Science fiction in particular offers students a way to cultivate their capacity for moral imagination.

BY EMANUELLE BURTON, JUDY GOLDSMITH, AND NICHOLAS MATTEI

How to Teach Computer Ethics through Science Fiction

COMPUTER SCIENCE FACULTY have a responsibility to teach students to recognize both the larger ethical issues and particular responsibilities that are part and parcel of their work as technologists. This is, however, a kind of teaching for which most of us have not been trained, and that faculty and students approach with some trepidation. In this article, we explore the use of science fiction as a tool to enable those teaching artificial intelligence to engage students and practitioners about the scope and implications of current and future work in computer science. We have spent several years developing a creative approach to teaching computer

ethics, through a course we call "Science Fiction and Computer Ethics."[7–9,18,28] The course has been taught five times at the University of Kentucky and two times at the University of Illinois at Chicago and has been successful with students, as evidenced by increasing and full enrollments; high teaching-evaluation numbers; positive anonymous comments from students; nominations and awards for good teaching; and invitations to speak about the course on conference panels and in talks.

Computer science, as a field, already recognizes that some ethics education is essential; the Accreditation Board for Engineering and Technology (http://www.abet.org/), one of the largest U.S.-based accreditors of engineering and technology programs, requires instruction on professional ethics. Indeed, some in computer science have gone so far as to require students in undergraduate courses to perform ethics consultations for local industry.[24] However, educating students to engage ethical challenges is often left to the cross-disciplinary portions of university curricula, especially in the U.S.[12] We, as well as others, argue that spending time focused on how these issues apply to both our own research and our students' future work is important and necessary within computer science.[30,36]

In fields with a strong practical component and established body of

» **key insights**

■ It is important to teach students to understand the difference between normative ethics—or what is the right answer or normal mode of thought—and descriptive ethics—or using the language of ethical theory to understand and describe a situation.

■ Using fiction to teach ethics allows students to safely discuss and reason about difficult and emotionally charged issues without making the discussion personal.

■ A good technology ethics course teaches students how to think, not what to think, about their role in the development and deployment of technology, as no one can foresee the problems that will be faced in a future career.

knowledge (such as medicine, engineering, and the undergraduate levels of many sciences) there is a temptation to teach solely through the transmitting of facts, rather than encouraging discussion and dissent.[11] This approach, which many undergraduates have seen, can condition them to interpret what they learn in terms of an authority-based view of "truth" that in turn leaves them unequipped to reason about situations involving no single correct answer or think cogently about ethical trade-offs.[23,34] We want to teach our students to move past this authority-based view and find the best, most efficient solution to technical problems; we argue that the same skills must be developed to engage with ethical challenges that arise from the substance of their work as well.

Many courses focused on both research and ethical considerations taught through fiction have been offered worldwide, including at Humboldt University in Berlin, Germany,[a] and a version focused on legal issues at Stanford University.[b,1–3,19,20] Courses in other fields use literature (including science fiction) in non-majors courses as both a "hook" and a platform for exploring core ethical issues.[3,13] Scholars in other humanistic disciplines (such as history and philosophy) have also argued that literature is an invaluable teaching tool for ethics and other topics; see Copp,[16] Davis,[17] and Pease.[35] The common observation is that a fiction-based approach makes it much easier to push beyond a review of best practices toward a more in-depth education in ethical reasoning; Nevala-Lee[33] said: "[...] fiction often removes the intellectual and emotional resistance some students might at first feel towards the subject of ethics."

Ethics and Values in Computer Science

Researchers in computing, as in all professions, hold multiple and often conflicting sets of values, as well as different ways to approach living up to one's values. It is important to be clear that the purpose in teaching ethics is not to unify the field around a particular value

> A key part of ethics education is helping students see beyond their own reflexive assumptions about what is true or right.

system but to encourage reflection and precision of thought among all computer professionals. Teaching this way will, we hope, lead to an openness and exchange of ideas about both core values and best practices.

The very idea of a universally applicable ethical doctrine has serious problems. As anthropologist Melville Herskovits wrote in protest of the United Nation's Universal Declaration of Human Rights, the declaration—although intended "to be applicable to all human beings ... [is] conceived only in terms of the values prevalent in countries of Western Europe and America."[15] That is, any attempt to codify a universal definition of the "right" way to be human cannot, by definition, take account of the particular social and ethical context of individual cultures. Cultures that have historically been most oppressed would thus be the most likely to be ignored or delegitimized by a "universal" declaration.

Although the precise status and possibilities of human rights discourse continue to be debated, scholars in both ethics and anthropology agree there is no way to formulate universal precepts of this kind that do not, on some level, reinforce the very kinds of social inequality they are designed to combat. The idea that a single code of laws or duties would solve all problems, and that our responsibility as teachers is to transmit those laws to students, is appealing but ultimately false. As Callahan[10] says, "No teacher of ethics can assume that he or she has such a solid grasp on the nature of morality as to pretend to know what finally counts as good moral conduct. No society can assume it has any better grasp of what so counts as to empower teachers to propagate it in colleges and universities. The premise of higher education is that students are at an age where they have to begin coming to their own conclusions and shaping their own view of the world. It is the time and place to teach them intellectual independence, and instill in them a spirit of critical inquiry."[10]

The responsibility of an ethics instructor is to train students to engage in understanding and reasoning. The students are thus prepared to navigate situations that offer no clean solutions and engage other computer science practitioners in discussion about what and

how to choose. Callahan[10] also endorses the idea of helping " ... students develop a means and a process for achieving their own moral judgments" when confronted with challenging situations.

