

# 1. Modified Asymmetric Edit Distance [10 points]

For this problem, we're going to redefine edit distance as the smallest sum of penalties for the operations below to convert one string into another but our definition won't be symmetric since inserting a character will be less of a penalty than deleting a character:

- 1) • 3 points for deleting a character from the first string  $x$ . *analogous to adding a gap to  $y$*
- 2) • 2 points for inserting a character into the first string  $x$ . *analogous to adding a gap to  $x$*
- 3) • 3 points for inserting two characters into  $x$  in consecutive locations. *analogous to adding 2 consecutive gaps to  $x$*
- 4) • 4 points for substituting a character. *analogous to mismatch*

(a) Update the recurrence to calculate the new edit distance

(b) Build the table to evaluate the recurrence and state the distance from the string "pair" to the string "spared".

a) New recurrence:

$$OPT(i, j) = \begin{cases} \infty & \text{if } i < 0 \text{ or } j < 0 \\ 3\lfloor j/2 \rfloor + 2\lfloor j/2 \rfloor & \text{if } i = 0 \text{ ie: no more letters remaining in } x \\ 3i & \text{if } j = 0 \text{ ie: no more letters remaining in } y \\ \min \begin{cases} OPT(i-1, j-1) + 4 & \text{operation ④} \\ OPT(i-1, j-1) + 0 & \text{if } x_i = y_i \\ OPT(i-1, j) + 3 & \text{operation ①} \\ OPT(i, j-1) + 2 & \text{operation ②} \\ OPT(i, j-2) + 3 & \text{operation ③} \end{cases} \end{cases}$$

*note: can't have 2 consecutive gaps in  $x$  in this case (out of bounds)*

if  $i < 0$  or  $j < 0$

if  $i = 0$  ie: no more letters remaining in  $x$   
if  $j = 0$  ie: no more letters remaining in  $y$  } base cases

operation ④

if  $x_i = y_i$

operation ①

operation ②

operation ③

b) "pair" → "spared"

$x = \text{"pair"}$     $y = \text{"spared"}$

$2 + 3 + 3 = 8$   
pair  
spared

Distance = 8

j	0	1	2	3	4	5	6
i		s	p	a	r	e	d
0	0	2	3	5	6	8	9
1 p	3	4	2	4	5	7	8
2 a	6	7	5	2	4	5	7
3 i	9	10	8	5	6	8	9
4 r	12	13	11	8	5	7	8