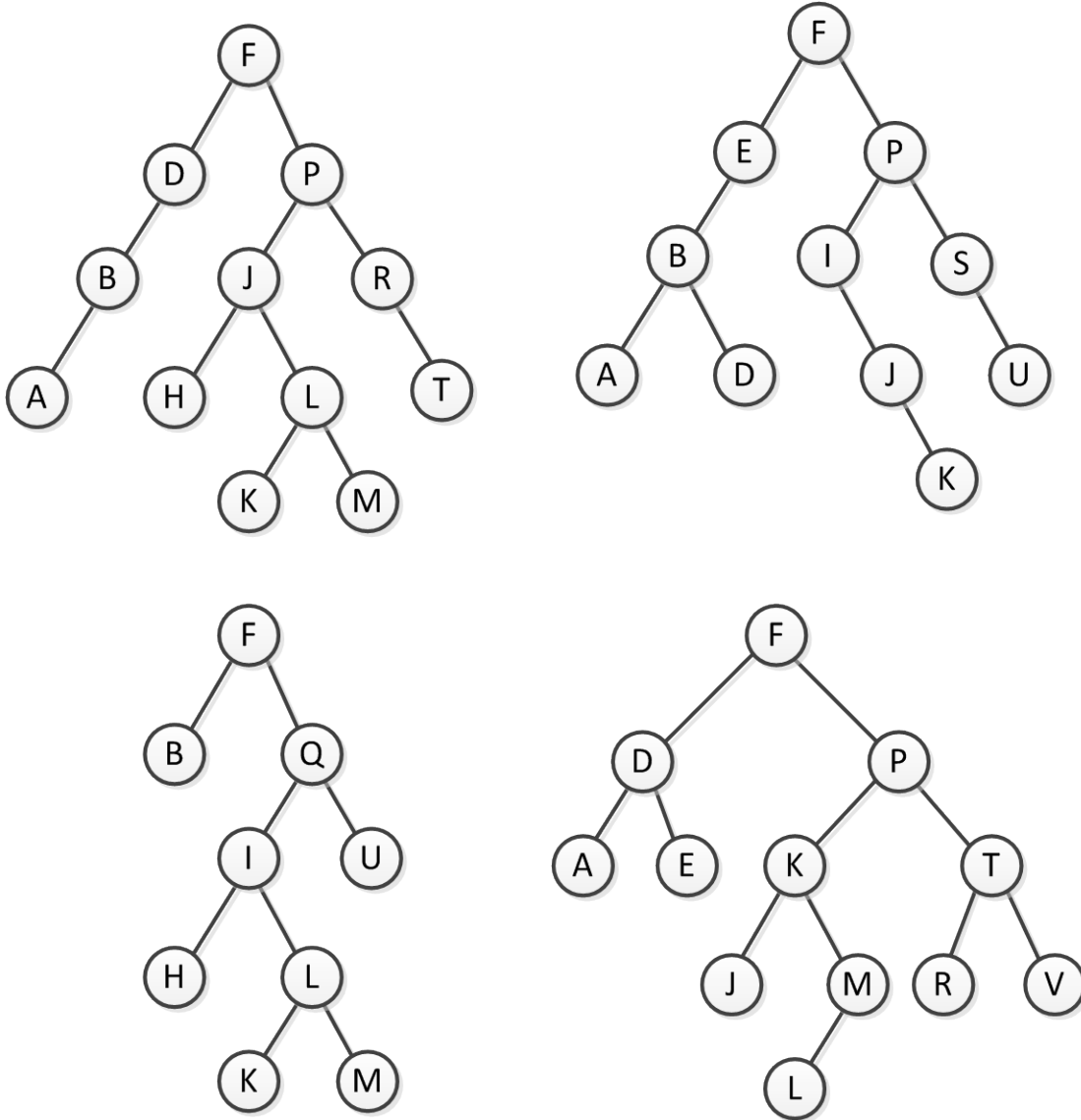


Bài tập phần AVL, Splay, 2-3, Red-Black Tree

Bài 1. Trong các cây nhị phân sau, cây nào là cây AVL, cây nào không phải. Với các cây không phải là AVL, hãy chỉ ra các đỉnh vi phạm.



Bài 2. Số nút tối thiểu trên cây nhị phân cân bằng AVL có chiều cao h là bao nhiêu?

