Teaching Ethics by Teaching Ethics Pedagogy
A Proposal for Structural Ethics Intervention

Victoria Dean
Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
vdean@cs.cmu.edu

Illah Nourbakhsh
Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
illah@cs.cmu.edu

ABSTRACT
We report on a reformulated general Ethics and Robotics course, in which we aim to address the twin curricular challenges of exposing computer science students to ethics discourse and establishing a pathway for ethics-oriented modules to be designed into computer science courses across an institution. Given computer science instructors’ lack of time and expertise to build ethics modules themselves, we tasked our students with creating ethics modules for instructors of 11 computer science courses at our university. Our course participants represented a diverse range of backgrounds and perspectives that catalyzed lively discussions and creative ideas for ethics pedagogy innovation. We report on course details, including in-class activities, assignments, and the project. We discuss our findings, including reception from students and computer science instructors and planned updates for the next course iteration. Given the course’s overall success, we share with the hope that others may learn from or adopt our course approach. Materials are available on our website: https://vdean.github.io/16-735-ethics-robotics.html.

CCS CONCEPTS
• Social and professional topics → Computing education.

KEYWORDS
Ethics; course design; pedagogy; robotics

ACM Reference Format:

1 INTRODUCTION
As ethical questions around technology in society arise, we have witnessed a growing interest in ethics integration into computer science (CS) education. Prior work on ethics in computing education typically falls into one of two categories: a) a standalone, full-length course (elective or required); or b) an ethics module integrated into an extant CS course. Each of these serves an important purpose. A standalone course allows for deep discussion and dissection of ethics concepts at a level not feasible in an episodic module. By contrast, a module makes the concepts contextually relevant, showing how ethical thinking interrelates with other aspects of CS design and algorithm work. Neither an introductory course nor a module alone is sufficient, because a truly effective pedagogical solution needs to combine a broad field-introducing analysis with disciplinary relevance through multiple deep dives.

Each of these categories also has challenges. Standalone courses seem independent, and some students may struggle to connect their learnings with areas of CS they will use on a daily basis. On the other hand, developing modules that are relevant to a course requires time and support that instructors frequently lack.

We aim to bridge these strategies through a full-length ethics course that includes a module design project that directly engages our students with instructors of other technical courses throughout the university. In this way, we mitigate the two challenges described above. Students see how ethics concepts relate to other CS areas by designing modules for these courses. Instructors of these courses are afforded the structure and support needed to create an ethics module by collaborating with students on the module design. This project bridges the categories in a win-win situation, introducing pedagogical design to our students in a practical, immediately relevant manner.

In this report, we describe our experience piloting this course in Spring 2021 in a virtual setting via Zoom due to COVID-19. We hope that others may learn from this curricular experiment, perhaps leading to localized implementations at other institutions.

2 RELATED WORK
A 2008 survey [27] of 251 U.S. undergraduate CS programs showed that these institutions do take computer ethics seriously and are teaching it. We focus coverage of related work on those with substantial similarity to our course: innovative full-length ethics course design and ethics module integration in other CS courses.

Ethics Courses. The need and demand for ethics courses, as both electives and requirements in computing degrees, has grown in recent years. Fiesler et al. [11] surveys 115 such technology ethics courses from 94 universities around the world. A number of efforts have innovated on ethics course structure. A computer ethics course in 1999 [29] incorporated active learning through student-led discussions, presentations, field trips, and more. In 2003, a Miami University computing ethics course [25] centered around discussion and class participation. Around the same time, an Ithaca College course [2] incorporated contemporary media (news, videos, and film) and reflective writing for a student-centered approach. More
recently, a Stanford course [23] took a multidisciplinary approach, with instructors from philosophy and political science in addition to CS. Last year, Rice University launched Deep Tech Ethics [10], a course focused on technology as it relates to social justice. Subsets of these courses have similar facets to ours, including the use of current events, discussion-centric class time, reflective writing assignments, and guest lectures.

**Ethics Modules and Course Integration.** In addition to full-length courses, prior work has examined the incorporation of ethics into existing CS courses. Narayanan and Vallor [20] motivates the need for ethics coverage in software engineering courses. A study of syllabi for machine learning courses [24] found that just 12% (22 of 186) of courses analyzed explicitly included ethics content. Followup work [12] found that the majority of such courses that did integrate ethics did so in the last two classes.

