Further Processing of Converted Text

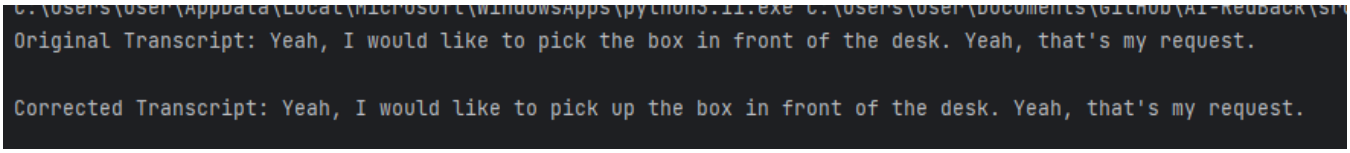
We decided to use the ChatGPT API to further process the transcription.

1. As a language model, ChatGPT has a superior grasp of context compared to basic transcription tools. This means it can potentially spot and rectify transcription mistakes by comprehending the broader context of the transcription.
2. After transcription, we want to summarize, categorize, or extract specific information from the text. ChatGPT can assist in these tasks, providing more structured or concise outputs based on the transcribed content.
3. ChatGPT can be used to add additional information or context to the transcribed content, making it more informative or easier to understand for the end-users.
4. After transcribing with the Whisper API, we can easily pass the content to ChatGPT for further processing, creating a streamlined workflow.

The prompts we used to process transcription text.

Grammar check:

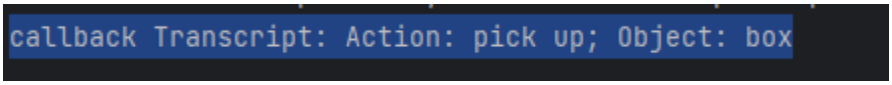
a. prompt=f"Please correct the grammar of the following text: \"{transcript}\""

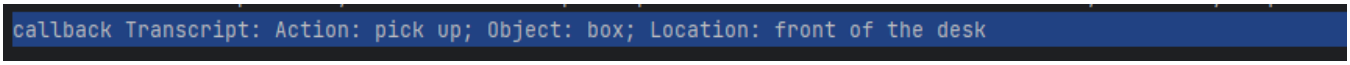


Extract specific information from the text:

The ChatGPT model we used is "text-davinci-003",

a. prompt=f"Translate the sentence into action and object, for example, a sentence 'move that cup' should be 'Action: move; Object: cup'. Now please translate {sentence}"
b. prompt=f"Translate the sentence into action, object and location, for example, a sentence 'slowly move that red cup to the top of the table' should be 'Action: slowly move; Object: red cup; Location: top of the table'. Now please translate {sentence}"





Commands paradigm

We recommend that users' voice commands conform to a certain paradigm. As the command like "I spy word starting with "M""-> The robot scans for objects in the environment that respect that constraints and if reachable, it would pick it up and place it in a box.

The users shall apply the Action+Object+Location paradigm. "I spy word starting with "M"""I want to pick up box with word starting with "M"'