

KESKI-SUOMEN MAAKUNNAN JA LÄHIKUNTIEN LUKIOIDEN TIETOTEKNIIKAN II KILPAILU

KESKI-SUOMEN MAAKUNNAN JA LÄHIKUNTIEN LUKIOIDEN FYSIKAN, KEMIAN JA MATEMATIIKAN XXIII KILPAILUN OSANA

23.1.2014

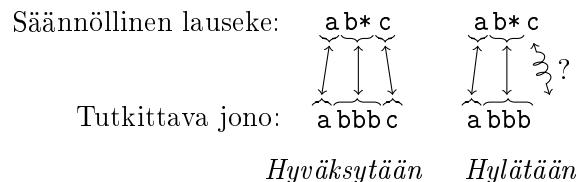
Huom: Tehtävä 2 (ohjelmointitehtävä) pohjautuu Tehtävän 1 kysymykseen 6. Kannattaa arvioida mihin kysymyksiin ehtii vastaamaan rajoitetun kilpailujan puitteissa.

Tehtävä 1: Säännölliset lausekkeet

Teoriaa

Kun tietokoneohjelmalla luetaan käyttäjän syötettä, olisi syöte aina syytä tarkistaa. Syötteessä voi olla vääriä merkkejä tai merkkejä väärillä paikoilla (*syntaktinen* vika), tai se voi olla sisällöllisesti mieletön (*semanttisesti* väärin, esim. päivämäärä 32.2.2013). Merkkijonon (jatkossa käytetään pelkkää sana *jono*) syntaktisessa oikeellisuustarkistuksessa voidaan käyttää säännöllisiä lausekkeita (englanniksi *regular expression*).

Säännöllisissä lausekkeissa tietyllä merkinnällä voidaan mm. sanoa, että joku merkki toistuu nolla tai useampia kertoja. Esimerkiksi tarkastellaan säännöllistä lauseketta ja kahta tutkittavaa jonoa:



Hyväksytään-esimerkissä säännöllisen lausekkeen alusta löytyy *a*, joka vastaa jonon *a*:ta. Lausekkeen *b** vastaa jonoa *bbb*, koska *** monistaa edellisen merkin 0-n kertaa. Lopussa molemmissa on *c*. Kaikille jonon merkeille löytyy siis vastine säännöllisestä lausekkeesta. Esimerkin säännöllinen lauseke hyväksyisi myös jonot *ac*, *abc*, *abbc* ja vaikkapa *abbbbbbbbbbbbc*.

Hylätään-esimerkissä on muuten samat vastaavuudet, mutta lausekkeen *c* ei vastaa mitään tutkittavan jonon merkkiä, jolloin lauseke ei siis hyväksy jonoa. Jos lauseke olisi *ab**, tutkittavaa jonoa *abbbc* ei hyväksyttäisi, koska tällöin jonon *c*:lle ei olisi vastaava merkkiä lausekkeessa.

Useimmissa ohjelmointikielissä on valmiit kirjastot säännöllisiä lausekkeita varten. Oletetaan, että käyttäjän syöte (esimerkiksi WWW-lomakkeelta) olisi luettu merkkijonomuuttujaan *jono*. Kellonajan tarkistaminen muotoon *hh:mm* voitaisiin tehdä esimerkiksi eräällä ohjelmointikielillä seuraavalla ohjelmakoodilla:

```
if ( Pattern.matches("[0-2][0-9]:[0-5][0-9]", jono) )
    ... jono oikein ...
```

Tämä tarkistus jättää vielä toki paljon mahdollisuuksia syötön sisällöllisille virheille. Palataan niihin myöhemmin.

Taulukossa 1 annetaan osajoukko eräistä säännöllisten lausekkeiden merkinnöistä. Esimerkikilausekkeen jälkeen on esimerkkejä jonoista, jotka lauseke hyväksyy ja esimerkkejä jonoista, jotka hylätään.

Merkintä	Selitys ja esimerkki	Hyväksyy mm. jonot	Hylkää mm. jonot
.	Mikä tahansa merkki. ki.a	kiva, kisa	kuva
*	Edellinen 0 kertaa tai miten monta kertaa tahansa. ab*a k.*a	aa, aba, abba kissa, koira, kuva	abca, abc kettu
+	Edellinen vähintään yhden kerran. ab+a	aba, abba, abbba	aca, aa
?	Edellinen 0 tai 1 kertaa. ab?a	aa, aba	abba
[lista]	Hakasuluissa voidaan antaa lista merkeistä, joista vain yksi käytetään. Tässä lista tai sen osa voidaan ilmoittaa myös muodossa alkumerkki-loppumerkki. a[bf]a a[a-ek]a a[a-e]*a	aba, afa aba, aea, aka aa, aaa, aea, aaba, abbeba	abfa, ada afa, aeka afa
{n}	Aaltosuluilla voidaan määrittää kerroin, jolla edellinen tulee tasan n kertaa. ab{2}a	abba	aba
(ryhmä)	Kaarisuluilla voidaan yhdistää merkkejä ryhmäksi, johon mm. *, + tai kerroin vaikuttaa. a(ab)*c a(ab)+c a(ab){2}c	ac, aabc, aababc aabc, aababc aababc	aac, adc, abac ac, aac, adc, abac ac, aabc, ababac
	Jompi kumpi vaihtoehdoista. a(bb k)a	abba, aka	akka, abka

Taulukko 1: Säännöllisten lauseiden merkintöjä ja esimerkkejä.

