บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัญหาที่สำคัญประการหนึ่งของการเรียนรู้แบบขั้นตอนวิธีเชิงพันธุกรรม(genetic algorithm) [1] ที่มีการกล่าวถึงจากงานวิจัยเป็นจำนวนมากคือการเกิดปรากฏการณ์ของการลู่เข้าก่อนกำหนด(premature convergence) ในระหว่างขั้นตอนวิวัฒนาการ(evolutionary process) [2-11] การลู่เข้าก่อนกำหนดเป็นลักษณะของกระบวนการวิวัฒนาการที่ประชากรของหน่วยชีวิต(individual)ลู่เข้าสู่คำตอบใดคำตอบหนึ่งอย่างรวดเร็วเกินไปซึ่งเป็นสาเหตุให้ประชากรสูญเสียความสามารถในการปรับปรุงเพื่อเป็นคำตอบที่ดียิ่งขึ้นต่อไปในอนาคต สาเหตุสำคัญของการลู่เข้าก่อนกำหนดนี้เป็นผลมาจากการขาดความหลากหลาย(diversity) ที่เพียงพอในประชากรของหน่วยชีวิตประกอบกับการขาดประสิทธิภาพของการไขว้เปลี่ยน(crossover) ในการสร้างคำตอบรุ่นใหม่ที่แตกต่างจากเดิม

ในขั้นตอนการทำงานของขั้นตอนวิธีเชิงพันธุกรรมแบบปกติจะมิได้มีส่วนของการควบคุมความหลากหลายของประชากรโดยตรง เมื่อความหลากหลายของประชากรมีไม่เพียงพอจึงอาจส่งผลให้กระบวนการเรียนรู้ไม่ประสบผลสำเร็จ ทั้งนี้เนื่องจากหน่วยชีวิตแต่ละหน่วยล้วนมีลักษณะที่ใกล้เคียงกัน การสร้างคำตอบที่ดียิ่งขึ้นจึงเป็นไปได้ยากซึ่งเป็นผลมาจากส่วนที่แตกต่างกันของประชากรอันเป็นวัตถุทางพันธุกรรม(genetic material) สำคัญที่จำเป็นต่อการสร้างคำตอบใหม่ได้ถูกขจัดทิ้งไปเป็นอันมากในระหว่างขั้นตอนวิวัฒนาการ ซึ่งส่งผลให้กระบวนการวิวัฒนาการมีแนวโน้มที่จะติดอยู่กับจุดที่ดีที่สุดท้องถิ่น(local optimum) [10,12,13] บริเวณใดบริเวณหนึ่งโดยขาดความสามารถในการพัฒนาเป็นคำตอบที่ดีที่สุดที่แท้จริงได้(global optimum)

การรักษาความหลากหลายของประชากรได้รับการยอมรับว่าเป็นส่วนสำคัญซึ่งทำให้กระบวนการหาคำตอบของขั้นตอนวิธีเชิงพันธุกรรมมีประสิทธิภาพมากขึ้น [14-17] โดยเฉพาะกับปัญหาที่มีสภาพแวดล้อมไม่คงที่ [18,19] อย่างไรก็ดีหากความหลากหลายของประชากรมีมากจนเกินไปก็อาจส่งผลให้ประสิทธิภาพในการหาคำตอบของขั้นตอนวิธีเชิงพันธุกรรมด้อยลงเนื่องจากต้องใช้เวลาในการหาคำตอบมากเกินไปหรือกระบวนการวิวัฒนาการมีอัตราการลู่เข้า(convergence rate) สู่คำตอบที่ต่ำ

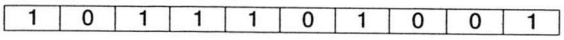
ได้มีงานวิจัยหลายงานได้เสนอวิธีการต่างๆเพื่อปรับปรุงประสิทธิภาพของขั้นตอนวิธีเชิงพันธุกรรมโดยใช้กลไกของการรักษาความหลากหลายของประชากรในระหว่างกระบวนการวิวัฒนาการในรูปแบบต่างๆกัน อย่างไรก็ดีการใช้วิธีการเรียนรู้ที่ปรับปรุงใหม่เหล่านี้จำเป็นต้องกำหนดค่าของพารามิเตอร์(parameter) ในการใช้งานที่เหมาะสมก่อนเริ่มทำงานซึ่งอาจจำเป็นต้องทดลองใช้โดยการลองผิดลองถูกหลายๆครั้ง ทั้งนี้หากใช้ค่าพารามิเตอร์ที่ไม่เหมาะสมกับปัญหาจะส่งผลให้ความหลากหลายของประชากรในกระบวนการวิวัฒนาการมีมากหรือน้อยเกินไปซึ่งจะนำไปสู่ปัญหาต่างๆดังที่กล่าวมาข้างต้น ด้วยเหตุนี้ระบบควบคุมความหลากหลายของประชากรที่สามารถปรับตัวให้เหมาะสมกับปัญหาเองได้จึงเป็นประเด็นที่น่าสนใจในการทำการศึกษา

