Exercise sheet 3: Sequence Alignment

Exercise 1 - Alignments

For the given examples which ones can be called alignments:

1a)

```
AGTTTTTT
AGGTTTTT
```

Hide

Solution   True

1b)

```
CCGTTTTTT
-AGGTTTTT
```

Hide

Solution   True

1c)

```
CCCGTTTTTTGC
-CGGTTTTT
```

Hide

Solution False. In an alignment both strings must have the same length

1d)

```
AG--TTTTTT
AG-GTTTTTT
```

Hide

Solution False. Gaps cannot be aligned with gaps

Exercise 2 - Hands on 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 S1, S2 and the given cost function complete the tasks A-D.

$$S1 = TACCGCGC \tag{1}$$
$$S2 = TCCGA \tag{2}$$

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

2a)

Complete the provided table with the correct initialization step.

Hide

Hint1: Formulae

$$D_{0,0} = 0 \tag{3}$$

$$\forall i \leq |S1| : D_{i,0} = \sum_{k=0}^{i} w(S1_i, -) \tag{4}$$

$$\forall j \leq |S2| : D_{0,j} = \sum_{k=0}^{j} w(-, S2_j) \tag{5}$$

Solution

2b)

Hide   Using dynamic programming technique fill in all values in the matrix.

Hint1: Formulae

$$\forall\, i,j > 0 \; : D_{i,j} = min \begin{cases} D_{i-1,j-1} + w(S1_i, S2_j) \\ D_{i,j-1} + w(-, S2_j) \\ D_{i-1,j} + w(S1_i, -) \end{cases}$$

Solution

2c)

Using the matrix from 1B) find the optimal alignment of the given sequences.

Hide

Solution

```
TACGCGC
| | ||:
T-C-CGA
```

2d)

Find an optimal alignment of the given sequences, while assuming that the first G character in each sequence has to be matched/aligned.

Hide

Hint1: You can split the alignment at the G and treat it as two separate alignments

Solution

```
TACGCGC     TACGCGC     TACGCGC
|:|||       |:||  :     |:||   :
TCCGA--     TCCG-A-     TCCG--A
```

Exercise 3 - Hirschberg recursion

Which statements about Needleman-Wunsch and the Hirschberg recursion are True and which are False.

3a)

Statements

☐ Hirschberg computes global alignment in $O(n^2)$ space.
☐ Needleman-Wunsch computes global alignment in $O(n^2)$ time.
☐ The Hirschberg recursion is a space optimized version of the Needleman-Wunsch algorithm.
☐ Hirschberg computes global alignment in $O(n)$ time.

Solution

☐ Hirschberg computes global alignment in $O(n^2)$ space.
☒ Needleman-Wunsch computes global alignment in $O(n^2)$ time.
☒ The Hirschberg recursion is a space optimized version of the Needleman-Wunsch algorithm.
☐ Hirschberg computes global alignment in $O(n)$ time.

Exercise 4 - Programming assignment

Programming assignments are available via Github Classroom and contain automatic tests.

We recommend doing these assignments since they will help you to further understand this topic.

Access the Github Classroom link: Programming Assignment: Sheet 03.