

# Algoritmi 2019/20

## Domača naloga 2

Nalogo morate natipkati in rešitve oddati v pdf obliki na učilnico.

**Problem 1.** *Ujemanje vzorcev.* Rešite nalogo A) in B) ali A) in C). V prvem primeru dobite do 7 točk, v drugem pa do 9 točk.

A) Izračunajte KMP preponsko funkcijo  $\pi$  za naslednji vzorec:

ababbabbabbabababbabb .

3 točk(e)

B) Trije zaporedni nukleotidi (primer: AUU) v genomu določijo katera amino kislina se bo tvorila v proteinu. Takšno trojico imenujemo *kodon*. Zaporedje nukleotidov torej določa zaporedje amino kislin v proteinu. Iskali boste ali je v genomu zaporedje , ki pripada določeni amino kislini.

- Izračunajte prehodno funkcijo  $\delta$  (tabelarično) za končni avtomat, ki bo iskal metionin (AUG) in triptofan (UGG). To pomeni, da iščemo končni avtomat, ki bo v genomu poiskal vse pojavitve vzorca AUGUGG. 2 točk(e)
- V prejšnem delu je bila amino kislina unikatno določena z enim zaporedjem kodonov. Včasih amino kislino določa tudi več različnih zaporedji kodonov. Izračunajte prehodno funkcijo  $\delta$  (tabelarično) za končni avtomat, ki bo iskal zaporedje asparagina (AAU, AAC) in metionina (AUG). To pomeni, da iščemo končni avtomat, ki bo v genomu poiskal vse pojavitve vzorcev AAUAUG in AACAUG. 2 točk(e)

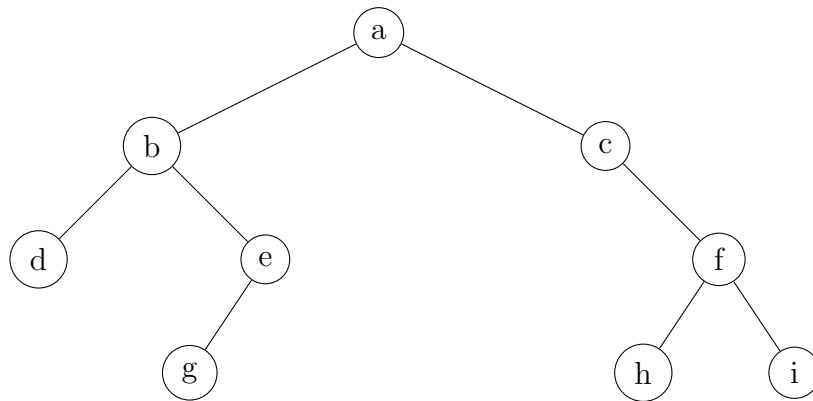
- C) Glej točko B) za podrobnejša navodila. Sestavite končni avtomat za iskanje asparagina in metionina in napišite program, ki po prebral genom in s pomočjo avtomata vrnil lokacije vseh pojavitev zelenih amino kislin. Oddajte izvorno kodo in delujoč primer. 6 točk(e)

**Problem 2.** *IP posredovanje in vEB drevesa.*

- A) Uporabite teorem Master za izračun časovne zahtevnosti operacije **vEB-Tree-Successor** vEB drevesa. Uporabite teorem Master in pokažite, da je časovna zahtevnost operacije **vEB-Tree-Successor** vEB drevesa  $O(\lg \lg M)$ , kjer je  $M$  velikost univerzuma. Predpostavite, da je  $M$  oblike  $2^k$ . Utemeljite svoj odgovor. 3 točk(e)
- B) Naivni pristop pri šteju bitov s pomočjo povzetkovne tabele z Luela algoritmom potrebuje tri pomnilniške reference. Razložite, kako lahko prva dva dostopa združimo v enega. 3 točk(e)
- C) Sestavite fixed stride številsko drevo s stride velikostjo 2 za omrežja **100\***, **01\***, **001\***, **11\***, **1011\***, and **1\***. 2 točk(e)

**Problem 3.** *Kompaktne podatkovne strukture*

Podano imate spodnje drevo:



- A) Predpostavite, da je drevo na sliki ordinalno in ga zapišite v obliki BP in LOUDS. 2 točk(e)

- B) Predpostavite, da je drevo na sliki kardinalno in ga zapišite v obliki BP in LOUDS.

*2 točk(e)*

- C) Za kardinalno drevo v obliki BP zgradite rmM-drevo za  $b = 8$ .

*2 točk(e)*

- D) Napišite pseudo kodo za funkcijo Izberi( $T, i$ ), ki vrne  $i$ -ti element v razširjenem iskalnem dvojiškem drevesu  $T$ . Drevo  $T$  v vsakem vozlišču hrani moč levega poddrevesa  $|L|$ .

*2 točk(e)*