It is essential that open ethical debates between well-informed practitioners take place. Computer science does not take place in a vacuum; to an ever-increasing degree, the IT systems and platforms, from search engines to smartphones, that are built by computer scientists and engineers are creating and redefining the social, political, and individual contexts in which human beings understand themselves.[21] Whatever principles and norms are adopted by computer scientists, and reinforced through the design and deployment of their systems, will have profound ethical and societal implications. Teachers and leaders in the field have a responsibility to drive the discussion about the effects of their own work and the work of their students. Indeed, Boyer[6] argued that academics have a responsibility to engage students and the public with their research.

We have started to see this engagement through a number of initiatives in the computer science community, including the International Joint Conference on Artificial Intelligence 2015 letter on autonomous weapons research[c] and the 2017 follow-on letter signed by CEOs of tech companies around the world;[d] ACM statement on algorithmic accountability;[e] development of the IEEE standard for algorithmic bias considerations;[f] and new conferences and research groups focused on fairness, accountability, and transparency,[g] as well as conferences focusing on the effect of artificial intelligence on society.[h] These debates are important for shaping the direction of the field, even though they rarely result in consensus. The utility of the debates is not that they result in standardized practices but rather that individual practitioners be-

come more thoughtful and better informed about their work and its long-term effects.

As in other areas of thought, this viewpoint diversity is a strength when it can be harnessed toward a productive exchange of ideas and perspectives. An example of such an exchange is the ongoing debate within the artificial intelligence research community about the appropriate value systems on which to build artificial intelligence systems. The goal of teaching ethics is to foster the debates and equip practitioners to participate productively. It does so, not by imposing a value system on students, but by informing them about the range of ethical descriptive and evaluative tools available to them.

At the same time, educators should make students and professionals aware of the social ramifications of their work, that research, development, and implementation can be carried out in a variety of ways and for a variety of ends. Computer science educators should dedicate significant time to ethics education, helping enable students to make informed, thoughtful, ethical choices about technology and its applications and implications.

What is ethics? Ethics can be understood as the task of answering "What should I do?" which is never a simple matter. Ethics includes both thought and practice, an organized and intentional reflection on morality and the effort to live in ways that are good, just, and/or right. Although many people use the words morality and ethics interchangeably, many ethicists understand them to be different. One common way of drawing the distinction is to define "morality" as a set of values or a worldview and "ethics" as the practice of reflecting on those values and their foundations and applications.[4,22]

There are many different, often conflicting, ways to understand how to be moral. The clashes are sometimes between people who share the same fundamental premises and method of inquiry into how to be moral but disagree about conclusions. Other times, the clashes are between people whose basic ideas of how to answer the question of how to be moral conflict with one another. Most approaches to morality can be understood in terms of the three major traditions of ethical thought—deontological

ethics, virtue ethics, and utilitarianism—with each growing out of different core questions and ways of seeing the world.

Ethics is typically understood to be normative; that is, it is aimed at establishing norms of thought, values, or conduct. This assumption is especially prevalent in professional ethics courses that are typically used as a means to steer students' future behavior toward a set of professionally agreed-upon values (such as professionalism and honesty).[26] But ethics is also a tool for description, furnishing decision makers with a critical framework that enables them to understand what is happening in a given situation and what is at stake in any action they might take. The boundary between normative and descriptive functions is sometimes fuzzy; for example, it is often the case that different details of a situation will appear salient depending on which ethical approach one adopts. This malleability of relevant details can make ethics itself seem murky or imprecise. However, teaching students to appreciate this difference, understand the modes of reasoning that they or others might employ in making an ethical decision, and move between these reasoning structures themselves is the goal of a good ethics course.

Educating students in the descriptive functions of ethics is as important as communicating to them the professional norms of computer science. Computer science is a field in which everyday practice and problem solving takes place in a context that could barely be imagined the decade before. Educators cannot predict the ethical quandaries their students will face. With an education in ethical description, the students will be better able to engage in subtle and substantive ethical reasoning when new and challenging problems confront them.

Practical challenges of teaching ethics. Ethics education is a notable challenge for two reasons. First, in the absence of any ideal universal ethics program, students must be taught how to approach problems as distinct from being led to particular pre-ordained conclusions that might narrow their vision and exclude important elements of a given problem. Second is how to achieve this goal while overcoming the biases students often bring to the classroom.

c http://futureoflife.org/AI/open_letter_autonomous_weapons

d https://futureoflife.org/autonomous-weapons-open-letter-2017

e https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf

f http://sites.ieee.org/sagroups-7003/

g http://www.fatml.org/

h http://www.aies-conference.com/

Teaching how, not what, to think. It is important to consider what it means for us as educators to say we want to inform our students how to think instead of what to think. It is tempting to assume we can formulate a set of rules in natural language, refine them until we agree on them, and then proceed as if these rules can be applied without further reflection. However, the real world is messy, and rules that may seem reliable under one set of conditions can falter under others. Furthermore, language is not always identical to the world it is intended to describe. Different people often describe the same experience in different ways or understand the same phrase to refer to different phenomena. At minimum, such universal rules would require everyone who relied on them to engage in ongoing reflection about their own understanding and application of the rules to the world.

Both the appeal of this rule-based approach and its limits can be seen with respect to the question of programming robots against concrete actions (such as law enforcement robots never to shoot a human). While this operating principle seems at first like a straightforward way to ensure the preservation of human life, it is not difficult to imagine scenarios in which shooting a person, perhaps even lethally, can be expected to save the lives of others. But how should a robot calculate the risks and values at stake in such a scenario? What sorts of input should it use when ascertaining if it should shoot a human? What sorts of input should it ignore? And what are the social costs or benefits of using robots that will shoot a human under certain circumstances? Another example is the ongoing recent discussion about the classic trolley problem in light of the rapid advance of self-driving cars.[5]

Ethics education often requires a different kind of education from understanding and applying an established body of knowledge. In computer science, knowledge usually leads to action; if one chooses to create or program a system to solve a problem, and know how to do it, there is little reason not to solve the problem in the most direct and efficient way possible. Ethical understanding, however, requires an additional layer of commitment. One must overcome both the temptation to adopt an easier or more self-serving course and the distractions that might prevent someone recognizing an ethical problem in the first place. It is not difficult to imagine a student who can get 100% on an exam correctly identify terms and offer cogent and sensible solutions to hypothetical scenarios but then enter the work world and act in ways that ignore ethical consequences or even violate their[i] own values. This student might not stop to think they have acted wrongly; or such a student might notice but consider practical or professional pressures to be more important. An ethics course is successful only if it goes beyond equipping such a student with information and knowledge they can use but also prepares students to scrutinize their use of that knowledge, even when doing so is not convenient or comfortable.

