

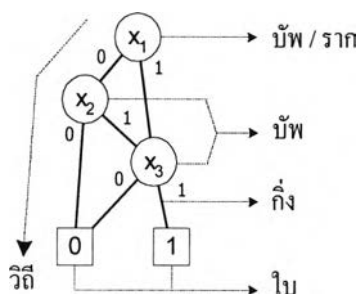
## บทที่ 2

### ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีและงานวิจัยต่างๆ ที่นำมาประยุกต์ใช้ เพื่อออกแบบขั้นตอนการปรับปรุงแผนภาพตัดสินใจทวิภาคที่สร้างโดยเทคนิคการเรียนรู้ต้นไม้ตัดสินใจ ด้วยการใช้ขั้นตอนวิธีพันธุกรรมร่วมกับวิธีการพัฒนาที่ละชั้น โดยจะกล่าวถึงลักษณะของแผนภาพตัดสินใจทวิภาค [1-3] การสร้างแผนภาพตัดสินใจทวิภาคโดยเทคนิคการเรียนรู้ต้นไม้ตัดสินใจ[20] การพัฒนาที่ละชั้น[6,7] และขั้นตอนวิธีพันธุกรรม

#### 2.1 แผนภาพตัดสินใจทวิภาค (Binary Decision Diagrams : BDDs)

แผนภาพตัดสินใจทวิภาคเป็นโครงสร้างข้อมูลแบบกราฟไม่มีวงที่ระบุทิศทาง (Directed Acyclic Graph : DAG) ซึ่งใช้แทนฟังก์ชันบูลีน [1-4] โดยมีส่วนประกอบดังรูปที่ 2.1 คือ



รูปที่ 2.1 แผนภาพตัดสินใจทวิภาค ของฟังก์ชัน  $f(x_1, x_2, x_3) = (x_1 + x_2) \cdot x_3$

- (1) บัพ (node / non-terminal) แสดงถึง ตัวแปรของฟังก์ชันบูลีน โดยเรียกบัพแรกว่าราก
- (2) กิ่ง (branch / edge) แสดงถึง ค่าของตัวแปร โดยแต่ละตัวแปรจะมี 2 กิ่งเสมอ คือ กิ่ง 0 และ 1
- (3) ใบ (leaf / terminal) แสดงถึง ค่าคำตอบของฟังก์ชันบูลีน ซึ่งมีค่าที่เป็นไปได้คือ 0 และ 1

- (4) วิถี (path) คือ เส้นทางที่เชื่อมต่อกันของบัพและกึ่งจากรากถึงใบ โดยทุกวิถีของแผนภาพตัดสินใจทวิภาคจะต้องมีลำดับของตัวแปรที่เหมือนกันทั้งหมด
- (5) ขนาด (size) คือ จำนวนบัพทั้งหมดของแผนภาพตัดสินใจทวิภาค

การสร้างแผนภาพตัดสินใจทวิภาคอย่างง่าย สามารถทำได้โดยการเปลี่ยนรูปแบบการแทนตารางค่าความจริงที่ได้จากฟังก์ชันบูลีนไปเป็นต้นไม้ ซึ่งในทุกวิถีของต้นไม้จะต้องมีลำดับของตัวแปรที่เหมือนกัน จากนั้นก็นำต้นไม้ดังกล่าวมาลดทอนส่วนที่ซ้ำกัน ได้เป็นแผนภาพตัดสินใจทวิภาค โดยการใช้กฎการลดทอน (reduction rule) ดังนี้

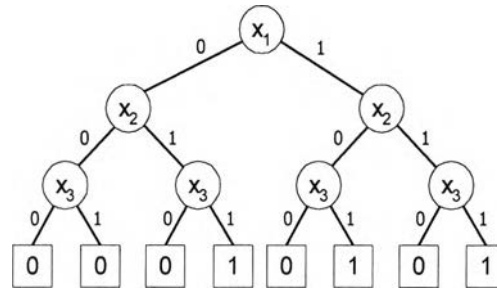
1. กำจัดใบที่ซ้ำกัน (remove duplicate terminals) เป็นการยุบใบที่ซ้ำกันให้เหลือเพียงใบเดียว และเปลี่ยนทิศทางของกิ่งที่เชื่อมต่อกับใบที่ถูกยุบไปเชื่อมต่อกับใบที่เหลือแทน
2. กำจัดบัพที่ซ้ำกัน (remove duplicate non-terminals) เป็นการยุบบัพที่ซ้ำกันให้เหลือเพียงบัพเดียว และเปลี่ยนทิศทางของกิ่งที่เชื่อมต่อกับบัพที่ถูกยุบไปเชื่อมต่อกับบัพที่เหลือแทน
3. กำจัดส่วนที่ซ้ำซ้อน (remove redundant tests) เป็นการตัดบัพที่กิ่ง 0 และ 1 ของมันเชื่อมต่อกันเหมือนกัน และเปลี่ยนทิศทางของกิ่งที่เชื่อมต่อกับบัพที่ถูกตัดให้เชื่อมต่อกันโดยตรงแทน

ตารางที่ 2.1 ตารางค่าความจริงของฟังก์ชัน  $f(x_1, x_2, x_3) = (x_1 + x_2) \cdot x_3$

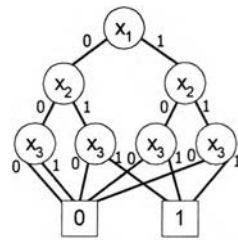
$x_1$	$x_2$	$x_3$	$f$
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

จากตัวอย่างตารางค่าความจริงในตารางที่ 2.1 สามารถเปลี่ยนรูปแบบการแทนข้อมูลเป็นต้นไม้ที่มีลำดับตัวแปร คือ  $(x_1, x_2, x_3)$  ได้ดังรูปที่ 2.2 โดยตารางค่าความจริงจะแสดงทุกค่าค่า

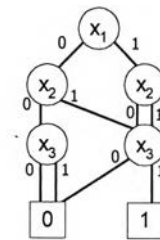
ตอบของฟังก์ชันสำหรับตัวแปร  $x_1$ ,  $x_2$  และ  $x_3$  ตามแนวตั้ง แต่แผนภาพตัดสินใจจะแสดงทุกค่าคำตอบของฟังก์ชันสำหรับตัวแปร  $x_1$ ,  $x_2$  และ  $x_3$  ตามแนวนอน หลังจากนั้นจะทำการลดรูปตามกฎลดทอนได้แผนภาพตัดสินใจวิภาคดังรูปที่ 2.3



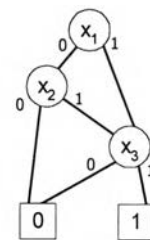
รูปที่ 2.2 ต้นไม้ที่มีลำดับตัวแปร  $(x_1, x_2, x_3)$  ของฟังก์ชัน  $f(x_1, x_2, x_3) = (x_1 + x_2) \cdot x_3$



