Speech based Tic Tac Toe

CS 566

Group 2:

Aadil Hoda 160101001
Abhinav Hinger 160101004
Paranjay Bagga 160101050
Akul Agrawal 160101085

Contents

Application Flow

1. The game will start only after either one of the user utters a specific word '**BEGIN**'.
2. Each utterance of the user will be considered only after a '**Start Recording**' is displayed to the user, which will work as a marker for start of recognition and '**Stop Recording'** which will be end of speech recognition. This duration is kept 3 seconds.
3. The game begins by showing a grid of 3 X 3 filled by '-' (Unmarked). As the game progresses, users speak the positions one by one. Condition of winning is checked after each move.
Numbers spoken from 0 to 8 determine the position on the grid.
4. Both the players speak and hence play alternately, as in the game, with '**X**' and '**O**' symbol assigned to them accordingly.
5. A player's turn comprises of uttering a single grid-position (from **0-9**) to be marked, followed by a confirmation from the application side, to which he/she replies either '**yes**' or '**no**'.
6. If 'Yes' is detected and the spoken digit position is valid, then the game smoothly proceeds ahead.
7. While, if 'no' is detected or the spoken position is invalid, in the sense that that position was already covered or out of the scope of the game, then the given turn is repeated for that player until he utters a valid 'yes' in future.
8. After each successful turn, the grid is updated on the console screen (corresponding to the uttered position) according to the mark represented by the player of that turn. (either X or O)
9. Also, any user can say '**Pause**' at any valid point of time during the game for pausing the game. After that, 3 options are shown to the user - '**Resume**', **'Begin'**(Restart) & '**Quit**'. The user can speak any one of them for executing it.
10. 'Resume' : resumes the game, 'Begin'(Restart) : restarts the whole game afresh from the beginning, and 'Quit' : ends the game completely.

Models

Since this is a 2 player game, we developed 2 pre-trained voice models for each word listed below. Turn alternates between 2 players and appropriate model is applied for recognition.

1. Begin
2. Digits : 0-9
3. Yes
4. No
5. Pause
6. Resume
7. Begin(Restart)
8. Quit

Trainable Module

On the start of the application, user can select option to build his/her own model. On selecting this, user will be prompted to speak 10 iterations of one of the word listed above after further

selecting it from the menu shown on the console. This goes through the model building pipeline. This involves Preprocessing and Clipping, Cepstral coefficients, Observation Sequence and finally HMM modelling. This new set of 10 observations are trained using one of the previous pre-trained models for that word so that it just fine-tunes that word according to the user and make it better for correct recognition.
User can then start the game and select **Custom User** option, when the user selection screen is displayed, to use this custom built model.