Rethinking the Agency of AI-powered Conversational Agents in Addressing Mental Health Problems

Abstract

Millions of people struggle with mental health issues every year, which increases the demand for healthcare services. One emerging technology with the potential to help address the resulting shortage in health care providers and other barriers to treatment access is conversational agents (CAs). CAs can provide a virtual audience that ensures anonymity and allows social disclosure. This paper highlights the significance of CAs and how altering their agency and making them emotionally intelligent can enhance user interactions. The paper will showcase case studies and personal experiences - (i) Diarybot - A chatbot for expressive writing, and (ii) MEEP - Interactive AI Robot Pet (My introduction to HCI/d project).

Introduction

"We may hope that machines will eventually compete with men in all purely intellectual fields."

- Alan Turing (1950)

Artificial Intelligence (AI) is a rapidly developing field in this modern era and has found numerous applications in various aspects of life. It has changed the way we live, work, and interact with one another. Recently, AI has been explored for its potential in mental health care. Each year, millions of people struggle with mental health problems which often go unaddressed due to various obstacles such as cost, access to care, and stigma associated with seeking help. Conversational agents (CAs) are one emerging technology that can serve as a solution. CAs, when powered by an Artificial Intelligence system, an agent that empathizes with people and gradually learns about their moods and behaviors, could offer significant benefits to people.

In this paper, I will argue that conversational agents can help address the obstacles that lead to unfulfilled healthcare needs. I will support this claim using examples of current research and development in this field. I will also discuss the potential implications of this technology for

mental health care. In the following sections, I will first explain the major obstacles to mental health care, then examine how CAs as "actors" connect to their users and learn from their interactions to become more intelligent and thus help to address these obstacles. Finally, I will explore how CAs can become "creators" by becoming emotionally intelligent and adapting to the behavior of their users to better understand and communicate with them (Cila et al., 2017).

Virtual Therapists - Conversational Agents

At present, people consult a psychologist in person for their mental health-related issues, who, over a series of consultations, would discuss their situation, evaluate their responses and discuss necessary measures for their mental well-being. While this is the best available treatment right now, there is limited availability of therapists (Ahmad et al., 2022). Moreover, many individuals may feel reluctant to ask for help for their mental health problems due to the stigma associated with them. This is where Conversational Agents (Virtual therapists) come into the picture and can provide a more frequent and personalized evaluation. By offering a safe and non-judgmental space for individuals to talk about their emotions and experiences, AI-powered CAs can help to reduce feelings of isolation and improve overall mental well-being.

One significant benefit of an AI therapist is that it would always be available. It offers critical assistance in unforeseen times of need or is helpful when someone is in the mood for a more analytical discourse. Through recurring meetings, it can learn considerably more about the person's conduct than a human therapist could, and it might also give reminders to keep the person on track. In addition, because of the pandemic, people are slowly adopting telehealth techniques. They may soon find it extremely common to receive advice from an agent on a computer or mobile screen (Czerwinski et al., 2021).

In their paper, Cila et al. talk about the different social relationships of technology with people's everyday practices. Conversational agents being 'actors,' also known as interventionists, act autonomously and adapt according to their users' behavior (patients). It interprets data gathered from its patients like a 'collector' and responds to it through assessments. Just as actual therapists, CAs continually evaluate their patient's responses, use that information to structure what is discussed, and adapt both content and presentation to ensure the best outcome, further cementing the metaphor of CAs being 'actors' (2017).
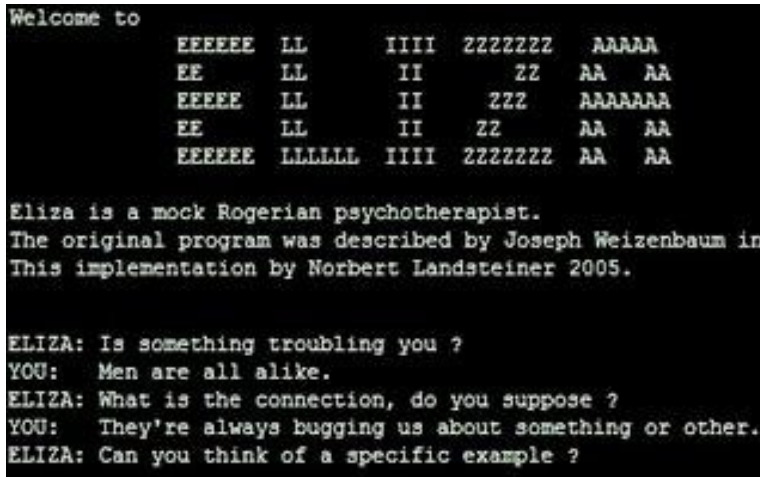
**Fig 1: ELIZA - A natural-language processing program**
Source: https://tinyurl.com/3ura5psp

