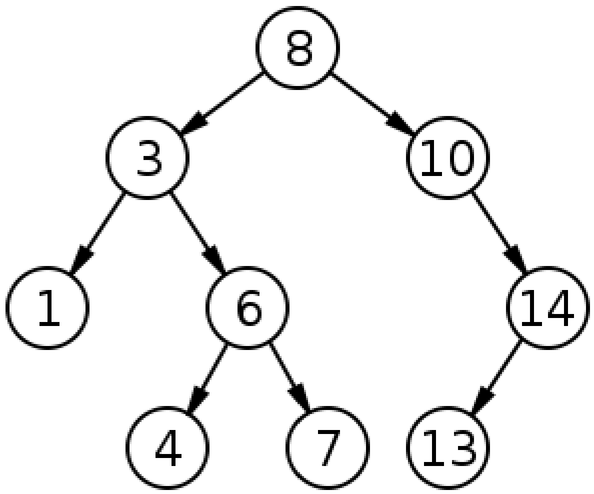
Zrównoważone drzewa BST

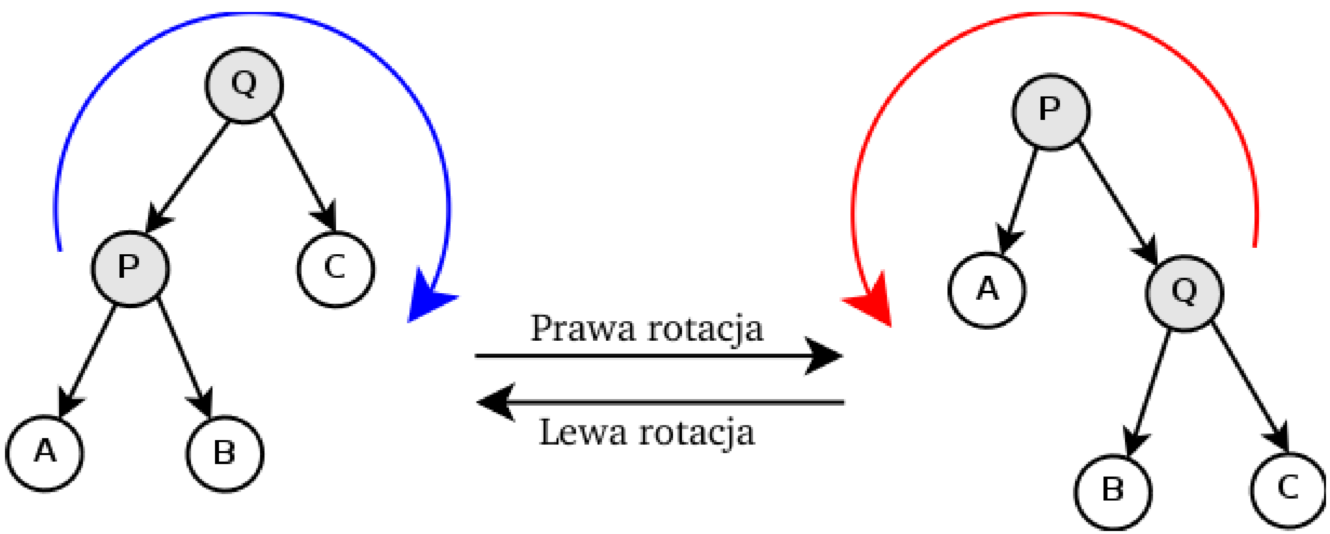
Drzewo BST

- Drzewo binarnych poszukiwań (ang. binary search tree), w skrócie BST, to drzewo binarne, w którym przechowujemy elementy z pewnego uniwersum z porządkiem liniowym i w drzewie tym jest zachowany porządek symetryczny.
- W drzewie binarnym jest zachowany porządek symetryczny, gdy elementy mniejsze od korzenia znajdują się w lewym poddrzewie, elementy większe od korzenia w prawym poddrzewie oraz w lewym i w prawym poddrzewie też jest zachowany porządek symetryczny.
- Liczba różnych n-elementowych drzew BST wynosi $C_n = \binom{2n}{n} / (n+1)$, czyli n-ta liczba Catalana.



Rotacje w drzewie BST

- Rotacja w lewo
- Rotacja w prawo
- W dowolnym wewnętrznym węźle drzewa BST można zrobić rotację (w lewo albo w prawo) i nie zostanie zniszczony porządek symetryczny w tym drzewie.

