

Individual Final Project Reflection

3. III. Problem 3 [LO5, LO6 and LO7] - 15 pts [Individual work] Individual student software engineering reflection on their final project. For this part, every student must submit an individual PDF and place this file in the root of the student's individual repository. The student's reflection must cover the following points:

3.1. A description of their role in the project.

Originally, my role was role 2: compute services. After Bianca created our first version of our general purpose library, I implemented functions to update the cell states based on the definition of different states. I coordinated a meeting with Dr. Drummond to fix a segmentation fault error, but during the meeting we realized our general purpose library wasn't general enough since it used our genotype rules and implementations. After much input, I recreated our general purpose library following the setup instructions in the pdf along with Dr. Drummond's insight, so Bianca and I ended up switching roles. I created the new general purpose library with declarations in the header file 'CA_library.h' and their implementations in the source file 'CA_library.cpp'. After completing part 1 with my team I started creating the test file 'test_genotype.cpp' to get our CA model set up.

3.2. How much do they contribute to the successful project completion (e.g. lead work, help others, coordinated meetings, etc),

I would like to say I had a good contribution to the success of our project completion. When we realized our setup was not what was expected, Bianca was out of state so I prioritized meeting with Dr. Drummond to get a clear understanding of her part since it had to be fixed before any other steps could be completed. I had to recreate our general purpose library so that it was up to standards before the compute step could be done. I initiated the changes as soon as we found out and helped Bianca fix her part when she was unable. This ended up being successful because we got good feedback on it and Bianca was able to successfully complete her new compute part afterwards! We had a group chat that we would use to send updates for all the feedback we got on our own, changes we made to our repository, and other communication about our final project. I contributed also by coordinating Zoom meetings to fill Bianca and Korede in on our new general purpose library and other changes I made as well as coordinating meetings/emails with Dr. Drummond to check in on the progress of those changes. Lastly, I got everything started for part 2 of the final in regards to implementing our CA model as well as summarizing it in a shared Google doc with my team.

3.3. Describe any challenges/problems you and/or your team dealt with and how you solved them.

The first challenge our team faced was what kind of neighbors we would be using to implement the allele frequency in our CA model. We solved it by meeting with Dr. Drummond after class to discuss possible solutions which gave us a clearer understanding of the layout of our model.

Another challenge was the different time zones in our group. Korede lives 9 hours ahead of Bianca and me, so it was difficult finding times to meet via Zoom or to even reply to each other's text

messages in our group chat, and when Bianca was out of state she was on a different time as well. To solve this, we would mostly meet and message each other in the morning so that it wasn't too late for Korede, or we would simply work on our own parts on our own time while updating each other through text as we go. We had a very good communication system considering we were all on different time zones!

A problem I faced on my end was git related. Towards the end of our project, I was unable to pull or push from Github because my local branch was a commit behind the remote branch. To try to resolve this, I reached out to Dr. Drummond and the GSI's, and searched for resources online to help. Manish attempted to help via email, but I could not get it fixed so I ended up just sending Bianca my changes on my end so that she could paste them in our repository and make the push from her laptop.

Again, when I met with Dr. Drummond to go over my original compute functions, we realized our entire general purpose library was not actually general, which was a major problem considering we had to basically start over with only four days before the due date. We solved this problem by consulting with Dr. Drummond and redoing our setup so that Bianca could later do compute services and Korede the output. Considering all the difficulties we faced, our team worked extremely well together to fix them all!

3.4. Comment on the algorithmic and performance analysis of your library (you can base this on the data structures and functions you included). You do not need to provide a detail of every function that you implemented but rather provide a high level description of the design considerations that make you choose the data structures and algorithmic implementations that you used in this project. In particular, the parts of the project that you lead.

Our cellular automata library was created using particular data structures and algorithms that would make it easy and adaptable for users of the application. We utilized a 2D grid to represent the population of our simulation which allowed for an efficient visual print out of the genetic makeup of the population along with how it changed over each generation in accordance with the neighborhood system. We implemented it using a vector of vectors so that each cell would be easily accessible. To make sure a random set of integers were generated at the initialization of each setup, we utilized the 'std::rand()' function so that users start with a completely random grid. Users are allowed to define their own custom setup for their model, making it a general yet versatile application. This is done through methods like 'setup_dimension()', 'setup_boundaries', etc. so that users can lay out the details of their own model using our application. These design considerations and implementations make our general purpose library simple yet functional so that it is suitable for further research in the realm of genetic modeling.

My role in the process was the creation of the general purpose library and facilitating updating my team and setting up the next steps. I lead the design of our test_genotype.cpp file where we implemented our CA model using the general purpose library I created (with Bianca's compute functionality) as well as the creation of our FinalProject_Assessment.pdf.

3.5. Comment on how effective your showcase CA application from Problem 2 worked (i.e., how realistic was the model, any suggestions on how to improve its accuracy, etc)

Our showcase CA application effectively demonstrated a model of changes in allele frequency and genetic traits over a given period of time or generations. It was a good representation of genetic inheritance at a simpler level, with the initial grid setting up the genetic makeup of the population given a specific starting frequency. The reality of our simulation is implemented through the conditional transition rule, which allows the genotypes of the next generations to be changed based on the neighbors of the current cells. To improve its accuracy, we could implement more real-world biological factors like reproductive rates or mutation rates for better representation of differences in genetics. We are aware that

the changes in allele frequency over time is not as simple as our model, so learning more about other factors that could contribute to the evolution of such genotypes could give us more insight on how to better our application. Overall, our model reveals a solid foundation as a graphical representation of changes in allele frequency, but there is always room for improvement!

3.6. What you would have done differently :

3.6.1. Library development

We definitely should have reached out to Dr. Drummond in the early stages of creating our general purpose library so that he could have guided us in the right direction before we spent too much time and energy doing it wrong. Our repository only made use of the main Bin, Include, Lib, Source, Tests, and Utils directories. In past assignments we usually utilize subdirectories in Source and Utils as well, so we could have improved our code by placing some pieces in the applicable subdirectories to keep the repository more organized. We also could have cleaned up our source code if possible since it ended up being around 500 lines.

3.6.2. Software project management

In a professional setting, we probably should have been more descriptive in our pull requests on Github so that if anyone outside of our team were to look at our repository, they would have a clear understanding of each software developmental step. We also should have done a bit more research when it came to the use of Github and a shared repository because we ran into issues that in the end were not easily fixable.

3.6.3. Product improvements

To improve our product, we could have found a way to include a feature that could let the user customize their parameters in real-time versus resetting all of them in the application at once. It probably would make it more realistic and easier to work with when studying the evolution of genetics in the real-world. Lastly, if we had more time we could have added more detailed documentation to lay out how exactly each part of the library would be used as well as the application in our test file. It would have made it more user-friendly considering outside users would go into it having no prior knowledge of our setup. These improvements would make our general purpose library and CA model easier to understand and utilize to its full potential!