(1) กำจัดใบที่ซ้ำกัน



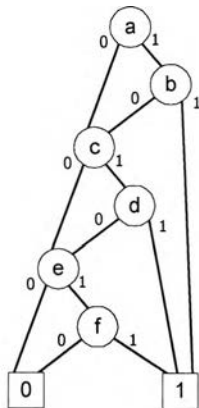
(2) กำจัดบัพที่ซ้ำกัน



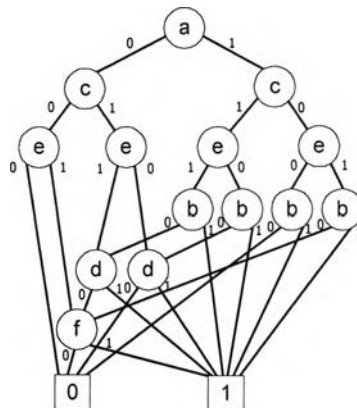
(3) กำจัดส่วนที่ซ้ำซ้อน

รูปที่ 2.3 การลดทอนส่วนที่ซ้ำของกราฟสำหรับฟังก์ชัน  $f(x_1, x_2, x_3) = (x_1 + x_2) \cdot x_3$

การสร้างแผนภาพตัดสินใจวิภาคด้วยวิธีดังกล่าวสามารถสร้างได้หลายแบบ ตัวอย่างเช่น แผนภาพตัดสินใจวิภาคสำหรับฟังก์ชัน  $f(a,b,c,d,e,f) = a \cdot b + c \cdot d + e \cdot f$  เดียวกัน แต่ใช้ลำดับของตัวแปรที่ต่างกันนั่นคือ แผนภาพตัดสินใจวิภาคที่มีลำดับตัวแปร คือ  $(a,b,c,d,e,f)$  จะมีขนาดเท่ากับ 6 บัพ ส่วนแผนภาพตัดสินใจวิภาคมีลำดับตัวแปร คือ  $(a,c,e,b,d,f)$  จะมีขนาดของแผนภาพตัดสินใจเท่ากับ 14 บัพ ดังรูปที่ 2.4(1) และ 2.4(2) ตามลำดับ ซึ่งจะเห็นว่าลำดับตัวแปรมีผลกระทบต่อขนาดของแผนภาพตัดสินใจวิภาค ดังนั้นสิ่งที่สำคัญของการสร้างแผนภาพตัดสินใจวิภาคที่มีขนาดเล็กก็คือ การหาลำดับตัวแปรที่ดี



(1) ลำดับตัวแปรคือ (a, b, c, d, e, f)



(2) ลำดับตัวแปรคือ (a, c, e, b, d, f)

รูปที่ 2.4 แผนภาพตัดสินใจทวิภาคของฟังก์ชัน  $f(a,b,c,d,e,f) = a \cdot b + c \cdot d + e \cdot f$

## 2.2 การเรียนรู้ต้นไม้ตัดสินใจ (Decision Tree Learning : DTL)

การเรียนรู้ต้นไม้ตัดสินใจ (Decision Tree Learning - DTL) เป็นการเรียนรู้จากตัวอย่างที่เรียกว่าข้อมูลสอน (training data) และแสดงความรู้ที่ได้ในรูปของต้นไม้ตัดสินใจ [13-14]

ข้อมูลสอนมีลักษณะเป็นตาราง ในแต่ละแถวแสดงถึงตัวอย่าง และสดมภ์แสดงถึงคุณลักษณะซึ่งมีอยู่ 2 ชนิดคือ

1. คุณลักษณะแบ่งพวก (category attribute) เป็นคุณลักษณะที่กำหนดว่าตัวอย่างถูกจัดอยู่พวกใด (class) และมีเพียงคุณลักษณะเดียวในแต่ละชุดข้อมูลสอน
2. คุณลักษณะไม่แบ่งพวก (non-category attribute) เป็นคุณลักษณะที่บอกถึงลักษณะต่างๆ ของตัวอย่าง และสามารถมีได้หลายคุณลักษณะ

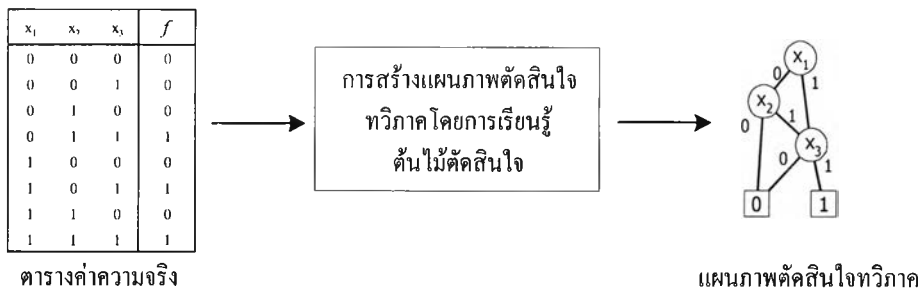
ต้นไม้ตัดสินใจสามารถสร้างขึ้นโดยการเลือกคุณลักษณะไม่แบ่งพวกเป็นบัพ เพื่อแยกตัวอย่างเป็นกลุ่มย่อยๆ และตัวอย่างส่วนใหญ่ในแต่ละกลุ่มย่อยเป็นพวกเดียวกันมากที่สุด (ในกรณีของแผนภาพตัดสินใจทวิภาค ตัวอย่างคือข้อมูลในแต่ละแถวของตารางค่าความจริงของฟังก์ชันบูลีนและพวกคือ ค่าของฟังก์ชัน (0 หรือ 1)) โดยใช้ฟังก์ชัน Gain เป็นเครื่องมือวัดความสามารถในการแยกตัวอย่าง ค่า Gain ที่น้อยแสดงถึงการแยกตัวอย่างที่ดี ฟังก์ชัน Gain แสดงดังด้านล่างนี้

$$Gain(node) = \sum_b \left[ \left( \frac{n_b}{n_t} \right) \times \left( \sum_c - \frac{n_{bc}}{n_b} \log_2 \frac{n_{bc}}{n_b} \right) \right]$$

โดยที่  $n_b$  เป็นจำนวนตัวอย่างในกิ่ง  $b$ ,  $n_t$  เป็นจำนวนตัวอย่างทั้งหมดของทุกกิ่งรวมกัน และ  $n_{bc}$  เป็นจำนวนตัวอย่างของกลุ่ม  $c$  ในกิ่ง  $b$  ในแต่ละขั้นตอน ตัวแปรทุกตัวจะถูกทดลองและตัวแปรที่ให้ค่า Gain น้อยที่สุดจะถูกเลือกเป็นบัพในการสร้างต้นไม้ตัดสินใจ กระบวนการนี้จะทำไปจนกระทั่งตัวอย่างในแต่ละกลุ่มย่อยเป็นกลุ่มเดียวกันทั้งหมด

### 2.3 การสร้างแผนภาพตัดสินใจทวิภาคโดยเทคนิคการเรียนรู้ต้นไม้ตัดสินใจ

เมื่อนำการเรียนรู้ต้นไม้ตัดสินใจมาประยุกต์กับการสร้างแผนภาพตัดสินใจทวิภาคสามารถทำได้ดังรูปที่ 2.5 โดยการแทนข้อมูลสอนของการเรียนรู้ต้นไม้ตัดสินใจด้วยตารางค่าความจริง [20] ซึ่งใช้แสดงทุกตัวอย่างที่เป็นไปได้ทั้งหมดของฟังก์ชันบูลีน และแทนต้นไม้ตัดสินใจที่ได้จากการเรียนรู้ด้วยแผนภาพตัดสินใจทวิภาค



