

## *Philosophy*

### **Module:**

The student-staff pairing from Philosophy worked on a third-year module that focuses on the philosophy of artificial intelligence in our culture.

### **Learning Outcomes:**

Some of the learning outcomes for this module are as follows:

*On successful completion of this module, students should be able to:*

- *Describe and explain the major philosophical approaches to AI*
- *Evaluate the major objections to AI in general, and also to specific examples*
- *Integrate considerations about AI with major questions in other branches of philosophy, including philosophy of mind, philosophy of science and ethics.*

### **Current Assessment Design:**

The assessment design for this module is **summative assessment** in the form of a 2500-word **take-home essay**.

### **Assignment Task:**

The student-staff pairing compared two styles of essay question: one “traditional” essay question from the 2021-22 academic year and one updated question from the 2022-23 academic year when the use of ChatGPT was required for the assignment. One question addressed the topic of the Turing Test, and the other addressed René Descartes’ views about the nature of the mind.

#### ***Essay questions 2021-22***

*Write an essay of approximately 2500 words on one of the following:*

1. *Explain and evaluate the “Imitation Game” as described in Turing’s (1950) paper ‘Computing Machinery and Intelligence.’ What exactly does Turing think it is for? Is it appropriate to call it the ‘Turing Test’? What, in your view, is the most significant objection to it, and what is the best response to that objection?*
2. *Explain and evaluate Descartes’s contention (from his 1637 Discourse on the Method) that it is ‘inconceivable’ that a machine could use language in the way that humans can, and that therefore a machine could never have a mind. Is he right or wrong? Of all the AI models we have looked at, which would have impressed him the most, and why?*

#### ***Essay questions 2022-23***

*Write an essay of approximately 2500 words on one of the following:*

1. *Conduct a Turing Test (i.e., play the 'imitation game') with ChatGPT. Does it pass or fail? How, and why? What kinds of conversations would make it more likely to pass, or more likely to fail? What does this show about the Turing Test? What does this show about ChatGPT?*
2. *Descartes thought that the ability to use language was an 'irreducible' feature of the mind, that it was the marker of human mentality, and that it was 'not conceivable' that a machine could do so. Does ChatGPT prove that he was wrong?*

*Have a conversation with ChatGPT about the topic to back up your view (you could even ask it to role-play the 'voice' of Descartes) and see whether there are features of your conversation that support your view, or Descartes's. Does your conversation suggest any better candidates for an irreducible mark of human mentality (i.e., that you have, but ChatGPT lacks)?*

### **Associated Risk of GenAI:**

As is the case with all take-home essays, the risk of academic misconduct through the use of GenAI is high.

### **Mitigation:**

In this instance, the lecturer chose to integrate the use of GenAI in their assessment design in an attempt to prevent or discourage the use of GenAI in a way that would be academically dishonest. This integration also provided students with a guided introduction to the technology, enabling them to learn best practice in using GenAI technology.

Students were specifically told to include screenshots and transcripts of their interaction with ChatGPT, as well as provide a rationale for the prompts that they used with respect to the philosophical issues and texts discussed in class.

### **Breakdown of process:**

The student partner was asked to compare and contrast 2021-22 assignments with the updated 2022-23 versions. Of particular interest to the lecturer was whether the student would adopt different approaches, whether they thought they had still learned the material, whether they were less likely or tempted to engage in academic misconduct, and whether either approach was more enjoyable.

The updated assignments, those that ethically integrated the use of GenAI in their design, enabled the student not only to learn about the topic of the assignment, but also about the technology itself (e.g., its limitations, uses to which it could be put in other contexts).

It was found that the student was still able to demonstrate their mastery of the intended learning outcomes.

### **Observations:**

This is a good example of the ethical integration of GenAI into assessment design. The lecturer felt that both approaches aligned with the learning outcomes, but in slightly different ways. Their previous approach was a standard explain-and-evaluate type of essay, which assessed the student's grasp of the material by asking them to report on it. The newer approach is somewhat more practical — since they demonstrate their understanding of the approach by including screenshots of their interaction with ChatGPT and explaining why they chose the prompts that they did, based on the material studied. However, both approaches can fulfil the learning outcomes.