In order to avoid causing great harm in the world, any field that involves practice requires not only technical proficiency of its practitioners but also ethical proficiency, as manifest not only in a command of the relevant knowledge but also the inclination and ability to let that knowledge take precedence over laziness or self-interest. That is, a successful professional ethics education does not just offer resources to indicate how problems can be identified and addressed; it also trains students to avail themselves of those resources, even when it is possible and easier not to. Teaching such skills and habits to students is a challenging task that cannot be successfully realized through cross-disciplinary requirements alone but must be integrated into their computer science education.[38] The number of recent professional society calls to deal with algorithmic bias and the disparate effects of information technology systems makes clear that computer science departments must engage directly with this responsibility.

Negotiating student biases. A key part of ethics education is helping students see beyond their own reflexive assumptions about what is true or right. Our classroom experience shows us that introducing students to three of the major schools of ethical theory—deontology, virtue ethics, and utilitarianism—helps broaden students' ability to recognize and reflect on those assumptions. While all three schools have proponents among philosophers, theologians, and other scholars who work in ethics, broader cultural discourse about ethics tends to adopt a utilitarian approach, often without being aware that there are other ways to frame ethical inquiry. This larger cultural reliance on utilitarianism may help explain why it consistently seems, to students, to be the most crisply defined and "usable" of the ethical theories. But there are significant critical shortcomings to this popular version of utilitarianism. The concept of "the greatest good" is notoriously ill-defined in utilitarianism, and while trained philosophers struggle to identify or formulate a suitable definition, the gap typically goes unnoticed in less-philosophical circles, enabling agents to plug in their own definition of "the good" without submitting it to scrutiny. Furthermore, it is very easy to try to apply the basic formula of utilitarianism—the greatest good for the greatest possible number—to a decision without thorough consideration of all those who will be affected. This move enables agents to declare they have pursued a morally reasoned course when, in fact, they have calculated the benefits only to themselves and those in their immediate sphere. This difficulty attaining a sufficiently broad understanding of the effects of actions, and thus in appropriately computing the utility of those actions, can curtail the ability to have a substantive ethical discussion, even insofar as everyone assents to utilitarianism.

In our experience teaching ethics courses under the auspices of computer science departments, we find that students are often drawn first to utilitarianism, perhaps because it seems more computational than the alternatives. One of the most important aspects of the course is to broaden their experience to help them see past the non-rigorous version of utilitarianism to which they were previously exposed. The aim is not to demonstrate the superiority of one approach over another but rather to help them understand the uses and limits of each approach. This limitation can be exemplified by the question of whether to replace factory

---

i   In the spirit of inclusion, and in deference to shifting usage norms, we use "they" as a singular, non-gendered pronoun, as well as for the more traditional third-person plural; we trust context disambiguates.

workers with robots. They may focus on the happiness of the factory owners, shareholders, and those who will pay less for manufactured goods, without considering the utility of the human factory workers and those whose jobs depend on factory workers having money to spend; or even the more high-level question about whether or not it is reasonable to consider human beings and machines as interchangeable. Indeed, the three approaches can be complementary, or even mutually informative; for example, recent theorists have argued that virtue ethics is best seen as part of successful deontology.[27]

Why fiction to teach ethics? Stories—literature, plays, poetry, and other narrative forms—have always been a way to talk about the world as it is, telling us what it is like and what effect our choices will have. Whether they are transmitted in print or through other media, stories play a potent role in shaping the thoughts and ideas of individuals, as well as the cultural norms of the societies in which they live.

Scholars of ethics have, in the past several decades, embraced fiction as an ideal way to think about and teach ethics, because, as philosopher Martha Nussbaum[32] writes, fiction " ... frequently places us in a position that is both like and unlike the position we occupy in life; like, in that we are emotionally involved with the characters, active with them, and aware of our incompleteness; unlike, in that we are free of the sources of distortion that frequently impede our real-life deliberations." By offering the reader both immersion and distance, an ethics course based in fiction helps students perceive the degree to which ethical quandaries are tangled up in other aspects of life while furnishing a context that keeps them connected to abstract principles and questions. As such, fiction-based ethics education helps them cultivate the capacity to recognize ethically complex situations as they arise or extract an ethical dilemma from a larger context. This combination of qualities also helps students develop the moral imagination that is a key component of successful ethics education.[10] The common alternative is to provide them with a prepackaged case studies in which the particular ethical dilemma under study is cleanly identified for the student.

Science fiction is particularly well

Spring 2018 Syllabus.

Week 1. Using ethical language; in-class discussion
Read Isaac Asimov's short story "The Dead Past"

Week 2. Professional ethics
Read the ACM's and the IEEE's codes of ethics

Week 3. Utilitarianism
Read Harlan Ellison's short story "'Repent, Harlequin!' Said the Ticktockman"
Ethical description exercise #1 due in class

Week 4. Deontology
Read Elizabeth Bear's short story "Dolly"
Ethical description exercise #2 due in class

Week 5. Virtue ethics
Read E.M. Forster's short story "The Machine Stops"
Ethical description exercise #3 due in class

. . .