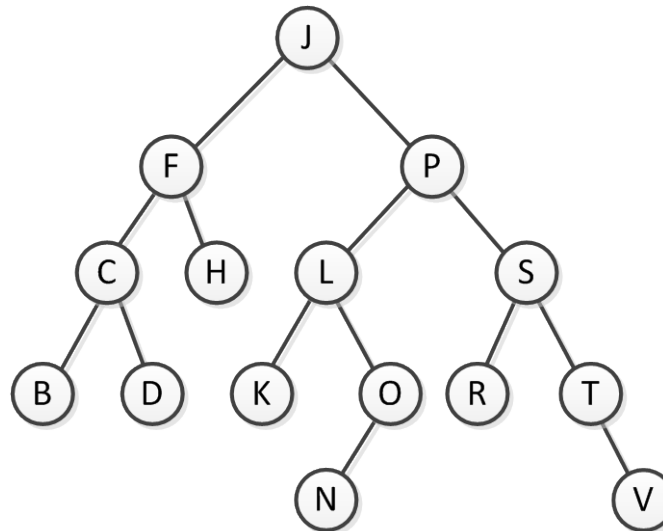
Bài 3. Vẽ cây AVL tạo thành bằng cách thêm lần lượt các khóa sau (vẽ cây sau mỗi lần 1 khóa được thêm vào)

- 3, 5, 7, 23, 12, 76, 87, 54, 19, 4
- D, H, A, G, T, C, D, I, Q, V, B, E
- 12, 34, 53, 76, 15, 21, 18, 45, 16, 55, 11

Bài 4. Với các cây AVL kết quả trong bài 3, hãy vẽ cây thu được sau khi ta xóa lần lượt các nút đã thêm vào theo

- Thứ tự các nút được thêm vào trước sẽ bị xóa trước (FIFO)
- Thứ tự các nút được thêm vào sau sẽ bị xóa trước (LIFO)

Bài 5. Vẽ cây thu được sau khi xóa các nút sau từ cây AVL ở dưới



- | | |
|------|------|
| a. J | d. L |
| b. F | e. S |
| c. P | f. R |

Bài 6. Xây dựng hàm để tìm và trả về chiều cao trên cây AVL, bằng cách đi từ gốc đến 1 nút lá thay vì đến tất cả nút lá

Gợi ý: dựa vào thông tin về trạng thái cân bằng của nút

Bài 7. Xây dựng hàm để tìm và trả về con trỏ tới nút bên trái nhất của cây con phải của một cây AVL khác rỗng

Bài 8. Chứng minh rằng số lượng các phép xoay khi thực hiện xóa một nút trên cây không vượt quá $\frac{1}{2}$ chiều cao của cây

Bài 9. Viết hàm để xóa một nút trên cây AVL

Bài 10. Hoàn thiện chương trình mô phỏng cây tìm kiếm nhị phân cân bằng AVL

Bài 11. Giả sử chúng ta có một dãy số liên tiếp a_1, \dots, a_n và chúng ta cần trả lời nhanh cho câu hỏi dạng: cho i, j hãy tìm và trả về giá trị lớn nhất trong đoạn a_i, \dots, a_j

- Hãy xây dựng cấu trúc dữ liệu để lưu trữ dãy với chi phí $O(n^2)$ về bộ nhớ nhưng thời gian tìm và trả lời câu hỏi chỉ là $O(1)$
- Xây dựng cấu trúc mà chỉ cần chi phí bộ nhớ cỡ $O(n)$, và thời gian tìm và trả lời cho câu hỏi là $O(\log n)$

Bài 12. Cho mảng các số thực gồm n phần tử $A[1, \dots, n]$. Hãy thiết kế thuật toán để thực hiện các công việc sau:

- $Add(i, y)$ cộng giá trị y vào phần tử thứ i trong dãy

- $Partial_sum(i)$ trả về tổng của i phần tử đầu tiên trong dãy $\sum_{j=1}^i a_j$

Các thao tác trên cần có thời gian thực hiện cỡ $O(\log n)$. Ở đây ta không phải thêm hay xóa các phần tử trong dãy, chỉ thực hiện cộng giá trị các phần tử. Có thể sử dụng thêm bộ nhớ phụ nếu cần.

Bài 13. So sánh ưu nhược điểm của các cấu trúc cây AVL, Splay và cây 2-3.

Bài 14. Vẽ cây Splay thu được khi ta thực hiện thêm lần lượt các khóa sau vào cây ban đầu rỗng

