

EE / CPR E / SE 491 - sddec20-22

DNA to Feature Models

Week 3 Report

2/17/2020 - 3/1/2020

Client: Dr. Myra Cohen

Faculty Advisor: Dr. Myra Cohen

## **DNA To Feature Models**

The following report presents the progress of the senior design team for the project, DNA to Feature Models, in the past two weeks. The team took the appropriate time over the past few weeks to understand software product lines, families of software product lines and feature models. Collectively, the team took a big step forward in creating a strong foundation for the plugin.

### **Team Members:**

- Abdul Rahman El Moughrabi - Developer/Documentation Management
- Ahmad Nazar - Team Leader/Developer
- Ahmed Alketbi - Developer/Debugger
- Hyegeun Gug - Developer/Web Management
- Prathik Nair - Debugger/Developer

### **Past Week Accomplishments:**

- Met with Mikaela, Professor Myra's research assistant, to discuss her work on scraping BioBricks Repository.
  - Learned two approaches on how we can scrape BioBricks Repository:
    - XML web scraping
    - MySQL database
  - XML provides more information about BioBricks parts which is why we decided to take this approach.
- Completed design document first draft
  - Received input on how to improve it for next version
- Experimented with web scraping for a specific category of parts (To build a start).

### **Pending Issues:**

- Executing the shell web scraping program from the plugin itself
- Receiving proper input from the web scraping program to parse through

### Individual Contributions:

#### **Abdul Rahman El Moughrabi**

In the past two weeks, I made a bit of progress regarding this project. I used 3 hours of the time to understand the project fully and fill all the gaps that I had. I taught myself different methods of data scraping and how to be able to read that scraped data more. In addition, I learned a bit about using the extracted data to input into our plugin. The information learned doesn't specifically relate a lot to what we are supposed to do, rather I used this as a building block to get some practice on what we will be doing. I looked at my teammate's code for the backend server and understood how everything will work together. The file parser is another thing I familiarized myself with and how this parser will take all the scraped data and be implemented into the plugin.

#### **Ahmad Nazar**

Since the last report, I have made significant progress; taking the time to understand how a model structure's expected XML file helped me analyze how to approach an XML builder. In addition, the web scraping experience proved useful. After realizing the output of the web scraper and the manipulations used in extracting specific information, the task of building a part was made clear. These two weeks have been spent on initializing a backend server to be used for dumps of web scraping output files. These dumps will be used in a file parser that the plugin uses to create parts which later stores in the database for later extraction. I am currently working on the parser that analyzes the web scraped file and stores strings of parsed objects in a collection for later analysis

#### **Ahmed Alketbi**

During the past two weeks, we had the chance to meet with Prof Cohen's research assistant (Mikaela Cashman) who worked on web scraping an older database of BioBricks repository. She demonstrated her approach for getting parts information from an XML export and a database export. We discussed how we can start a small portion of the project by only scraping one category of parts from the BioBricks repo. During the first week, I tried to build some experimental feature models using XML formats just to see and learn how it behaves. It's a necessary future step because once we scrape the repository and parse the information we will convert the acquired data to feature models in XML format. As for the second week we went over our design document and reviewed it with our client to get feedback on our first project foundation.

**Hyegeun Gug**

For the last 2 weeks, I tried to modify and run live running data scraping from Biobrick repository which Mikaela, Research assistant of Professor Myra, provided to us. Which gave me a chance to gain web scraping experience. However, there was an unexpected Unicode breakdown. Therefore I couldn't analyze or there was an error on the code I built. As there was a discussion about the server and putting all the data to MySQL. This will provide users to successfully access the Biobricks Repository without an internet connection when they download the whole repository on previous use. This could be done after our team could fully understand web scraping, XML, and data structure.

**Prathik Nair**

After learning more about extracting XML data, and scraping the BioBricks, the team and I have a better understanding of how to get started. We also learned that we need to essentially, piece-by-piece, bring in data to our own database. I was tasked with figuring out how to go about sorting all the data we bring into our database. While this is not what we initially expected, this allows users to use the plugin without an internet connection. And it eliminates the need for us to scrape the web. Other tasks this week involved starting the first few sections of our design document, as well as our other weekly duties. I expect us to push and review more code in the upcoming weeks. I also plan to put in more hours starting next week.

Team Member	Weekly Hours	Total Hours
Abdul Rahman El Moughrabi	5	14
Ahmad Nazar	6	16.5
Ahmed Alketbi	5.5	14.5
Hyegeun Gug	5.5	15
Prathik Nair	5	14

### Plans of Action:

By the next report, the team hopes to achieve the following goals:

1. Develop solidified models related to parts.
2. Parse through a file and create objects stored in a database based on file contents.
3. Build a live web scraping protocol with storing data to the database.