

Egzamin na magisterskie studia uzupełniające (2005 r.)

1. **[Programowanie]** Napisać program w języku ANSI C, który dla podanego pliku tekstowego wypisuje na standardowym wyjściu 10 najczęściej występujących w nim słów. Dodatkowo przy każdym wypisanym słowie należy podać w procentach stosunek liczby jego wystąpień do liczby wystąpień wszystkich słów w pliku. Słowem nazywamy tu maksymalny ciąg kolejnych znaków w pliku nie zawierający znaków odstępu, tabulacji i zmiany wiersza. Program powinien działać poprawnie przy następujących założeniach: wielkość pliku do $1GB$ i długość słowa nie większa 1000 znaków. Pozostałe wymagania są następujące:

- Słownik przeczytanych słów powinien być trzymany w strukturze o dynamicznie zmieniającej się wielkości; struktura powinna umożliwiać szybkie wyszukiwanie i wstawianie nowych słów.
- Nazwa pliku powinna być podawana jako argument wywołania programu; program powinien sensownie reagować przy złym wywołaniu.

Punktacja: **30 % punktów** za deklarację odpowiedniej struktury słownika oraz za procedury wyszukiwania i wstawiania słów; **30 % punktów** za procedurę czytania pliku i zliczania liczb wystąpień; **20 % punktów** za procedurę wyznaczania 10 najczęściej występujących słów oraz **20 % punktów** za program główny i wypisywanie wyników.

2. **[Matematyka dyskretna]** Niech G będzie grafem, a T - jego drzewem spinającym. Niech

$$\{e_1, e_2, \dots, e_k\}$$

będzie zbiorem wszystkich krawędzi G , nienależących do T . Niech C_i będzie cyklem otrzymanym przez dodanie e_i do T . Pokaż, że dowolny cykl C zawarty w G jest różnicą symetryczną cykli C_i , dla których e_i należy do C .

3. **[Algorytmy i struktury danych]** Dany jest graf nieskierowany $G = (V, E, d)$ z określonymi dodatnimi wagami krawędzi (funkcja $d : E \mapsto \mathcal{R}^+$). Należy znaleźć długości najkrótszych ścieżek pomiędzy wszystkimi parami wierzchołków, czyli wyznaczyć dwuwymiarową tablicę S , która na pozycji $S[u, v]$ dla $u, v \in V$ zawiera długość najkrótszej ścieżki w grafie G pomiędzy tymi wierzchołkami.

- (a) Podaj algorytm dynamiczny pozwalający wyznaczyć tablicę S (**20% punktów**).

Oszacuj złożoność algorytmu i uzasadnij jego poprawność (**20% punktów**).

Zmodyfikuj (jeśli to konieczne) algorytm tak, by pozwalał także dla zadanych u i v wypisać wierzchołki tworzące najkrótszą ścieżkę (**10% punktów**).

- (b) Podaj zarys algorytmu pozwalającego obliczyć tablicę S opartego na przeglądaniu grafu (**20% punktów**).

Przeanalizuj, jakich struktur danych można użyć w tym algorytmie i jak zależy od nich jego złożoność (**30% punktów**).