Week 6. Selfhood and technological mediation
Read James Patrick Kelly's short story "Itsy Bitsy Spider" and
Tom Sorell's and Heather Draper's paper "Robot Carers, Ethics, and Older People"
Ethical description exercise #4 due in class

. . .

Week 9. Privacy
Read Ken Liu's short story "Here-and-Now" and
Articles: Helen Nissenbaum's "Privacy as Contextual Integrity,"
Manan Kakkar's "A Case Against Online Privacy," and Adam D. Moore's "Privacy, Speech and Values: What We Have No Business Knowing"
. . .

Week 14. What is ethical warfare?
Read Linda Nagata's short story "Codename: Delphi," Ronald C. Arkin's essay "Ethical Robots in Warfare," Jean Elshtain's paper "The Problem of Dirty Hands," and Emerson T. Brooking's and Peter Singer's "War Goes Viral"
Ethical argument assignment #3 due in class

Last week. Professional ethics, the importance of integrity
Read E. Saxey's short story "Not Smart, Not Clever"
Ethical argument assignment #4 due in class
Reread Asimov's "The Dead Past"

---

suited to teaching computer ethics. As Alec Nevala-Lee[31] says, "Science fiction has been closely entwined with military and technological development from the very beginning. The first true science fiction pulp magazine, *Amazing Stories*, was founded by editor Hugo Gernsback expressly as a vehicle for educating its readers about future technology." Our project builds on this long-recognized insight—that science fiction is, in key respects, better able than "realistic" fiction to reflect the near future (or possible futures) in which computer professionals work. Science fiction thus permits a curricular design that hews more closely to the concerns and quandaries of computer-related fields of study and work. A successful ethics course will reframe the task of ethical engagement so students understand the

ongoing responsibility to ask ethical questions of themselves and their work; and further, that they are equipped to perceive, describe, and understand the challenges as they arise. We find that science fiction makes the key ethical questions of technology development and use more vivid and engaging and the critical resources for addressing ethical questions more intelligible.

We take science fiction in its broadest sense, as the fantastical worlds or even the futuristic technology gives us a starting platform for discussion. The category of science fiction was first described by Hugo Gernsback, for whom the prestigious Hugo Prize is named, in the editorial of the first issue of *Amazing Stories* in 1926 as, " ... I mean the Jules Verne, H.G. Wells, and Edgar Allan Poe type of story—a charming romance intermin-

gled with scientific fact and prophetic vision." Using this broad definition, almost any fiction dealing with sufficiently advanced technology is science fiction. Though the majority of the literary and philosophical establishment has not, until recently, seen science fiction as a venue for serious ethical thinking, this fact reflects longstanding biases in the field rather than the merits or possibilities of science fiction itself.

Fiction allows educators to reframe recognizable human situations and problems in terms of unfamiliar settings and technology. Hence, any fiction, and especially science fiction in the case of technology, can be an ideal medium for raising and exploring ethical concerns. By presenting a familiar problem (such as conflicts between different social groups or the invasion of privacy in unfamiliar terms and settings), a work of science fiction can mitigate a reader's tendency to defend, reflexively, their own previously held views. As Nussbaum[32] writes, "Since the story is not ours, we do not get caught up in the vulgar heat of our personal jealousies or angers or the sometimes blinding violence of our loves." In this way, science fiction creates an opportunity for students to gain fresh insight into, and even empathy for, ethical positions and people whose real-world analogues are not embraced by their values or politics.

We thus advocate science fiction for several reasons in addition to the ones outlined here. First, the use of futuristic or alien settings allows students to detach from political preconceptions and experience the dilemmas of plot and characters as something fresh. Second, it has so far proven popular and effective with students. One student wrote the following on a Spring 2017 anonymous course evaluation: "Going into this course, there were several times that I could acknowledge an ethical situation and had my own ideas as to whether it was 'right' or 'wrong,' but I couldn't necessarily articulate why. This course gave me the tools to be able to have a meaningful discussion about these topics. It was also a productive way to get out of the coding mindset, take a step back, and consider what other results might come from the technologies that we will be making. Phenomenal course, and phenomenal instructor." Finally, some

of the science fiction we chose also posits new science infrastructure and allows students to think about doing research and development outside the fairly rigid industrial and academic boxes, driven by something other than current funding paradigms. This creative thinking about practical problems, according to some philosophers[29,37] and educators,[14] is a crucial component in developing the ethical reasoning abilities of students. All these reasons, along with the distance from the material that can be created through fiction, have led to a very successful course taught more than eight times as of August 2018 and that has won us multiple teaching awards.

The Course

The aim of the course is to prepare our students to recognize ethical problems in their present and future work as technologists, focusing on methods of applied ethical reasoning (for the future), as well as on particular current problems. During class discussion and in homework assignments, they analyze both science fiction stories and brief articles, using the major ethical theories not only as evaluative tools but as a descriptive apparatus to enable them to recognize problems and consider possible solutions from multiple perspectives. As we have seen, this focus on ethical theory as a descriptive tool, combined with the use of science fiction stories as an arena for ethical description and analysis, sharpens the students' ability to perceive and describe ethical challenges and expands their capacity to address them with creativity and nuance. An abbreviated example syllabus is outlined in the figure here.

The class opens with a crash course on ethical theories and a review of the IEEE and the ACM codes of ethics. Students consider the different modes of ethical engagement invited by each code and discuss whether, and in what ways, either one is likely to affect their decision making. Although this discussion typically evinces varying opinions on the usefulness or relevance of either code, there is near-universal consensus that the codes are not, by themselves, sufficient to help an IT professional address the challenging problems that may arise. We, the instructors, stress this is a problem common to all codes

of ethics and the solution is not a more-perfect code but rather IT professionals better prepared to engage in ethical reasoning, and thus to make use of professional codes.

