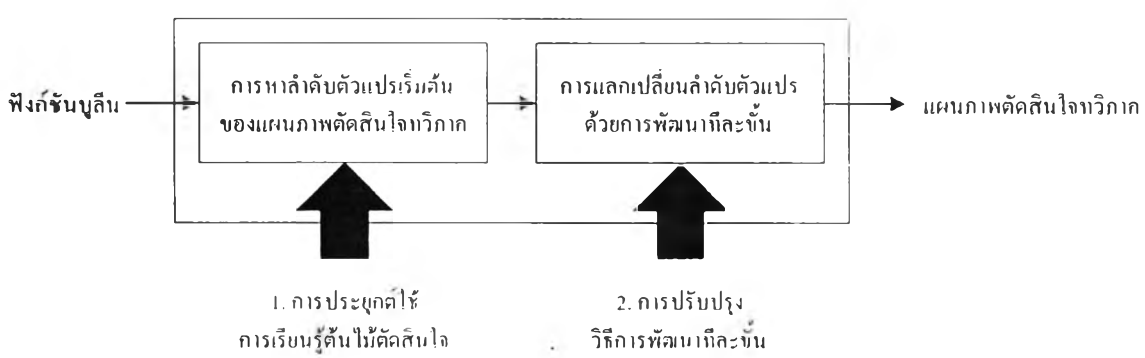
บทที่ 3

การสร้างแผนภาพตัดสินใจทวิภาคโดยการเรียนรู้ต้นไม้ตัดสินใจ และการปรับปรุงวิธีการพัฒนาทีละขั้น

วิทยานิพนธ์นี้ได้นำเสนอวิธีการสร้างแผนภาพตัดสินใจทวิภาคขนาดเล็ก โดยการหาลำดับตัวแปรเริ่มต้นของแผนภาพตัดสินใจทวิภาค จากนั้นจึงใช้การแลกเปลี่ยนลำดับตัวแปรของการพัฒนาทีละขั้นเพื่อลดขนาดของแผนภาพตัดสินใจทวิภาคให้เล็กลง โดยได้นำเสนอแนวทางในการพัฒนา 2 ส่วน ดังรูปที่ 3.1 คือ

(1) การประยุกต์ใช้การเรียนรู้ต้นไม้ตัดสินใจในการหาลำดับตัวแปรเริ่มต้น เพื่อให้ได้แผนภาพตัดสินใจทวิภาคเริ่มต้นที่มีขนาดเล็ก

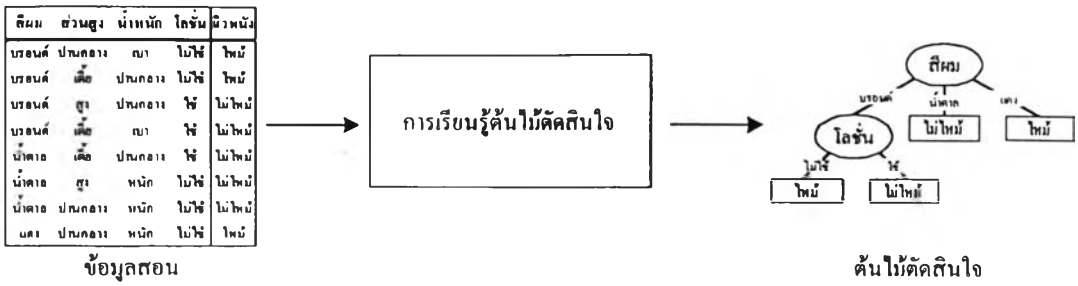
(2) การปรับปรุงวิธีการพัฒนาทีละขั้น เพื่อเพิ่มความสามารถในการลดขนาดของแผนภาพตัดสินใจทวิภาค



รูปที่ 3.1 การสร้างแผนภาพตัดสินใจทวิภาค

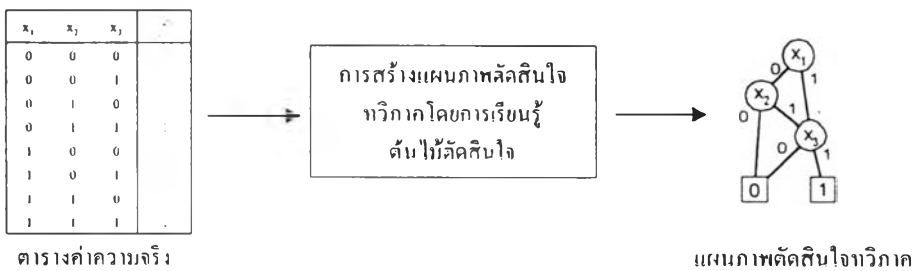
3.1 การประยุกต์ใช้การเรียนรู้ต้นไม้ตัดสินใจ

