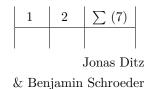
Bioinformatics I

WS 15/16

Tutor: Alexander Seitz



Assignment 6

(Handed in 23. November 2015)

Theoretical Assignment - Sequence-profile alignment and expected patterns in sequences

The profile as a PSWM for the given MSA would look like table 1.

Table 1: PSWM of the given MSA

			p_3		p_5
A	$0.\overline{3}$	0	$0.\overline{3}$	0	0
\mathbf{C}	0	0	0	1	0
\mathbf{G}	0	0	0	0	1
${f T}$	٥	1	$0.\overline{6}$	0	0
-	$0.\overline{6}$	0	0.0	0	0

Using this PSWM we can now compute an optimal semiglobal alignment of our profile with the sequence A = CATTCCGTTC. First we calculate the scoring matrix using as a scoring function s(a,b) = -1, s(a,a) = 3 and d = 2:

	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}
	С	A	${ m T}$	${\bf T}$	\mathbf{C}	\mathbf{C}	G	${ m T}$	${ m T}$	\mathbf{C}
p_1	$-1.\overline{6}$	$-0.\overline{3}$	$-1.\overline{6}$							
p_2	-1	-1	3	3	-1	-1	-1	3	3	-1
p_3	-1 -1	$0.\overline{3}$	$1.\overline{6}$	$1.\overline{6}$	-1	-1	-1	$1.\overline{6}$	$1.\overline{6}$	-1
p_4	3	-1	-1	-1	3	3	-1	-1	-1	3
p_5	-1	-1	-1	-1	-1	-1	3	-1	-1	-1

Now we fill the DP matrix using that scoring matrix:

One can see that the optimal alignment (colored in red) is:

Theoretical Assignment - Practice writing an introduction / background for a paper