Work from 2005 [3] presents four success stories of timely and relevant ethics integration in programming courses. More recent work [26] presents example ethics activities and pilot results from a Human-Computer Interaction course. A 2020 SIGCSE Special Session [7] describes five example assignments blending ethics and technology. A 2019 SIGCSE workshop [19] provided structure for participants to design their own ethics modules, which they could take back to their home institutions. The Embedded EthICS program [13] piloted ethics modules in 14 Harvard CS courses, as created and taught by philosophy PhD students and postdocs. Our approach also leads to ethics module development but does so via our course’s students, who have the time necessary for both pedagogy training and the module undertaking.

3 OUR BACKGROUND

Our teaching team includes specific experience teaching ethics and robotics formally beginning in the early 1990s, initially focusing on a course module approach to embedding ethics discourse in a Robot Program Laboratory class at two institutions nationally. Due to strong student engagement, we expanded the material to a whole-semester Ethics and Robotics course concentrating, initially, on questions of power hegemony and the role of computational and robotics technologies in warmaking and in state control. Following the success of this course strategy, which was open-sourced online and used by a number of institutions thereafter, our more recent experiment combined the forces of the Humanities and CS Colleges to design and deploy an ethics course on technology’s influence on society, focusing specifically on AI, as material specifically designed for first-semester incoming college undergraduates. This course, presented during the students’ first semester of college, provided pathways of analysis and discourse to inform their career orientation and college trajectory decision-making throughout their subsequent four years. The course was popular enough to warrant placement as a multi-year alternative to Freshman English and became available internationally as a textbook.

4 COURSE OVERVIEW

For this reinvention of Ethics and Robotics, we took a student-centered approach to course design. Similar to many ethics courses [2, 10, 23, 25, 29], we aimed to maximize student engagement during class time through discussions and other active learning activities.

Unique to our course, we teach pedagogy and task students with developing modules to convey ethics concepts to others. Learning by teaching can deepen students’ understanding of the material [18, 22, 28]. Thus by having our students design ethics pedagogy, they metacognitively understand ethics in a more profound way. Below, we describe the specifics of our course: first, an overview in this section; followed by in-class activities (Section 5); and finally, assignments and the project (Section 6).

4.1 Learning Objectives

Our goals for students of this course included:

- Critically evaluate the ethics of technology and its societal ramifications.
- Articulate future implications of a new technology through writing and creative storytelling.
- Create an ethics reflection module for curricular deployment in a technical course.
- Identify real-world examples of ethical implications of robotics research and evaluate the impact of related decisions.

4.2 Assessment Breakdown

The primary recurring assessment was journaling (Section 6.1) and reading analyses, worth 40% of the semester grade. The remaining 60% was split equally between class participation in discussion, peer reviews, presentations, and the module project.

4.3 Participants

The 22 course participants included undergraduate, master’s, and PhD students as well as staff from 12 departments, a diverse range of backgrounds and perspectives leading to lively discussion and creative ideas for ethics pedagogy. 16 of the students were in CS.

5 IN-CLASS ACTIVITIES

5.1 Active Learning Exercises

Every class session involved an active learning activity. Our most common format was breakout rooms, as they were easy to implement via Zoom and allowed students to share while interacting with peers. We gave a prompt before each breakout room, most commonly in the form of discussion questions. Breakout rooms varied from 3 to 7 students and lasted between 5 and 20 minutes. Students took notes in a shared Google Doc with a section for each room. The instructors were not present in breakout rooms; instead, we monitored the notes and identified points to surface as a group.

The following are additional one-off active learning activities we used. Students created concept diagrams depicting a corporation and the agents and relationships involved, touching on themes like privacy, profit, and regulation. During our pedagogy unit, students critiqued the learning objectives for our own course, a fun and self-referential way to prepare students for their own ethics module design. Later in the unit, students created their own active learning exercises to deepen understanding of a provided poem, and one breakout group ran their activity with the class. One take-away from this exercise was to show students how class session outlines might not go according to plan and that instructors will frequently need to adapt. Finally, during the week on futuring, as an extension
of the Black Mirror assignment (described in Section 6.2) students created and performed 5-minute adaptations of their screenplays.

