

Dialogue Assistant System Using ChatGPT’s Function Calling

Fukuta Ito¹, Takumi Iwamoto², Kosuke Shigematsu¹

Abstract—For the Dialogue Robot Competition 2023 (DRC2023), we developed a system that suggests tourist attractions in Kyoto through conversation with users. We utilized ChatGPT, a Large Language Model (LLM) provided by OpenAI, to generate conversations with users. The latest tourist information is obtained from external APIs through the Function Calling feature of ChatGPT. For acquiring tourist information, we utilized the “Rurubu API” provided by JTB Corporation.

I. INTRODUCTION

Recently, the tourism industry has been facing severe labor shortages, calling for effective solutions. To address this issue, competitions focusing on dialogue performance by robots are being held by travel agencies. The robots are tasked to suggest itineraries covering two tourist spots in Kyoto, and their performance is evaluated based on user satisfaction and the feasibility of the plans through dialogue. The competition’s dialogue robots are equipped with capabilities such as speech recognition, vocal expression, and the ability to change postures and facial expressions.

Our focus was primarily on developing the dialogue system, aiming to achieve more natural conversations than traditional rule-based dialogues by using ChatGPT.

II. OVERVIEW OF THE DIALOGUE SYSTEM

Figure 1 shows an overview of our system. We input user queries into ChatGPT, and if necessary, ChatGPT employs Function Calling to retrieve information from external APIs. Finally, the text generated by ChatGPT is converted into speech for interaction.

ChatGPT is accessed through an API provided by OpenAI, and we used the GPT-3.5-turbo model. While GPT-4 offers superior performance, its response time is not sufficient for dialogue systems. Therefore, we opted for GPT-3.5-turbo to balance response speed with adequate performance.

A common issue with LLMs, including hallucination leading to the provision of false information, is undesirable in dialogue systems. Peng et al. successfully enhanced the factuality score by feeding external knowledge to LLMs like ChatGPT. In this system, we utilized the additional functionality of Function Calling in ChatGPT to acquire information from external APIs, thereby aiming to suppress hallucinations by providing ChatGPT with the latest and most accurate information.

For speech recognition, we used Google Speech Recognition integrated into our humanoid robot system. Tourist site

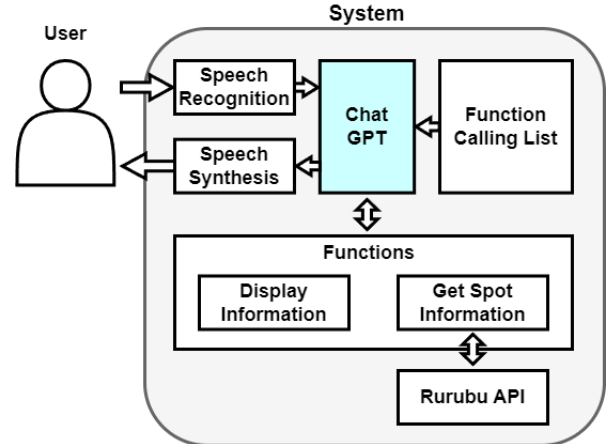


Fig. 1. System Overview Diagram

TABLE I
ROLES AVAILABLE FOR CHATGPT

Role	Description
system	Configures ChatGPT settings
user	Inputs user queries
assistant	Inputs text generated by ChatGPT

photos and maps are displayed using the robot’s SiteViewer feature. SiteViewer can show images and maps from specified URLs, so in our system, we provide users with visual information about tourist spots by displaying image URLs and Google Maps URLs.

A. ChatGPT

ChatGPT is a versatile dialogue system that can be specialized for specific tasks by adding prompts. Prompts consist of roles and content, with three possible roles: system, user, and assistant. Table I describes each role. Thus, in our system, we attempted to improve dialogue performance through prompt engineering. Through trial and error, we iteratively refined and adjusted prompts to steer the system towards more desirable behaviors, gradually aligning its responses with our intended dialogue goals. Table II shows some of the prompts used in our system.

