

Sprawozdanie: Tablica asocjacyjna

Mateusz Król 226400

June 19, 2017

1 Wstęp oraz krótki opis programu

Celem obecnych zajęć było zapoznanie się z nową strukturą danych; tablicą asocjacyjną. Badano złożoność obliczeniową odczytu z tablicy dla 2 różnych funkcji haszujących. Badania prowadzono przy różnej ilości bucketów oraz różnej ilości elementów.

Tablica asocjacyjna ogólnie przechowuje zawartość w postaci klucz i wartość, co umożliwi szybszy dostęp do podanej wartości po podaniu klucza. Spodziewana złożoność obliczeniowa to $O(n)$.

Wybrano 2 funkcje haszujące; Pierwsza została zaprojektowana podczas zajęć i polega na zamianie dwóch pierwszych znaków klucza na kod ASCII oraz zsumowaniu ich wartości. Funkcja na wyjściu podaje resztę z dzielenia obliczonej wcześniej wartości przez liczbę bucketów

Dруга funkcja haszująca została zaczerpnięta z biblii Cormena 11.3. Nazywała się haszowanie przez mnożenie. Funkcja na wyjściu podaje resztę z dzielenia obliczonej wcześniej wartości przez liczbę bucketów. Obie funkcje znajdują się w pliku hashtable.cpp linie 47-72.

2 Uśrednione czasy dla 100 pomiarów

.....Zliczanie ASCII				
ilosc elementow	1 bucket [ms]	10 bucketów [ms]	100 bucketów	1000 bucketów
10	0.00017	0.00016	0.00009	0.000065
100	0.00042	0.0002	0.00021	0.00011
1000	0.00485	0.00065	0.00033	0.0003
10000	0.03358	0.0052	0.00152	0.00105
100000	0.41	0.1018	0.0167	0.00299
.....Mnożenie				
1	0.00011	0.00016	0.00015	0.00009
10	0.000455	0.00033	0.00014	0.00017
1000	0.00375	0.00085	0.00038	0.00038
10000	0.03517	0.00534	0.00141	0.001
100000	0.3415	0.11428	0.01844	0.01128

Table 1: Zestawienie czasów odczytu z tablicy asocjacyjnej w zależności od bucketów i liczby elementów.

3 Wnioski

Jak możemy zauważyć z tabeli powyżej i z wykresów poniżej (skala log-log) wraz ze wzrostem liczby bucketów zmniejsza się czas odczytu z tablicy.

Możemy również zauważyć, że w przypadku małej ilości danych złożoność obliczeniowa jest mniejsza niż $O(n)$ przy dużej ilości danych zaczyna ona już przypominać złożoność $O(n)$.

Z pewnością na czas miała wpływ ilość kolizji, dodatkowo brano do wyszukania losowy element, co nawet przy 100 pomiarach mogło zaburzyć znacznie średnią, ponieważ prezentowane czasy są bardzo niewielkie.

Lepsza okazała się funkcja haszująca numer 2, ale czasy przydzielania kluczy są bardzo zbliżone.

