Homework 1

Alessio Scaravetti, Sterpin Enrico, Radinovic Jovana

October 21, 2023

1 First Exercise

1.1 Read Turing's original paper on AI (Turing, 1950). In the paper, he discusses several objections to his proposed enterprise and his test for intelligence. Which objections still carry weight?

Turing introduced the Turing Test to determine if a machine can have intelligent behavior indistinguishable from humans. He also had some objections and the objections that still carry weight are:

Theological Objections(questions about whether machines can possess a soul or consciousness)

Mathematical Objection (claims that there are mathematical theorems that prove limitations on what a computer can do)

The Argument from Consciousness (suggests that because machines lack consciousness, they cannot be truly intelligent)

1.2 Are his refutations valid?

Judging by understanding of computing in his time they were valid although some of his refutations were subjective.

1.3 Can you think of new objections arising from developments since he wrote the paper?

Some of the objections arising since his time are

Ethics and Moral Implications(new objections have emerged regarding the ethical and moral implications of creating intelligent machines)

Safety and Control(new objections have emerged regarding the ethical and moral implications of creating intelligent machines)

Technological Unemployment(new objections have emerged regarding the ethical and moral implications of creating intelligent machines)

1.4 In the paper, he predicts that, by the year 2000, a computer will have a 30 percent chance of passing a five-minute Turing Test with an unskilled interrogator. What chance do you think a computer would have today?

Now we can see that his predictions did not occur, although AI has made some significant progress, achieving human-level natural language understanding and general intelligence remains a challenging problem that still haven't been resolved.

1.5 In another 25 years?

It's hard to predict what is going to happen in the future although AI progress is rapid predicting this kind of thing is uncertain. We are going to see significant progress in AI capabilities but achieving human-intelligence is a question no one has an answer yet.

2 Second Exercise

Summarize the pros and cons of allowing the development, deployment, and use of lethal autonomous weapons.

Pros

1. ***Less mortality*** (Replacing human soldiers with autonomous weapons on dangerous missions, by doing that we reduce the risk of casualties)
2. ***Advancement in technology***(A lot of inventions that raised the quality of human life were first used in the military eg. the internet)
3. ***Military advantages:*** (Autonomous weapons systems act as a force multiplier, fewer warfighters are needed for a given mission, and the efficacy of each warfighter is greater, reducing human casualties, also autonomous weapons systems expand the battlefield, allowing combat to reach into areas that were previously inaccessible. Robots are better suited than humans for long-duration sorties)
4. ***Moral justification*** (Roboticist Ronald C. Arkin believes autonomous robots in the future will be able to act more "humanely" on the battlefield because they do not need to be programmed with a self-preservation instinct, potentially eliminating the need for a "shoot-first, ask questions later" attitude, and also the judgments of autonomous weapons systems will not be clouded by emotions such as fear or hysteria, and the systems will be able to process much more information than humans without discarding or distorting it to fit preconceived notions)

Cons

1. ***Lack of Emotional thinking*** (Autonomous weapon don't have empathy so it's difficult to give them life-or-death decisions)
2. ***Moral and Ethic problems*** (It's a violation of human rights letting the autonomous weapon decide whether someone should live or die)
3. ***Hacking and misuse are possible*** (It's possible to hack an autonomous weapon and use it for the wrong purpose)
4. ***Reduced incentive for diplomacy***(It's easier for countries to give instructions to a machine than to negotiate with other countries)
5. ***Lethal autonomous targeting:*** (Autonomous weapons systems will find it very hard to determine who is a civilian and who is a combatant, which is difficult even for humans. Allowing AI to make decisions about targeting will most likely result in civilian casualties and unacceptable collateral damage.)

3 Third Exercise

Github link with a solution to the third exercise