

การวิเคราะห์หาลำดับความสัมพันธ์ของกลุ่มยีน และแสดงผลในรูปแบบของ Interval Graph โดยใช้ PC -Trees.

รวินทร์ วิรัชพินทุ
ดร. นพดล คีรีเพชร

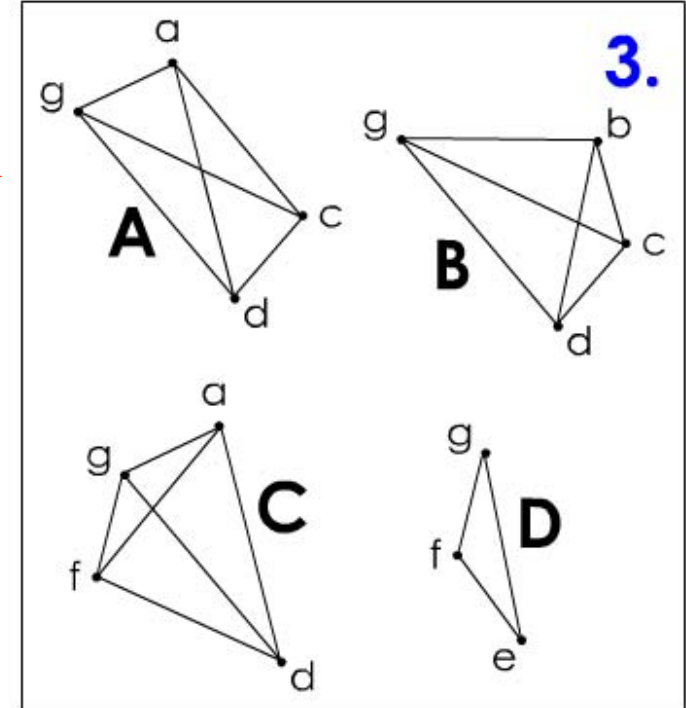
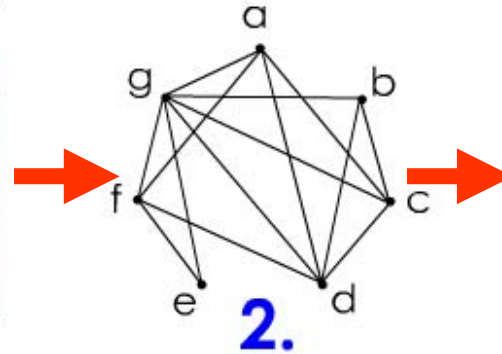
ฝ่ายวิจัยและพัฒนาเทคโนโลยีคอมพิวเตอร์เพื่อการคำนวณ
ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ



ข้อมูลในงานวิจัยด้านชีวสารสนเทศศาสตร์ส่วนใหญ่ มีปริมาณมากและซับซ้อน การนำเสนอข้อมูลแบบตารางสื่อความหมายได้น้อยและยากต่อการนำไปใช้วิเคราะห์

1.

	a	b	c	d	e	f	g
a	1	0	1	1	0	1	1
b	0	1	1	1	0	0	1
c	1	1	1	1	0	0	1
d	1	1	1	1	0	1	1
e	0	0	0	0	1	1	1
f	1	0	0	1	1	1	1
g	1	1	1	1	1	1	1



4.

	A	B	C	D
a	1	0	1	0
b	0	1	0	0
c	1	1	0	0
d	1	1	1	0
e	0	0	0	1
f	0	0	1	1
g	1	1	1	1

4.