5.2 Guest Lectures
Five class sessions had guest lecturers: an emeritus professor in economics and public policy discussing the 4 industrial revolutions, the future of labor, and universal basic income; a robotics professor from another university discussing machine bias, an ethicist discussing their experience with developing course ethics modules; a philosophy professor discussing autonomy and the military; and a design professor discussing futuring and design fiction. Many of these guest visits were preceded by readings and media related to the lecture topic, allowing for the “lectures” to consist heavily of Q&A and active learning, a structure we found to be effective.

5.3 Topics and Readings
The course covered a different set of topics each week with a combination of reading and class discussion. Chronologically, the topics (and corresponding readings) were: surveillance and humanity, labor [14, 21], the industrial revolutions [21], algorithmic bias [1, 9, 15], safety and trust [8], curricular integration [13], learning objectives [5, 17], robotics and society [4], visionaries and public reception [16], and autonomy and military [6].

6 ASSIGNMENTS
6.1 Journal Entries
Journal entries were common assignments throughout the semester, tied to many of the above readings and videos. Students wrote analysis of the assigned media, answering questions focused on student exploration of key concepts. These entries were due 24 hours before class, allowing instructors to surface themes and examples from the submissions to enhance discussion. Each entry received feedback from an instructor to encourage growth. For one such journal entry, students also conducted peer reviews of a classmate’s writing.

6.2 Individual Assignments
The course featured 4 one-off assignments, allowing students to practice and demonstrate skills beyond the written analysis assessed through journal entries. Students gave “from research to news” presentations, juxtaposing a research paper and a news article on a topic of their choice. Another assignment tasked students with creating a concept diagram representing organizational structure in higher education as it relates to embedded ethics in technology learning. In the pedagogy unit, students practiced creating learning objectives for an introductory CS course. Finally, as a fun end to the course, students created Black Mirror-style screenplay outlines to develop their futuring skills.

6.3 Module Project
The course culminated in an ethics module design project in collaboration with an instructor of another robotics or CS course. Each pair of students studied a course and designed complete ethics and reflection module tailored to that course, with suitable rubrics, activity design, reading, and homework assignments. These materials went through multiple iterations incorporating feedback from that course’s instructor. The goal is for the modules to be used by these instructors in future semesters of each course.

We needed instructors interested in creating or updating ethics modules for their own courses. We recruited instructors through announcements to all CS faculty and targeted outreach to teaching faculty. 13 faculty volunteered a total of 16 courses, 11 of which we were able to match with the 11 pairs of students in the course.

7 RESULTS
We use a number of sources to evaluate the course outcomes. Section 7.1 has an example project showing the quality of the resulting module materials. Section 7.2 presents anonymized student feedback from week 7 of the course. Section 7.3 shows the final course evaluations. Finally, Section 7.4 shows the instructor reception.

7.1 Module Examples
Here we provide 2 concrete examples of the modules created by student teams. The full set of module materials are on our website.

7.1.1 Data Science and Machine Learning at Scale: 23andMe Case Study. The students in this smaller, upper-level course come in with a baseline understanding of machine learning and, to some extent, ethics from prior courses. The module is a single class session centered around a real-world case study on 23andMe and genetic data, chosen because there was no clear villain, there were unforeseen consequences, and it would be less familiar than more commonly-used case studies. The module begins with breakout groups taking on roles of case study actors, followed by class discussion on generating alternative actions. The session culminates in an assignment: each student must write an email to their boss at the fictional company BioNano expressing concerns about the company direction. The developed materials include learning objectives, the session outline, readings, a reading quiz, slides, discussion questions, and the email assignment rubric. The instructor was enthusiastic about the module: “[the team] really embraced the challenge... I am going to be stealing this letter assignment for other assignments in class... It’s really nice to give students practice actually acting on your concerns as opposed to just identifying the concerns. It’s been a really great experience for me so thanks.”