1.2 ทฤษฎีที่เกี่ยวข้อง

1.2.1 ขั้นตอนวิธีเชิงพันธุกรรม

ขั้นตอนวิธีเชิงพันธุกรรมเป็นวิธีการหนึ่งที่ใช้สำหรับการเรียนรู้ (learning) ซึ่งคิดค้นโดย Holland [1] วิธีการนี้เป็นการค้นหาคำตอบโดยใช้การจำลองแบบตามกฎการคัดเลือกโดยธรรมชาติ (natural selection) ประกอบกับการใช้ตัวปฏิบัติการทางพันธุกรรม(genetic operator) ต่างๆ ขั้นตอนวิธีเชิงพันธุกรรมเป็นวิธีการที่มีประสิทธิภาพสำหรับปัญหาหลายประเภท แต่เป็นวิธีการที่ต้องใช้การคำนวณสูง การค้นหาคำตอบโดยขั้นตอนวิธีเชิงพันธุกรรม อาจแบ่งออกเป็นขั้นตอนหลักคือ การสร้างประชากรของหน่วยชีวิตเริ่มต้น การตรวจสอบประสิทธิภาพของหน่วยชีวิต การสร้างประชากรของหน่วยชีวิตรุ่นใหม่ และการหาผลลัพธ์

1. การสร้างประชากรของหน่วยชีวิตเริ่มต้น ขั้นตอนนี้เป็นขั้นตอนเริ่มต้นของขั้นตอนวิธีเชิงพันธุกรรมซึ่งจะทำการสร้างประชากรของหน่วยชีวิตแบบโดยใช้การสุ่ม(random) ในขั้นตอนวิธีเชิงพันธุกรรมแบบดั้งเดิมหน่วยชีวิตแต่ละหน่วยจะถูกแทนด้วยสายอักขระของเลขฐานสอง(binary string) โดยจำนวนประชากรของหน่วยชีวิตและความยาวของสายอักขระเป็นพารามิเตอร์ซึ่งต้องกำหนดก่อนการทดลอง สายของอักขระสามารถเทียบเคียงได้กับโครโมโซม(chromosome) ของสิ่งมีชีวิต โดยแต่ละบิตเปรียบเสมือนจีน(gene) บนโครโมโซมนั้น ผลลัพธ์ที่ได้จากขั้นตอนนี้คือหน่วยชีวิตจำนวนมากซึ่งมีลักษณะของโครโมโซมที่แตกต่างกันอันเป็นผลมาจากการสุ่มสร้าง ซึ่งหน่วยชีวิตเหล่านี้จะถูกนำไปตรวจสอบประสิทธิภาพในขั้นตอนต่อไป ตัวอย่างของหน่วยชีวิตในขั้นตอนวิธีเชิงพันธุกรรมแสดงดังรูปที่ 1.1

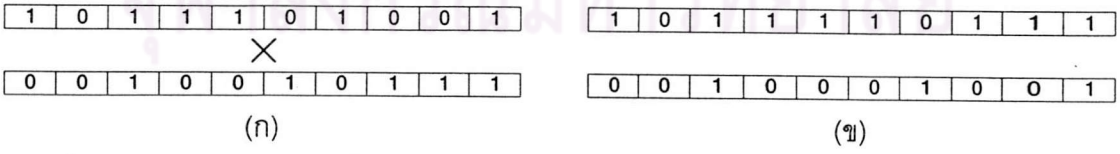


รูปที่ 1.1 ตัวอย่างหน่วยชีวิตในขั้นตอนวิธีเชิงพันธุกรรม

2. <u>การตรวจสอบประสิทธิภาพของหน่วยชีวิต</u> ในขั้นตอนนี้จะทำการตรวจสอบประสิทธิภาพของหน่วยชีวิตแต่ละหน่วย โดยจะมีหน่วยชีวิตส่วนน้อยเท่านั้นที่มีประสิทธิภาพดีซึ่งจะได้รับการคัดเลือกเพื่อนำไปสร้างเป็นประชากรของหน่วยชีวิตในรุ่น(generation) ใหม่ต่อไป การตรวจสอบประสิทธิภาพของหน่วยชีวิตนี้จะตรวจวัดโดยการคำนวณค่าความเหมาะ(fitness value) โดยใช้ฟังก์ชันความเหมาะ(fitness function) ในกรณีที่ค่าความเหมาะสูงแสดงถึงประสิทธิภาพที่ดีโดยกลุ่มของหน่วยชีวิตที่มีค่าความเหมาะสูงจะมีโอกาสถูกนำไปสร้างประชากรของหน่วยชีวิตในรุ่นใหม่ต่อไปมากกว่ากลุ่มที่ค่าความเหมาะต่ำ การนิยามฟังก์ชันความเหมาะที่มีความเหมาะสมกับปัญหานับเป็นส่วนที่ซับซ้อนและสำคัญที่สุดส่วนหนึ่งของวิธีการเรียนรู้แบบขั้นตอนวิธีเชิงพันธุกรรม โดยมีเพียงปัญหาส่วนน้อยเท่านั้นที่ฟังก์ชันความเหมาะมีลักษณะที่ง่าย