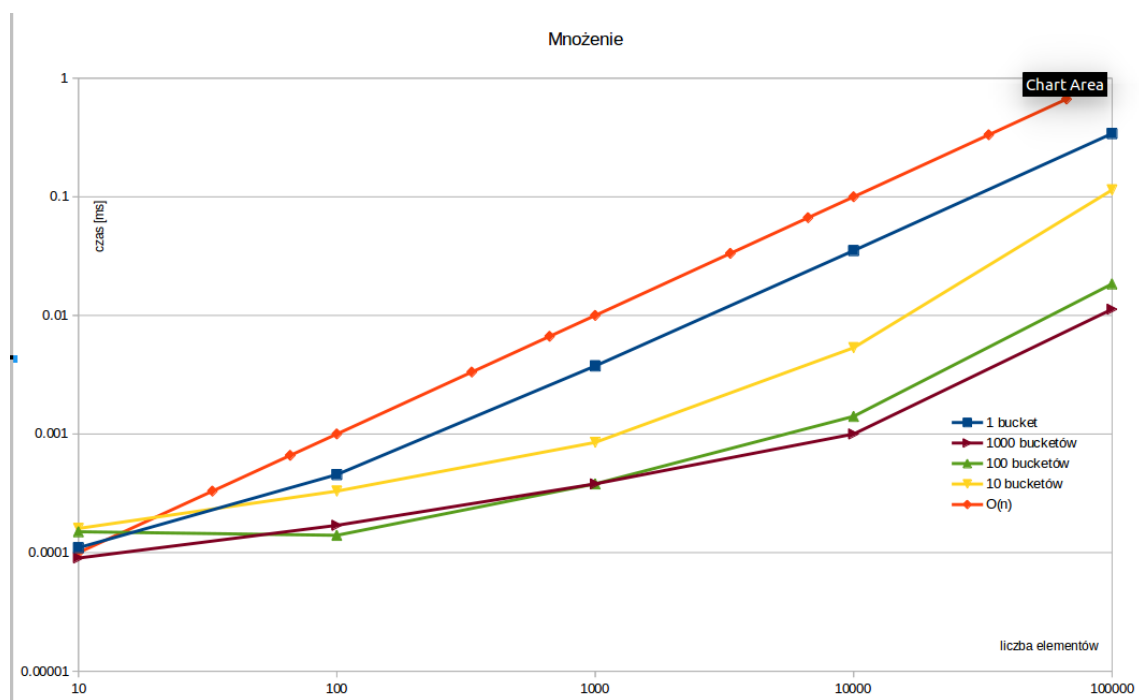


Figure 1: Zestawienie czasów przeszukiwanie tablicy

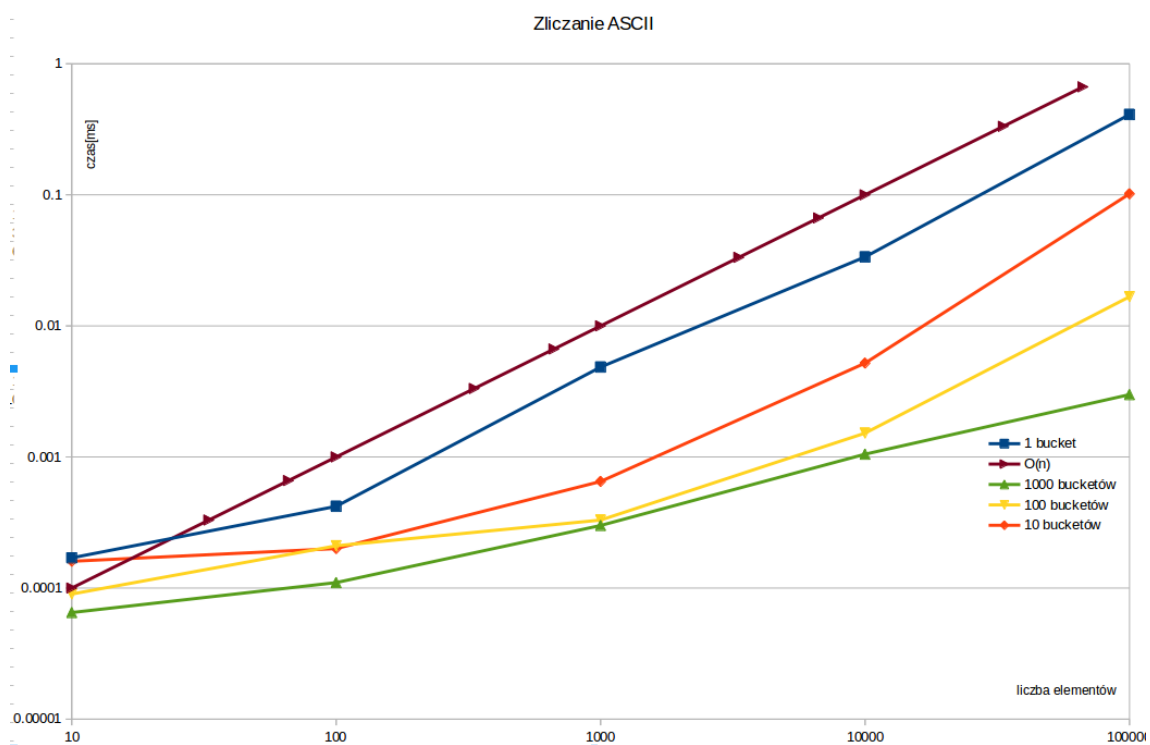


Figure 2: Zestawienie czasów przeszukiwanie tablicy