27.03.2018r.

Struktury danych i złożoność obliczeniowa

Projekt: Struktury danych

Prowadzący: dr inż.. Dariusz Banasiak  
Autor: Wojciech Wójcik  
Indeks: 235621  
Wtorek 15:15 TN

1. Wstęp

Głównym celem było samodzielne zaimplementowanie danych struktur i przeprowadzenie na nich odpowiednich eksperymentów.

Zaimplementowane zostały następujące struktury danych:

* Tablica
* Lista dwukierunkowa
* Kopiec binarny(typu maksimum – element maksymalny w korzeniu)

Eksperymenty miały na celu bezpośrednie porównanie wyżej wymienionych struktur poprzez mierzenie czasu wykonywania podstawowych operacji. Ma to na celu pokazanie do czego dana struktura nadaje się najlepiej. W tym celu został też stworzony interfejs umożliwiający masowe testowanie i generowanie danych. Zostało przeprowadzone ponad 500 eksperymentów, w których został zmierzony czas każdej operacji każdej struktury.

1. Plan eksperymentu:

Wykorzystane środowisko programistyczne: Visual Studio 2017

Język: C i C++

W programie zostały zaimplementowane struktury danych., oraz interfejs umożliwiający testowanie.  
Dane są generowane do i wczytywane z pliku.  
Testy wykonane dla następującej ilości elementów: 100, 1000, 20000, 50000 oraz 1000000.  
Mierzenie czasu zostało wykonane za pomocą QueryPerformanceCounter.   
Mierzenie czasu było uruchamiane zaraz przed i odczytywanie zaraz po uruchomieniu mierzonej funkcji.

Dla każdej struktury została zaimplementowana metoda dodawania elementu, usunięcia elementu i wyszukanie elementu. Dodatkowo dla tablic i listy rozpatrzono osobne operacje dodawania na początku, w dowolnym miejscu i końcu tablicy. Z każdym pomiarem miejsce wstawiania, usuwania elementu oraz dane były generowane losowo.

1. Krótki opis struktur

* **Lista dwukierunkowa** – posiada wskaźnik na poprzedni i następny element. Usuwanie i dodawanie odbywa się poprzez modyfikację wskaźników poprzedniego i następnego elementu w liście. Wskaźniki pierwszego i ostatniego elementu wskazują na wskaźnik zerowy (null pointer).  
  Tworzenie tej struktury odbywa się poprzez dodawanie elementu na początek, jest to dużo szybsze (podczas prób przy dodawaniu elementu na koniec przy tworzeniu struktury czas jej tworzenia wynosił ponad 10s, dlatego porzuciłem dalsze testowanie.)
* **Tablica dynamiczna** – przechowuję wskaźnik na tablicę danych. Realizuje operację modyfikacji struktury zawsze poprzez utworzenie nowej tablicy i przepisanie do niej wartości uwzględniając daną operację.
* **Kopiec typu maksimum** – jest zaimplementowany na podstawie tablicy dynamicznej. Zostały dodane metody pozwalające przechodzić po strukturze w poprawny sposób (wyliczanie indeksu dziecka itp.). Został zaimplementowany w niej interfejs umożliwiający dodanie, usunięcie oraz szukanie, jednak korzysta on z metod tablicy dynamicznej. Interfejs ten również czuwa by zawsze były spełnione warunki kopca.

1. Wyniki

Średnie wyniki wczytywań struktur (wyniki w milisekundach):

|  |  |  |  |
| --- | --- | --- | --- |
| Rozmiar | Średnia z dynamicznej tablicy | Średnia z kopca | Średnia z listy dwukierunkowej |
| 100 | 0,4857398 | 0,5518034 | 0,4360011 |
| 1000 | 5,454463 | 5,706945 | 4,326826 |
| 20000 | 234,0677 | 253,5184 | 93,57667 |
| 50000 | 1088,252 | 1099,091 | 234,0363 |
| 100000 | 3805,921111 | 3818,308 | 461,2496 |

