

Rae Jones

CSCE 240

Prog4 - User Intent To Query Mapper Documentation

In this version of my project, I focused on cleaning and reorganizing much of my old code. This resulted in the creation of two new classes: `DataLoader` and `Responses`. As expected, `DataLoader` handles all of the information from the CDC and WebMD txt files. `Responses` is a collection of methods formed to translate the user's question into a query that the program knows how to handle. The function of the `Responses` class begins with the `queryMatch()` method. It takes in a single word of the user's question and checks if any of the known queries contain that string of characters. If they do, it returns a one. The method `queryConfidence()` takes in user input and uses `split()` to cut the string into individual words. It then uses a for loop to compare each of the words to a list of known queries using `queryMatch()`. It then adds together the number of matches to each known query that the user's input has and returns that number. This method is utilized by `chosenQuery()`. `chosenQuery()` contains a 2D array of known keywords and their corresponding phrase to search the data for. It also contains a variable called "closestQuery" to hold the best match and a variable "highMatchNum" to hold the highest number of matches found. `chosenQuery()` uses a for loop to go through the known keywords in row 0 of the 2D array and calls `queryConfidence()` to find if it is a match. If it finds a higher matchNumber than it currently has, the `chosenQuery` is set to equal a new `closestQuery` and `highMatchNum` is updated. Once a known query has been found by `chosenQuery`, `handleQuerys` takes it in to search my QnA objects for the right response using a for loop. This method has a special way to

handle “get everything”. I also have a tellMore() method to figure out if there are multiple answers and split them for the “tell me more” part of the ui.

To help me with my pattern matching in dataLoader, I used information from w3schools at https://www.w3schools.com/java/java_regex.asp.

To help with patternMatching in my confidence estimation method, I used <https://stackoverflow.com/questions/7929495/java-finding-the-number-of-word-matches-in-a-given-string>

In the future, I would like to find a way to get away from the semi-hardcoded 2D array in the Responses class. I also feel that some of the code in the Responses methods could be condensed in the next version of my project.