การเรียนรู้ต้นไม้ตัดสินใจเป็นวิธีการเรียนรู้จากตัวอย่างและแทนความรู้ที่ได้ในรูปของต้นไม้ตัดสินใจ ซึ่งมีลักษณะโครงสร้างดังรูปที่ 3.2 โดยมีข้อมูลเข้าเป็นข้อมูลสอน และข้อมูลออกเป็นต้นไม้ตัดสินใจ



รูปที่ 3.2 โครงสร้างของการเรียนรู้ต้นไม้ตัดสินใจ

จากโครงสร้างการเรียนรู้ต้นไม้ตัดสินใจในรูปที่ 3.2 เมื่อนำมาประยุกต์กับการสร้างแผนภาพตัดสินใจทวิภาคสามารถทำได้ดังรูปที่ 3.3 โดยการแทนข้อมูลสอนของการเรียนรู้ต้นไม้ตัดสินใจด้วยตารางค่าความจริง ซึ่งใช้แสดงทุกตัวอย่างที่เป็นไปได้ทั้งหมดของฟังก์ชันบูลีน และแทนต้นไม้ตัดสินใจที่ได้จากการเรียนรู้ด้วยแผนภาพตัดสินใจทวิภาค



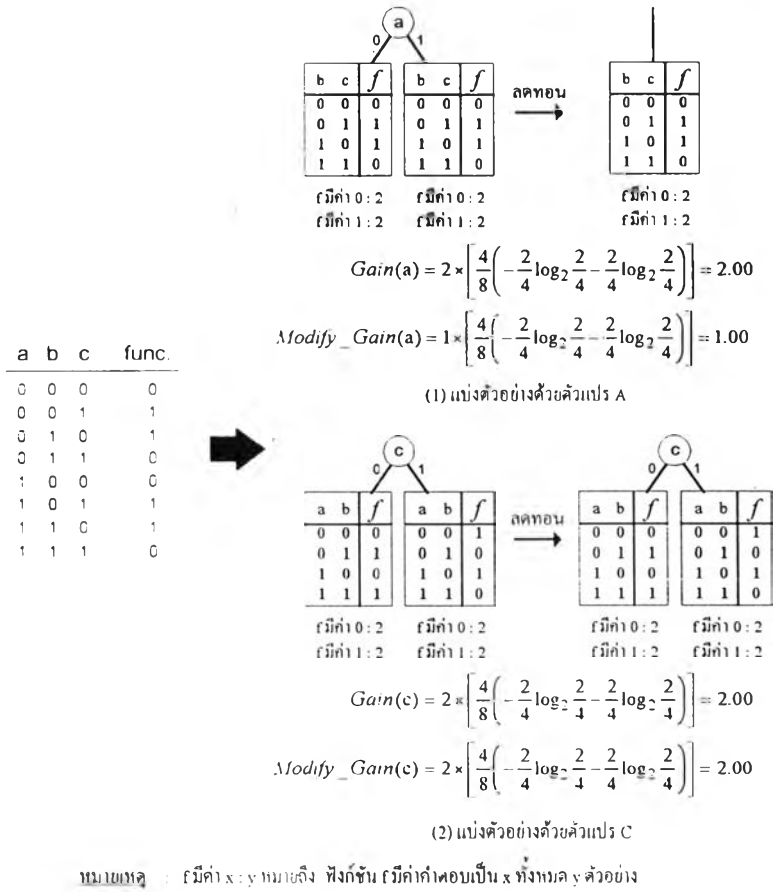
รูปที่ 3.3 โครงสร้างของการสร้างแผนภาพตัดสินใจทวิภาคโดยการเรียนรู้ต้นไม้ตัดสินใจ