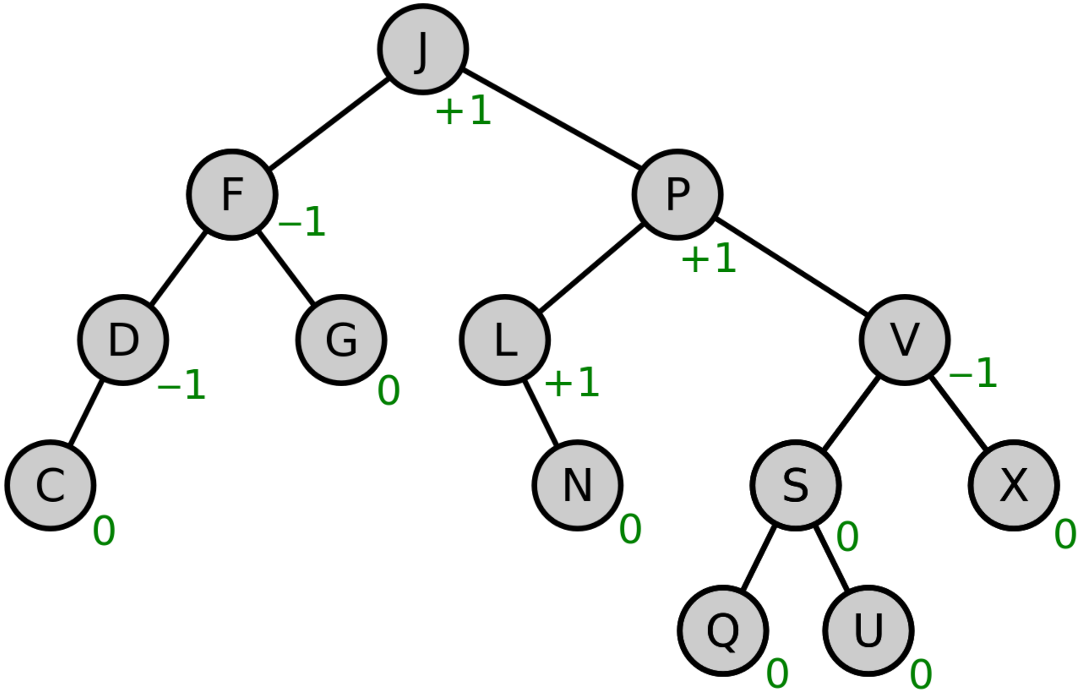


Drzewa AVL

- Drzewo AVL to opracowane przez Adelsona-Velskiego oraz Landisa w 1962 roku drzewo BST, w którym każdy węzeł ma dodatkowy atrybut: balans (2 bity na zaprezentowanie wartości -1, 0 lub 1).
- Warunek zrównoważenia w drzewie AVL: w każdym węźle drzewa AVL wysokość lewego i prawego poddrzewa różni się co najwyżej o jeden (w prostszej implementacji w węzłach takiego drzewa jest pamiętana wysokość poddrzewa zakorzenionego w danym węźle).
- Taka definicja zrównoważenia gwarantuje, że wysokość n-elementowego drzewa AVL wynosi O(log n).

Drzewa AVL

- Można udowodnić, że n-elementowe drzewo AVL ma wysokość $\leq 1.4405 \cdot \log_2(n+1)$, przez co elementarne operacje słownikowe będą wykonywać się w czasie O(log n).



Drzewa AVL

- Jaka jest minimalna i maksymalna liczba węzłów w drzewie AVL o wysokości h?
- Maksymalna liczba węzłów w drzewie AVL o wysokości h wynosi 2^h-1 (drzewo pełne o wysokości h).
- Drzewo AVL o wysokości h i minimalnej liczbie węzłów składa się z korzenia, jednego poddrzewa o minimalnej liczbie węzłów o wysokości h-1 i drugiego poddrzewa o minimalnej liczbie węzłów o wysokości h-2.
- Liczba M_h węzłów w minimalnym drzewie AVL o wysokości h wynosi: 0 dla drzewa pustego, 1 dla drzewa z jednym węzłem oraz $M_{h-1} + 1 + M_{h-2}$ w drzewie o wysokości h>1.
- Można indukcyjnie udowodnić, że $M_h = F_{h+2} – 1$ (F_h to liczby Fibonacciego).
- Więc $h = \log_{(1+\sqrt{5})/2} n + O(1)$.

Drzewa AVL

- Drzewo AVL jest drzewem zrównoważonym, którego wysokość jest rzędu $\Theta(\log n)$, gdzie n to liczba węzłów w całym drzewie.
- Zrównoważenie drzewa osiąga się, przypisując każdemu węzłowi współczynnik wyważenia, który jest równy różnicy wysokości lewego i prawego poddrzewa. Może wynosić 0, +1 lub -1.
- Wstawiając lub usuwając elementy drzewa (tak, by zachować własności drzewa BST), modyfikuje się też współczynnik wyważenia, a gdy przyjmie on niedozwoloną wartość, wykonuje specjalną operację rotacji węzłów, która przywraca zrównoważenie.

Drzewa AVL – wyszukiwanie

- Wyszukiwanie w drzewie AVL jest wykonywane tak samo jak w niezrównoważonym drzewie BST dzięki serii porównań wartości wyszukiwanej z wierzchołkami, poczynając od korzenia.
- Wynik porównania oraz istnienie poddrzew pozwala stwierdzić, czy element szukany jest wierzchołkiem, znajduje się w lewym bądź prawym poddrzewie, czy też nie należy do drzewa. Zachowanie struktury AVL pozwala jednak na redukcję pesymistycznego czasu wyszukiwania do O(lg n).
- Operacja wyszukiwania nie modyfikuje struktury drzewa, więc nigdy nie pociąga wyważania drzewa (rotacji węzłów).

