

Lekcija 9 – Genetski kod sa fiksnim pozicijama

Genetski kod DNK sastavljen je od nukleotida (Adenin, Guanin, Timin i Citozin) koje predstavljamo slovima A, C, T i G. Pored toga definisamo genetsku sekvencu kao niz od tri nukleotida (bilo koja). Npr. "CTA" ili "GGC" su primjeri genetskih sekvenci.

Napisati program koji iz datoteke "genetski_kod.txt" učitava opis genetskog koda te pronalazi i na ekran ispisuje genetsku sekvencu koja se uzastopno ponavlja najviše puta. . Npr. ako datoteka glasi:

TGATGACGCC**CGGCGGCGG**CAGTGATGA

Najviše puta uzastopno se ponavlja sekvenca "CGG" (tri puta – označeno boldom), dok se sekvenca "TGA" ponavlja po dva puta uzastopno (na početku i kraju datoteke). Datoteka se sastoji od niza genetskih sekvenci - sekvenca počinje na svakom trećem mjestu tako da npr. ne možete posmatrati "GAT" ili "ATG" kao sekvence u primjeru datom iznad.

U slučaju da format datoteke ne odgovara datom iz bilo kojeg razloga, na ekranu treba ispisati poruku "Pogresan format" i prekinuti rad.

Napomena: U zadatku je strogo zabranjeno koristiti bilo kakve pomoćne nizove ili stringove! Rješenje sa pomoćnim nizovima ili stringovima će dobiti 0 bodova. Dozvoljeno je koristiti pojedinačne promjenljive tipa int ili char.