

Tablica asocjacyjna - Benchmark

Konrad Łakomy

24 04 2014

Zadanie do wykonania

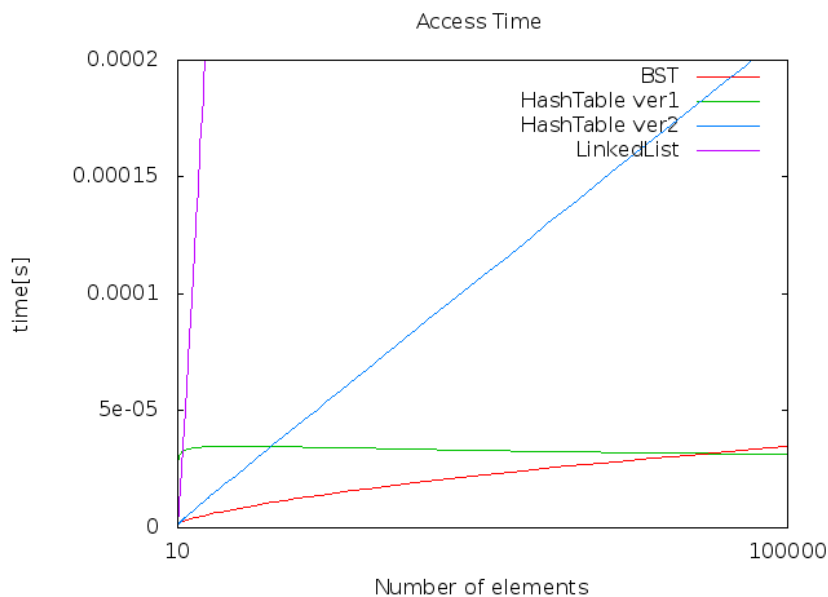
Implementacja i benchmark tablicy asocjacyjnej na wymienionych niżej strukturach:

1. Drzewo Binarne (ang. Binary Search Tree)
2. Tablica Haszująca (ang. Hash Table)
 - (a) Wersja1: Rozmiar tablicy równy liczbie elementów
 - (b) Wersja2: Rozmiar tablicy stały równy 100
3. Lista jednokierunkowa (ang. Linked List)

Porównanie czasu dostępu do pojedynczego elementu

LiczbaElementów	BST	HashTable v1	HashTable v2	LinkedList
10	1.67614e-06[s]	2.5747e-05[s]	9.14772e-07[s]	1.50405e-06[s]
100	2.26206e-06[s]	3.6347e-05[s]	1.85324e-06[s]	6.00731e-06[s]
1000	3.22015e-06[s]	4.3055e-05[s]	5.03528e-06[s]	2.06877e-05[s]
10000	1.46007e-05[s]	2.536e-05[s]	2.40773e-05[s]	0.000471269[s]
100000	8.8078e-05[s]	2.7515e-05[s]	0.000212257[s]	-
1000000	9.3665e-05[s]	2.8618e-05[s]	-	-

Wykres czasu dostępu w zależności od rozmiaru tablicy



Rysunek 1: Czas dostępu do pojedynczego elementu

Wnioski

W celu zapewnienia rzetelności wyników każdy pomiar został wykonany kilkunastokrotnie a następnie wynik został uśredniony. Na podstawie tabeli z wynikami oraz wykresu można wyciągnąć następujące wnioski:

1. Efektywność tablicy asocjacyjnej w dużej mierze zależy od tego w jakiej strukturze danych jest zaimplementowana. W zależności czy jest to drzewo binarne, tablica haszująca czy lista jednokierunkowa to implementacja tablicy asocjacyjnej będzie charakteryzować się różnym czasem dostępu do elementu.

2. W przypadku drzewa binarnego czas dostępu do pojedynczego elementu jest $O(\log n)$. Duży wpływ na to jak szybko znajdziemy dany klucz będzie zależało od jego umiejscowienia w drzewie tj. czy będzie na początku czy na końcu drzewa.

3. Tablica asocjacyjna charakteryzuje się praktycznie stałym czasem dostępu w przypadku gdy liczba elementów jest równa rozmiarowi tablicy. Natomiast w przypadku gdy rozmiar tablicy jest mniejszy niż liczba elementów czas dostępu zaczyna rosnąć wraz ze wzrostem liczby elementów ponieważ pojawiają się kolizje, tzn. pod jednym indeksem może znajdować się kilka elementów zapisanych na strukturze listy. Co powoduje że aby znaleźć dany element trzeba przeszukać całą listę.

4. Struktura listy jednokierunkowej charakteryzuje się bardzo dużym czasem dostępu do elementu, ze względu na to iż aby znaleźć zadany element trzeba przeszukać całą listę. W tym przypadku wyszukiwanie liniowe zależy od ilości elementów znajdujących się w liście.

5. Podsumowując można stwierdzić że najbardziej wydajną strukturą do implementacji tablicy asocjacyjnej jest struktura tablicy haszującej, natomiast najmniej wydajna jest struktura listy jednokierunkowej.