

# Sequence Alignment Algorithm Visualization

## Notes for Developers

By Hannah Coleman

## Table of Contents

<b>General Advice</b>	<b>2</b>
<b>Remaining Tasks</b>	<b>2</b>
Multiple Paths	2
Definition of Current Coordinates	2
Max Formula Display	2
Alignment Algorithms	2
Score Matrix Controller	3
Cell Element Positioning	4
Show/Hide Arrows Toggle	4
Screen Resolution	4

Any questions about my contribution to this project may be sent to me at [hbsc@uw.edu](mailto:hbsc@uw.edu), and I will do my best to answer them.

## General Advice

I would like to pass along some advice that might help you save some time and energy.

Firstly, be sure to keep a weekly meeting with Dr. Kim. I recommend against spacing meetings further apart than that. The meetings are extremely valuable because Dr. Kim can help guide you in the right direction.

Second, requirements gathering is one of your responsibilities. Even if a client is more of an expert in your field than you are, they will not necessarily tell you everything they want; you have to ask them. When you notice you could approach a problem more than one way, it is a good idea to ask Dr. Kim if she has an opinion.

Third, try to avoid creating multiple bugs in the same part of the program, as this makes the debugging process more complicated. Testing the program incrementally or creating small side programs to test ideas can help.

## Remaining Tasks

### **Multiple Paths**

Currently, if there is a tie for max score when a cell is to be filled, tiebreakers are applied immediately, in the grid-filling phase. This results in only one path direction per cell. A future version of the program should allow multiple arrows in the event of a tie; tiebreakers should then be applied in the tracing phase.

### **Definition of Current Coordinates**

Currently, a yellow inner border marks the cell that is about to be acted on. This is because each step consists of acting on a cell and then modifying `DPModel.x` and `DPModel.y`, and the yellow marks those “current” coordinates. While this is not exactly wrong, it might be better to mark the cell that has just been acted on, instead. Dr. Kim may have an opinion about this.

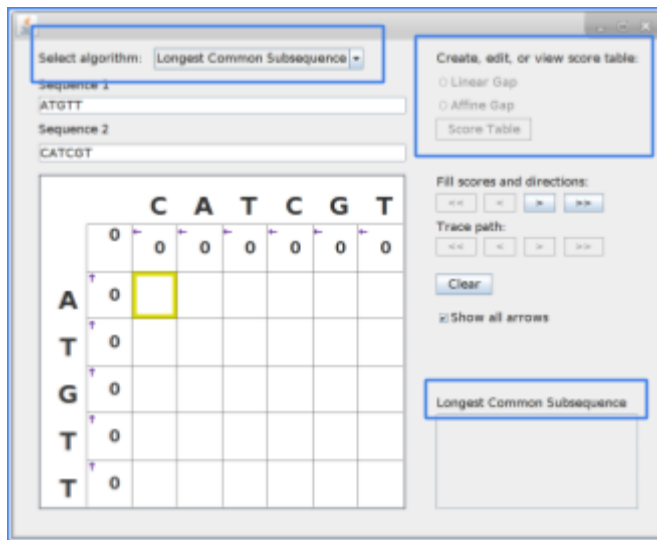
### **Max Formula Display**

The program should display the max formula for the current cell in the filling phase and bold (or otherwise emphasize) the line in the max formula that applies to the current cell. See B.A.B.A. for one approach to this.

### **Alignment Algorithms**

At minimum, the program must have algorithms for global and local alignment in addition to the LCS algorithm. Selecting “Global Alignment” or “Local Alignment” in the menu must enable the controls for the score-table settings; selecting “Longest Common Subsequence” must disable these controls. (Alternatives: add/remove the controls or, if possible, make the controls visible/invisible.) In addition, selecting “Global Alignment” or “Local Alignment” should make

the label above the alignment view read “Alignment”; selecting “Longest Common Subsequence” should make the label read, “Longest Common Subsequence”.



## Score Matrix Controller

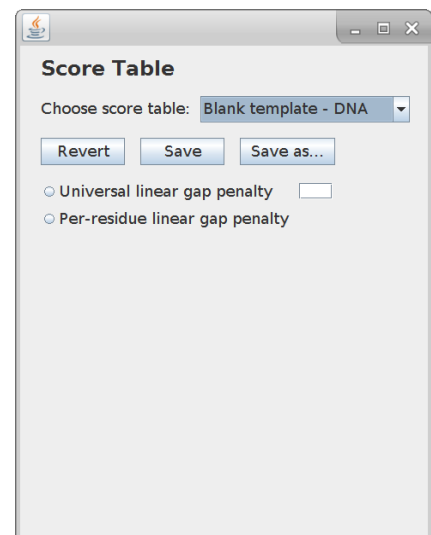
Controls for score-matrix and gap-penalty settings have been partially designed, but the design will need revision.

Selecting the “Linear Gap” radio button and clicking the “Score Table” button in the main window should bring up a separate window that allows for editing the gap penalty and score table. The program should have a menu for selecting a score table and some options available, stored in a folder with the program. (I have not added those options yet.)

Blank/all-zero templates for DNA and protein score tables should also be available; these are for the user to make a score table from scratch.

The program should allow the user to set a single, universal linear gap penalty or a custom linear gap penalty for each residue. For the latter, the program should provide an extra row in the score table to add the gap penalties.

A grid with the score table should be displayed according to the selection made in the menu. The cells should be text fields so they can be edited. The program should allow the user to save changes to a new file (“Save as”), save changes to a current file if the user did not start from a blank template (“Save”), or revert to the last saved version (“Revert”). If the user presses the “Save” button, the program should provide a popup dialog message asking if the user really wishes to overwrite the previous file. (Another option may be to disable the “Save” button for certain cases, such as for the original score tables provided with the program, making it impossible to write over these.)



Selecting the “Affine Gap” radio button from the main controller should allow the user to set a gap-opening penalty and a gap-extension penalty. The settings for this are currently in a separate window, but they can be placed in the main window.

### Cell Element Positioning

The inner borders of the current cell (yellow) and the cells in the traced path (green) displace the score and arrow. The absence of an arrow also displaces the score in the cell in the upper left corner of the grid. Ideally, all cells’ scores and arrows should have the same relative positions.

		C	A
	0	0	0
A	0	0	1

In addition, the grid-view scroll pane is always the same size even if there are very few cells inside the grid, with the cells being larger. It is better if the cell size is constant and the scroll pane can adjust to be smaller.

### Show/Hide Arrows Toggle

A future version of the program may have an option to show or hide the arrows in the cells. Dr. Kim may have further specifications for this. I have already included a “Show arrows” checkbox. It can be removed if this functionality is not to be included in the final program.

### Screen Resolution

I designed the GUI to be slightly small but comfortable on a 2560x1440-pixel laptop screen. It is likely too large for lower resolutions. This could be fixed by detecting the device’s display dimensions, by making multiple versions, or by making some compromises to the sizing of GUI elements.