3. <u>การสร้างประชากรของหน่วยชีวิตรุ่นใหม่</u> ขั้นตอนนี้จะเป็นการสร้างประชากรของหน่วยชีวิตรุ่นใหม่โดยการไขว้เปลี่ยน(crossover) และการกลาย(mutation)

<u>การไขว้เปลี่ยน</u> เป็นกระบวนการที่สำคัญที่สุดกระบวนการหนึ่งของขั้นตอนวิธีเชิงพันธุกรรม กระบวนการนี้จะทำการสุ่มหน่วยชีวิตรุ่นเดิมครั้งละ 2 หน่วยชีวิต(parents) มาทำการสร้างเป็นหน่วยชีวิตใหม่ 2 หน่วยชีวิต การสุ่มเลือกจะมีผลมาจากค่าความเหมาะของแต่ละหน่วยชีวิต ทำให้โอกาสที่จะถูกเลือกของแต่ละหน่วยชีวิตไม่เท่ากัน โดยหน่วยชีวิตที่มีค่าความเหมาะสูงกว่าจะมีโอกาสได้รับเลือกมากกว่า อัตราการเกิดการไขว้เปลี่ยนเป็นผลมาจากการกำหนดพารามิเตอร์ P_c ซึ่งต้องกำหนดก่อนการทดลอง ตัวอย่างของการไขว้เปลี่ยนแสดงดังรูปที่ 1.2



รูปที่ 1.2 ตัวอย่างการไขว้เปลี่ยนระหว่างหน่วยชีวิตรุ่นเดิม 2 หน่วย (ก) ซึ่งได้ผลลัพธ์เป็นหน่วยชีวิตใหม่ 2 หน่วย (ข)

การกลาย เป็นกระบวนการที่สำคัญอีกกระบวนการหนึ่งซึ่งเปรียบเทียบได้กับการกลายในทางพันธุศาสตร์ หน่วยชีวิตแต่ละหน่วยที่ผ่านจากขั้นตอนการไขว้เปลี่ยนจะถูกสุ่มตำแหน่งในโครโมโซมเพื่อทำการกลายโดยการเปลี่ยนค่าของบิตนั้น อัตราการเกิดการกลายเปลี่ยนเป็นผลมาจากการกำหนดพารามิเตอร์ P_m ตัวอย่างของการกลายแสดงดังรูปที่ 1.3

(ก) (ข)

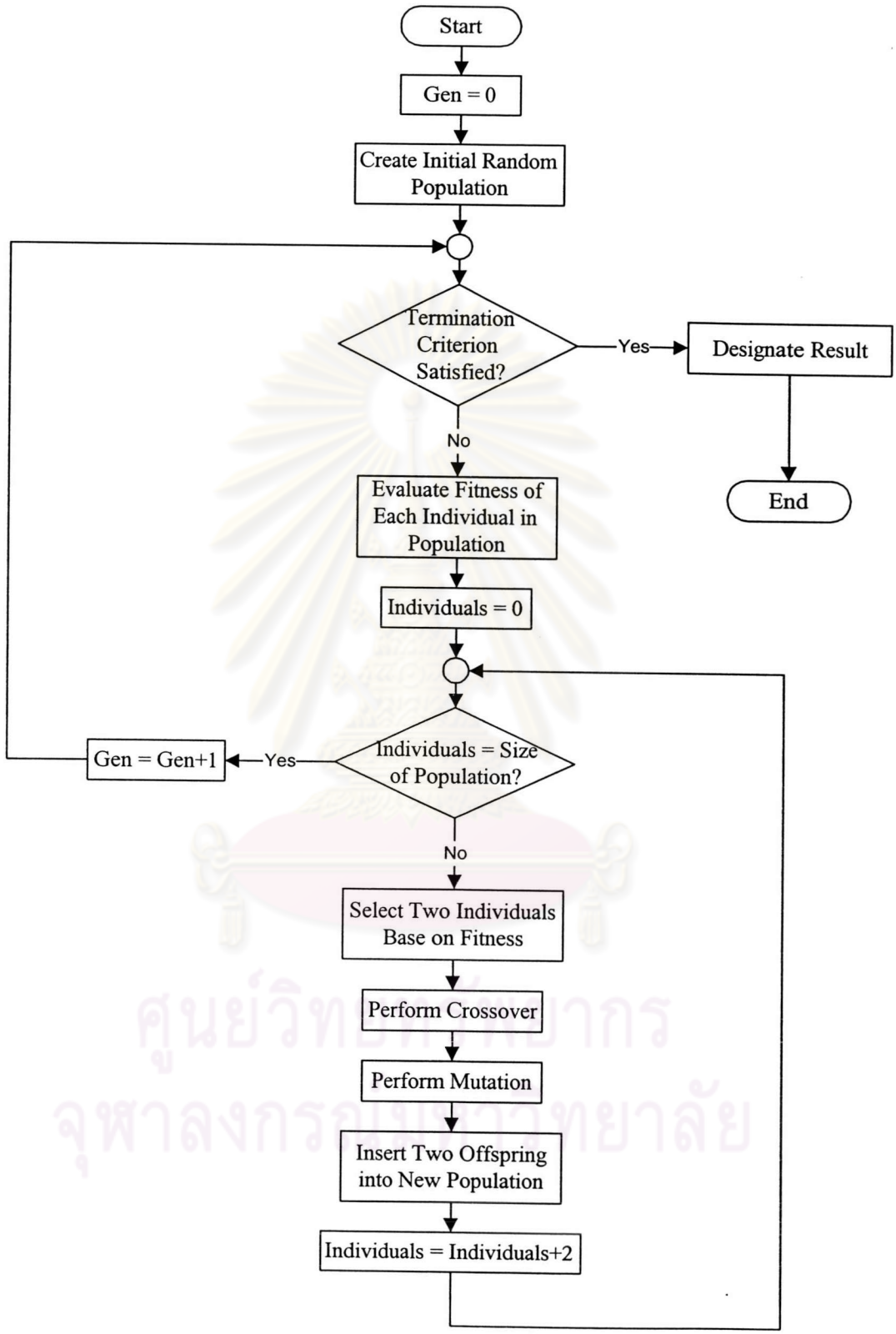
รูปที่ 1.3 ตัวอย่างการกลายจากหน่วยชีวิตเริ่มต้น (ก) และได้ผลลัพธ์เป็นหน่วยชีวิตใหม่ (ข)

