

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

นับตั้งแต่อดีตเป็นต้นมา การพัฒนาโปรแกรมส่วนใหญ่ถูกพัฒนาขึ้นโดยใช้แนวความคิดที่เรียกว่า วิธีการเขียนโปรแกรมเชิงโครงสร้าง (Structured Programming) ซึ่งมีข้อดี คือ โปรแกรมจะแบ่งลักษณะการทำงานออกเป็น ส่วน ๆ โดยแต่ละส่วนนั้นจะทำงานภายใต้ขอบเขตการทำงานที่ชัดเจน และเรียงลำดับการทำงาน แต่ในทางตรงกันข้ามกลับส่งผลกระทบต่อการพัฒนาโปรแกรมที่มีขนาดใหญ่ เช่น เสียค่าใช้จ่ายสูงและใช้ระยะเวลานานในการพัฒนาโปรแกรมแต่ละโปรแกรม รวมทั้งเกิดความยุ่งยากในการนำเทคโนโลยีสมัยใหม่เข้ามาประยุกต์ใช้ ทำให้โปรแกรมที่พัฒนาขึ้นขาดความน่าเชื่อถือและยากแก่การบำรุงรักษา เมื่อเปรียบเทียบการเขียนโปรแกรมในลักษณะนี้กับวิธีการเขียนโปรแกรมเชิงวัตถุ (Object-Oriented Programming) พบว่า การเขียนโปรแกรมเชิงวัตถุอำนวยความสะดวกในการบำรุงรักษาโปรแกรมมากกว่าการเขียนโปรแกรมเชิงโครงสร้าง ทำให้ในปัจจุบันนี้แนวคิดเกี่ยวกับเรื่องการออกแบบและพัฒนาโปรแกรมเชิงวัตถุได้รับความสนใจ การระบุวัตถุซอฟต์แวร์ที่เป็นไปได้ (Candidate Object Identification) เป็นอีกวิธีหนึ่งที่อำนวยความสะดวกในการปรับย้ายโครงสร้างโปรแกรมจากแบบเชิงโครงสร้างมาเป็นเชิงวัตถุ การระบุวัตถุซอฟต์แวร์ที่เป็นไปได้อาจมีขั้นตอนการทำงานดังนี้ 1) คัดแยกส่วนประกอบเดิมออกจากโปรแกรมเก่า 2) หาความสัมพันธ์ระหว่างส่วนประกอบเดิม 3) จัดเรียงส่วนประกอบเดิมเป็นกลุ่ม และ 4) เลือกกลุ่มที่มีความหมายมาประกอบเป็นโครงสร้างที่เรียกว่า วัตถุซอฟต์แวร์ที่เป็นไปได้ (สมศักดิ์ ภัทรศุกุล, 2545 :1) ในปัจจุบันได้มีงานวิจัยทางด้านการระบุวัตถุซอฟต์แวร์ที่เป็นไปได้ โดยงานวิจัยดังกล่าวเน้นไปที่การศึกษาผลที่ได้จากการนำวิธีการวิเคราะห์ข้อมูลชนิดต่าง ๆ มาประยุกต์ใช้ ซึ่งวิธีการวิเคราะห์ข้อมูลเป็นส่วนสำคัญของการระบุวัตถุซอฟต์แวร์ที่เป็นไปได้เพราะวิธีการวิเคราะห์ข้อมูลช่วยให้การระบุวัตถุซอฟต์แวร์ที่เป็นไปได้ในขั้นตอนของการจัดเรียงส่วนประกอบเดิมและเลือกกลุ่มที่มีความหมายมีความรวดเร็ว และง่ายยิ่งขึ้น แต่ทั้งนี้ก็ยังขึ้นอยู่กับว่าจะเลือกวิธีการวิเคราะห์ข้อมูลมาประยุกต์ใช้ให้เหมาะสมได้เพียงใด วิธีการวิเคราะห์ที่ได้รับความสนใจในปัจจุบันนี้ ได้แก่ วิธีการวิเคราะห์แบบแนวคิด (Concept Analysis) วิธีการจัดกลุ่มข้อมูลแบบลำดับชั้น (Hierarchical Clustering Analysis)

วิธีการวิเคราะห์แบบแนวคิด (Concept Analysis) เป็นวิธีการวิเคราะห์ข้อมูลชนิดหนึ่งที่เป็นทฤษฎีทางด้านปัญญาประดิษฐ์ ใช้จัดเรียงข้อมูลที่มีคุณสมบัติตรงกันให้อยู่ในกลุ่มเดียวกัน ซึ่งมีข้อดีหลายประการ คือ 1) สามารถจัดเรียงส่วนประกอบเดิมได้อย่างถูกต้องและรวดเร็ว 2) บอกรายละเอียดของการจัดกลุ่มได้อย่างชัดเจน โดยแสดงในรูปของแผนภูมิต้นไม้ ซึ่งทำให้ง่ายแก่การทำทำความเข้าใจ อย่างไรก็ตามได้พบว่าวิธีการดังกล่าวนี้มีข้อเสียเช่นเดียวกัน คือ 1) ไม่สามารถถ่วงน้ำหนักการทำงานได้ ทำให้ได้สารสนเทศทั้งที่มีประโยชน์และไม่มีประโยชน์ต่อการใช้งาน 2) ส่วนประกอบเดิมที่ถูกจัดอยู่ในคอนเซ็ปต์หนึ่ง สามารถอยู่ในอีกคอนเซ็ปต์หนึ่งได้ ทำให้เกิดการซ้อนทับกัน และไม่สามารถนำมาใช้เป็นตัวแทนของกลุ่มข้อมูลได้ (Siff and Reys, 1977 : 170-179) และ 3) เมื่อนำมาใช้กับโปรแกรมที่มีโครงสร้างแบบผูกติด วิธีการดังกล่าวไม่สามารถแยกแยะส่วนประกอบเดิมได้อย่างถูกต้อง (S. Phattarsukol and P. Muenchaisri, 2001 : 381-389)