	A	B	C	D
a	1	0	1	0
b	0	1	0	0
c	1	1	0	0
d	1	1	1	0
e	0	0	0	1
f	0	0	1	1
g	1	1	1	1



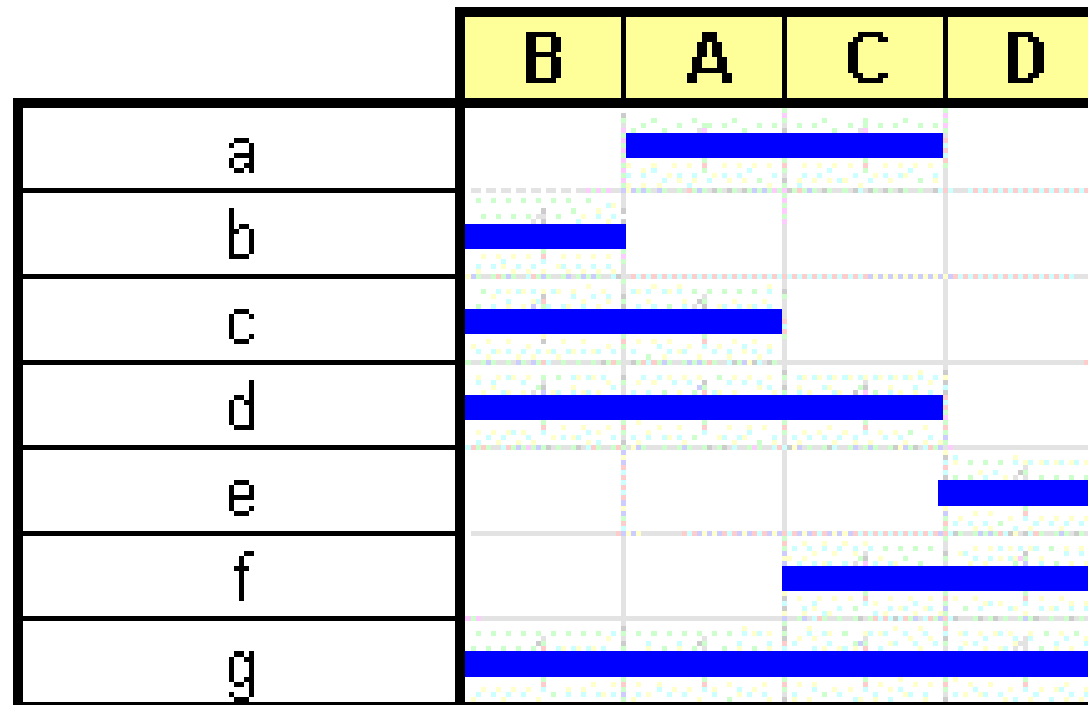
5.

	B	A	C	D
a		■		
b	■			
c		■		
d	■	■		
e				■
f			■	
g	■	■	■	

Interval graph

Interval graph

- กราฟแสดงความสัมพันธ์ของข้อมูลที่อยู่ในรูปของเส้นตรงที่ต่อเนื่องกัน



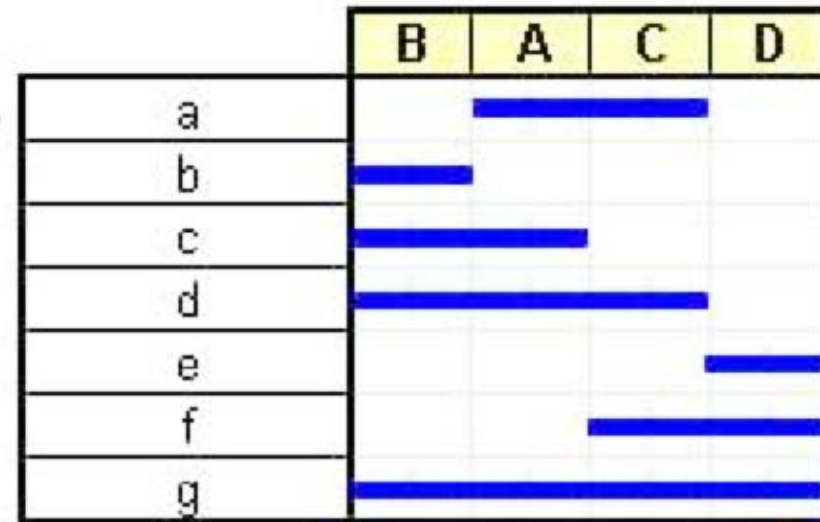
วิธีการสร้าง Interval graph

- สลับลำดับคอลัมน์ให้ข้อมูลที่มีค่า อยู่ติดกันทั้งหมด
- แทนข้อมูลที่มีค่าด้วยเส้นตรง

4.