รูปที่ 2.5 โครงสร้างของการสร้างแผนภาพตัดสินใจทวิภาคโดยการเรียนรู้ต้นไม้ตัดสินใจ

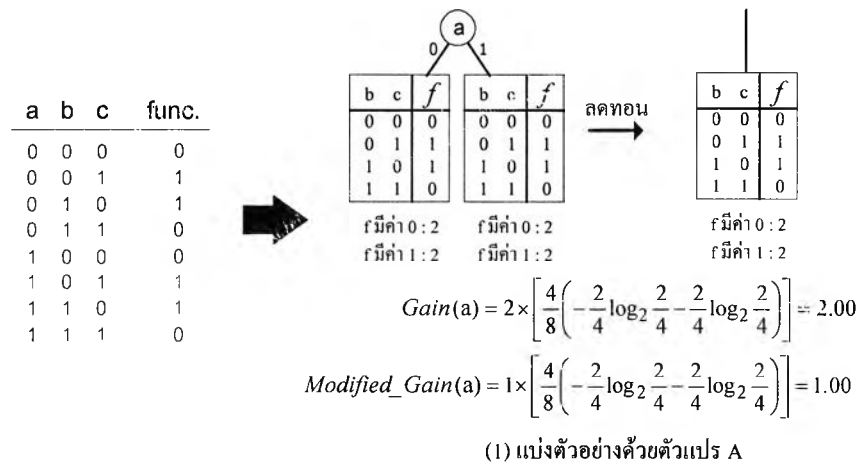
การประยุกต์ใช้การเรียนรู้ต้นไม้ตัดสินใจกับการสร้างแผนภาพตัดสินใจทวิภาค มีสิ่งที่จะต้องพิจารณาเพิ่มขึ้น คือ ในกรณีที่แยกตัวอย่างออกเป็นกลุ่มย่อย แล้วมีกลุ่มย่อยมากกว่าหนึ่งกลุ่มที่เหมือนกัน ซึ่งในกรณีนี้แผนภาพตัดสินใจทวิภาคจะแสดงเพียงกลุ่มย่อยเดียวเท่านั้น (ตามคุณสมบัติการลดรูปในหัวข้อที่ 2) ดังนั้นเมื่อคำนวณค่าความสามารถในการแยกตัวอย่างจะพิจารณาเพียงกลุ่มย่อยเดียวเช่นกัน และเรียกค่าที่คำนวณใหม่นี้ว่า Modified\_Gain ดังรูปที่ 2.6

ขั้นตอนของการสร้างแผนภาพตัดสินใจทวิภาคโดยการเรียนรู้ต้นไม้ตัดสินใจ สามารถทำได้ดังนี้

**ขั้นตอนที่ 1** ทดลองแบ่งตารางค่าความจริงและคำนวณค่า Modified\_Gain สำหรับทุกตัวแปร โดยแต่ละตัวแปรจะถูกใช้แบ่งทุกฟังก์ชันของตารางค่าความจริง ออกเป็นตารางค่าความจริงย่อยตามค่าของตัวแปร จากนั้นจะลดทอนส่วนที่ซ้ำด้วยการกำจัดตารางค่าความจริงย่อยที่ซ้ำกันและบันทึกซ้ำซ้อนกันออกไป แล้วนำตารางค่าความจริงย่อยที่ได้จากการลดทอนไปคำนวณค่า Modified\_Gain สำหรับตัวแปรนั้น

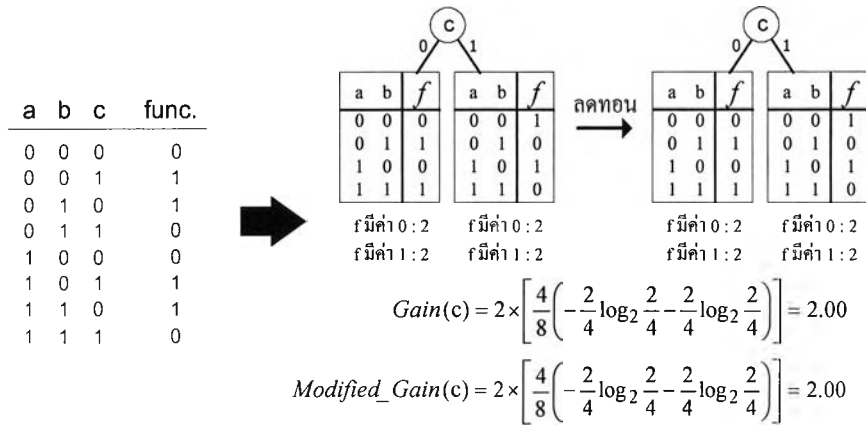
**ขั้นตอนที่ 2** เลือกตัวแปรที่มีค่า Modified\_Gain น้อยที่สุดเป็นลำดับถัดไปของแผนภาพตัดสินใจทวิภาค

โดยทั้ง 2 ขั้นตอนการสร้างข้างต้นจะถูกทำซ้ำ จนกระทั่งในแต่ละตารางค่าความจริงย่อยมีค่าคำตอบของฟังก์ชันเหมือนกันทั้งหมด



**หมายเหตุ :** f มีค่า x : y หมายถึง ฟังก์ชัน f มีค่าคำตอบเป็น x ทั้งหมด y ตัวอย่าง

รูปที่ 2.6 ตัวอย่างการคำนวณค่าฟังก์ชัน Gain และ ฟังก์ชัน Modified Gain



(2) แบ่งตัวอย่างด้วยตัวแปร C

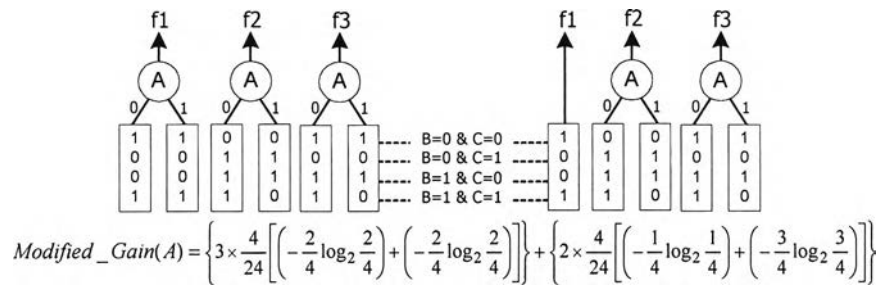
หมายเหตุ : f มีค่า x : y หมายถึง ฟังก์ชัน f มีค่าคำตอบเป็น x ทั้งหมด y ตัวอย่าง

รูปที่ 2.6 ตัวอย่างการคำนวณค่าฟังก์ชัน Gain และ ฟังก์ชัน Modified Gain (ต่อ)

