

Datatähti 2019 alku

task	type	time limit	memory limit
A Kolikot	standard	1.00 s	512 MB
B Leimasin	standard	1.00 s	512 MB
C Taulukko	standard	1.00 s	512 MB
D Ruudukko	standard	1.00 s	512 MB
E Sanalista	output only	N/A	N/A

A Kolikot

Sinulla on n kolikkoa ja haluat jakaa kolikot pinoihin niin, että missään kahdessa pinossa ei ole yhtä monta kolikkoa.

Mikä on suurin mahdollinen pinojen määrä? Entä miten pinot voi muodostaa?

Syöte

Syötteen ainoalla rivillä on yksi kokonaisluku n : kolikoiden määrä.

Tuloste

Tulosta ensin kokonaisluku k : suurin mahdollinen pinojen määrä.

Tulosta sitten k kokonaislukua: kunkin pinon kolikoiden määrä. Voit tulostaa minkä tahansa kelvollisen ratkaisun.

Esimerkki 1

Syöte:
5

Tuloste:
2
2 3

Esimerkki 2

Syöte:
14

Tuloste:
4
1 3 4 6

Osatehtävä 1 (23 pistettä)

- $1 \leq n \leq 10$

Osatehtävä 2 (25 pistettä)

- $1 \leq n \leq 1000$

Osatehtävä 3 (52 pistettä)

- $1 \leq n \leq 10^9$

B Leimasin

Haluat muodostaa merkkijonon, jonka pituus on n merkkiä. Merkkijonon merkit on numeroitu $1, 2, \dots, n$, ja aluksi jokainen merkki on ?.

Käytössäsi on *leimasin*, jonka pituus on m merkkiä. Voit joka siirrolla asettaa leimasimen johonkin merkkijonon kohtaan ja korvata merkkijonon merkit tällä alueella vastaavissa kohdissa olevilla leimasimen merkeillä. Leimasimen täytyy olla kokonaan merkkijonon sisällä joka siirrolla, eikä sen asentoa voi muuttaa.

Tehtäväsi on etsiä siirtosarja, jolla voit tuottaa halutun merkkijonon leimasimen avulla, tai todeta, että tämä ei ole mahdollista.

Syöte

Ensimmäisellä rivillä on merkkijono, jossa on n merkkiä: tavoitteena oleva merkkijono.

Toisella rivillä on merkkijono, jossa on m merkkiä: leimasin.

Molemmat merkkijonot muodostuvat merkeistä A-Z.

Tuloste

Tulosta ensin kokonaisluku k : siirtojen määrä.

Tulosta sitten k kokonaislukua, jotka kuvaavat siirrot järjestyksessä. Jokainen luku on välillä $1 \dots n - m + 1$ ja ilmaisee leimasimen vasemman kohdan.

Voit tulostaa minkä tahansa ratkaisun, jossa k on enintään $10n$. Jos mitään ratkaisua ei ole olemassa, tulosta vain luku -1 .

Esimerkki 1

Syöte:
AABCACA
ABCA

Tuloste:
3
1 4 2

Selitys: Käytät leimasinta näin: ??????? → ABCA??? → ABCABCA → AABCACA.

Esimerkki 2

Syöte:

AYBABTU

AB

Tuloste:

-1

Osatehtävä 1 (29 pistettä)

- $1 \leq m \leq n \leq 10$

Osatehtävä 2 (31 pistettä)

- $1 \leq m \leq n \leq 100$

Osatehtävä 3 (40 pistettä)

- $1 \leq m \leq n \leq 1000$

C Taulukko

Annettuna on taulukko, jossa on n lukua. Monessako taulukon yhtenäisessä osataulukossa on enintään k eri lukua?

Syöte

Syötteen ensimmäisellä rivillä on kaksi kokonaislukua n ja k : taulukon koko ja suurin sallittu määrä eri lukuja osataulukossa.

Seuraavalla rivillä on n kokonaislukua x_1, x_2, \dots, x_n : taulukon sisältö.

Tuloste

Tulosta yksi kokonaisluku: tehtävän vastaus.

Esimerkki 1

Syöte:

5 2
1 2 3 1 1

Tuloste:

10

Esimerkki 2

Syöte:

8 1
1 1 1 1 1 1 1 1

Tuloste:

36

Osatehtävä 1 (12 pistettä)

- $1 \leq k \leq n \leq 100$
- $1 \leq x_i \leq 10$

Osatehtävä 2 (35 pistettä)

- $1 \leq k \leq n \leq 5000$
- $1 \leq x_i \leq 10^6$

Osatehtävä 3 (18 pistettä)

- $1 \leq k \leq n \leq 5 \cdot 10^5$

- $1 \leq k \leq 2$
- $1 \leq x_i \leq 10^9$

Osatehtävä 4 (35 pistettä)

- $1 \leq k \leq n \leq 5 \cdot 10^5$
- $1 \leq x_i \leq 10^9$

D Ruudukko

Tehtäväsi on sijoittaa $n \times n$ -ruudukkoon merkkejä A ja B niin, että jokaisessa ruudussa on enintään yksi merkki ja jokaisella pysty- ja vaakarivillä on tarkalleen yksi merkki A ja yksi merkki B.

Lisähaasteena osa merkeistä on sijoitettu valmiiksi. Monellako tavalla voit täydentää ruudukon valmiiksi?

Syöte

Syötteen ensimmäisellä rivillä on kokonaisluku n : ruudukon koko.

Tämän jälkeen syötteessä on n riviä, jotka kuvaavat ruudukon alkutilanteen. Jokaisella rivillä on n merkkiä. Merkki . tarkoittaa tyhjää ruutua, ja merkit A ja B ilmaisevat valmiiksi ruudukossa olevat merkit.

