

## Praktikum, Genomorientierte Bioinformatik

Blatt 1, 17.10.2012

Abgabe 31.10.2012

# 1 Gotoh - Einzelabgabe

Implementieren Sie den Algorithmus von Gotoh zum Erstellen von globalen, lokalen und Free-Shift Alignments zweier Aminosäuresequenzen. Dazu benötigte Substitutionsmatrizen liegen als Dateien im Verzeichnis

```
/home/proj/biosoft/praktikum/genprakt-ws13/assignment1/matrices
```

vor. Geben Sie ein von Kommandozeile aufrufbares Programm ab, das folgende Parameter akzeptiert:

Alignment aller Paare in pairfile:

```
java -jar gotoh.jar -pairs <pairfile> -seqlib <seqlibfile> <optional params>
```

seqlib enthaelt Zeilen der Form "id:sequence"

Optionale Parameter sind:

-m	matrixname (Standard dayhoff)
-go	gapopen (Standard -12)
-ge	gapextend (Standard -1)
-mode	eines aus local global freeshift (Standard freeshift)
-printali	gibt auch jedes Alignment aus
-printmatrices	eines aus txt html, gibt auch die Gotoh-Matrizen aus, entweder als tab separiert oder html
-check	ueberprueft die berechneten Scores anhand des Alignments

Beispiel-Aufruf:

```
java -jar gotoh.jar -pairs cathscop.inpairs -seqlib domains.seqlib -mode freeshift -matrixname blosum50
```

Das Ausgabeformat für den Aufruf ohne die -printali Option ist (für jede Zeile im Pairfile ein Eintrag):

```
1j2xA00 1wq2B00 33.900
```

Das Ausgabeformat für den Aufruf mit der -printali Option ist (für jede Zeile im Pairfile ein Eintrag):

```
>1j2xA00 1wq2B00 33.900
1j2xA00: GPLDVQVTEDAVRRYLTRKPMTTKDLLKKFQTKKTGLSSEQTVNVLAQILKRLNPERKMINDKMHFSLK-
1wq2B00: ----MEEAKQKVVDFLNSKSK-SKFYFNDFDLFPDMKQREVKKILTALVNDEVLEYWSSGSTTMYGLKG
```

Die Ausgaben sind auf die Standardausgabe zu schreiben. Achten Sie darauf, dass die Eingabe/Ausgabe Ihrer Implementierung genau diese Vorgabe erfüllt!

## 2 Korrektheit - Einzelabgabe

Überprüfen Sie die Korrektheit Ihrer Implementierung mit Hilfe der Referenz-Daten in

`/home/proj/biosoft/praktikum/genprakt-ws13/assignment1/out`

Da die optimalen Alignments nicht eindeutig sein müssen, kann die Korrektheit des Tracebacks nicht durch einen einfachen Vergleich mit der Ausgabe der Referenzimplementierung durchgeführt werden. Die Check-Option soll daher aus der Repräsentation des Alignments selbst den score berechnen und mit dem Eintrag in der Gotoh-Matrix vergleichen. Falsche Alignments müssen in eine Log-Datei geschrieben werden.

**Die schnellste korrekte Gotoh-Implementierung bekommt einen Buch-Preis.**

**Viel Erfolg.**