

Sprawozdanie Struktury Baz Danych Projekt 1

Sebastian Kwaśniak

2024-12-08

Wprowadzenie

Wylosowane przeze mnie typ rekordu to:

29. File records: Right circular cylinders - the radius of the base and the height of the cylinder. Sorting by volume.

Implementacja w języku C++. Przyjąłem, że jeden rekord jest podzielony na cztery liczby, rozmiar rekordu to 16 bajtów (4 bajty dla klucza, 4 bajty dla podstawy, 4 bajty dla wysokości, 4 bajty dla wskaźnika).

Opis struktury kodu

Kod został głównie przeniesiony z projektu 1, w którym:

- Klasa `Tape` zajmuje się obsługą zarówno głównej taśmy oraz przepełnienia
- Klasa `Index` zajmuje się trzymaniem indeksów
- Klasa `Cylinder` implementuje typ rekordu

Zasada działania

Łańcuch przepełnień

Łańcuch przepełnień działa na zasadzie podobnej do struktury `linked list`, gdzie w moim wypadku, wskaźnikami jest offset w pliku dodany o 1 (wartość 0 jest u mnie wartością specjalną - wskaźnik nie istnieje).

Insert

Gdy próbujemy umieścić nowy rekord w taśmie, najpierw przeszukujemy `index`. `Index` posiada w sobie informacje na której stronie zaczynają się poszczególne klucze, dlatego wystarczy że znajdziemy poprzednika od pierwszego większego znalezionej klucza od tego który chcemy wstawić. Mając stronę, nie musimy przeszukiwać całego pliku a tylko skoczyć do wybranej strony i odczytać ją. W niej szukamy poprzednika i umieszczamy go zaraz po poprzedniku. Jeśli nie ma miejsca w głównej taśmie, to umieszczamy go w łańcuchu przepełnień.

Reorganise

1. Tworzymy dwa tymczasowe pliki: dodatkową taśmę i indeks, ze wzoru niżej wyliczamy liczbę stron głównych, gdzie N, V - liczba rekordów w taśmie głównej i przepełnieniu, b - liczba rekordów danych na stronę, α - średnie zapełnienie strony po reorganizacji pliku.

$$\lceil \frac{N + V}{b * \alpha} \rceil \quad (1)$$

2. Przechodzimy kolejno przez rekordy zgodnie z rośnięciem kluczy i umieszczamy je na kolejnych stronach (respektując α)
3. Usuwamy stare pliki i zamieniamy tymczasowe na nie.

Prezentacja wyników programu

Po włączeniu programu użytkownikowi zostają pokazane wszystkie możliwości:

TODO: przekopiować output tutaj

Głównie są to 2 komendy:

- `insert <key> <base> <height>` - dodanie rekordu do bazy
- `file` - wczytanie komend z pliku (domyślnie plik z nazwą `input.txt`)
- `dump` - wypisanie całej bazy
- `reorganise` - reorganizacja całej bazy

Przykładowe wyjście z programu:

TODO: przykładowe wyjście

Eksperyment

Przeprowadzono eksperymenty na zasadzie wczytywania danych z pliku, dane użyte do testów są w folderze `tests`.