วิธีการจัดกลุ่มข้อมูลแบบลำดับชั้น (Hierarchical Cluster Analysis) เป็นวิธีการวิเคราะห์ข้อมูลประเภทหนึ่งที่ได้รับความนิยมและนำไปประยุกต์ใช้ในงานวิจัยด้านต่าง ๆ อย่างกว้างขวาง เช่น งานวิจัยด้านชีววิทยา เศรษฐศาสตร์ ถึงแม้ว่าวิธีการจัดกลุ่มข้อมูลแบบลำดับชั้นจะมีความยืดหยุ่นในการทำงาน สามารถปรับเปลี่ยนส่วนประกอบต่าง ๆ ได้มากมาย (Wiggerts, 1997 : 33-43) และมีการแสดงผลเป็นรูปแผนภูมิต้นไม้ ซึ่งง่ายแก่การทำทำความเข้าใจ แม้ว่าวิธีการดังกล่าวจะมีข้อดี แต่ก็มีข้อเสียเหมือนกัน กล่าวคือ 1) ข้อมูลเมื่อถูกจัดกลุ่มแล้ว ไม่สามารถเปลี่ยนกลุ่มได้อีก ทำให้ไม่สามารถแสดงจำนวนของกลุ่มข้อมูลทั้งหมดที่เป็นไปได้ 2) ต้องใช้คนเป็นผู้พิจารณาผลของการจัดกลุ่มอีกครั้ง

นอกจากนี้ยังพบว่าวิธีการจัดกลุ่มข้อมูลแบบเออาร์ที 1 (ART1) และเออาร์ที 2 (ART 2) สามารถนำมาใช้ในการจัดกลุ่มข้อมูลด้วยเช่นกัน ซึ่งแม้ว่าจะเป็นวิธีการหนึ่งที่ได้รับความนิยมและนำไปประยุกต์ใช้งานวิจัยอย่างกว้างขวาง เช่น ทางด้านวิศวกรรมศาสตร์ แต่ก็พบว่าวิธีการดังกล่าวนี้ยังมีข้อเสียเช่นเดียวกัน กล่าวคือ วิธีการจัดกลุ่มข้อมูลแบบเออาร์ที 1 นั้น ไม่สามารถวิเคราะห์และจัดกลุ่มข้อมูลที่มีลักษณะของกลุ่มข้อมูลแบบแอนะล็อก (Analog) ได้ สำหรับวิธีการวิเคราะห์ข้อมูลแบบเออาร์ที 2 นั้นมีข้อเสียเช่นเดียวกัน กล่าวคือ เออาร์ที 2 นั้นมีลักษณะการทำงานที่ค่อนข้างซับซ้อนและยากแก่การทำทำความเข้าใจ และเริ่มไม่ได้รับความนิยมตั้งแต่มีการนำวิธีการจัดกลุ่มข้อมูลแบบฟิชเชอร์ เออาร์ที มาใช้งาน

วิธีการจัดกลุ่มข้อมูลแบบฟัซซี่ เออาร์ที (Fuzzy ART) เป็นวิธีการวิเคราะห์ข้อมูลอีกประเภทหนึ่งที่ได้รับการยอมรับและนำไปประยุกต์ใช้ในงานวิจัยด้านต่าง ๆ ที่เกี่ยวข้องกับโครงข่ายประสาท ซึ่งพบว่าสามารถใช้ในการจัดกลุ่มข้อมูลที่มีลักษณะของกลุ่มข้อมูลแบบไบนารีและแอนาล็อกได้

วิทยานิพนธ์ฉบับนี้ จัดทำขึ้นเพื่อนำเสนอแนวทางในการระบุหาวัตถุและความสัมพันธ์แบบสืบทอดโดยอัตโนมัติ โดยใช้เทคนิคฟัซซี่ เออาร์ที (Fuzzy ART) มาประยุกต์ใช้ในการจัดกลุ่มข้อมูลเพื่อระบุหาวัตถุ และเทคนิคของกราฟเพื่อระบุหาความสัมพันธ์แบบสืบทอดระหว่างวัตถุ โดยตัวอย่างที่ใช้ศึกษานั้นเป็นโปรแกรมที่พัฒนาขึ้นมาจากภาษาซีและอาศัยหลักการเขียนโปรแกรมเชิงโครงสร้าง และโปรแกรมที่มีลักษณะการเขียนแบบสืบทอดคุณสมบัติ (Inheritance)

1.2 วัตถุประสงค์ของการวิจัย

เพื่อนำเสนอแนวทางในการระบุหาวัตถุและความสัมพันธ์แบบสืบทอดโดยอัตโนมัติ

1.3 ขอบเขตการวิจัย

1. โปรแกรมต้นฉบับเชิงโครงสร้างที่ใช้ศึกษาพัฒนามาจากภาษาซี ที่มีลักษณะแตกต่างกัน 4 โปรแกรม คือ (1) โปรแกรมสัตว์เลี้ยง ซึ่งมีลักษณะการเขียนแบบสืบทอด (Inheritance) (2) โปรแกรมแถวลำดับและแถวคอย แบบส่วนจำเพาะ (Modular Case) (3) โปรแกรมแถวลำดับและแถวคอยแบบผูกติด (Tangled Case) (4) โปรแกรมการลงทะเบียนของพนักงาน (Employee Case)
2. ซอฟต์แวร์ที่ใช้ในวิทยานิพนธ์ ได้แก่ ชุดพัฒนาโปรแกรม MATLAB รุ่น 6.5