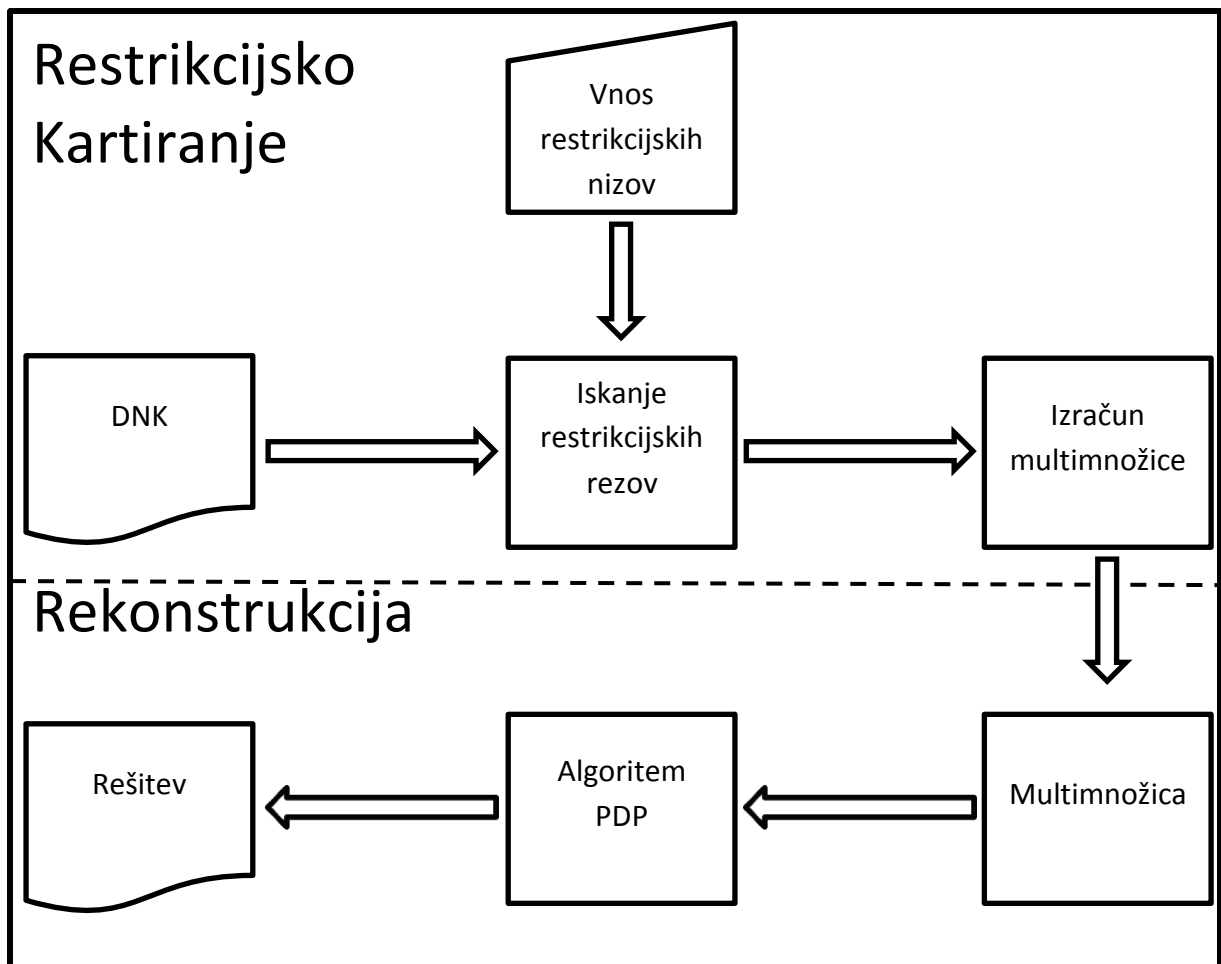


Restriksijsko kartiranje

Izdelajte preprosto aplikacijo za restriksijsko kartiranje poljubnega DNK zaporedja.

Aplikacija naj omogoča naslednje:

- Branje tekstovnih datotek z zapisi DNK ([DNK1](#), [DNK2](#), [DNK3](#))
- Iskanje poljubnih restriksijskih rezov (nukleotidov) v molekulah DNK
- Izračun vseh razdalj med restriksijskimi rezi (multimnožica)
- Rekonstrukcija restriksijskih rezov iz multimnožice
 - Naivni pristop
 - Razveji in omeji
- Izvoz rešitev v tekstovno datoteko
- Meritve



Naivni pristop:

```
ANOTHER_BRUTE_FORCE( $L, n$ )  
   $M \leftarrow$  največji elemnt v  $L$   
  for  $\forall$  množica  $n - 2$  naravnih št.  $0 < x_2 < \dots < x_{n-1} < M$  iz  $L$   
     $X \leftarrow \{0, x_2, \dots, x_{n-1}, M\}$   
    Tvor  $\Delta X$  iz  $X$   
    if  $\Delta X = L$   
      return  $X$   
  output »Ni rešitve«
```

Razveji in omeji:

```
PARTIAL_DIGEST(L)
  širina ← največji element v L
  DELETE(širina, L)
  X ← {0, širina}
  PLACE(L, X)
```

```
PLACE(L, X)
  if L je prazna
    output X
    return
  y ← največji element v L
  if  $\Delta(y, X) \in L$ 
    dodaj y v X in briši razdalje  $\Delta(y, X)$  iz L
    PLACE(L, X)
    briši y iz X in dodaj razdalje  $\Delta(y, X)$  v L
  if  $\Delta(\text{širina} - y, X) \in L$ 
    dodaj širina - y v X in briši razdalje  $\Delta(\text{širina} - y, X)$  iz L
    PLACE(L, X)
    briši širina - y iz X in dodaj razdalje  $\Delta(\text{širina} - y, X)$  v L
  return
```