The course then spends several weeks on in-depth study of each of the three major ethical theories—utilitarianism, deontology, and virtue ethics—with one day for each on a critical reading assignment that introduces the theory in detail and another day analyzing and discussing a short story from within the perspective of that theory. To prepare for these discussions, students write "ethical description exercises," answering guided questions about how the story world can be understood through that week's ethical lens. Some of these stories, particularly Elizabeth Bear's "Dolly," which is used to teach deontology, and E.M. Forster's "The Machine Stops," which is used to teach virtue ethics, end up as touchstones for the course, resurfacing in student discussions about later subjects.

After helping build the students' analytic competency in ethical theory, the course moves to a consideration of major ethical concerns in IT, including surveillance, the interrelationship between news and social media, and self-driving cars. On the strength of the assigned science fiction stories, students consider both immediate practical problems and deep underlying issues that recur in IT ethics past, present, and possibly future.

Each story touches on multiple core issues, enabling the students to appreciate, and grapple with, the interconnectedness of the various challenges they will confront. Stories like James Patrick Kelly's "Itsy Bitsy Spider" and Paul Shoemaker's "Today I Am Paul," both focusing on carebots looking after aging parents with dementia, serve as the basis for a discussion of carebots in particular but also inspire broader discussions on how technological interventions can change the conditions in human relationships. Paolo Bacigalupi's "The Gambler" helps frame a discussion of new media and the attention economy, highlighting the particular hurdles this new information environment creates for minority experience and positions.

Ken Liu's "Here-and-Now" offers a potent view of the personal and social stakes of the post-privacy era, particularly in the context of the mostly un-

regulated gig economy that depends so heavily on IT innovations. And Michael Burstein's "Teleabsence" explores how technological innovations designed to address social inequality can in fact exacerbate it while raising probing questions about the powers and limits of how one might redefine oneself on the Internet. Although the reading list has changed with each iteration, these stories and others like them have formed the backbone of each version of the course.

In each such iteration, our students have emerged from the semester's reading inspired, troubled, and invigorated by the new perspectives they have gained on their future work.

The assignments in the course help develop their capacity for attention and critical thought in a manner intended to serve them well throughout their professional lives. By working descriptively with three different ethical theories, they develop a rich critical vocabulary for recognizing ethically fraught situations as they arise. The questions given to the students for a particular story are deliberately open-ended, requiring them to identify and formulate the problem from the ground up, an approach that addresses a practical gap created when they are taught using only case studies. This open-endedness also fosters a wider range of responses than a more closely tailored set of questions, thus creating a more varied class discussion.

Through the multiple writing assignments, the students not only become aware of a range of potential ethical challenges in their work in computer science but also alert to the variety of ways these problems might initially emerge. They are thus able to identify potential ethical risks in a given technology or model or in a company's and the public's use of the technology or model. They are better prepared to articulate their arguments for why a given approach is the most (or least) ethical choice and see past incomplete or specious defenses of potentially unethical projects.

Example Story Materials

Here, we include an example of the pedagogical materials we have developed to capitalize on the lively accessibility of the fiction-reading experience while also helping the students come

> **Science fiction makes the key ethical questions of technology development and use more vivid and engaging and the critical resources for addressing ethical questions more intelligible.**

to grips with the complexity of considering a problem in the context of the wider world where it takes place. These materials include both a story frame to introduce the stories to students and a pedagogy guide to help instructors. The stories we have collected for the course (and, no doubt, many others) are engaging enough to spark energetic debate about ethical questions on their own and reward sustained scrutiny along ethical lines with several layers of productive challenge beyond an initial encounter. Once the problems illustrated by the narrative are described and conceptualized, the full ethical implications and challenges can be understood by "re-embedding" the problem back into its narrative context. The students should then consider how the world of the story created the conditions for both the external problems and the internal struggles addressed by the related characters.

The story frame furnishes the students with light guidelines, preparing them to pay attention to particular issues without instructing them how to answer, or even ask, ethical questions. The story frame thus leaves room for the students to discover the questions for themselves and grapple with the challenge of identifying and naming the problems at hand. This choice not only helps preserve the excitement of discovery that comes with reading good fiction but also requires the students to undertake these tasks on their own. While their own initial attempts to frame, define, and address ethical problems are likely inadequate, their attempts to do so both individually and collectively are an essential part of the learning in an ethics course, as the real-world problems they encounter will not come with a set of pre-formulated guidelines to steer them toward the nominal best answer.

The pedagogy guide, in addition to offering generalized tips for stimulating and sustaining productive discussion about fiction and ethics, also points the instructor toward relevant themes, details, and patterns in the text. These details and patterns do not, by themselves, constitute an "answer" to any of the core ethical questions raised by the stories. As a list of facts, they are not especially helpful for students grappling with the

core ethical challenges of a given story. In the context of an ongoing discussion, the instructor can introduce these details to raise new questions or challenge provisional explanations about how the world of the story works or why characters make the choices they do. In story discussions—and, indeed, in discussions concerning the real world—students often begin the course by wanting to find tidy answers for challenging ethical problems. To counter this impulse, future instructors will find it useful to interject into the discussion details that complicate the students' explanations. In this way, discussion of the story worlds can help train students to perceive complexity in the real world.

Story frame for "Here-and-Now." The story under study here is Liu's "Here-and-Now," a short story that has sparked lively and productive discussion among students in previous versions of the course. Liu, a trained computer scientist, has written several excellent stories in recent years that directly address issues in computer ethics, material circulated to students, along with the story text itself; "Here-and-Now" is available for free at http://www.kasmamagazine.com/here-and-now.cfm

The story itself[25] begins with this sentence: "It's amazing what you can get, just by asking."

