

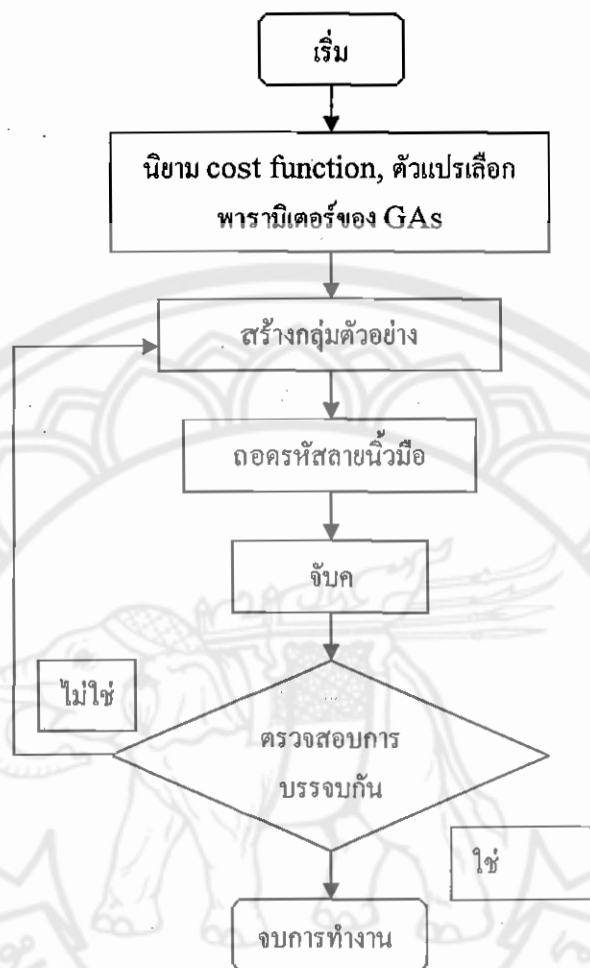
บทที่ 2

หลักการและทฤษฎี

ในการสืบทอดยีนนั้นได้ใช้ลักษณะเฉพาะของยีนเป็นตัวแปรสำคัญ ซึ่งในแต่ละคนนั้นต่างก็มียีนที่เป็นลักษณะเฉพาะของตัวเอง การสืบทอดยีนนั้นจำเป็นต้องมีกระบวนการอย่างมีระบบแบบแผน โดย Genetic Algorithms (GAs) เป็นกระบวนการหนึ่งที่มีระบบขั้นตอนที่เป็นแบบแผน สามารถใช้กับรูปแบบการสืบทอดยีนได้ ซึ่งในบทนี้ได้กล่าวถึงลักษณะโดยรวมของ GAs การเลือกเทคนิคที่ใช้ในการสืบทอดยีนเพื่อให้ได้ผลดีที่สุด และข้อได้เปรียบของ GAs ว่าดีกว่าวิธีอื่นอย่างไร

2.1 Genetic Algorithms (GAs)

Genetic Algorithms เป็นรูปแบบการคัดเลือกทางธรรมชาติ การทำงานของ GAs จะเกี่ยวกับการให้กำเนิดประชากร (Population) ปัญหาส่วนใหญ่เป็นการเปรียบเทียบโครโมโซม (Chromosomes) และแสดงสาย Binary Members โดย GAs จะเป็นการคาดหวังหาสถานะที่น้อยสุดในแต่ละสาเหตุที่ฟังก์ชันสุดท้าย ซึ่งเราจะรู้ค่าเริ่มต้นหรือจำนวนประชากรเริ่มต้นที่แน่นอน ระบบการทำงานของ GAs หรือเรียกว่า Binary Genetic Algorithm ซึ่งจะมีหลักการทำงานแสดงใน flowchart ในรูปที่ 2.1 โดยมีรายละเอียดดังต่อไปนี้



รูปที่ 2.1 ขั้นตอนการทำงานของ Genetic Algorithms

2.1.1 การเลือก Cost Function และตัวแปร

โดยทั่วไป cost function จะเป็นการหาค่าที่เราต้องการทราบ ในกรณีของการหาลายนิ้วมือที่ต้องการนั้น การหา cost function จะต้องเลือกตัวแปรให้เหมาะสม โดยตัวแปรในกรณีนี้ได้แก่ ขนาดของลายนิ้วมือ ลายเส้นแต่ละเส้นของลายนิ้วมือ เป็นต้น

