

PAMSI - Sprawozdanie 6

Filip Guzy 218672

25 kwietnia 2016

Drzewo czerwono-czarne

Drzewo binarne to struktura danych w której poszczególne elementy są ze sobą w relacjach typu ojciec-syn. Jedną z odmian drzew binarnych są drzewa czerwono-czarne, w których liście nie przechowują danych, a zamiast tego stosuje się koncepcję tzw. strażnika, czyli węzła pełniącego rolę wszystkich liści. Wszystkie węzły takiej struktury są określane kolorem czerwonym lub czarnym (pokolenia sąsiadujące nie mogą mieć tego samego koloru), co umożliwia utrzymanie odpowiedniej wysokości drzewa, a co za tym idzie odpowiedniej złożoności obliczeniowej wykonywanych na nim operacji. Dokonując zapisu elementu drzewo czerwono-czarne ulega reorganizacji w przypadku naruszenia porządku kolorów. Poniżej przedstawiono złożoności obliczeniowe zapisu i odczytu zaimplementowanego w celach testowych drzewa.

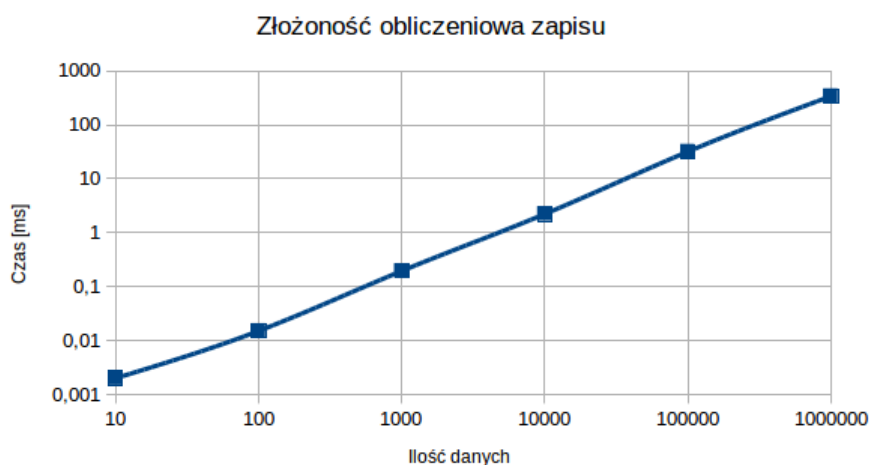
Pomiary zapisu i odczytu danych

W celu sprawdzenia złożoności obliczeniowej zapisu danych do drzewa binarnego wczytywano kolejno następujące ilości danych: 10, 100, 1000, 10000, 100000 oraz 1000000 elementów. Zgodnie z założeniami teoretycznymi pojedyncza operacja wstawiania do drzewa binarnego powinna wykonywać się w czasie $\log n$ dla przypadku średniego oraz w czasie n dla przypadku pesymistycznego, czyli złożoność wstawiania n ilości danych powinna być rzędu $n \log n$ w przypadku średnim i n^2 w pesymistycznym. W poniższej tabeli przedstawiono czasy wykonania operacji wstawiania dla różnych ilości danych wejściowych.

Tabela 1: Zapis elementów do drzewa czerwono-czarnego

Ilość	10	100	1000	10000	100000	1000000
Czas [ms]	0,00200009	0,015	0,19	2,181	32,414	334,678
	0,00199997	0,015	0,212	2,186	29,565	348,181
	0,000999928	0,015	0,189	2,179	29,804	345,687
	0,00199997	0,015	0,191	2,261	33,979	340,951
	0,00199997	0,015	0,191	2,209	30,934	340,452
	0,00200009	0,016	0,193	2,264	33,027	334,354
	0,00199997	0,0139999	0,189	2,239	32,828	343,533
	0,00199997	0,015	0,189	2,182	32,946	336,066
	0,00300002	0,016	0,217	2,309	31,383	340,965
	0,00199997	0,015	0,19	2,229	32,53	339,658
Średnia	0,0019999948	0,01509999	0,1951	2,2239	31,941	340,4525

