

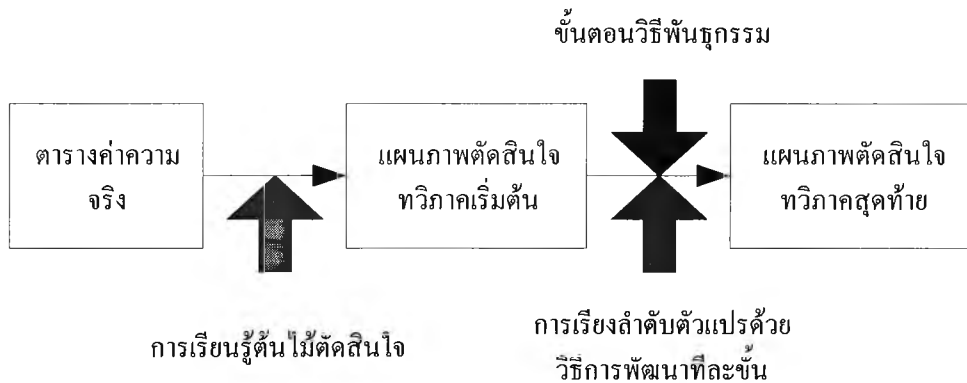


บทที่ 3

การสร้างลำดับของวิธีการพัฒนาที่ละชั้นจากขั้นตอนวิธีพันธุกรรม

ในบทนี้จะกล่าวถึงการปรับปรุงการสร้างแผนภาพตัดสินใจทวิภาคที่ได้ผ่านการหาลำดับตัวแปรเริ่มต้นจากการใช้เทคนิคการเรียนรู้ต้นไม้ตัดสินใจ[20] โดยการปรับปรุงในส่วนของการแลกเปลี่ยนลำดับตัวแปรของวิธีการพัฒนาที่ละชั้น เพื่อลดขนาดของแผนภาพตัดสินใจทวิภาคให้เล็กลง เนื่องจากการพัฒนาที่ละชั้นแต่ละวิธีเหมาะสมกับวงจรเชิงผสมต่างกัน โดยได้นำเสนอแนวทางในการพัฒนาคือ

การนำขั้นตอนวิธีพันธุกรรมมาเรียนรู้ลำดับของวิธีการพัฒนาที่ละชั้น เพื่อเพิ่มประสิทธิภาพในการหาลำดับตัวแปร ดังรูปที่ 3.1

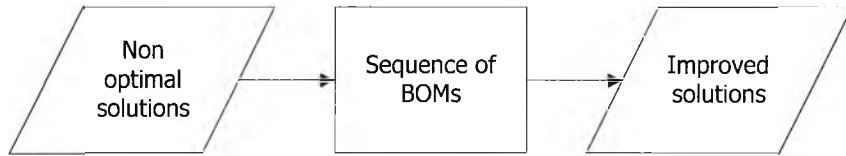


รูปที่ 3.1 การสร้างแผนภาพตัดสินใจทวิภาค

3.1 แบบจำลองการเรียนรู้ (Learning Model) และการประยุกต์ใช้ขั้นตอนวิธีพันธุกรรม

แบบจำลองการเรียนรู้ที่ตั้งอยู่ในสมมติฐานที่ว่า มีชุดของวิธีการที่สามารถปรับปรุงผลเฉลยที่ยังไม่ปรับปรุง (non optimal solution) ให้มีคุณภาพเพิ่มขึ้นได้ โดยเรียกวิธีการเหล่านี้ว่าหน่วยปรับปรุงพื้นฐาน (Basic Optimization Modules หรือ BOM) [16] ซึ่งประกอบด้วยวิธีศึกษาลำบาก (heuristics) มาเรียงลำดับกันเป็นขั้นเป็นตอน เป้าหมายของการศึกษาก็คือการค้นหาลำดับ

ของหน่วยปรับปรุงพื้นฐานที่ดีที่สุดที่ทำให้คุณภาพโดยรวมของตัวอย่างที่ใช้ในการเรียนรู้ทั้งหมดดีขึ้นดังรูปที่ 3.2



รูปที่ 3.2 การทำงานของหน่วยปรับปรุงพื้นฐาน (BOMs)

เมื่อใช้ขั้นตอนวิธีพันธุกรรมในการหาคำตอบจะแทนลำดับขั้นตอนของหน่วยปรับปรุงพื้นฐานซึ่งเป็นผลเฉลยด้วยสายโครโมโซม และใช้ฟังก์ชันความเหมาะสมวัดความเหมาะสมของแต่ละผลเฉลยดังสมการ

$$fit(s) = \sum_{i=0}^n quality(example_optimized_i)$$

โดยที่ $example_optimized_i$ แสดงตัวอย่างที่ i หลังจากผ่านกระบวนการปรับปรุงโดยใช้ลำดับขั้นตอน s ของหน่วยปรับปรุงพื้นฐาน (BOMs)

