CSC-349 Design and Analysis of Algorithms
Programming Assignment 3

When a new gene is discovered, a standard approach to understanding its function is to look through a database of known genes and find close matches. The closeness of two genes is measured by the extent to which they are aligned. To formalize this, think of a gene as being a long string over an alphabet $\Sigma = A,C,G,T$.Consider two genes (strings) x = ATGCC and y = TACGCA. A - alignment of x and y is a way of matching up these two strings by writing them in columns, for instance:

```
- & A & T & - & G & C & C
T & A & - & C & G & C & A
```

Here the - indicates a gap. The characters of each string must appear in order, and each column must contain a character from at least one of the strings. The score of an alignment is specified by a scoring matrix δ of size $|\Sigma|+1 \times |\Sigma|+1$, where the extra row and column are to accommodate gaps.

For instance, the preceding alignment has the following score: δ(-,T) + δ(A,A) + δ(T,-)+ δ(-,C), + δ(G,G) + δ(C,C) + δ(C,A). Give a dynamic programming algorithm that takes as input two strings x and y and a scoring matrix δ, and returns the highest scoring alignment. Reconstruct the solution to find the optimal alignment.

Example scoring matrix:

	A	**C**	**G**	**T**	**-**
A	2	-1	-1	-1	-2
C	-1	2	-1	-1	-2
G	-1	-1	2	-1	-2
T	-1	-1	-1	2	-2
-	-2	-1	-2	-2	0

Example score (not the optimal solution):

δ(-,T) + δ(A,A) + δ(T,-)+ δ(-,C), + δ(G,G) + δ(C,C) + δ(C,A) = -2 + 2 + -2 + -1 + 2 + 2 + -1 = 0

The grading criteria for this lab are:

The implementation **correctly** solves the task for all test cases. The algorithm uses **dynamic programming strategy** and follows the best coding practices.	**100%**
The implementation **correctly** solves the task for all test cases but does not use dynamic programming strategy.	**60%**
The implementation solves the task but has several failing edge cases.	**50%**
The implementation attempts to solve the task but fails in many cases.	**30%**
The implementation does not correctly solve the task.	**15%**
No attempt was made.	**0%**