เนื่องจากต้นไม้ตัดสินใจและแผนภาพตัดสินใจทวิภาคเป็นโครงสร้างข้อมูลแบบกราฟเหมือนกัน แต่มีลักษณะบางอย่างที่แลกต่างกันด้วย คือ

(1) ต้นไม้ตัดสินใจไม่มีลำดับของตัวแปร แต่แผนภาพตัดสินใจทวิภาคมี
(2) ต้นไม้ตัดสินใจไม่มีการลดทอนส่วนที่ซ้ำ แต่แผนภาพตัดสินใจทวิภาคมี

ด้วยลักษณะของการลดทอนส่วนที่ซ้ำที่แตกต่างกัน จึงส่งผลให้การคำนวณค่าความสามารถในการแยกตัวอย่างแตกต่างกันด้วย ดังรูปที่ 3.4 นั่นคือ ต้นไม้ตัดสินใจจะคำนวณจากทุกๆกลุ่มย่อย แต่แผนภาพตัดสินใจทวิภาคจะคำนวณจากกลุ่มย่อยที่ไม่ซ้ำกัน (กลุ่มย่อยที่ถูกลดทอนแล้ว) เท่านั้น และเรียกค่าที่คำนวณได้ว่า *Modify_Gain*

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 3.4 การคำนวณค่าความสามารถในการแยกตัวอย่างของ *Gain* และ *Modify_Gain*

การคำนวณค่า *Gain* และ *Modify_Gain* สำหรับแผนภาพตัดสินใจทวิภาค สามารถทำได้เช่นเดียวกับต้นไม้ตัดสินใจ ดังสมการที่ 2.2 โดย n_t เป็นจำนวนแถวทั้งหมดของตารางค่าความจริงก่อนการแยกตัวอย่าง สำหรับทุกฟังก์ชันรวมกัน (จำนวนแถวคูณด้วยจำนวนฟังก์ชัน) n_b เป็นจำนวนแถวของตารางค่าความจริง b ที่ได้ภายหลังจากการแยกตัวอย่างแล้ว และ n_{bc} เป็นจำนวนแถวที่มีค่าคำตอบของฟังก์ชันเป็น c ในตารางค่าความจริง b

ขั้นตอนของการสร้างแผนภาพตัดสินใจทวิภาคโดยการเรียนรู้ต้นไม้ตัดสินใจ สามารถทำได้ดังนี้

<u>ขั้นตอนที่ 1</u> ทดลองแบ่งตารางค่าความจริงและคำนวณค่า *Modify_Gain* สำหรับทุกตัวแปร

โดยแต่ละตัวแปรจะถูกใช้แบ่งทุกฟังก์ชันของตารางค่าความจริง ออกเป็นตารางค่าความจริงย่อยตามค่าของตัวแปร จากนั้นจะลดทอนส่วนที่ซ้ำด้วยการกำจัดตารางค่าความจริงย่อยที่ซ้ำกันและบัพที่ซ้ำซ้อนแล้วนำตารางค่าความจริงย่อยที่ได้จากการลดทอนไปคำนวณค่า *Modify_Gain* สำหรับตัวแปรนั้น

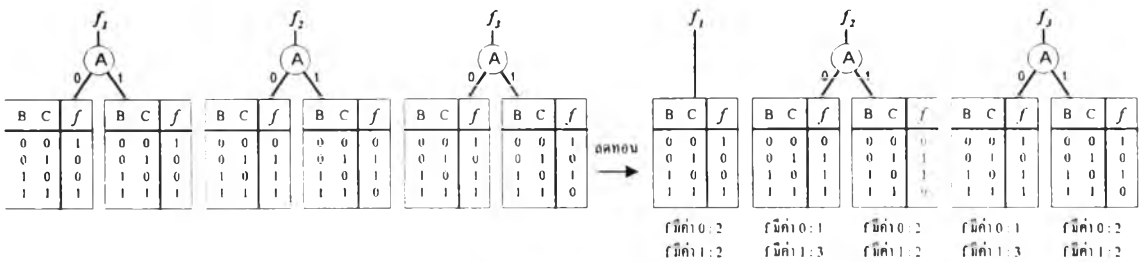
<u>ขั้นตอนที่ 2</u> เลือกตัวแปรที่มีค่า *Modify_Gain* น้อยที่สุดเป็นลำดับถัดไปของแผนภาพตัดสินใจทวิภาค