	A	B	C	D
a	1	0	1	0
b	0	1	0	0
c	1	1	0	0
d	1	1	1	0
e	0	0	0	1
f	0	0	1	1
g	1	1	1	1

5.



Interval graph

คำจำกัดความ

Consecutive Ones Property (COP)

คุณสมบัติของข้อมูลเมทริกซ์ที่มีค่าหนึ่งในแต่ละแถวติดกันเป็นช่วงเดียวใน แนวเส้นตรง

Circular Ones Property (CROP)

คุณสมบัติของข้อมูลเมทริกซ์ที่มีค่าหนึ่งในแต่ละแถวติดกันเป็นช่วงเดียวใน แนววงกลม

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

NO COP, NO CROP

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

NO COP, CROP

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

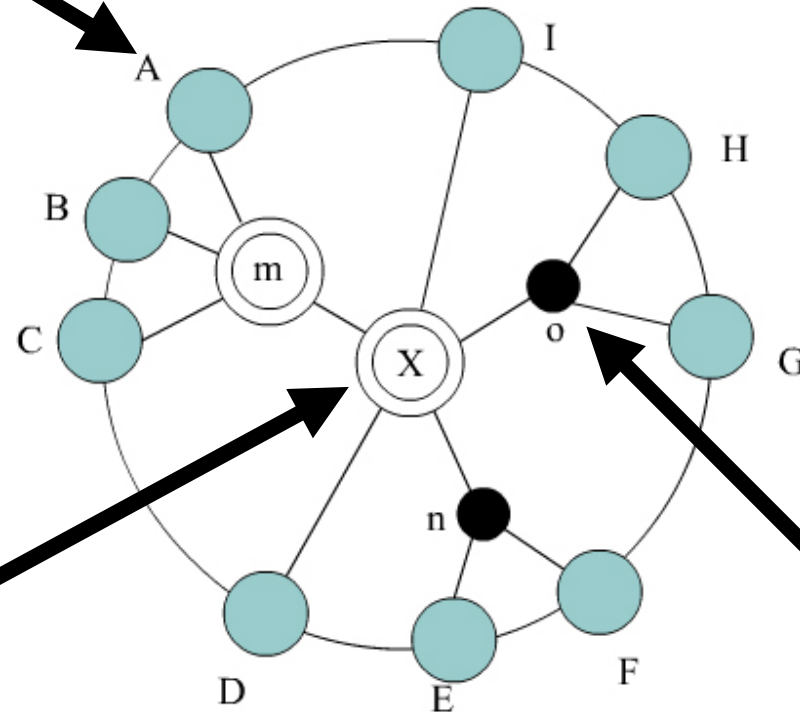
COP, CROP

PC-Tree

Leaf node

C node

