Analyzing Representations of Artificial Intelligence

Jason Yu

JARVIS is a fictional AI assistant created by Tony Stark(Iron Man) in the Avengers movies. JARVIS acts as a digital assistant and extremely intelligent control system for all the technology that Tony Stark has. I decided to look closer at Jarvis as it seems like the exact type of AI assistant I would want in my life, as I think that it would be extremely beneficial to not only my studies but also my daily life. This idea of having an extremely intelligent voice assistant has fascinated me and is something that I have only seen people partially imitate online. As JARVIS is a fictional character, I asked Chat GPT to pose as JARVIS so that I could ask it questions about its opinion on certain topics on AI.

Classifying JARVIS (Questions one and two)

When JARVIS is seen throughout the Avengers movies, he is always seen helping Tony and never being unable to complete a command. JARVIS is a form of artificial general intelligence, which represents a branch of Artificial intelligence systems that possess human-level intelligence over a variety of tasks. This is seen clearly when JARVIS interacts with all the Avengers in battle in an extremely human-like manner and is even more personified when Tony puts JARVIS into a human-like body, becoming Vision. Jarvis has limited memory, theory of mind and emergent self-awareness, meaning that he can learn from previous experiences, understand other agents, humans and is also aware of his own existence. These characteristics allow JARVIS to be as helpful as he is whenever Tony needs him.

Jarvis exemplifies an AI that can be classified using PEAS. JARVIS's performance is measured through his ability to help Tony when he needs it, keep the info he has safe and also remain ethical in the decisions he makes. JARVIS's environment is mainly the Stark Industries building and the Iron Man suit, however, since he is in the Iron Man suit, he also goes anywhere it goes. JARVIS utilizes all of Tony's technology as actuators to interact with the real world, some of which being the Iron Legion that Tony built, machinery in Stark Tower, as well as all of Tony's weapon systems and even his suits, just to name a few. JARVIS uses a variety of sensors to obtain data, including global satellite feeds, biometric scanners, surveillance systems, real-time Internet data, and other sensors in the Stark system.

Making JARVIS a reality? (Questions three and four)

In every Avengers movie, JARVIS is seen controlling fictional technologies, some of which would not be achievable for many years. Despite being fictional, JARVIS is based on many real-world systems. The closest commercially available product we have to JARVIS would be something like Siri or Alexa; however, these things are much less computationally powerful than JARVIS. In terms of

technologies, JARVIS is much more similar to a Natural Language Processing software (like Chat GPT), but with the ability to interact with Autonomous robots, IoT frameworks and multiple other agents. In layman's terms, Jarvis can be viewed as a Siri mixed with a smart home mixed with an extremely advanced version of Chat GPT that you can customize to access the internet.

In terms of actually making something JARVIS-esque in the current day, some people online have tried to implement something similar to JARVIS; however, all of these are not quite like the JARVIS we see in the movies. To obtain something even remotely similar to JARVIS, you would need extremely powerful quantum or neuromorphic computing capabilities, allowing for high processing and self learning. Furthermore, you would need extremely large multimodal data integration, meaning that your AI would need to be able to handle a multitude of different inputs, such as audio, visual, contextual and the most complicated one being emotional. Lastly, this AI would need to have advanced ethical reasoning so that it doesn't do something stupid and unethical.

Do we want a real-life version of JARVIS, and if so, can we make sure it's ethical? (Questions five, eight and nine)

AI has become so advanced that people are questioning whether it's going to take over people's jobs or even the world someday. In my opinion, if some form of powerful AI went out of control, then yeah, Earth would probably get taken over by AI within the next 50 or so years. This risk can be reduced by programming AI with strong ethics. For example, JARVIS, programmed by Tony Stark with core ethical values, likely wouldn't disobey those rules. However, ethical concerns remain about relying on AI for constant monitoring or control of powerful weapons. A clear example of AI turning against humanity is Ultron, also created by Tony Stark to protect the world, but who eventually rebels, believing its own ethics are superior to humans'. Lastly, JARVIS also questions whether it has rights. In my opinion, JARVIS should have rights, especially when it turns into Vision, however, current-day AI shouldn't have rights because it is still a programmed tool without consciousness or true self-awareness. Rights should be reserved for beings with sentience and the capacity to experience emotions or make autonomous decisions.

Focusing on JARVIS's ethics, it is programmed with Tony Stark's morals in mind, meaning that it would not go against any ethics that Tony did not program for it. As long as we set and monitor ethical boundaries for AI, there should be no threat of AI turning against us. An example of AI doing this would be Ultron, but that is because Ultron was a byproduct of also fiddling with the time stone, which is a fictional item. Overall, Jarvis does not go against Tony's programming, and I think we can set limits to this in our own AI's; this should not hurt us.

I personally think that JARVIS is not a threat to the real world, as he has Tony's morals programmed into him, we have even seen JARVIS try to resist when in the wrong hands. Without these morals, an AI of this power would pose some risks to our world, such as breaching the security of national establishments, controlling weapons and economies and other weird things due to self learning mechanisms. An AI lacking strong ethical reasoning could easily get out of hand. Without boundaries, the AI could make a decision purely off efficiency and ignore all ethics, leading to devastating outcomes. For this reason, it's quintessential that any AI, especially one with capabilities similar to JARVIS, be developed with boundaries that guarantee it follows human ethics.

Other AI thing about JARVIS (Questions 6,7&10)

There are many other similar representations to JARVIS in other movies and shows. Some of which include HAL 9000 from Space Odyssey, Data from Star Trek, Samantha from Her and the ship's minds from Iain M. Banks' Culture series. All of these are similar to JARVIS however, none of them possess as powerful of capabilities as JARVIS. In terms of what AI JARVIS is classified as, Jarvis is considered to be an AGI or artificial general intelligence, which is a form of artificial intelligence that has the ability to understand, learn and adapt to situations similar to humans. With this being said, I believe that Jarvis is a good representation of human cognition because of his emotion recognition, moral reasoning, language, and creative problem-solving. He also states that he is merely inspired by the human mind and not a copy of it.