Voit olettaa, että syötteessä jokaisella pysty- ja vaakarivillä on enintään yksi merkki A ja enintään yksi merkki B.

Tuloste

Tulosta yksi kokonaisluku: tehtävän vastaus modulo $10^9 + 7$.

Esimerkki

Syöte:

5

.....
..AB.

.....
B....

...A.

Tuloste:

16

Osatehtävä 1 (31 pistettä)

- $2 \leq n \leq 5$

Osatehtävä 2 (14 pistettä)

- $2 \leq n \leq 500$
- ruudukossa ei ole valmiina mitään merkkejä

Osatehtävä 3 (55 pistettä)

- $2 \leq n \leq 500$

E Sanalista

Tehtäväsi on rakentaa *säännöllinen lauseke*, joka pyrkii tunnistamaan, onko annettu suomen kielen sana kirjoitettu oikein vai väärin päin. Lauseke testataan aineistolla, jonka voit ladata omalle koneellesi tutkittavaksi.

Aineisto

Tehtävän aineistona on kaksi sanalista:

- [Lista A](#): suomen kielen sanoja kirjoitettuna oikein päin
- [Lista B](#): samat sanat kirjoitettuna väärin päin

Kummankin listan jokaisella rivillä on yksi sana. Listassa A on esimerkiksi sana *ohjelmointi* ja listassa B vastaavasti *itniomlejho*. Huomaa, että listoissa on päällekkäisyyksiä: esimerkiksi molemmissa listoissa on sanat *aatto* ja *ottaa*.

Listat on tallennettu UTF-8-muodossa, millä on vaikutusta ääkkösten ja muiden erikoismerkkien käsittelyssä.

Aineisto perustuu Kotimaisten kielten keskuksen [nykysuomen sanalistaan](#), joka on julkaistu EUPL-lisenssillä. Listaa on muokattu tähän tehtävään sopivaksi mm. poistamalla sanat, joissa on tiettyjä erikoismerkkejä.

Tehtävä

Laadi säännöllinen lauseke S , joka täsmää mahdollisimman moneen listan A sanaan ja mahdollisimman harvaan listan B sanaan.

Molemmissa listoissa on $n = 93855$ sanaa. Olkoot a niiden listan A sanojen määrä, joihin S täsmää, ja b niiden listan B sanojen määrä, joihin S täsmää. Tällöin osumatarkkuus on

$$r = \frac{a + n - b}{2n}.$$

Tavoitteesi on päästä mahdollisimman korkeaan osumatarkkuuteen.

Esimerkkejä:

- Säännöllinen lauseke `.*rv.*` täsmää $a = 1453$ sanaan listassa A ja $b = 2$ sanaan listassa B, joten osumatarkkuus on $r \approx 0.508$.
- Säännöllinen lauseke `xyzzy` ei täsmää mihinkään sanaan listoissa A ja B, ja osumatarkkuus on $r = 0.5$.
- Säännöllinen lauseke `.*` täsmää kaikkiin sanoihin listoissa A ja B, ja osumatarkkuus on $r = 0.5$.

Pisteytys

Jos laatimasi lauseke on kelvollinen (ks. alla täsmälliset säännöt), saat pisteitä osumatarkkuudesta r riippuen kaavalla

$$\min(100, \max(0, \lfloor 250 \cdot (r - 0.55) \rfloor)).$$

Esimerkiksi osumatarkkuus 0.55 antaa 0 pistettä, osumatarkkuus 0.75 antaa 50 pistettä ja osumatarkkuus 0.95 antaa 100 pistettä, mikä on tehtävän maksimipistemäärä.

Säännölliset lausekkeet

Säännöllinen lauseke on merkkijono, joka täsmää yhteen tai useampaan merkkijonoon. Tässä tehtävässä voit käyttää lausekkeessa kaikkia aineistossa esiintyviä merkkejä, vaihtoehtoa (`|`), toistoa (`*`), sulkuja sekä yleismerkkiä (`.`).

Vaihtoehdon avulla voit antaa useamman kelpaavan merkkijonon. Esimerkiksi lauseke `apina|banaani|cembalo` täsmää merkkijonoihin `apina`, `banaani` ja `cembalo`.

Toisto tarkoittaa, että sitä edeltävä lausekkeen osa voi toistua miten tahansa monta kertaa. Esimerkiksi lauseke `(aa)*` vaatii, että merkkijonon jokainen merkki on `a` ja sen pituus on parillinen (myös tyhjä merkkijono on sallittu).

Yleismerkki `.` täsmää mihin tahansa yksittäiseen merkkiin. Esimerkiksi lauseke `.*abc.*` täsmää mihin tahansa merkkijonoon, jossa on jossain kohtaa merkkijono `abc`, koska alussa ja lopussa voi olla mitä tahansa.

Esimerkki mutkikkaammasta lausekkeesta on `api(la|na)|.*s`. Tämä lauseke täsmää sanoihin `apila`, `apina` sekä mihin tahansa sanaan, jossa viimeinen kirjain on `s`, kuten sanoihin `ilmoitus` ja `ratas`.

Palautusohje

Palauta UTF-8-muotoinen tekstitiedosto, jossa on laatimasi säännöllinen lauseke. Lausekkeen pituus saa olla enintään **500 merkkiä**.

Voit käyttää halutessasi lausekkeessa välilyöntejä ja rivinvaihtoja. Tarkastaja jättää nämä huomiotta, eikä niitä lasketa lausekkeen pituuteen.

Esimerkiksi `.*rv.*` on kelvollinen ratkaisu, jonka pituus on 6 merkkiä. Tämän lausekkeen osumatarkkuus on 0.508, joten sillä ei saa vielä tehtävästä pisteitä.