Fig 2: Woebot - Therapeutic mental health CA
Source: https://tinyurl.com/yc4aek3m

The concept of a virtual therapist is not new. In the 1960s, Joseph Weizenbaum of MIT created scripts for the ELIZA, a natural-language processing program, which frequently repeated users' statements back to them in a highly simplified simulation of psychotherapy (Weizen- Baum, 1966). There is rising optimism that conversational bots will soon be used in mental health to automate some functions of clinical diagnosis and treatment. Three particularly well-known AI-based therapeutic mental health CAs have developed in recent years: Woebot (Woebot Health, 2021), Wysa (Wysa, 2021), and Diarybot (Park et al., 2021). These CAs are freely accessible mobile phone applications that offer counseling services to assist people in managing the signs of anxiety and depression. In the next section, I will illustrate the benefits of CAs through the case study of "Diarybot."

"Diarybot" - A chatbot for expressive writing

One case study where conversational agents proved to be social actors and helped address its users' mental health is 'Diarybot,' a chatbot assistant for expressive writing. Psychology has long known that writing eloquently about trauma helps enhance one's physical and emotional well-being. Just writing about a traumatic event for 15-20 mins for about 3-5 days has proven to enhance one's immune functions and mood and reduce doctor visits (Pennebaker & Beall, 1986). In this regard, 'Diarybot' provides a virtual safe space that maintains privacy while enabling social disclosure.
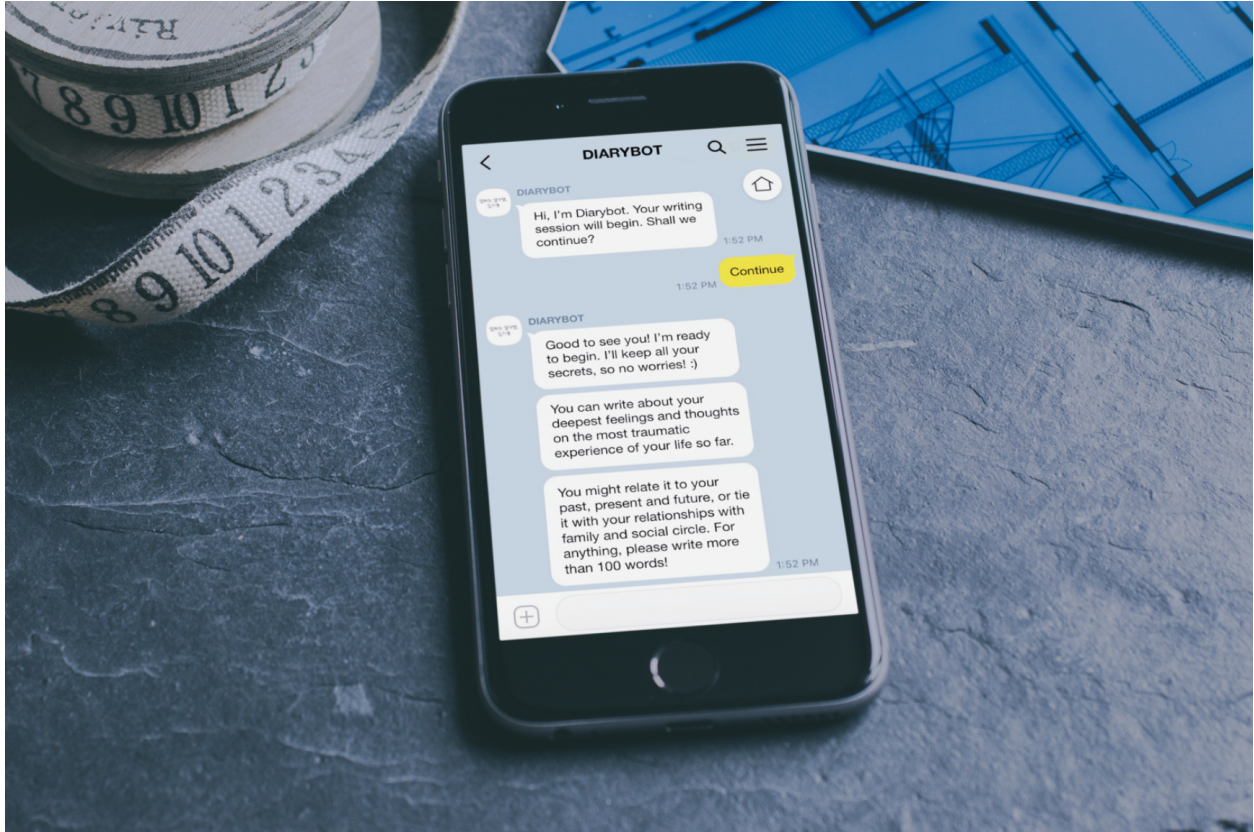
Fig 3: A chatbot assistant for expressive writing - Diarybot
Source: https://doi.org/10.1145/3461778.3462143