จากสมการ ข้างบนความเหมาะสมของผลเฉลยคือ ผลรวมคุณภาพของลำดับขั้นตอนการศึกษาสำนึกที่อยู่ในสายอักขระ s ที่ได้กระทำต่อชุดตัวอย่างจำนวน n ตัวอย่าง จะเห็นว่าการเลือกตัวอย่างเพื่อใช้สอนมีอิทธิพลต่อคุณภาพของการศึกษาสำนึกมาก ในการศึกษานี้จะแทนการศึกษาสำนึกด้วยวิธีการพัฒนาที่ละชั้น

3.2 การเข้ารหัสของคำตอบ (Representation)

ผลเฉลยในที่นี้คือวิธีการพัฒนาที่ละชั้นในแต่ละวิธีมาเรียงต่อกันเป็นสายลำดับ โดยผลเฉลยจะถูกเข้ารหัสแบบหลายค่า (multi-valued encoding) แทนการใช้สายอักขระเลขฐานสองเพื่อสร้างโครโมโซมที่ง่ายต่อการดำเนินการและความถูกต้องของผลเฉลย ดังรูปที่ 3.3 โดยแต่ละตำแหน่งในสายอักขระจะแทนวิธีการพัฒนาที่ละชั้น ถ้าสายอักขระมีทั้งหมด n อักขระจะมีวิธีการ

พัฒนาที่ละชั้นทั้งหมดที่เป็นไปได้ n วิธี ความยาวของโครโมโซมจะวัดจากจำนวนอักขระที่อยู่บนสายอักขระนั้น

วิธีการพัฒนาที่ละชั้นที่ใช้ในสายโครโมโซม ได้แก่ SIFTING, AD2, AD3 และ AD4 ในแต่ละวิธีได้กล่าวไว้แล้วในบทที่ 2 โดยจะแทนวิธีพัฒนาที่ละชั้นด้วยตัวอักษรภาษาอังกฤษและเลขฐานสิบ S, 4, 3, 2 และ N ดังต่อไปนี้ S แทน SIFTING, 4 แทน AD4, 3 แทน AD3 2 แทน AD2 และ N (no operation) หมายถึงไม่ต้องทำแลกเปลี่ยนตัวแปรด้วยวิธีใด ทำให้ลดการทำงานของชั้นตอนลง

S	S	3	N	4	2	3	N
---	---	---	---	---	---	---	---

รูปที่ 3.3 ตัวอย่างสายโครโมโซมแบบหลายค่า

3.3 การประเมินค่าความเหมาะสม (fitness evaluation)

ในการศึกษานี้ ตัวอย่างที่ใช้ในการเรียนรู้คือวงจรเปรียบเทียบสมรรถนะ MCNC ที่ใช้ในงานวิจัยของ ศิริพรรณ วิญญูรัตน์ทุกุล [20] เพื่อให้สามารถเปรียบเทียบผลการทดลองได้ โดยจะนำชุดวงจรเปรียบเทียบสมรรถนะมาสร้างแผนภาพตัดสินใจภาคด้วยเทคนิคการเรียนรู้ต้นไม้ตัดสินใจ ดังนั้นค่าความเหมาะสมของโครโมโซมจะวัดจากคุณภาพของสายลำดับวิธีพัฒนาที่ละชั้นที่สามารถทำให้บัพของแผนภาพตัดสินใจมีจำนวนน้อยลง ดังสมการ

$$quality = \frac{1}{nodes}$$

จากสมการคุณภาพ (quality) จะแปรผกผันกับจำนวนบัพ (node) เพื่อผลักดันให้ขั้นตอนวิธีพันธุกรรมหาลำดับขั้นตอนวิธีการพัฒนาที่ละชั้นที่เป็นผลให้ชุดแผนภาพตัดสินใจภาคโดยรวมมีขนาดเล็กลง

3.4 ตัวดำเนินการพันธุกรรม (Genetic Operators)

ตัวดำเนินการพันธุกรรมที่ใช้สร้างประชากรใหม่คือ การคัดเลือก (selection) การไขว้เปลี่ยน (crossover) และการกลายพันธุ์ (mutation)