โดยทั้ง 2 ขั้นตอนการสร้างข้างต้นจะถูกทำซ้ำ จนกระทั่งในแต่ละตารางค่าความจริงย่อยมีค่าคำตอบของฟังก์ชันเหมือนกันทั้งหมด

ตารางที่ 3.1 ตารางค่าความจริงของฟังก์ชัน $f_1 = (\overline{B \oplus C})$ $f_2 = (B \oplus C) + \overline{A}B$ และ $f_3 = \overline{C} + \overline{A}B$

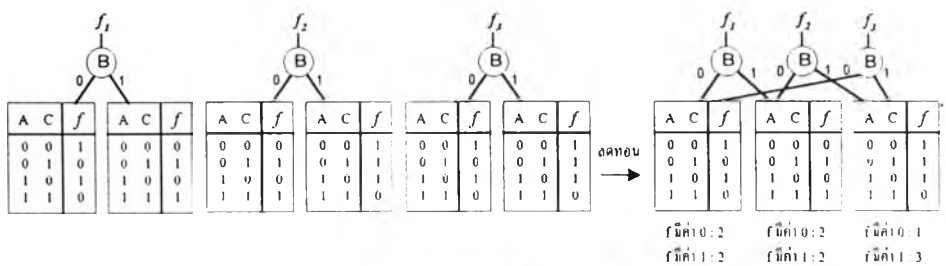
A	B	C	f_1	f_2	f_3
0	0	0	1	0	1
0	0	1	0	1	0
0	1	0	0	1	1
0	1	1	1	1	1
1	0	0	1	0	1
1	0	1	0	1	0
1	1	0	0	1	1
1	1	1	1	0	0

จากตัวอย่างของตารางค่าความจริงในตารางที่ 3.1 สามารถนำมาสร้างแผนภาพตัดสินใจทวิภาคโดยการเรียนรู้ต้นไม้ตัดสินใจได้ โดยการทดลองแบ่งตัวอย่างสำหรับแต่ละตัวแปร A B และ C จะได้ดังรูปที่ 3.5



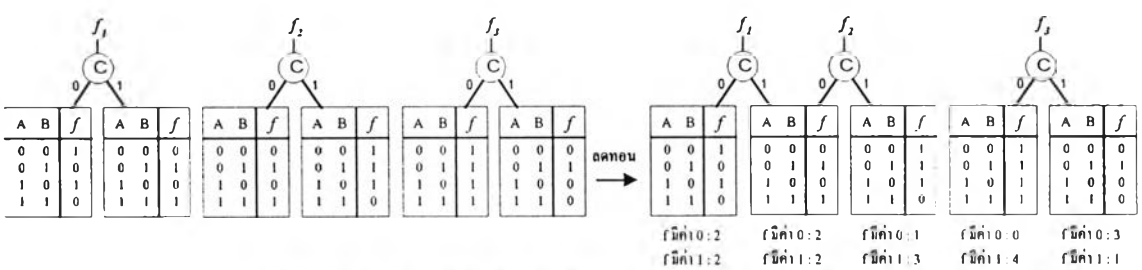
$$Modify_Gain(\mathrm{A}) = \left\{3 \times \frac{4}{24} \times \left(-\frac{2}{4}\log_2\frac{2}{4} - \frac{2}{4}\log_2\frac{2}{4}\right)\right\} + \left\{2 \times \frac{4}{24} \times \left(-\frac{1}{4}\log_2\frac{1}{4} - \frac{3}{4}\log_2\frac{3}{4}\right)\right\} = 0.7705$$

(1) แบ่งตัวอย่างด้วยตัวแปร A



$$Modify_Gain(\mathrm{B}) = \left\{2 \times \frac{4}{24} \times \left(-\frac{2}{4}\log_2\frac{2}{4} - \frac{2}{4}\log_2\frac{2}{4}\right)\right\} + \left\{1 \times \frac{4}{24} \times \left(-\frac{1}{4}\log_2\frac{1}{4} - \frac{3}{4}\log_2\frac{3}{4}\right)\right\} = 0.4686$$

