Andrea Youwakim
Edward DeCoste

# Final Project Proposal - Multi-Step Vocal Command

**Task:**

The problem we wish to resolve is multi-step voice recognition delivered orally by any user. This task has a multitude of applications because it would allow the robot to interpret a set of tasks instead of accepting only a single task at a time. Our project's goal is to build upon the BWI Bot's existing speech to command functionality by creating a framework to allow the bots to perform sequential operations based off of natural speech patterns. By enhancing the speech-to-text recognition abilities, we are enabling the robots to operate on a more autonomous level.

**Goals:**

By the end of the semester, our goal is to have the robots processing speech and performing actions independently once the initial input has been processed. Initial performance testing will include a series of phrases that will be delivered at one time. Each sentence or command will be separated by key delimiters. Such phrases may include "Go to xyz room and say hello" and "Say xyz then spin 90 degrees." If the robot does not understand one or all of the commands, it will prompt the user to repeat themselves. If the robot cannot understand the commands after asking the user to repeat themselves twice, it will  perform a predefined action to let the user know that it was not able to recognize the command.

**Testing:**

We will test the robot by giving it a series of commands that will be parsed using the program that we will implement. For instance, the robot should be able to receive commands such as "Go find Jivko then Bring him a cup of coffee then return to the charging station and await further instruction." Notice, each task is to be separated by a key delimiter phrases— in this case, the word "then." Other delimiters may include

Andrea Youwakim
Edward DeCoste

"and", "afterward," and "followed by." If the robot is unable to properly parse the series of commands and therefore cannot complete the given tasks, then we will know that the tests failed and revisions need to be made.

In reference to long term applications, the most efficient way to test this system is to use different voices and dialects with each available function that the robot has mastered. This will allow for the most accurate speech detection and operation as well as broaden the field of accessibility to various non english-speaking users. Another testable feature is that the robot is able to complete each task in order and keep track of an accurate list of tasks that are to be completed.

## Resources:

**Cloud Speech API**s will allow us to utilize enterprise grade speech to text software. The current top platforms are Google Speech and Alexa Speech Service.

**rqt** is a library for creating ROS GUI programs. This will allow us to create the visuals to show that the robot is listening and provide the user a way to tell the robot to start processing the commands.

## Timeline:

April 3  -  April 17 : Work on getting robot to understand and parse a series of commands.

April 17 - 31: Refine speech parsing and action algorithms.

April 31 - May 7 : Perfect final touches on program.

May 7 - May 12th: Write paper

May 13 - 16th: Presentation