หน่วยชีวิตที่ผ่านทั้งการไขว้เปลี่ยนและการกลายแล้วจะถูกบันทึกไว้เป็นประชากรของหน่วยชีวิตรุ่นใหม่ (children หรือ offspring) จากนั้นประชากรของหน่วยชีวิตรุ่นใหม่เหล่านี้ จะถูกดำเนินการซ้ำในขั้นตอนตรวจสอบประสิทธิภาพใหม่ และนำไปสร้างเป็นประชากรหน่วยชีวิตในรุ่นต่อๆไปจนพบเงื่อนไขการจบการทำงานซึ่งต้องกำหนดไว้ก่อนเริ่มการทดลอง ตัวอย่างเงื่อนไขการจบการทำงานได้แก่ จบการทำงานเมื่อครบจำนวนรุ่นที่กำหนด หรือจบการทำงานเมื่อพบหน่วยชีวิตแรกที่เป็นคำตอบของปัญหาเป็นต้น

4. การหาผลลัพธ์ ขั้นตอนนี้เป็นการหาผลลัพธ์ของขั้นตอนวิธีเชิงพันธุกรรม โดยหน่วยชีวิตหรือกลุ่มของหน่วยชีวิตที่มีประสิทธิภาพดีที่สุด(ประเมินด้วยฟังก์ชันความเหมาะ)ที่ปรากฏในขั้นตอนวิวัฒนาการจะเป็นผลลัพธ์ของขั้นตอนวิธีเชิงพันธุกรรม

ผังงานสรุปการทำงานของขั้นตอนวิธีเชิงพันธุกรรมแสดงดังรูปที่ 1.4

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 1.4 ผังงานของขั้นตอนวิธีเชิงพันธุกรรม

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

1.2.2 ความหลากหลายของประชากร

ความหลากหลายของประชากรเป็นค่าที่สามารถตรวจวัดได้หลายรูปแบบ ทั้งการตรวจวัดจากลักษณะที่แสดงออก(phenotype) ของประชากรของหน่วยชีวิต และการตรวจวัดจากลักษณะของจีน(genotype) ตัวอย่างของการตรวจวัดความหลากหลายของประชากรโดยดูจากลักษณะที่แสดงออกเช่นการตรวจวัดจากความแตกต่างของค่าความเหมาะเฉลี่ยและค่าความเหมาะมากที่สุดในแต่ละรุ่น(generation) ซึ่งความแตกต่างที่มากหรือน้อยจะแสดงถึงค่าความหลากหลายของประชากรที่มากหรือน้อยไปด้วยเช่นกัน

การตรวจวัดความหลากหลายของประชากรจากลักษณะของจีนเป็นวิธีการตรวจวัดซึ่งพิจารณาจากลักษณะของจีนที่แตกต่างกันของประชากรของหน่วยชีวิตโดยตรง โดยมักนิยามการตรวจวัดเป็นสมการทางคณิตศาสตร์ อย่างไรก็ดีการตรวจวัดวิธีนี้มีลักษณะที่ขึ้นกับลักษณะการแทน(representation) คำตอบสูง หากลักษณะการแทนคำตอบเปลี่ยนไป (เช่นกำหนดให้แต่ละจีนบนโครโมโซมเก็บค่าจำนวนเต็มแทนการเก็บเป็นเลขฐานสองแบบขั้นตอนวิธีเชิงพันธุกรรมดั้งเดิม) ก็จำเป็นต้องนิยามสมการที่ใช้ในการตรวจวัดใหม่