7.1.2 Introduction to Computer Systems: Applying Frameworks to Lecture Topic Scenarios. This course is required for all computer science majors; 1500 students take the course each year. The ethics module content integrates into lectures throughout the semester, prioritizing takeaways for future use across 6 topics. Each of these 6 sessions has the same format. Before the lecture, students read a scenario related to the lecture topic and answer “what would you do” questions on a discussion board (expected to take 20 minutes). The lecture features a review of the scenario and group discussions with neighbors (20 minutes of class). After class, students complete a Google form with what they would do in the scenario and how their choices changed based on peer discussion. Form statistics can be shared back with the class and also give the instructor a sense of how the students are responding to the ethics module. The developed materials include the written scenarios and objectives, lecture slides with scenario review, and the post-class form.
7.2 Early Course Feedback

In week 7 of the course, an external teaching consultant from our university led a 30-minute student feedback session without the instructors present. This session involved breakout rooms for surfacing independent feedback followed by a full-group debrief to determine consensus. The consultant anonymized this feedback for the instructors. Below are excerpts from the consultant’s report.

7.2.1 Strengths That Assist in Learning. Below are relevant strengths, the number of students that surfaced each, and anonymized quotes.

Students get a lot out of the discussions, especially small group discussions (16 of 21 students):

- "Discussions are the most helpful to our learning. We get to hear a lot of different perspectives from classmates, professors, and guest speakers."
- "Small discussion breakout rooms are a huge plus. It’s very challenging to get ideas out in the large room, so the small rooms are conducive for good discussions. We really appreciate the zoom chat that promotes discussion."

Journals help students engage with the content and discussion (9 of 21 students):

- "Journals help us engage with the content and discussion."
- "Journaling gives us a chance to think deeper and reflect about the teaching."

The instructors help create a welcoming class environment even while remote (all 21 students):

- "The professors are very relaxed and don’t make you feel less intelligent if you don’t know the background of the topic. It’s nice to be surrounded by students who feel safe enough to share their perspectives but both professors have done a good job to mitigate rude remarks. They encourage healthy discourse, but make sure there isn’t disrespect between the students."
- "This is the first class I have attended that actually cares about making people heard, and consider the suggested opinions at value."
- "[The instructors] almost never dismisses an opinion and make it a welcome environment."
- "Many students don’t feel intimidated to share their thoughts because the professors are very relaxed."

Another teaching consultant observed that we "valued and encouraged students’ contributions to the class by responding to each student’s comments in a positive manner and highlighting important points or connections they brought up."

7.2.2 Suggestions to Improve Learning. Below are opportunities for improvement and, if applicable, how we incorporated the feedback.

Provide more time for student discussion (12 of 21 students):

- "There is so much content being spoken in the discussions, but there isn’t enough time to cover everything... It would be great to have monitoring of the zoom chat to ensure that we don’t get off topic but also bring the zoom chat into the actual discussion, to incentive cross communication." After this feedback, we encouraged use of a class Slack channel for offline discussions and made more use of the Zoom chat.
- "As it is now, the structure of the breakout rooms doesn’t allot enough time (solution: maybe either extend breakout room time or remove them and continue the large discussion)." In the second half of the semester, we increased the time for most breakout room sessions.
- "Timing on reading breakout room results - there’s a lot of repetition: we could... have it structured more of a discussion on the differences and the highlights.” Near the beginning of the semester, we had each breakout room share main points from their discussion. Since this turned out to be repetitive, we replaced it with more time in breakout rooms and the occasional surfacing of highlights by instructors.

Some students would like to see more contemporary views from the field to frame discussions as well as more challenging discussion questions (9 of 21 students):

- "Read more scholarly work on the topic and/or formal ethics reading, just to see 1) what’s going on right now 2) have a grounding element.” The consultant added: "Students clarified that they don’t mean solely scholarly readings (which may not be accessible to all), but that it could include different kinds of contemporary media."
- "We believe maybe discussing more solutions to ethical concerns can be helpful to include."

7.3 Student Course Evaluations

17 of 22 (77%) students responded to the official course evaluation. They rated the overall quality of the course a 4.94 +/- 0.24 (out of 5). The department average is 4.41, the college average is 4.30, and the university average is 4.29. The Freshman ethics seminar described in Section 3 had an average of 4.63 over the 3 years it was offered.

