย่างมีประสิทธิภาพ

บทที่ 3

กรอบงานสำหรับสร้างคำถามแบบปรนัย

กรอบงาน (framework) ที่นำเสนอในงานวิจัยนี้แบ่งออกเป็น 4 ขั้นตอน ประกอบด้วย ขั้นตอนที่ 1 เกี่ยวกับการเขียนข้อกำหนดภววิทยา (ontology specifications) ที่ว่าด้วยการกำหนดแนวคิด (concepts) และความสัมพันธ์ของแนวคิด (concepts' relations) ที่อยู่ในโดเมนความรู้ (knowledge domain)[2] ที่สนใจ ขั้นตอนที่ 2 – 4 โดยสรุปแล้วเกี่ยวกับการใช้งานภววิทยาที่ถูกกำหนดในขั้นตอนที่ 1 เพื่อสร้างประโยคคำถาม และสร้างตัวเลือกแบบปรนัย 4 ตัวเลือก ประกอบด้วยตัวเลือกที่เป็นคำตอบที่ถูกต้อง (correct answer choice) หนึ่งตัวเลือก และตัวลวง (distractor choices) อีกสามตัวเลือก รายละเอียดทั้ง 4 ขั้นตอนของกรอบงาน เป็นดังนี้

ขั้นตอนที่ 1 – การกำหนดภววิทยา: หลักการออกแบบภววิทยาถูกนำมาใช้ในขั้นตอนนี้ ตามหลักความสัมพันธ์ของแนวคิดในภววิทยานั้น ความสัมพันธ์จะเป็นแบบ association เมื่อแนวคิดมีน้ำหนักพอ ๆ กัน อยู่ในระนาบเดียวกัน หรือมักเกิดขึ้นพร้อม ๆ กันในโลกแห่งความเป็นจริง เช่น ความสัมพันธ์ของแนวคิดในรูปแบบเชิงพื้นที่ (spatial) หรือความสัมพันธ์ของแนวคิดในรูปแบบเชิงเวลา (temporal)

ในกรณีแนวคิดมีความสัมพันธ์แบบแนวคิดทั่วไป (generic concept) และแนวคิดเฉพาะเจาะจง (specific concept) หรือแนวคิดทั่วไปกับแนวคิดที่เป็นตัวอย่าง (instance concept) จะแสดงความสัมพันธ์ในรูปแบบของ generalization

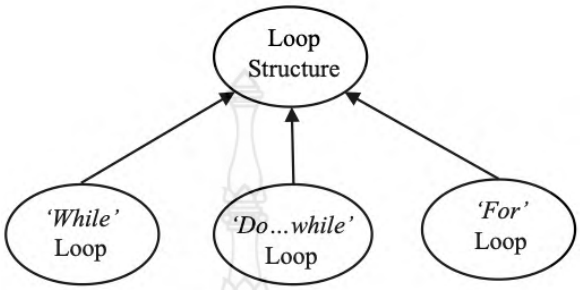
นอกจากนี้ แนวคิดยังสามารถอยู่ในลักษณะของความสัมพันธ์ระหว่างแนวคิดส่วนรวมทั้งหมด (whole) และแนวคิดที่เป็นส่วนประกอบ (part) เช่น แนวคิดแบบ whole คือเซลล์ของสิ่งมีชีวิตและแนวคิดแบบ part ได้แก่ นิวเคลียส (nucleus) ไซโตปลาสซึม (cytoplasm) และเมมเบรน (membrane) เป็นต้น ความสัมพันธ์ในรูปแบบนี้จะถูกเรียกว่า parts and whole

ขั้นตอนที่ 2 – การสร้างคำถาม: ดังที่ได้แนะนำในงานวิจัยของ Ragonis (2012) ว่าการตั้งคำถามที่ดีในลักษณะหนึ่ง ควรสนับสนุนให้นักเรียนคิดถึงแนวคิดการแก้ปัญหาที่หลากหลาย แนวคิดนี้จะนำมาใช้ในขั้นตอนนี้ กล่าวคือ สมมติมีการกำหนดภววิทยาดังแสดงในภาพที่ 3.1 ที่นิยามไว้ว่า

---

[2] โดเมนความรู้ หมายถึงความรู้จำเพาะในสาขาวิชา คำว่าโดเมนความรู้มีความหมายตรงกันข้ามกับความรู้โดยทั่วไป เช่น วิศวกรรมซอฟต์แวร์จัดว่าเป็นการกล่าวถึงความรู้โดยทั่วไปที่มีโดเมนความรู้จำเพาะย่อยลงมา เช่น โดเมนความรู้เกี่ยวกับการโปรแกรมคอมพิวเตอร์ เป็นต้น

โครงสร้างของการเขียนโปรแกรมแบบลูป (loop) ซึ่งเป็นแนวคิดโดยทั่วไป (generic concept) ของคำสั่งให้ทำซ้ำ สามารถอธิบายในรายละเอียดจำเพาะ (specific concept) ของวิธีที่ทำซ้ำประกอบด้วย while loop, do...while, และ for loop



ภาพที่ 3.1: ตัวอย่างการกำหนดภววิทยาของโครงสร้างการเขียนโปรแกรมแบบลูป

ดังนั้น การออกแบบคำถามตามกรอบงานที่นำเสนอนี้ แท้จริงแล้ว เริ่มต้นจากการคิดถึงปัญหาที่ต้องแก้ด้วยการโปรแกรมแบบ loop และในประโยคคำถามต้องมีการสร้างคำตอบเริ่มต้น (initial answer) จากหนึ่งแนวคิดที่กำหนดในภววิทยา เช่นในภาพที่ 3.2 ที่มีการใช้คำสั่ง while loop ในการสร้างคำตอบเริ่มต้น และมีการถามให้นักเรียนหาวิธีการอื่น ๆ ที่สามารถใช้เป็นแนวทางการแก้ปัญหาได้เช่นเดียวกับคำตอบเริ่มต้นที่แนะนำในโจทย์คำถาม

```
int i <-- 1
int j <-- 1

while ( i < 11 )
    j <-- j * i
    i <-- i + 1
    if ( j == x )
        break
    end if
end while
```

ภาพที่ 3.2: ตัวอย่างการสร้างคำตอบเริ่มต้นจากหนึ่งแนวคิดที่กำหนดในภววิทยา

ขั้นตอนที่ 3 – การสร้างตัวเลือกของคำตอบที่ถูกต้อง: ขั้นตอนนี้เป็นการสร้างคำตอบที่ถูกต้องจากแนวคิดสัมพันธ์ที่กำหนดในภาววิทยา ดังแสดงในภาพที่ 3.1 แนวคิดสัมพันธ์ หรือแนวคิดที่เกี่ยวข้องกันประกอบด้วย while, do...while และ for การสร้างคำตอบที่ถูกต้องคือการสร้างคำตอบบนพื้นฐานแนวคิดที่แตกต่างจากแนวคิดที่เสนอในคำถาม เช่น ในคำถามใช้ while ในการสร้างคำตอบเริ่มต้น ดังนั้น ตัวเลือกของคำตอบที่ถูกต้องจึงใช้ for loop (ดังแสดงในภาพที่ 3.3) ในการสร้างคำตอบ ซึ่งในที่นี้จะเห็นว่าการใช้คำสั่ง while ในโจทย์คำถามและ for ในตัวเลือกให้ผลลัพธ์ของคำตอบที่เหมือนกัน

```
int i <-- 1
int j <-- 1

for (i = 1; i < 11; i++)
    j <-- j * i
    if (j == x)
        // program is terminated
    end if
end for
```

ภาพที่ 3.3: ตัวอย่างการสร้างคำตอบที่ถูกต้องจากหนึ่งในแนวคิดสัมพันธ์

ขั้นตอนที่ 4 – การสร้างตัวลวง: ขั้นตอนนี้จะอ้างอิงคำแนะนำในเอกสารวิจัยก่อนหน้า (Haladyna and Downing 1993; Cizek and O'Day 1994; Sidick et al. 1994; Delgado and Prieto 1998; Roger and Harley 1999; Woodford and Bancroft 2005; Shizuka et al. 2006; Tarrant et al. 2009) ในการสร้างตัวลวง (distractor) ใน 3 ตัวเลือกของแต่ละชุดคำถาม ที่ว่า ตัวลวงจะต้องไม่เป็นตัวเลือกที่ไม่ทำงาน (non-functioning distractor)