ตัวอย่างสมการที่ใช้ตรวจวัดความหลากหลายของประชากรจากลักษณะของจีนได้แก่

$$Diversity = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} H(I_i, I_j)}{n^2 \cdot l} \tag{1.1}$$

เมื่อ I_i, I_j แทนหน่วยชีวิตลำดับที่ i และ j ในกลุ่มประชากรตามลำดับ

$H(I_i, I_j)$ แทนระยะทางฮัมมิง(hamming distance)ของหน่วยชีวิตลำดับที่ i และ j ในกลุ่มประชากร

n แทนจำนวนประชากร และ

l แทนความยาวโครโมโซมของหน่วยชีวิต

โดยค่าความหลากหลายที่คำนวณได้โดยสมการนี้เป็นค่าเฉลี่ยของระยะทางฮัมมิงของหน่วยชีวิตทุกคู่ที่เป็นไปได้ (ในที่นี้หน่วยชีวิตหน่วยแรกและหน่วยที่ 2 อาจเป็นหน่วยชีวิตเดียวกันได้) โดยทำการปรับค่าโดยการหารด้วยความยาวโครโมโซม ค่าความหลากหลายของประชากรที่คำนวณได้มีค่าน้อยที่สุดเป็น 0 (หน่วยมีชีวิตทั้งหมดมีความเหมือนกันทุกประการ) และมีค่ามากที่สุดที่ 0.5 (ในทุกๆตำแหน่งของโครโมโซมมีจำนวนหน่วยชีวิตที่มีจีน ณ ตำแหน่งนั้นเป็นค่า 0 และค่า 1 เท่ากัน) สำหรับงานวิจัยซึ่งเสนอวิธีการการตรวจวัดความหลากหลายของประชากรจากลักษณะของจีนในรูปแบบอื่นๆได้แก่งานวิจัย [2,14,15] เป็นต้น

1.2.3 ความเพียรพยายามเชิงคำนวณ

เนื่องจากกระบวนการของขั้นตอนวิธีเชิงพันธุกรรมเป็นกระบวนการที่ใช้การสุ่ม ซึ่งทำให้ในแต่ละครั้งของการทดลองได้ผลลัพธ์ที่ไม่เหมือนกัน การวัดประสิทธิภาพของกระบวนการเรียนรู้แบบขั้นตอนวิธีเชิงพันธุกรรมจึงต้องใช้การคำนวณทางสถิติโดยใช้ข้อมูลจากผลของการทดลองซ้ำหลายๆครั้ง โดยในที่นี้จะใช้คำว่า *รอบของการทดลอง* แทนการทดลองซ้ำในแต่ละครั้ง

ตามหนังสือ Genetic Programming II ของ Koza [20] ได้กล่าวถึงความเพียรพยายามเชิงคำนวณซึ่งเป็นวิธีหนึ่งในการวัดประสิทธิภาพของวิธีการเรียนรู้แบบกำหนดการเชิงพันธุกรรม (genetic programming) (โดยสามารถนำมาใช้ในการวัดประสิทธิภาพของวิธีการเรียนรู้แบบขั้นตอนวิธีเชิงพันธุกรรมได้เช่นกัน เนื่องด้วยกระบวนการทำงานที่ไม่แตกต่างกัน) ความเพียรพยายามเชิงคำนวณเป็นการคำนวณทางสถิติเพื่อหาจำนวนหน่วยชีวิตอย่างน้อยที่ต้องประมวลผลจนพบหน่วยชีวิตที่ประสบความสำเร็จ(หมายถึงหน่วยชีวิตที่มีประสิทธิภาพซึ่งยอมรับว่าเป็นคำตอบของปัญหา)ภายใต้ความมั่นใจที่กำหนด ดังสรุปได้ดังนี้