Drzewa AVL – wstawianie

- Wstawianie do drzewa AVL może odbywać się tak, jak gdyby było ono zwyczajnym drzewem poszukiwań binarnych, następnie zaś są odtwarzane kroki w kierunku korzenia i wykonywane aktualizacje wyważeń węzłów.
- Zmiana wartości współczynnika wyważenia na 0 oznacza, iż wysokość poddrzewa nie została zmieniona. Operacja wstawiania jest wówczas zakończona.
- Zmiana współczynnika wyważenia na 1 lub -1 oznacza, iż poddrzewo, którego korzeniem jest rozważany wierzchołek, zachowało własność drzewa AVL, lecz jego wysokość uległa zwiększeniu. Informacja o tym przekazywana jest do ojca tego węzła.
- Jeśli współczynnik został zmieniony na 2 lub -2, to drzewo straciło własność AVL. Aby ją przywrócić, potrzebna jest rotacja. Maksymalnie będą potrzebne dwie rotacje, po których nie ma potrzeby wykonywania dalszych aktualizacji.

Drzewa AVL – usuwanie

- Jeśli usuwany węzeł jest liściem, zostaje usunięty. Jeśli nie jest liściem, musi zostać zastąpiony największym elementem z jego lewego poddrzewa lub najmniejszym z jego prawego poddrzewa.
- Wyszukany największy lub najmniejszy element ma co najwyżej jedno dziecko, które zostaje teraz dzieckiem rodzica wyszukanego elementu. Po jego usunięciu odtwarzana jest ścieżka od rodzica wyszukanego elementu do korzenia, zaś współczynniki wyważenia są aktualizowane.
- Odtwarzanie może być zatrzymane, jeśli współczynnik wyważenia zostaje zmieniony na -1 lub 1, oznacza to bowiem, iż wysokość poddrzewa pozostaje niezmieniona.
- Zmiana współczynnika wyważenia na 0 oznacza zmniejszenie wysokości poddrzewa, aktualizowanie współczynników musi być kontynuowane.
- Jeśli współczynnik zostanie zmieniony na -2 lub 2, to wykonywana jest rotacja w celu przywrócenia struktury AVL. W przeciwieństwie do operacji wstawiania wystąpienie rotacji nie musi oznaczać, iż drzewo zostało wyważone. W pesymistycznym przypadku rotacje będą wykonywane aż do korzenia drzewa.

Drzewa R-B (czerwono-czarne)

- Drzewo czerwono-czarne (ang. red-black tree), lub krótko drzewo RB, to opracowane przez Rudolfa Bayera w 1972 roku drzewo BST, w którym każdy węzeł ma dodatkowy atrybut: kolor (1 bit na zaprezentowanie koloru czarnego albo czerwonego).
- Warunek zrównoważenia w drzewie RB: w każdym węźle drzewa RB ścieżka do dowolnego wskaźnika pustego w poddrzewie zaczepionym w tym węźle przechodzi przez tyle samo węzłów czarnych, a węzły czerwone nigdy ze sobą nie sąsiadują.
- Taka definicja zrównoważenia gwarantuje, że wysokość n-elementowego drzewa RB wynosi O(log n).

Drzewo R-B (czerwono-czarne)

- Drzewo RB jest drzewem zrównoważonym, którego wysokość jest rzędu $\Theta(\log n)$, gdzie n to liczba węzłów w całym drzewie.
- Zrównoważenie drzewa RB osiąga się dzięki przestrzeganiu następujących reguł:
  - każdy węzeł jest czerwony lub czarny;
  - korzeń zawsze jest czarny;
  - każdy liść jest czarny (traktujemy wszystkie wskaźniki puste null jako liście);
  - jeśli węzeł jest czerwony, to jego synowie muszą być czarni;
  - każda ścieżka z korzenia do liścia zawiera tyle samo czarnych węzłów.
- Wstawiając elementy do lub usuwając z drzewa RB (tak, by zachować własności drzewa BST), wykonuje się operacje rotacji węzłów i przekolorowywania, aby przywrócić wszystkie własności drzewa RB (zrównoważenie).

Drzewo R-B (czerwono-czarne)

- Warunki narzucone na drzewa czerwono-czarne gwarantują, że najdłuższa ścieżka od korzenia do liścia będzie co najwyżej dwukrotnie dłuższa, niż najkrótsza – wynika to wprost z ostatniej własności.
- Można udowodnić, że n-elementowe drzewo czerwono-czarne ma wysokość $\leq 2 \cdot \log_2(n+1)$, przez co elementarne operacje słownikowe będą wykonywać się w czasie $O(\log n)$.

Drzewo R-B (czerwono-czarne)

- Można udowodnić, że n-elementowe drzewo czerwono-czarne ma wysokość $\leq 2 \cdot \log_2(n+1)$, przez co elementarne operacje słownikowe będą wykonywać się w czasie O(log n).