เพื่อเป็นการหลีกเลี่ยงตัวเลือกที่ไม่ทำงาน การสร้างตัวเลือกควรสร้างความคล้ายคลึงระหว่างคำตอบที่ถูกต้องและตัวลวง กรอบงานที่นำเสนอในงานวิจัยนี้กำหนดให้ใช้แนวคิดสัมพันธ์ในการสร้างตัวลวงซึ่งเป็นสาเหตุทำให้คำตอบที่ถูกต้องมีความคล้ายคลึงกับตัวลวงในแง่ของแนวคิดที่มีความสัมพันธ์กัน และ/หรือมีแนวคิดรวบยอด (conceptual idea) ที่เป็นไปในทิศทางเดียวกัน ดังตัวอย่างของภาววิทยาที่แสดงในภาพที่ 3.1 แนวคิดรวบยอดคือการเขียนโปรแกรมให้ทำงานซ้ำเป็นลูป แต่แนวคิดสัมพันธ์ในกลุ่มเดียวกันที่สามารถทำให้โปรแกรมทำงานวนลูปได้มี 3 แนวคิด ได้แก่ while, do...while และ for ซึ่งในระดับหนึ่งแนวคิดเหล่านี้มีความคล้ายกัน แต่หากพิจารณาในรายละเอียดแล้ว แต่ละแนวคิดก็มีเอกลักษณ์เฉพาะตัว การจะตอบคำถามที่ออกแบบในลักษณะได้ถูกต้องจึงต้องมีความเข้าใจแต่ละแนวคิดอย่างชัดเจนจนสามารถเห็นความเหมือนและความแตกต่างของแนวคิดสัมพันธ์เหล่านี้ได้

บทที่ 4

การทดลอง

เพื่อเป็นการทดสอบประสิทธิภาพของกรอบงานที่นำเสนอ งานวิจัยนี้จึงออกแบบการทดลองโดยการใช้กรอบงานที่นำเสนอเพื่อสร้างคำถาม 11 ข้อ ในโดเมนความรู้เกี่ยวกับการโปรแกรมคอมพิวเตอร์ โดยคำถามทั้ง 11 ข้อนี้ถูกนำไปถามนักศึกษาปีที่ 1 จำนวน 55 คนที่มีวิชาเอกวิศวกรรมคอมพิวเตอร์ ก่อนเรียนวิชานี้นักศึกษาส่วนใหญ่ไม่มีความรู้ทางด้านการโปรแกรมคอมพิวเตอร์มาก่อน มีเพียงนักศึกษา 3 คนเท่านั้นที่เคยเรียนการโปรแกรมคอมพิวเตอร์เบื้องต้นมาก่อนในระดับมัธยมศึกษาตอนปลาย ดังนั้นจึงสามารถสรุปได้ว่านักศึกษาส่วนใหญ่ในกลุ่มทดลอง มีเบื้องหลังในการศึกษาวิชาการโปรแกรมคอมพิวเตอร์ใกล้เคียงกัน และนักศึกษากลุ่มนี้ได้รับความรู้ทางการโปรแกรมคอมพิวเตอร์โดยใช้ภาษาซี (C programming language) ในชั้นเรียนโดยเสมอภาคกันจากผู้สอนคนเดียวกัน รายละเอียดของการสร้างคำถามทั้ง 11 ข้อที่เกี่ยวกับการโปรแกรมคอมพิวเตอร์โดยใช้กรอบงานที่นำเสนอ เป็นดังนี้

ขั้นตอนที่ 1 – การกำหนดภววิทยา: ภววิทยาที่กำหนดในขั้นตอนนี้อยู่ในโดเมนการโปรแกรมคอมพิวเตอร์ที่ประกอบด้วย 11 แนวคิด (concepts) ที่สามารถจัดกลุ่มได้เป็น 5 กลุ่ม สำหรับแนวคิดที่อยู่ในกลุ่มที่ 1 – 4 ในตารางที่ 4.1 เป็นแนวคิดที่มีความสัมพันธ์แบบ generalization และในกลุ่มที่ 5 เป็นแนวคิดสัมพัทธ์ที่เชื่อมโยงกันด้วยความสัมพันธ์แบบ composition

ตารางที่ 4.1 แสดงภววิทยาของโดเมนการโปรแกรมคอมพิวเตอร์ ประกอบด้วย 4 คอลัมน์ ได้แก่ คอลัมน์กลุ่ม (Group) หมายถึงกลุ่มแนวคิด (concept) แบ่งออกเป็น 4 กลุ่มคือ Groups 1 – 4 คอลัมน์ความคิดรวบยอด (Conceptual Idea) อธิบายกลุ่มแนวคิดที่อยู่ในโดเมนที่สนใจ ประกอบด้วยแนวคิดของการสร้างเงื่อนไข (conditional structure) แนวคิดการสร้างประโยค (sentence construction) แนวคิดโครงสร้างลูป (loop structure) แนวคิดวิธีการเรียกใช้ฟังก์ชัน (methods for calling a function) และแนวคิดอาร์เรย์ (array)

ตารางที่ 4.1: แนวคิดและความสัมพันธ์ระหว่างแนวคิดในโดเมนการโปรแกรมคอมพิวเตอร์[3]

Group	Conceptual Idea	Relative Concepts	Relation Type
1	Conditional structure	'if' and 'switch...case' statements	Generalization
2	Sentence construction	String (a variable type) and an array of characters	Generalization
3	Loop structure	'for', 'while', and 'do...while' structures	Generalization
4	Methods for calling a function	calling by value and calling by reference	Generalization
5	Array	pointer and address	Composition

คอลัมน์แนวคิดสัมพัทธ์ (Relative Concepts) ในกลุ่ม Group 1 แนวคิดสัมพัทธ์ประกอบด้วยแนวคิดของโครงสร้างเงื่อนไข if และ switch...case ในกลุ่ม Group 2 แนวคิดสัมพัทธ์ประกอบด้วยแนวคิดของการสร้างประโยคประกอบด้วย string และ array ของตัวอักษร (an array of characters) ในกลุ่ม Group 3 แนวคิดสัมพัทธ์ประกอบด้วยแนวคิดสำหรับสร้างลูปทำงานซ้ำประกอบด้วย for, while และ do...while ในกลุ่ม Group 4 แนวคิดสัมพัทธ์ประกอบแล้วแนวคิดการเรียกใช้งานฟังก์ชันแบบ calling by value และ calling by reference และสุดท้ายในกลุ่ม Group 5 แนวคิดสัมพัทธ์เกี่ยวกับ array ประกอบด้วยแนวคิดที่เป็นองค์ประกอบ ได้แก่ พอยท์เตอร์ (pointer) และแอดเดรส (address) คอลัมน์สุดท้ายเกี่ยวกับประเภทความสัมพันธ์ (Relation Type) สำหรับโดเมนนี้มี 2 แบบด้วยกันคือ แบบ generalization และแบบ composition

ขั้นตอนที่ 2 – การสร้างคำถาม: การสร้างประโยคคำถามประกอบด้วยแนวคิดการแก้ปัญหาเบื้องต้นและคำถามที่ให้นักเรียนค้นหาแนวทางการแก้ปัญหาอื่น ๆ ที่ให้ผลลัพธ์ที่เหมือนกัน ภาพที่ 4.1 เป็นตัวอย่างของการสร้างแนวทางการแก้ปัญหาเบื้องต้นโดยใช้ for ซึ่งเป็นแนวคิดใน

[3] ข้อมูลการทดลองในรายงานโครงการวิจัยนี้เป็นชุดเดียวกันกับข้อมูลที่ปรากฏในเอกสารวิจัยที่เผยแพร่ในงานประชุมวิชาการ Sirithumgul, Pornpat, Prasertsilp, Pimpaka, Suksa-Ngiam, Watanyoo & Olfman, Lorne. An Ontology-Based Framework as a Foundation of an Information System for Generating Multiple-Choice Questions. In Proceedings of the 25th Americas Conference on Information Systems, Cancun, Mexico. August 15 – 17th, 2019.