ในการดำเนินการทดลองของกำหนดการเชิงพันธุกรรมซึ่งมีจำนวนประชากรของหน่วยชีวิตในแต่ละรุ่นคงที่เท่ากับ M และมีจำนวนรุ่นเท่ากับ G หากสามารถพบหน่วยชีวิตที่ประสบความสำเร็จในรุ่น (generation) เดียวกันทุกรอบของการทดลอง จำนวนหน่วยชีวิตอย่างน้อยที่ต้องประมวลผลจนพบหน่วยชีวิตที่ประสบความสำเร็จจะมีค่าเท่ากับผลคูณของจำนวนประชากร M กับจำนวนรุ่นที่ใช้ในการหาหน่วยชีวิตที่ประสบความสำเร็จ หากไม่พบหน่วยชีวิตที่ประสบความสำเร็จในรุ่นเดียวกันในแต่ละรอบของการทดลอง แต่สามารถพบหน่วยชีวิตที่ประสบความสำเร็จภายใน G รุ่นทุกรอบของการทดลอง จำนวนหน่วยชีวิตอย่างน้อยที่ต้องประมวลผลจนพบหน่วยชีวิตที่ประสบความสำเร็จจะมีค่าเท่ากับผลคูณของจำนวนประชากร M กับค่าเฉลี่ยของจำนวนรุ่นที่ใช้จนพบหน่วยชีวิตที่ประสบความสำเร็จในแต่ละรอบของการทดลอง แต่เนื่องจากกำหนดการเชิงพันธุกรรมเป็นขั้นตอนวิธีที่ใช้หลักของความน่าจะเป็นซึ่งไม่สามารถรับประกันได้ว่าในการทดลองทุกรอบจะสามารถพบหน่วยชีวิตที่ประสบความสำเร็จได้ภายในจำนวนรุ่นที่กำหนด จึงต้องใช้การคำนวณทางสถิติเพื่อหาจำนวนหน่วยชีวิตอย่างน้อยที่ต้องประมวลผลจนพบหน่วยชีวิตที่ประสบความสำเร็จภาพใต้ความมั่นใจที่กำหนด

หากกำหนดให้ $Y(M,i)$ แทนความน่าจะเป็นที่จะพบหน่วยชีวิตที่ประสบความสำเร็จในรุ่นที่ i ซึ่งคำนวณได้จากจำนวนรอบของการทดลองที่พบหน่วยชีวิตที่ประสบความสำเร็จในรุ่นที่ i หารด้วยจำนวนรอบของการทดลองทั้งหมด จะสามารถคำนวณหาค่าความน่าจะเป็นสะสมของ

ความสำเร็จ $P(M,i)$ ได้ โดย $P(M,i)$ จะเป็นฟังก์ชันที่ไม่ลดลง หากในทุกรอบของการทดลองพบหน่วยชีวิตที่ประสบความสำเร็จภายใน G รุ่นดังกล่าวมาข้างต้นจะได้ว่า $P(M,G)$ จะมีค่าเท่ากับ 1.0 แต่ในความเป็นจริงจะพบว่า $P(M,G)$ มักจะมีค่าน้อยกว่า 1.0 เนื่องจากจะมีบางรอบของการทดลองที่ไม่สามารถพบหน่วยชีวิตที่ประสบความสำเร็จได้ภายใน G รุ่น

หากกำหนดให้ $R(M,i,z)$ แทนจำนวนรอบของการทดลองเพื่อให้สามารถพบหน่วยชีวิตที่ประสบความสำเร็จภายในรุ่นที่ i โดยมีความมั่นใจของความสำเร็จเท่ากับ z (เช่น z เท่ากับ 99 เปอร์เซ็นต์) จะได้ว่า

$$z = 1 - \left[1 - P(M,i)\right]^{R(M,i,z)} \tag{1.2}$$

ซึ่งสามารถหา $R(M,i,z)$ ได้จากสมการ

$$R(M,i,z) = \left\lceil \frac{\log(1-z)}{\log(1-P(M,i))} \right\rceil \tag{1.3}$$

หากกำหนดให้ $I(M,i,z)$ แทนจำนวนหน่วยชีวิตอย่างน้อยที่ต้องประมวลผลจนพบหน่วยชีวิตที่ประสบความสำเร็จภายในรุ่นที่ i โดยมีค่าความมั่นใจของความสำเร็จ z จะได้ว่า

$$I(M,i,z) = M \times i \times R(M,i,z) \tag{1.4}$$

สำหรับรุ่นที่มีค่า $I(M,i,z)$ น้อยที่สุดเรียกว่ารุ่นที่ดีที่สุด (best generation, i^*) ซึ่งค่า $I(M,i,z)$ ของรุ่นที่ดีที่สุดนี้เรียกว่า ความเพียรพยายามเชิงคำนวณ(E)โดย

