Luke Hudspeth

COSC 072

Final Project Proposal:

A simple “Siri/Alexa” named Levi:

For my final project, I plan on implementing a naive and simple interactive voice to text processor that can handle a few small tasks, this processor will be named Levi. Levi will be capable of taking an auditory snippet that will be prompted by saying its name “Levi.” Levi will then segment this auditory sequence and parse out potential words from the audio. I plan on using our min-edit distance algorithm to correct potential words from the audio that might be misspelled. For that algorithm, I will either use the provided dictionary or generate my own Dataset that will only contain words related to what Levi can process, and if the distance is too far, then I will return a hard-coded response, such as “sorry Levi has not developed the skillset to do such a task yet, or please speak more clearly.” I anticipate the majority of the work on this project being related to the audio to text parsing, and assuming I can accomplish that sufficiently, I will implement a naive response algorithm, that will be able to answer a few generic questions such as “What day/time is it?”, a greeting, and perhaps a few generic questions dependent upon how well the initial portion of the project proceeds.

I anticipate having to generate a dataset and a mapping dataset that will parse out words to garner the semantic nature of a question. I.e. Using “wh-“ words to interpret if a question is being posed. I will evaluate the performance primarily on the parsing of audio to text. The other performance aspects will be secondary and limited to my ability to have Levi perform simple tasks (perhaps multiplication or division). The baseline I will compare this to will be quite competitive (the current abilities of Siri, Alexa, and Google, to parse the same verbal sequences). I know python has some libraries already dedicated to auditory parsing so I will explore those and implement my own methods specific to Levi.