5.1. Cây và các khái niệm liên quan

5.1.1. Định nghĩa cây

* Định nghĩa 1: Cây là một đồ thị liên thông không có chu trình.
* Định nghĩa 2: Một cây là tập hợp hữu hạn các nút trong đó có một nút đặc biệt gọi là gốc (Root). Giữa các nút có mối quan hệ phân cấp gọi là quan hệ cha- con.

A diagram of a network

Description automatically generated

* Bậc của một nút: Là số nút con của nút đó.
* Bậc của một cây: Là bậc của nút có bậc lớn nhất trên cây đó. Cây có bậc n thì gọi là cây n – phân.

• Nút gốc: Là nút đặc biệt, không có nút cha –

• Nút lá: Là nút có bậc bằng 0 (không có nút con). –

• Nút nhánh: Là nút có bậc khác 0 và không phải là nút gốc.

• Mức của một nút .

* Gốc có mức 1. .
* Nếu nút cha có mức i thì các nút con có mức i+1.

• Chiều cao của cây: Là mức của nút có mức lớn nhất có trên cây.

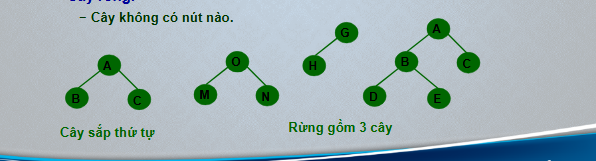
* Đường đi: Dãy các nút N1, N2, ... Nk, được gọi là đường đi nếu Ni là cha của Ni+1 (1 <=i<=k- 1). .
* Độ dài của đường đi: Là số nút trên đường đi trừ đi 1.

• Cây con: Là cây có gốc là một nút nhánh, lá.

* Cây được sắp thứ tự: Các nút được sắp theo một thứ tự nhất định.

• Rừng: Là tập hợp hữu hạn các cây phân biệt. –

• Cây rỗng: Cây không có nút nào.



* Ứng dụng: Cây thường được sử dụng trong nhiều ứng dụng như tổ chức dữ liệu (hệ thống tệp), biểu diễn cấu trúc phân cấp (như cây thư mục), và trong các thuật toán tìm kiếm.

5.2. Cây nhị phân

• Định nghĩa: Là cây mà mỗi nút không có quả 2 nút con, hai nút con (nếu có) được gọi là con trái và con phải.

• Cây con trái: Là cây có gốc là nút con trái.

Cây con phải: Là cây có gốc là nút con phải.

A diagram of a house

Description automatically generated

* Tính chất:
* Số nút tối đa ở mức i trên cây nhị phân là 2(i-1) (i ≥ 1).
* Số nút tối đa trên cây nhị phân chiều cao h là 2h – 1 (h >=1)
* Cấu trúc:
* Nút: Mỗi nút chứa một giá trị và hai tham chiếu đến nút con trái và nút con phải.
* Nút con trái (Left Child): Nút con bên trái của một nút.
* Nút con phải (Right Child): Nút con bên phải của một nút.
* Cây nhị phân đầy (Full Binary Tree): Tất cả các nút đều có 0 hoặc 2 nút con.
* Cây nhị phân hoàn chỉnh (Complete Binary Tree): Tất cả các cấp trừ cấp cuối cùng đều đầy và các nút ở cấp cuối cùng được đặt từ trái sang phải.
* Cây nhị phân cân bằng (Balanced Binary Tree): Chiều cao của hai cây con trái và phải không chênh lệch quá 1.
* Ứng dụng: Cây nhị phân thường được sử dụng để tổ chức dữ liệu, trong các thuật toán tìm kiếm và trong biểu diễn các biểu thức toán học.

5.3. Cây nhị phân tìm kiếm

2.1. Định nghĩa

Cây nhị phân tìm kiếm (CNPTK) là cây nhị phân hoặc rỗng hoặc không rỗng thì phải thoả mãn đồng thời các điều kiện sau:

- Khoá của các nút thuộc cây con trái nhỏ hơn khoá nút gốc

- Khoá của nút gốc nhỏ hơn khoá của các nút thuộc cây con phải của nút gốc

- Cây con trái và cây con phải của gốc cũng là cây nhị phân tìm kiếm

A diagram of numbers and circles

Description automatically generated

* Nút: Mỗi nút trong BST chứa một giá trị, một tham chiếu đến nút con trái và một tham chiếu đến nút con phải.
* Hoạt động:
* Tìm kiếm (Search): Để tìm một giá trị trong BST, bắt đầu từ nút gốc, so sánh giá trị cần tìm với giá trị của nút hiện tại và tiếp tục tìm kiếm trong cây con trái hoặc cây con phải tùy thuộc vào kết quả so sánh.
* Chèn (Insert): Tương tự như tìm kiếm, bắt đầu từ nút gốc và tìm vị trí thích hợp để chèn nút mới theo quy tắc BST.
* Xóa (Delete): Có ba trường hợp: xóa nút lá (không có con), xóa nút có một con, và xóa nút có hai con (thay thế bằng giá trị lớn nhất từ cây con trái hoặc nhỏ nhất từ cây con phải).
* Ưu điểm: Cây tìm kiếm nhị phân cho phép tìm kiếm, chèn và xóa với độ phức tạp trung bình là O(log n) nếu cây được cân bằng.
* Nhược điểm: Nếu không duy trì tính cân bằng, độ phức tạp có thể trở thành O(n) trong trường hợp xấu nhất (ví dụ, nếu cây trở thành một danh sách liên kết).