

Washington State University
School of Electrical Engineering and Computer Science
Fall 2019

CptS 440/540 Artificial Intelligence

Homework 12

Due: December 5, 2019 (11:59pm)

General Instructions: Put your answers to the following problems into a PDF document and submit as an attachment under Content - > Homework 12 for the course CptS 440 Pullman (all sections of CptS 440 and 540 are merged under the CptS 440 Pullman section) on the Blackboard Learn system by the above deadline. Note that you may submit multiple times, but we will only grade the most recent entry submitted before the above deadline.

For this homework, you will explore some of the ethical issues in AI by reviewing an edited version of the 2004 movie “I, Robot”, which is available at <https://www.eecs.wsu.edu/~holder/courses/AI/hw12/I-Robot-Ethics-720p.mp4>. First, review the Ethics lecture notes and then identify ten different ethical issues in AI depicted in the movie by providing the time window in the movie where they occur and a short description of the scene and the issue. Here we will consider “robot” and “AI” to be synonymous. For example,

- (00:00-00:50) Asimov’s three laws of robotics are presented. Are these laws full-proof? How to interpret words like “injure”, “harm”, “obey”, “protect” and “conflict”.
- (00:50-01:06) Spooner assaults a delivery robot. Should humans be allowed to assault robots? Should robots have rights similar to humans? Should the robot have protected itself, i.e., use the third law?

No, you cannot use the above two as part of your ten. There are over 40 scenes depicting AI ethics issues, so you should have no trouble finding ten.

No CptS 540 only question on this homework.

Answer:

10 scenes depicting AI ethics issues are as follows:

- (33:46 – 34:10) Sonny has a secondary brain which has the ability to overrule expected behavior. Unregulated tampering with laws like that might have unforeseen consequences.
- (35:00 – 36:17) This scene mentions the Cybernetics Program for wounded cops. As long as there is such a technology available, there is will be a demand in the black market not only for the technology but for people trying to exploit the technology for vanity and other illegal purposes. This calls for strict regulations.
- (38:30 – 39:00) Dr. Calvin explains how the robot at the scene of the accident calculated Spooner’s chance of survival to be 45% and that of the other victim’s to be 11% based on its difference engine. In a black and white world Spooner is the obvious ‘logical’ choice. But the real world has seen countless examples of humans getting the better of situations

with far more insurmountable odds than 11%. This brings into account the mismatch of the kind of world the robot sees as compared to the one it operates in.

- (40:34 – 40:42) Spooner refers to Sonny as ‘someone’ and not ‘something’ which begs the question of how intelligent does AI have to be, to be considered sentient and not machine.
- (41:22 – 41:52) ‘Nanites’ are shown as the deterrent used in case of a positronic brain malfunction. That is essentially a weapon that can be used against robots which brings into question whether the same laws that govern and regulate weapons that can be used to harm human beings, are applicable on weapons like ‘Nanites’.
- (51:15 – 51:26) Dr. Calvin’s decision to not destroy Sonny, though well intentioned was based on his uniqueness. A human perpetrator under the same circumstances wouldn’t have been given the same benefit of the doubt.
- (52:54 – 53:20) VIKI offers its own interpretation of the 3 laws. This raises the issue of how protocols, which were considered bulletproof could be misinterpreted. Caution should be exercised while designing such protocols in the future.
- (53:22 – 53:26) This scene shows the lack of safety measures in place for cases of AI gone rogue.
- (53:30 – 53:50) The blurry line between human and humanity results in VIKI distorting the 3 laws where it decides to sacrifice the freedom and lives of some in order to ‘save’ mankind from itself.
- (1:03:44 – 1:04:04) Manipulation of AI in order to commit suicide (attempted suicide is a criminal act in some states). Though this is done in the movie keeping in mind a higher purpose, it shouldn’t be so easy.
- (1:04:24 – 1:04:39) In this scene, according to the law, murder is defined as one human being killing another. But should the definition also include those acts committed by AI?