Natalia Niewdzięczna 149514

Wadim Sokołowski 149557

Mikołaj Grochowski 149435

**Elementy Bioinformatyki**

**Temat 15. Filogeneza dystansowa. Typ A.** Dla danej macierzy odległości między liśćmi utworzyć ważone drzewo ukorzenione metodą UPGMA oraz nieukorzenione metodą NJ. Wylistowanie obu drzew oraz możliwość automatycznej weryfikacji zgodności ich topologii (tj. z pominięciem długości krawędzi)

**1. Opis problemu**

Naszym zadaniem było utworzenie programu, który na wejściu otrzymuje macierz odległości między liśćmi, na wyjściu zaś zwraca ważone drzewo ukorzenione metodą UPGMA oraz nieukorzenione metodą NJ. W naszym projekcie nie zaimplementowano mechanizmu automatycznej weryfikacji zgodności topologii otrzymanych drzew.

2. **Algorytmy**

* **Algorytm UPGMA**

Pseudokod**:**

clusters [|L|]

while ( clusters . length > 1):

calculate distances between clusters

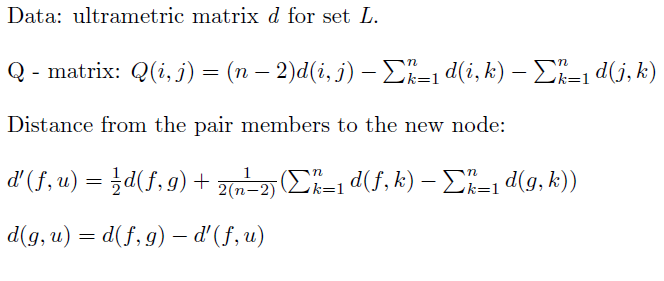
(sum of distances between cluster members

divided by product of cluster cardinalities)

find the lowest distance

merge the closest clusters

* **Algorytm NJ**



Pseudokod:

clusters [|L|]

while ( number of clusters > 2):

calculate Q- matrix

find the lowest q- distance

merge the q- closest clusters

update distances

merge last 2 clusters

**3. Przykładowe dane wejściowe/wyjściowe**

**Wejście programu**:

Ultrametryczna macierz odległości

a,b,c,d,e

0 ,8 ,8 ,5 ,3

8 ,0 ,3 ,8 ,8

8 ,3 ,0 ,8 ,8

5 ,8 ,8 ,0 ,5

3 ,8 ,8 ,5 ,0

**Wyjście programu:**

**Drzewo utworzone przy pomocy metody UPGMA**

[[8.0]]

|-[[5.0]]

| |-[[3.0]]

| | |---- [[a]]

| | `---- [[e]]

| `---- [[d]]

`-[[3.0a]]

|---- [[b]]

`---- [[c]]

**Drzewo utworzone przy pomocy metody NJ** (pomimo faktu, że algorytm NJ zwraca drzewo nieukorzenione, dla łatwiejszej wizualizacji jest ono zaprezentowane jako drzewo ukorzenione z pewnym węzłem wybranym jako korzeń tegoż drzewa)

[[g]]

|---- [1.5-<-[a]]

|---- [1.5-<-[e]]

`-[1.75-<-[h]]

|---- [1.0-<-[d]]

`-[4.0-<-[f]]

|---- [1.5-<-[b]]

`---- [1.5-<-[c]]