Ethics & Robotics Final Project

Integrating Ethics Discussion in Human-Robot Interaction Class

Pancho Cabrera & Lynn Kirabo

Assigned Course: Human-Robot Interaction (16467)

We were assigned the Human-robot interaction class. This class offers students an introduction to the field of human-robot interaction. The class follows the following format, lectures in the first third of the semester focus on foundational concepts, while the last two thirds focus on the integration of those concepts. At the end of the semester, the students must conduct a replication study on a published HRI paper. This section of this class is usually taken by undergraduate students who come from both technical and nontechnical backgrounds.

Class Constraints

We took the following constraints into consideration while designing our modules:

First, the course already has an established format (i.e., foundations vs application), our recommendations keep this structure in mind and seek to support it.

Secondly, the class size usually leans towards a little over 60 students. Each course TA is usually responsible for about 25 students. It is therefore important that any modifications do not add additional strain on the teaching staff.

Learning Objectives & Justification

In order to support the existing structure of the course, we chose to implement multiple small modules that would lead up to the course's Ethics lecture. We kept the following learning goals in mind while designing the modules:

1. Students should be able to identify any ethical topics embedded in HRI concepts, applications, or conversations.

2. Students should be able to develop vocabulary to express ethical concerns they observe.

3. Students should be able to develop the ability to observe/identify comparisons between ethical events in current news and future implications.

Instructions for use

The assignments accompanying these modules are designed to be low stakes. This is an effort to encourage fluency in Ethics discussions over fighting for grades.

All assignments are designed to be taken by students before the lecture. The lecture will usually start with a 15minute discussion on their reflections. To encourage student interaction, some reflections will include discussion board posts with students given the option to comment on another classmate's post. Course TAs will post the initial prompt on the courses' chosen discussion board (e.g., Piazza or Canvas). This responsibility can be rotated between the TAs.

We have also included a number of quiz samples. These are to be administered in the week after the lecture has been taught. This will follow the format currently used by the class that includes a 10minute quiz slot on the previous week's content.

Sample Module Recommendations

Module 1: What is Ethics Anyway?

We recommend running this module in Lecture 2 (Survey). The course lecture introduces the different topics that will be covered over the course.

Pre-Class Activity: Students will have to read the assigned New Yorker article. After which, they will post a response to the prompt, "*What is one question regarding ethics and human-robot collaborations you have?*"

These should be posted in a Google Document by 11.59PM the day before the lecture. Students can also add +1 to other students' questions.

Class Outline: The Professor can use the questions asked as class discussion opportunities or choose to respond to them in the lecture content.

Assignment: This assignment is worth 2 points.
Assigned Reading: [Did Uber Steal Google's Intellectual Property](#)
[Link to Google Document](#) for Responses.

Module 2: Robots of the Future

We recommend running this module in Lecture 3 (Autonomy). The course lecture focuses on discussing three types of autonomy - teleoperation, autonomy, and shared-autonomy.

Pre-Class Activity: Students will watch the assigned YouTube video and answer the following prompt on Piazza/Canvas: *Do you agree that integration of robots in the workforce will increase HRI? Why/Why not?*

The course TA will start the thread by posting the prompt, the students should submit their responses as comments to the thread. Students should also be encouraged to respond to other students comments by replying to those specific responses.

Class Outline: First 10 minutes of class will be devoted to discussion from the Piazza/Canvas discussion. Here are the [sample slides](#) that can be used to facilitate that portion of the lecture.

Assignment. This assignment will be worth 5 points. ([Sample Prompt](#))
Assigned Video: [Robots of the Future at Boston Dynamics](#)

Module 3: Racial Divide in Speech Recognition Systems

We recommend running this module in Lecture 8 (Survey). The course lecture teaches how robots generate and understand speech as well as the concepts behind expressing and recognizing emotion.

Pre-Class Activity: Students will have to read the assigned NYT article. After which, they will post a response to the prompt on Piazza/Canvas, "*Prompt Question: Eric Schmidt is quoted as saying, 'We know the data has bias in it. You don't need to yell that as a new fact. Humans have bias in them, our systems have bias in them. The question is: What do we do about it?'*

Do you agree that the challenge of bias is an application problem? Why/Why not"

The course TA will start the thread by posting the prompt, the students should submit their responses as comments to the thread. Students should also be encouraged to respond to other students comments by replying to those specific responses.

Class Outline: 15 minutes of class will be devoted to discussion from the Piazza/Canvas discussion. Here are the [sample slides](#) that can be used to facilitate that portion of the lecture.

Assignment: This assignment will be worth 5 points. ([Sample Prompt](#))
Assigned Reading: [There is a racial divide in speech recognition systems, researchers say](#)

Module 4: Predictive Policing

We recommend running this module in Lecture 12 (Data & Analysis). The course lecture teaches students how data is analyzed. The lecture is not specific to human-robot interaction topics.

Pre-Class Activity: Students will have to read the assigned NYU Law review article. This is similar to students doing class readings before the lecture.

Class Outline: First 10 minutes of class is usually devoted to quiz time. Students will respond to two questions from the reading.

