

# Algoritmi in podatkovne strukture – 2

## Drugi kolokvij (2014/15)

Kolokvij morate pisati posamič. Pri reševanju je literatura dovoljena. Pri odgovarjanju bodi natančni in: (i) odgovarjajte *na zastavljena* vprašanja; in (ii) odgovorite na *vsa* zastavljena vprašanja.

Čas pisanja izpita je 75 minut.

Veliko uspeha!

NALOGA	TOČK	OD TOČK	NALOGA	TOČK	OD TOČK
1			3		
2			4		

IME IN PRIIMEK: \_\_\_\_\_

ŠTUDENTSKA ŠTEVILKA: \_\_\_\_\_

DATUM: \_\_\_\_\_

PODPIS: \_\_\_\_\_

**1. naloga:** Iskanje vzorca.

## VPRAŠANJA:

1. Pri opisu algoritmov za iskanje vzorca v besedilu smo uporabljali pojem koncept končnega avtomata. (i.) Za vzorec TCGAC sestavite iskalni končni avtomat. (ii.) Ali je vaš avtomat determinističen ali nedeterminističen? Utemeljite odgovor.
2. Vzorci v bioloških zaporedjih niso vedno enolično določeni. Tako je lahko na posameznem mestu v vzorcu možnih več različnih znakov. Primer:  

0	1	2	3	4	5	6	
T	C/G	G	A/G	G	A/C/G/T	C	sta na mestu s možna

znaka bodisi C ali G in podobno naprej. (i.) Peter Zmeda je dobil idejo za definicijo nedeterminističnega končnega avtomata za iskanje tovrstnih vzorcev. Narišite ustrezni končni avtomat, pri čemer jasno označite njegova stanja in prehode. (ii.) Kakšna je časovna zahtevnost iskanja vzorca v besedilu dolžine  $n$  z vašim nedeterminističnim končnim avtomatom? Utemeljite odgovor.
3. V bioinformatiki imamo pogosto opravka s problemom iskanja podobnosti med dvema vzorcema. Eden od načinov je z uporabo dinamičnega programiranja, ki smo ga že spoznali (razdalja urejanja in njemu sorodna algoritma za lokalno ter globalno poravnavo). Težava tega algoritma je, da ima časovno zahtevnost  $O(n^2)$ , kar je v praksi pogosto neuporabno, saj sta vzorec in/ali besedilo lahko dolga več milijard znakov. Zato potrebujemo hitrejši algoritem. Ena od možnosti je izračun točkovne matrike, kot smo jo spoznali v eni domačih nalog. Opišite učinkovito podatkovno strukturo in algoritem, ki pri danem vzorcu  $p$  in besedilu  $t$  poišče najpopularnejšo diagonalo. Kakšna je časovna in prostorska zahtevnost tvoje rešitve? Utemeljite odgovor.

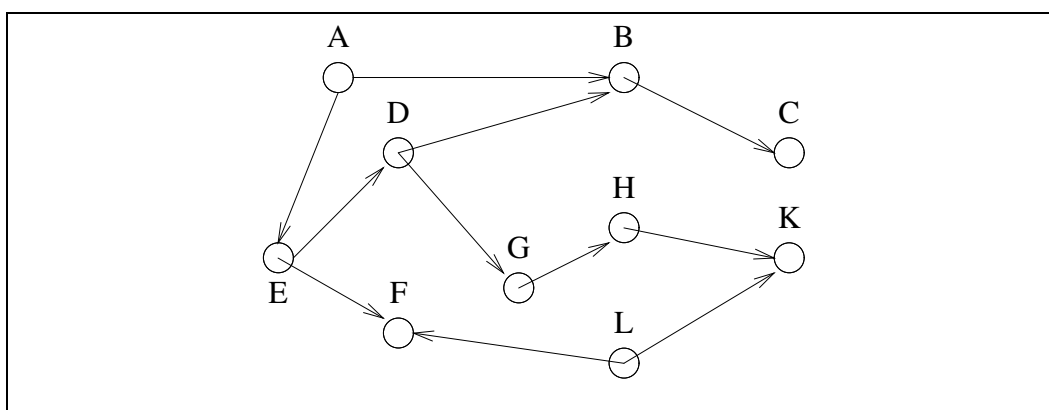
**2. naloga:** Pri Petru Zmedi v službi so dobili nove pametne telefone, ki bi jih želeli nadgraditi tako, da bi znali sami dopolnjevati imena iz imenika, ko bi jih uporabnik vpisoval. Na primer, naj bosta v imeniku dve imeni, ki se pričneta na črko M: MATEJ in MARKO. Ko bi uporabnik vpisal M, se bi samodejno izpisalo MA(2), kjer številka 2 pomeni, da sta v imeniku dve imeni, ki se pričneta na MA. Če nato vpiše še T, se bi izpisalo MATEJ.

