

Zaawansowane struktury danych - Projekt 3

26 stycznia 2024

Plan zajęć

- 1 Plan projektu
- 2 Wyjście programu
- 3 Termin

Drzewa trie i Patricia

Napisz drzewo trie i Patricia wyposażone w następujące funkcje:

- insert wstawiającą wyraz do drzewa
- find znajdującą wyraz w drzewie
- delete usuwającą wyraz w drzewa
- add dodającą wszystkie wyrazy z pliku do drzewa
- save zapisującą drzewo do pliku
- load wczytującą drzewo z pliku
- getNuberOfWords zwracającą liczbę słów w drzewie

teksty można pobrać ze strony: <https://datatracker.ietf.org/>

Program przygotowanie

Przed uruchomieniem programu należy przygotować następujące pliki:

- pobieramy kilka plików tekstowych RFC i wstawiamy je do obu drzew (te same pliki do jednego i drugiego drzewa) tak aby liczba słów w drzewie D_W (liści) osiągnęła następujące wartości: 1000, 2000, ...
- należy zapisać oba drzewa w postaci plików, tak aby zgromadzić pliki: D_W_1000.tri, D_W_2000.tri, ... i D_W_1000.pat, D_W_2000.pat, ...
- pobieramy kilka innych niż w punkcie pierwszym plików tekstowych RFC i łączymy je tworząc pliki wejściowe IN_W_10.txt, IN_W_100.txt, ... o określonej w nazwie pliku liczbie słów (słowa mogą się powtarzać)

Wyjście programu dla każdego drzewa

trie	10	100	...
1000W[1212B]	[100%, 0s]	[98%, 2s]	...
2000W[2480B]	[98%, 2s]	[87%, 4s]	...
...

- 1000W - liczba słów w słowniku
- 1212B - rozmiar słownika w bajtach
- 100% - pokrycie badanego ciągu słów
($\frac{\text{LICZBA_ZNALEZIONYCH_SLOW}}{\text{LICZBA_WSZYSTKICH_SLOW}}$)
- ,0s - czas wyszukiwania wszystkich słów w sekundach

Termin

Program należy przesłać do 4 lutego 2024 do godziny 12:00 na adres e-mail:

`marcin.zurowski@amu.edu.pl`

tytuł maila ZSD 11 s123456 - Projekt 3

gdzie s123456 jest numerem indeksu osoby studiującej.

Punktacja to 5 punktów za każde drzewo (max 10 pkt. więc jak ktoś zrobi 4 drzewa to zdobędzie 200% :-)).