2.1.2 การสร้างกลุ่มตัวอย่างเริ่มต้น (Generate initial population)

กลุ่มตัวอย่างของลายนิ้วมือที่มีอยู่ที่ถูกสุ่มขึ้นมาเป็นกลุ่มตัวอย่างเพื่อใช้ในการคัดเลือก

2.1.3 การเปลี่ยนลายนิ้วมือให้เป็นรหัสและถอดรหัส (Variable Encoding and Decoding)

ก่อนที่จะเริ่มทำงานโดยใช้โปรแกรม MATLAB นั้นเราจะต้องเข้ารหัสและถอดรหัสก่อน ในการทดสอบนี้ ลายนิ้วมือจะถูกแปลงออกมาแล้วนำเข้าเครื่องคอมพิวเตอร์ แต่ค่าที่ถูกแปลงออกมานี้จะยังไม่สามารถคำนวณในโปรแกรมได้เลย เราจึงต้องทำการถอดรหัสก่อน (รหัส 0 กับ 1) เพื่อให้เครื่องสามารถที่จะรับค่าไปคำนวณต่อไปได้

2.1.4 การจับคู่ (Selection mate)

วิธีการจับคู่ของลายนิ้วมือจะมีด้วยกัน 2 วิธี ดังนี้

2.1.4.1 การคัดสรรโดยธรรมชาติ (Natural Selection)

จะเป็นการเลือกค่าที่ดีที่สุด (optimum) เพื่อดำเนินการต่อไป ส่วนค่าที่ไม่ดีที่สุดจะถูกตัดทิ้ง โดยข้อดีของการคัดสรรโดยธรรมชาติคือ ไม่ต้องทำการคัดเลือกกลุ่มตัวอย่างเพียงแต่นำค่าที่มีอยู่แล้ว มาเลือกค่าที่ดีที่สุด

2.1.4.2 การคัดสรรแบบทั่วไป

จะเป็นการเลือกคู่ของลายนิ้วมือ โดยการใช้วิธีการคัดเลือกลายนิ้วมือที่จะมาเข้าคู่กัน มาด้วยกัน 4 วิธี คือ

- การจับคู่จากบนลงล่าง (Pairing form top to bottom)
จะเป็นการจับคู่กันของลายนิ้วมือของแต่ละแถว เช่น การจับคู่ระหว่างแถวพ่อ 1, 3, 5,... กับแถวแม่ 2, 4, 6,... เป็นต้น ซึ่งวิธีนี้จะเป็นรูปแบบการคัดเลือกที่ไม่ดีนักแต่จะง่ายต่อการทำโปรแกรม
- การจับคู่แบบสุ่ม (Random pairing)
จะใช้การสุ่มในการเลือกคู่ของลายนิ้วมือแต่ละลายนิ้วมือ
- การจับคู่แบบสุ่มถ่วงน้ำหนักอันดับ (Rank weighted random pairing)
เป็นการออกแบบการจับคู่ของลายนิ้วมือ โดยเป็นการคัดเลือกที่อิสระ คือเราสามารถเลือกค่าความน่าจะเป็นในการคัดเลือกได้จาก

$$P = \frac{N_{keep} - n + 1}{\sum_{n=1}^{N_{keep}} n}$$

P คือ ค่าความน่าจะเป็น

N_{keep} คือ จำนวนตัวอย่างลายนิ้วมือทั้งหมดใน 1 กลุ่ม

n คือ ลำดับลายนิ้วมือ

ตัวอย่างการสุ่มเลือกค่าความน่าจะเป็นแสดงในตารางที่ 2.1

ตารางที่ 2.1 แสดงการจับคู่แบบสุ่มล่วงหน้าหน้าอันดับ

n	finger code	P_n	$\sum P_n$
1	1010101	0.4	0.4
2	1001000	0.3	0.7
3	1011001	0.2	0.9
4	1011000	0.1	1.0

ถ้าเลือกค่าความน่าจะเป็น $P_n = 0.6$ แล้ว ค่า P_n นั้นตกอยู่ในช่วง $0.4 < P_n < 0.7$ ซึ่งเราทำการทดสอบเพียงช่วง $n = 1$ และ 2 เท่านั้นถ้าเป็นลายนิ้วมือที่ตรงกันก็จะหยุดการคัดเลือก แต่ถ้าเป็นลายนิ้วมือแตกต่างกันก็จะทำการเลือกค่าความน่าจะเป็นใหม่แล้วทำการทดสอบซ้ำเดิม จนกว่าจะได้ลายนิ้วมือที่ตรงกัน