(2) แบ่งตัวอย่างด้วยตัวแปร B



$$Modify_Gain(\mathrm{C}) = \left\{2 \times \frac{4}{24} \times \left(-\frac{2}{4}\log_2\frac{2}{4} - \frac{2}{4}\log_2\frac{2}{4}\right)\right\} + \left\{2 \times \frac{4}{24} \times \left(-\frac{1}{4}\log_2\frac{1}{4} - \frac{3}{4}\log_2\frac{3}{4}\right)\right\} + \left\{1 \times \frac{4}{24} \times \left(-\frac{0}{4}\log_2\frac{0}{4} - \frac{4}{4}\log_2\frac{4}{4}\right)\right\} = 0.6038$$

(3) แบ่งตัวอย่างด้วยตัวแปร C

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

รูปที่ 3.5 การสร้างแผนภาพตัดสินใจทวิภาคโดยการเรียนรู้ต้นไม้ตัดสินใจ (ระดับที่ 1)

เมื่อพิจารณาการแบ่งตัวอย่างด้วยตัวแปร A ในรูป 3.5(1) พบว่าฟังก์ชัน f_1 แยกตัวอย่างเป็นกลุ่มย่อยที่เหมือนกัน แต่ฟังก์ชัน f_2 และ f_3 แยกตัวอย่างเป็นกลุ่มย่อยที่ไม่เหมือนกัน ดังนั้นเมื่อลดทอนแผนภาพด้วยการกำจัดส่วนที่ซ้ำซ้อนของฟังก์ชัน f_1 ทำให้ฟังก์ชัน f_1 ถูกตัดบัพและกลุ่มย่อยที่ซ้ำกันทิ้ง จากนั้นก็จะนำกลุ่มย่อยที่เหลือจากการลดทอนมาคำนวณค่า *Modify_Gain* ดังสมการที่ 3.1

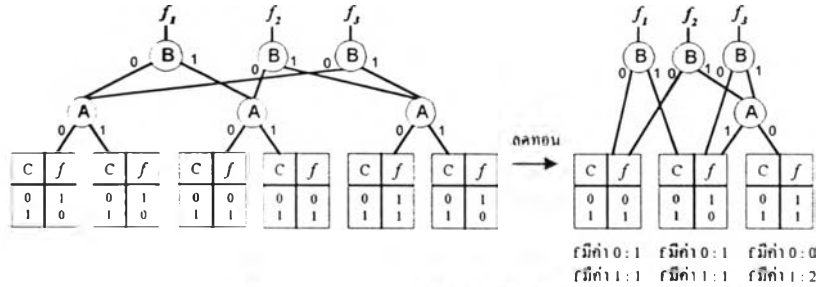
$$Modify_Gain(A) = \left\{\frac{4}{24}\times\left(-\frac{2}{4}\log_2\frac{2}{4}-\frac{2}{4}\log_2\frac{2}{4}\right)\right\}+\left\{\frac{4}{24}\times\left(-\frac{1}{4}\log_2\frac{1}{4}-\frac{3}{4}\log_2\frac{3}{4}\right)\right\}+\left\{\frac{4}{24}\times\left(-\frac{2}{4}\log_2\frac{2}{4}-\frac{2}{4}\log_2\frac{2}{4}\right)\right\}+$$
$$\left\{\frac{4}{24}\times\left(-\frac{1}{4}\log_2\frac{1}{4}-\frac{3}{4}\log_2\frac{3}{4}\right)\right\}+\left\{\frac{4}{24}\times\left(-\frac{2}{4}\log_2\frac{2}{4}-\frac{2}{4}\log_2\frac{2}{4}\right)\right\}=0.7705 \qquad (3.1)$$

เมื่อพิจารณาการแบ่งตัวอย่างด้วยตัวแปร B และ C ในรูปที่ 3.5(2) และ 3.5(3) พบว่ามีกลุ่มย่อยที่ซ้ำกันระหว่างสองฟังก์ชัน ดังนั้นเมื่อลดทอนแผนภาพด้วยการกำจัดกลุ่มย่อยที่ซ้ำกันแล้ว จากนั้นก็นำกลุ่มย่อยที่เหลือจากการลดทอนมาคำนวณค่า *Modify_Gain* ดังสมการที่ 3.2 และ 3.3 ตามลำดับ

$$Modify_Gain(B) = \left\{\frac{4}{24}\times\left(-\frac{2}{4}\log_2\frac{2}{4}-\frac{2}{4}\log_2\frac{2}{4}\right)\right\}+\left\{\frac{4}{24}\times\left(-\frac{2}{4}\log_2\frac{2}{4}-\frac{2}{4}\log_2\frac{2}{4}\right)\right\}+\left\{\frac{4}{24}\times\left(-\frac{1}{4}\log_2\frac{1}{4}-\frac{3}{4}\log_2\frac{3}{4}\right)\right\}+$$
$$= 0.4686 \qquad (3.2)$$

