

A simple chat bot on health counsel

Maximillian Paulus

I. DOMAIN

First ideas on the domain arose from the lack of motivation in people to care about their health, especially when they conduct a rich working life and do not see the necessity to involve more reasonable thinking in the after-work life.

A first topic-wise conception of the chat bot is shown in Figure 1. Depending on the well-being of the user, the system decides whether to propose athletic activities (well-being) first or to directly switch to food and drinking advice.

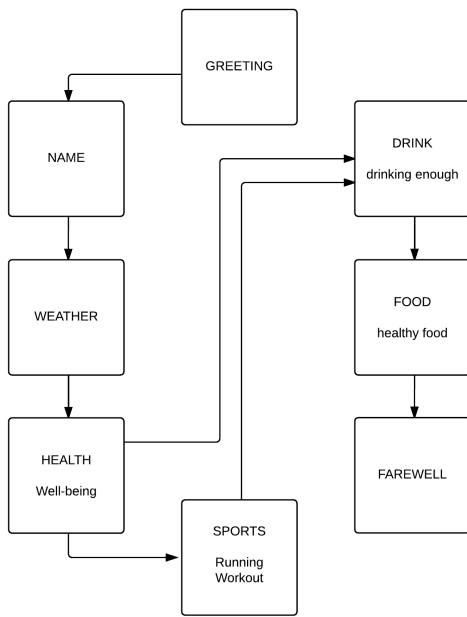


Fig. 1. Domain conception

II. SCENARIOS

On the basis of the domain conception, two possible scenarios can be considered:

(1) A full-time student who spends his day at uni and comes back home at some point in the evening. He might already have had the opportunity to do sports. His time is limited, but still he might want some quick advice on nutrition.

(2) An employee who works a lot and does not prioritize his well-being. The best thing that could happen to him is someone telling him what to do to stay healthy. A system that takes the initiative would be the most suitable choice in this scenario.

III. WoZ EXPERIMENT

I conducted a Human-Human dialog using the Wizard of Oz technique to accumulate first ideas on how the chat bot

should look like. The original dialog is shown in the Appendix. I applied distilling, i.e., not all phrases were used in the first implementation and most of them were transformed, so as to lead the conversation into the required direction.

IV. IMPLEMENTATION

Results from domain selection, scenarios and WoZ experiment were merged into the first implementation of the chat bot. The most difficult part was to connect domains with each other in order to obtain a coherent system.

The implementation of the chat bot is mostly system-initiative with some excursions towards a mixed initiative, e.g., when the user has the option to decide for a workout exercise. The policy of the chat bot is very strict, i.e., the context is easy to determine at any time. Questions by the system are very specific and mostly allow only for a small range of variance. This fact is due to the division of the domain (health counsel) in smaller sub-domains: conversation needs to follow a certain track from which it should not deviate too much.

A sample conversation with the chat bot is shown in Table I. It shows the turn structure and a variety of dialog acts that are encoded in the conversation. The sample conversation was conducted with the first implementation of the dialog system, but in general, a similar sequence of dialog acts can be observed with the latest version.

The system is also able to handle certain grounding mechanisms (understanding and using), e.g., the system shows turn-management abilities like providing a sequence of related contributions connected by feedback (acknowledgment) by the user: *You should drink a lot (S) - Thanks for the information (U) - About 2.5 liters a day is best (S)*. The system is also able to understand continued attention by the user, e.g., when it provides descriptions of workout exercises, continuing with suggestions until the user shows that he wants to change the focus of his attention (*That was enough*). Moreover, the chat bot is equipped with the functionality to demonstrate its understanding of the situation as a whole by providing a summary at the end with reference to the current emotional or physical state of the user (*Chin up* or *You seem to be in a good mood*). This very specifically shows that the system is involved in grounding.

For the implementation, I did not follow a specific grammar scheme. Taking into account the size of the dialog system, I came to the conclusion that it would be logical to follow the natural flow of the WoZ dialog. A grammar could have been useful for a more complex system with an artificially created structure.

