

BÁO CÁO THỰC NGHIỆM ĐO CHIỀU CAO CÂY NHỊ PHÂN

Họ và tên: Lê Bình Nguyên

MSSV: 22520969

Gmail: 22520969@gm.uit.edu.vn

Mở đầu

- AVL Tree và RBTTree là hai ứng dụng tiêu biểu của Binary Search Tree (cây tìm kiếm nhị phân).
- Cả AVL Tree và Red Black Tree đều là cây nhị phân tự cân bằng (thừa hưởng các tính chất của một BST và thêm những tính chất riêng).
- Các tính chất chung:
 - Mỗi Node (trừ Node là lá) luôn có tối đa 2 Child (Node con).
 - Node con bên trái luôn nhỏ hơn Node cha.
 - Node con bên phải luôn lớn hơn Node cha.
 - Chiều cao của cây xấp xỉ $\log(N)$ (với N là số Node của cây).
- Với cây AVL:
 - Cây con trái và phải đều là AVL.
 - Độ cao của 2 cây con lệch nhau không quá 1.
- Với cây RB:
 - Các cạnh được tô màu ĐỎ hoặc ĐEN.
 - Trên mỗi đường đi từ gốc đến lá không có 2 cạnh ĐỎ kề nhau.
 - Cạnh nối với nút ngoài (NULL) luôn có màu đen.
- Bài Thực nghiệm dưới đây sẽ tiến hành đo chiều cao của hai loại cây trong các trường hợp.

Tiến hành

Input

Các 10 bộ dữ liệu được sinh ngẫu nhiên gồm các số nguyên dương thuộc đoạn $[0; 10^9]$ trong đó:

- 8 test case đầu, các số được sắp xếp lộn xộn.
- 1 test case, các số được sắp xếp tăng dần.
- 1 test case, các số được sắp xếp giảm dần.

Output

Độ cao của hai loại cây với từng test case.

Dự đoán

Cây AVL: $h \leq 1.45 \log(N)$

Cây RB: $\log(N + 1) \leq h \leq 2 \log(N + 1)$

Quy trình

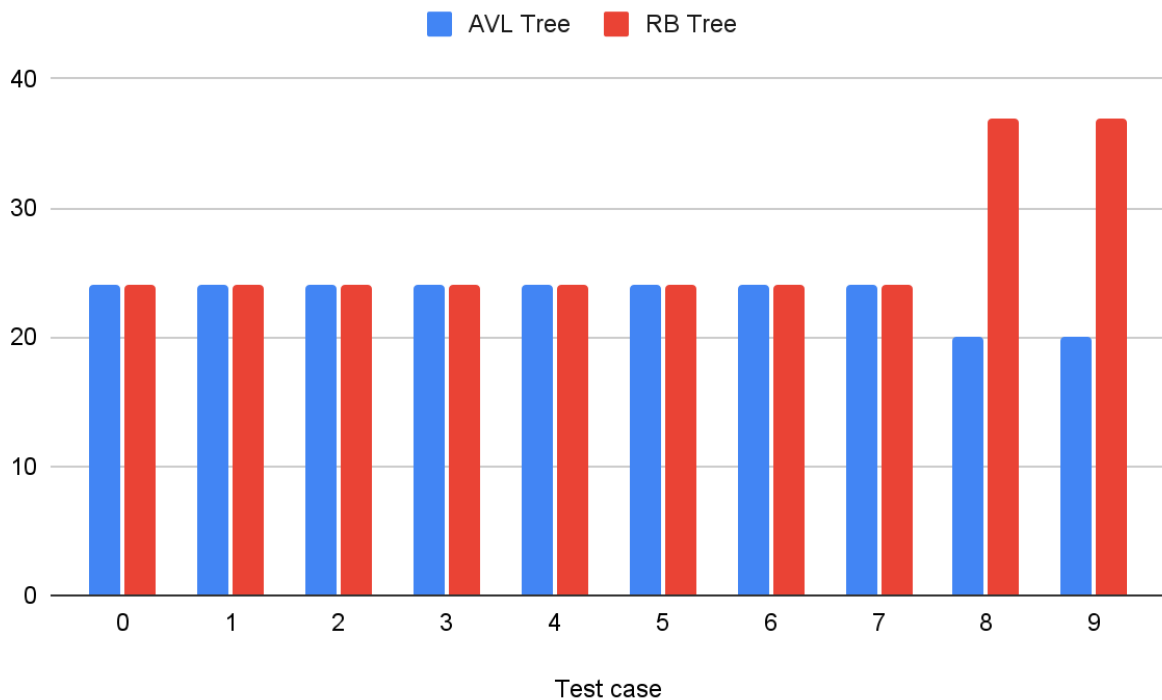
B1: Chạy file testGenerate.cpp để sinh các test case ngẫu nhiên.

B2: Chạy lần lượt các file AVLTree.cpp và RedBlackTree.cpp và thu được dữ liệu đầu ra từ các file AVLTree.out và RedBlackTree.out.

B3: Dùng các số liệu vừa thu được để vẽ biểu đồ trực quan.

B4: Rút ra nhận xét.

Biểu đồ



Nhận xét

Với $N = 10^6$, ta có

- $\log(N) = 19.93$
- $1.45\log(N) = 28.9$
- $2\log(N + 1) = 39.8$

Nhìn chung, các số liệu thu được từ việc thực nghiệm đều nằm trong khoảng dự đoán ban đầu.

Đối với những test case có các số được sắp xếp ngẫu nhiên:

- Các test case này đều có các điểm chung là:
 - Có cùng số node.
 - Các node hầu như không bị trùng lặp nhau.
 - Cây AVL và cây RB đều sẽ cố gắng để tối ưu hóa chiều cao của cây và thu được kết quả khá trung tính (không tốt nhất và cũng không tệ nhất).
- Điều này giải thích vì sao hầu hết kết quả thu được đều cùng một giá trị.

Đối với những test case có các số được sắp xếp theo thứ tự:

- Đây là trường hợp tối ưu đối với cây AVL và khá tệ đối với cây RB:
 - Cây AVL thu được chiều cao là 20 - xấp xỉ giá trị tốt nhất đối với một cây nhị phân cân bằng hoàn toàn nhờ cơ chế của nó.
 - Cây RB cần chiều cao lớn hơn - là 37 - gần bằng với giá trị cực đại có thể thu được để duy trì điều kiện ràng buộc về các nút đỏ và nút đen.
-