

# Project Proposal

## Sequence Alignment Learning Tool

Tanzhen Li(tli67)      Chenliang Li(cli94)

---

### **Brief Motivation**

Sequence alignments might be difficult for beginners to understand in the bioinformatics field. We have personally struggled to fully understand the procedures for the variations of sequence alignment(Global alignment, Local alignment, Fitting alignment). Thus, we believe an interactive learning tool which guides the user through detailed steps would be beneficial.

### **Tool Description**

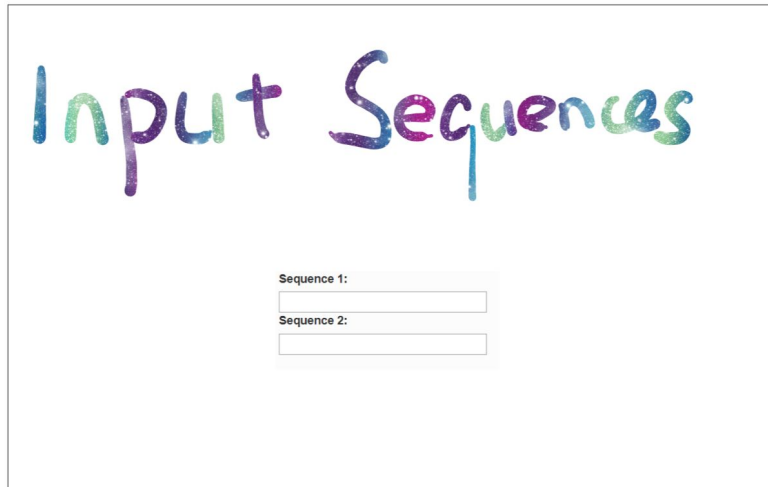
This tool is a web tool which interactively gives the user step by step tutorial based on the alignment model and the sequence inputted by the user. It asks the user to complete steps such as initializing dp-table, choosing table filling direction, actual table filling, score calculation and backtracking. The user can only proceed to the next step if she correctly completes the current step. Each steps completion is verified by our own algorithm implementation.

---

---

## Example Walkthrough

As the user enters the tool, it asks for the sequences the user wishes to align as illustrated below:



The screenshot shows a web form titled "Input Sequences" in a colorful, hand-drawn font. Below the title, there are two input fields. The first is labeled "Sequence 1:" and the second is labeled "Sequence 2:". Both fields are empty and have a light gray border.

Then the tool asks for the scoring matrix

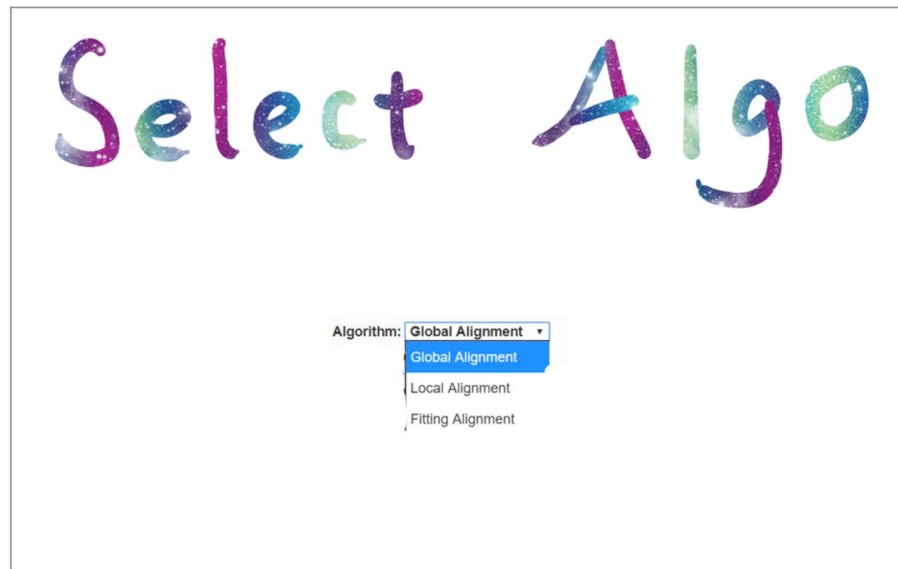


The screenshot shows a web form titled "Input Scoring Matrix" in a colorful, hand-drawn font. Below the title, there is a table for the scoring matrix. The table has four columns labeled A, T, C, and G, and four rows labeled A, T, C, and G. The values in the table are as follows:

	A	T	C	G
A	1	0	0	0
T	0	1	0	0
C	0	0	1	0
G	0	0	0	1

Below the table, there is a label "Gap penalty:" followed by an input field containing the value "0".