Group 3 จากนั้นมีประโยคคำถาม *"Which one in the following choice does give the same answer of the code given in Figure 7?"* หมายความว่า ตัวเลือกใดต่อไปนี้ให้คำตอบเดียวกับโค้ดที่ให้ในภาพ Figure 7

Question 7: Which one in the following choice does give the same answer of the code given in Figure 7?

```c
#include <stdio.h>
void main()
{
    int a;
    for(a=10; a<=19; a++){
        printf("a = %d \n", a);
    }
}
```

Figure 7

ภาพที่ 4.1: ตัวอย่างการสร้างประโยคคำถามที่ประกอบด้วยคำตอบเริ่มต้นและประโยคคำถามให้ค้นหาคำตอบโดยวิธีอื่น

ขั้นตอนที่ 3 – การสร้างตัวเลือกของคำตอบที่ถูกต้อง: ในขั้นตอนนี้เป็นการนำแนวคิดสัมพัทธ์มาใช้เพื่อสร้างคำตอบที่ถูกต้อง ดังแสดงในภาพที่ 4.2 เป็นตัวอย่างของการใช้แนวคิดสัมพัทธ์ do...while เพื่อสร้างอีกทางเลือกของคำตอบที่ให้ผลลัพธ์เหมือนกับการใช้แนวคิด for ที่ปรากฏในประโยคคำถาม

```c
#include <stdio.h>
void main()
{
    int a = 10;
    do{
        printf("a = %d \n", a);
        a += 1;
    }while(a < 20);
}
```

ภาพที่ 4.2: ตัวอย่างการสร้างตัวเลือกของคำตอบที่ถูกต้อง

ขั้นตอนที่ 4 – การสร้างตัวลวง: แนวคิดสัมพัทธ์ในกลุ่มเดียวกันถูกนำมาสร้างตัวลวงในขั้นตอนนี้ ดังแสดงในภาพที่ 4.3 แนวคิด while ถูกนำมาสร้างตัวลวง หากพิจารณาความคิดรวบยอด while สามารถสั่งให้คอมพิวเตอร์ทำงานแบบวนลูปได้เหมือนกับคำสั่ง for ในประโยคคำถามและคำสั่ง do...while ในตัวเลือกของคำถามที่ถูกต้อง อย่างไรก็ดี ตัวลวงที่ปรากฏในภาพที่

4.3 มีข้อบกพร่องของการนำคำสั่ง while ไปใช้ทำให้ผลของการทำงานของคอมพิวเตอร์แตกต่างกันออกไป กล่าวโดยสรุป การสร้างตัวลวงในขั้นตอนนี้ แท้จริงแล้วคือการใช้แนวคิดสัมพัทธ์ในการสร้างวิธีการแก้ปัญหาแต่มีการเพิ่มข้อผิดพลาด (error) แฝงเข้าไปในการประยุกต์ใช้แนวคิดด้วย

```
1  #include <stdio.h>
2  void main()
3  {
4      int a = 10;
5      while(a < 20){
6          a = a + 1;
7          printf("a = %d \n", a);
8      }
9  }
```

ภาพที่ 4.3: ตัวอย่างการสร้างตัวเลือกของคำตอบที่เป็นตัวลวง

ตารางที่ 4.2 เป็นข้อมูลสรุปเกี่ยวกับคำถามที่สร้างขึ้น 11 ข้อ (Q1 - Q11) ในการทดลองนี้ ตารางแสดงข้อมูลของคำถาม 11 ข้อใน 5 คอลัมน์ ประกอบด้วย

1) ความคิดรวบยอด (Conceptual Idea) ของแนวคิดที่ใช้เพื่อสร้างคำถาม
2) คำถาม (Questions) ประกอบด้วยคำถาม Q1 – Q11
3) แนวคิดที่ใช้เพื่อสร้างแนวทางการแก้ปัญหาเบื้องต้น (Initial Solutions)
4) แนวคิดที่ใช้เพื่อสร้างตัวเลือกของคำตอบที่ถูกต้อง (Correct Answers)
5) แนวคิดที่ใช้เพื่อสร้างตัวเลือกที่เป็นตัวลวง (Distractors)

ตัวอย่างของข้อมูลคำถามเช่น คำถามข้อ Q1 ใช้แนวคิด switch...case ในการสร้างแนวทางการสร้างคำตอบเบื้องต้น ส่วนคำตอบในตัวเลือกที่ถูกต้องใช้แนวคิด if และ แนวคิด switch...case และ if ถูกนำมาสร้างตัวลวงที่แฝงข้อผิดพลาดไว้ในตัวลวงใน 3 ตัวเลือก

ตารางที่ 4.2: ข้อมูลสรุปเกี่ยวกับคำถาม Q1 - Q11 ที่สร้างขึ้น [4]

Conceptual Idea	Questions	Initial Solutions	Correct Answers	Distractors
Conditional structure	Q1	switch…case	if	switch…case, if
	Q2	if	switch…case	switch…case, if
Sentence construction	Q3	String (variable type)	character array	String and character array
	Q4	character array	String (variable type)	String and character array
Loop structure	Q5	for	while	for, while, do…while
	Q6	while	for	for, while, do…while
	Q7	do…while	for	for, while, do…while
Methods for calling a function	Q8	calling by value	calling by reference	calling by value, calling by reference
	Q9	calling by reference	calling by value	calling by value, calling by reference
Array	Q10	pointer	address	pointer, address
	Q11	address	point	pointer, address

[4] ข้อมูลการทดลองในรายงานโครงการวิจัยนี้เป็นชุดเดียวกันกับข้อมูลที่ปรากฏในเอกสารวิจัยที่เผยแพร่ในงานประชุมวิชาการ Sirithumgul, Pornpat, Prasertsilp, Pimpaka, Suksa-Ngiam, Watanyoo & Olfman, Lorne. An Ontology-Based Framework as a Foundation of an Information System for Generating Multiple-Choice Questions. In Proceedings of the 25th Americas Conference on Information Systems, Cancun, Mexico. August 15 – 17th, 2019.

บทที่ 5

ผลการทดลอง

ผลลัพธ์จากการทดลองคือ คำถามที่สร้างขึ้นตามแนวทางของกรอบงานที่นำเสนอทั้ง 11 คำถาม (Q1 – Q11) ถูกนำไปทดสอบกับนักศึกษา 55 คน เพื่อประเมินคุณภาพของคำถามใน 3 ด้าน ได้แก่ คุณภาพของตัวลวงซึ่งถูกพิจารณาจากการปรากฏของตัวเลือกที่ไม่ทำงาน (non-functioning distractor) ความยากของคำถามโดยใช้ดัชนีวัดความยาก (Difficulty Index หรือ P-scores) และความสามารถของคำถามในการแบ่งแยกนักเรียนที่เรียนดีและเรียนอ่อนออกจากกันที่เรียกว่า ดัชนีอำนาจจำแนก (Discrimination Index หรือ D-scores)

1. **คุณภาพของตัวลวงซึ่งถูกพิจารณาจากการปรากฏของตัวเลือกที่ไม่ทำงาน** ดังที่ได้แนะนำในงานวิจัยก่อนหน้า (Haladyna and Downing 1993; Cizek and O'Day 1994; Sidick et al. 1994; Delgado and Prieto 1998; Roger and Harley 1999; Woodford and Bancroft 2005; Shizuka et al. 2006; Tarrant et al. 2009) ได้ระบุว่า ***"ตัวลวงที่ไม่ทำงาน"*** (non-functioning distractor) หมายถึงตัวลวงที่ถูกเลือกน้อยกว่า 5 เปอร์เซนต์ของจำนวนนักเรียนทั้งหมดที่ตอบคำถาม คำถามแบบปรนัยจะถูกพิจารณาว่ามีคุณภาพถ้า 1 ใน 3 ของตัวลวงไม่เป็นตัวเลือกที่ไม่ทำงาน

   เฉพาะสำหรับการทดลองในงานวิจัยนี้ เนื่องจากมีนักเรียนทั้งหมด 55 คน ตัวลวงที่ไม่ทำงานจึงหมายถึง ตัวลวงที่มีนักเรียนเลือกน้อยกว่า 3 คน (คือ 5 เปอร์เซ็นต์ของ 55 คน) ผลลัพธ์ออกมาแสดงให้เห็นว่าทั้ง 11 คำถามน่าเชื่อถือเนื่องจากทุกคำถามมีอย่างน้อย 1 ตัวลวงที่ไม่เป็นตัวเลือกที่ไม่ทำงานซึ่งสามารถจำแนกได้ว่า 8 ใน 11 คำถาม ไม่มีตัวเลือกที่ไม่ทำงานปรากฏอยู่เลย 2 ใน 11 คำถามปรากฏ 1 ตัวลวงที่ไม่เป็นตัวเลือกที่ไม่ทำงาน และ 1 ใน 11 คำถามปรากฏ 2 ตัวลวงที่ไม่เป็นตัวเลือกที่ไม่ทำงาน

   จึงสามารถสรุปได้ว่า คุณภาพของตัวลวงที่อยู่ในคำถามที่สร้างอยู่บนพื้นฐานของกรอบงานที่เสนอมีประสิทธิภาพใช้งานได้ (valid) อีกทั้ง ตัวลวงที่ไม่เป็นตัวเลือกที่ไม่ทำงานสามารถนำมาใช้เพื่อวิเคราะห์ความเข้าใจผิดของนักเรียนที่ตอบคำถามได้ ซึ่งหากตีความโดยหลักของภววิทยา (ontology) แล้วมีความเป็นไปได้สูงว่าความสับสนของนักเรียนเกิดขึ้นระหว่างแนวคิดสัมพัทธ์ที่อยู่ในโดเมนความรู้เดียวกัน