Lista dwukierunkowa(wyniki w milisekundach):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Rozmiar | Średnia z dodanie na start | Średnia z usuniecie ze startu | Średnia z dodanie na końcu | Średnia z usuniecie z końca |
| 100 | 0,00145 | 0,000798 | 0,00158 | 0,00143 |
| 1000 | 0,00145 | 0,000767 | 0,00910 | 0,00779 |
| 20000 | 0,00161 | 0,000873 | 0,1973 | 0,178 |
| 50000 | 0,00189 | 0,000878 | 1,22 | 1,25 |
| 100000 | 0,00225 | 0,000991 | 4,43 | 4,85 |

|  |  |  |  |
| --- | --- | --- | --- |
| Rozmiar | Średnia z dodanie na pozycji | Średnia z usuniecie z pozycji | Średnia z wyszukiwania |
| 100 | 0,00100 | 0,00100 | 0,000544 |
| 1000 | 0,00465 | 0,00467 | 0,00282 |
| 20000 | 0,0516 | 0,0473 | 0,0676 |
| 50000 | 0,186 | 0,0856 | 0,416 |
| 100000 | 0,583 | 0,155 | 1,92 |

Kopiec typu maksimum (wyniki w milisekundach):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Rozmiar | Średnia z wyszukiwania | Średnia z usuwania | Średnia z dodawania | Średnia z wyszukiwania |
| 100 | 0,00809 | 0,00357 | 0,00462 | 0,00809 |
| 1000 | 0,0547 | 0,00456 | 0,00682 | 0,0547 |
| 20000 | 1,42 | 0,0465 | 0,0218 | 1,42 |
| 50000 | 4,97 | 0,152 | 0,0452 | 4,97 |
| 100000 | 11,1 | 0,304 | 0,0881 | 11,1 |

Tablica dynamiczna (wyniki w milisekundach):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Rozmiar | Średnia z dodanie na start | Średnia z usuniecie ze startu | Średnia z dodanie na końcu | Średnia z usuniecie z końca |
| 100 | 0,00260 | 0,00113 | 0,00107 | 0,000979 |
| 1000 | 0,00423 | 0,00206 | 0,00176 | 0,00168 |
| 20000 | 0,0360 | 0,0183 | 0,0142 | 0,0136 |
| 50000 | 0,0911 | 0,0398 | 0,0341 | 0,0372 |
| 100000 | 0,196 | 0,0818 | 0,0738 | 0,0929 |

|  |  |  |  |
| --- | --- | --- | --- |
| Rozmiar | Średnia z dodanie na pozycji | Średnia z usuniecie z pozycji | Średnia z wyszukiwania |
| 100 | 0,00120 | 0,00106 | 0,000402 |
| 1000 | 0,00190 | 0,00177 | 0,00177 |
| 20000 | 0,0164 | 0,0171 | 0,0286 |
| 50000 | 0,0438 | 0,0363 | 0,0675 |
| 100000 | 0,0944 | 0,0755 | 0,149 |

1. Wnioski i analiza wyników
   1. Wczytywanie

Przy czytywaniu najwydajniejsza okazała się lista dwukierunkowa, ale tylko w przypadku, gdy tworzenie struktury odbywało się poprzez dodanie na początek. W innych strukturach nie miało to widocznego znaczenia.

Z uwagi na to że kopiec dziedziczy po dynamicznej tablicy, czas jego tworzenia jest minimalnie dłuższy, jest to spowodowanie dodatkowymi warunkami, które kontrolują poprawność kopca.

Ponadto, czasy tworzenia kopca i tablicy dynamiczna nie rosną liniowo, tak jak lista dwukierunkowa – jest to spowodowane, tym że, przy każdym dodaniu trzeba przekopiować całą zawartość struktury.

* 1. Operacja na strukturach

Wedle przypuszczeń najszybszą metodą dodawania elementów do struktury jest dodawanie na start w liście dwukierunkowej. Ciekawa jest różnica w tablicy dynamicznej przy dodawaniu na początek, a dodawaniu na koniec. Jest ono ponad 2 razy szybsze od dodania na początek (w przypadku rozmiaru struktury o rozmiarze 100000) – takiego wyniku nie spodziewałem się.

W przypadku Listy dwukierunkowej najszybszą operacjami są te wykonywane jak najbliżej początku.

Najszybsze wyszukiwanie daje nam tablica dynamiczna, jest to prawdopodobnie spowodowane ułożeniem danych w pamięci.   
Ciekawe za to są dość długie czasy wyszukiwani w kopcu, oczekiwałem dużo mniejszych. Jak widać dużo szybsze jest przeszukanie całej tablicy. Co prawda kopiec eliminuje pewne gałęzie drzewa, lecz dodatkowe warunki, jak i pobieranie danych z różnych miejsc pamięci skutecznie ograniczają tą strukturę.

W przypadku dodawania na określoną pozycję tablica dynamiczna również okazała się najszybsza.

1. Bibliografia
   1. Materiał z ćwiczeń dr inż. Jarosława Mierzwy
   2. <http://eduinf.waw.pl/inf/>