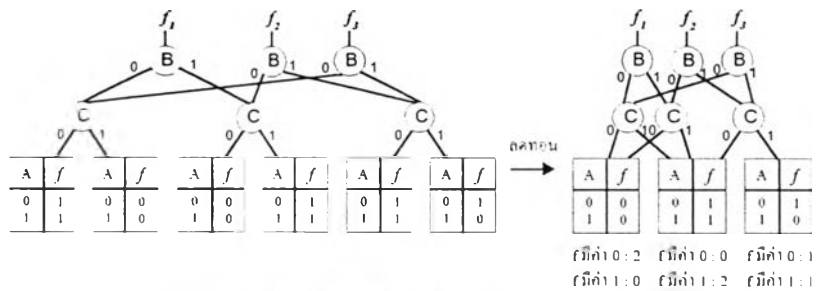
$$Modify_Gain(C) = \left\{\frac{4}{24}\times\left(-\frac{2}{4}\log_2\frac{2}{4}-\frac{2}{4}\log_2\frac{2}{4}\right)\right\}+\left\{\frac{4}{24}\times\left(-\frac{2}{4}\log_2\frac{2}{4}-\frac{2}{4}\log_2\frac{2}{4}\right)\right\}+\left\{\frac{4}{24}\times\left(-\frac{1}{4}\log_2\frac{1}{4}-\frac{3}{4}\log_2\frac{3}{4}\right)\right\}+$$
$$\left\{\frac{4}{24}\times\left(-\frac{0}{4}\log_2\frac{0}{4}-\frac{4}{4}\log_2\frac{4}{4}\right)\right\}+\left\{\frac{4}{24}\times\left(-\frac{3}{4}\log_2\frac{3}{4}-\frac{1}{4}\log_2\frac{1}{4}\right)\right\}=0.6038 \qquad (3.3)$$

จากค่า *Modify_Gain* ของตัวแปร A B และ C ที่คำนวณได้ จะเห็นว่าค่า *Modify_Gain* ของตัวแปร B มีค่าน้อยที่สุด ดังนั้นตัวแปร B จะถูกเลือกเป็นลำดับแรกของแผนภาพตัดสินใจทวิภาค และแผนภาพในรูปที่ 3.5(2) ก็จะถูกเลือกใช้ทดลองแบ่งตัวอย่างของกลุ่มย่อยต่อไปสำหรับการหาลำดับตัวแปรที่ 2 และ 3 ของแผนภาพตัดสินใจทวิภาค ดังรูปที่ 3.6 และ 3.7(4) ตามลำดับ โดยแผนภาพตัดสินใจทวิภาคที่ได้จากการเรียนรู้ต้นไม้ตัดสินใจมีลำดับตัวแปร คือ (B,C,A) และมีขนาดเท่ากับ 7 บัพ



$$Modify_Gain(A) = \left\{2 \times \frac{2}{12} \times \left(-\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2}\right)\right\} + \left\{1 \times \frac{2}{12} \times \left(-\frac{0}{2}\log_2\frac{0}{2} - \frac{2}{2}\log_2\frac{2}{2}\right)\right\} = 0.3333$$

(1) แบ่งตัวอย่างด้วยตัวแปร A

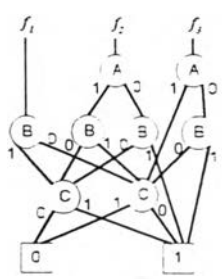


$$Modify_Gain(C) = \left\{2 \times \frac{2}{12} \times \left(-\frac{0}{2}\log_2\frac{0}{2} - \frac{2}{2}\log_2\frac{2}{2}\right)\right\} + \left\{1 \times \frac{2}{12} \times \left(-\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2}\right)\right\} = 0.1667$$