$$E = I(M,i^*,z) = M \times i^* \times R(M,i^*,z) \tag{1.5}$$

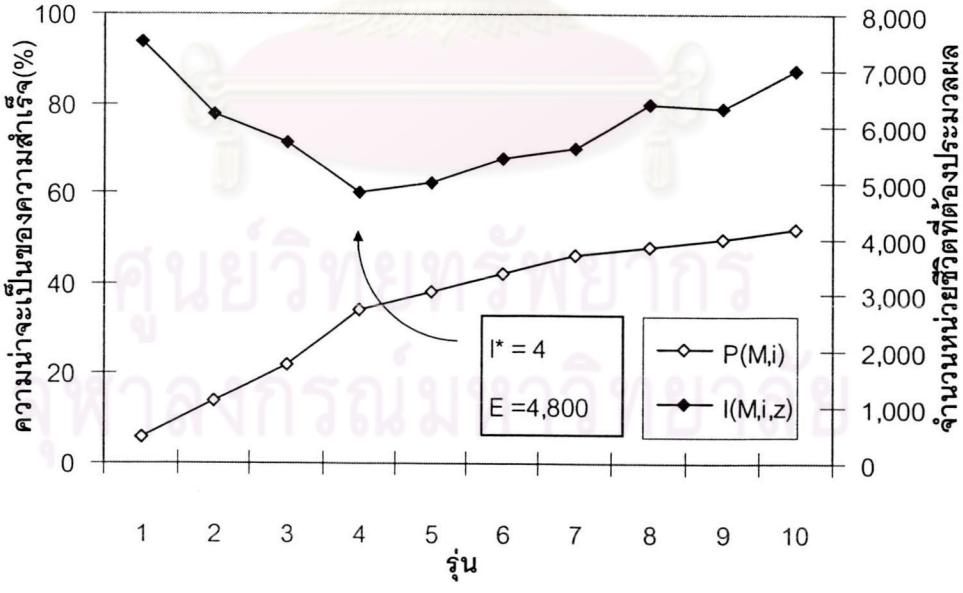
ซึ่งค่าความเพียรพยายามเชิงคำนวณที่คำนวณได้นี้ไม่จำเป็นจะต้องเป็นค่าความเพียรพยายามเชิงคำนวณน้อยที่สุดจริงๆของปัญหานั้น

ตัวอย่างการคำนวณค่าความเพียรพยายามเชิงคำนวณของตัวอย่างปัญหาแสดงดังตารางที่ 1.1 โดยกำหนดจำนวนประชากรของหน่วยชีวิตในแต่ละรุ่นเท่ากับ 100 จำนวนรุ่นเท่ากับ 10 รุ่น จำนวนรอบของการทดลองเท่ากับ 50 รอบ ค่าความมั่นใจของความสำเร็จเท่ากับ 99 เปอร์เซ็นต์ ซึ่งพบว่าค่าความเพียรพยายามเชิงคำนวณของตัวอย่างปัญหานี้เท่ากับ 4,800 โดยรุ่นที่ดีที่สุดคือรุ่นที่ 4

กราฟที่แสดงความน่าจะเป็นสะสมของความสำเร็จ $P(M,i)$ และจำนวนหน่วยชีวิตอย่างน้อยที่ต้องประมวลผลจนพบหน่วยชีวิตที่ประสบความสำเร็จ $I(M,i,z)$ ในแต่ละรุ่น i เรียกว่า *เส้นโค้งสมรรถนะ*(performance curve) สำหรับเส้นโค้งสมรรถนะของตัวอย่าง ปัญหาแสดงดังรูปที่ 1.5

ตารางที่ 1.1 ค่าของตัวแปรต่างๆในการหาค่าความเพียรพยายามเชิงคำนวณของตัวอย่างปัญหา

รุ่นที่	จำนวนรอบของการทดลองที่พบหน่วยชีวิตที่ประสบความสำเร็จ	$Y(M,i)$	$P(M,i)$	$R(M,i,z)$	$I(M,i,z)$
1	3	0.06	0.06	75	7,500
2	4	0.08	0.14	31	6,200
3	4	0.08	0.22	19	5,700
4	6	0.12	0.34	12	4,800
5	2	0.04	0.38	10	5,000
6	2	0.04	0.42	9	5,400
7	2	0.04	0.46	8	5,600
8	1	0.02	0.48	8	6,400
9	1	0.02	0.50	7	6,300
10	1	0.02	0.52	7	7,000



รูปที่ 1.5 เส้นโค้งสมรรถนะของตัวอย่างปัญหา

ในการเปรียบเทียบโดยใช้ค่าความเพียรพยายามเชิงคำนวณ หากใช้กับวิธีการหาคำตอบเดียวกันในการหาคำตอบของปัญหาที่ต่างกัน ปัญหาที่มีค่าความเพียรพยายามเชิงคำนวณสูงกว่าจะแสดงถึงปัญหานั้นมีความยากกว่าปัญหาที่มีค่าความเพียรพยายามเชิงคำนวณต่ำกว่า และหากใช้วิธีการหาคำตอบต่างวิธีกันในการหาคำตอบของปัญหาเดียวกัน วิธีการหาคำตอบที่คำนวณค่าความเพียรพยายามเชิงคำนวณได้ต่ำกว่าแสดงถึงความมีประสิทธิภาพในการทำงานมากกว่าวิธีการเรียนรู้ที่คำนวณค่าความเพียรพยายามเชิงคำนวณได้สูงกว่า

1.3 วัตถุประสงค์ของการวิจัย

เสนอวิธีการควบคุมความหลากหลายของประชากรในขั้นตอนวิธีเชิงพันธุกรรมโดยการจับคู่แบบโอนเอียง

1.4 ขอบเขตของการวิจัย

ทำการวิจัยในกลุ่มของปัญหาซึ่งนิยมใช้เป็นเกณฑ์ทดสอบสำหรับขั้นตอนวิธีเชิงพันธุกรรม ทั้งนี้เพื่อประโยชน์ในการเปรียบเทียบผลการทดลองที่ได้กับงานวิจัยเดิม ปัญหาที่ใช้ในการทดลองได้แก่

