Binding Site Accessibility With Motion Planning Guided By Annotated Skeletons

Nancy M. Amato[1,4], Diane Uwacu[1,4], Priscilla Saarah[2], Vasu Chalsani[1]

I. INTRODUCTION

Proteins are one of the most important biological molecules. Proteins interact with another group of molecules, ligands which can be ions, or other biomolecules like nuclei acids, and this interaction results in a change in the structure (conformation) and function of the protein. This change in form determines the cellular processes and how efficient they would be performed by the protein. As such, knowing where the binding site is, which atoms make up the binding site, and the ligands for a particular protein is helpful for understanding the actual intermolecular interactions that occur and also to look into diseases and medicines (2).

Research studying the interactions of proteins and ligands is continuously being pursued, especially computational methods to predict protein and ligand interaction. My mentors had worked and developed a motion planning algorithm - a “topological skeleton-guided rapidly-exploring random graph (RRG)” that focuses on finding paths to the binding site of a protein based on energy scores (1). This summer, I worked on a follow-up study to test the algorithm on more protein cases.

[1]Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL, USA
[2]Department of Biology, Dillard University, New Orleans, LA, USA
[3]Texas A&M University, College Station, Texas
[4]Parasol Lab

namato@illinois.edu
duwacu@illinois.edu
psaarah@dillard.edu
sc83@illinois.edu

II. METHODOLOGY

As this project is a continuation of the previous work by my mentors, most of the code is already written so I started off the summer with a crash course. I had a training on reading scientific papers, after which I read and had discussions on Probabilistic Roadmap (PRM) articles and also articles on the previous work on this project; then I run the ‘Box Folding Problem’ - a motion planning benchmark testing model available on the Parasol Lab’s website as a motion planning practice.

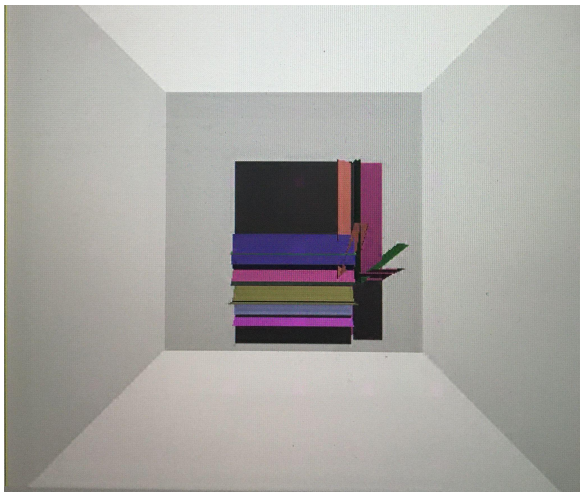

Figure 1: showing the visualization of the box folding motion planning problem.

I used the Parasol lab Motion Planning Library (PMPL), now known as the Parasol lab Planning Library (PPL) code to generate roadmaps and visualized them through Vizmo++.

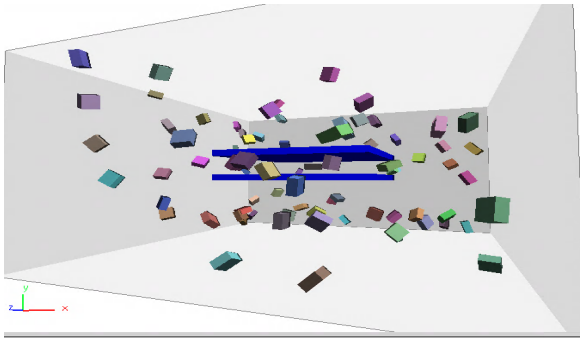

Figure 2: showing the nodes generated by one of the PRM samplers I run.