How much is information worth? That is the question Aaron, the protagonist of "Here-and-Now," is forced to confront over the course of one complicated afternoon and evening. Aaron is one of thousands (if not millions) of people using a new app called *Tilly Here-and-Now* that allows users to pose anonymous requests for "information" of any kind. The story asks deceptively simple questions as to why information matters. It also points out that some kinds of information are much more meaningful or valuable to some people than to others, asking readers to consider whether that difference should matter, and how.

The world of the story is not quite the same as ours but is similar in many ways. It appears that Centillion, the app's parent company, has achieved data-management capabilities that are not yet available in the real world, though we recognize the possibility is certainly on the horizon. Likewise, nothing exactly like the *Here-and-Now* app exists yet, but it is a plausible amalgam of many apps

Discussing ethics in the context of fiction can make it easier for instructors to adopt an open-ended approach required for a good ethics course.

and services that do exist, including *TaskRabbit*, *Pokémon Go*, and *YikYak*. Indeed, the app in the story was based on one described in a 2013 academic paper.[39] Still, we are fast approaching a world like the one in the story, and it is not difficult to imagine an app like *Here-and-Now* existing here, and now.

Study questions. Among the many essential ingredients of *Tilly Here-and-Now*'s economy are money and information but also interest on the part of users, information requesters and information gatherers alike. What are the sorts of interest that might lead someone to use the app in either of these roles? Are any of these interests in tension with the others?

Does it matter that Tilly's request function is anonymous? Why or why not?

Early in the story, Aaron decides "*Tilly Here-and-Now* made you more aware of the world around you … more connected to your community." How do the events of the story itself confirm or challenge that conclusion? Characters in the story you can use to think about this question include Aaron's acquaintance, Lucas, Aaron's parents, the unnamed people whose bounties are being fulfilled, the girls in the video Lucas has purchased, and Aaron himself.

The reward for fulfilling an information request is called a "bounty," rather than, say, a "fee," "one-time payment," or other possible term; you can probably think of others. How does that choice of word affect the way the reader think about the relationship between the information-requester and the information gatherer? Does it affect how the reader thinks about the relationship between either of these individuals and the information that is gathered? Do you think the choice of the word "bounty" has an effect on the characters in the story, as well? If so, in what way?

Moreover, who has access to the requests, and how is it controlled?

Instructor's guide. This material is available to instructors to help them guide in-class discussion of the story.

The *Tilly Here-and-Now* app exists in a world that is just different enough from our own to be provocative but similar enough to feel intuitive. Students may be tempted to jump straightaway to talking about the app itself, independent of the story. But

this particular narrative provides an exceptionally effective window onto Liu's slightly reimagined world, and the discussion will likely be more focused and productive if you dedicate at least 20 minutes to 30 minutes to discussing Aaron's experiences and reflections before moving onto the more general implications of the Tilly app.

As always, the best approach is a Socratic one, in which you guide your students toward discovering things for themselves. Here are some observations and details about the story. You can use them to ask "fishing" questions if you think your students are missing important details or to prompt them to reassess their view of the story if they have settled on a version that ignores such details.

Aaron. Aaron is interested in information about others. He likes claiming bounties and furnishing others with the information they want but also likes trying to figure out why it is that people want them. When Lucas baits him, saying, "I got something cool," Aaron cannot help asking about it.

On the other hand, Aaron hates giving up information about himself to the people he knows. He does not want his mother to know about his part in the school play and will not tell Lucas how much money he earned. In the entire story, we learn of only one instance in which Aaron willingly shares information with another character, when he teaches his mother (before the beginning of the story) about *Tilly Here-and-Now.* By the end of the story, Aaron regrets having shared that information, since his mother is now using the app "against" him to learn things about him and about his father.

The individuality of knowledge. At several points in the story, both major and minor, the reader's attention is directed to the ways information matters more to the people it touches directly. The story thus adds a new layer to frequently expressed concerns about privacy, focusing on the damage done to the character(s) whose information is known or made available. As the story explains, the person who knows can be just as affected or damaged by that knowledge as the subjects about whom it is known.

Lucas is happy with his video of two girls kissing (which strikes Aaron as invasive of the girls' privacy) but "Would have been even better if they're people I know," as Lucas says. "Next time I'm going to raise the bounty and limit the range more. It's amazing what you can get, just by asking."[25] Lucas anticipates that knowing the girls involved would make the video more satisfying. The invasion of private space is part of the pleasure.

Aaron and Lucas respond very differently to the license-plate request, not only because Aaron recognizes the plate number but because he has something personal at stake in the fulfillment of the request, and in its asking. Previously, the reader has seen Aaron wonder why he is being asked to fulfill this or that request, but never whether he should. Only when the request touches him personally does he realize the damage that might be done if it is fulfilled.

Aaron himself is later undermined (in a small way) by another *Tilly Here-and-Now* user, fulfilling another of his mother's requests, when she discovers he has been cast in the school play. On the surface, this plot point lines up primarily with more typical concerns about privacy. Aaron, who had hoped to conceal the information about his being cast in the play, is the one who has been injured but, insofar as Aaron trusts his mother less, she is also damaged.

Information control and performance. The story repeatedly touches on the theme of people pretending to be who they are not, as signaled at the opening of the story, when the reader learns Aaron has been cast in a play. A play is a performance, but the "deception" is a matter of mutual consent; the audience knows it is watching actors, and in this sense the play does not represent a miscarriage of knowledge.

This non-deceptive deception differs from the way Aaron's parents talk to each other over dinner toward the end of the story. Aaron knows by then that his mother suspects his father of cheating and he halfway suspects his father as well, but they treat each other normally, as if nothing is wrong. "He couldn't hear anything different in their tones. His mother acted like she had never asked the question. His father acted like he had nothing to hide."[25] Aaron's mother, and possibly his father as well, perform with an intent to mislead. But whom are they misleading—Aaron, each other, or both? And when did the deception begin?

