

## Tehtävä 2: Loppuosataulukko - Vastaukset

### Kysymys 1.

1. Alkuarvot  $alku = 1$  ja  $loppu = 16$ .
2.  $keski = (1 + 16)/2 = 9$ ,  $p > t[S[9]..16]$ ,  $alku = keski + 1 = 10$ .
3.  $keski = (10 + 16)/2 = 13$ ,  $p < t[S[13]..16]$ ,  $loppu = keski - 1 = 12$ .
4.  $keski = (10 + 12)/2 = 11$ ,  $p < t[S[11]..16]$ ,  $loppu = keski - 1 = 10$ .
5.  $keski = (10 + 10)/2 = 10$ ,  $p > t[S[10]..16]$ ,  $alku = keski + 1 = 11$ .
6. Nyt  $alku > loppu$ , joten haku päättyi tuloksettomana. Hahmo  $p =$  "babaa" ei esiinny tekstissä.

### Kysymys 2.

Loppuosa	$i$	$S[i]$
a	1	13
alinta	2	8
eisvalinta	3	4
hteisvalinta	4	2
inta	5	10
isvalinta	6	5
linta	7	9
nta	8	11
svalinta	9	6
ta	10	12
teisvalinta	11	3
valinta	12	7
yhteisvalinta	13	1

Vastauksessa riittää, että arvoja  $S[i]$  kuvaava sarake on annettu.

Kysymys 3.

a) Loppuosa	$i$	$S[i]$
aacat	1	16
aacatcgatagctagaacat	2	1
acat	3	17
acatcgatagctagaacat	4	2
agaacat	5	14
agctagaacat	6	10
at	7	19
atagctagaacat	8	8
atcgatagctagaacat	9	4
cat	10	18
catcgatagctagaacat	11	3
cgatagctagaacat	12	6
ctagaacat	13	12
gaacat	14	15
gatagctagaacat	15	7
gctagaacat	16	11
t	17	20
tagaacat	18	13
tagctagaacat	19	9
tcgatagctagaacat	20	5

Vastauksessa riittää, että arvoja  $S[i]$  kuvaava sarake on annettu.

- b)
1. Alkuarvot  $alku = 1$  ja  $loppu = 20$ .
  2.  $keski = (1 + 20)/2 = 11$ ,  $p > t[S[11]..20]$ ,  $alku = keski + 1 = 12$ .
  3.  $keski = (12 + 20)/2 = 16$ ,  $p < t[S[16]..20]$ ,  $loppu = keski - 1 = 15$ .
  4.  $keski = (12 + 15)/2 = 14$ ,  $p < t[S[14]..20]$ ,  $loppu = keski - 1 = 13$ .
  5.  $keski = (12 + 13)/2 = 13$ ,  $p < t[S[13]..20]$ ,  $loppu = keski - 1 = 12$ .
  6.  $keski = (12 + 12)/2 = 12$ ,  $p$  täsmää loppuosan  $t[S[12]..20]$  kanssa.  
Hahmo  $p = \text{“cga”}$  löydettiin alkaen tekstin kohdasta  $S[12] = 6$ .

**Kysymys 4.**

Tehtävän avainkohta lienee havaita, että annetun loppuosataulukon arvot ovat laskevassa järjestyksessä eli mitä aikaisemmaksi loppuosa sijoittuu aakkosjärjestyksessä, sitä suurempi sen indeksi on. Tai toisin päin: mitä pienempi loppuosan indeksi on, sitä myöhempänä se on aakkosjärjestyksessä. Tämän ominaisuuden sekä tehtävän muut vaatimukset täyttäväksi merkkijonoksi käy mikä tahansa suomalaisen aakkoston merkeistä koostuva 8 merkin pituinen merkkijono  $t$ , jossa mikään merkki ei ole aakkosjärjestyksessä suurempi kuin sitä edeltävä merkki. Tämän voi ilmaista täsmällisemmin esimerkiksi niin, että kaikilla  $i = 2, \dots, 8$  pitää päteä  $t[i - 1] \geq t[i]$ . Tämä on sekä riittävä että välttämätön ehto minkä tahansa tehtävässä sallituista merkeistä koostuvan vastauksen laillisuudelle.

Vastaukseksi kelpaa siten esimerkiksi mikä tahansa merkkijonoista “hgfedcba”, “öäwwnkkk” ja “mmmmmmmm”.

**Kysymys 5.**

Tehtävän ratkaisu on yksikäsitteinen:  $t = \text{“bbaababbababaaba”}$ . Tämä merkkijono voidaan muodostaa esimerkiksi seuraavanlaisella päättelyketjulla:

1. Jos merkkijono olisi  $t = \text{“bbbbbbbbbbbbbbbb”}$ , olisivat loppuosataulukon arvot, kysymyksen 4 tapaan, laskevassa järjestyksessä. Koska tämä ei päde, on merkkijonossa oltava ainakin yksi merkki “a”.
2. Päätelmän 1 perusteella aakkosjärjestyksessä ensimmäisen loppuosan ensimmäisen merkin on oltava  $t[S[1]] = t[16] = \text{“a”}$ .
3. Jos pätesi  $t[15] = \text{“a”}$ , olisi loppuosa  $t[15..16] = \text{“aa”}$  aakkosjärjestyksessä toisena. Näin ei kuitenkaan ole, sillä  $S[2] \neq 15$ , joten  $t[15] = \text{“b”}$ .
4. Kun ainoastaan merkit “a” ja “b” ovat sallittuja, olisi yhden merkin pituinen merkkijono “b” ainoa merkillä “b” alkava loppuosa, joka on aakkosjärjestyksessä aikaisemmin kuin loppuosa  $t[15..16] = \text{“ba”}$ . Koska jo tiedetään, että  $t[16] = \text{“a”}$  eli merkkijono  $t$  ei sisällä loppuosaa “b”, on jokaisen loppuosaa  $t[15..16] = \text{“ba”}$  aakkosjärjestyksessä aikaisemman loppuosan ensimmäisen merkin oltava “a”. Koska  $S[9] = 15$  eli “ba” on aakkosjärjestyksessä 9:s loppuosa, sitä pienemmillä indekseillä  $i = 1, \dots, 8$  pätee  $t[S[i]] = \text{“a”}$ . Nyt saatiin täydennettyä merkkijonoon merkit  $t[13] = \text{“a”}$ ,  $t[3] = \text{“a”}$ ,  $t[14] = \text{“a”}$ ,  $t[11] = \text{“a”}$ ,  $t[9] = \text{“a”}$ ,  $t[4] = \text{“a”}$  ja  $t[6] = \text{“a”}$ .
5. Toisaalta jokaisen loppuosan, joka on aakkosjärjestyksessä loppuosan  $t[15..16] = \text{“ba”}$  jälkeen, on alettava merkillä “b”. Siten indekseillä  $i = 10, \dots, 16$  pätee  $t[S[i]] = \text{“b”}$ . Nyt saatiin asetettua loput puuttuvat merkit  $t[12] = \text{“b”}$ ,  $t[2] = \text{“b”}$ ,  $t[10] = \text{“b”}$ ,  $t[8] = \text{“b”}$ ,  $t[5] = \text{“b”}$ ,  $t[1] = \text{“b”}$  ja  $t[7] = \text{“b”}$ . Lopputuloksena syntyi merkkijono  $t = \text{“bbaababbababaaba”}$ .