1 Problem 2

1.1 Algorithm

1. Array $A[0...m, 0...n]$
2. Initialize $A[i, 0] = i\delta$ for each i
3. Initialize $A[0, j] = j\delta$ for each j
4. For $j = 1, ...n$
   - For $i = 1, ..., m$
     - Use the recurrence relation to compute $A[i, j]$
   - Endfor
5. Endfor
6. Return $A[m, n]$

Recurrence Relation The minimum alignment cost satisfy the following recurrence for $i \geq 1$ and $j \geq 1$:

$$OPT(i,j) = min \begin{cases} \alpha_{x_i y_j} + OPT(i-1, j-1) & (1) \\ \delta + OPT(i, j-1) & (2) \\ \delta + OPT(i-1, j) & (3) \\ \beta_{x_i y_j} + OPT(i-2, j-2) & (4) \end{cases}$$

in which, given an optimal alignment M,
(1) : $(i, j) \in M$ and we pay $\alpha_{x_i y_j}$ and then align $x_1...x_{i1}$ as well as possible with $y_1...y_{j1}$
(2) : j^{th} position of Y is not matched and we pay δ and then align $x_1...x_i$ as well as possible with $y_1...y_{j1}$
(3) : i^{th} position of X is not matched and we pay δ and then align $x_1...x_{i-1}$ as well as possible with $y_1...y_j$
(4) : we swap $x_{i-1}x_i$ to match $y_{j-1}y_j$ (or vice versa) and we pay $\beta_{x_i y_j}$ and then align $x_1...x_{i-1}$ as well as possible with $y_1...y_j$

1.2 Runtime

Initializing matrix A takes $O(mn)$ time as the OPT calculation takes $O(1)$ time with a total of $O(mn)$ iterations to fill the matrix. Thus, overall running time is $O(mn)$.

1.3 Correctness

Lemma 1. *Let $f(i, j)$ denote the minimum cost of aligning strings X and Y. Then for all i, j, we have $f(i, j) = OPT(i, j)$.*

Proof of Lemma 1. We will prove Lemma 1 using a proof by induction.
If $i = j = 0$, then $f(0, 0) = OPT(0, 0) = 0$.

Now consider arbitrary values of i and j, and suppose the lemma is true for all pairs (i', j') with $i' + j' < i + j$. The last character of the previous subalignment from (i, j) is either $(i-2, j-2), (i-1, j-1), (i, j-1),$ or $(i-1, j-1)$. Thus, we have

$$f(i,j) = min \begin{cases} \alpha_{x_i y_i} + f(i-1, j-1) \\ \delta + f(i, j-1) \\ \delta + f(i-1, j) \\ \beta_{x_i y_i} + f(i-2, j-2) \end{cases}$$

$$= min \begin{cases} \alpha_{x_i y_i} + OPT(i-1, j-1) \\ \delta + OPT(i, j-1) \\ \delta + OPT(i-1, j) \\ \beta_{x_i y_i} + OPT(i-2, j-2) \end{cases}$$

$$= OPT(i, j).$$

In which we pass from the first to the second line using the induction hypothesis, and the definition of the recurrence relation to from the second to the third line. Hence, we have proved Lemma 1 as needed.