2. **ความยากของคำถามที่สร้างขึ้นโดยกรอบงานที่นำเสนอ** การวัดความยากของคำถามจะใช้ดัชนี P-scores ซึ่งก็คือ อัตราส่วนระหว่างจำนวนนักเรียนที่อยู่ในกลุ่มเรียนดีที่ตอบคำถามถูกต้องต่อจำนวนนักเรียนในกลุ่มเรียนอ่อนที่ตอบคำถามถูกต้อง

จากการทดลองพบว่า 6 คนจาก 55 คนได้รับเกรด A และ B^+ จากวิชาการโปรแกรมคอมพิวเตอร์ นักเรียนทั้ง 6 คนนี้จัดได้ว่าอยู่ในกลุ่มนักเรียนเรียนดี และนักเรียน 11 คนจาก 55 คนได้รับเกรด F จากวิชานี้ จึงกล่าวได้ว่า นักเรียนทั้ง 11 คนนี้จัดได้ว่าเป็นกลุ่มนักเรียนที่เรียนอ่อน ดังนั้นในการคำนวณหาค่า P-scores จึงใช้คะแนนของนักเรียนทั้งสองกลุ่ม คือ กลุ่มเรียนดีและเรียนอ่อนทั้ง 11 คนในการคำนวณ ผลลัพธ์การคำนวณแสดงในตารางที่ 5.1

ตารางที่ 5.1: จำนวนของคำถามที่ปรากฏใน 3 ระดับความยาก[5]

Difficulty Index (P-scores)		
Difficult ($P \leq 30$)	Moderate ($30 < P < 70$)	Easy ($P \geq 70$)
2	6	3

งานวิจัยของ Pande et al. (2013) กล่าวว่าระดับความยากควรแบ่งออกเป็น 3 ระดับ ได้แก่ ระดับยาก (difficult) มีค่า P-scores น้อยกว่าหรือเท่ากับ 30 ระดับกลาง (moderate) มีค่า P-scores ระหว่าง 30 และ 70 และระดับง่าย (easy) มีค่า P-scores มากกว่าหรือเท่ากับ 70 ในการทดลองพบว่า คำถามที่สร้างขึ้นมี 2 ข้อจาก 11 ข้ออยู่ในระดับยาก 6 ข้อจาก 11 ข้ออยู่ในระดับกลาง และ 3 ข้อจาก 11 ข้ออยู่ในระดับง่าย

ผลลัพธ์จากการทดลองสามารถสรุปได้ว่า 6 คำถามในระดับกลาง เหมาะสมที่สุดที่จะใช้เป็นคำถามมาตรฐานสำหรับวัดความรู้ในโดเมนการโปรแกรมคอมพิวเตอร์ เพราะเป็นกลุ่มคำถามที่ไม่ยากและไม่ง่ายจนเกินไป อย่างไรก็ตาม ความเชื่อถือได้ของคำถามกลุ่มนี้จะมีค่าสูงสุดเมื่อถูกใช้วัดความรู้ของนักเรียนที่มีพื้นหลังการเรียนคล้ายกันกับนักศึกษาในกลุ่มที่อยู่ในการทดลองของงานวิจัยนี้ เช่น มีการเรียนวิชาการโปรแกรม

[5] ข้อมูลการทดลองในรายงานโครงการวิจัยนี้เป็นชุดเดียวกันกับข้อมูลที่ปรากฏในเอกสารวิจัยที่เผยแพร่ในงานประชุมวิชาการ Sirithumgul, Pornpat, Prasertsilp, Pimpaka, Suksa-Ngiam, Watanyoo & Olfman, Lorne. An Ontology-Based Framework as a Foundation of an Information System for Generating Multiple-Choice Questions. In Proceedings of the 25th Americas Conference on Information Systems, Cancun, Mexico. August 15 – 17th, 2019.

คอมพิวเตอร์มาก่อน และ/หรือเป็นนักเรียน/นักศึกษาที่เรียนวิชาเอกในสาขาวิศวกรรมคอมพิวเตอร์ เป็นต้น

คำถามที่ถูกประเมินให้อยู่ในกลุ่มยากและง่ายนั้นมีความสมเหตุสมผล เมื่อถูกนำไปใช้เพื่อสะท้อนผลการเรียนของนักเรียนในแง่ของความเข้าใจของนักเรียนที่เกี่ยวเนื่องกับแนวคิดในการเรียน กล่าวคือ นักเรียนมีแนวโน้มที่จะมีความรู้ความเข้าใจดีสำหรับแนวคิดที่ถามโดยคำถามที่อยู่ในระดับง่าย และยังขาดความรู้ความเข้าใจสำหรับแนวคิดที่ถามโดยคำถามที่อยู่ในระดับยาก (Pande et al. 2013)

3. **ดัชนีอำนาจจำแนก (discrimination index หรือ D-scores)** ถูกนำมาใช้เพื่อวัดความสามารถของคำถามในการแบ่งแยกกลุ่มนักเรียนเรียนดีและเรียนอ่อนออกจากกัน D-scores คืออัตราส่วนของความแตกต่างระหว่างจำนวนนักเรียนในกลุ่มเรียนดีและเรียนอ่อนที่สามารถตอบคำถามได้ถูกต้องต่อจำนวนนักศึกษาทั้งสองกลุ่ม

ข้อมูลของนักศึกษา 17 คน ประกอบด้วยนักเรียนกลุ่มเรียนดี 6 คน และกลุ่มเรียนอ่อน 11 คนที่ใช้เพื่อการคำนวณ P-scores ถูกนำมาใช้เพื่อคำนวณ D-scores ด้วย ข้อมูล D-scores ดังแสดงในตารางที่ 5.2 แบ่งออกเป็น 3 ช่วง ได้แก่ ระดับต่ำ (minimum) ระดับกลาง (moderate) และระดับดีเยี่ยม (exceptional)

ตารางที่ 5.2: จำนวนของคำถามที่ปรากฏใน 3 ระดับของดัชนีอำนาจจำแนก[6]

Discrimination Index (D-scores)		
Minimum $(0 < D \leq 0.19)$	Moderate $(0.20 \leq D \leq 0.29)$	Exceptional $(D \geq 0.30)$
4	2	5

ในเอกสารวิจัยของ Pande et al. (2013) ระบุว่าคำถามหนึ่งจะเชื่อถือได้เมื่อมีค่า D-scores มากกว่าศูนย์ ซึ่งถูกตีความได้ว่านักเรียนในกลุ่มเรียนดีสามารถตอบคำถามได้ถูกต้องมากกว่านักเรียนในกลุ่มเรียนอ่อน ในทางกลับกันคำถามหนึ่งจะมีความน่าเชื่อถือน้อยและถูกมองว่าผิดปกติ (anomaly) เมื่อมีค่า D-scores น้อยกว่าหรือเท่ากับศูนย์ ซึ่งสามารถ

[6] ข้อมูลการทดลองในรายงานโครงการวิจัยนี้เป็นชุดเดียวกันกับข้อมูลที่ปรากฏในเอกสารวิจัยที่เผยแพร่ในงานประชุมวิชาการ Sirithumgul, Pornpat, Prasertsilp, Pimpaka, Suksa-Ngiam, Watanyoo & Olfman, Lorne. An Ontology-Based Framework as a Foundation of an Information System for Generating Multiple-Choice Questions. In Proceedings of the 25th Americas Conference on Information Systems, Cancun, Mexico. August 15 – 17th, 2019.

ตีความได้ว่าคำถามหนึ่งจะมีประสิทธิภาพน้อย เมื่อนักเรียนในกลุ่มเรียนดีและเรียนอ่อนสามารถตอบคำถามข้อนั้นได้ถูกต้องพอ ๆกัน หรือมีจำนวนนักเรียนในกลุ่มเรียนอ่อนมากกว่าจำนวนนักเรียนเรียนดีที่สามารถตอบคำถามข้อนั้นได้

ค่า D-scores ในตารางที่ 5.2 แสดงว่าคำถามทั้ง 11 ข้อ (Q1 – Q11) ที่สร้างจากกรอบงานที่นำเสนอมีความเชื่อถือได้ โดยคำถาม 4 ข้อมีค่า D-scores มากกว่า 0 แต่ไม่เกิน 0.19 ซึ่งหมายความว่า มีความแตกต่างในระดับน้อย ระหว่างจำนวนนักเรียนในกลุ่มเรียนดีและนักเรียนในกลุ่มเรียนอ่อนที่ตอบคำถามได้อย่างถูกต้อง

มีอยู่ 2 คำถามที่ค่า D-scores มากกว่าหรือเท่ากับ 0.20 แต่น้อยกว่าหรือเท่ากับ 0.29 ซึ่งหมายความว่า มีความแตกต่างในระดับกลาง ระหว่างจำนวนนักเรียนในกลุ่มเรียนดีและนักเรียนในกลุ่มเรียนอ่อนที่ตอบคำถามได้อย่างถูกต้อง