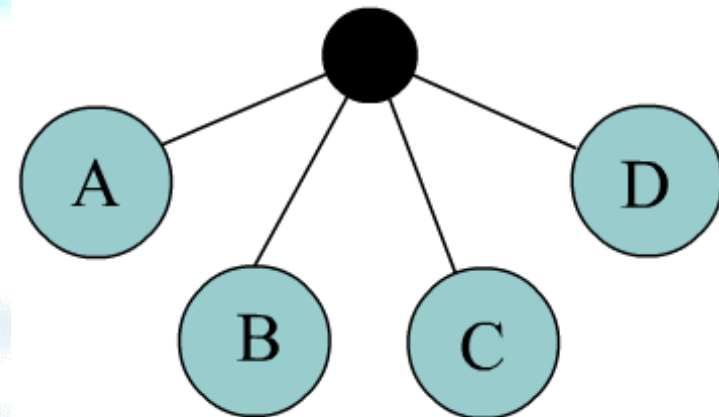
P node



ขั้นตอนการสร้าง PC-Tree

กำหนดโครงสร้างเริ่มต้น

	A	B	C	D
a	1	0	1	0
b	0	1	0	0
c	1	1	0	0
d	1	1	1	0
e	0	0	0	1
f	0	0	1	1
g	1	1	1	1



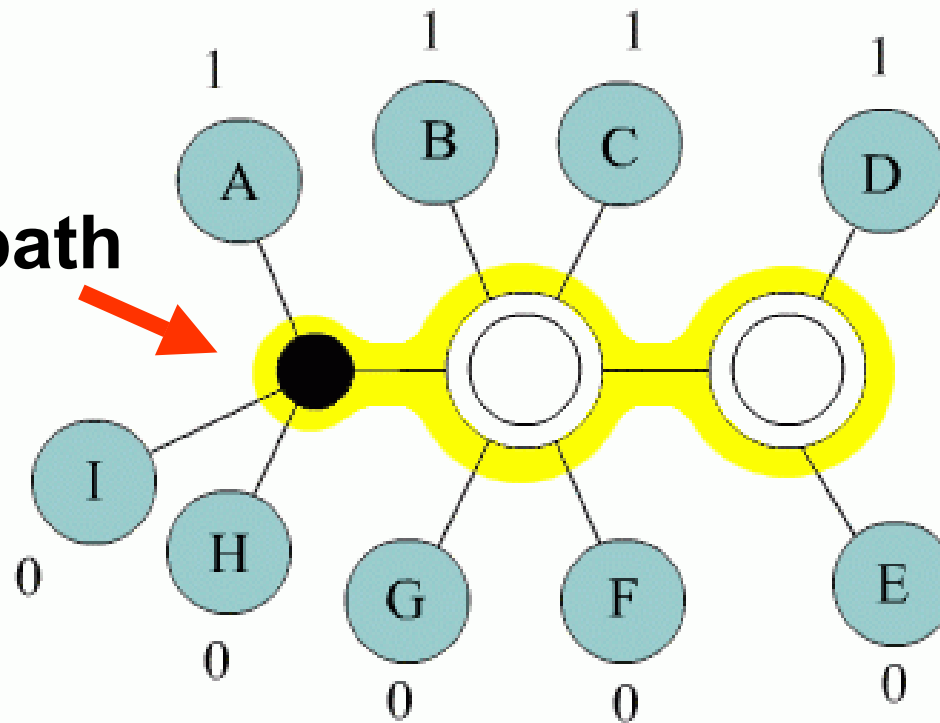
การทำงานของ PC-Tree

ขั้นตอนหลังการนำข้อมูลเข้าที่ละแถว

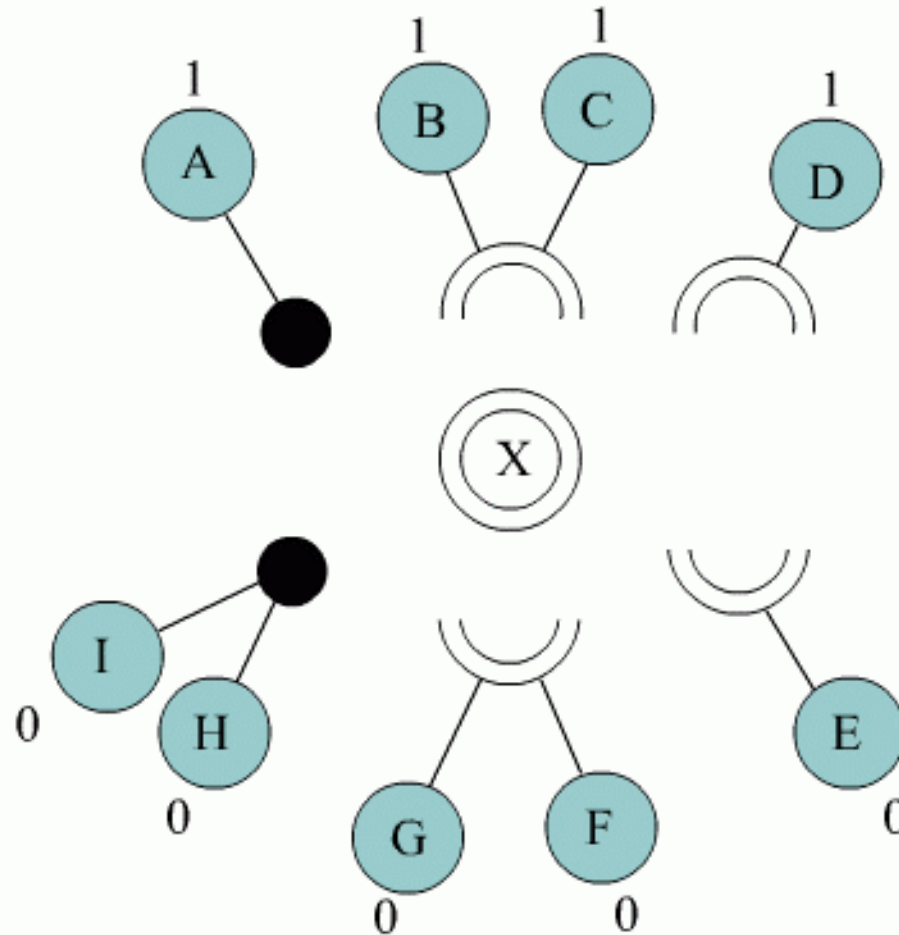
1. หา terminal path และสลับลำดับลูกของ P nodes เพื่อให้ค่า 1 อยู่ฝั่งเดียวของ path
2. แบ่งแต่ละ node บน path ออกเป็น 2 nodes, โดยแต่ละ node จะมี leaves ค่าเหมือนกันทั้งหมด
3. ลบ node บน path ทิ้ง, สร้าง C node ขึ้นมาใหม่ เพื่อเก็บ nodes ที่ได้จากการแบ่งตามลำดับบน path
4. เชื่อมต่อ C node ใหม่เข้ากับทุก nodes