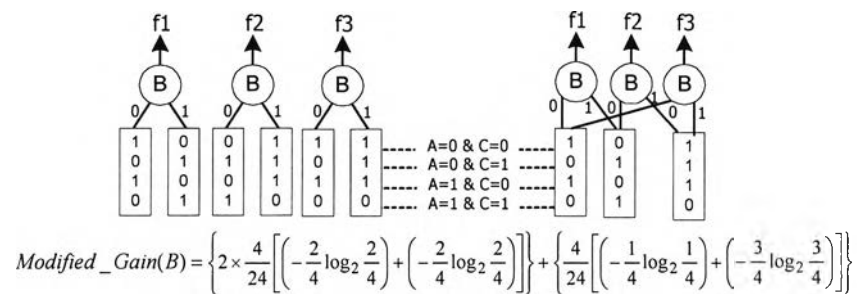
ตารางที่ 2.2 ตารางค่าความจริงของฟังก์ชัน  $f_1 = \overline{(B \oplus C)}$   $f_2 = (B \oplus C) + \overline{A}B$  และ  $f_3 = \overline{C} + \overline{A}B$

A	B	C	$f_1$	$f_2$	$f_3$
0	0	0	1	0	1
0	0	1	0	1	0
0	1	0	0	1	1
0	1	1	1	1	1
1	0	0	1	0	1
1	0	1	0	1	0
1	1	0	0	1	1
1	1	1	1	0	0

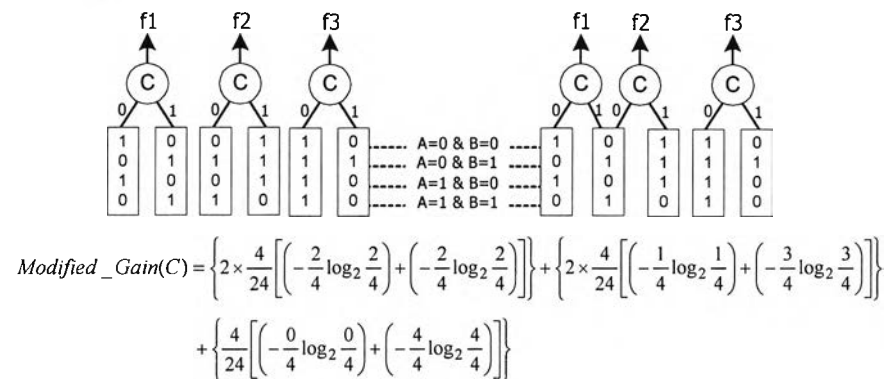
จากตัวอย่างของตารางค่าความจริงในตารางที่ 2.2 สามารถนำมาสร้างแผนภาพตัดสินใจ ทวิภาคโดยการเรียนรู้ต้นไม้ตัดสินใจได้ โดยการทดลองแบ่งตัวอย่างสำหรับแต่ละตัวแปร A B และ C จะได้ดังรูปที่ 2.7



Modified\_Gain(A) = 0.7705 และจำนวนโหนด = 2



Modified\_Gain(B) = 0.4686 และจำนวนโหนด = 3



Modified\_Gain(C) = 0.6038 และจำนวนโหนด = 3

รูป 2.7 ตัวอย่างการประยุกต์ใช้การเรียนรู้ต้นไม้ตัดสินใจ

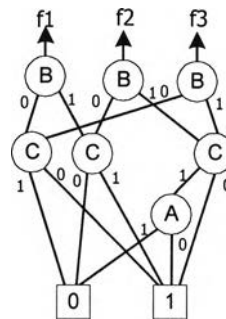
ในการเลือกบัพแรกของแผนภาพตัดสินใจทวิภาค โดยการทดลองสร้างต้นไม้ตัดสินใจ และคำนวณค่า Modified\_Gain สำหรับแต่ละตัวแปร เช่น

ในกรณีที่ทดลองใช้ตัวแปร A พบว่าฟังก์ชัน f1 แยกข้อมูลเป็นกลุ่มย่อยที่เหมือนกัน แต่ฟังก์ชัน f2 และ f3 แยกข้อมูลเป็นกลุ่มย่อยที่ไม่เหมือนกันดังนั้น



$$\begin{aligned} \text{Modified\_Gain}(A) &= \left\{ \frac{4}{24} \times \left( -\frac{2}{4} \log_2 \frac{2}{4} - \frac{2}{4} \log_2 \frac{2}{4} \right) \right\} + \left\{ \frac{4}{24} \times \left( -\frac{1}{4} \log_2 \frac{1}{4} - \frac{3}{4} \log_2 \frac{3}{4} \right) \right\} + \\ & \left\{ \frac{4}{24} \times \left( -\frac{2}{4} \log_2 \frac{2}{4} - \frac{2}{4} \log_2 \frac{2}{4} \right) \right\} + \left\{ \frac{4}{24} \times \left( -\frac{1}{4} \log_2 \frac{1}{4} - \frac{3}{4} \log_2 \frac{3}{4} \right) \right\} + \left\{ \frac{4}{24} \times \left( -\frac{2}{4} \log_2 \frac{2}{4} - \frac{2}{4} \log_2 \frac{2}{4} \right) \right\} = 0.7705 \end{aligned}$$

หลังจากที่ทดลองกับทุกๆ ตัวแปร แล้วตัวแปร B จะถูกเลือกเป็นตัวแปรลำดับแรกของ BDD ด้วยค่า Modified\_Gain ที่น้อยที่สุด และทำเช่นนี้ต่อไปเพื่อหาตัวแปรลำดับที่ 2 และ 3 ของ BDD ซึ่งจะได้เป็น C และ A ตามลำดับโดย BDD ที่สมบูรณ์สามารถแสดงได้ดังรูปที่ 2.8



รูปที่ 2.8 แผนภาพตัดสินใจทวิภาคที่ได้จากการสร้างด้วยการเรียนรู้ต้นไม้ตัดสินใจ

## 2.4 วิธีการพัฒนาทีละขั้น (Gradual Improvement)

การพัฒนาทีละขั้นเป็นวิธีที่ใช้หาลำดับของตัวแปรที่ดี ภายหลังจากการสร้างแผนภาพตัดสินใจทวิภาคโดยใช้การแลกเปลี่ยนลำดับของตัวแปร เพื่อลดขนาดของแผนภาพตัดสินใจทวิภาคให้เล็กลงอีก โดยขั้นตอนวิธีสำหรับการพัฒนาทีละขั้นนั้นมีหลายวิธีคือ AD2 AD3 AD4 และ SIFTING [18]

1. ขั้นตอนวิธี AD2, AD3 และ AD4 (AD<sub>m</sub> (m=2, 3 และ 4)) เป็นการแลกเปลี่ยนตัวแปร m ตัวที่อยู่ติดกันของแผนภาพตัดสินใจ โดยการเรียงสับเปลี่ยนลำดับตัวแปร m ตัวนั้น ซึ่งจำนวนของลำดับตัวแปรทั้งหมดมีค่าเป็น m! จากนั้นจะเลือกแผนภาพที่เล็กที่สุดใน m! นั้น กระบวนการนี้จะถูกทำซ้ำจนกระทั่งไม่สามารถหาแผนภาพตัดสินใจที่เล็กกว่าได้ [6]