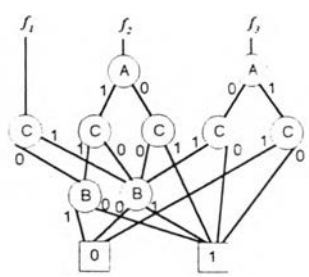
(2) แบ่งตัวอย่างด้วยตัวแปร C

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

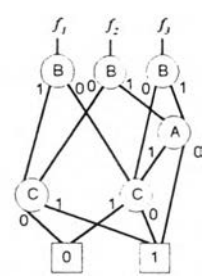
รูปที่ 3.6 การสร้างแผนภาพตัดสินใจทวิภาคโดยการเรียนรู้ต้นไม้ตัดสินใจ (ระดับที่ 2)



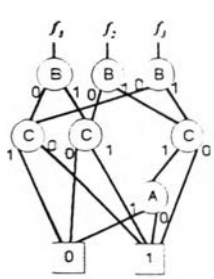
(1) ลำดับตัวแปร คือ (A, B, C)
และขนาด = 8 บัพ



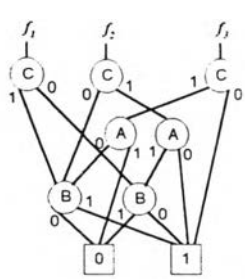
(2) ลำดับตัวแปร คือ (A, C, B)
และขนาด = 9 บัพ



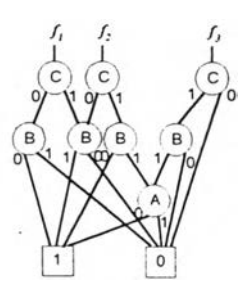
(3) ลำดับตัวแปร คือ (B, A, C)
และขนาด = 6 บัพ



(4) ลำดับตัวแปร คือ (B, C, A)
และขนาด = 7 บัพ



(5) ลำดับตัวแปร คือ (C, A, B)
และขนาด = 7 บัพ



(6) ลำดับตัวแปร คือ (C, B, A)
และขนาด = 8 บัพ

รูปที่ 3.7 แผนภาพตัดสินใจทวิภาคทั้งหมดของฟังก์ชัน $f_1 = \overline{(B \oplus C)}$ $f_2 = (B \oplus C) + \overline{A}B$ และ $f_3 = \overline{C} + \overline{A}B$

จากรูปที่ 3.7 แสดงถึงแผนภาพตัดสินใจทวิภาคของทุกลำดับตัวแปรสำหรับฟังก์ชันในตารางที่ 3.1 สามารถนำมาเรียงลำดับขนาดของแผนภาพตัดสินใจทวิภาคจากน้อยไปมาก คือ 6 7 7 8 8 และ 9 บัพ ซึ่งแต่ละขนาดจะสอดคล้องกับลำดับตัวแปร คือ (B,A,C) (B,C,A) (C,A,B) (A,B,C) (C,B,A) และ (A,C,B) ตามลำดับ โดยแผนภาพตัดสินใจทวิภาคที่ได้จากการเรียนรู้ต้นไม้ตัดสินใจเป็นแผนภาพที่มีลำดับตัวแปรดีเป็นลำดับที่ 2 นั่นคือ การใช้การเรียนรู้ตัดสินใจในการสร้างแผนภาพตัดสินใจทวิภาคจะนำไปสู่การหาลำดับตัวแปรที่ดี

เมื่อวิเคราะห์ความซับซ้อนของการประยุกต์ใช้การเรียนรู้ต้นไม้ตัดสินใจกับการสร้างแผนภาพตัดสินใจทวิภาค สำหรับตัวอย่างที่มีตัวแปรทั้งหมด n ตัว จะพบว่า

ระดับที่ 1 : จะต้องทดลองสร้างแผนภาพตัดสินใจทวิภาค n ต้น
ระดับที่ 2 : จะต้องทดลองสร้างแผนภาพตัดสินใจทวิภาค n –1 ต้น
...
ระดับที่ n-1 : จะต้องทดลองสร้างแผนภาพตัดสินใจทวิภาค 2 ต้น
ระดับที่ n : จะต้องทดลองสร้างแผนภาพตัดสินใจทวิภาค 1 ต้น