It is also worth raising the question of whether, and how, Aaron's own actions qualify as misrepresentations, as in his desire to keep his role in the play a secret, and his own Tilly request, which is designed to distract Lucas from fulfilling his mother's request.

Additional topics. Liu's "Here-and-Now" also raises issues of access control and information integration, or combining different possibly innocuous sources to complete more complex, thorough, and possibly invasive records. At one point, the reader's perception shifts when Aaron recognizes his father's license plate. How do the different ethical theories frame the possibility of deanonymization in the story, either deliberately or accidental? Discussing deanonymization can lead to further discussion of hacking and Wikileaks, trust and distrust in data scrubbing, as well as other directions.

Ethical description writing assignment. The purpose of this assignment is description. Addressing the points cited in the following paragraphs, describe Liu's story in terms of one of the three major theories of ethics. (You will receive separate instructions telling you which theory to use.) Be sure to title your assignment "Here-and-Now: [name of ethical theory]."

Assigned theory. Using the concepts and worldview of your assigned theory, give a two-to-four sentence summary of the central ethical problem(s) in the story.

Ethical problems. What is at stake in the ethical problem(s) so described? That is, what possible goods could be gained or lost or what kinds of harm could occur or be prevented? Using the language of your theory, explain why these costs or benefits are significant.

Characters. What character(s) is/are in a position to take meaningful action with respect to the problem? What about their character or circumstances positions them to take such action?

Course of action. Choose one such character from your answer. Using the language and concepts of your assigned theory, describe the course of action this character takes in the story. Are there other possible courses of action the story suggests the character might have

taken? Describe them, again using the language of your assigned theory. According to that theory, what might be a better course of action, and why?

Argument. What argument do you think the ending of the story intends to make? You are still describing, rather than arguing. Use the language of your assigned theory to describe Liu's argument.

Students will bring these assignments to class on the day they are due. They are welcome to make notes on them, over the course of discussion, for their own edification and turn them in to the professor at the end of class.

Conclusion

Teaching ethics to computer science students is a pressing responsibility for computer science faculty but also a challenge. Using fiction as the basis for an ethics course offers several advantages beyond its immediate appeal to many students and some faculty. First, fiction offers students a way to engage with ethical questions that helps them cultivate their capacity for moral imagination; science fiction in particular can make the ethical stakes of blue-sky projects vivid, pressing, and immediate. Second, stories offer students the chance to develop their writing and verbal skills in ethical description. And finally, discussing ethics in the context of fiction can make it easier for instructors to adopt an open-ended approach required for a good ethics course. A course built around fiction enables instructors to incorporate the best and most useful aspects of a humanistic approach to ethics education while remaining close to the central technological concerns within computer science.

