บทที่ 2

ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีต่างๆ ที่จำเป็นและเกี่ยวข้องกับงานวิจัยนี้ ซึ่งมีโครงสร้างของข้อมูลแบบต่างๆ ที่จำเป็นต้องใช้ในงานวิจัยนี้

โครงสร้างของข้อมูล

1. รายการ

รายการเป็นลักษณะการเก็บชุดของข้อมูลที่มีชนิดเดียวกันโดยมีขนาดจำกัด (Finite) และมีลำดับ (Sequence) รวมทั้งการกระทำ (Operation) ที่สามารถทำได้กับข้อมูลชุดนี้ดังต่อไปนี้

1) การตั้งค่าเริ่มต้นให้เป็นรายการว่าง

2) การตรวจสอบว่ารายการนี้เป็นรายการว่างหรือไม่

3) การตรวจสอบว่ารายการนี้เต็มหรือยัง

4) การหาความยาวของรายการ

5) การดึงเอาค่าของข้อมูลตัวใดๆ ในรายการมาได้เมื่อรายการนั้นไม่เป็นรายการว่าง

6) การเพิ่มข้อมูลตัวใหม่เข้าแทนที่ข้อมูลตัวเดิมในตำแหน่งใดๆ ภายในรายการเมื่อรายการนั้นไม่เป็นรายการว่าง

7) การเพิ่มข้อมูลชุดใหม่ลงไปในตำแหน่งใดๆ เมื่อรายการนั้นยังไม่เต็ม

8) การลบข้อมูลตัวใด ออกจากรายการได้เมื่อรายการนั้นไม่เป็นรายการว่าง

(Kruse, R. L., 1991)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

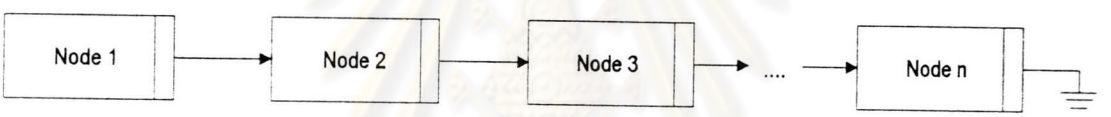
หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ข้อมูลแบบรายการสามารถแบ่งออกได้เป็น 2 แบบคือ

1.1 รายการแบบเรียงอันดับ (Ordered List) ซึ่งตำแหน่งและค่าข้อมูลของสมาชิกภายในรายการประเภทนี้มีความสำคัญ ดังนั้นการสลับที่หรือเปลี่ยนตำแหน่งของสมาชิกมีผลทำให้ความหมายของข้อมูลที่รายการนั้นแสดงอยู่เปลี่ยนไป

1.2 รายการแบบไม่เรียงอันดับ (Unordered List) เป็นการแทนค่าข้อมูลโดยการสลับหรือเปลี่ยนที่ตำแหน่งของสมาชิกในรายการไม่มีผลต่อการแปลความหมายของข้อมูลรายการนั้นๆ (Kruse, R. L., 1991)

รูปของโครงสร้างข้อมูลแบบรายการจะแทนด้วยส่วนประกอบสองส่วนคือ ส่วนที่เก็บค่าของข้อมูลและส่วนของตัวชี้ (Pointer) เพื่อบ่งบอกถึงตำแหน่งของสมาชิกตัวต่อไปในรายการ ดังรูปที่ 2.1



รูปที่ 2.1 โครงสร้างข้อมูลแบบรายการ

2. ต้นไม้

มีความหมายกว้างๆ ของข้อมูลแบบต้นไม้คือ เป็นเซตของจุด (Vertice) และเซตของเส้นเชื่อมคู่จุดที่อยู่ห่างจากกัน (กิ่ง - Edge หรือ Branch) ดังนั้น (1) จึงมีลำดับของกิ่ง (เส้นทาง - Path) จากจุดๆ หนึ่งไปยังจุดอื่นๆ และ (2) จะไม่มีการครบรอบเป็นวงจร (Circuit) นั่นคือไม่มีเส้นทางใดที่เริ่มต้นจากจุดๆ หนึ่งแล้วย้อนกลับมาที่จุดๆ นั้นอีก โดยบนคอมพิวเตอร์เรามักจะไม่ใช้ต้นไม้ในรูปแบบทั่วไปเช่นนี้ แต่เราจะเจาะจงกับรูปแบบของต้นไม้ที่มีจุดๆ หนึ่งซึ่งเรียกว่าจุดราก (Root) และเรียกต้นไม้แบบนี้ว่าต้นไม้มีราก (Rooted Tree) (Kruse, R. L., 1991)

นิยามของต้นไม้อีกนิยามหนึ่งคือต้นไม้แบบไดเรก (Directed Tree) ซึ่งมีคุณสมบัติดังนี้

1) มีจุดยอดอยู่หนึ่งจุด โดยเรียกว่าโหนดราก (Root Node)

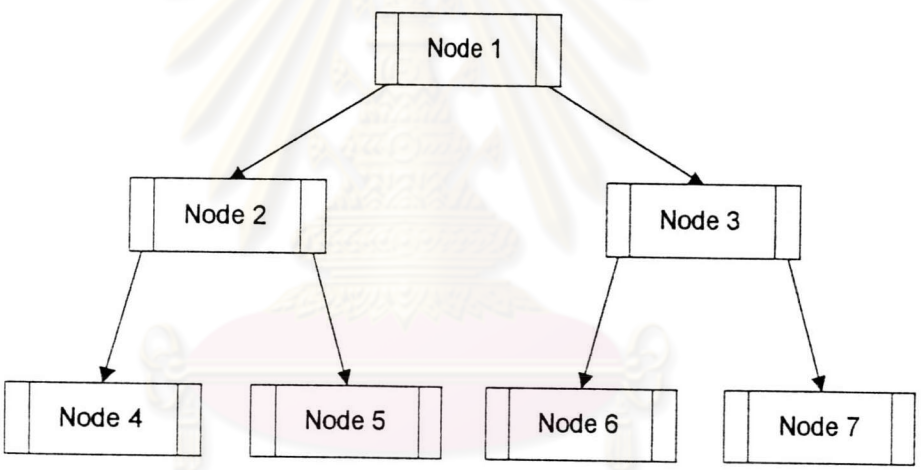
2) มีส่วนที่เรียกว่าต้นไม้ย่อย (Subtree)

