

## Wstęp do bioinformatyki

### ZADANIE 5: Dendrogram UPGMA

Odnośnik do repozytorium:  
<https://github.com/236807/WBI>

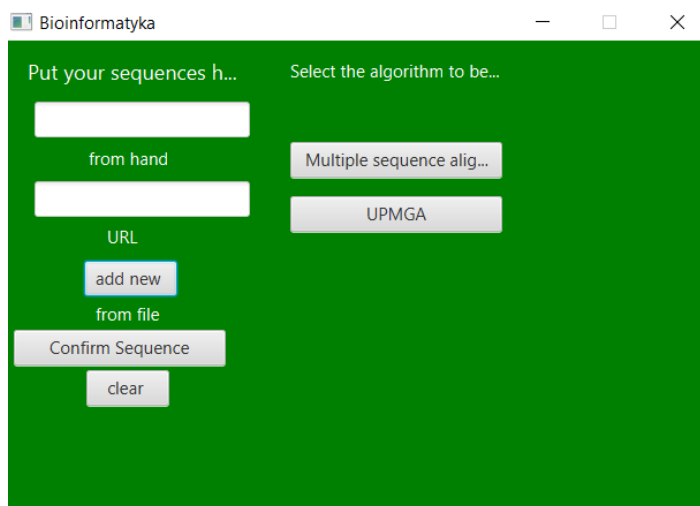
#### 1) Cel ćwiczenia

Stworzenie dendrogramu metodą UPGMA dla sekwencji nukleotydowych i/lub białkowych.

#### 2) Zasada działania programu

Jest to podprogram dodany do zadania 4.

Po uruchomieniu widzimy graficzny interfejs użytkownika, aby dodać sekwencje należy wpisać ją w odpowiednim polu, lub link do niej (wersja z URL) i potwierdzić przyciskiem, lub aby dodać z pliku kliknąć odpowiedni przycisk i wyznaczyć ścieżkę do naszej sekwencji. Po dodaniu odpowiedniej ilości sekwencji klikamy **UPMGA**, który generuje nam plik z podobieństwem sekwencji w formacie Newick.



Wynik programu znajduje się w pliku Newick.txt:



#### 3) Złożoność czasowa

Podana metoda jest rzędu  $\Theta(n*m) + 2 \Theta(n^2)$ .

#### 4) Złożoność pamięciowa

Macierz którą implementuje z n sekwencji, więc złożoność pamięciowa jest rzędu  $3\Theta(n)$ .

#### 5) Testy akceptacyjne

Zrzuty ekranu w celu sprawdzenia poprawności każdego z testów zostały wrzucone poniżej tabeli.

Id	Cel testu	Wprowadzone dane	Spodziewany wynik
1	Sprawdzenie poprawności formatu Newick	1) CCCTGATAGCTAGAT 2) GTCAACTTGACT 3)AGTCATGTACTA 4) CGTGTATCAGT	(( (CCCTGATAGCTAGA-T:4.0, -CGTG-TATC-AG--T:4.0):1.5, -GTCA-A-CTTGACT:5.5):1.75, -AGTCAT-G-TACT-A:7.25)
2	Sprawdzenie wyniku przy 2 takich samych sekwencjach	1)CCCTGATAGCTAGAT 2)CCCTGATAGCTAGAT	0.0,0.0, 0.0,0.0
3	Poprawność dopasowanie globalnego	1) CCCTGATAGCTAGAT 2) GTCAACTTGACT 3)AGTCATGTACTA 4) CGTGTATCAGT	CCCTGATAGCTAGA-T --GTCA-A-CTTGACT -AGTCAT-G-TACT-A -CGTG-TATC-AG--T
4	Poprawne budowanie macierzy podobieństwa sekwencji	1) CCCTGATAGCTAGAT 2) GTCAACTTGACT 3)AGTCATGTACTA 4) CGTGTATCAGT	0.0, 11.0, 15.0, 8.0, 11.0, 0.0, 14.0, 11.0 15.0, 14.0, 0.0, 15.0 8.0, 11.0, 15.0, 0.0
5	Sprawdzenie dwóch różnych w każdym nukleotydzie sekwencji	1) ACTG 2) GTCA	0.0,8.0, 8.0,0.0
6	Sprawdzenie dwóch podobnych lecz nie identycznych do siebie sekwencji	1)AGTCATGTACTA 2)CGTGTATCAGT	(AGT-CATGTACTA:5.5, CGTGTAT-CAGT-:5.5)
7	Sprawdzenie poprawności wprowadzania sekwencji do tablicy	1) ACTG 2) ACTG 3)AGTCATGTACTA 4) CGTGTATCAGT	1) ACTG 2) ACTG 3)AGTCATGTACTA 4) CGTGTATCAGT
8	Sprawdzenie poprawności	1) ACTG 2) ACTG	4

	policzenia ilości wprowadzanych sekwencji	3)AGTCATGTACTA 4) CGTGTATCAGT	
--	---	----------------------------------	--

- 1) Zrzut ekranu:

```
Phylogenetic Tree (Newick format): (((CCCTGATAGCTAGA-T:4.0, -CGTG-TATC-AG--T:4.0):1.5, |--GTCA-A-CTTGACT:5.5):1.75, -AGTCAT-G-TACT-A:7.25);
```

TEST ZALICZONY

- 2) Zrzut ekranu:

```
... 10 more
Caused by: java.lang.NullPointerException
    at msa.Calculator.multipleSequenceAlignment(Calculator.java:135)
    at api.WindowController.upmga(WindowController.java:115)
    ... 58 more
```

---

TEST NIEZALICZONY

- 3) Zrzut ekranu:

Multiple Sequence Alignment:

KOSZT CALKOWITY =: 146

```
CCCTGATAGCTAGA-T
--GTCA-A-CTTGACT
  * * * * *
--GTCA-A-CTTGACT
--GTCA-A-CTTGACT
*****
--GTCA-A-CTTGACT
--GTCA-A-CTTGACT
*****
--GTCA-A-CTTGACT
-AGTCAT-G-TACT-A
* * * * *
-AGTCAT-G-TACT-A
-CGTG-TATC-AG--T
* * * * *
```

TEST ZALICZONY

- 4) Zrzut ekranu:

```
[[0.0, 11.0, 15.0, 8.0], [11.0, 0.0, 14.0, 11.0], [15.0, 14.0, 0.0, 15.0], [8.0, 11.0, 15.0, 0.0]]
```

TEST ZALICZONY

- 5) Zrzut ekranu:

```
[[0.0, 8.0], [8.0, 0.0]]
```

TEST ZALICZONY

6) Zrzut ekranu:

`(AGT-CATGTACTA:5.5, CGTGTAT-CAGT-:5.5)`

!

TEST ZALICZONY

7) Zrzut ekranu:

`[AGTCATGTACTA, CGTGTATCAGT, ACTG, ACTG]`

TEST ZALICZONY

8) Zrzut ekranu:

`WARNING: Loading FXML document with JavaFX API of version 10.0.1 by JavaFX runtime of version 8.0.191`  
`4`

TEST ZALICZONY