

CS/PHIL 201 Discussion Prompt Response

Name:	Elita Danilyuk
Other Group Members:	Dan Butcher, Amanda Elbaz, Kevin Finger, Chris LaBerge

Response to Prompt for Module # 2 on topic Programing Ethics in Robots

Assume it is possible for a machine to have consciousness and make moral decisions. Make an argument about which moral theory (Utilitarianism, Deontology, Virtue) you would prefer for a machine to be programmed with.

If it was possible to program a robot to subjectively learn from its program, have consciousness, and make moral decisions, I would program it with the deontological moral theory. Through deontology or rights-based ethics, the robot would be programmed with universal moral laws. Thus, the robot would have the fundamental rights and, therefore, the wrongs of universally excepted moral laws. The robot could then be further programed with deontological ethics based on the cultural norms, if preferred. Deontological ethics would be the best moral theory to program into a consciousness robot because it is the best starting point for ethical building blocks.

By giving the robot these fundamental universal ethical rules, they can then make the best ethical decisions based on which acts do not violate the rights. Although, this goes to show why the deontology would be the best to moral theory to program into a robot, I do not think a single moral theory would be the best choice to develop into a conscious robot. I think diverse developers from vast cultural backgrounds and ethical ideas should come together to discuss the best possible solution. By having vast backgrounds and ethical norms, the robot could learn from not just one perspective, but many.

Play devil's advocate. What reasons speak against your preferred moral theory and in favor of the others? You might list a few reasons that speak in favor of the other moral theories, or you might list some replies to the reasons you offered in part 1 of this assignment.

One of the reasons deontology may not be the best moral theory to program into a robot is that not everyone agrees on the 'rights and wrongs.' This is why I argue that it would be best to develop with a diverse group and come to a conclusive set of rights.

Utilitarianism is another moral theory that could be programed into the robot. Utilitarianism could be seen as a good choice because it would allow the robot to make decisions for the highest value of happiness, welfare, and well-being. However, this would be extremely difficult to program because of the differences in interpersonal comparisons and possible inconsistencies with justice.

Lastly, virtue ethics would program the robot to focus on acting through virtues and living a good life. This moral theory could be argued the best choice if the purpose of the robot is to live out its own best life or to allow humans to live their best lives. This moral theory could lead to vast issues in the future. The robot may do things to help one human, but it could hurt many others in the process. This would also be difficult to program because, although some virtues are agreed on by most, there may be others that are significantly subjective.

These are some examples of why utilitarianism and virtue ethics could be better candidates to program into the robot.