ขั้นตอนวิธี AD2 ( $m = 2$ ) เป็นการแลกเปลี่ยนตัวแปร 2 ตัวที่อยู่ติดกัน ซึ่งจำนวนลำดับตัวแปรของการเรียงสับเปลี่ยนมีค่าเท่ากับ 2 ลำดับ รูปที่ 2.9 แสดงตัวอย่างการเรียงสับเปลี่ยนของขั้นตอนวิธี AD2 สำหรับตัวแปร  $x_2$  และ  $x_3$

$$\text{การเรียงสับเปลี่ยนของตัวแปร } (x_2, x_3) = \{ (x_2, x_3) (x_3, x_2) \}$$

$x_1, x_2, x_3, x_4, x_5, x_6, x_7$	เริ่มต้น
$x_1, x_3, x_2, x_4, x_5, x_6, x_7$	สลับที่ $(x_2, x_3)$

รูปที่ 2.9 การเรียงสับเปลี่ยนของขั้นตอนวิธี AD2 สำหรับตัวแปร  $x_2$  และ  $x_3$

ขั้นตอนวิธี AD3 ( $m = 3$ ) เป็นการแลกเปลี่ยนตัวแปร 3 ตัวที่อยู่ติดกัน ซึ่งจำนวนลำดับตัวแปรของการเรียงสับเปลี่ยนมีค่าเท่ากับ 6 ลำดับ รูปที่ 2.10 แสดงตัวอย่างการเรียงสับเปลี่ยนของขั้นตอนวิธี AD3 สำหรับตัวแปร  $x_3$   $x_4$  และ  $x_5$

$$\begin{aligned} \text{การเรียงสับเปลี่ยนของตัวแปร } (x_3, x_4, x_5) = \{ & (x_3, x_4, x_5) (x_3, x_5, x_4) (x_5, x_3, x_4) \\ & (x_5, x_4, x_3) (x_4, x_5, x_3) (x_4, x_3, x_5) \} \end{aligned}$$

$x_1, x_2, x_3, x_4, x_5, x_6, x_7$	เริ่มต้น
$x_1, x_2, x_3, x_5, x_4, x_6, x_7$	สลับที่ $(x_4, x_5)$
$x_1, x_2, x_5, x_3, x_4, x_6, x_7$	สลับที่ $(x_3, x_5)$
$x_1, x_2, x_5, x_4, x_3, x_6, x_7$	สลับที่ $(x_3, x_4)$
$x_1, x_2, x_4, x_5, x_3, x_6, x_7$	สลับที่ $(x_5, x_4)$
$x_1, x_2, x_4, x_3, x_5, x_6, x_7$	สลับที่ $(x_5, x_3)$

รูปที่ 2.10 การเรียงสับเปลี่ยนของขั้นตอนวิธี AD3 สำหรับตัวแปร  $x_3$   $x_4$  และ  $x_5$

ขั้นตอนวิธี AD4 ( $m = 4$ ) เป็นการแลกเปลี่ยนตัวแปร 4 ตัวที่อยู่ติดกัน ซึ่งจำนวนลำดับตัวแปรของการเรียงสับเปลี่ยนมีค่าเท่ากับ 24 ลำดับ รูปที่ 2.11 แสดงตัวอย่างการเรียงสับเปลี่ยนของขั้นตอนวิธี AD4 สำหรับตัวแปร  $x_1$   $x_2$   $x_3$  และ  $x_4$

$$\begin{aligned} \text{การเรียงสับเปลี่ยนของตัวแปร } (x_1, x_2, x_3, x_4) = \{ & (x_1, x_2, x_3, x_4) (x_2, x_1, x_3, x_4) (x_2, x_1, x_4, x_3) \\ & (x_1, x_2, x_4, x_3) (x_1, x_4, x_2, x_3) (x_1, x_4, x_3, x_2) \\ & (x_4, x_1, x_3, x_2) (x_4, x_1, x_2, x_3) (x_4, x_2, x_1, x_3) \} \end{aligned}$$

$(X_2, X_4, X_1, X_3)$   $(X_2, X_4, X_3, X_1)$   $(X_4, X_2, X_3, X_1)$   
 $(X_4, X_3, X_2, X_1)$   $(X_4, X_3, X_1, X_2)$   $(X_3, X_4, X_1, X_2)$   
 $(X_3, X_4, X_2, X_1)$   $(X_3, X_2, X_4, X_1)$   $(X_2, X_3, X_4, X_1)$   
 $(X_2, X_3, X_1, X_4)$   $(X_3, X_2, X_1, X_4)$   $(X_3, X_1, X_2, X_4)$   
 $(X_3, X_1, X_4, X_2)$   $(X_1, X_3, X_4, X_2)$   $(X_1, X_3, X_2, X_4)$  }

$X_1, X_2, X_3, X_4, X_5, X_6, X_7$	เริ่มต้น
$X_2, X_1, X_3, X_4, X_5, X_6, X_7$	สลับที่ $(X_1, X_2)$
$X_2, X_1, X_4, X_3, X_5, X_6, X_7$	สลับที่ $(X_3, X_4)$
$X_1, X_2, X_4, X_3, X_5, X_6, X_7$	สลับที่ $(X_2, X_1)$
$X_1, X_4, X_2, X_3, X_5, X_6, X_7$	สลับที่ $(X_2, X_4)$
$X_1, X_4, X_3, X_2, X_5, X_6, X_7$	สลับที่ $(X_2, X_3)$
$X_4, X_1, X_3, X_2, X_5, X_6, X_7$	สลับที่ $(X_1, X_4)$
$X_4, X_1, X_2, X_3, X_5, X_6, X_7$	สลับที่ $(X_3, X_2)$
$X_4, X_2, X_1, X_3, X_5, X_6, X_7$	สลับที่ $(X_1, X_2)$
$X_2, X_4, X_1, X_3, X_5, X_6, X_7$	สลับที่ $(X_4, X_2)$
$X_2, X_4, X_3, X_1, X_5, X_6, X_7$	สลับที่ $(X_1, X_3)$
$X_4, X_2, X_3, X_1, X_5, X_6, X_7$	สลับที่ $(X_2, X_4)$
$X_4, X_3, X_2, X_1, X_5, X_6, X_7$	สลับที่ $(X_2, X_3)$
$X_4, X_3, X_1, X_2, X_5, X_6, X_7$	สลับที่ $(X_2, X_1)$
$X_3, X_4, X_1, X_2, X_5, X_6, X_7$	สลับที่ $(X_4, X_3)$
$X_3, X_4, X_2, X_1, X_5, X_6, X_7$	สลับที่ $(X_1, X_2)$
$X_3, X_2, X_4, X_1, X_5, X_6, X_7$	สลับที่ $(X_4, X_2)$
$X_2, X_3, X_4, X_1, X_5, X_6, X_7$	สลับที่ $(X_3, X_2)$
$X_2, X_3, X_1, X_4, X_5, X_6, X_7$	สลับที่ $(X_4, X_1)$
$X_3, X_2, X_1, X_4, X_5, X_6, X_7$	สลับที่ $(X_2, X_3)$
$X_3, X_1, X_2, X_4, X_5, X_6, X_7$	สลับที่ $(X_2, X_1)$
$X_3, X_1, X_4, X_2, X_5, X_6, X_7$	สลับที่ $(X_2, X_4)$
$X_1, X_3, X_4, X_2, X_5, X_6, X_7$	สลับที่ $(X_3, X_1)$
$X_1, X_3, X_2, X_4, X_5, X_6, X_7$	สลับที่ $(X_4, X_2)$