1. ปัญหา one-max เป็นปัญหาพื้นฐานที่ง่ายและนิยมใช้ในการเปรียบเทียบประสิทธิภาพของขั้นตอนวิธีเชิงพันธุกรรม งานวิจัยที่มีการใช้ปัญหา one-max ได้แก่งานวิจัย [21,22]
2. ปัญหาฟังก์ชัน multimodal เป็นกลุ่มของปัญหาที่ประกอบไปด้วยบริเวณที่ดีที่สุดหลายบริเวณ ด้วยเหตุนี้ปัญหานี้จึงมีความยากในการหาบริเวณที่มีคำตอบที่ดีที่สุดที่แท้จริงของปัญหา งานวิจัยที่มีการใช้ปัญหาฟังก์ชัน multimodal ได้แก่งานวิจัย [10,16,23-31]
3. ปัญหาฟังก์ชัน deceptive เป็นปัญหาที่จัดอยู่ในกลุ่มปัญหายากสำหรับขั้นตอนวิธีเชิงพันธุกรรม ทั้งนี้เนื่องจากฟังก์ชันความเหมาะของปัญหาไม่ได้นำกระบวนการวิวัฒนาการไปถูกทาง การหาคำตอบที่ดีที่สุดอย่างแท้จริงจึงเป็นไปได้ยาก งานวิจัยที่มีการใช้ปัญหาฟังก์ชัน deceptive ได้แก่งานวิจัย [10,22-23,32-36]

1.5 งานที่นำเสนอในการวิจัย

งานที่นำเสนอในการวิจัยมีดังนี้

1. การคิดค้นการจับคู่แบบโอนเอียง (preference mating)

การจับคู่แบบโอนเอียงเป็นวิธีที่เพิ่มขยายมาจากการจับคู่แบบมีข้อกำหนด (restricted mating) วัตถุประสงค์ของการออกแบบการจับคู่แบบโอนเอียงคือความต้องการระบบที่ไม่ซับซ้อน สามารถที่จะเข้าใจได้ง่าย และต้องเป็นระบบที่สามารถสร้างความหลายหลากของประชากรมากน้อยต่างๆกันได้โดยสะดวก ทั้งนี้เพื่อประโยชน์ในการนำไปใช้สร้างระบบควบคุมความหลากหลายของประชากรต่อไป การจับคู่แบบโอนเอียงสามารถสร้างความหลากหลายของประชากรที่แตกต่างกันได้โดยการปรับเปลี่ยนเพียงค่า *ระดับโอนเอียง* เพียงค่าเดียว รายละเอียดของการจับคู่แบบโอนเอียงแสดงในหัวข้อที่ 3.1

2. การคิดค้นการสร้างประโยชน์ (contribution)

การสร้างประโยชน์เป็นมาตรวัด(measurement)ที่มีประสิทธิภาพในการชักนำระบบควบคุมความหลากหลายของประชากรให้สามารถวิวัฒนาการไปยังทิศทางของความหลากหลายของประชากรที่ถูกต้อง การออกแบบการสร้างประโยชน์เป็นผลมาจากการค้นพบว่ามาตรวัดประสิทธิภาพพื้นฐานของขั้นตอนวิธีเชิงพันธุกรรมได้แก่ค่าเฉลี่ยของค่าความเหมาะและค่าดีที่สุดของค่าความเหมาะมีข้อจำกัดในการใช้งานกับระบบควบคุมความหลากหลายของประชากร รายละเอียดของการสร้างประโยชน์แสดงในหัวข้อที่ 3.8

3. การคิดค้นระบบควบคุมความหลากหลายของประชากรที่สามารถปรับตัวได้ (adaptive diversity control system)

การคิดค้นระบบควบคุมความหลากหลายของประชากรที่ปรับตัวได้เป็นผลมาจากสมมติฐานว่า ปัญหาแต่ละปัญหามีความต้องการระดับของความหลากหลายของประชากรที่เหมาะสมแตกต่างกันไป งานวิจัยจำนวนมากได้เสนอวิธีการซึ่งมีกลไกของการรักษาความหลากหลายของประชากรเพื่อเพิ่มประสิทธิภาพในการหาคำตอบของปัญหา [4,9,10,37-39] หากแต่วิธีการเหล่านี้จำเป็นต้องใช้ความรู้ในการปรับตั้งค่าพารามิเตอร์ต่างๆซึ่งเกี่ยวข้องกับความหลากหลายของประชากรที่เหมาะสมกับปัญหาก่อนเริ่มการทดลอง ด้วยเหตุนี้ระบบควบคุมความหลากหลายของประชากรที่สามารถปรับตัวตามปัญหาเองได้จึงมีความน่าสนใจในการทำวิจัย รายละเอียดของระบบที่ออกแบบแสดงในหัวข้อที่ 3.9