WE NEED TO TALK: CREATING SPEECH DIALOGUES AND PERSONALITY IN HUMANIZED-BOTS

STATEMENT OF RESEARCH PROJECT

Chatbots, talk/voice bots, chatterbots, and all other AI-powered bots and artificial conversational entities (ACE) are trending. Whether you are fascinated by the most common manifestation of virtual agents such as Siri (Apple) and/or Alexa (Amazon), or whether you are compelled to use the chat assistant in smartphone apps or online messages platforms in seeking information for banking/commerce, healthcare, or public sectors service or not, chatbots have become an integral part of our daily life. We all have experienced excitement and frustration when navigating information for customer service or counselling, or for getting entertainment or education.

The goal of this project is to enhance the experience, quality, efficiency and safety when engaging in dialogues with bots. No doubt that there is a huge untapped market profit behind intelligent virtual agents worldwide¹ and virtual celebrities are getting widely used in e-commerce (Strait times), researchers are alerted by the misinformation and unconstructiveness of using bots to resort to various kinds of communications that may bring harm more than benefits. A few examples that have reminded us of the research imperatives are: in the United Kingdom, Alexa tells a 10-year-old girl to touch a live plug with a penny happened in December 2021 (BBC, 2021)². In Singapore, users have found the chat with bots mechanical and impersonal after using the recently launched MOE Mental Health Portal in Singapore (The Must Share News Team, 2022) These incidents are just the tip of the iceberg pinpointing scientists should channel more energy and effort into this research area.

The objectives of this project are:
1) To create dialogue systems with consistent persona and suitable speech styles for bots, and:
2) to mitigate the widely circulated problems in the existing dialogue systems in virtual agents.

The two objectives above are intertwined as it is inescapable for modern dwellers and smart-nation citizens to use chatbots in all walks of life.

Case studies illustrating research questions:

As of today, most users are, indeed, holding a speculative attitude toward chatbots, and even virtual agents as a whole. In a study by Intercom in 2021, 87% of customers would rather speak to a human than a chatbot. In a report by CGS in 2019, 57% of the users expressed negative sentiment toward chatbots, describing them as “useless,” “impersonal,” and “irrelevant.” (CGS, 2019). The attitude of the users sent us thinking about the major problems virtual agents are facing. We conclude that there are two major problems:

1. The inability to provide useful information.
2. The distance felt by people during the interaction, whether it is emotional or conversational.

For problem 1, the inability to provide useful information may come from the inability to comprehend the query. Taking counselling chatbots (Wysa) used by MOE as an example:
A user may come with a sophisticated situation with a sentence formed by a complex syntactic structure and ask for the bot’s help. The bot is not well-versed to understand freetext and can only provide help based on the categories of common emotional problems. Similarly, if we take another mental health bot, Woebot, as another example, you may find the answer is also annoying. The tone in his speech is not what the bot is trained for, so the bot is not able to extract useful information from the destitute mood but can only provide rhetorical and repetitive answers.

Let us now investigate another example. XiaoIce, an open-domain chatbot widely deployed in society in 2014, is suffering from inconsistency in personality, which is displayed in its inconsistency in responses to questions, especially to questions that ask about the opinion of XiaoIce. Examples of this are shown below in (a) and (b) of Figure 2.

When the Mandarin version of XiaoIce on WeChat asked about its opinion on “self-esteem” and “a peaceful world,” it provided contradictory opinions in consecutive responses.
Figure 2 (a)(b)(c): Inconsistent responses produced by XiaoIce
In the example of Figure 2, the disengagement felt by the users may come from at least 2 aspects

a) Providing inconsistent responses within context 

b) Lack of engagement, or improper engagement

For aspect (a), inconsistency problems are especially evident in open-domain chatbots. For aspect (b), on the other hand, lack of proper engagement is a common problem that haunts chatbots for counselling and daily mental health practice. As the chatbots are not designed to be open-domain chatbots, many find the built-in responses insipid, ridiculous and meaningless as they do not react properly to the information they provided.

1 According to Polaris Market Research in 2020, the total market value for intelligent virtual agents worldwide was estimated at around USD 2.5 billion in 2019 (Polaris Market Research). Whereas another research report projected that the market revenue will go up to USD 3.99 billion by 2030 (Grand View Research, 2022)

2 Other example includes a case when user asked for Covid measure protocol at MOH’s bot, it provides advice on having safe sex. (Strait Times, 2021)
This project is looking for students from Engineering School/computer science Schools who are
(a) bilingual (in both Chinese and English);
(b) biscriptal (in both traditional and simplified Chinese characters) who could navigate various
kinds of gaming and communication platforms in the digital world;
(c) An exposure to translation studies (broadly defined including speech recognition system) and
linguistics research or knowledge would be a plus.

This project also welcomes Arts/Humanities/Social Science students who have been properly
(d) trained in machine learning, Natural Language Processing (NLP) and AI

In either case, students must be able to conduct quantitative research and qualitative research to
ensure the academic rigor of this project when handling, collecting, interpreting and analyzing
data.