รูปที่ 2.11 การเรียงสับเปลี่ยนของขั้นตอนวิธี AD4 สำหรับตัวแปร  $X_1$   $X_2$   $X_3$  และ  $X_4$

2. ขั้นตอนวิธี SIFTING เป็นวิธีการหาตำแหน่งที่เหมาะสมสำหรับตัวแปรหนึ่งๆ ของแผนภาพตัดสัจใจทวิภาคที่มีตัวแปรทั้งหมด  $n$  ตัว โดยการเลื่อนตำแหน่งของตัวแปรนั้นไปยัง  $n$  ตำแหน่งที่เป็นไปได้ แล้วสร้างแผนภาพตัดสัจใจทวิภาคที่ได้จากการเลื่อนทั้งหมด  $n$  แผนภาพ จากนั้นจะเลือกแผนภาพตัดสัจใจทวิภาคที่เล็กที่สุดภายใน  $n$  แผนภาพนั้น กระบวนการนี้จะถูกทำซ้ำจนกระทั่งไม่สามารถหาแผนภาพตัดสัจใจทวิภาคที่เล็กกว่าได้ [7]

การเลื่อนตำแหน่งของตัวแปรตัวหนึ่งไปยัง  $n$  ตำแหน่ง เพื่อหาตำแหน่งที่ดีที่สุดของตัวแปรนั้น สามารถทำได้โดยแบ่งการเลื่อนตำแหน่งออกเป็น 2 ส่วนตามตำแหน่งเดิมของตัวแปรที่สนใจ ส่วนแรกเป็นการแลกเปลี่ยนตำแหน่งกับตัวแปรก่อนหน้ามันไปเรื่อยๆ จนกระทั่งตัวแปรนั้นกลายเป็นตัวแปรแรก นั่นก็คือตัวแปรนั้นได้เลื่อนขึ้น (sift up) เป็นบัพแรกหรือรากของแผนภาพตัดสัจใจ ส่วนที่สองเป็นการแลกเปลี่ยนตำแหน่งกับตัวแปรถัดจากมันไปเรื่อยๆ จนกระทั่งตัวแปรนั้นกลายเป็นตัวแปรตัวสุดท้าย นั่นก็คือตัวแปรนั้นได้เลื่อนลง (sift down) เป็นบัพล่างสุดของแผนภาพตัดสัจใจ ดังรูปที่ 2.12

$X_1, X_2, X_3, X_4, X_5, X_6, X_7$	เริ่มต้น
$X_1, X_2, X_4, X_3, X_5, X_6, X_7$	สลับที่ ( $X_3, X_4$ )
$X_1, X_4, X_2, X_3, X_5, X_6, X_7$	สลับที่ ( $X_2, X_4$ )
$X_4, X_1, X_2, X_3, X_5, X_6, X_7$	สลับที่ ( $X_1, X_4$ )

(1) การเลื่อนตัวแปร  $x_4$  ขึ้น

$X_1, X_2, X_3, X_4, X_5, X_6, X_7$	เริ่มต้น
$X_1, X_2, X_3, X_5, X_4, X_6, X_7$	สลับที่ ( $X_4, X_5$ )
$X_1, X_2, X_3, X_5, X_6, X_4, X_7$	สลับที่ ( $X_4, X_6$ )
$X_1, X_2, X_3, X_5, X_6, X_7, X_4$	สลับที่ ( $X_4, X_7$ )

(2) การเลื่อนตัวแปร  $x_4$  ลง

รูปที่ 2.12 การเลื่อนตำแหน่งตัวแปรของขั้นตอนวิธี SIFTING สำหรับตัวแปร  $x_4$

จะเห็นว่าขั้นตอนวิธี SIFTING สามารถเคลื่อนย้ายตัวแปรเป็นระยะทางไกลจากตำแหน่งเดิมได้ โดยไม่สนใจต่อขนาดของแผนภาพตัดสินใจที่เพิ่มขึ้นระหว่างการเลื่อนตัวแปรไปยังตำแหน่งที่ดีที่สุด เมื่อสมมติให้ตัวแปรอื่นอยู่ในตำแหน่งคงที่

## 2.5 ขั้นตอนวิธีทางพันธุกรรม(Genetic Algorithm : GA)

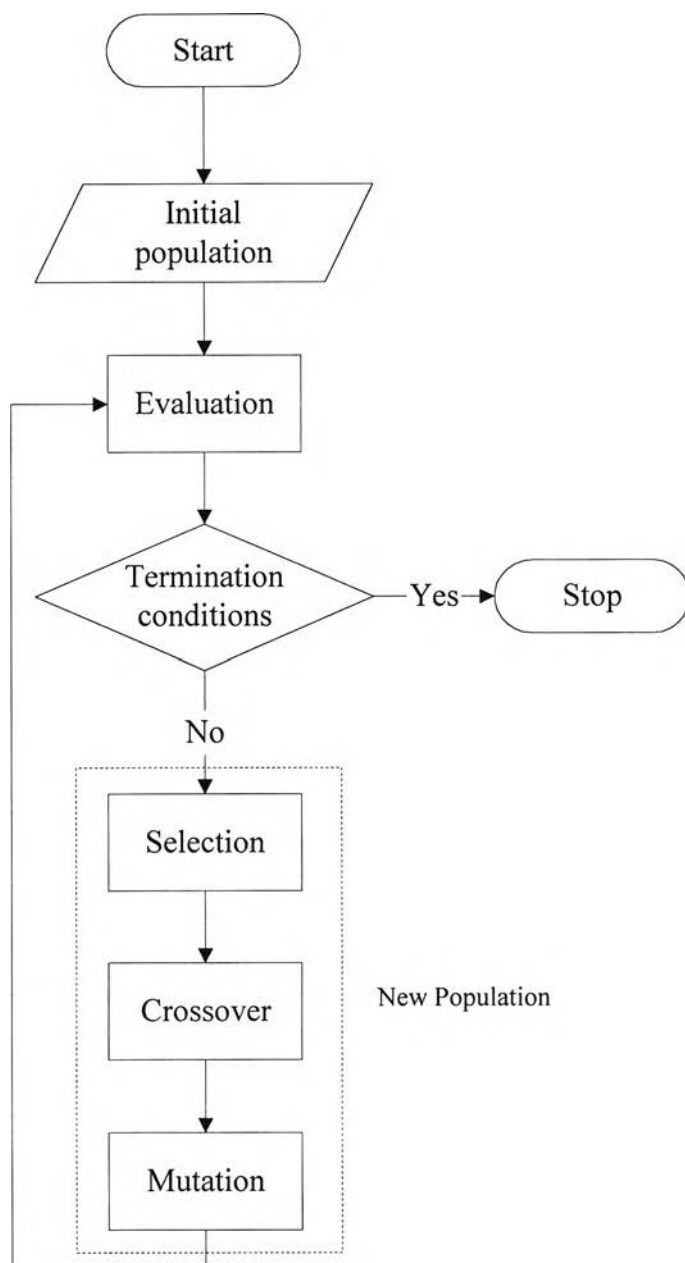
ขั้นตอนวิธีพันธุกรรม (Genetic Algorithm หรือ GA) เป็นวิธีการในการหาคำตอบซึ่งเลียนแบบกระบวนการวิวัฒนาการที่เกิดขึ้นในธรรมชาติ [9,15] วิธีนี้ได้รับการเสนอโดย Holland (1975) การค้นหาคำตอบของขั้นตอนวิธีพันธุกรรมอาศัยการเข้ารหัสผลเฉลยของปัญหาไปเป็นโครงสร้างข้อมูลที่มีลักษณะคล้ายโครโมโซม (chromosome) ซึ่งเป็นหน่วยพันธุกรรมของสิ่งมีชีวิต โครโมโซมมักอยู่ในรูปของสายอักขระฐานสอง (binary string) หรือสายอักขระ (string) การทำงานของขั้นตอนวิธีพันธุกรรมทำให้เกิดการแข่งขันกันระหว่างโครโมโซมที่อยู่ในกลุ่มประชากร (population) เดียวกัน โครโมโซมที่แก้ปัญหาได้ดี และใกล้ที่จะแก้ปัญหาสำเร็จจะอยู่รอด เพื่อเป็นต้นแบบในการสร้างโครโมโซมใหม่ ส่วนโครโมโซมที่แก้ปัญหาได้ไม่ดีก็จะถูกละทิ้งไป ประชากรกลุ่มนี้จะถูกพัฒนาเป็นรุ่นๆ จนกระทั่งพบโครโมโซมที่แก้ปัญหาได้สำเร็จ หรือพบเงื่อนไขที่ตั้งไว้