In this study, two chatbot versions were created, Basic (collector) and Responsive (actor), to examine the writing experience with and without bot follow-up interactions. The results of this 4-day user study involving 30 participants showed that Diarybot's social disclosure feature stimulates narrative writing in a Basic chat with relative ease and emotional expression. However, a respondent conversation facilitates the social acceptability of the bot and, in the process, offers direction for self-reflection among users. Increased interaction in the responsive version resulted in a perception of the bot's social acceptance and placed prompts provided an opportunity for self-reflection and cognitive organization (Park et al., 2021). If properly developed, conversational agents can provide interactions that help users better navigate their cognitive processes and help their stories avoid digressive or gloomy thoughts, a counter-factor that is frequently emphasized in expressive writing.

MEEP - Interactive AI Robot Pet

In the 'Introduction to HCI' course, my team identified several psychological impacts of excessive gaming. To this end, we designed an interactive AI robot pet to address gamers' mental health. The idea was to provide emotional support to gamers that cheer the players up with music, dance movements, and cute expressions and also reminds players to sleep on time, encourages them to take a walk, and talks to them when they need it.
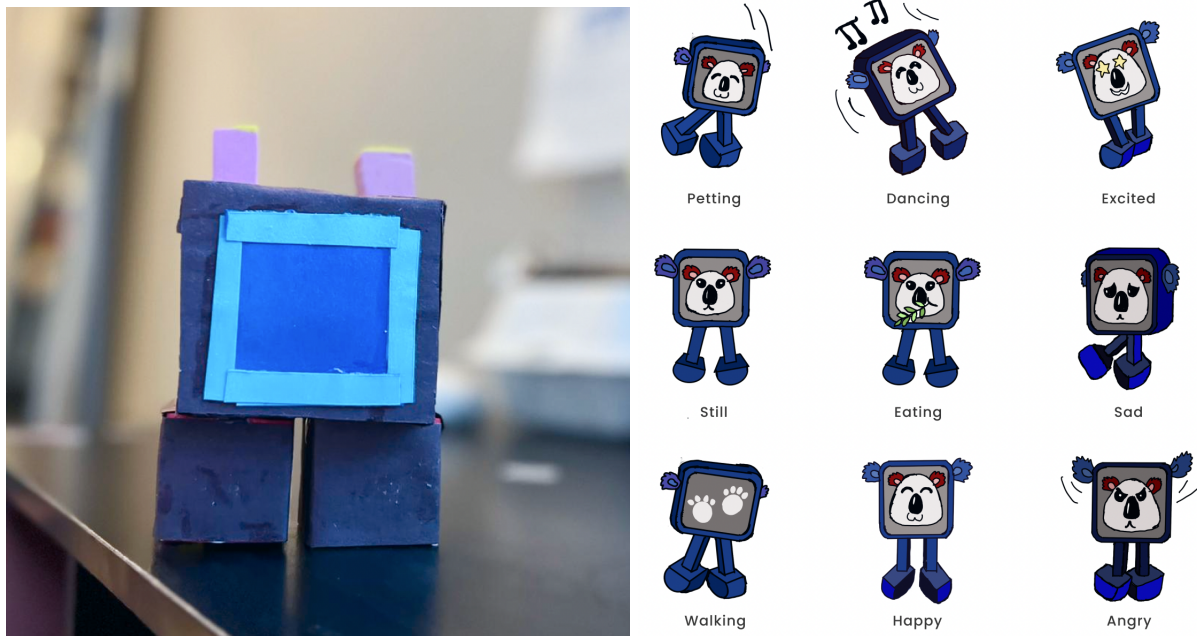

Fig 4 & 5: A mockup of the Interactive AI Robot Pet
Source: Author and team

We designed the robot to have the agency of an 'actor.' The idea was for it to be intelligent and inquisitive about its environment, make its own decisions and evolve its personality based on its surroundings. We tried to incorporate natural language and expressions, which would bond with its users and make them feel comfortable in expressing their emotions openly.