In addition to the quantitative results, the student feedback at the end of the semester was overwhelmingly positive. Below we have grouped student comments thematically.

Instructor Interest in Student Growth.

- "The Professors were very interested in learning about every student, their ideas, etc. It was a very welcoming and personal feeling compared to many SCS Classes at CMU."
- "I had an immense amount of respect for [the instructors'] teaching style which emphasizes student learning over student assessment. This greatly increases the enjoyment of class in a world where we already have enough things to be anxious about."
- "[Instructor 1] is insanely patient, responsive, intelligent, and also truly cares about the student’s growth in this class."

Discussion Moderation.

- "It is easily my favorite course I’ve ever taken at [our university]... I’ve never met professors who were able to engage students in such honest dialogue without making the discussion feel draining to be a part of. Both of [the instructors] really cared about the student’s growth in this course."
- "Initially, I felt like I didn’t have anything to contribute and I appreciate how [the instructors] moderated conversation. It felt like that... Conversation."
- "[Instructor 1] is early in her teaching career, which made her able to easily relate to student concerns and experiences."
She complimented [Instructor 2] extremely well and was a leader in class discussions and organization.

- “I learned a lot from the instructors’ anecdotes.”

Interaction With Peers.
- “This is the first class where I had any meaningful exchange with other students at CMU.”
- “I learned a lot from… the class set up that allowed for conversation across classmates of interdisciplinary backgrounds.”
- “The class had a large variety of degree types in the class increasing the perspectives.”
- “My collaboration skills have improved ever since taking this class due to the group activities.”

In-Class Structure.
- “I also appreciated the thought and effort put into details that allowed the class to run so smoothly, like pre-created Google Docs for teamwork.”
- “I enjoyed the intentionality that was put into the break out room sessions… This class was simply fantastic.”
- “Class content can be messy with no slides and a lot of discussions but it was the complete opposite: it was super organized.”
- “FUN FUN FUN activities!”

Assignments.
- “I really appreciate [Instructor 1] and all her effort of assignment crafting and details of objectives for each assignment she created.”
- “I really enjoyed the journal assignments requiring us to probe our ethical opinions.”
- “Super in-depth analyses”
- “The course is really wonderful - it helped me get the ‘vernacular’ I want to talk about ethics and it helped me with other stuffs as well - how to structure my thoughts in writing, how to design engagements.”
- “I also appreciate the assignments were always graded with comments.”

Module Project.
- “The ethics module… is the coolest project I’ve ever worked on. I really do love this class.”
- “The awesome Ethics module project… really pushed me at least out of my comfort zone.”

Miscellaneous.
- “This class teaches critical skills for anyone who wants to work in the technological world.”
- “The reason that I took the class was because the syllabus mentioned a diversity of content (e.g. very specifically the Black Mirror).”

7.4 Instructor Response
The response from instructors involved in the project was equally positive. One instructor emailed us to say that “[the student team] did a great job!” Another instructor said: “I am teaching computer vision next semester and would love to incorporate the ideas when I teach the course.”
Class Size. Our class had 22 students, which was quite ideal. With 22 students, each could feasibly contribute to each group discussion. Each instructor graded assignments for half of the students, and we swapped halves mid-way through the course. This allowed us to follow student trajectories across multiple assignments and personalize the feedback. With a larger class size, getting to know the students and their work at this level might be challenging.

Module Project Collaborations. A few aspects of the module project made it effective. First, the students were working on materials to be used in a real course with that course’s instructor. Had the students been tasked with designing a module for a course in the abstract, the module would not have the same grounding or impact. By collaborating with the instructor, the course could be optimized for its actual use case. Second, the students were motivated to make these materials great and were given the time and skills to do so. In other scenarios, ethics module development could be an additional task for graduate students or the instructors themselves. Putting in multiple weeks of work on module design could be infeasible for these individuals. Even if they had the time, they might not have the pedagogical skills to do so. The course aimed to equip students with both the time and the skills to make these modules excellent.

Peer Interaction. The variety of perspectives allowed students to learn more from their peers. In addition to gender and racial diversity, our students spanned many stages of life and academia across many fields. For example, our class included both active military members and people who refused to take military research funding. Within our class environment, these perspectives combined in positive, constructive ways. These peer interactions could be further developed by fostering offline discussions on a class forum and by increasing peer review of assignments.