Then the tool asks for the algorithm the user wish to learn(global, local, fitting)



Then the tool asks the user to select the current recurrence which fits the algorithm the user chose before

STEP1: Choose the correct Recurrence for Global Alignment

$$s[i, j] = \max \begin{cases} 0, & \text{if } i = 0 \text{ and } j = 0, \\ s[i-1, j] + \delta(v_i, -), & \text{if } i > 0, \\ s[i, j-1] + \delta(-, w_j), & \text{if } j > 0, \\ s[i-1, j-1] + \delta(v_i, w_j), & \text{if } i > 0 \text{ and } j > 0. \end{cases}$$

deletion

insertion

match/  
mismatch

$$s[i, j] = \max \begin{cases} 0, & \text{if } i = 0, \\ s[i-1, j] + \delta(v_i, -), & \text{if } i > 0, \\ s[i, j-1] + \delta(-, w_j), & \text{if } i > 0 \text{ and } j > 0, \\ s[i-1, j-1] + \delta(v_i, w_j), & \text{if } i > 0 \text{ and } j > 0. \end{cases}$$

$$s^* = \max\{s[m, 0], \dots, s[m, n]\}$$

$$s[i, j] = \max \begin{cases} 0, & \text{if } i = 0 \text{ and } j = 0, \\ s[i-1, j] + \delta(v_i, -), & \text{if } i > 0, \\ s[i, j-1] + \delta(-, w_j), & \text{if } j > 0, \\ s[i-1, j-1] + \delta(v_i, w_j), & \text{if } i > 0 \text{ and } j > 0. \end{cases}$$

$$s^* = \max_{i,j} s[i, j]$$

Then the tool asks the user to initialize the DP-Table

STEP2: Initialize DP-Table. Fill in the blue blanks

Then the tool asks the user to select the direction to fill in the table

STEP3: Select the direction to fill in the table according to the recurrence

$$s[i, j] = \max \begin{cases} 0, & \text{if } i = 0 \text{ and } j = 0, \\ s[i - 1, j] + \delta(v_i, -), & \text{if } i > 0, \\ s[i, j - 1] + \delta(-, w_j), & \text{if } j > 0, \\ s[i - 1, j - 1] + \delta(v_i, w_j), & \text{if } i > 0 \text{ and } j > 0. \end{cases}$$

deletion  
insertion  
match/  
mismatch

---

Then the tool asks the user to fill the table

STEP4: Complete the table

		C	A	G	C
		0	0	0	0
C	0				
A	0				
A	0				
C	0				
C	0				

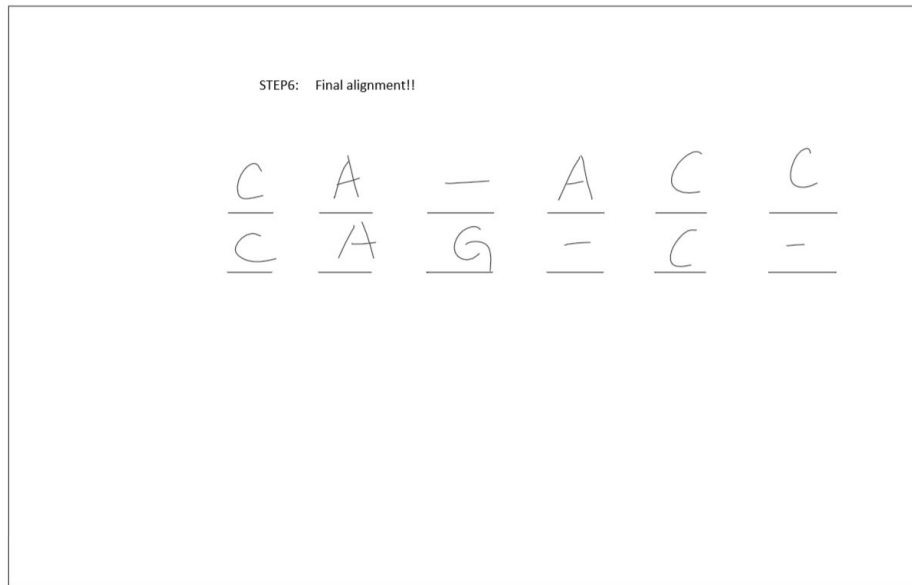
Then the tool asks the user to backtrace

STEP5: Backtrace, highlight the steps

		C	A	G	C
		0	0	0	0
C	0	1	1	1	1
A	0	1	2	2	2
A	0	1	2	2	2
C	0	1	2	2	2
C	0	1	2	2	2

---

Finally, it asks th user to give the alignment result.



Each step will have detailed instructions.

## Datasets

Our tool does not require any dataset. The user inputs the data on her own and the tool process those data to create the tutorial.

---

## Planned method/experimentations

We plan to use React as the framework for the web development. React is one of the most popular front-end languages in the industry and it works great with GitHub pages for demoing purpose.

## Project time line

TIME	TASK
11/20	Complete design of each process
11/25	Algorithm implementation
11/30	Interface implementation
12/5	Configuration & Testing
12/14	Project Due