# WST540: Exercise 5

**Problem 1.** Let $s =$ `solver` and $t =$ `lovely`. Answer the following questions.

(i) Recall that, to compute the edit distance between $s$ and $t$, we learned a dynamic programming algorithm which works by filling in a 2d array $A$, such that $A[i, j]$ $(0 \le i, j \le 6)$ equals the edit distance between $s[1..i]$ and $t[1..j]$. Give the entire $A$ in its final form.

(ii) Remember that each cell in $A$ is *determined* by at least one other cell (where the notion of "determine" is as defined in Lecture 13). Give all the cells that can determine $A[3, 4]$. Repeat this for $A[2, 5]$ and $A[4, 3]$.

(iii) Give a trace for $s$ and $t$ that corresponds to an editing path that changes $s$ to $t$ with the minimum operations. Also explain what are these operations.

**Solution.** (i)

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 1 | 1 | 2 | 3 | 4 | 5 | 6 |
| 2 | 2 | 2 | 1 | 2 | 3 | 4 | 5 |
| 3 | 3 | 2 | 2 | 2 | 3 | 3 | 4 |
| 4 | 4 | 3 | 3 | 2 | 3 | 4 | 4 |
| 5 | 5 | 4 | 4 | 3 | 2 | 3 | 4 |
| 6 | 6 | 5 | 5 | 4 | 3 | 3 | 4 |

(ii) $A[3, 4]$ can be determined by $A[2, 3]$ or $A[3, 3]$. $A[2, 5]$ is determined by $A[2, 4]$. $A[4, 3]$ is determined by $A[3, 2]$.

(iii) Trace: $\{(1, 1), (2, 2), (4, 3), (5, 4), (6, 6)\}$. Operations: substitute `s` with `l`, remove `l`, insert `l`, and substitute `r` with `y`.

**Problem 2.** Let $s$ and $t$ be as defined in Problem 1. Suppose that we only want to verify whether the edit distance between $s$ and $t$ is greater than 1. Give the values of the cells of $A$ that need to be computed by the algorithm in Lecture 14.

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 1 | 2 | - | - | - | - |
| 1 | 1 | 1 | 2 | 2 | - | - | - |
| 2 | 2 | 2 | 1 | 2 | 2 | - | - |
| 3 | - | 2 | 2 | 2 | 2 | 2 | - |
| 4 | - | - | 2 | 3 | 3 | 3 | 2 |
| 5 | - | - | - | 2 | 3 | 4 | 3 |
| 6 | - | - | - | - | 2 | 3 | 4 |

**Problem 3.** Let $s =$ `father` and $t =$ `feather`. Answer the following questions for $q = 3$:

(i) List all the positional $q$-grams of $s$ and $t$.

(ii) Let $d = 1$. Give the number of positional $q$-grams of $t$ that $d$-match at least one positional $q$-gram of $s$. List those $q$-grams of $t$.

**Solution.** (i)
For $s$: $\{(1, \texttt{fat}), (2, \texttt{ath}), (3, \texttt{the}), (4, \texttt{her}), (5, \texttt{er\#}), (6, \texttt{r\#\#})\}$.
For $t$: $\{(1, \texttt{fea}), (2, \texttt{eat}), (3, \texttt{ath}), (4, \texttt{the}), (5, \texttt{her}), (6, \texttt{er\#}), (7, \texttt{r\#\#})\}$

(ii) Five: $(3, \texttt{ath}), (4, \texttt{the}), (5, \texttt{her}), (6, \texttt{er\#}), (7, \texttt{r\#\#})$.

**Problem 4.** Let $s$ be a string of length 6, and $t$ a string of length 7. Fix $d = 2$ and $q = 2$. Let $x$ be the number of positional $q$-grams of $s$ that $d$-match at least one positional $q$-gram of $t$. Answer the following questions.

(i) If $x = 2$, can the edit distance of $s$ and $t$ be at most $d$? If not, explain why; otherwise, justify your answer with an example of $s$ and $t$.

(ii) If $x = 3$, can the edit distance of $s$ and $t$ be at most $d$? If not, explain why; otherwise, justify your answer with an example of $s$ and $t$.

**Solution.** (i) No. This is because, by the Lemma in Lecture 15, $x$ must be at least $7 - 2 \times 2 = 3$.

(ii) Yes. Here is an example: $s = \texttt{aaaaaa}, t = \texttt{abaabaa}$.