Dialogue Systems: Lab 2 VG part

Viktorija Buzaitė

1 Limitations and solutions

First of all, a huge limitation of the app is grammar. I extended it to incorporate more people, days of the week and timings, but it becomes very loaded and it is not an efficient way to do it. For example, if you want a meeting at eight, you have to specify whether it is 8 am or 8 pm, but all of these variations have to be listed in the grammar for every timing. I left *at 8* option which is connected to *8:00* instead of *20:00* due to the reason that it is more likely that a person would like to have an appointment in the morning rather than in the evening. With weekdays there are also limitations, because it can be both *on Monday*, *Monday*, etc., and if the user says something similar but with the wrong preposition which is not listed in the grammar, the tool will not recognise it. In addition, the names are also limited because you can know more than one person with the same name so they would have to be encoded in a different way and it is not user-friendly. It would be better if in the case of a user saying that they want a meeting with John, the app would double-check and ask whether it is John Appleseed or John Melonseed. Also, another limitation that has to do with names is pronunciation. For example, if a user wants a meeting with Joe, the app confirms a meeting with Joe Biden according to the grammar, but the last name is pronounced incorrectly. This is not a big limitation, but name distortion can be annoying to some users.

Another limitation which I encountered was the app not recognising some of the utterances which had to be repeated multiple times. I am aware it depends on the English accent and some words might be pronounced not entirely correct by the user who is not a native speaker, so a better speech recognition system is needed for the app. Another advantage would be recognizing voice even with the background noise. Same goes for the intents part. The algorithm assigns the most probable one according to the utterances, but it not always correct, so a certain threshold should be added to better understand users' inquiries. 75-85% threshold would be good enough to proceed with the intent.

What I also find limiting is that the user is not able to change their mind in the middle of the process and choose another tool, so an option should be added that the user should be able to quit the dialogue in any state and return to the beginning to choose another task, or even change their mind about the person or timing. In this context it can also be mentioned that in a normal

human dialogue you are not only able to return back at any possible moment, but you can also use phrases like *I am not sure* instead of merely *yes* or no so an algorithm could encompass such phrases also (they could be inserted in the grammar) and then the app should check with the user again after hearing such an utterance. Also, it is rather irritating that in order to proceed with at least two tasks, the user would have to do them separately which is not efficient and time consuming. The app could be developed to recognize such words as *and*, *in addition*, etc. to be able to go back to the start and add another command in the same dialogue instance. To go with this even further, after choosing the task, the app could come up with suggestions to the user of further possible actions. For example, in case of making an appointment, the app could ask whether you wanted to call or message the person you just made an appointment with.

In addition, an issue which is limiting for the app is the Internet connection. If there is no connectivity, the app simply will not work and the user will not be able to either create an appointment, use the to do list or the timer, which should also be developed further. The app's features should be stored on the device so some functionality would remain, otherwise it is very limiting in places where the Internet connectivity is poor or when the user has no data.