

VIENKĀRŠĀ DATU SASPIEŠANA

Uzdevums 1. projektam studiju kursā "Risinājumu algoritmizēšana un programmēšana (2)"

Uzdevums

Nukleotīdi, kas veido gēnus DNS molekulās, tiek apzīmēti ar vienu no četrām vērtībām: A, C, G vai T. Ja gēna apraksts tiks saglabāts mainīgā tipa String, tad katra nukleotīga kodēšanai būs nepieciešami divi baiti, jeb 16 biti, jo String tipa virknēs valodā Java sastāv no Unicode simboliem. Taču četru iespējamo vērtību kodēšanai var izmantot divus bitus, piemēram, A iekodēt ar bitiem 00, C ar bitiem 01, G ar 10, bet T ar 11. Tad atmiņas apjoms, kas ir vajadzīgs gēna apraksta glabāšanai, var būt samazināts 8 reizes (no 16 bitiem līdz 2 bitiem katra nukleotīda kodēšanai). [David Kopec "Classic Science Problems in Python"]

Izstrādāji programmu, kas izpilda sekojošas darbības:

- Pārveido simbolu virkni no simboliem A, C, G un T, par masīvu no baitiem, iekodējot katru no simboliem ar diviem bitiem.
- Pārveido masīvu no baitiem, par simbolu virkni no simboliem A, C, G un T, aizvietojo katru bitu pāru ar atbilstošu simbolu.

Izstrādāto programmu augšupielādējiet ORTUSā un pārbaudiet ar testiem replit.com portālā.

Prasības programmai

- programma ir jāizstrādā kā konsoles lietotne (bez grafiskiem logiem), prasības lietotāja saskarei ir aprakstītas zemāk;
- programmā jāiekļauj lietotāja nepareizu darbību apstrāde;
- Programmas pirmkodu izvietot vienā failā ar nosaukumu Main.java;
- Failā Main.java sākumā ievietot komentāru, kas satur informāciju par programmas autoru (apliecības numurs, vārds, uzvārds, grupa);
- Operatoru package neizmantot.

Prasības lietotāja saskarnei:

Programmai ir jāapstrādā vaicājumi comp, decomp, about un exit. Kamēr lietotājs neievadīs komandu exit programmai bezgalīgā ciklā jāievada komandas un jāizpilda atbilstošas darbības. Visu darbību apraksts un komandu formāts ir parādīts sekojošā tabulā:

Vaicājums	Komandas formāts un programmas darbības apraksts
comp 1 punkts	Komandas formāts: comp str <i>kur str ir simbolu virkne</i> Darbības: <ul style="list-style-type: none">• Pārbaudīt vai ievadīta simbolu virkne <i>str</i> satur tikai burtus A, C, G un T (lielos un mazos burtus neatšķirt). Ja simbolu virkne satur arī citus simbolus, tad izvadīt paziņojumu "wrong command format" un pabeigt komandas <i>comp</i> izpildi.• Pārveido ievadīto simbolu virkni par masīvu no elementiem tipa byte, iekodējot simbolu A ar bitiem 00, C ar bitiem 01, G ar 10, bet T ar 11. Pirmā masīva elementā ierakstīt burtu skaitu, neizmantotos bitus pēdējā elementā aizpildīt ar nullēm.• Izvadīt baitu masīvu ekrānā. Masīva elementus atdalīt ar atstarpes simboliem, elementu vērtības izvadīt heksadecimālā skaitīšanas sistēmā (ar lieliem burtiem).

Vaicājums	Komandas formāts un programmas darbības apraksts
decomp	<p>Komandas formāts: decomp n b1 b2 b3 ... bn</p> <p>kur:</p> <ul style="list-style-type: none"> • <i>n b1 b2 b3 ... bn</i> ir veseli skaitļi tipa <i>byte</i> • <i>n</i> ir skaitļu <i>b1 b2 b3 ... bn</i> daudzums <p>Darbības:</p> <ul style="list-style-type: none"> • Izvadīt baitu masīvu heksadecimālā skaitīšanas sistēmā un izvadīt ekrāna atdalot elementus ar vienu atstarpi. • Pārveidot baitu masīvu <i>b1 b2 b3 ... bn</i> par simbolu virkni no simboliem A, C, G un T un izvadīt iegūto simbolu virkni ekrānā. • Gadījumā, ja aiz komandas decomp tika ievadīts teksts vai tika ievadīti pārāk lieli skaitļi, kurus nav iespējams saglabāt tipa <i>byte</i> mainīgā, vai skaitļu ir pārāk maz vai daudz, izvadīt paziņojumu "wrong command format" un pabeigt komandas decomp izpildi.
about	<p>Komandas formāts: about</p> <p>Darbības:</p> <ul style="list-style-type: none"> • Izvadīt informāciju par programmas autoru (studenta apliecības numurs, vārds, uzvārds, grupa).
exit	<p>Komandas formāts: exit</p> <p>Darbības:</p> <ul style="list-style-type: none"> • Pārtraukt programmas darbību. <p>Kamēr nav ievadīta komanda <i>exit</i> programmai ciklā ir jāievada un jāizpilda lietotāja vaicājumus, tādus kā <i>comp</i>, <i>decomp</i>, <i>about</i>. Ja lietotājs ievadīja neeksistējošu komandu, tad izvadīt paziņojumu "wrong command" un turpināt darbu.</p>

Piemēram, ja lietotājs ievadīja:

comp CGATAAG

Tad programmai ir jāizvada:

7 63 8

Kur 7 ir simbolu skaits.

Bet, ja lietotājs ievadīja:

decomp 4 10 30 -127 32



Šeit 4 ir ievadīto baitu skaits, bet 10 ir nukleotīdu skaits

Tad programmai ir jāizvada:

A 1E 81 20

ACTGGAACAG