

# Übungen zur Vorlesung Sequenzanalyse II

Universität Bielefeld, SoSe 2008

Prof. Dr. Jens Stoye · Dipl.-Inform. Peter Husemann

<http://gi.cebitec.uni-bielefeld.de/teaching/2008summer/sequenzanalyse>

**Blatt 4 vom 09.05.2008**

**Abgabe in einer Woche zu Beginn der Vorlesung.**

Gegeben sind die vier Sequenzen  $s_1 = \text{ACCA}$ ,  $s_2 = \text{TCCA}$ ,  $s_3 = \text{ATC}$ ,  $s_4 = \text{ATCCA}$ , sowie linear-homogene Gapkosten von 2 und die folgende Substitutions-Scorematrix:

	A	C	G	T
A	3	-2	-1	-2
C	-2	5	-2	-1
G	-1	-2	3	-2
T	-2	-1	-2	5

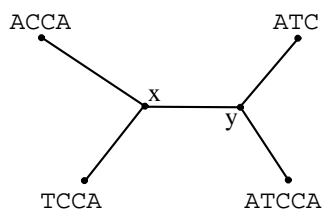
## Aufgabe 1 Sum-of-Pairs Multiple Sequence Alignment

Berechne den *Sum-of-Pairs-Score* des folgenden multiplen Alignments:

```
A - C C A
- T C C A
A T C - -
A T C C A
```

## Aufgabe 2 Baumalignment

Desweiteren sei ein ungewurzelter Baum gegeben:



Ermittle sinnvolle, idealerweise scoremaximierende, Sequenzen für die gemeinsamen Vorfahren  $x$  von  $s_1$  und  $s_2$ , sowie  $y$  von  $s_3$  und  $s_4$ . Berechne danach den Baum-Alignment-Score.

**Aufgabe 3** Bei einem Alignment von Aminosäure-Sequenzen kann die **PAM250** Ähnlichkeitsmatrix benutzt werden. Warum ist diese Score-Matrix nicht in allen Schritten eines Baumalignments empfehlenswert?