

**Lista 10 Termin oddania: czwartek: 20.05.2021r.****Zadanie (4p)**

Napisz program, który odczytuje plik tekstowy, następnie oblicza częstotliwość występowania poszczególnych znaków w odczytanym tekście. Na tej bazie utwórz drzewo Huffmana. Kod Huffmana wykorzystaj do zakodowania tekstu z pliku, zakodowany tekst zapisz w pliku tekstowym. Sposób kodowania poszczególnych znaków wraz z częstością ich występowania wyświetl na ekranie.

Do prezentacji zadania na zajęciach przygotuj pliki tekstowe: z dowolnym tekstem oraz z tekstem przykładowym podanym poniżej.

Wskazówka: Każdy węzeł drzewa Huffmana zawiera dwa pola danych: wagę i znak. Waga węzła będącego liściem określa częstotliwość występowania znaku, a waga węzła wewnętrznego jest sumą wag obu poddrzew danego węzła.

Do konstrukcji drzewa Huffmana wykorzystaj kolejkę priorytetową.

Algorytm:

1. Utwórz zbiór drzew z korzeniami zawierającymi poszczególne znaki i ich wagi.
2. Umieść zbiór drzew w kolejce priorytetowej.
3. while kolejka priorytetowa zawiera więcej niż jeden element
  - a. usuń dwa drzewa o najmniejszych wagach
  - b. połącz je w jedno drzewo binarne, w którym waga korzenia jest sumą wag jego dzieci
  - c. wstaw utworzone w ten sposób drzewo z powrotem do kolejki priorytetowej

Aby określić kod znaku, tworzy się ciąg binarny, przechodząc od korzenia do liścia z wybranym znakiem. Za każdym razem, gdy wybiera się lewą stronę, dodaje się 0, a prawą stronę, to dodaje się 1. Zapamiętana sekwencja zer i jedynek stanowi kod znaku.

**Przykład:**

Tekst: **ala i ola.**

Zakodowany tekst: **1000101111101110110010**

a	- 3	10
l	- 2	00
spacja	- 2	111
i	- 1	110
o	- 1	011
.	- 1	010

Uwaga: Kody znaków powinny być jak najkrótsze, w tym przykładzie po 3 znaki. Jeśli dla tego przykładu kod będzie 4-znakowy, tzn. algorytm nie jest optymalny i za to będzie odjęty 1 punkt.

**Zadanie za 1 punkt dodatkowy.**

Napisz metodę odszyfrowania tego tekstu, wykorzystując utworzone drzewo