Esimerkiksi 2000-luvulla syntyneen suomalaisen henkilötunnuksen merkkien oikeellisuus voitaisiin tarkistaa lausekkeella

$$[0-9]\{6\}A[0-9]\{3\}[0-9A-FHJ-NPR-Y]$$

Kyseisen säännöllisen lausekkeen toiminta voidaan selittää tarkastelemalla sen osien toimintaa seuraavasti:

$[0-9]\{6\}$	Numeroita väliltä 0–9 kuusi kappaletta,
A	merkki A (koska 2000-luvulla syntyneillä erotinmerkki on A),
$[0-9]\{3\}$	numeroita väliltä 0–9 kolme kappaletta,
$[0-9A-FHJ-NPR-Y]$	ja lopuksi hetun tarkistusmerkiksi jokin numeroista 0–9 tai kirjaimista A–Y, paitsi ei G, I, O tai Q (koska olisivat käsin kirjoitettuna sotkettavissa numeroihin 6, 1 tai 0).

Esimerkki hyväksyttävästä jonosta voisi olla 230506A123B. Toisaalta jono 230506-123B hylättäisiin, koska 6. numeron jälkeen ei ole A-kirjainta. Jono 230506A123 hylättäisiin, koska kolmen loppunumeron jälkeen ei ole vaadittua tarkistusmerkkiä. Vastaavasti 230506A123I hylättäisiin, koska kirjain I ei ole sallittujen tarkistusmerkkien joukossa.

Tämän syntaktisen tarkistuksen lisäksi on toki ohjelmallisesti tarkistettava, että päiväysosan kuusi numeroa muodostavat semanttisesti järkevän päiväyksen sekä tarkistusmerkki noudattaa sovittua muista merkeistä laskettavaa sääntöä. Tähän emme kuitenkaan puutu enempää tehtävässä. Kyseinen tarkistaminen olisi kuitenkin helppoa, koska em. syntaktisen tarkistuksen jälkeen voidaan luottaa siihen, että kaikki numerot ovat oikeassa paikassa.

Joskus voidaan hyväksyä myös tyhjä jono. Esimerkiksi lauseke $a*b*$ hyväksyy mm. jonot a, b, aa, ab, aaaaaabbbbb. Toisaalta a voi olla nolla kertaa ja b voi olla nolla kertaa, joten lauseke hyväksyy myös tyhjän jonon eli jonon, jossa ei ole yhtään merkkiä (merkkien määrä on nolla; katso * taulukossa 1).

Kysymys 1. Vastaa kohtiin (a) ja (b) täyttämällä vastauspaperille kysymyksessä (a) oleva mallitaulukko ruksaamalla kunkin jonon osalta oikea vastaus.

(a) Olkoon säännöllinen lauseke muotoa

$$a. (t|mu)^*e$$

Huomaa, että **mu** muodostaa tässä kokonaisuuden.

Mitkä seuraavista jonoista lauseke hyväksyy ja mitkä hylkää?

Jono	Hyväksyy	Hylkää
ate		
aate		
atue		
aatutue		
aamute		
atmut		

(maksimipistemäärä 2)

(b) Mitkä kohdan (a) jonoista hyväksytään ja mitkä hylätään, jos lauseke on muotoa

$$a. ((t|m)u)^*e$$

(maksimipistemäärä 2)

Kysymys 2.

(a) Kellonajan tarkistava säännöllinen lauseke voisi olla seuraavaa muotoa:

$$[0-2]?[0-9]:[0-5][0-9]$$

Anna yksi esimerkki kellonajasta, jonka lauseke hyväksyy, muttei ole todellisuudessa kellovollinen kellonaika.

(maksimipistemäärä 1)

(b) Miksi seuraava lauseke korjaisi (a)-kohdan ongelmia?

$$(([01]?[0-9]) | (2[0-3])) : [0-5][0-9]$$

(maksimipistemäärä 1)

Kysymys 3. Olkoon säännöllinen lauseke muotoa

$$((1*0)\{3\})^*1^*$$

(a) Anna kolme jonoa pituudeltaan kymmenen merkkiä, jotka lauseke hyväksyy.

(maksimipistemäärä 1)

(b) Anna kolme jonoa pituudeltaan korkeintaan kymmenen merkkiä, joissa on joko merkkejä 0 tai merkkejä 1 (ei molempia) ja jotka lauseke hylkää.

