**Developing Simulations to Train for Ethical-Decision Making in Computer Science Workplaces**

**Trevor Hubbard**

College of Information and Computer Sciences

tmhubbard@umass.edu

May 30th, 2019

**Advisor:** Professor Leon Osterweil

**Other Committee Members:** Professor Peter Haas

**Research Type:** Thesis

### Section 1: Introduction

The project that I plan to work on throughout the Fall 2019 semester is an interactive ethics simulation geared towards Computer Science students. This simulation would place players in the shoes of a professional in a CS environment, and challenge their ability to make decisions in an ethically challenging situation. During the simulation, players can choose from a number of different actions (talking to coworkers, consulting upper management, developing their code, etc.) that will influence the flow of the narrative; the simulation will ultimately conclude in such a way that highlights the impact of the player’s decisions!

The goal of the simulation is not to teach players a “correct” way to think about these situations; instead, it’s to try and facilitate some critical thinking about the situation, and to introduce players to the ethical complexity of the professional world. Dispersed throughout the experience are “reflection points”, where players will have to respond to writing prompts about the events of the simulation. Professors administering the simulation would then be able to read and evaluate the player responses, which would hopefully be a source of deep insight into the strength of a player’s ethical thinking. (Developing alternative methods of evaluating player engagement is another focus of our research.)

Throughout the Spring 2019 semester, I’ve been working with Professor Lee Osterweil, Professor Peter Haas, Professor Michelle Trim, Research Associate Heather Conboy, and two additional undergraduate students (Kyle Tan and Jonathan Trott) on researching and designing the scenario for the simulation. In its original conception, Professor Osterweil envisioned different ethically challenging scenarios being designed specifically for certain Computer Science classes. We’ve kept this idea in mind throughout the semester, and have divided our work accordingly: each of the undergraduates working on the project are designing a *different* scenario for this simulation.

The scenario I’ve been developing is one we’ve called “Testedness” - in it, the avatar character the player is controlling works at a medical devices company, and is working on developing the communications technology of an artificial pancreas device. The device is near its launch date, but the avatar begins to question whether the device is as robustly tested as it ought to be; after further investigation, it becomes apparent that there’s been no penetration testing done on the device, but management doesn’t seem concerned with the device’s security. Should the avatar follow the boss’s directive and try to finish testing on time, or try to convince the boss that the company ought to invest more time into security? The choice is up to the player.

Most of the work that’s been done during this 499Y semester was geared towards planning *what* exactly we’d build, and *how* we’d go about building it. After several rounds of storyboarding and countless roundtable discussions, I’ve developed a fairly robust outline of the flow of this simulation. During the beginning Fall 2019 semester, I plan on using Little-JIL (a process coordination language created by UMass’s LASER) to implement this outline into a workable simulation prototype. (The prototype will consist of graphics, text prompts, and a simple UI to facilitate player choice making; the flow of the simulation will be guided by Little-JIL’s agent-coordination logic.) Later in the semester, I’ll research using Little-JIL as a form of backend control for Unity, a game development tool; if this is feasible, my next iteration of the simulation prototype could benefit from an added layer of realism (as it would look much more like a modern video game, where players would be able to move around and make their choices in a 3D virtual office space.)

After developing an initial prototype, we’ll spend time evaluating players’ reactions to it; this work will be done through surveying volunteer playtesters. Professor Osterweil fully intends for the development of this system to continue development after this Fall 2019 semester. With that in mind, our main goal is to figure out how players interact with the simulation, and to develop systems in order to facilitate the sort of interaction we’re hoping for. One mechanic I want to devote particular focus to is the “rewind” ability - this mechanic would allow players to “turn back the clock” on decisions they’d made in order to see the results of navigating the scenario differently. I believe that a player’s desire to explore the range of outcomes of a situation speaks towards their ethical thinking abilities; naturally, then, the implementation of this rewind mechanic would allow this exploration, as well as allowing for further evaluation of player engagement. This “rewind” ability will be the focus of my thesis - after creating the simulation prototype, I’ll shift my efforts to researching players’ engagement with this mechanic.