- Funkcija $DELETE(y, L)$ briše y iz L .
- $\Delta(y, X)$ predstavlja multimnožico razdalj med y in vsemi točkami v X .

$$\Delta(2, \{1, 3, 4, 5\}) = \{1, 1, 2, 3\}$$

Primer:

$$L = \{2, 2, 3, 3, 4, 5, 6, 7, 8, 10\}$$

$$|L| = 10 = \frac{n(n-1)}{2} \rightarrow n = 5$$

$$x = \{0, x_2, x_3, x_4, x_5\}$$

- Vstavimo največji element in ga odstranimo iz množice L

$$x = \{0, 10\}; L = \{2, 2, 3, 3, 4, 5, 6, 7, 8\}$$

- Poiščemo največji element iz množice L . Imamo dve možnosti:

$$x_2 = 2 \rightarrow |x_2 x_5| = 8$$

$$x_4 = 8 \rightarrow |x_1 x_4| = 8$$

- Izberemo eno možnost in izločimo razdalje iz množice L

$$x_2 = 2$$

$$|x_1 x_2| = 2, |x_2 x_5| = 8$$

$$x = \{0, 2, 10\}; L = \{2, 3, 3, 4, 5, 6, 7\}$$

- Poiščemo največji element iz množice L . Imamo dve možnosti:

$$x_3 = 3 \rightarrow |x_2 x_3| = 7$$

$$x_4 = 7 \rightarrow |x_1 x_4| = 7$$

- Izberemo eno možnost in izločimo razdalje iz množice L

$$x_3 = 3$$

$$|x_2 x_3| = 1$$

$$|x_2 x_3| \notin L$$

- Izberemo drugo možnost

$$x_4 = 7$$

$$|x_1 x_4| = 7, |x_2 x_4| = 5, |x_4 x_5| = 3$$

$$x = \{0, 2, 7, 10\}; L = \{2, 3, 4, 6\}$$

- Poiščemo največji element iz množice L . Imamo dve možnosti:

$$x_3 = 4 \rightarrow |x_3 x_5| = 6$$

$$x_3 = 6 \rightarrow |x_1 x_3| = 6$$

- Izberemo eno možnost in izločimo razdalje iz množice L

$$x_3 = 6$$

$$|x_3 x_4| = 1$$

$$|x_3 x_4| \notin L$$

- Izberemo drugo možnost

$$x_3 = 4$$

$$|x_1 x_3| = 4, |x_2 x_3| = 2, |x_3 x_4| = 3, |x_3 x_5| = 6$$

$$x = \{0, 2, 4, 7, 10\}; L = \emptyset$$

- Rešitev

$$x = \{0, 2, 4, 7, 10\}$$

DNK1: **GTGTG** (1)

Multimnožica: 168,831,999

Rešitev: 0,831,999

0,168,999

DNK1: **TTCC, CTCTCT** (7)

Multimnožica: 15,47,66,73,73,88,103,113,118,139,150,165,186,189,191,216,231,238,289,
304,304,377,433,506,572,619,622,695,722,737,761,808,810,911,926,999

Rešitev: 0,189,622,695,761,808,911,926,999

0,73,88,191,238,304,377,810,999

DNK1: **AAAA, CCCC, TTTT, GGGG** (8)

Multimnožica: 1,1,47,49,50,73,74,105,119,120,120,121,153,193,200,202,203,240,249,250,
273,274,322,323,323,324,393,442,443,451,452,501,556,557,606,654,701,759,774,775,806,
879,880,894,999

Rešitev: 0,119,120,193,240,393,442,443,894,999

0,105,556,557,606,759,806,879,880,999

DNK2: **ACTACT, GGAGGA, GAGGCC, CTCTCT** (10)

Multimnožica: 192,203,252,292,381,395,452,502,705,744,807,897,999,1059,1202,1251,
1454,1501,1704,1793,1882,1956,2126,2174,2245,2378,2626,3185,3291,3377,3580,3672,
4082,5173,5417,5465,5669,5798,5917,6050,6476,6668,6857,6871,7049,7252,7299,7373,
7551,7591,7754,7843,8043,8295,8358,8550,8650,8753,8842,9045,9102,9255,9294,9497,
9547,9999

Rešitev: 0,502,705,897,1704,1956,4082,7373,7754,9255,9547,9999

0,452,744,2245,2626,5917,8043,8295,9102,9294,9497,9999

DNK3: **TTTTTTT, GTGTCGT, ACACACA** (13)

Multimnožica: 1,1,28,548,549,807,835,1305,2112,2140,3758,3759,4307,5074,5467,5758,
6772,7579,7607,9516,9517,10065,11225,12530,13337,13365,14983,14984,15532,16288,
16289,16837,17017,17018,17095,17096,17123,17124,17565,17566,17566,17567,17644,
17672,18618,18619,21324,21325,23692,23693,27082,27083,32549,32550,33854,33855,
34661,34662,34689,34690,35636,36184,36185,39943,40710,41258,41259,41617,41645,
42452,43757,45017,45701,49224,50775,51168,52473,53280,53308,54982,56242,57547,
58354,58382,58740,58741,59289,76306,76307,94925,99999

Rešitev:

0,41617,41645,42452,43757,49224,54982,58740,58741,59289,76306,76307,94925,99999

0,5074,23692,23693,40710,41258,41259,45017,50775,56242,57547,58354,58382,99999