สุดท้าย มีอยู่ 5 คำถามที่มีค่า D-scores มากกว่าหรือเท่ากับ 0.30 ซึ่งหมายความว่า มีความแตกต่างในระดับสูง ระหว่างจำนวนนักเรียนในกลุ่มเรียนดีและนักเรียนในกลุ่มเรียนอ่อนที่ตอบคำถามได้อย่างถูกต้อง

บทที่ 6

บทสรุปงานวิจัย

บทนี้จะกล่าวถึงประเด็นถกเถียง (discussion) ที่ใช้เพื่อการตีความผลที่ได้รับจากงานวิจัย การพัฒนางานวิจัยต่อเนื่องในอนาคต (future work) และบทสรุปงานวิจัย (conclusion) ดังนี้

6.1 ประเด็นถกเถียงและการพัฒนางานวิจัยต่อเนื่องในอนาคต

ประเด็นที่ใช้วิเคราะห์ถกเถียงในที่นี้เกี่ยวข้องกับระดับการเรียนของนักเรียน การนำความรู้ที่ได้จากงานวิจัยไปใช้ และการตีความงานวิจัยนี้เพื่อนำไปใช้ในอนาคต

6.1.1 การวิเคราะห์ความสัมพันธ์ระหว่างดัชนีสองตัวได้แก่ discrimination index (D-scores) และ difficulty index (P-scores) เป็นดังตารางที่ 6.1

ตารางที่ 6.1: ความสัมพันธ์ของข้อมูล D-scores และ P-scores[7]

Discrimination Index (D-scores)	Difficulty Index (P-scores)		
	Difficult	Moderate	Easy
Minimum	Q9	Q10	Q3, Q8
Moderate	Q1	Q2	-
Exceptional	-	Q5, Q6, Q7, Q11	Q4

ตารางที่ 6.1 แสดงให้เห็นความสัมพันธ์ระหว่างความยากของคำถามที่วัดโดยดัชนี P-scores และความสามารถของคำถามในการแบ่งแยก (discrimination) นักเรียนที่มีผลการเรียนดีและเรียนอ่อนออกจากกันที่วัดโดยดัชนี D-scores ทั้งนี้ ดังแสดงในตารางที่ 6.1 จะเห็นว่าระดับความยากของคำถามถูกแบ่งเป็น 3 ช่วงคือ ระดับยาก (difficult) ระดับกลาง (moderate) และระดับง่าย (easy)

[7] ข้อมูลที่ใช้เพื่อการวิเคราะห์ในรายงานโครงการวิจัยนี้เป็นชุดเดียวกันกับข้อมูลที่ปรากฏในเอกสารวิจัยที่เผยแพร่ในงานประชุมวิชาการ Sirithumgul, Pornpat, Prasertsilp, Pimpaka, Suksa-Ngiam, Watanyoo & Olfman, Lorne. An Ontology-Based Framework as a Foundation of an Information System for Generating Multiple-Choice Questions. In Proceedings of the 25th Americas Conference on Information Systems, Cancun, Mexico. August 15 – 17th, 2019.

และระดับการแบ่งแยกของคำถาม 3 ช่วง ได้แก่ ระดับน้อย (minimum) ระดับกลาง (moderate) และระดับดีเยี่ยม (exceptional) ความหมายของข้อมูลในตารางเป็นดังนี้

1) **คำถามที่มีความยากในระดับกลาง และระดับการแบ่งแยกอยู่ในระดับกลางไปจนถึงระดับดีเยี่ยมควรถูกใช้เป็นคำถามมาตรฐานที่ใช้สำหรับวัดความรู้** คำถาม Q2, Q5 – Q7, และ Q11 ที่ถูกสร้างตามแนวทางของกรอบงานที่งานวิจัยนำเสนอ ถือว่าเป็นคำถามในกลุ่มนี้ นอกจากนี้ หากพิจารณาในอีกมุมหนึ่ง คำถามเหล่านี้สามารถใช้ในการวัดเพื่อแบ่งกลุ่มนักเรียนที่ประสบความสำเร็จมากและน้อยออกจากกันได้

2) **จุดอ่อนของนักเรียนกลุ่มเรียนดีสามารถสังเกตได้จากคำถามยากที่มีความสามารถในแบ่งแยกกลุ่มนักเรียนที่อยู่ในระดับน้อยถึงปานกลาง** คำถาม Q1 และ Q9 ในตารางที่ 6.1 อยู่ในกลุ่มนี้ การที่ค่า D-scores ของคำถามอยู่ในระดับกลาง - ต่ำ หมายถึงคำถามเหล่านี้มีความยาก และไม่สามารถจะแบ่งแยกกลุ่มของนักเรียนที่มีผลการเรียนดี – ไม่ดีได้อย่างชัดเจน อย่างไรก็ดี การที่นักเรียนที่มีระดับผลการเรียนดีส่วนใหญ่ไม่สามารถตอบคำถามเหล่านี้ได้ แสดงให้เห็นว่า คำถามในกลุ่มนี้แสดงให้เห็นถึงจุดอ่อนของนักเรียนในกลุ่มเรียนดี ซึ่งมีความเป็นไปได้ว่านักเรียนที่มีระดับผลการเรียนดีอาจจะยังมีความสับสนกับแนวคิดที่ถูกถามโดยคำถามข้อ Q1 และ Q9

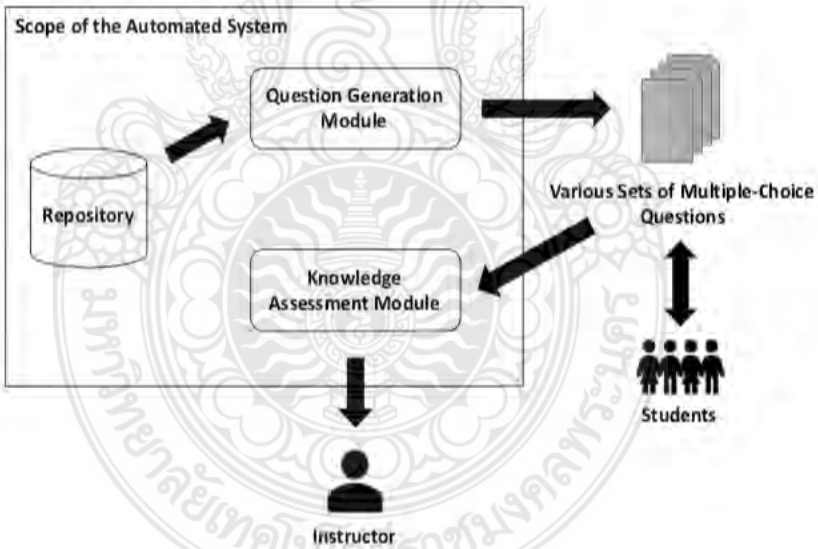
3) **จุดอ่อนของนักเรียนกลุ่มเรียนอ่อนสามารถสังเกตได้จากคำถามที่อยู่ในระดับง่ายและมีความสามารถในการแบ่งแยกกลุ่มนักเรียนอยู่ในระดับดีเยี่ยม** คำถาม Q4 อยู่ในเงื่อนไขข้อนี้ ซึ่งหมายความว่าการที่คำถามสามารถแบ่งแยกกลุ่มนักเรียนเรียนดีและเรียนอ่อนได้อย่างชัดเจนทั้งที่เป็นคำถามระดับง่าย แสดงให้เห็นว่า นักเรียนที่เรียนอ่อนอาจจะยังมีอุปสรรคในการทำความเข้าใจแนวคิดที่ถามโดยคำถามที่อยู่ในเงื่อนไขข้อนี้

4) **จุดเด่นหรือการประสบความสำเร็จในการเรียนของนักเรียนส่วนใหญ่ให้สังเกตจากคำถามระดับกลาง – ง่ายและไม่สามารถแบ่งแยกนักเรียนกลุ่มเรียนดีและเรียนอ่อนออกจากกันได้** ในที่นี้หมายถึงคำถาม Q3, Q8 และ Q10 ในตาราง 6.1 การที่คำถามสามารถแบ่งแยกกลุ่มเรียนดีและเรียนอ่อนได้ในระดับต่ำ ทั้งที่คำถามมีความยากอยู่ในระดับกลางถึงง่าย สามารถตีความได้ว่านักเรียนทั้งสองกลุ่ม – ทั้งกลุ่มเรียนดีและเรียนอ่อน มีแนวโน้มที่จะเข้าใจแนวคิดที่ถามโดยคำถามเป็นอย่างดี