Acknowledgments

Ⓒ

References

1. Bates, R., Goldsmith, J., Berne, R., Summet, V., and Veilleux, N. Science fiction in computer science education. In *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education* (Raleigh, NC, Feb. 29–Mar. 3). ACM Press, New York, 2012, 161–162.
2. Bates, R., Goldsmith, J., Summet, V., and Veilleux, N. Using science fiction in CS courses. In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education* (Atlanta, GA, Mar. 5–8), ACM Press, New York, 2014, 736–737.
3. Bates, R.A. AI & SciFi: Teaching writing, history, technology, literature and ethics. In *Proceedings of the ASEE Annual Conference & Exposition* (Vancouver, B.C., Canada, June 26–29). American Society for Engineering Education, Washington D.C., 2011, 1–12.
4. Beard, M. Ethics, morality, law: What's the difference? The Ethics Centre, Sydney, NSW, Australia, Aug. 27, 2012; http://www.ethics.org.au/on-ethics/blog/september-2016/ethics-morality-law-whats-the-difference
5. Bonnefon, J.-F., Shariff, A., and Rahwan, I. The social dilemma of autonomous vehicles. *Science 352*, 6293 (June 24, 2016), 1573–1576.
6. Boyer, E.L. *Scholarship Reconsidered: Priorities of the Professoriate.* Jossey-Bass, San Francisco, CA, 1997.
7. Burton, E., Goldsmith, J., Koenig, S., Kuipers, B., Mattei, N., and Walsh, T. Ethical considerations in artificial intelligence courses. *AI Magazine 38*, 2 (Summer 2017), 22–35.
8. Burton, E., Goldsmith, J., and Mattei, N. Teaching AI ethics using science fiction. In *Proceedings of the First International Workshop on AI, Ethics and Society at the 29th AAAI Conference on Artificial Intelligence* (Austin, TX, Jan. 25–30). Association for the Advancement of Artificial Intelligence, Palo Alto, CA, 2015, 33–37.
9. Burton, E., Goldsmith, J., and Mattei, N. Using 'The Machine Stops' for teaching ethics in artificial intelligence and computer science. In *Proceedings of the Second International Workshop on AI, Ethics and Society at the 30th AAAI Conference on Artificial Intelligence* (Phoenix, AZ). Association for the Advancement of Artificial Intelligence, Palo Alto, CA, 2016, 80–88.
10. Callahan, D. Goals in the teaching of ethics. Chapter 2 in *Ethics Teaching in Higher Education.* Plenum Press, New York, 1980, 61–80.
11. Colbeck, C.L. Merging in a seamless blend: How faculty integrate teaching and research. *The Journal of Higher Education 69*, 6 (Nov./Dec. 1998), 647–671.
12. Davis, B.G. *Tools for Teaching, Second Edition.* Jossey-Bass, San Francisco, CA, 2009.
13. Dils, L.S. Science fiction and the future. Curriculum Unit, Yale-New Haven Teachers Institute, New Haven, CT, 1987; http://teachersinstitute.yale.edu/curriculum/units/1987/2/87.02.04.x.html
14. Edmiston, B. Ethical imagination: Choosing an ethical self in drama. Chapter 3 in *Imagining to Learn: Inquiry, Ethics, and Integration Through Drama.* B. Edmiston and J.D. Wilhelm, Eds. Heinemann Drama, Portsmouth, NH, 1998.
15. Executive Board of the American Anthropological Association. Statement on Human Rights. *American Anthropologist 49*, 4 (Oct.-Dec. 1947), 539–543.
16. Garcia Iommi, L. Let's watch a movie!: Using film and film theory to teach theories of international politics from a critical perspective. In *Proceedings of the American Political Science Association 2011 Annual Meeting* (Seattle, WA, Sept. 1–4). American Political Science Association, Washington, D.C., 2011; https://ssrn.com/abstract=1903282
17. Goering, S. Using children's literature as a spark for philosophical discussion: Stories that deal with death. Chapter 15 in *Ethics and Children's Literature*, C. Mills, Ed. Ashgate, Farnham, U.K., 2014, 233–247.
18. Goldsmith, J. and Burton, E. Why teaching ethics to AI practitioners is important. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence* (San Francisco, CA, Feb. 4–9). Association for the Advancement of Artificial Intelligence, Palo Alto, CA, 2017, 4836–4840.
19. Goldsmith, J. and Mattei, N. Science fiction as an introduction to AI research. In *Proceedings of the Second AAAI Symposium on Educational Advances in Artificial Intelligence* (San Francisco, CA, Aug. 7–11). Association for the Advancement of Artificial Intelligence, Palo Alto, CA, 2011, 1717–1722.
20. Goldsmith, J. and Mattei, N. Fiction as an introduction to computer science research. *ACM Transactions on Computer Science Education 14*, 1 (Mar. 2014), 1–14.
21. Greenfield, A. *Radical Technologies: The Design of Everyday Life.* Verso, London, U.K., and New York, 2017.
22. Haidt, J. *The Righteous Mind: Why Good People are Divided by Politics and Religion.* Vintage, New York, 2012.
23. Haworth, J.G. and Conrad, C.F. Curricular transformations: Traditional and emerging voices in the academy. In *Revisioning Curriculum in Higher Education*, J.G. Haworth and C.F. Conrad, Eds. Simon & Schuster Custom Publishing, New York, 1995, 191–202.
24. Huff, C. and Furchert, A. Toward a pedagogy of ethical practice. *Commun. ACM 57*, 7 (July 2014), 25–27.
25. Liu, K. Here-and-Now. *Kasma Online* (Nov. 1, 2013); http://www.kasmamagazine.com/here-and-now.cfm
26. Martin, C.D. and Weltz, E.Y. From awareness to action: Integrating ethics and social responsibility into the computer science curriculum. *ACM SIGCAS Computers and Society 29*, 2 (June 1999), 6–14.
27. McNaughton, D. and Rawling, P. Deontology. Chapter 15 in *The Oxford Handbook of Ethical Theory*, D. Copp, Ed. Oxford University Press, Oxford, U.K., 2006.
28. Mihail, R., Rubin, B., and Goldsmith, J. Online discussions: Improving education in CS? In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education* (Atlanta, GA, Mar. 5–8). ACM Press, New York, 2014.
29. Murdoch, I. The idea of perfection. Chapter 1 in *The Sovereignty of Good*, Routledge & Kegan Paul, Abingdon, U.K., and New York, 1970.
30. Narayanan, A. and Vallor, S. Why software engineering courses should include ethics coverage. *Commun. ACM 57*, 3 (Mar. 2014), 23–25.
31. Nevala-Lee, A. Deadline: John W. Campbell in World War II. *WorldCon 2016 Academic Track* (Kansas City, MO, Aug. 18, 2016) and personal communication.
32. Nussbaum, M. *Love's Knowledge: Essays on Philosophy and Literature.* Oxford University Press, Oxford, U.K., 1990.
33. Pease, A. Teaching ethics with science fiction: A case study syllabus. *Teaching Ethics: The Journal of the Society for Ethics Across the Curriculum 9*, 2 (Spring 2009), 75–82.
34. Perry, W.G. Cognitive and ethical growth: The making of meaning. Chapter 3 in *The Modern American College*, A.W. Chickering and Associates, Eds. Jossey Bass, San Francisco, CA, 1980, 76–109.
35. Rodwell, G. *Whose History?: Engaging History Students Through Historical Fiction.* University of Adelaide Press, Adelaide, Australia, 2013.
36. Rogaway, P. *The Moral Character of Cryptographic Work.* Cryptology ePrint Archive, Report 2015/1162, 2015; http://eprint.iacr.org/
37. Smith, B. Analogy in moral deliberation: the role of imagination and theory in ethics. *Journal of Medical Ethics 28*, 3 (Aug. 2002), 244–248.
38. Spradling, C.L. *A Study of Social and Professional Ethics in Undergraduate Computer Science Programs: Faculty Perspectives.* Ph.D. thesis, University of Nebraska-Lincoln, 2007; https://digitalcommons.unl.edu/dissertations/AAI3255458/
39. Von Der Weth, C. and Hauswirth, M. Finding information through integrated ad hoc socializing in the virtual and physical world. In *Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies* (Atlanta, GA, Nov. 17–20). IEEE Press, Washington, D.C., 2013, 37–44.

Emanuelle Burton (enburton@uic.edu) is a faculty member in the Computer Science Department at the University of Illinois at Chicago, Chicago, IL, USA.

Judy Goldsmith (goldsmit@cs.uky.edu) is a professor of computer science in the Computer Science Department at the University of Kentucky, Lexington, KY, USA.

Nicholas Mattei (n.mattei@ibm.com) is a research staff member with the IBM Research AI-Reasoning Lab at the T.J. Watson Research Center, Yorktown Heights, NY, USA.

Watch the authors discuss their work in this exclusive *Communications* video. https://cacm.acm.org/videos/how-to-teach-computer-ethics-with-science-fiction