

CS 4610 – Robotic Science and Systems – VSD Homework

Due by April 1, 2022 at 11:59 PM ET

Note: Complete this assignment with the other members of your project group.

Note 2: Please direct questions to Prof. Vance Ricks (vance@ccs.neu.edu)

In this assignment, you will take some of the *preliminary* steps that constitute a value-sensitive design (VSD)-based approach to technology ethics. This is an exercise in applied ethics, critical thinking, and creativity. Your responses will be in the form of thoughtful answers, rather than code, pseudocode, or mathematical formulas. You **do not** need to be familiar with VSD to complete this assignment successfully. (There is general information available about VSD and its methods at the following website: <https://vsd.ccs.neu.edu>.)

Summary of the issue: Social robots, defined here as “robots that can sense, plan, and act in the context of social interactions with humans”, are the “beneficiaries” of our tendency to anthropomorphize. For example, for a 2007 article in *The Washington Post*, the journalist Joel Garreau interviewed members of the United States military about their relationships with robots. He heard of cases in which robots were awarded Purple Hearts (a US military decoration awarded to those wounded and/or killed while serving). He heard soldiers’ emotional distress over destroyed robots. One robot was built to walk on and detonate land mines. The colonel overseeing the testing exercise for the machine ended up stopping it, saying the sight of the robot dragging itself along the minefield with its one remaining functional leg was too “inhumane”. Military robots have been given funerals with gun salutes. The researcher Julie Carpenter conducted a 2013 in-depth study on explosive ordinance disposal robots in the military. She found that the robots’ operators sometimes interacted with the robots as they would interact with a pet. There are even accounts of soldiers risking their lives to save the robots they work with.

More generally: social robots elicit our built-in tendencies to project intentions and feelings onto robots’ behaviors. That projection, in turn, makes us likelier to respond emotionally to social robots similarly to the ways we respond to nonhuman animals (or even to other humans).

Much of our behavior towards animals and other humans is regulated by many strong and longstanding (though always evolving) ethical norms, cultural traditions, and civil laws. That is not yet the case when it comes to our behavior towards robots. There are many ongoing disagreements about which norms, traditions, and laws ought to apply, and why.

The questions: (a) What, if anything, is unethical about intentionally damaging or destroying a present-day social robot? (b) What are the main ethical values or ethical principles that you’re using to support your answer?

Refinements to the Questions:

Zeroth, this is a discussion only about social robots (e.g., hitchBOT, Paro, Pepper, Siri) *as they currently are*. That means you are writing about objects lacking consciousness, sentience, affect (emotional capacity), etc.

First, assume that you own the robot (or that no one does). This is not a question about the wrongness of destroying someone else's property.

Second, assume it's possible to damage or destroy a robot without also damaging or destroying any bystanders or ecosystems. This is not a question about the wrongness of causing possibly harmful side effects to "third parties".

Third, think about the kinds of empirical evidence that might – or might not! – be available to support your answer. This is crucial if your answer rests on any cause-and-effect claims. For instance, "Beating up a robot is unethical because it makes you more likely to beat up humans (which is unethical)" is an empirical, cause-and-effect claim. It says engaging in one behavior increases the likelihood of engaging in a different behavior. As such, it requires support from validated studies or similarly convincing observational evidence. If there are no such studies or similar evidence, then your answer will be, at best, extremely speculative, not persuasive.

Fourth, "ethical values" are those values that express our ideals and aspirations for how we should live – both individually and collectively. Ethical values include justice, truthfulness, accountability, care, privacy, and respect – among others. "Ethical principles" state ways that we should act on our ethical values. Some ethical principles include, "Give everyone what they are due", "Show respect to your 'elders'", and "Don't manipulate other people", among others.

Format:

Write your answer and your explanation for your answer in **at least two full pages and at most four full pages** (1.5 or double spaced). Write conversationally, but still grammatically correctly. You do not necessarily need to engage in outside research, but if you do use any outside sources, you must credit/cite them in your paper. Because your paper is a group-written document, its answers must reflect the consensus of *all* the people in the group.

Evaluation:

Total possible points: 4

See the scoring criteria on the next page.

4 points	The response provides very clear and strong support for its answer to question (a)	The response provides a very clear statement of at least 4 supporting ethical values and/or principles (total) in its answer to question (b)	The responses to both questions show clear awareness of the “Refinements” constraints by not violating any of them	The total response is between 2 – 4 full pages long, and has almost no grammatical, typographical, or spelling glitches
3 points	The proposed outcome provides clear and moderately strong support for its answer to question (a)	The response provides a clear statement of at least 3 supporting ethical values and/or principles (total) in its answer to question (b)	The responses to both questions show clear awareness of the “Refinements” constraints by not violating any of them	The total response is between 2 – 4 full pages long, and has few grammatical, typographical, or spelling glitches
1 - 2 points	The proposed outcome provides weak support for its answers to question (a)	The response provides a clear statement of 1 or 2 supporting ethical value and/or principles (total) in its answer to question (b)	The responses to both questions show slight or inconsistent awareness of the “Refinements” constraints by violating 1 or 2 of them	The total response is between 1 – 3 full pages long, and has several grammatical, typographical, or spelling glitches
0 – 0.5 points	Only one question answered; excessive writing glitches			