### 

### Section 2: Significance

As time moves forward and computing technology becomes more and more powerful, it’s becoming increasingly clear that Computer Science students desperately need an opportunity to develop their ethical thinking skills. New technologies are creating new ethical quandaries for graduates. Software is being built and sold with a “minimum viable product” ethos, only to have any glaring flaws or lack of features fixed weeks or months after release - is this ethical? Machine learning has allowed companies to analyze and act on immense amounts of data in a cheap and efficient way - are these methods built with fairness in mind, though? The ubiquity of the Internet has encouraged development of the“Internet of Things”; while convenient and beneficial to the users, do these technologies have the privacy and safety of the user in mind?

Ethical, socially-conscientious computer scientists ought to ask questions like these throughout every stage of development; the world will be a much better off with more focus devoted towards discussing the potential issues that may arise from a product’s creation. One doesn’t need to look too far to find examples of projects or research where ethical thinking *wasn’t* as much of a focus as it should have been. (Any recent scandal on Facebook’s mismanagement of user data could serve as poignant food for thought.) By training students from an early age at how to navigate some ethically questionable situations, universities will be more positively shaping the computing workforce (and subsequently, the world at large).

Despite having a couple of courses that touch upon this sort of ethical thinking (namely, COMPSCI 320H and COMPSCI 305 - JYW), UMass’s CS department has no formal ethics course. Implementation of these ethics simulations into each core course would redefine the department’s attitude towards ethics; instead of being somewhat of an afterthought, the importance of ethical thinking would be a common thread throughout a large number of courses. With a motto like “Computing for the Common Good”, UMass has expressed a vested interest for training its students to grapple with ethical quandaries; the creation of these simulations would allow for this to happen on a widespread level in an easily adoptable fashion. The world of computing as a whole could benefit from this project in similar ways, whether its through direction adoption of this style of ethical training, or through an increased desire to flesh out their own ethics training programs.

My particular simulation scenario takes inspiration from two different areas. Most CICS graduates will experience some degree of “crunch” within their first few years out of school - testing can fall to the wayside in a developer’s effort to complete code according to a deadline, and this lack of testing could produce some unfortunate consequences to users of the product. By exposing students to this reality (and how one might choose to navigate it), hopefully they’ll be able to deal with it more gracefully once they enter the workplace.

My scenario also highlights a growing issue in the field of computing: the security (or lack thereof) of medical devices. While it may seem quite specific, this topic is encapsulated by a broader issue: the security of Internet of Things devices. Developers create these tools are concerned, first and foremost, with making a working product - implementing basic cybersecurity measures is an afterthought, if it’s even a thought at all. (For instance: botnets entirely composed of various IoT cameras have been created and used for DDoS attacks; the technique used to hack these cameras was as simple as running through a list of default device passwords.) Concerns for wrongdoing are multiplied exponentially when considering medical devices - studies have shown them to be just as vulnerable as any other IoT tech, and the health of users is directly at risk for manipulation if they are breached. Hopefully, through exposing students to an issue they might not have heard of, the next generation of medical devices workers will have a much higher level of concern about the security of their product.

### Section 3: Background

The research that I’ve done throughout this Spring 2019 semester has been fairly multi-faceted; I’ve looked into a couple of disparate topics of study, slowly but surely learning about the study of ethics, the practice of teaching it, the process of designing an informative narrative, the development and uses of an artificial pancreas, and some of the technologies that we’re going to use in order to build the simulation.

Because I’ve studied a wider range of topics, I’ve broken up the literature review into a couple of different subcategories:

#### Section 3.1: Ethics and Computer Science

*Entries in this category revolve around ethics as a general area of study, investigations into* how *to teach ethics, and the call for further emphasis on ethics within Computer Science education.*

[***ACM Code of Ethics and Professional Conduct***](https://www.acm.org/code-of-ethics)

As is stated at the beginning of the document, “The Code is not an algorithm for solving ethical problems; rather it serves as a basis for ethical decision-making.” This encapsulates the intentions of the ethics simulation we’ve been designing, and we’ve used some of the tenets of this document as guidelines for what ethical decision-making looks like. Some of the ethical responsibilities outlined in this document that are particularly relevant to my scenario include, “avoiding doing harm”, “knowing the existing rules pertaining to your work” (and challenging them when they’re unethical), and “designing and implementing systems that are robustly and usably secure.”