V. EVALUATION

To allow external evaluation, it was necessary to design a script that would handle annotations by a user. I first

| Dialog Act | Understand | Generate |
|----------------|-------------------------------|---|
| Greeting | Hi | Hi John |
| Correction | My name is ... | |
| Self-error | | Oh, my fault. Hi ... |
| Negative-check | | Isn't it hot today? |
| Agreement | It is! | |
| Yes/No | | Are you feeling well? |
| Answer | Yes | |
| Social | | That is good to hear! |
| Yes/No | | Did you already do sports today? |
| Answer | No | |
| Directive | | I would recommend to go for a nice run |
| Inform | There is no forest nearby | |
| Social | | Don't worry |
| Inform/Offer | | I've got some workout tips for you |
| Wish | I want to work on my legs | |
| Inform | | This exercise is called ... |
| Wish | I want to do another exercise | |
| Inform | | This exercise is called ... |
| Directive | This is enough | |
| Committing | | Then we stop |
| Yes/No | | Are you exhausted? |
| Answer | No | |
| Suggest | | You should drink a lot |
| Acknowledge | Thanks, I know that | |
| Inform | | About 2.5 liters of water is best |
| Acknowledge | Thanks again! | |
| Yes/No | | Did you already have dinner? |
| Answer | Yes, I had! | |
| Inform | | Grab a snack ... |
| Acknowledge | Thanks | |
| Bilancing | | You ate already. Good! I think that is all I can do for you |
| Farewell | Thanks. Have a nice one! | See you! |

TABLE I
DIALOG ACTS IMPLEMENTED IN THE CHAT BOT

applied the function to mark an utterance by the system as inappropriate while at the same time counting their frequency. This was done to test the coherence of the dialog.

In a next step, I applied the function to log the number of turns in a sample dialog to open up the possibility of a complexity measure.

Finally, the user obtained an option to rate his conversation with the chat bot on a likert scale from 1-5 (unsatisfied-satisfied).

The first three testing sessions revealed that the chat bot lacked some essential synonym handling. The average rating was 1.7 (out of 5); the average number of inappropriate utterances by the chat bot was 1.3; and the average number of turns was 20.7, including empty utterances after not finding appropriate matches. The number of inappropriate utterances increased after manually checking the log files of the conversation. The occurrence of empty responses after not finding matches gave reason for changing the script in a way that the user is told that he would need to restart the conversation.

VI. REFINEMENTS AND FUTURE IMPROVEMENTS

In the latest version of the chat bot, a standard rule for each topic was added that would match any unhandled response by the user bringing him back to the beginning of the topic. This reduced annoyance by not showing system errors when a response could not be matched. Furthermore, the user was not brought back to the beginning of the dialog, but was able to stay in the current topic.

I conducted another block of test sessions with 6 participants where I achieved an average rating of 3.8. The average number of inappropriate utterances was 0.8 and the average number of turns 28. I.e., all three variables indicated improvements from the first version of the system.

The major issue I encountered was that it is impossible to capture all possible meaning synonyms. In that sense, the functionality of this simple chat bot can never be maximal. However, to some extent the desired variation within the predefined domains was accomplished.

In further improvements of the chat bot I would extend the existing domains in such a way that the user is more flexible in his responses. This would mean to apply a more open policy. The chat bot would have access to a much richer knowledge base.

VII. APPENDIX

U Hello
S How are you?
U Fine, thanks
S Are you feeling well?
U Very well
S That is good to hear. Did you already do sports today?
U Yes, I went to the gym at 8 o'clock
S That is excellent! Are you exhausted? Do you have any sort of pain?
U No, I'm feeling good
S That's great! Isn't it hot today?
U Very hot!
S Yeah! Did you drink enough?
U I did, about 1 liter
S That's good! Did you drink water or anything else?
U Just water, but at the moment I have a coffee
S Water, that is very good! You mentioned coffee, how much coffee do you drink?
U Hmm, about two cups a day maybe?!
S That is alright, you should not exceed that though!
U I won't
S Excellent! Water is best, about 2.5 liters a day. You can also have tea and juice, but you should mix it with water!
U Good to know
S I appreciate you acknowledge my advice. Are you planning on working out today (again)?
U No, it's just too hot! I'll stay inside and work
S I figured that! You mentioned you worked out today, did you do muscle work using weights or did you go for a run?
U Yes, I used some weights and machines and did some condition workout
S You shouldn't run too much, it's hot. Do you feel like some parts of your body are much more exhausted than others?
U My arms are feeling a bit heavy, I think they are more tired than my legs or my back
S Perfect! That means you are going to relax your arms a bit and I would recommend training your legs and your back tomorrow. you should talk to me again tomorrow so we can figure out which exercise would serve you best!
U That sounds good to me, thanks!!