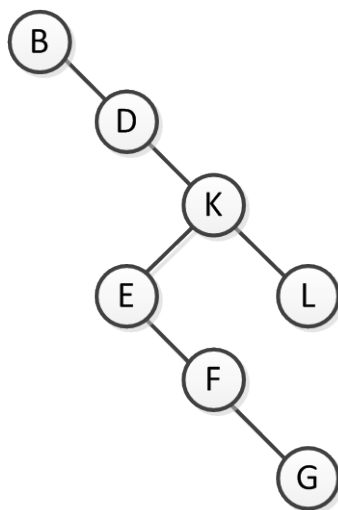
a. 12, 45, 23, 76, 3, 47, 12, 15, 32

b. A, B, D, C, G, H, F, E

c. A, B, C, D, E, F, G, H, I, J

d. 15, 13, 12, 11, 9, 8, 7, 5, 1

Bài 15. Cho cây nhị phân sau



Thực hiện splay tại các trường hợp khóa sau: G, F, L, E

Bài 16. Vẽ cây 2-3 thu được khi ta thực hiện thêm lần lượt các khóa sau vào cây ban đầu rỗng

a. 12, 45, 23, 76, 3, 47, 12, 15, 32

b. A, B, D, C, G, H, F, E

c. A, B, C, D, E, F, G, H, I, J

d. 15, 13, 12, 11, 9, 8, 7, 5, 1

Bài 17. Cài đặt hàm thực hiện các phép xoay trên cây Splay trong trường hợp Zig-Zig và Zag-Zag

Bài 18. Viết hàm tìm kiếm trên cây Splay (sau khi tìm kiếm phải thực hiện Splay)

Bài 19. Viết hàm thực hiện xóa một nút trên cây Splay

Bài 20. Viết hàm thực hiện tìm kiếm trên cây 2-3

Bài 21. Viết hàm thực hiện thêm một nút trên cây 2-3

Bài 22. Viết hàm thực hiện xóa nút trên cây 2-3

Bài 23. Giả sử bạn phải cài đặt một chương trình tra từ trong từ điển, bạn đã có từ điển khoảng 100,000 từ và ý nghĩa của mỗi từ. Bạn sẽ dùng cấu trúc nào để lưu trữ các từ để thực hiện tìm kiếm nhanh nhất (giả

sử sẽ không có thao tác thêm, xóa từ trong từ điển này). Các cấu trúc bạn có thể chọn là mảng, cây AVL, cây Splay và cây 2-3. Hãy giải thích lựa chọn của bạn một cách ngắn gọn.

Bài 24. Giả sử bạn có một tập các văn bản mà bạn lấy được từ web (chẳng hạn tập các bài báo trên vnexpress.net, hoặc voanews.com). Bạn phải tìm cách lưu trữ các văn bản này sao cho khi người dùng nhập vào một từ bất kỳ thì bạn phải đưa ra được các văn bản chứa từ đó (sắp xếp theo thứ tự tần số xuất hiện nhiều nhất trước). Hãy mô tả cấu trúc dữ liệu của bạn dùng để lưu trữ và thuật toán mà bạn dùng để tìm kiếm trong 2 trường hợp:

- a. Số lượng văn bản là cố định (không có thêm và xóa các văn bản trong tập)
- b. Số lượng văn bản có thể biến động (có sự thêm và xóa các văn bản thường xuyên)

Bài 25. Giả sử bạn cần quản lý thông tin là mối quan hệ của mỗi người dùng (kiểu như quản lý mối quan hệ giữa mọi người trong facebook), để tìm ra mối quan hệ bạn bè giữa 2 người

Ví dụ. Quan hệ bạn bè giữa tôi và bạn Tùng là Tôi>Hùng>Duy>Thành>Tùng

Hãy mô tả cấu trúc dữ liệu mà bạn dùng để thực hiện thao tác này.

Bài 26. Giả sử bạn có file dữ liệu về thông tin các thành viên của một mạng xã hội khoảng 5GB, mà bạn chỉ có một máy tính với dung lượng bộ nhớ trong khoảng 500MB. Bạn đang phải xây dựng một chương trình tìm kiếm các thành viên theo số CMND (là số nguyên gồm 9 chữ số và phân biệt cho mỗi thành viên). Hãy mô tả cấu trúc dữ liệu mà bạn dùng để lưu trữ các thành viên trên sao cho việc tìm kiếm diễn ra nhanh nhất có thể.

Chú ý: Trong trường hợp này bạn không đủ bộ nhớ trống để nạp thông tin tất cả các thành viên vào bộ nhớ trong máy tính để tìm kiếm. Bạn chỉ có thể nạp từng phần vào bộ nhớ trong máy tính, và thời gian nạp vào bộ nhớ trong là lâu.

Bài 27. Giả sử bạn có một file đầu vào chứa khoảng 4 tỷ số nguyên, hãy xây dựng một thuật toán để sinh ra số nguyên mà không có trong file này. Với giả sử bạn có 1GB bộ nhớ (10 MB bộ nhớ).

Bài 28. Giả sử bạn đang xây dựng một chương trình download tự động các trang web trên mạng – crawler. Bạn thu được khoảng một tỷ các url, hãy mô tả cách mà bạn quản lý các url này sao cho chúng không trùng nhau.