Exercise 1 - Needleman-Wunsch algorithm

The Needleman-Wunsch algorithm enables the calculation of the optimal pairwise sequence alignment with linear gap cost. Given the following two sequences $S_1, S_2$ and the cost function in Eq. (1):

$$S_1 = \text{TCCGA}$$
$$S_2 = \text{TACGCGC}$$

$$w(x, y) = \begin{cases} 
+1 & \text{if } x = '-', \lor y = '-' \\
-1 & \text{if } x = y \\
0 & \text{else}
\end{cases}$$  \hspace{1cm} (1)

a) Calculate the distance matrix $D$, the optimal alignment score, and the optimal alignment for $S_1, S_2$.

b) Calculate the optimal alignment score while assuming that the first G character in each sequence has to be matched/aligned! Compare the resulting alignment with the optimal solution.
c) Why is this gap-scoring scheme named *linear gap cost*? What is the advantage of such a scoring scheme?

d) Implement Needleman-Wunsch in a spreadsheet. Here is the link to an editable google spreadsheet with some instructions: [http://goo.gl/MVCD11](http://goo.gl/MVCD11)

**Exercise 2 - Hirschberg recursion**

Starting from the pseudocode for the Hirschberg algorithm from the lecture for computing the optimal (prefix) alignment score in 2n space (two arrays each of length n), develop the pseudocode to compute the same in using only one array of length n and an additional variable (as shortly discussed within the lecture).

**Exercise 3 - Hirschberg algorithm (homework for next session)**

The Hirschberg algorithm solves the global sequence alignment problem in linear space by employing a divide and conquer approach. The two sequences $S_1$ and $S_2$ and the cost function are given in Eq. 2.

\[
S_1 = \text{TACGCGC} \quad \quad \quad S_2 = \text{TCCGA}
\]

\[
w(x, y) = \begin{cases} 
+1 & \text{if } x = '-' \lor y = '-' \\
-1 & \text{if } x = y \\
0 & \text{else}
\end{cases} \quad (2)
\]

a) Compute in linear space in which cell the traceback intersects row 3 of the dynamic programming matrix.

b) Which parts of the dynamic programming matrix are considered in the next step of the Hirschberg algorithm?