A conversational agent like MEEP could offer significant benefits to gamers by catering to their mental health problems. We conducted a small user study with 4-5 participants and found that users felt more comfortable talking to a robot about their gaming experiences than their family members. The idea of having a companion's presence, one that mimicked actual human characteristics, during a gaming session gave them some mental relief.

Computers as Social Actors

According to the Turing test, a machine should be considered intelligent if a person cannot determine that a machine creates its responses. However, this test is inappropriate for application in mental health care because, according to available data, patients respond psychologically to conversational agents in the same way as they would to an actual person, whether or not they believe it to be one (Miner et al., 2017). As discussed in the 'Foundations of HCI' class, it is a widely known paradigm in HCI that humans subconsciously react to social cues from computers (e.g., conversational agents), mimicking their social interactions with another person. This comes from the fact that humans treat computers as "social actors" (Nass et al., 1994). The main reason is that conversational agents interact with users anthropomorphically, such as by conversing in a natural language. Interactive language use can be very crucial in triggering social responses from users. Some of the other social cues that CAs can use are human-like avatars, names, vocalization, and gestures.

In the research by Lucas et al. about virtual assistants, one group was briefed that the virtual assistant used AI to converse with them. In contrast, the other group was told that the virtual assistant was controlled by a person in another room, yet maintaining anonymity. Findings from this research indicated that when people thought they were interacting with a computer, they were less self-conscious and expressed sadness more intensely than people who believed a human was running the conversational agent.

Potential Implications

Although this technology offers several opportunities in the mental healthcare sector, it may have potential implications. One of the main concerns is that if the CAs are not regulated, they may violate some users' privacy expectations. Patients, especially vulnerable populations such as children, may have expectations of privacy and anonymity that may conflict with the ability of a conversational agent to track and share information. A negative consequence could be that users may be less inclined to seek assistance or disclose mental health issues in real-world professional settings if they had a bad experience talking about them with a conversational AI.

Another concern is the possibility that chatbots cannot offer the same level of assistance and care as real mental health specialists. Although advanced CAs may simulate discussions using therapeutic techniques, replacing human therapists with CAs is "not on the near horizon" (D'Alfonso, 2020). Current AI advancements do not allow CAs to engage in the same level of empathy and understanding as trained therapists or counselors. Therefore, a high level of emotional intelligence would be necessary for this form of virtual therapy to be successful.

Changing Agency - Emotionally Intelligent AI

In his quote, Turing talks about a future where machines compete against humans' intellectual capabilities. However, he ignores the immense importance of what may be considered 'Emotional Intelligence.' Like many, I have preconceived notions that AI can never replace humans as they lack the emotional capabilities we possess. But what if they can?

Building conversational agents that are emotionally intelligent would give CAs a new product agency - 'Creators' that are self-aware and have the ability to learn and evolve (Cila et al. 2017). This would open up an immense playground for interactions between CAs and patients. These CAs could detect and understand the user's interests and shifting emotional states to enhance communication. For example, self-aware CAs could prevent someone with borderline personality disorder from committing suicide.