ความแตกต่างระหว่างขั้นตอนวิธีพันธุกรรม และวิธีอื่นๆ ทางปัญญาประดิษฐ์ อยู่ที่ขั้นตอนวิธีพันธุกรรมมีการทำงานแบบความน่าจะเป็น (Probability) และไม่ต้องมีการสร้างความรู้พื้นฐานของปัญหาที่สนใจ แต่จะทำงานโดยเข้ารหัสพารามิเตอร์ของปัญหาแทน

ดังนั้นขั้นตอนแรกในการใช้ขั้นตอนวิธีพันธุกรรมคือ จะเข้ารหัสปัญหาที่ศึกษาอย่างไร หรือจะสามารถแทนพารามิเตอร์ของปัญหาได้อย่างไร เพราะขั้นตอนวิธีพันธุกรรมกระทำกับประชากรของผลเฉลยที่เป็นไปได้ทั้งหมด ขั้นตอนถัดมาสามารถแบ่งออกได้เป็น การสร้างประชากรผลเฉลยเริ่มต้น การวัดค่าความเหมาะสมของผลเฉลย และการสร้างกลุ่มประชากรของผลเฉลยใหม่ ดังรูปที่ 2.13

### 2.5.1 การแทนโครโมโซม (Chromosome Representation)

ขั้นตอนวิธีพันธุกรรมโดยทั่วไปจะแทนโครโมโซมในรูปแบบสายอักขระ เพราะเหมาะสมต่อการทำงานของตัวดำเนินการพันธุกรรม วิธีที่ใช้แทนค่าจะมีผลกระทบอย่างมากต่อความถูกต้องและเวลาในการประมวลผลของขั้นตอนพันธุกรรม



รูปที่ 2.13 ผังงานของลักษณะการทำงานของขั้นตอนวิธีพันธุกรรม

วิธีที่ใช้แทนโครโมโซมโดยทั่วไปในปัญหาเชิงตัวเลขมี 2 วิธี คือการแทนโครโมโซมด้วยสายอักขระเลขฐานสอง และการแทนโครโมโซมด้วยสายอักขระเลขจำนวนเต็มหรือเลขจำนวนจริง ดังรูปที่ 2.14 โดยสายอักขระจะต้องสามารถใช้แทนค่าผลเฉลยที่เป็นไปได้ทั้งหมดในปัญหาที่สนใจ

1	0	0	1	0	0	1	1
---	---	---	---	---	---	---	---

สายอักขระเลขฐานสอง

2	5	0	1	6	4	3	7
---	---	---	---	---	---	---	---

สายอักขระเลขจำนวนเต็ม

รูปที่ 2.14 ตัวอย่างโครโมโซมที่เป็นสายอักขระ

### 2.5.2 การสร้างประชากรเริ่มต้น

เมื่อเริ่มต้นค้นหาคำตอบ ขั้นตอนวิธีพันธุกรรมจำเป็นต้องใช้ประชากรของผลเฉลยเริ่มต้นหรือกลุ่มของโครโมโซมเริ่มต้น การสร้างประชากรเริ่มต้นมีอยู่ด้วยกัน 2 วิธี วิธีแรกเป็นการสร้างผลเฉลยแบบสุ่มโดยใช้ตัวกำเนิดจำนวนแบบสุ่ม (random number generator) วิธีนี้เหมาะกับปัญหาที่ไม่มีความรู้พื้นฐานอยู่เลย

วิธีที่สองใช้ความรู้พื้นฐานของปัญหาที่เคยมีการศึกษาไว้ ในการสร้างประชากรผลเฉลยเริ่มต้นที่มีคุณภาพดีพอสมควร ทำให้ขั้นตอนวิธีพันธุกรรมหาผลเฉลยที่แก้ปัญหาสำเร็จในเวลาที่น้อยกว่าวิธีแรก

### 2.5.3 การวัดค่าความเหมาะสมของผลเฉลย

เมื่อได้ประชากรผลเฉลยเริ่มต้นแล้ว ขั้นตอนต่อไปจะเป็นการนำผลเฉลยแต่ละตัวมาประมวลผล และวัดค่าความเหมาะสม (Fitness Value) ในการแก้ปัญหา ขั้นตอนนี้มีความสำคัญต่อขั้นตอนวิธีพันธุกรรม เนื่องจากค่าความเหมาะสมนี้เป็นสิ่งที่ผลักดันให้เกิดกระบวนการวิวัฒนาการเพื่อค้นหาคำตอบ

ในกระบวนการวิวัฒนาการทางธรรมชาติ สิ่งมีชีวิตที่มีความเหมาะสมต่อสภาพแวดล้อมจะดำรงอยู่ ขณะที่สิ่งมีชีวิตที่มีความเหมาะสมน้อยหรือปรับตัวเข้ากับสภาพแวดล้อมไม่ได้จะสูญหายไป ขั้นตอนวิธีพันธุกรรมเลียนแบบหลักการนี้ โดยใช้ฟังก์ชันความเหมาะสม (Objective function) เป็นเครื่องมือวัดความเหมาะสมของแต่ละผลเฉลย ค่าที่ได้จะเป็นเกณฑ์ในการตัดสินว่าผลเฉลยใดมีประสิทธิภาพในการแก้ปัญหาสมควรอยู่รอด และได้รับการพัฒนาเป็นรุ่นต่อไป