1. หา terminal path และสลับลำดับลูกของ P nodes เพื่อให้ค่า 1 อยู่ฝั่งเดียวของ path

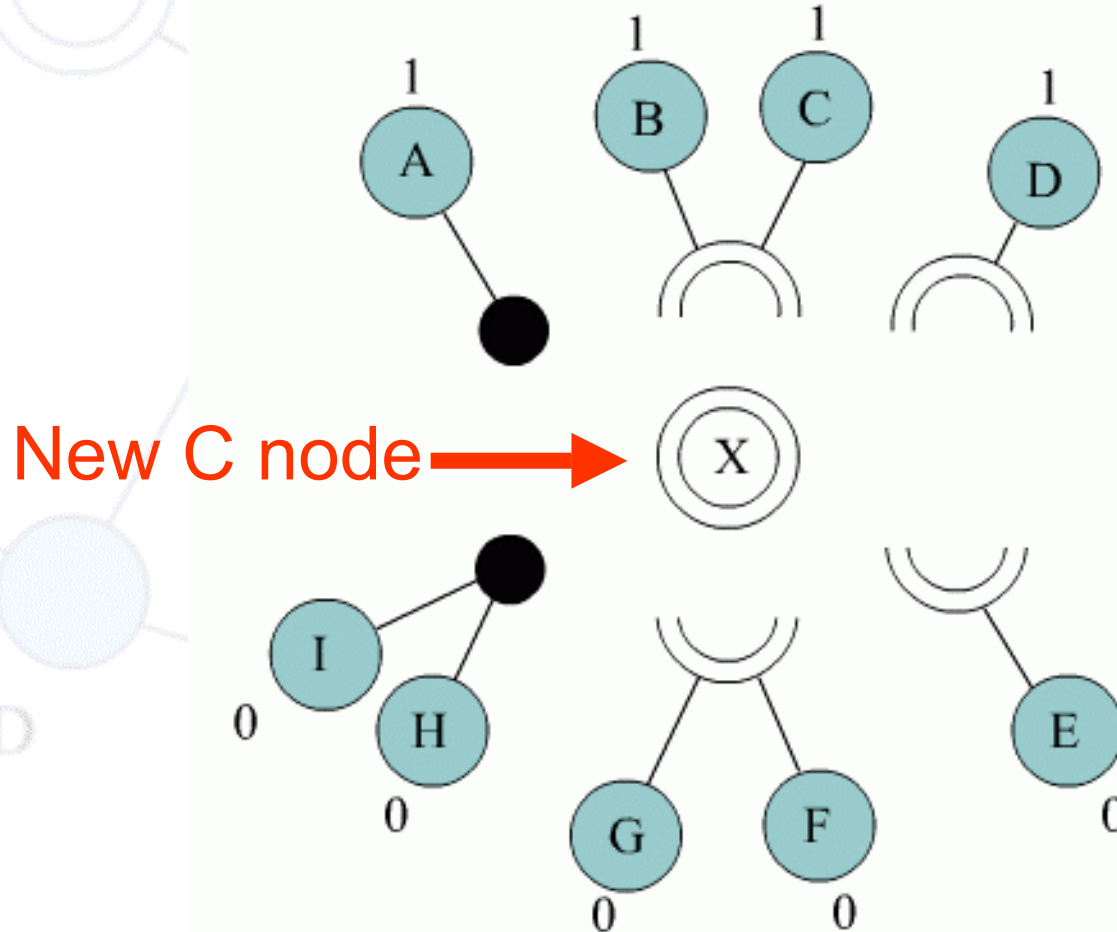
Terminal path