## VPRAŠANJA:

1. i.) Opišite podatkovno strukturo, ki naj jo uporabi Peter, da bo učinkovito rešil nalogo. ii.) Opišite postopek vstavljanja v podatkovno strukturo.

2. Petrov predstojnik se je ravno vrnil z izpopolnjevanja, kjer je izvedel, da so razpršilne tabele zelo učinkovita podatkovna struktura za hranjenje podatkov ter je zato zahteval od Petra, da jih uporabi v svoji rešitvi (navkljub predlogu iz prvega vprašanja). Kako naj Peter uporabi razpršilno tabelo za vseeno še relativno učinkovito rešitev.
3. Predlagana aplikacija na telefonu je bila Petru všeč, le motilo ga je, da se je izpisala samo številka. Zato se je odločil, da aplikacijo nadgradi tako, da, če se uporabnik dotakne številke (v zgornjem primeru MA(2) številke 2), se mu v meniju ponudijo obe imeni ter bi lahko izbral želeno ime. Seveda imen je lahko zelo veliko, a s tem se Peter še ne ukvarja. Predlagajte nadgradnjo vaše podatkovne strukture iz prvega vprašanja, da bo učinkovito pomagala Petru nadgraditi aplikacijo.

**3. naloga:** V tej nalogi se bomo poukvarjali s čim hitrejšim zaključkom dela, ki je razdeljeno na več opravil. Na sl. 1 vsako vozlišče predstavlja eno opravilo in



**Slika 1:** Primer grafa.

povezave prikazujejo katero opravilo mora biti prej zaključeno (vozlišče, iz katerega izhaja puščica) predno se lahko naslednje opravilo izvede (vozlišče, kamor kaže puščica). V splošnem imamo graf  $G(V, E)$  ( $|V| = n, |E| = m$ ).

**VPRAŠANJA:**

1. Katero opravilo mora biti zaključeno prvo in katero je lahko zaključeno zadnje v primeru na sl. 1? Utemeljite odgovor.
2. Predpostavimo, da vsako opravilo zahteva enako časa, da se ga izvede na enem procesorju. Recimo, da imamo na voljo en sam procesor in nas zanima, koliko časa potrebujemo, da opravimo vse delo. Podajte algoritem in njegovo časovno zahtevnost. Utemeljite odgovor.

3. Na koncu predpostavimo, da imamo poljubno število procesorjev, ki lahko hkrati izvajajo opravila in nas zanima, koliko najmanj časa potrebujemo sedaj pri poljubnem grafu  $G(V, E)$ , da opravimo vse delo? Podajte algoritem in njegovo časovno zahtevnost. Utemeljite odgovor.

**4. naloga:** Fiziki so res čudni tiči. Ti, s katerimi ima tokrat opravka naš Peter Zmeda so posebej čudni. Namreč še med seboj si ne zaupajo. Tako je prof. Bučka odkril neko novo obliko letalskega krila, vendar noče nikakor izdati njegove oblike, a kljub temu želi da mu krilo prebarvajo na ksanadu barvo. V barvarni bi radi naročili barvo, ki pa ni poceni ter zato ne želijo naročiti preveč barve, da bi ostajala. Zaposili so za pomoč Petra.

VPRAŠANJA:

1. Peter je od dr. Bučke želel zgolj implementacijo naslednjega modula:

```
interface Krilo {
    // skrajni x-koordinati krila
    float Levo();
    float Desno();
    // skrajni y-koordinati krila
    float Zgoraj();
    float Spodaj();
    // ali je točka (x, y) na krilu
    boolean NaKrilu(float x, float y);
}
```

Pri tem je predpostavka, da je krilo položeno plosko na tla. Kaj menite, kakšen algoritem si je zamislil Peter za izračun površine krila? Opišite postopek vključno z zapisom algoritma.

2. Na predavanjih smo spoznali problem iskanja najcenejšega vpetega drevesa v grafu  $G(V, E)$ . To je optimizacijski problem. (i.) Preoblikujte ga v odločitveni problem. (ii.) Ali je vaš odločitveni problem v NP? Utemeljite odgovor.
3. Eden od zanimivih neuporabnih algoritmov je *Bogosort*, katerega psevdokoda je:

```
WHILE NOT urejeno(polje):
    NakljucnoPermutiraj(polje)
```

- (i.) Ali gre za Monte Carlo ali za Las Vegas algoritem? Utemeljite odgovor. (ii.) Kakšna je: (a) najboljša, (b) najslabša in kakšna (c) pričakovana časovna zahtevnost algoritma? Utemeljite odgovor.