- การคัดสรรแบบแข่งขัน (Tournament Selection)

เป็นการแบ่งกลุ่มจำนวนลายนิ้วมือออกเป็นกลุ่มแล้วทำการเปรียบเทียบค่าที่ใกล้เคียงกับลายนิ้วมือทดสอบมากที่สุด จากนั้นแยกออกมาแล้วนำไปรวมกับค่าที่ใกล้เคียงของกลุ่มย่อยอื่นๆทำเช่นนี้ไปเรื่อยๆจนกว่าจะได้ลายนิ้วมือที่มีความใกล้เคียงที่สุด

จากวิธีการเลือกที่ได้ศึกษามานั้นการคัดสรรแบบแข่งขันและการจับคู่แบบสุ่มล่วงหน้าหน้าอันดับเป็นที่นิยมใช้ใน GAs มากที่สุดเนื่องจากเหมาะในการหาค่าที่ optimum จากกลุ่มตัวอย่างที่มีขนาดใหญ่

2.1.5 การตรวจสอบการบรรจบกัน (Convergence check)

เป็นการเปรียบเทียบลายนิ้วมือว่ามีความใกล้เคียงกับลายนิ้วมือตัวอย่างหรือไม่ ถ้าตรงกันก็จะทำการหยุดการทดสอบ แต่ถ้าไม่มีลายนิ้วมือใดๆ ในกลุ่มตัวอย่างที่ตรงกับลายนิ้วมือตัวอย่างเลข เราก็จะทำการสุ่มกลุ่มตัวอย่างใหม่แล้วทำการทดสอบซ้ำอีกครั้ง

2.2 การเลือกเทคนิคเพื่อให้ได้ผลสูงที่สุด

เทคนิคในการคัดเลือกค่าที่ดีที่สุด (optimization) มีหลายชนิดด้วยกัน ได้แก่

2.2.1 Exhaustive techniques

สามารถค้นหาได้ทั่วทั้งหมดแต่การคำนวณที่ได้มีการยับยั้งเพราะขอบเขตของการค้นหาและการพิจารณาที่เวลาจริง

2.2.2 Calculus-based techniques

เป็นวิธีการค้นหาทางคณิตศาสตร์โดยฟังก์ชันที่ใช้ในการค้นหานั้นใช้ได้กับรูปแบบที่ไม่ซับซ้อนมาก ซึ่งความไม่ต่อเนื่องและความซับซ้อนของรูปร่างนั้นเป็นอุปสรรคต่อการค้นหาผลลัพธ์

2.2.3 Partial knowledge techniques

เป็นเทคนิคที่รู้ค่าเป็นบางส่วน โดยจะรู้เพียงค่าเริ่มต้นของการค้นหาเท่านั้นและในกระบวนการค้นหานั้นไม่มีความแม่นยำเพียงพอ ซึ่งเทคนิคนี้ไม่สามารถกำหนดระยะจุดสิ้นสุดของเป้าหมายที่ชัดเจนได้ รวมทั้งค่าเครื่องมือคำนวณมีราคาแพง

2.2.4 Knowledge-based techniques

เป็นเทคนิคความรู้พื้นฐาน ซึ่งระบบนี้จำกัดขอบเขตของการนำไปใช้งานจริง มีรูปแบบของสมการที่กระจัดกระจายและการหาสมการที่นำไปใช้ได้ยาก รวมทั้งลักษณะเฉพาะที่ใช้นั้นไม่สามารถใช้กับในระบบนี้

2.2.5 Hierarchical techniques

เป็นเทคนิคที่ใช้การปรับแก้ค่าการกระจายเริ่มต้นเพื่อให้ระยะการกระจายของผลลัพธ์ลดลง ซึ่งผลลัพธ์ที่ได้นั้นเป็นช่วงของการกระจายที่แคบที่สุด แต่อย่างไรก็ตามเทคนิคนี้ไม่สามารถหาค่าที่ดีที่สุดของการแก้ปัญหาได้

2.2.6 Genetic algorithms

วิธีนี้เป็นวิธีที่ดีที่สุดในการค้นหาค่าที่ใกล้เคียงที่สุดซึ่งจะดีกว่า 5 ข้อที่กล่าวมาข้างบนและจะไม่มีข้อเสียของทั้ง 5 ข้อข้างบนนั้นด้วย ในการศึกษาจึงเลือกเทคนิคนี้ในการหาเลขนี้ว่ามี