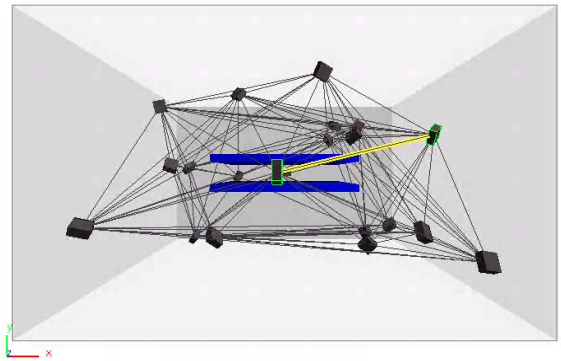

Figure 3: showing a generated nodes and roadmap after running a query

We got started on the project about the sixth week. The project has four stages: running and troubleshooting the existing code, designing data analysis and performance matrics, testing five new protein cases, experimenting and writing. I run the code on four tested proteins and have been fixing the issues and improving the code. I downloaded the protein and ligand data from the protein data bank (PDB) and used CHIMERA to obtain the protein geometric models in higher order structure; I translated the ligand data into pdb using a translator tool from the CADD Group Chemoinformatics Tools and User Services (CACTUS). I run the code to construct a planning workspace and configuration space, create the skeleton, build the roadmap.

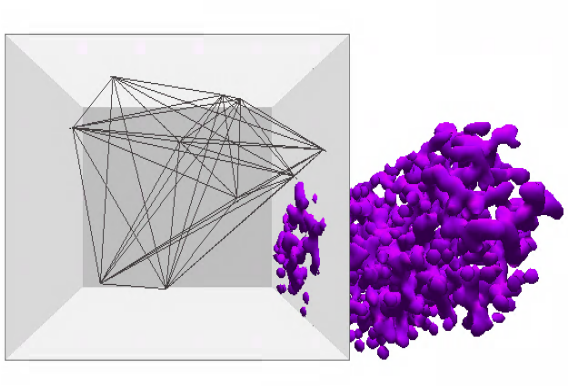

Figure 4: showing an example of the issues we had to fix: the protein was outside its environment**.**

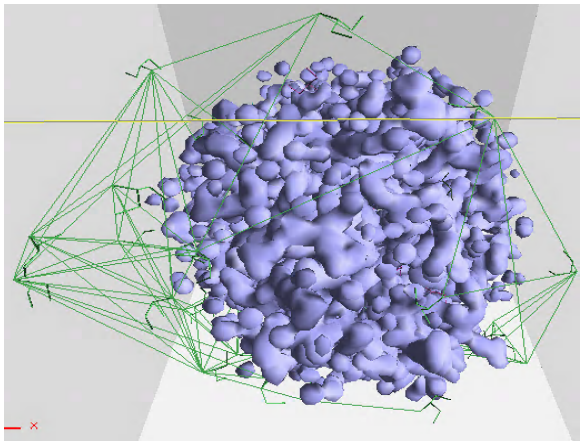

Figure 5: showing a protein, generated ligand configurations and roadmap.

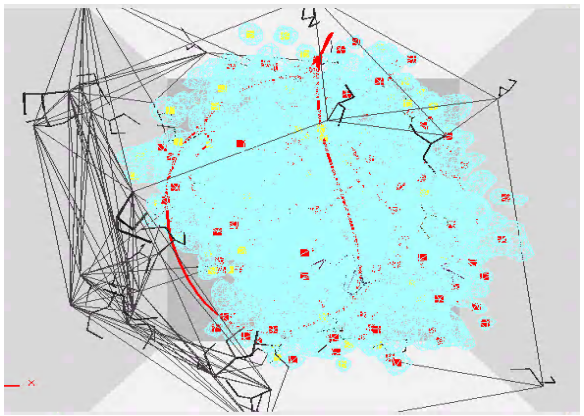

Figure 6: showing a generated skeleton (in red), ligand configurations and roadmap.

III. CONCLUSION

During the ten weeks I worked on this project, I had the opportunity to learn about applications of robotics in biology, which is very important for me as a biology major, motion planning and even programming. I interacted with PhD students who are working on cutting edge research and other amazing undergraduates from different places. I had much fun and learned a lot this summer that I am continuing to work on the project even though my time with DREU is over.

REFERENCES

1. Diane Uwacu, Abigail Ren, Shawna Thomas, and Nancy M. Amato. 2020. Using Guided Motion Planning to Study Binding Site Accessibility. In *Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (BCB '20), September 21-24, 2020, Virtual Event, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3388440.3414707

2. Zhao, J., Cao, Y., & Zhang, L. (2020). Exploring the computational methods for protein-ligand binding site prediction. *Computational and structural biotechnology journal, 18, 417-426. https://doi.org/10.1016/j.csbj.2020.02.008*