6.1.2 ข้อสรุปโดยทั่วไปของการค้นพบในงานวิจัย ข้อสรุปโดยทั่วไปของการนำงานวิจัยไปใช้ กล่าวได้ว่า คำถาม Q1 – Q11 ที่สร้างขึ้นในงานวิจัยนี้สามารถนำไปใช้ซ้ำเพื่อวัดความรู้ได้ หากนักเรียนมีพื้นฐานความรู้ที่เรียนมาในระดับเดียวกัน แต่สิ่งที่ควรพิจารณานำไปประยุกต์ใช้ แท้จริงแล้วคือ กรอบ

งานที่นำเสนอสำหรับสร้างคำถาม และแนวความคิดวิเคราะห์ผลการเรียน ดังที่แสดงในตารางที่ 6.1 เพื่อใช้ประโยชน์ในการพิจารณาความสามารถในการเรียนรู้ของนักเรียน หลักการสร้างกรอบงานและตารางพิจารณาความสามารถในการเรียนรู้จึงถือว่ามี *"คุณสมบัติของหลักการโดยทั่วไป"* (generalizability) ที่สามารถนำไปใช้ได้กับทุกโดเมนความรู้ ในทุกศาสตร์สาขาวิชา

6.1.3 การพัฒนางานวิจัยต่อเนื่องในอนาคต เนื่องจากผลการทดลองแสดงให้เห็นแล้วว่ากรอบงานที่นำเสนอมีแนวโน้มของการนำไปใช้ได้จริง ในฐานะที่เป็นกลไกปัญญาประดิษฐ์ที่อยู่ภายในระบบสารสนเทศ (information system) ตามภาพที่ 6.1 แสดงสถาปัตยกรรมระดับบน (high-level architectural design) ของระบบสารสนเทศสำหรับสร้างคำถามและประเมินผลการเรียนอัตโนมัติในฐานะงานในอนาคตที่สืบเนื่องจากงานวิจัยนี้ โดยขอบเขตของระบบอัตโนมัติ (scope of the automated system) ที่แสดงในภาพประกอบด้วย คลังข้อมูล (repository) มอดูลสำหรับสร้างคำถาม (question generation module) และมอดูลสำหรับประเมินความรู้ (knowledge assessment module)



ภาพที่ 6.1: สถาปัตยกรรมระดับบนของระบบสารสนเทศสำหรับสร้างคำถามและประเมินผลการเรียนอัตโนมัติ

องค์ประกอบของระบบอัตโนมัติที่นำเสนอมีดังต่อไปนี้

1) ***คลังข้อมูล*** ใช้สำหรับเก็บส่วนประกอบของคำถามแบบปรนัยทั้งหมด ได้แก่ คำตอบเบื้องต้น หรือแนวทางการแก้ปัญหาเบื้องต้น (initial answer) ในประโยคคำถาม คำตอบของตัวเลือกที่ถูกต้อง (correct answer choice) และตัวลวง (distractor) หรือคำตอบที่ผิด (incorrect answer) อัน

ที่จริงแล้ว คำตอบที่ถูกต้องและแนวทางการแก้ปัญหาเบื้องต้นสามารถรวมกลุ่มเป็นข้อมูลชุดเดียวกัน ในการสร้างคำถามแบบปรนัยแบบสุ่มเลือกข้อมูล มีความเป็นไปได้ที่คำตอบของตัวเลือกที่ถูกต้องในคำถามข้อหนึ่งสามารถใช้เป็นแนวทางการแก้ปัญหาเบื้องต้นของคำถามอีกข้อหนึ่งได้

2) ***มอดูลสำหรับสร้างคำถาม*** เป็นส่วนการทำงานที่สืบค้น (retrieve) ข้อมูลจากคลังข้อมูลเพื่อสร้างคำถามตามรูปแบบของกรอบงาน (framework) ที่งานวิจัยนี้นำเสนอ ผลลัพธ์ของมอดูลสำหรับสร้างคำถามนี้จะได้เซตของคำถามแบบปรนัยที่มีอยู่หลากหลายรูปแบบ (various sets of multiple-choice questions) เซตของคำถามเหล่านี้จะถูกนำไปใช้เพื่อวัดความรู้ของนักเรียน (students) ตามที่ปรากฏในภาพที่ 6.1

3) ***มอดูลสำหรับประเมินความรู้*** เป็นส่วนวิเคราะห์คำตอบที่นักเรียนที่ตอบคำถามที่มอดูลสำหรับสร้างคำถามสร้างขึ้น มอดูลสำหรับประเมินความรู้นี้มีโครงสร้างการวิเคราะห์ข้อมูลเหมือนกับตารางเมตริกที่แสดงในตารางที่ 6.1 ที่มีโครงสร้างหลักเป็นดัชนี D-scores และ P-scores สำหรับวิเคราะห์ความสามารถการเรียนรู้ของนักเรียนในกลุ่มที่เรียนดี และเรียนอ่อน รวมถึงมีการสรุปการประมวลผลให้กับผู้สอน (instructor) ได้รับทราบอีกด้วย

6.1 บทสรุป

งานวิจัยนี้เสนอกรอบงานสำหรับสร้างคำถามที่มีตัวเลือกแบบปรนัย การทดลองในงานวิจัยนี้สาธิตให้เห็นการสร้างคำถามในโดเมนวิชาการโปรแกรมคอมพิวเตอร์โดยใช้กรอบงานที่นำเสนอ ดังนั้นอีกนัยหนึ่ง กรอบงานที่นำเสนอนี้จัดได้ว่าเป็นเครื่องมือสำหรับวัดความรู้และความเข้าใจของแนวคิดในบทเรียนที่นักเรียนได้เรียน ท้ายที่สุด ในงานวิจัยนี้มีการนำเสนอวิธีการนำกรอบงานไปใช้เป็นกลไกพื้นฐานของระบบสารสนเทศที่สามารถสร้างคำถามและประมวลผลความรู้ของนักเรียนได้อย่างอัตโนมัติ

บรรณานุกรม

Massive open online course. (n.d.). In Wikipedia. Retrieved March 28, 2018, from https://en.wikipedia.org/wiki/Massive_open_online_course

Coursera. (n.d.). In Wikipedia. Retrieved March 28, 2018, from https://en.wikipedia.org/wiki/Coursera

Budanitsky, A., & Hirst, G. (2001). Semantic distance in WordNet: An experimental, application-oriented evaluation of five measures. In Proceedings of the Workshop on WordNet and Other Lexical Resources, Second Meeting of the North American Chapter of the Association for Computational Linguistics (NAACL-2001), pp. 29–24, Pittsburgh, PA.

D. Bollegala, Y. Matsuo, and M. Ishizuka: Measuring semantic similarity between words using Web search engines, Proceedings of the 16th WWW, Banff, Alberta, Canada, May 8 – 12, 2007, pp 757 – 766.

Sirithumgul, Pornpat; Olfman, Lorne. A model for measuring knowledge constructions of students in online discussions. Proceedings of the 24th Australasian Conference on Information Systems (2013): RMIT University

Ragonis, N. (2012). Type of questions – the case of computer science. Olympiads in Informatics, 6, 115–132.

K. Woodford, P. Bancroft, "Multiple choice questions not considered harmful," Proceedings of the 7th Australasian conference on Computing education, p.109-116, January 01, 2005, Newcastle, New South Wales, Australia.

Tarrant, M., Ware, J., & Mohammed, A. M. (2009). An assessment of functioning and nonfunctioning distractors in multiple-choice questions: A descriptive analysis. BMC British Medical Education, 9, 40. http://dx.doi.org/10.1186/1472-6920-9-40

Cizek GJ, O'Day DM: Further investigations of nonfunctioning options in multiple-choice test items. Educ Psychol Meas 1994, 54(4):861-872.

Haladyna TM, Downing SM: How many options is enough for a multiple-choice test item? Educ Psychol Meas 1993, 53(4):999-1010.

Delgado AR, Prieto G: Further evidence favoring three-option items in multiple-choice tests. Eur J Psychol Assessment 1998, 14(3):197-201.

Shizuka T, Takeuchi O, Yashima T, Yoshizawa K: A comparison of three- and four-option English tests for university entrance selection purposes in Japan. LangT 2006, 23(1):35-57.

Sidick JT, Barrett GV, Doverspike D: Three-alternative multiple choice tests: An attractive option. Pers Psychol 1994, 47(4):829-835.

Rogers WT, Harley D: An empirical comparison of three- and four-choice items and tests: susceptibility to testwiseness and internal consistency reliability. Educ Psychol Meas 1999, 59(2):234- 247.

Pande SS, Pande SR, Parate VR, Nikam AP, Agrekar SH. Correlation between difficulty and discrimination indices of MCQs in formative examination in Physiology. South East Asian Journal of Medical Education 2013; 7(1):45-50.

Rao C, Kishan Prasad HL, Sajitha K, Permi H, Shetty J. Item analysis of multiple choice questions: assessing an assessment tool in medical students. Int J EducPsychol Res. 2016; 2:201-4.

