

Sprawozdanie 1

Teoretyczne opracowanie metody heurystycznej

Mateusz Babiaczyk, Bartosz Nawrotek

2018-04-08

1 Wprowadzenie teoretyczne

Do rozwiązania problemu zastosowaliśmy algorytm genetyczny.

1.1 Kodowanie

Osobniki są kodowane jako wektory liczb całkowitych w których indeks oznacza kolejność występowania oligonukleotydu w sekwencji:

$$X = [x_1, x_2, \dots, x_m], \quad (1)$$

gdzie:

1. m - wielkość zbioru dostępnych oligonukleotydów
2. x_i dla $i \in \langle 1, m \rangle$ - indeks oligonukleotydu w liście dostępnych oligonukleotydów

Jako rozwiązanie będziemy traktować sekwencje zbudowaną z oligonukleotydów począwszy od x_1 nie przekraczającą długości n .

1.2 Funkcja oceny

Jako minimalizowaną funkcję oceny osobnika przyjęliśmy następującą postać addytywną:

$$f_{min}(X) = \sum_{i=1}^{i=k-1} [2len(x_i, x_{i+1}) - l + 1] + \sum_{i=k}^{i=m-1} [len(x_i, x_{i+1}) - l + 1] \quad (2)$$

gdzie:

1. k - taka liczba całkowita, dla której długość sekwencji $[x_1, x_2, \dots, x_k]$ będącej złożeniem k pierwszych oligonukleotydów osobnika X jest mniejsza lub równa n
2. $len(x, y)$ jest długością sekwencji otrzymanej z połączenia oligonukleotydów x oraz y
3. l - długość oligonukleotydu

Przyjmując funkcję kosztu powyższej postaci pragnęliśmy zwrócić szczególną uwagę na pierwsze k oligonukleotydów z osobnika, które zawierają szukane rozwiązanie. Druga część sumy ma na celu nie losowe uporządkowanie pozostałych oligonukleotydów, które będą wykorzystane w operacji krzyżowania