**Accelerating Biomolecular Nuclear Magnetic Resonance Assignment with A\***

Joel Venzke, Paxten Johnson, Rachel Davis, John Emmons, Katherine Roth, David Mascharka, Leah Robison, Timothy Urness (Mentor), Adina Kilpatrick (Mentor)

Department of Mathematics and Computer Science, Arts and Sciences

Nuclear magnetic resonance (NMR) spectroscopy is a method for studying the 3D structure of molecules such as proteins. The post-genomic era has created the need for information of unknown proteins encoded by discovered genes. Current manual techniques for analyzing NMR datasets can take days to months to assign and are error prone. The current goal of this research is to develop an algorithm to automate the process of assigning NMR datasets, minimize human error and accelerate this task.

The algorithm utilizes A\*search to sequence amino acids produced from NMR datasets of proteins. The program is given a set of amino acid nodes and each node is placed in the initial position of the protein chain. Unplaced nodes are sequentially added to the end of a list to create child nodes. The node list with the smallest error is then expanded. The error is a running total of how well the nodes fit the reference chain, coupled with the match accuracy of the placed child. This process of assigning child nodes and calculating costs repeats until every node is placed, and the best node list contains the lowest cost possible.

Assignments of test sets have shown this method can be used to assign NMR datasets, but need for improvement remains. Research is ongoing as we continue refining our algorithm. Future goals include adjusting for missing data and testing larger datasets. Upon completion, this algorithm will allow for great advancements in the areas of structural biology and proteomics.