3.4.1 การคัดเลือก

วิธีที่ใช้คือการคัดเลือกตามสัดส่วนของค่าความเหมาะสม (roulette wheel selection) นั่นคือแต่ละโครโมโซมจะมีโอกาสถูกเลือกตามสัดส่วนค่าความเหมาะสมของมัน นอกจากนี้โครโมโซมที่มีค่าความเหมาะสมมากที่สุดในประชากรรุ่นปัจจุบันจะถูกเลือกเก็บไว้ในประชากรรุ่นถัดไปเสมอ โดยไม่มีการเปลี่ยนแปลงใดๆ เพื่อรับประกันว่าผลเฉลยที่ดีที่สุดจะไม่สูญหายไป

3.4.2 การไขว้เปลี่ยน

ในการศึกษานี้ใช้วิธีการไขว้เปลี่ยนแบบหนึ่งจุดและการไขว้เปลี่ยนแบบสองจุดร่วมกันในอัตราส่วนหนึ่งต่อหนึ่งของอัตราการไขว้เปลี่ยน โดยอัตราการไขว้เปลี่ยนที่ใช้คือ 60% ของประชากรที่ได้รับการคัดเลือกทั้งหมด ดังนั้นอัตราการไขว้เปลี่ยนแบบหนึ่งจุดและอัตราการไขว้เปลี่ยนแบบสองจุดเท่ากับ 30% ของประชากรที่ได้รับการคัดเลือกทั้งหมด เหตุผลที่กำหนดอัตราการไขว้เปลี่ยนไว้สูงก็เพื่อให้เกิดโครโมโซมที่มีลักษณะใหม่เพิ่มขึ้น ส่วนขั้นตอนการไขว้เปลี่ยนจะเริ่มจากสุ่มเลือกโครโมโซมที่ไม่ซ้ำกันขึ้นมา 2 สาย จากนั้นทำการสุ่มเลือกจุดตัดแล้วแลกเปลี่ยนส่วนที่ถูกแบ่งออกจากโครโมโซมที่เป็นคู่กันดังที่กล่าวไว้ในบทที่ 2

3.4.3 การกลายพันธุ์

การทำงานจะเริ่มด้วยการสุ่มเปลี่ยนค่าอักขระของสายโครโมโซมในช่วงค่าอักขระที่กำหนดไว้ ตัวอย่างเช่น ถ้ากำหนดชุดของอักขระเป็น S, 4, 3, 2 และ N ค่าอักขระใหม่ที่จะเปลี่ยนจะต้องอยู่ในชุดอักขระนี้เท่านั้น และค่าใหม่จะต้องไม่ซ้ำกับค่าเดิม

โครโมโซมสุ่มเปลี่ยนค่าจากโครโมโซมที่คัดเลือกไว้ ประชากรที่ใช้มีขนาดกลางจึงตั้งอัตราการกลายพันธุ์ไว้ต่ำคือ 3.3% ของประชากรที่คัดเลือกไว้ เพื่อไม่ให้สูญเสียส่วนที่ดีของโครโมโซมไป และยังคงความหลากหลายไว้ [13]

3.5 ขั้นตอนวิธี

1. ประชากรเริ่มต้นสร้างจากสายอักขระแบบหลายค่าแบบสุ่ม สายอักขระมีความยาวคงที่ โดยความยาวของสายอักขระจะทดลองในบทต่อไป และกำหนดขนาดประชากรเท่ากับ 30 แล้วทำการวัดค่าความเหมาะสมของสายอักขระแต่ละตัว
2. เก็บสายอักขระที่ดีที่สุดไว้ในประชากรใหม่ และใช้ตัวดำเนินการพันธุกรรม การคัดเลือก การไขว้เปลี่ยน และการกลายพันธุ์สร้างประชากรใหม่ตามค่าความน่าจะเป็นที่กำหนดไว้ แล้วทำการวัดค่าความเหมาะสมของสายอักขระใหม่แต่ละตัว
3. นำประชากรใหม่แทนที่ประชากรเดิม
4. กลับไปทำซ้ำข้อ 2 และข้อ 3 จนกระทั่งครบ 100 รุ่น จึงหยุดการทำงาน ดังรูปที่ 3.4

```
genetic_algorithm(benchmark){
    generate_initial_population();
    calculate_fitness();
    {
        apply_operators_with_corresponding_probabilities();
        calculate_fitness();
        update_population();
    } while (generation < 100);
    return ;
}
```

รูปที่ 3.4 โครงสร้างพื้นฐานของขั้นตอนวิธี

คำตอบคือสายอักขระที่มีค่าความเหมาะสมมากที่สุดในประชากรรุ่นสุดท้าย