(maksimipistemäärä 1)

(c) Hylkääkö vai hyväksyykö lauseke tyhjän jonon (ei yhtään merkkiä)? Perustele.

(maksimipistemäärä 1)

Kysymys 4.

- (a) Kirjoita viisi jonoa, joissa on vain merkkejä 0 ja 1 sekä jonoista löytyy osajono 00.
(maksimipistemäärä 1)
- (b) Kirjoita viisi jonoa, joissa on vain merkkejä 0 ja 1, mutta joista ei löydy osajonoa 00.
(maksimipistemäärä 1)

Vinkki seuraaviin kysymyksiin

Seuraavissa tehtävissä pyydetään säännöllistä lauseketta, joka **hyväksyy vain** sellaiset jonot, joissa kysymyksessä annetut ehdot täyttyvät, eli se **hylkää** kaikki muut.

Merkkijonojen tapauksessa termi *osajono* tarkoittaa jonon sellaista osaa, jonka merkit löytyvät samassa järjestyksessä peräkkäin itse jonosta. Esimerkiksi *ab* on jonojen *saab*, *abba* ja *ab* osajono, mutta *ab* ei ole jonon *samba* osajono.

Selventääksesi tehtävää itsellesi, kirjoita aina ensin itsellesi muutamia erilaisia esimerkkijonoja, jotka kysymyksen asettelun mukaan pitää hyväksyä, ja muutamia, jotka pitää hylätä. Muista sitten verrata vastauksiasi niihin. Muista silti, että kaikki muutkin säännön täyttävät jonot pitää hyväksyä.

Kysymys 5. Valitse alla olevista viidestä säännöllisestä lausekkeesta yksi, joka **hyväksyy vain** jonot, joissa kaikki seuraavat ehdot täyttyvät:

- on vain merkkejä 0 tai 1,
- jonosta löytyy osajono 00.

Lauseke	Hyväksyy
0^*1^*	
$[01]^+$	
$[01]^*00[01]^*$	
$00[01]^*$	
$[01]^+00[01]^+$	

(maksimipistemäärä 2)

Kysymys 6. Kirjoita säännöllinen lauseke, joka **hyväksyy vain** jonot, joissa kaikki seuraavat ehdot täyttyvät:

- on vain merkkejä 0 tai 1,
- löytyy ensin jostakin kohdasta osajono 00 ja sitten heti tai myöhemmin osajono 11. Huomautus: Jonon ei tarvitse alkaa osajonolla 00 eikä loppua osajonoon 11.

(maksimipistemäärä 2)

Kysymys 7. Kirjoita säännöllinen lauseke, joka **hyväksyy vain** jonot, joissa kaikki seuraavat ehdot täyttyvät:

- on vain merkkejä 0 tai 1,
- löytyy jostakin kohdasta osajono 00 tai osajono 11 (tai molemmat).

(maksimipistemäärä 2)

Kysymys 8. Kirjoita säännöllinen lauseke, joka **hyväksyy vain** jonot, joissa kaikki seuraavat ehdot täyttyvät:

- on vain merkkejä 0 tai 1 tai jono on tyhjä,
- ei ole kahta peräkkäistä merkkiä 0 (eli ei osajonoa 00).

(maksimipistemäärä 2)

Kysymys 9. Kirjoita säännöllinen lauseke, joka **hyväksyy vain** jonot, joissa kaikki seuraavat ehdot täyttyvät:

- on vain merkkejä 0 tai 1 tai jono on tyhjä,
- ei ole osajonoa 00, eikä osajonoa 11.

(maksimipistemäärä 3)

Tehtävä 2: Ohjelmointitehtävä

Kysymys 10. Alla on saman toiminnallisuuden sisältävä esimerkkialiohjelma kahdella eri ohjelmointikielellä tukemaan näiden kielten syntaksin muistamista. Voit myös käyttää jotain muuta ohjelmointikieltä seuraavan kysymyksen vastauksessasi. Mitkä ohjelmointikielet ovat kyseessä? (2 pistettä)

Esimerkki 1.

```
public static int summaaPositiiviset(int[] taulu){
    int summa = 0;
    for (int i = 0; i < taulu.length; i++) {
        if ( taulu[i] > 0 ) {
            summa = summa + taulu[i];
        }
    }
    return summa;
}
```

Esimerkki 2.

```
def summaa_positiiviset(taulu):
    summa = 0
    for alkio in taulu:
        if alkio>0:
            summa = summa + alkio
    return summa
```

Kysymys 11. Tee valitsemallasi ohjelmointikielellä aliohjelma (rutiini, proseduri, funktio tai metodi), joka hyväksyy samat jonot kuin edellisen tehtävän kysymyksessä 6. Aliohjelma ottaa syötteenä olevan tarkastettavan jonon parametrinaan joko kokonaislukutaulukko tai merkkijono. Aliohjelman tulee palauttaa boolean-arvo tosi jos parametrina tuotu syöte hyväksytään tai muuten epätosi. (20 pistettä)