Personified Robotic Chatbot Based On Compositional Dialogues

Michal Vavrecka
Czech Institute of Informatics, Robotics and Cybernetics
Czech Technical University
Prague, Czech Republic
michal.vavrecka@cvut.cz

Gabriela Sejnova
Czech Institute of Informatics, Robotics and Cybernetics
Czech Technical University
Prague, Czech Republic
gabriela.sejnova@cvut.cz

Petr Schimperk
Czech Institute of Informatics, Robotics and Cybernetics
Czech Technical University
Prague, Czech Republic
schimpet@fel.cvut.cz

2022 International Conference on Interactive Media, Smart Systems and Emerging Technologies (IMET) | 978-1-6654-7016-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/IMET54801.2022.9929602

***Abstract*—We present a novel chatbot architecture implemented in the humanoid robot Pepper that is capable to communicate in natural language in a compositional fashion. Compared to the previous solutions based either on question-answering pairs or dialogue trees, we propose a novel method which constructs the dialogue from atomic blocks that can be applied interchangeably. It thus allows the creation of hundreds of meaningful dialogues from only a few building blocks. Most of the dialogues are focused on the user preferences and opinions which are stored in a database and translated into feature vectors. The chatbot is able to compare these preferences with other users or with data from the internet so that it can be used as a crowd-based recommendation system. To improve the user experience, the chatbot has its own personification in the form of predefined opinions and preferences for each question asked during the communication. The embodiment of the chatbot in the humanoid robot enables further improvements in the interaction. We report results from a preliminary evaluation with a group of volunteers that compared our chatbot with a different system.**

***Index Terms*—conversational agents, humanoid robots, chatbots, personification**

I. INTRODUCTION

Communication in natural language is considered a crucial requirement in the development of artificial intelligence systems (chatbots) and has been a subject of intense research over the past decades [1]. The recent chatbots are mostly based on approaches using precoded dialogue trees (e.g. Alquist [2]) or neural networks that are able to generate answers based on the dialogues they were trained on (e.g. GPT3 [3]).

The most recent survey [4] presents a systematic overview of factors that are crucial for successful conversation in natural language. The authors postulate three key areas: conversational intelligence, social intelligence, and personification. In the conversational intelligence area, proactivity is an important aspect of communication. It increases the engagement of the user as the chatbot introduces new topics and contexts [5]. The next area postulated in [4] is social intelligence which focuses on habitual social protocols. From the social point of view, it is also very important to keep a communication balance between actors during the dialogue as the turn-taking communication strategy allows the chatbot to adopt a human-like dialogue style [6]. The last area mentioned in [4] is dedicated to personification that reflects the identity of the chatbot and the personality representation of the user. The personality representation (personalization) highly improves cooperation and engagement [7]. Complementary to the identification of the user, it is also desirable to develop the personality of the chatbot. Overall, we believe that chatbots equipped with a set of specific modules can exhibit more complex and human-like behavior than simple bots with only a few modules. The list of abilities that we incorporated into our proposed architecture is following. The chatbot proactively seeks for partners and discusses with them various topics and offers novel topics that are compositionally created from templates. The chatbot can compare users and recommend similar ideas based on collaborative filtering. The turn-taking strategy is based on interchangeable communication blocks. The chatbot has handmade consistent personality and opinion on every topic and it is embodied in a humanoid robot.

II. ARCHITECTURE

Here we focus on the conversation core which consists of structural modules: User database, Dialogue manager, Topic Selection, Recommendation system and Game/Quiz module. The connections among individual structural modules are shown in Figure 1. As the architecture is fully modular, it is possible to run the chatbot with any subset of these components (with the others turned off) except for the Dialogue manager which is necessary for the action selection process. The whole architecture is implemented in the humanoid robot Pepper based on an API for Python developed by our research team. This API is a wrapper around the qi framework by Aldebaran which controls Pepper using Python 2.7.

A. User Database

The main purpose of the User database is to store information collected about individual users, their topic preferences and topic-related answers. It also contains the robot personality in the form of a predefined set of opinions and answers for individual topic questions. The database is an SQL.

B. Dialogue manager

The Dialogue manager controls the whole conversation and decides what should be the next action. These decisions should

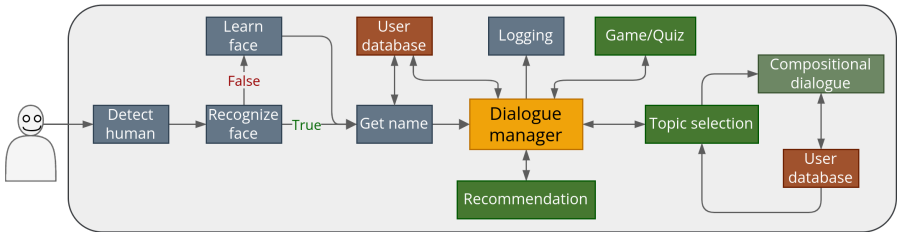

Fig. 1. Overview of the chatbot architecture. Once a human is detected, the robot tries to recognize the face. The chatbot uses information about the person stored in the User database to provide personalised content. The Dialogue manager module chooses the next most appropriate action - either a Recommendation, a topic-related Game/Quiz or Topic Selection which is followed by a Compositional dialogue.

be primarily based on the current user's conversation history. If there is not enough data about the user, we use the average preferences of other users or, if there is not enough data from other users, we use random decisions.

