

Referat – lista 7

Specyfikacja urządzenia:

Procesor	Intel(R) Core(TM) i5-9300HF CPU @ 2.40GHz 2.40 GHz
Zainstalowana pamięć RAM	16,0 GB

Przykładowe rezultaty:

Po kolei podawane wyniki dla metod _a, _d, _c

Result1.txt:

Zad 2 results in milliseconds:

Approximate time needed to accomplish the task: 3.32104206085205078125

Approximate time needed to accomplish the task: 0.03796529769897460938

Approximate time needed to accomplish the task: 1.32086896896362304688

Zad 3 results in milliseconds:

Approximate time needed to accomplish the task: 3.02398633956909179688

Approximate time needed to accomplish the task: 2.02387857437133789062

Approximate time needed to accomplish the task: 0.03601241111755371094

Zad 4 results in milliseconds:

Approximate time needed to accomplish the task: 1.51126050949096679688

Approximate time needed to accomplish the task: 0.01200532913208007812

Approximate time needed to accomplish the task: 0.00602412223815917969

Result4.txt:

Zad 2 results in milliseconds:

Approximate time needed to accomplish the task: 3.30406594276428222656

Approximate time needed to accomplish the task: 0.04101729393005371094

Approximate time needed to accomplish the task: 1.38609099388122558594

Zad 3 results in milliseconds:

Approximate time needed to accomplish the task: 3.05812120437622070312

Approximate time needed to accomplish the task: 1.99864625930786132812

Approximate time needed to accomplish the task: 0.02999019622802734375

Zad 4 results in milliseconds:

Approximate time needed to accomplish the task: 1.68015122413635253906

Approximate time needed to accomplish the task: 0.01096105575561523438

Approximate time needed to accomplish the task: 0.00200009346008300781

Wnioski:

W każdym przypadku najgorzej sprawiała lista krotek all_cases. Brak jakiegokolwiek uporządkowania uniemożliwia w większości przypadków wcześniejsze zakończenie przeszukiwania listy, gdyż:

- W przypadku zad 2 nie możemy wcześniej zakończyć działania, gdyż układ wczytanego pliku został najpierw posortowany przez nazwy państw, a dopiero potem przez daty
- W przypadku zad 3 nie wiemy, ile dni z przypadkami i zgonami zostało zarejestrowanych dla danego państwa w tym pliku, więc musimy sprawdzić wszystkie (ewentualnie wykorzystać

posortowanie przez państwa, tak, że kiedy znaleziony zostanie pierwszy rekord z poprawnym państwem, to przy pierwszym rekordzie z niepoprawnym państwem wyszukiwanie zostanie zakończone

- W przypadku zad 4 wiemy, że para data + państwo jest unikatowa, więc po znalezieniu można od razu zakończyć poszukiwania, bo na pewno nie będzie innego rekordu z tą samą datą i państwem na raz, w dalszym ciągu jednak, jeżeli rekord znalazł się na końcu pliku, to trzeba będzie przejść przez prawie cały plik

Słownik `by_date` jest dość skuteczny. Uporządkowanie po dacie pomaga we wcześniejszym zakończeniu poszukiwań po słowniku:

- W przypadku zad 2 po prostu sumujemy wszystkie potrzebne nam elementy listy krotek, na którą wskazują nam argumenty, bardzo proste
- W przypadku zad 3 w każdej liście krotek powiązanej z datą szukamy rekordu z pasującym krajem i po znalezieniu odpowiedniego rekordu w liście zatrzymujemy szukanie w tej dacie i idziemy do następnej, prędkość zależy od tego które w kolejności wypada podane państwo, zawsze szybsze niż `all_cases` (chyba, że wprowadzić modyfikację)
- W zad 4 od razu mamy datę i tylko szukamy pasującego państwa

Słownik `by_country` jest dość skuteczny. Uporządkowanie po państwie pomaga we wcześniejszym zakończeniu poszukiwań po słowniku:

- W przypadku zad 2 w każdym państwie szukamy interesującej nas daty i sumujemy odpowiednie elementy, po znalezieniu daty w jednym państwie, zatrzymujemy poszukiwanie i szukamy w kolejnym, prędkość zależy w której kolejności jest data
- W przypadku zad 3 sumujemy to co trzeba z danego państwa, które łatwo znajdujemy, gdyż jest to słownik
- Zad 4 od razu mamy państwo i szukamy pasującej daty

Najczęściej zad 4 jest wykonywane szybciej przez `by_country` niż przez `by_date`, prawdopodobnym powodem jest to, że elementy daty zapisywane są w postaci liczb całkowitych, a nazwa państwa jako łańcuch znaków, przez co mimo że dat jest więcej niż państw, to porównywania państw może być bardziej czasochłonne

Zdaje się, że gdyby spośród tych trzech trzeba było podać najlepszy sposób przechowywania danych by móc je znaleźć (na sposoby podane w zadaniach), to najlepszy jest `by_country`. Jednakże najlepiej byłoby mieć też słownik `by_date` i adekwatnie do sytuacji korzystać tego bardziej skutecznego.