3) มีส่วนที่เป็นใบ (Leaf Node หรือ Terminal Node) คือจะมีจำนวนต้นไม้ย่อยเป็นศูนย์หรือบางครั้งเรียกว่าดีกรีศูนย์ (Degree 0 หรือ Outdegree) ดังนั้นอาจกล่าวได้ว่าดีกรีคือจำนวนต้นไม้ย่อยนั่นเอง

4) มีจุดที่มีดีกรีไม่เท่ากับศูนย์เรียกว่าจุด (สมาชิก) ภายใน (Internal Node หรือ Branch Node)

(ประกาศิต ชาติบุรุษ, 2521)

ดังตัวอย่างในรูปที่ 2.2 แสดงให้เห็นถึงโครงสร้างของข้อมูลต้นไม้แบบทวิภาคที่มีความสูงเป็น 2



รูปที่ 2.2 โครงสร้างของต้นไม้แบบทวิภาคที่มีความสูงเป็น 2

รูปแบบของต้นไม้ยังสามารถแบ่งย่อยออกเป็นแบบต่างๆ ได้มากมายดังนี้

2.1 ต้นไม้แบบทวิภาค (Binary Tree) คือต้นไม้ว่างหรือต้นไม้ที่ประกอบด้วยจุดยอดจุดหนึ่งเรียกว่ารากา (Root) และต้นไม้แบบทวิภาคอีกสองต้นคือต้นไม้ย่อยทางซ้ายและต้นไม้ย่อยทางขวาของราก

2.2 ต้นไม้ค้นแบบทวิภาค (Binary Search Tree) คือต้นไม้ที่ว่างหรือ ต้นไม้ที่ประกอบด้วยสมาชิกที่มีคีย์ (Key) ซึ่งเป็นไปตามเงื่อนไขดังนี้

1) ค่าคีย์ทุกตัว (ถ้ามี) ในต้นไม้ย่อยทางซ้ายของรากต้องเป็นค่าที่น้อยกว่าค่าคีย์ที่ราก

2) ค่าคีย์ที่รากจะต้องเป็นค่าที่น้อยกว่าค่าคีย์ทุกตัว (ถ้ามี) ของต้นไม้ย่อยทางขวา

3) ต้นไม้ย่อยทั้งทางซ้ายและทางขวา ต้องเป็นต้นไม้ค้นแบบทวิภาค ด้วย

2.3 ต้นไม้แบบเอวีแอล (AVL Tree) คือต้นไม้ค้นแบบทวิภาคที่มีความสูงของต้นไม้ย่อยทางซ้ายและขวาต่างกันอย่างมาก 1 ระดับ และต้นไม้ย่อยทั้งทางซ้ายและขวาต่างก็เป็นต้นไม้แบบเอวีแอล

2.4 ต้นไม้ค้นแบบหลายทาง (Multiway Search Tree หรือ m-way tree) คือต้นไม้ซึ่งเมื่อกำหนด m คืออันดับ (Order) ของต้นไม้เท่ากับ m คือการกำหนดว่าสมาชิกแต่ละตัวในต้นไม้จะมีลูกได้ไม่เกิน m ตัว และถ้า k คือจำนวนลูกที่มี และ $k \leq m$ แล้วจะได้ว่าสมาชิกตัวนั้นจะเก็บค่า คีย์จำนวน $k - 1$ ตัว ซึ่งจะแบ่งคีย์ทั้งหมดออกเป็นเซตย่อย (Subset) จำนวน k เซตย่อย ถ้าเซตย่อยใดๆ เป็นเซตว่าง (Empty Set) ก็จะหมายความว่าไม่มีลูกในต้นไม้นั้น

2.5 ต้นไม้แบบบี (B-Tree หรือ Balanced Multiway Tree) คือต้นไม้ค้นแบบหลายทางแบบหนึ่งซึ่งมีคุณสมบัติเพิ่มเติมขึ้นดังนี้คือ

1) สมาชิกที่เป็นใบทุกตัวอยู่ในระดับเดียวกัน

2) ถ้ากำหนดให้อันดับของต้นไม้นี้เท่ากับ m แล้ว จะได้ว่าสมาชิกภายในทุกตัวยกเว้นรากมีลูกได้มากที่สุด m ตัว (ถ้าไม่ว่าง) และมีสมาชิกอย่างน้อย $\lceil m / 2 \rceil$ ตัว (ที่ไม่ว่าง

3) จำนวนของคีย์ภายในสมาชิกภายในแต่ละตัว จะมีน้อยกว่าจำนวนของลูกอยู่หนึ่งเสมอและคีย์เหล่านี้จะแบ่งคีย์ของลูกออกจากกันแบบเดียวกับต้นไม้ค้นแบบหลายทาง

4) รากจะมีลูกได้มากที่สุด m ตัว และมีลูกได้น้อยที่สุดเพียงสองตัวถ้าลูกไม่ใช่ใบหรือไม่มีลูกเลยถ้าต้นไม้มีเพียงจุดรากเพียงจุดเดียว