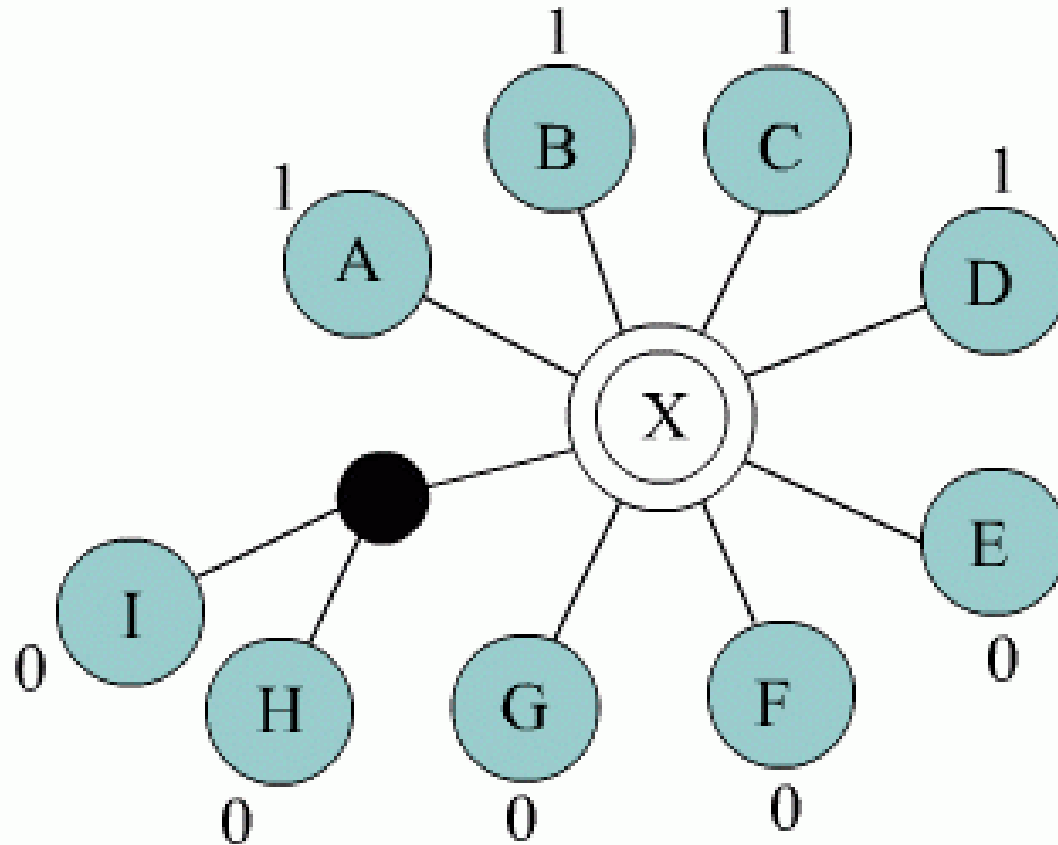
2. แบ่งแต่ละ node บน path ออกเป็น 2 nodes,
โดยแต่ละ node จะมี leaves ค่าเหมือนกันทั้งหมด