Forehand, M. (2005). Bloom's taxonomy: Original and revised. In M. Orey (Ed.), Emerging perspectives on learning, teaching, and technology. Retrieved <insert date>, from http://projects.coe.uga.edu/epltt/

H. Kunichika, T. Katayama, T. Hirashima, and A. Takeuchi. 2004. Automated question generation methods for intelligent English learning systems and its evaluation. Proceedings of ICCE2004, Hong Kong.

Michael Heilman and Noah A. Smith. Good question! Statistical ranking for question generation. In Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the ACL, pages 609–617, Stroudsburg, PA, USA, 2010. ACL.

Brown, J., Frishkoff, G., Eskenazi, M. 2005. Automatic Question Generation for Vocabulary Assessment. Proceedings of the Annual Human Language Technology Meeting: 249-254.

R. Mitkov, L. A. Ha, and N. Karamanis. 2006. A computer-aided environment for generating multiple choice test items. Natural Language Engineering, 12(2).

Goto, T., Kojiri, T., Watanabe, T., Iwata, T., & Yamada, T. (2010). "Automatic Generation System of Multiple choice Cloze Questions and its Evaluation", Knowledge Management & E-Learning: An International Journal (KM&EL), Vol 2, No 3, 2010

Girish Kumar, Rafael E. Banchs, and Luis F. D'Haro. 2015. RevUP: Automatic gap-fill question generation from educational texts. In Proceedings of the 10th Workshop on Innovative Use of NLP for Building Educational Applications, pages 154–161.

Agarwal, M and Mannem, P 2011, "Automatic gap-fill question generation from text books," Proceedings of the 6th Workshop on Innovative Use of NLP for Building Educational Applications, pp. 56–64, Portland, Oregon, USA.

Simon Smith, P.V.S Avinesh and Adam Kilgarriff. 2010. Gap-fill Tests for Language Learners: Corpus-Driven Item Generation.

T. Alsubait, B. Parsia and U. Sattler, Ontology-based multiple choice question generation, Künstliche Intelligenz 30(2) (2016), 183–188.

C. Kwankajornkiet, A. Suchato and P. Punyabukkana, "Automatic multiple-choice question generation from Thai text," 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), Khon Kaen, 2016, pp. 1-6.

ภาคผนวก

ผลของโครงการวิจัยที่รายงานอย่างเต็มรูปแบบในเอกสารฉบับนี้ ส่วนหนึ่งได้ถูกตีพิมพ์และเผยแพร่ในงานประชุมวิชาการระดับนานาชาติ โดยสามารถอ้างอิงเอกสารที่เผยแพร่ได้จาก ***Sirithumgul, Pornpat***, Prasertsilp, Pimpaka, Suksa-Ngiam, Watanyoo & Olfman, Lorne. ***An Ontology-Based Framework as a Foundation of an Information System for Generating Multiple-Choice Questions***. In Proceedings of the 25th Americas Conference on Information Systems, Cancun, Mexico. August 15 – 17th, 2019.

ประวัติผู้วิจัย

1. ชื่อ-สกุล นางสาวพรภัทร์ ศิริธรรมกุล

2. ตำแหน่งทางวิชาการ อาจารย์

3. ตำแหน่งปัจจุบัน อาจารย์ประจำสาขาวิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์
มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร

4. วุฒิการศึกษา

ระดับการศึกษา	ชื่อวุฒิ	วิชาเอก	สถาบันการศึกษา	ปีที่สำเร็จการศึกษา
ปริญญาตรี	วศ.บ.	วิศวกรรมคอมพิวเตอร์	สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง	2548
ปริญญาโท	วท.ม.	วิศวกรรมซอฟต์แวร์	จุฬาลงกรณ์มหาวิทยาลัย	2552
ปริญญาโท	M.S.	Information Systems and Technology	Claremont Graduate University, USA	2557
ปริญญาเอก	Ph.D.	Information Systems and Technology	Claremont Graduate University, USA	2559

5. สถานที่ติดต่อ

ที่ทำงาน

สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์
มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร โทรศัพท์ (02) 02-836-3000 ต่อ 4183 หรือ 4184
อีเมล pornpat.s@rmutp.ac.th และ sirithumgul.p@gmail.com
ที่อยู่ 1381 ถนนประชาราษฎร์ 1 แขวงวงศ์สว่าง เขตบางซื่อ กทม. 10800

6. ความชำนาญ และความสนใจพิเศษ Software Engineering, Text-mining, Ontology-based Algorithms, Natural Language Processing (NLP), Artificial Intelligence (AI) for Knowledge Processing, and Online Learning Technology

7. ทุนการศึกษา

ทุนรัฐบาลไทยสังกัดกระทรวงวิทยาศาสตร์และเทคโนโลยี (ปี 2552 - 2559) เพื่อศึกษาต่อในระดับปริญญาโท – เอก ณ ประเทศสหรัฐอเมริกา

8. ผลงานทางวิชาการ

8.1 งานประพันธ์ตำราเรียน

พรภัทร์ ศิริธรรมกุล (2563). หน่วยที่ 4. ทฤษฎีการรับรู้. ใน *ตำราเรียนชุดวิชาการออกแบบส่วนปฏิสัมพันธ์บนเว็บและโมบาย* หน่วยที่ 1–7. นนทบุรี: สาขาวิชาวิทยาศาสตร์และเทคโนโลยี สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมาธิราช.

พรภัทร์ ศิริธรรมกุล (2563). หน่วยที่ 8. การออกแบบประสบการณ์ผู้ใช้. ใน *ตำราเรียนชุดวิชาการออกแบบส่วนปฏิสัมพันธ์บนเว็บและโมบาย* หน่วยที่ 8–13. นนทบุรี: สาขาวิชาวิทยาศาสตร์และเทคโนโลยี สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมาธิราช.

พรภัทร์ ศิริธรรมกุล (2563). หน่วยที่ 13. เซอร์วิสบนแอนดรอยด์. ใน *ตำราเรียนชุดวิชาการโปรแกรมประยุกต์บนอุปกรณ์เคลื่อนที่* หน่วยที่ 8–13. นนทบุรี: สาขาวิชาวิทยาศาสตร์และเทคโนโลยี สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมาธิราช.

พรภัทร์ ศิริธรรมกุล และภรณี ศรีสุทธิ์ **(2563).** หน่วยที่ 7. วิธีการเพื่อการพัฒนาระบบสารสนเทศ. ใน *ตำราเรียนชุดวิชาเทคโนโลยีเพื่อการจัดการสารสนเทศสำหรับนักศึกษาระดับปริญญาโท หน่วยที่ 6-10.* นนทบุรี: สาขาวิชาศิลปศาสตร์ แขนงวิชาสารสนเทศศาสตร์ สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมาธิราช.

พรภัทร์ ศิริธรรมกุล และชูจิตต์ ธนภัณฑ์ภาดา **(2563).** หน่วยที่ 11. การจัดหา การติดตั้ง การบำรุงรักษา และการประเมินระบบสารสนเทศ. ใน *ตำราเรียนชุดวิชาเทคโนโลยีสารสนเทศเบื้องต้น หน่วยที่ 11-15.* นนทบุรี: สาขาวิชาศิลปศาสตร์ แขนงวิชาสารสนเทศศาสตร์ สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมาธิราช.

พรภัทร์ ศิริธรรมกุล. (2562). หน่วยที่ 2. ระดับและกระบวนการด้านความมั่นคงปลอดภัยไซเบอร์. ใน *ตำราเรียนชุดวิชาความมั่นคงปลอดภัยไซเบอร์ หน่วยที่ 1-5.* นนทบุรี: สาขาวิชาวิทยาศาสตร์และเทคโนโลยี สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมาธิราช.

พรภัทร์ ศิริธรรมกุล (2562). หน่วยที่ 4. การเก็บรวบรวมและการเตรียมข้อมูลก่อนการประมวลผล. ใน *ตำราเรียนชุดวิชาการวิเคราะห์ข้อมูล หน่วยที่ 1-5.* นนทบุรี: สาขาวิชาวิทยาศาสตร์และเทคโนโลยี สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมาธิราช.

8.2 ผลงานวิจัยที่เผยแพร่

Sirithumgul, Pornpat, Prasertsilp, Pimpaka & Olfman, Lorne. An Algorithm for Generating Gap-Fill Multiple Choice Questions of an Expert System. In Proceedings of 2021 IEEE International Conference on Software Engineering (ICSE 2021), Madrid, Spain (forthcoming).

Sirithumgul, Pornpat, Prasertsilp, Pimpaka, Suksa-Ngiam, Watanyoo & Olfman, Lorne. An Ontology-Based Framework as a Foundation of an Information System for Generating Multiple-Choice Questions. In Proceedings of the 25th Americas Conference on Information Systems, Cancun, Mexico. August 15 – 17, 2019.

