

SBD Sprawozdanie z zadania 2.

Filip Gołaś 188776

17 grudnia 2023

1 Zadanie

W ramach zadania zaimplementowano indeksową organizację pliku w postaci B-drzewa z możliwością wstawiania, usuwania, wyszukiwania i modyfikowania rekordów. Kluczem jest sekwencja pięciu liczb naturalnych rozszerzona o klucz (liczbę naturalną).

1.1 Implementacja

Program w trakcie działania przechowuje w buforze sekwencję wierzchołków drzewa od korzenia do badanego liścia by uniknąć wielokrotnego odczytywania. Po każdej operacji na drzewie, bufor jest wpisywany do pliku by jego zawartość była aktualna. Wierzchołki nie są zapisywane jeżeli nie uległy zmianie. Struktura wierzchołka b-drzewa zawiera w sobie tablicę struktur BRecord, które zawierają klucz i wskaźnik do rekordu w pliku danych i tablicę wskaźników plikowych do wierzchołków potomnych. W celu usprawnienia działania programu, wierzchołki posiadają także: adres rodzica, liczbę wpisanych rekordów, swój własny adres oraz informację, czy są zgodne z plikiem. Wierzchołki są zapisywane do pliku poprzez serializację do postaci binarnej. Rekordy, do których odwołują się wierzchołki, są zapisywane w pliku danych w zserializowanej postaci binarnej. Rekordy nie są w tym pliku nigdy przesuwane, są zawsze dodawane na koniec pliku, a ich adresy są zapisywane w wierzchołkach.

1.2 Plik testowy

Testować program można z użyciem pliku testowego oraz poprzez tryb interaktywny. Polecenia w obydwu sposobach są identyczne: '+ n' - dodaj rekord o kluczu n, gdzie n to liczba naturalna. '-' n' - usuń rekord o kluczu n, gdzie n to liczba naturalna. '? n' - wyszukaj rekord o kluczu n, gdzie n to liczba naturalna. '=' n' - zmodyfikuj rekord o kluczu n, gdzie n to liczba naturalna. W pliku testowym polecenia są zapisane w kolejnych liniach, w trybie interaktywnym należy wpisywać je i zatwierdzać enterem.

1.3 Prezentacja wyników

Wierzchołki drzewa są reprezentowane w postaci tabelki, na przykład:

```
BNode & 0
&3
{K2 & 1}
&4
{}
—
```

Oznacza wierzchołek zapisany w pliku o adresie 0. Ma on jeden rekord o kluczu 2, który jest zapisany w pliku danych pod adresem 1. Wierzchołek ma dwoje dzieci. Pod adresem 3 w pliku z wierzchołkami znajduje się potomek o kluczach mniejszych od 2, pod adresem 4 potomek o kluczach większych od 2. Drugi rekord w wierzchołku jest pusty - {}, ostatnie dziecko jest puste (-).

W trakcie trybu interaktywnego można poruszać się po drzewie poprzez:

- 'n' - przejdź do n-tego dziecka.
- 'p' - przejdź do rodzica.
- 'r' - przejdź do korzenia.

W trakcie poruszania się wyświetlany jest aktualnie wybrany wierzchołek. Można wyświetlać zawartość drzewa poprzez:

- 't' - wyświetl całe drzewo.
- 's' - wyświetl rekordy w aktualnym wierzchołku.
- 'a' - wyświetl wszystkie rekordy występujące w drzewie zgodnie z uporządkowaniem

Można operować na drzewie poprzez:

- '? n' - wyszukaj rekord o kluczu n, gdzie n to liczba naturalna.
- '+ n' - dodaj rekord o kluczu n, gdzie n to liczba naturalna.
- '- n' - usuń rekord o kluczu n, gdzie n to liczba naturalna.
- '= n m' - zmodyfikuj rekord o kluczu n, tak by miał klucz m, gdzie n i m to liczby naturalne.

Operacje na drzewie wymagają powrotu do korzenia, więc aktywny wierzchołek zmieni się na korzeń.

2 Eksperymenty

Brak

3 Wnioski

B-drzewo jest strukturą, która pozwala na efektywne wyszukiwanie z ilością koniecznychostępów ograniczoną przez wysokość drzewa. Modyfikacja drzewa również wymaga dostępów do plików zależnych od wysokości drzewa, która rośnie wraz z powiększaniem się drzewa bardzo wolno, bo jest ograniczona poprzez $\log_d N$, gdzie d to stopień drzewa (ilość rekordów w wierzchołku), a N to całkowita ilość rekordów w drzewie. B-drzewo jest strukturą stabilną dzięki zagwarantowaniu równej długości ścieżki od korzenia do każdego z liści. Dodatkowo można mieć pewność, że w każdej chwili pamięć jest wykorzystywana co najmniej w 50% a średnio w 70%. Implementacja b-drzewa jest trudna, jednak w przypadku tworzenia systemu bazodanowego wysokiej jakości działającego na bardzo dużych zbiorach danych warto zaimplementować tę strukturę lub jej pochodne.