ดังนั้น แผนภาพตัดสินใจทวิภาคที่ทดลองสร้างทั้งหมด = $(n)+(n-1)+\ldots+2+1 = \frac{n(n+1)}{2} = O(n^2)$

3.2 การปรับปรุงวิธีการพัฒนาทีละขั้น

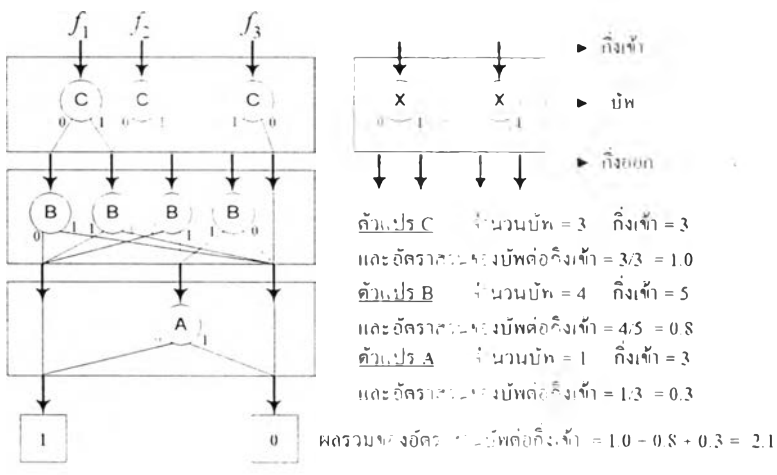
ขั้นตอนของวิธีการพัฒนาทีละขั้นจะประกอบไปด้วย (1) การเลือกตัวแปรที่ใช้ในการแลกเปลี่ยน (ยกเว้น AR และ ARSA ซึ่งเลือกตัวแปรแบบสุ่ม) และ (2) การเลือกแผนภาพตัดสินใจทวิภาคที่มีขนาดเล็กที่สุด

Ishiura [10] ได้เสนอขั้นตอนของวิธีการพัฒนาทีละขั้นแบบง่าย โดยการเลือกตัวแปรที่มีจำนวนบัพมากที่สุดก่อนในการแลกเปลี่ยน และในกรณีที่มีแผนภาพตัดสินใจทวิภาคที่มีขนาดเล็กที่สุดหลายแผนภาพ จะเลือกแผนภาพตัดสินใจที่พบในลำดับแรกก่อน

เมื่อพิจารณาขั้นตอนของวิธีการพัฒนาทีละขั้นดังกล่าว พบว่ามีจุดบกพร่องและควรปรับปรุงดังนี้

(1) การเลือกตัวแปรที่มีจำนวนบัพมากที่สุดก่อนในการแลกเปลี่ยน จะส่งผลให้ตัวแปรในลำดับแรกและสุดท้ายมักถูกเลือกหลังสุดเสมอ ดังนั้นเพื่อความเท่าเทียมกันสำหรับทุกตัวแปร จึงปรับปรุงวิธีการพัฒนาทีละขั้นด้วยการเลือกตัวแปรที่มีอัตราส่วนบัพต่อกิ่งเข้า (ดังแสดงในรูปที่ 3.8) ที่มากที่สุดก่อนในการแลกเปลี่ยน

(2) การเลือกแผนภาพตัดสินใจทวิภาคขนาดเล็กที่สุดที่พบในลำดับแรก จะส่งผลให้แผนภาพตัดสินใจทวิภาคที่มีขนาดเล็ก (ที่สุด) เท่ากัน แต่พบในลำดับถัดมาไม่ถูกเลือกเสมอ ดังนั้นเพื่อเพิ่มโอกาสของการลดขนาดสำหรับทุกๆแผนภาพตัดสินใจทวิภาคที่มีขนาดเล็ก (ที่สุด) เท่ากัน จึงปรับปรุงวิธีการพัฒนาทีละขั้นด้วยการเลือกแผนภาพตัดสินใจทวิภาคขนาดเล็กที่สุด ที่มีผลรวมของอัตราส่วนบัพต่อกิ่งเข้า (ดังแสดงในรูปที่ 3.8) ที่มากที่สุด



รูปที่ 3.8 การคำนวณอัตราส่วนบัพต่อกิ่งเข้าและผลรวม