Sirithumgul, P. (2016). An ontology-based algorithm as a foundation of an automated knowledge assessment tool applied in the scientific discussions. Available from ProQuest Dissertations & Theses Global.

Sirithumgul, Pornpat, & Olfman, Lorne. A model for measuring knowledge constructions of students in online discussions. In Proceedings of the 24th Australasian Conference on Information Systems, Melbourne, Australia. December 4-6, 2013.

Sirithumgul, Pornpat, Suchato, Atiwong, & Punyabukkana, Proadpran. Quantitative Evaluation for Web Accessibility with Respect to Disabled Groups. In Proceedings of the 6th International Cross- Disciplinary Conference on Web Accessibility (W4A2009), Madrid, Spain. April 20-21, 2009.

Sirithumgul, Pornpat, Suchato, Atiwong, & Punyabukkana, Proadpran. Elearning Transformation for the visually impaired. In Proceedings of the 8th Distance Learning and the Internet Conference (DLI2007), Bangkok, Thailand. December 12-15, 2007.

8.3 โครงการวิจัย

พรภัทร์ ศิริธรรมกุล (2563). หัวหน้าโครงการวิจัยเรื่อง *กรอบงานบนพื้นฐานของภววิทยาในฐานะที่เป็นพื้นฐานของระบบสารสนเทศสำหรับสร้างคำถามแบบปรนัย (An Ontology-based Framework as a Foundation of an Information System for Generating Multiple-*

choice Questions) สนับสนุนโดยทุนงบประมาณเงินรายได้หน่วยงาน มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร ประจำปี 2563

พรภัทร์ ศิริธรรมกุล (2562). หัวหน้าโครงการวิจัยเรื่อง *แบบจำลองสำหรับสร้างระบบผู้เชี่ยวชาญที่ใช้สำหรับผลิตคำถามแบบเติมคำในช่องว่างที่มีตัวเลือกแบบปรนัย (A Model for Constructing an Expert System Used for Generating Gap-fill Multiple Choice Questions)* สนับสนุนโดยโครงการสนับสนุนทุนนักวิจัยใหม่กระทรวงวิทยาศาสตร์ (วท.) สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ (สวทช.)

8.4 ประสบการณ์การเป็นวิทยากรและอาจารย์พิเศษ

พรภัทร์ ศิริธรรมกุล (2563). วิทยากรบรรยายพิเศษเรื่อง *การบริการเว็บและเอพีไอบนเว็บ (Web Services and Web APIs)* ซึ่งเป็นส่วนหนึ่งของชุดวิชา *การโปรแกรมเว็บ (Web Programming)* สำหรับนักศึกษาระดับปริญญาตรี

พรภัทร์ ศิริธรรมกุล (2562). วิทยากรรายการโทรทัศน์เพื่อการสอนทางไกลเรื่อง *ระดับและกระบวนการด้านความมั่นคงปลอดภัยไซเบอร์* ซึ่งเป็นส่วนหนึ่งของชุดวิชา *ความมั่นคงปลอดภัยไซเบอร์* แพร่ภาพผ่าน สถานีวิทยุโทรทัศน์ของมหาวิทยาลัยสุโขทัยธรรมาธิราช (STOU Channel) สถานีวิทยุโทรทัศน์ การศึกษาทางไกลผ่านดาวเทียม True Vision ช่อง DLTV14 ทางเครือข่ายอินเทอร์เน็ต www.stou.ac.th และสื่อสังคม YouTube

พรภัทร์ ศิริธรรมกุล (2562). วิทยากรรายการโทรทัศน์เพื่อการสอนทางไกลเรื่อง *การเก็บรวบรวมและการเตรียมข้อมูลก่อนการประมวลผล* ซึ่งเป็นส่วนหนึ่งของชุดวิชา *การวิเคราะห์ข้อมูล* แพร่ภาพผ่าน สถานีวิทยุโทรทัศน์ของมหาวิทยาลัยสุโขทัยธรรมาธิราช (STOU Channel) สถานีวิทยุโทรทัศน์ การศึกษาทางไกลผ่านดาวเทียม True Vision ช่อง DLTV14 ทางเครือข่ายอินเทอร์เน็ต www.stou.ac.th และสื่อสังคม YouTube

พรภัทร์ ศิริธรรมกุล (2562). อาจารย์อาสาและอาจารย์พิเศษวิชา *การโปรแกรมคอมพิวเตอร์เบื้องต้นโดยใช้ภาษาซี* ให้กับโครงการการศึกษาทางไกลผ่านดาวเทียมไกลกังวล ภายใต้การสนับสนุนของมูลนิธิการศึกษาทางไกลผ่านดาวเทียมในพระบรมราชูปถัมภ์ ที่เปิดให้ความรู้กับนักเรียนระดับอาชีวศึกษาและประชาชนทั่วไป แพร่ภาพทางช่อง สศทท.13 (dltv13) ระหว่างวันที่ 20 มิถุนายน ถึง 13 ตุลาคม 2562

พรภัทร์ ศิริธรรมกุล (2551). วิทยากรอบรมหัวข้อ Web Accessibility, WCAG 2.0 and Assistive Technology ในโครงการความร่วมมือระหว่าง คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยและกระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร

พรภัทร์ ศิริธรรมกุล (2551). วิทยากรอบรมการใช้งานโปรแกรมอ่านหน้าจอคอมพิวเตอร์ Non-Visual Desktop Access (NVDA) ให้กับนักเรียนตาบอด ณ ศูนย์การศึกษาพิเศษส่วนกลาง กรุงเทพมหานคร

8.5 ประสบการณ์การสอน

1. อาจารย์ผู้สอนวิชา *วิทยาศาสตร์ข้อมูล (Data Science)* สำหรับนักศึกษาชั้นปีที่ 4 สาขาวิชาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร (ตั้งแต่ปี 2562 - ปัจจุบัน)
2. อาจารย์ผู้สอนวิชา *พื้นฐานการออกแบบและการโปรแกรมเชิงวัตถุโดยใช้ภาษาไพธอน (Fundamentals of Object-Oriented Design and Programming Using Python)* สำหรับนักศึกษาชั้นปีที่ 2 ภาควิชาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร (ตั้งแต่ปี 2562 - ปัจจุบัน)
3. อาจารย์ผู้สอนวิชา *วิศวกรรมซอฟต์แวร์ (Software Engineering)* สำหรับนักศึกษาชั้นปีที่ 3 ภาควิชาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร (ตั้งแต่ปี 2559 - ปัจจุบัน)
4. อาจารย์ผู้สอนวิชาการโปรแกรมคอมพิวเตอร์ สำหรับนักศึกษาชั้นปีที่ 1 ภาควิชาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร (ตั้งแต่ปี 2559 - 2562)
5. ผู้ช่วยสอนวิชา Knowledge Management and Social Media (Transdisciplinary Course) สำหรับนักศึกษาระดับปริญญาเอก Claremont Graduate University (ภาคการศึกษา Fall 2014)
6. ผู้ช่วยสอนวิชาวิศวกรรมซอฟต์แวร์ สำหรับนักศึกษาชั้นปีที่ 3 ภาควิชาวิศวกรรมคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย (ปีการศึกษา 2550 - 2552)

8.6 งานบริการวิชาการ

1. **Program committee member and reviewer** for the International Multi-Conference on Computing in the Global Information Technology (ICCGI)
2. **Session chair** for the 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON 2020)
3. **Reviewer** for international journal and conferences including:
   3.1 Computer & Education: An International Journal
   3.2 International Conference on Information Systems (ICIS)
   3.3 Hawaii International Conference on System Sciences (HICSS)

8.7 Professional Society Membership

1. Association for Computing Machinery (ACM) (2009 – present)
2. Association for Information Systems (AIS) (2019 – present)

9. ประสบการณ์การทำงานภาคเอกชน

1. **ตำแหน่ง Senior software engineer** บริษัท Motif Technology Co., Ltd. เพื่อสร้างระบบการจัดการสารสนเทศ (DIS-MIS) ให้กับกรมสอบสวนคดีพิเศษแห่งราชอาณาจักรไทย (Department of Special Investigation: DSI)
2. **ตำแหน่ง Software engineer** บริษัท Motif Technology Co., Ltd. เพื่อสร้างระบบการจัดการบัญชีหลัก (Mainstream Account System: MAS) ให้กับบริษัท Bangkok Aviation Fuel Services Public Company Limited (BAFS)

10. อื่น ๆ

เป็นผู้แทนคณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร ทำหน้าที่ดูแลและฝึกสอนนักศึกษาแลกเปลี่ยนโครงการ International Association for the Exchange of Students for Technical Experience (IAESTE)

มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร