William Kluge

CMPT-220 Project 2 Proposal

autoDEC

My idea for project 2 is to create a program that utilities various speech-to-text and audio editing libraries to translate the lyrics in an MP3 file into a DECtalk script. DECtalk is a speech synthesizer and text-to-speech program developed in 1984 that is famous for the vibrato of its voices and the precision that can be obtained with it. I want to create a program that will take an MP3 and turn it into a script for DECtalk to use. I believe this project will be difficult, but fortunately a lot of the work has already been done for me. I will likely use a library called sphinx4 to assist with speech-to-text, or more accurately speech-to-phone. DECtalk operates using phones (the sounds words are made from), pitches, and duration. The pitch and duration should not be terribly hard to get using audio processing libraries, so really the phones are the hard part. Sphinx4 already has the capability to split speech into phones, so I just need to tweak how the library wants to be implemented a bit. Instead of using AI to select the phone with the highest probability of being next to its surrounding phones, I want to create a user interface that allows for playback and selection from all possible phones. Overall I think this is a fairly complex project, but it is appropriate for where I am in computer science and offers me a challenge worth pursuing. I do not expect autoDEC to work perfectly when it is done, but the opportunity to experiment with these libraries will be very valuable to me.