

CSC485 Final Assessment

Wilson Wei-Sheng Hsu

June 2020

1. (2 marks) What is the “Turing test,” according to the viewpoint of an AI enthusiast in the year 2020?

The modern day point of view of an AI enthusiast on the “Turing test” is that in order to say that an AI has passed the “Turing test”, it needs to be able to hold a conversation with human interrogators, such that the human interrogators cannot tell that they are talking to an AI. Such example can be found like Google Duplex, where the AI successfully made a reservation with a human hair dresser without them noticing that they are speaking to an AI. And as the time pass, modern day Turing test adapted the Winograd Schemas, which is a better way to assess a computer’s linguistic and cultural fluency.

2. (2 marks) What test did Turing originally propose in this 1950 article, and how does it differ from the answer to question (1)?

Turing proposed the imitation game test to answer the question “Can machines think?” In the set up of the test, there will be an human interrogator, human participant, and a digital computer participant. The human interrogator’s job is to ask questions and try to distinguish between the two participant. The human participant’s job is to let the human interrogator believe that they are human. The digital computer’s job is to provide answers that would naturally be given by human, such that the interrogator cannot distinguish the two.

It is different from the answer in question (1) in a way that human interrogator in the original Turing test knows about the existence of the digital computer participant, but the modern day Turing test is typically set up with only the human interrogator and the AI participant that tries to mimic human. As the modern Turing test adapted the Winograd Schemas, it is also way more challenging than the original Turing test that Turing proposed back in 1950.

3. (6 marks) Question Answering in 2020-style Turing Test

First of all, I will be collecting some questions from online forums with each one of their answer set using an online information retrieval program that would classify each question into factoid, narrative and hybrid type. And add some twist to the question set by collecting questions and responses from tourist spot information booths around the country, or maybe presidential debates. Weight those answers gotten from forums base on the relative up-vote or down-vote they have. Create a table for those most frequently used words in answers for each question and give them scores based on the weight. After having the question and score table for most frequently used words, we can start asking the contestants questions, marking their response base on the words they used in their response and the time they used to answer.