

Vaje pri predmetu Programiranje II

Teden 7: Histogram besed

Histogram besed

Naloga

Napišite program, ki za podano besedilo izpiše n najbolj pogostih besed. Besede morajo biti izpisane v padajočem vrstnem redu glede na število pojavitev, zraven besede izpišite njeno frekvenco. Če se več besed pojavi v besedilu enako pogosto, jih uredite po abecedi.

Nalogo bomo reševali po korakih. Imeli bomo tri (pod)naloge. Za vse naloge imate ločene teste.

Naloga1

Preberite besede (uporabite `%s` pri `scanf`) in jih izpišite. Pri tem upoštevajte:

- Vsa ločila je potrebno izločiti (uporabite funkcijo `ispunct`). Ena beseda ima lahko tudi več ločil.
- Velike in male črke se obravnavajo enako. Tako se v frazi "Gori na gori gori" beseda "gori" pojavi 3x, čeprav je 1x zapisana z % veliko začetnico in 2x z malo.
- Lahko predpostavite, da bodo v besedilih le ASCII znaki.

Naloga2

Isto kot naloga1, vendar se vsako besedo izpiše le enkrat. Če se beseda v tekstu ponovi, je ne izpisujte.

Naloga3

Rešite celotno nalogo. Namig: uporabite še eno tabelo za štetje besed, ki bo imela enako število elementov, kot jih ima tabela z besedami.

Vhod

V vhodnih podatkih bo najprej napisano zahtevano število najbolj pogostih besed, sledi besedilo. Vhodni podatki se nahajajo v direktoriju `tests`.

Testni vhod:

```
5
Levo, pa desno, pa spet levo in potem naravnost do semaforja, kjer zavijes spet
levo. Pa pazi na pesce!
```

Vhodne besede ne bodo daljše od 10^5 znakov. Besedilo je lahko poljubne dolžine.

Izhod

Izhodne datoteke za posamezne naloge se nahajajo v direktorijih out1-3.

Naloga1

Pričakovani izhod:

```
levo
pa
desno
pa
spet
levo
in
potem
naravnost
do
semaforja
kjer
zavijes
spet
levo
pa
pazi
na
pesce
```

Naloga2

Pričakovani izhod:

```
levo
pa
desno
spet
in
potem
naravnost
do
semaforja
kjer
zavijes
pazi
na
pesce
```

Naloga3

Vsako najbolj pogosto besedo skupaj s frekvenco izpišite v svoji vrstici, ki se mora zaključiti z izhodnim znakom za novo vrstico '\n'.

Pričakovani izhod:

```
levo 3
pa 3
spet 2
desno 1
do 1
```