

Monika Reguła 236689

- globalne dopasowanie (Needleman-Wunsch)
- lokalne dopasowanie (Smith-Waterman)

- dopasowania
- niedopasowania
- przerwy

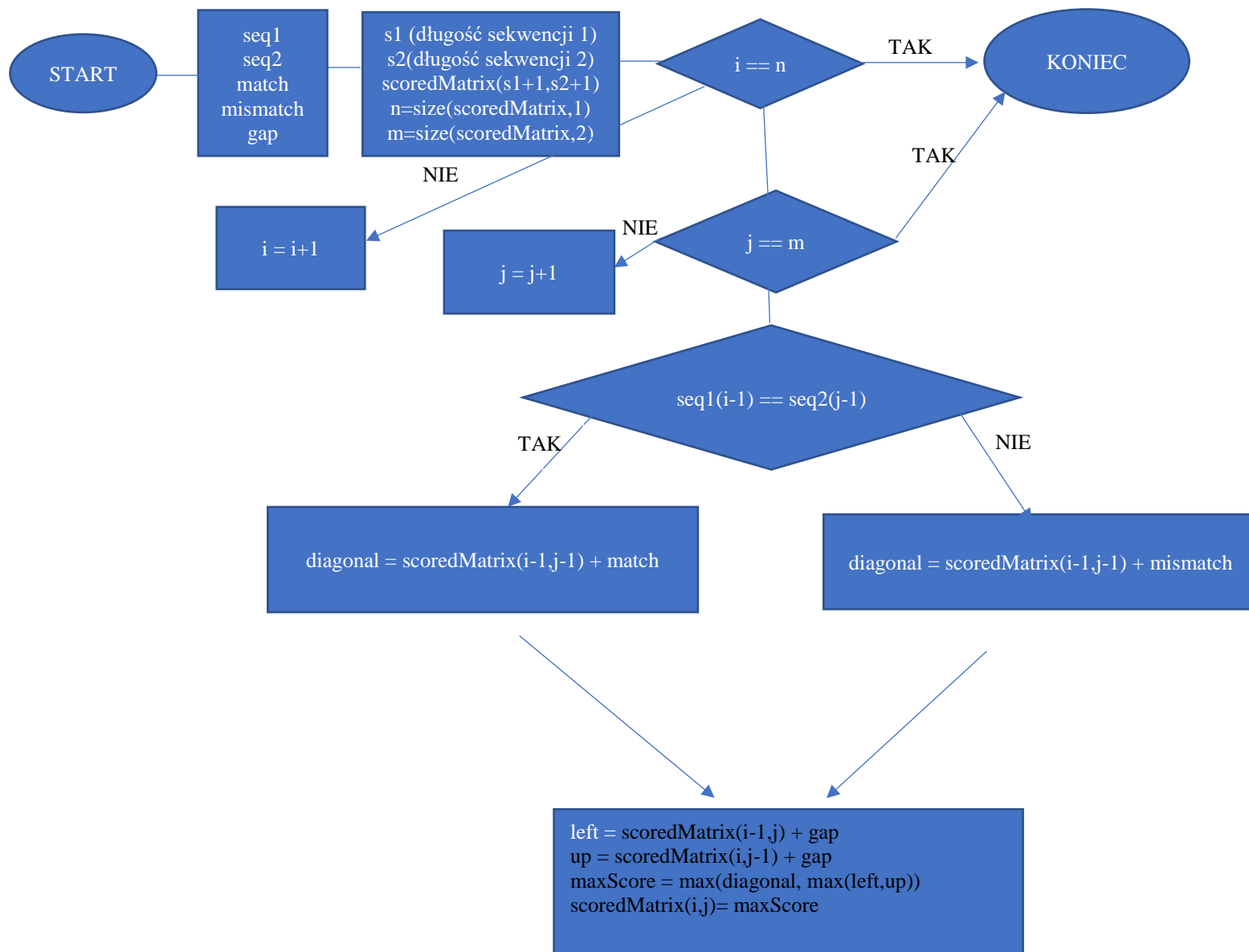
Przerwy są wstawiane, aby użyć wzrost liczby dopasowań w innych miejscach. Znajduję optymalne dopasowanie (jednak może ich być kilka). Tworzę macierz punktacji, w której każda komórka reprezentuje punktację dla najlepszego dopasowania kończącego się w danej pozycji. Wybieram maksymalną wartość z 3 możliwych (lewo, prawo, przekątna):

Po wykonaniu macierzy punktacji przechodzę do poszukiwania najbardziej optymalnej ścieżki

