# Exercise sheet 2: Edit operations and alignments

\_\_\_\_\_

### Exercise sheet 2 - Levenshtein Distance

Compute the minimal Levenshtein edit distance for the following pairs of sequences.

#### Question 1A

$$S_1 = A \tag{1}$$

$$S_2 = T (2)$$

 $\mathbf{Hint} \quad A \to T$ 

Correct Answer  $A \rightarrow T = 1$ 

Question 1B

$$S_1 = AGATATA \tag{3}$$

$$S_2 = TATATATA \tag{4}$$

 $\mathbf{Hint} \quad \mathrm{AGATATA} \rightarrow \mathrm{ATATATA} \rightarrow \dots$ 

 $\textbf{Correct Answer} \quad \text{AGATATA} \rightarrow \text{ATATATA} \rightarrow \text{TATATATA} = 3$ 

Question 1C

$$S_1 = AGTCCT (5)$$

$$S_2 = CGCTCA \tag{6}$$

 $\mathbf{Hint} \quad \mathrm{AGTCCT} \rightarrow \mathrm{AGCTCA} \rightarrow \dots$ 

 $\textbf{Correct Answer} \quad \text{AGTCCT} \rightarrow \text{CGTCCT} \rightarrow \text{CGCCCT} \rightarrow \text{CGCTCA} = 4$ 

#### Question 1D

$$S_1 = TGCATAT (7)$$

$$S_2 = ATCCGAT \tag{8}$$

 $\mathbf{Hint} \quad \mathrm{TGCATAT} \rightarrow \mathrm{AGCATAT} \rightarrow \dots$ 

 $\textbf{Correct Answer} \quad \text{TGCATAT} \rightarrow \text{AGCATAT} \rightarrow \text{ATCATAT} \rightarrow \text{ATCCGAT} = 4$ 

#### Question 1E

$$S_1 = ACGTATATAGCCCCGCG (9)$$

$$S_2 = ACGTTATATAGCCGCGC (10)$$

Hint You need to use all the possible operations

 $ACGTATATAGCCCCGCG \rightarrow ACGTTATATAGCCCCGCG \rightarrow \dots$ 

Correct Answer ACGTATATAGCCCCGCG  $\rightarrow$  ACGTTATATAGCCCCGCG  $\rightarrow$  ACGTTATATAGCCCGCGCG  $\rightarrow$  ACGTTATATAGCCGCGCC = 4

### Exercise 2 - Metric function

Check if the corresponding functions are metric.

#### Question 2A

$$w(x,y) = x - y \tag{11}$$

**Hint** What if x = -5 and y = -5?

Correct Answer Not metric

#### Question 2B

$$w(x,y) = |x - y| \tag{12}$$

Hint You need to check all the properties.

Correct Answer Metric

#### Question 2C

$$w(x,y) = x + y \tag{13}$$

**Hint** What if x = -5 and y = -5?

Correct Answer Not metric

#### Question 2D

$$w(x,y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{else} \end{cases}$$
 (14)

**Hint** You need to check all the properties.

Correct Answer Metric

## Exercise 3 - DNA and RNA