

Due Date: January 15, 2015 at 11:59pm

Submission Instructions: Assignments are to be submitted through LEARN, in the Dropbox labelled *Assignment 1 Submissions* in the Assignment 1 folder. Late assignments will be accepted up until January 17, 2015 at 11:59pm. Please read the course policy on assignments submitted after the official due date. *No assignment will be accepted, for any reason, after 11:59pm on January 17th.*

Announcements: The following exercises are to be done individually.

Overview

In this assignment, you will be asked to read two important early papers on the topic of artificial intelligence, as well as recent commentary on some of the social concerns around this technology. You will then be asked to write a series of brief responses. The purpose of the exercise is to have you spend some time considering the long term goals of AI research, and whether they are possible or even desirable.

1 Computing Machinery and Intelligence

Read Turing's original paper:

- A. M. Turing (1950), Computing machinery and intelligence, *Mind*, 59, 433-460.

This paper is available in the assignment 1 folder on the course webpage.

In this paper, Turing proposes an operational definition of artificial intelligence based on what is now called the “Turing Test”. He also discusses several objections to his definition, and to the project of creating artificial intelligence in general.

Write a brief (500-700 words total) response to the following questions:

- (2 pts) What is the imitation game?
- (4 pts) Why does Turing think that the imitation game is worth studying as a test of intelligence?

- iii (4 pts) Pick **two** objections from section 6 of the paper, and summarize both the objections and Turing's responses to them.
- iv (4 pts) Pick **two** objections from the paper which you think are still relevant today. Explain why you think these objections are relevant.
- v (4 pts) Find or invent an objection not considered by Turing. Describe it briefly (with citation if applicable), and offer a response to it.

2 Minds, Brains and Programs

Read Searle's paper:

- John R. Searle (1980), Minds, brains, and programs, Behavioral and brain sciences 3.03, 417-424.

This paper is available in the assignment 1 folder on the course webpage.

In this paper, Searle presents an argument against computers being able to think, called the "Chinese Room" argument.

Write a brief (400-600 words total) response to the following questions:

- i (4 pts) What are the differences between "strong" and "weak" AI?
- ii (2 pts) What is the Chinese Room?
- iii (2 pts) Why does Searle think that the Chinese Room invalidates strong AI?
- iv (4 pts) Pick **two** of replies from I-VI in the paper, and summarize both the reply and Searle's critique of it.
- v (3 pts) What does Searle think is the main difference between human cognition and Strong AI?
- vi (3 pts) Do you agree with Searle? Why or why not?

3 AI in the Media

There has been considerable recent discussion of the perils of AI in the media, starting with Stephen Hawking et al.'s piece in the Independent.

Read the original opinion piece, which can be found via the assignment 1 folder on the course webpage.

Write a brief (no more than 150 words total) response to the following questions:

- i (5 pts) What is the piece's main concern?
- ii (5 pts) How plausible do you find the scenarios Hawking poses?
- iii (4 pts) What do you think AI researchers **should** do, if anything, to avoid these scenarios?