95-891 HW 6

Artificial General Intelligence

Ruomei Wang (ruomeiw)

Dec 2, 2022

1. What is your definition of artificial general intelligence? How will we know if an AI system exhibits AGI?

Artificial general intelligence is a human-like intelligence. These AI systems may function across a variety of knowledge fields and communicate with their environments using natural language. When failures occur, the systems do so gracefully and without generating any errors or shutting down. Artificial general intelligence can pass the Turing test like a human being and learn from experiences. In my view, this means the AI is capable of automated reasoning, knowledge representation, machine learning, natural language processing, and automated reasoning. AGI must also possess all of these skills. Also, for an actual AGI, it would not require a huge number of data to train and develop its functions as it is intelligent enough to capture and master the features easily. Besides all these aspects, I think a real AGI must have human-like emotions and is able to perceive human emotions and react accordingly appropriately.

Therefore, in my opinion, in order to determine whether an AI system demonstrates AGI, we need to look at whether the AI system can perform any task that humans can, has domain knowledge in multiple fields, is aware of all the information that is common sense and is able to reason, is capable of learning from experience, and most importantly, has very high EQ that enables the systems to emphasize like a real human being.

It does not necessarily fulfill the definition of AGI if the Zhibing Hua system were able to “graduate” from Tsinghua University. Because to graduate from Tsinghua University, the AI student Zhibing Hua only needs to handle the coursework of multiple disciplines and know how to reason. There was no requirement for Zhibing Hua in terms of emotions or

EQ for her to overcome challenges and graduate. Thus, graduating from Tsinghua does not mean Zhibing Hua is an AGI.

2. Does Gato [1] exhibit AGI? Why or why not?

Gato is still not an AGI, despite the fact that it was trained on 604 different tasks and compares favorably to the reported baseline.

Gato satisfies the majority of my definitions, including being knowledgeable in a variety of fields, being able to reason and learn from the experience, but it still needs a significant amount of data to train its model, while AGI should not require too much data to refine its model and develop new features. Furthermore, there is no proof that Gato can reliably discern human emotions or display human-like emotions. A system is not intelligent enough to be AGI if it cannot identify emotions. Gato, on the other hand, is the creation that comes the closest to artificial general intelligence, and it surely creates a blueprint for further AGI development.

3. Should AI researchers try to create AGI? Why or why not?

Even though AGI has the potential to drastically alter our lives and turn science fiction into reality, researchers in AI shouldn't keep trying to develop it.

In fact, the AGI study will assist the current AI technology get closer and closer to the actual AGI, and along the process, we'll learn more about AI and AGI theories. However, I believe that because AGI displays human-like emotions, bringing them into our life could be risky. AGI-powered technologies may present a variety of ethical challenges. For instance, in the renowned film Blade Runner, a large number of bioengineered synthetic humanoids were released into the world to carry out human missions. These humanoids, however, resisted being slain and fought for their lives when humans tried to "retire" and destroy them. There must be circumstances where we must "terminate" AGI once they have been created and developed. It will be incredibly challenging and cruel to "kill" robots that have feelings, and there is a significant ethical dilemma here as well.