Pomiary przedstawiono także na poniższym wykresie:



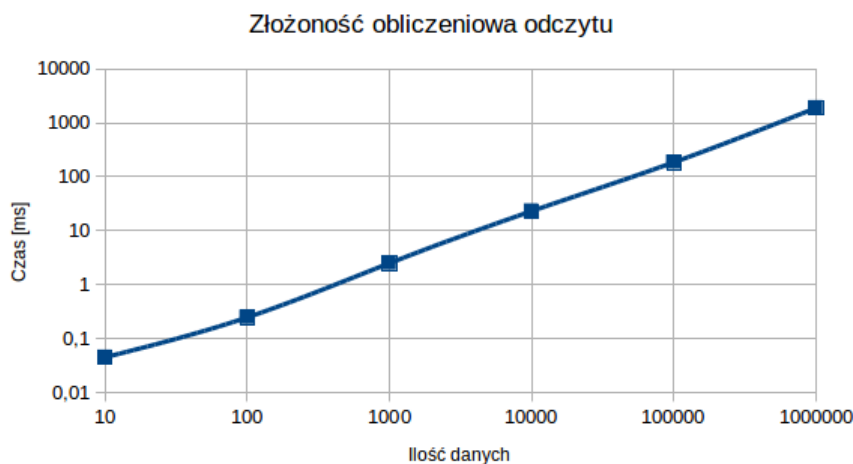
Jak można zauważyć z tabeli i wykresu, otrzymano złożoność zapisu $O(n \log n)$, zatem zapis pojedynczego elementu do drzewa wykonuje się w czasie $\log n$, co jest zgodne z założeniami teoretycznymi.

Aby sprawdzić złożoność obliczeniową odczytu danych z drzewa po dodaniu miliona losowych elementów wyszukiwano n razy nieistniejący w zbiorze element '-1', aby zminimalizować szansę wyszukiwania elementów na najniższych poziomach drzewa. Zgodnie z teorią odczyt pojedynczego elementu z drzewa powinien być rzędu $\log n$ w średnim przypadku, a n w pesymistycznym, zatem złożoność odczytu n elementów powinna być rzędu $n \log n$ dla przypadku średniego i n^2 dla pesymistycznego. Poniżej przedstawiono w tabeli oraz na wykresach czasy wyszukiwań 10, 100, 1000, 10000, 100000 i 1000000 elementów.

Tabela 2: Zapis elementów do drzewa czerwono-czarnego

Ilość	10	100	1000	10000	100000	1000000
Czas [ms]	0,046	0,251	2,456	24,853	184,724	1889,34
	0,046	0,26	2,486	19,365	183,036	1896,24
	0,046	0,236	2,465	24,776	187,088	1888,44
	0,044	0,238	2,413	22,776	184,434	1863,52
	0,044	0,252	2,502	20,457	191,771	1895,73
	0,043	0,255	2,492	21,492	183,752	1903,89
	0,0419999	0,24	2,729	24,94	185,014	1889,4
	0,044	0,239	2,483	25,289	185,209	1884,02
	0,046	0,253	2,431	19,562	183,83	1883,76
	0,046	0,237	2,42	26,775	185,064	1863,84
Średnia	0,04469999	0,2461	2,4877	23,0285	185,3922	1885,818

Pomiary przedstawiono także na poniższym wykresie:



Z obserwacji tabeli i wykresu można wywnioskować, że złożoność odczytu n ilości danych dokonuje się w czasie $n \log n$, zatem pojedyncze przeszukanie drzewa ma złożoność $O(\log n)$. Otrzymane złożoności są zgodne z założeniami teoretycznymi.

Wnioski

Drzewo czerwono-czarne jest strukturą danych o efektywnych średnich złożonościach obliczeniowych wyszukania i zapisu. Testy zaimplementowanego drzewa wykazały zgodność z założeniami teoretycznymi, czyli złożoność zapisu i odczytu pojedynczego elementu zawiera się w $O(\log n)$,