

## **Creating a User Interface for Quick and Easy Protein Analysis**

Joseph Moeller  
Hiram College

Computer technology today offers numerous ways for aiding users to accomplish what used to be tedious everyday tasks. These people can range from children who use computers for schoolwork to business workers relying on expensive software to make precise calculations. Fisch states that the “primary goal of the user interface is to maximize the user’s productivity and decrease the time the user must spend becoming acquainted with the system, or application” (Fisch, 1993). If the end user does not have a background in computer science, then chances are that they are in constant contact with a detailed interface to help them harness the power of the computer. Based on common knowledge, computer users in many areas of study, including statistics, psychology, accounting, and physics, rely on detailed interfaces to use the computer to their advantage.

The focus of this research is positioned on the simplification of protein data analysis for biologists. Many biologists have the ability to utilize free online protein analysis algorithms that are posted by several different research groups, including the Center for Biological Sequence Analysis. Although the process of using these web interfaces is rather simple, the common biologist does not always have a need for the massive data sets produced by these programs. To get rid of the unwanted data, the biologist may have to find a programmer to write simple scripts to parse out what he or she does not want. In addition to obtaining too much data, each type of protein analysis must be done separate from one another on a different web page, producing a new set of data for each analysis completed. As one can see, the process of obtaining this protein data is much more laborious than it needs to be.

With the main goal of simplifying protein data analysis, I propose to create a PHP-based website that will link several of these protein analysis algorithms together. Furthermore, the user will be able to select which types of data they would like returned to them for each analysis he or she has selected. The selected analyses will be run on the FASTA formatted input data file containing the protein sequences. The output of each selected analysis will be parsed to the user's specifications using PERL scripts and combined together with the output from each of the other selected analyses into MySQL for storage of the data into a central database. The MySQL database will allow for detailed sorting of the final set of data by the user. The data in the MySQL database, whether it has been sorted by the user or not, can be simply exported to other file types such as Microsoft Excel files or even plain Tab-Delimited files. These files will contain the well-organized protein data from all the different analyses that were selected on the initial PHP interface.

As compared to the current way of obtaining protein analysis data, the idea proposed here will provide a simpler, more user-friendly way for biologists to analyze protein data for their personal research purposes.

## **References**

Fisch, E. A. (1993, April). Understanding and Improving the User Interface Design Process. *Software Engineering Notes* , 27-29.