8.2 Improvements for Future Iterations

With this being the first offering of a newly-designed course, there are many potential areas for improvement. Below we describe a number of aspects we might change in future offerings.

Distributing Workload More Evenly. The amount of work assigned varied by week. We adjusted these throughout the semester by giving class-wide extensions during particularly busy weeks. In the next iteration, we’ll have a better idea of the workload and use this to inform the semester schedule in advance.

Spreading Pedagogy Learning Throughout the Semester. Some students found the 2 weeks focused on pedagogy in the middle of the course to be somewhat isolated from the ethics content in other parts of the course. Instead, we could weave pedagogy into other ethics topics and assignments to unify the course more effectively.

Supplementing Class Discussion With Other Forums. We frequently found that full-group discussions ran out of time to cover all the content and we would have to move on with students’ hands still raised, contributions left unsaid. At the beginning of the semester, we dissuaded students from using the Zoom chat for side discussions. However, as time went on, we saw the value of the chat feature in allowing students to make quick remarks or reference an outside source. In future iterations, especially if the course is taught in-person, we would encourage the continuation of discussions outside of class time on a forum such as Slack or Piazza. We did create a Slack channel for the course, but it was infrequently used.

Starting Module Project Earlier in the Semester. Many students were eager to start on the module project earlier than we had prepared. We announced the list of participating courses in week 6, students formed teams at the beginning of week 8, and we made team-instructor introductions by the end of week 8. In the future, we would provide the list of courses earlier in the semester and shorten the time provided to form teams and submit course preferences. By moving introductions earlier, each team might be able to meet with their instructor more and increase iterations on the module project.

Relax Participation Requirements. We began the semester with a quite rigid participation rubric. For example, an ‘exemplary’ rating for participation frequency involved contributing more than once each class session. In practice, it was infeasible for all 22 students to contribute twice to every group discussion. Nevertheless, we largely found that students were participating at a level that we found exemplary. We adjusted the final participation scores to reflect this and in the future would make a similar update to the initial rubric.

8.3 Potential for Adoption

This course structure aims to address two large needs in computing ethics. First, CS students would benefit from the depth of a full-length ethics course that feels intimately tied to their other coursework and eventual career. Second, there is a growing need for incorporating ethics into other CS courses to reinforce how ethics pervades all areas of technology. Our course was successful because it addressed both of these needs. It provided an engaging course experience for our students, grounded in an impactful final project. The project alone will have an impact on the 11 CS courses that now have carefully tailored ethics integration. The ethics integration will serve thousands of students across the 11 courses.

We anticipate even greater impact if this model can be replicated across the SIGCSE community. If a department already has an ethics course, adoption could be relatively straightforward. First, identify CS instructors interested in integrating ethics in their courses; second, incorporate pedagogy learning and practice in the course; third, add a final project which pairs student teams with the identified instructors. To scale the model to larger class sizes, in-class activities would need tailoring. For the module project, the size of student teams working with a single instructor could increase.

Applying pedagogy design as a learning tool has potential benefits beyond ethical analysis of robotics and CS. In a College of Science, technical advances like CRISPR provide similar affordances for the design of pedagogical interventions regarding their social ramifications. In a College of Arts and Social Sciences, technical coursework in Public Policy, Sustainability, Philosophy, and History of Science all yield similar opportunities for the analysis of technology that affects public good in complex ways. We believe that a cross-curricular approach to active learning can yield equally effective results in engaging students in faculty discussions, course design, and lifting student analyses from a theoretical island to a pragmatic, topic-oriented treatment of ethics and social ramifications from all manner of advancements.
ACKNOWLEDGEMENTS

We thank the 22 Ethics and Robotics students the 11 collaborating instructors for participating in the first-ever iteration of this course. We thank Jacqui Stimson from the Eberly Center for the early course feedback and Brian Railing for feedback on a draft of this paper. The instructors were supported by the NSF Graduate Research Fellowship Program and the K&L Gates Endowment for Ethics and Computational Technologies.

REFERENCES