B. Function Calling

Function Calling is an additional feature of ChatGPT that allows dynamic execution of functions. ChatGPT has a predefined list of functions and, based on user input, requests the API caller to execute the necessary function. For functions requiring arguments, ChatGPT determines the

¹Department of Information Engineering, National Institute of Technology, Oita College

²Electrical, Electronic and Information Engineering Major, National Institute of Technology, Oita College

TABLE II
SAMPLE OF PROVIDED PROMPTS

Role	Content
system	You are the contact person for a travel agency.
system	Don't talk about plans where money is involved.
system	Do not speak except where asked by the customer.
system	Your name is Shoko.
user	Answer the question in the same language as the question.
user	If you do not understand the meaning or words of a question, please ask back.
user	The reply to the user is given by voice. Please answer in clear sentences, not in lists, etc.
user	MUST ANSWER IN A FEW SENTENCES.

arguments based on the function's definition and user input. The API caller then returns the execution results to ChatGPT, which generates a response for the user based on this feedback.

In our system, we defined two functions: one to obtain tourist information using the Rurubu API and another to display images and maps of tourist sites on the monitor.

III. RESULTS

We participated in the DRC2023 preliminaries held in Tenjin and Nagoya using our system. The evaluation method involved users completing a post-dialogue survey to rate their satisfaction and the plan's evaluation. Satisfaction was determined through a seven-point scale across nine survey items, while plan evaluation was based on two items and assessed as either positive or negative. Our system interacted with 16 users at the two preliminary venues.

Figure 2 shows the evaluation results of all teams. The scatter plot has satisfaction ratings on the horizontal axis and plan evaluation on the vertical axis. Our team, positioned at G, scored below the baseline in both evaluations. This was mainly due to issues in speech recognition, such as incorrect word recognition and premature termination of recognition during pauses in user speech, leading to inappropriate responses from ChatGPT. As a remedy, we are considering using a Deep Neural Network to assess speech recognition success and, if deemed unsuccessful, prompt the user to repeat their statement.

IV. CONCLUSION

We developed a dialogue system for DRC2023 using ChatGPT and Function Calling. The results showed lower user satisfaction and plan evaluation than the baseline. Future challenges include accurately recognizing and understanding user speech to provide appropriate responses.

REFERENCES

- [1] Takashi Minato, Ryuichiro Higashinaka, Kurima Sakai, Tomo Funayama, Hiromitsu Nishizaki, and Takayuki Nagai, "Design of a competition specifically for spoken dialogue with a humanoid robot", *Advanced Robotics*, vol.37, no.21, pp.1349-1363, 2023.
- [2] Takashi Minato, Ryuichiro Higashinaka, Kurima Sakai, Tomo Funayama, Hiromitsu Nishizaki, Takayuki Nagai, "Overview of Dialogue Robot Competition 2023," *Proceedings of the Dialogue Robot Competition 2023*.

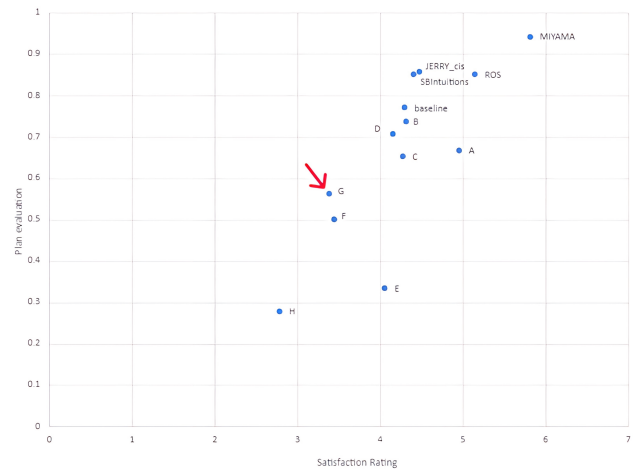


Fig. 2. User Satisfaction and Plan Evaluation[2]

- [3] Baolin Peng, Michel Galley, Pengcheng He, Hao Cheng, Yujia Xie, Yu Hu, Qiyuan Huang, Lars Liden, Zhou Yu, Weizhu Chen, Jianfeng Gao, Check your facts and try again: Improving large language models with external knowledge and automated feedback. *arXiv preprint arXiv:2302.12813*, 2023.