ฟังก์ชันความเหมาะสมเป็นตัวกลางที่ใช้ติดต่อระหว่างขั้นตอนวิธีพันธุกรรมกับปัญหาที่ศึกษา ค่าความเหมาะสมที่ได้จากฟังก์ชันนี้เป็นข้อมูลช่วยเหลือเพียงอย่างเดียวที่เข้าสู่ระบบ มีผลโดยตรงต่อกระบวนการวิวัฒนาการ ดังนั้นการกำหนดฟังก์ชันความเหมาะสมเป็นสิ่งสำคัญต้องพิจารณาให้ดี

#### 2.5.4 การสร้างประชากรใหม่

ในการกำเนิดกลุ่มประชากรรุ่นใหม่มีวิธีการดำเนินการทางพันธุกรรมโดยทั่วไปอยู่ 3 วิธี คือ การคัดเลือก (selection) การไขว้เปลี่ยน (crossover) และการกลายพันธุ์ (mutation)

##### ก. การคัดเลือก (selection)

ในการคัดเลือกผลเฉลยหรือโครโมโซมเพื่อนำไปผลิตกลุ่มผลเฉลยรุ่นถัดไป สามารถทำได้หลายวิธี โดยปกติจะใช้วิธีคัดเลือกตามสัดส่วนของค่าความเหมาะสม (roulette wheel selection) วิธีการนี้จะแทนค่าความเหมาะสมเป็นความกว้างของช่องในวงล้อเสี่ยงโชค เมื่อมีการหมุนวงล้อแบบสุ่มผลเฉลยที่มีความกว้างของช่องมากจะมีโอกาสสูงที่จะถูกเลือกสำหรับรุ่นถัดไป ดังนั้นผลเฉลยที่มีค่าความเหมาะสมสูงจะมีโอกาสถูกเลือกสูงกว่าผลเฉลยที่มีค่าความเหมาะสมต่ำ โดยสมการความน่าจะเป็นของการถูกเลือกของผลเฉลยแต่ละตัวสามารถแสดงได้ดังนี้

$$P_i = \frac{f_i}{\sum_{j=0}^n f_j}$$

$P_i$  : ความน่าจะเป็นที่ผลเฉลยแต่ละตัวจะถูกเลือก

$f_i$  : ค่าความเหมาะสมของผลเฉลยแต่ละตัว

$n$  : จำนวนผลเฉลย

##### ข. การไขว้เปลี่ยน (crossover)

ตัวดำเนินการนี้เป็นการสร้างโครโมโซมรุ่นใหม่ โดยการสุ่มเลือกโครโมโซมจำนวน 2 สาย จากกลุ่มประชากรที่ได้คัดเลือกแล้ว เพื่อนำมาเป็น พ่อ แม่ (parents) จากนั้นสุ่มเลือกตำแหน่งที่จะตัดโครโมโซม ส่วนที่ถูกตัดของโครโมโซมทั้งสองจะถูกสับเปลี่ยนกันเพื่อสร้างโครโมโซมใหม่สองสาย (รูปที่ 2.15 a - b) ในการไขว้เปลี่ยนกันนี้ได้ คาดหวังว่าอาจทำให้ส่วนที่ดีๆ ของผลเฉลยแต่ละตัวมาอยู่ด้วยกัน และส่วนที่ไม่ดีก็จะอยู่ด้วยกัน ซึ่งผลเฉลยที่มีแต่ส่วนที่ดีมาอยู่ด้วยกันก็จะมี



โอกาสถูกเลือกไปผลิตรุ่นถัดไป การไขว้เปลี่ยนสามารถทำได้หลายรูปแบบเช่น การไขว้เปลี่ยนแบบ  
หนึ่งจุด การไขว้แบบสองจุด เป็นต้น

สายอักขระเดิม 1	<b>1 0 0 0 1 0 0 1 1 1 1</b>
สายอักขระเดิม 2	<b>0 1 1 0 1 1 0 0 0 1 1</b>
สายอักขระใหม่ 1	<b>1 0 0 0 1 1 0 0 0 1 1</b>
สายอักขระใหม่ 2	<b>0 1 1 0 1 0 0 1 1 1 1</b>

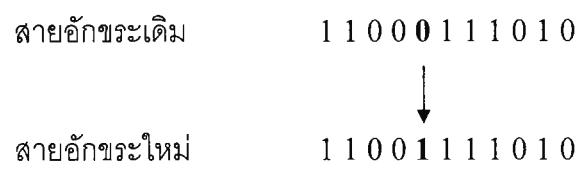
รูปที่ 2.15(a) การไขว้เปลี่ยนแบบหนึ่งจุด

สายอักขระเดิม 1	<b>1 0 1 0 0 0 1 1 0 1 0</b>
สายอักขระเดิม 2	<b>0 1 1 0 1 1 1 1 0 1 1</b>
สายอักขระใหม่ 1	<b>1 0 1 0 1 1 1 1 0 1 0</b>
สายอักขระใหม่ 2	<b>0 1 1 0 0 0 1 1 0 1 1</b>

รูปที่ 2.15(b) การไขว้เปลี่ยนแบบสองจุด

### ค. การกลายพันธุ์ (mutation)

การกลายพันธุ์เป็นการใส่ความเปลี่ยนแปลงโดยการสุ่มลงไปในโครโมโซม ถ้าโครโมโซม  
อยู่ในรูปของสายอักขระเลขฐานสองการกลายพันธุ์จะทำได้โดยการสุ่มบิตที่จะมีการเปลี่ยนแปลง  
แล้วทำการเปลี่ยนบิตจากค่า 0 เป็นค่า 1 หรือจากค่า 1 เป็นค่า 0 ลักษณะของการทำงานแสดงใน  
รูปที่ 2.16 ตัวดำเนินการกลายพันธุ์นี้ทำให้เกิดผลเฉลยในลักษณะใหม่ๆ ขึ้นมาบ้าง และช่วยเพิ่ม  
ความหลากหลายของประชากร โดยทั่วไปจะกำหนดค่าความน่าจะเป็นการกลายพันธุ์น้อย เพราะ  
ค่าความน่าจะเป็นสูงจะเป็นผลให้สูญเสียส่วนที่ดีของผลเฉลยไป



รูปที่ 2.16 การกลายพันธุ์

หลังจากที่สร้างประชากรด้วยตัวดำเนินการทั้ง 3 แบบแล้ว ประชากรรุ่นใหม่ที่ได้จะแทนที่ประชากรเดิม การทำงานของขั้นตอนวิธีพันธุกรรมจะกลับไปสู่การวัดค่าความเหมาะสม การสร้างประชากรใหม่ด้วยตัวดำเนินการ ในแต่ละรอบของการทำซ้ำเรียกว่า รุ่น (generation) กระบวนการหาคำตอบของขั้นตอนพันธุกรรมจะดำเนินต่อไปจนกระทั่งเงื่อนไขในการหยุดทำงานเป็นจริง ซึ่งได้แก่ การพบคำตอบที่ต้องการ หรือจำนวนรุ่นเกินกว่าที่กำหนดเพื่อไม่ให้เกิดการหาคำตอบใช้เวลานานเกินไป