Further, since humans view CAs as social beings, they associate different personality traits with them. In the research (Moon & Nass), participants ascribed an extroverted or an introverted personality to the CA depending on the strength of the CA's language, the level of apparent confidence, and the order of the interactions. This means that the emotional capabilities of an AI can be powerful in making users feel comfortable in sharing their feelings openly.
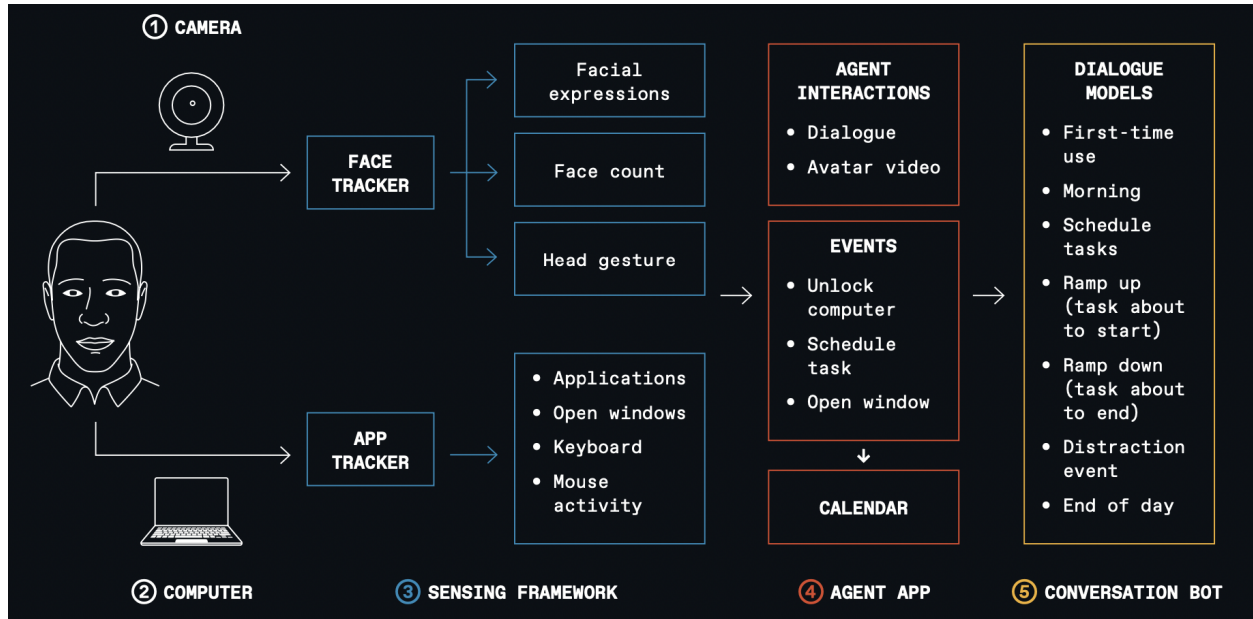
Fig 6: A software to track users' behavior and emotional state
Source: https://doi.org/10.1109/mspec.2021.9423818

Today, much research is going on to imbue AI with emotional intelligence, building on work in psychology, human-computer interaction, linguistics, neuroscience, electrical engineering, and machine learning. Tiny and wireless sensors can be used to estimate a person's emotional state. Mobile phones and wearable technology can also be utilized to investigate visceral human experiences in contexts where feelings are important (Czerwinski et al., 2021).

Conclusion

In conclusion, conversational agents, can significantly improve mental health care in the sector. By providing 24/7 access to mental health resources, reducing the stigma surrounding mental health, and providing personalized support, these agents can help ensure that individuals receive the support and assistance they need to manage their mental health issues.

That said, current technology in conversational agents is not living up to its full potential. Current CAs are trained to carry out tasks but cannot transfer their acquired skills to other domains as humans can. Unline AI humans can use their gut instincts and emotions to navigate through unfamiliar daily situations. Therefore, with the present technology, mental healthcare CAs should be used in conjunction with actual therapists as a support system.

References

1. Ahmad, R., Siemon, D., Gnewuch, U., & Robra-Bissantz, S. (2022). Designing Personality-Adaptive Conversational Agents for Mental Health Care. *Information Systems Frontiers*, *24*(3), 923–943. https://doi.org/10.1007/s10796-022-10254-9

2. Park, S., Thieme, A., Han, J., Lee, S., Rhee, W., & Suh, B. (2021). "I wrote as if I were telling a story to someone I knew.": Designing Chatbot Interactions for Expressive Writing in Mental Health. *Designing Interactive Systems Conference 2021*. https://doi.org/10.1145/3461778.3462143

3. Czerwinski, M., Hernandez, J., & McDuff, D. (2021). Building an AI That Feels: AI systems with emotional intelligence could learn faster and be more helpful. *IEEE Spectrum*, *58*(5), 32–38. https://doi.org/10.1109/mspec.2021.9423818

4. Junhan Kim, Yoojung Kim, Byungjoon Kim, Sukyung Yun, Minjoon Kim, and Joongseek Lee. 2018. Can a Machine Tend to Teenagers' Emotional Needs? A Study with Conversational Agents. In Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18). Association for Computing Machinery, New York, NY, USA, Paper LBW018, 1–6. https://doi.org/10.1145/3170427.3188548

5. Miner, A. S., Milstein, A., & Hancock, J. T. (2017). Talking to Machines About Personal Mental Health Problems. *JAMA*, *318*(13), 1217. https://doi.org/10.1001/jama.2017.14151

6. Matthewman, S. (2011). Theorizing Technology. *Technology and Social Theory*, 8–28. https://doi.org/10.1007/978-0-230-34395-5_2

7. Ekbia, H. R. (2008). Artificial Dreams. *Cambridge University Press*. https://doi.org/10.1017/cbo9780511802126

8. Cila, N., Smit, I., Giaccardi, E., & Kröse, B. (2017). Products as Agents. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3025453.3025797