3. ลบ node บน path ที่,สร้าง C node ขึ้นมาใหม่
เพื่อเก็บ nodes ที่ได้จากการแบ่งตามลำดับบน path

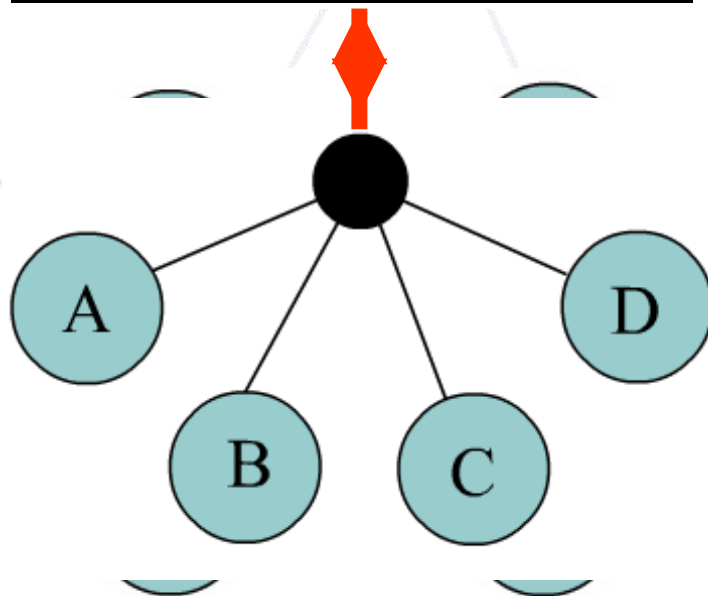


4. เชื่อมต่อ C node ใหม่เข้ากับทุก nodes



	A	B	C	D
a	1	0	1	0
b	0	1	0	0
c	1	1	0	0
d	1	1	1	0
e	0	0	0	1
f	0	0	1	1
g	1	1	1	1

	B	A	C	D
a	0	1	1	0
b	1	0	0	0
c	1	1	0	0
d	1	1	1	0
e	0	0	0	1
f	0	0	1	1
g	1	1	1	1



	B	A	C	D
a		█		
b	█			
c	█			
d	█			
e				█
f			█	
g	█			

แนวทางในการพัฒนาต่อ

- พัฒนาส่วนติดต่อผู้ใช้ (GUI)
 - ส่วนแสดงผล PC Tree ในรูปแบบของ Graphics
 - Reverse
- นำชุดข้อมูลจริงมาวิเคราะห์ทำ Interval graph

เอกสารอ้างอิง

- [1] John R. Jungck, Dept. of Biology, Beloit College, Beloit, WI and Vince Streif, **DELETION MAPPING OF GENETIC "FINE STRUCTURE": SUPPLEMENTING AD HOC PROBLEM SOLVING APPROACHES WITH ALGORITHMS AND HEURISTICS** ,Computer Center, University of Wisconsin-Eau Claire, Eau Claire, WI
- [2] Wen-Lian Hsu Ross M. McConnell , **"PC Trees and circular-ones Arrangements"**,Theoretical Computer Science 296(1), (2003), 99-116.
- [3] Wen-Lian Hsu , **"PC-Trees vs. PQ-Trees"** , Workshop on Graph Structures and Algorithms,also,appeared in Lecture Notes in Computer Science 2108, (2001),207-217.
- [4] MohammadTaghi Hajiaghayi , **"Consecutive Ones Property"**



Nikon D70 - Chiang Rai -
photo by tawan_ones_



เมื่อทำการเชื่อมโยงลูกที่เกิดจากการแบ่งทั้งหมดเข้ากับ C node ตัวใหม่เสร็จแล้ว อาจมีการยุบกิ่งที่ไม่มีสำคัญออก หรือลดรูปของต้นไม้ลง เพื่อให้อยู่ในโครงสร้างที่สั้นและกระชับ เช่น การยุบ C node ที่เป็นลูกของ C node เข้าด้วยกัน โดยยังคงลำดับของ leaves node ไว้ให้เหมือนเดิม เป็นต้น