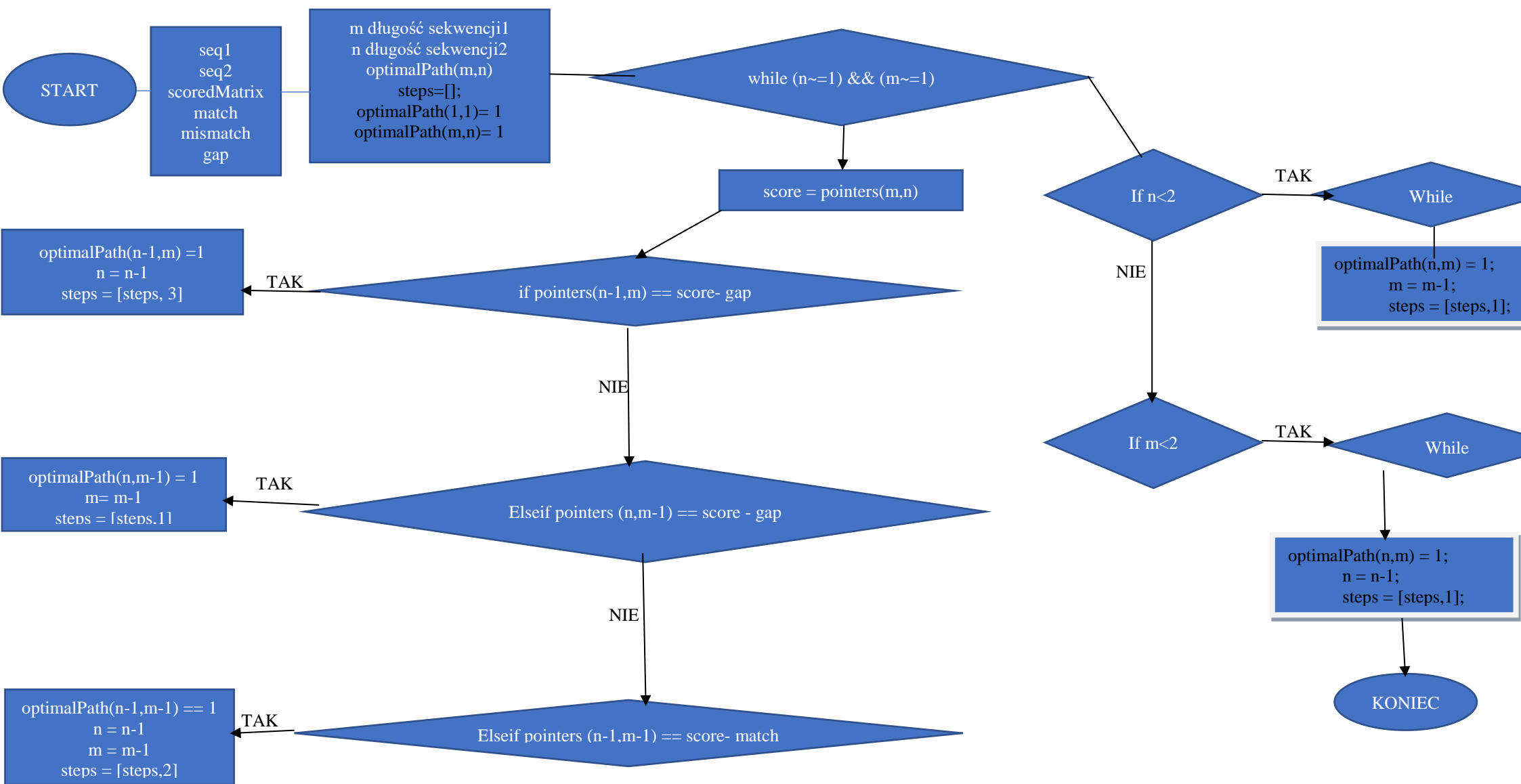
Następnie w pętli while, której warunkiem jest sprawdzanie czy komórka nie jest lewym górnym rogiem czy też prawym dolnym. W bloku while dochodzi do mapowania kierunku drogi.

```
optimalPath =
```

[illegible]



Schemat blokowy
globalMatching.m



Schemat blokowy
traceback.m

1. Analiza obliczeniowa:

- globalMatching.m
 - czasowa $O(m)$, przy założeniu, że $m \sim n$
 - pamięciowa $O(m \cdot n)$, $m \sim n$ $O(m^2)$
- traceback.m
 - czasowa $O(m \cdot n)$, $m \sim n$ $O(m^2)$
 - pamięciowa $O(m \cdot n)$, $m \sim n$ $O(m^2)$

Sumarycznie:

- pamięciowa $O(m^2) + O(m^2) = 2 O(m^2) \sim O(m^2)$
- czasowa $O(m^2) + O(m) \sim O(m^2)$

2. Porównanie przykładowych par sekwencji ewolucyjnie :

- Powiązanych

```
%insulina wariant 1 u krowy i insulina wariant 2 krowy |
globalMatching(getFromNCBI('NM_173926.2'),getFromNCBI('NM_001185126.1'),1,-1,-1)
```

Rysunek 2 Wywołanie test2.m

```
#Mode: distance
#Score: 461
#Length: 516
#Match: 1
#Mismatch: -1
#Gap: -1
#Identity: 497/516 96%
#Gaps: 31/516 6%
```

Rysunek 3 Plik tekstowy zapisany w wyniku wywołania test2.m

- Niepowiązanych

```
%insulina wariant 1 u człowieka i krowy
globalMatching(getFromNCBI('NM_000207.2'),getFromNCBI('NM_173926.2'),1,-1,-1)
```

Rysunek 4 Wywołanie test3.m

```
#Mode: distance
#Score: 267
#Length: 548
#Match: 1
#Mismatch: -1
#Gap: -1
#Identity: 416/548 76%
#Gaps: 77/548 14%
```

Rysunek 5 Plik tekstowy zapisany w wyniku wywołania test3.m

Dla obu wariantów insuliny u krowy (*bos taurus*) widać duże podobieństwo wynoszące 96%. Inny wariant insuliny w obrębie jednego gatunku wykazuje wysoką zgodność. Natomiast insulina wobec 2 różnych gatunków wykazuje znacznie niższą zgodność. Dla insuliny u człowieka oraz krowy podobieństwo jest mniejsze, wynoszące 76%.