Assignment:

We recommend the following sample quiz questions:

1. (1 point) Does more data lead to accurate results? Yes/No
2. (2 points) Identify the ways in which data is biased:
   a) data reflects historical practices and policies.
   b) feedback loops.
   c) data may omit certain types of crimes and certain types of criminals.
   d) algorithms should be modified to allow for more accurate predictions

Assigned Reading: [Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice](#)

Module 5: *But*, how far is too far?

We recommend running this module in Lecture 15 (MDPs). The course lecture teaches students about robot policies on decision making.

Class Outline & Activity: This activity will likely take up to 30minutes. It is best placed towards the end of the lecture. The student prompt should be given during class. This will allow for students to think on their feet and get creative.

Students will be presented with a hypothetical scenario:

> *A robot in the Newell Simon hallway is moving back and forth asking random people passing by for help. A student stops to inquire what is wrong, the robot replies that it urgently needs to test signal strength and needs help being hoisted*

up. Student A complies. The robot urges the Student to reach higher or grab a chair from iNoodle. Student A complies, Student B has been watching this exchange and offers hoisting advice to Student A. Now standing on the chair, the Robot tells the student they need to reach it higher and the student complies now, standing on tiptoe while trying to balance on the chair.

Divide the class into two groups: The For & Against Groups. The groups will then defend their stance based on one of the questions below. Students should be given 5 minutes to prepare.

1. Is it ever okay for human beings to implicitly trust robots?
2. Was the robot wrong to ask for help?

Assignment: There is no assignment for this module.

Module 6: Who has the right of way?

We recommend running this module in Lecture 19 (Social Navigation). The course lecture teaches students about robots navigating through social spaces.

Pre-Class Activity: Students will have to read the assigned Bloomberg article. After which, they will post an individual reflection on a Piazza/Canvas assignment. The following questions will guide the student reflection:

1. Identify the different players you think are involved in the creation and use of a robot like the starship robot.
2. When and Where should decisions about accessibility be made? Discuss why?

Class Outline: 15 minutes of class will be devoted to discussing student ideas. Here are the sample slides that can be used to facilitate that portion of the lecture.

Assignment: This assignment will be worth 5 points.
Assigned Reading: My Fight With A SideWalk Robot

Module 7: To drone, or not to drone

We recommend running this module in Lecture 23 (Collaboration). The course lecture teaches students about humans and robots working together to achieve common goals.

Pre-Class Activity: Students watch the Drones trailer & pick 1 of 2 articles to read (NYT & Times articles). After which, they will post an individual reflection on a Piazza/Canvas assignment. The following questions will guide the student reflection:

1. What are your thoughts on the application of drones on different aspects of life, i.e., Warfare, Healthcare?

Class Outline: 20 minutes of class will be devoted to discussing student ideas. Here are the sample slides that can be used to facilitate that portion of the lecture.

Assignment: This assignment will be worth 5 points.
Assigned Video & Readings: Drones Trailer, The Moral Case for Drones, Zipline Delivering Blood.

Module 8: Final Project

We recommend adding a required ethics reflection to the Final Project Report, and a question leading up to the reflection in Mini-Project 3. Below we have provided the materials with the necessary changes corresponding to our recommendations:

Mini-Project 3: Project Plan; **Final Project Rubric**; **Final Project Slides**

Module 9: Extra Credit

The following resources may be used as extra credit assignments. We felt that these assignments would simply be a Piazza/Canvas post consisting of the students' reactions and thoughts regarding the corresponding video(s) or reading(s). Students should be encouraged to respond with their opinion about the given reading.

Here is an outline of the resources/topics we presented:

Topic: **Robot Rights and Citizenship (for lecture on Appearance and Anthro)**
        Resources: Meet Sophia, World's First AI Humanoid Robot | Tony Robbins

Topic: **Ethics of Robot Policies - Tabletops and Evacuations (for lecture on MDPs)**
        Resources: see module 5 for hypothetical; People think robots are pretty incompetent and not funny, new study says

Topic: **Social Justice in Tech: Timnit Gebru vs. Google (for Ethics lecture)**
        Resources: We read the paper that forced Timnit Gebru out of Google. Here's what it says.

Topic: Algorithmic Bias: The Case of Robert Williams (for Robots in the Wild lecture)
      Resources: [Wrongfully Accused by an Algorithm](#); [That Ain't Right: AI Mistakes and Black Lives](#)

Honorable Mentions: While the resources below are not linked to a specific lecture, they may also serve as optional reading where you see fit.

Topic: The Origin of the Word 'Robot'
      Resources: [Rossum's Universal Robots](#); [Machine Morality and Human Responsibility](#)

Topic: Environmental Impact of NLP Models
      Resources: [Frontiers in Machine Learning: Climate Impact of Machine Learning](#); [AI and Climate Change: How they're connected, and what we can do about it](#)

Topic: Ubuntu Ethics
      Resources: [From Rationality to Relationality: Ubuntu as an Ethical and Human Rights Framework for Artificial Intelligence Governance](#)