Angela Alaisa

CSCE 240, Section 001

April 15, 2022

<p style="text-align:center">Chatbot Project Report</p>

Requirement

The instructor asked us to create a chatbot that answers questions and gives information about a representative from a chosen district in South Carolina state legislature. The chatbot was broken up into five programming assignments: an extractor, processor, user interface, query mapper, and session logger. We then combined all of these assignments together into one program to create the fully functioning chatbot.

Specification

I decided to code this project in Java as it is the language that I am most comfortable with. At first, I tried to use C++ but I found it too difficult since I have not taken a class on it or thoroughly taught myself that language yet. Another simplification I made was to extract data from a local copy of the webpage rather than pull directly from the site itself. I did extract the data using the URL class in Java for PA 1, but for the rest of the assignments I used a local copy. The reason why is because I felt that it would be more reliable when others are testing my code and it also allows me to test offline since I don't need an internet connection to access the data. To return content based on a given input, I used regular expressions to search through the file for the requested information. I also decided to not use the suggested approach for PA 4 to utilize a confidence estimate. I found it easier to create more regular expressions to handle more specific requests and if the chatbot still couldn't find any matches it would simply ask the user to reword their question.

Development Highlights

All of the code for each programming assignment was written in one java file. I felt like it would be easier to have everything in one spot rather than create multiple files or classes that might make things complicated for me. I had one overall class with variables for each information type which could be accessed by any method in the class rather than creating a separate class that had to be extending or implementing to class containing the main method. I did most of my testing through outputting certain things that I was looking for to make sure I was extraction the right data, creating the right file, etc., rather than having to look through every line to figure out where the problem might be occurring. I often faced problems regarding loops, there were times where too many files were created or an output was printed too many times but these were easily fixed by moving some code outside of the loop.

Reuse

I basically built off of each programming assignment which made my code quite reusable. The only differences between each of them would be the addition of new methods, adjustments to existing methods, and the main method in order to meet the requirements and functionality for the next assignment. I also made sure to not hard code or use paths that would only work for me so that others could be able to use my code with no problems. I ended up using a small portion of Darshan Patel's code for a couple reasons. First, he was using the same language that I used, and he was also the first person to reach out to me about using my code. After looking over his work, I saw a small portion that I used to help me get unstuck from a section of PA 5 that I was having a hard time figuring out how to do. Darshan reused my code for pretty much the same reasons. I don't recall having any major challenges regarding code reused besides having to change the names of variables to not conflict with one another.

Future Work

I feel like the one thing that would make my chatbot more useful would be to use the suggested approach that found the confidence of the query. Another way that would make it more useful would be if it could access the data from the linked sites to my representative's voting record and sponsored bills since it wasn't included in the html code for the main District 74 page. I think that another cool feature to have would be to have the chatbot suggest queries if it is unsure about the query. Over time, I feel like the code would need to be changed to incorporate additional committee assignments since the representative currently does not have any so the regex to search for that information may need to be altered.