C. Topic Selection

We use the Topic selection module to collect data about the users and to entertain them with web-scraped contextual answers and personal robot opinions. As depicted in Figure 1, the loop starts with the *section* and *topic* selection. This choice is constrained by several aspects: whether the user prefers the *section/topic* or not, whether the *topic* was already discussed during the current conversation and whether the user has already answered questions on that *topic* in the past.

D. Recommendation System

When the chatbot has collected a sufficient amount of data about the user, it can use it to provide personalised recommendations for the given area of interest. Due to the various information sources retrieved from conversations and robot sensors, we are able to combine three commonly used recommendation techniques: content-based, collaborative and demographic filtering [8].

E. Games and Quizzes

To avoid stereotype, we also incorporate a set of interactive games and quizzes related to various *topics*. These include for example the blank map labelling game (for the travel *topic*), song or movie recognition quiz (music/cinematography) or exercise with the robot (sport). All these applications take advantage of the robotic embodiment and employ gestures, sound or the robot's tablet.

III. EVALUATION

We evaluated the quality of interaction using a post-hoc questionnaire that covers basic aspects of social interaction: the physical aspect of communication (understanding and pronunciations), the emotional aspects (believability, empathy) the social aspects (quality of the dialogue, adequateness) and overall judgement (whether the interaction resembles the human-human communication). There were 20 volunteers in the sample. To compare the difference between chatbots, half of the users talked to the Alquist chatbot [2] implemented in the Pepper robot. Based on the evaluation of both chatbots, our system reached better average score (2.5 SD=(0.4)) compared to the Alquist chatbot (3.1 SD=(0.6)).

IV. CONCLUSION

We have developed a chatbot architecture embodied in a humanoid robot that is capable to mimic human communication skills. The main advantage of our architecture lies in the compositional fashion of question generation that allows us to create many dialogue combinations from few atomic parts. Also, the presence of interchangeable dialogue parts is a novel method in the area of chatbot conversation strategy. The presence of a consistent chatbot personality improves the dialogue quality. The recommendation system based on crowd-collected data is more reliable than anonymous opinions and preferences distilled from the internet. The robotic embodiment of the system improves the believability of the chatbot as it offers also nonverbal aspects of communication.

ACKNOWLEDGMENT

This work was supported by the TAČR agency project no. TL02000362, INAFYM project no. CZ.02.1.01/0.0/0.0/16.019/, by the CTU Student Grant Agency project no. SGS21/184/OHK3/3T/37 and by the GAČR project no. 21-31000S

REFERENCES

[1] S. A. Abdul-Kader and J. Woods, "Survey on chatbot design techniques in speech conversation systems," *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 7, 2015.

[2] J. Pichl, P. Marek, J. Konrád, M. Matulík, and J. Šedivỳ, "Alquist 2.0: Alexa prize socialbot based on sub-dialogue models," *arXiv preprint arXiv:2011.03259*, 2020.

[3] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell *et al.*, "Language models are few-shot learners," *arXiv preprint arXiv:2005.14165*, 2020.

[4] A. P. Chaves and M. A. Gerosa, "How should my chatbot interact? a survey on human-chatbot interaction design. corr abs/1904.02743 (2019)," *arXiv preprint arXiv:1904.02743*, 2019.

[5] H.-Y. Shum, X. He, and D. Li, "From eliza to xiaoice: challenges and opportunities with social chatbots," *arXiv preprint arXiv:1801.01957*, 2018.

[6] C. Toxtli, A. Monroy-Hernández, and J. Cranshaw, "Understanding chatbot-mediated task management," in *Proceedings of the 2018 CHI conference on human factors in computing systems*, 2018, pp. 1–6.

[7] Q. V. Liao, M. Davis, W. Geyer, M. Muller, and N. S. Shami, "What can you do? studying social-agent orientation and agent proactive interactions with an agent for employees," in *Proceedings of the 2016 acm conference on designing interactive systems*, 2016, pp. 264–275.

[8] P. B. Thorat, R. M. Goudar, and S. Barve, "Survey on collaborative filtering, content-based filtering and hybrid recommendation system," *International Journal of Computer Applications*, vol. 110, no. 4, pp. 31–36, 2015.