[***Weapons of Math Destruction***](https://weaponsofmathdestructionbook.com/)

This book takes an industry-wide look at some of the different ways that bias can be built into computing systems, and the dangers that this bias may present to system users if present. Additionally, the book advocates for developers to take more personal responsibility in ensuring that their systems are fair, and for legislators to create regulations for how these systems may be used. While medical device safety isn’t a topic explored in this book, there are still a number of investigations into ethical software development practices (and the consequences that could result without them).

Section 3.2: Artificial Pancreas Device Systems

*Entries in this category have helped me understand the technology and laws surrounding artificial pancreas device systems.*

[***Content of Premarket Submissions for Management of Cybersecurity in Medical Devices***](https://www.fda.gov/media/119933/download)

The FDA last updated their cybersecurity guidelines in 2014, but the landscape of medical technology has evolved enough in the time since that they put out another draft in 2018. In this draft, the FDA outlines a number of precautions that medical device developers ought to take into consideration when building their devices. Prevention of unauthorized use and consistent sustaining of code integrity and data confidentiality are primary concerns of the FDA, and they outline a number of methods developers can employ in order to ensure the aforementioned things. Somewhat unfortunately, though, all of the guidance provided in this document is simply that: guidance. The FDA has no official cybersecurity regulations, so medical device makers aren’t held to any specific standards when building their products.

[***Content of Investigational Device Exemption and Premarket Approval Applications for Artificial Pancreas Device Systems***](https://www.fda.gov/media/80644/download)

This is another set of guidelines and regulations from the FDA; these ones are geared specifically towards the development, testing, and release of artificial pancreas systems! This document provides a detailed breakdown on the design and functionality of these systems, including an exploration of the core components of the system, their communication, and how the control algorithm ought to work. Additionally, the document explained some of the standing regulations on the development of these sorts of systems - these regulations were mostly concerned with how to administer clinical trials, and had little mention of implementation of cybersecurity features.

#### Section 3.3: Building Simulations

*Entries in this category are resources that have taught me important information about what it’ll take to actually build my ethics simulation.*

[***Little-JIL 1.5 Language Report***](http://laser.cs.umass.edu/techreports/06-51.pdf)

This manual introduced me to Little-JIL, which is the agent-coordination language that we hope to use in building our simulation. Little-JIL functions, on a high-level, by assigning agendas (filled with tasks to be completed) to specific agents, which are defined by the programmer. Each task can have certain prerequisites and attributes associated with it, which’ll allow us to parameterize interactions within the simulation with ease. Since Little-JIL is a visual language, too, it’ll be well-suited for representing the flow between states of the simulation!

[***Designing and Using Games to Teach Ethics and Ethical Thinking***](https://dl.acm.org/citation.cfm?id=2811156)

In this paper, Karen Schrier examines a slew of games that’ve been used to try and teach ethics to players, as well as some frameworks that’ve been written for designing these games. Some of the best practices that were identified include making the consequences of choices clear, allowing time for player-character relationships to develop, and using realistic scenarios and contexts. When outlining my simulation scenario, I’ve tried to keep these practices in mind, as to impart the lessons I’m trying to teach as best as possible.

### Section 4: Methodology

As was mentioned earlier, I’ve spent the semester developing an understanding of ethical issues and artificial pancreas device systems. This research was supplemented with some invaluable storyboarding training from Professor Trim. Through this training, I’ve been taught how to develop stories from the ground up; Professor Trim pushed me to constantly ask important questions regarding the intent of the simulation, the structure of the exposition, the delivery of the ethical quandary, and much more. We also experimented with various storyboarding formats and story layouts; with each new attempt, my ability to conceptualize the complex, branching flow of the simulation grew stronger.

This training, coupled with my continuous research and discussion of these topics, has led me to create an extensively detailed outline of my simulation’s scenario as a whole. This outline begins a high-level overview of the simulation, which looks like this:



Each section of the overview is then further expanded upon, examining the purpose of any interactions that may occur during, examples of dialogue that may appear, and various mechanics that the player can take advantage of as they’re progressing through the simulation. Parameters that determine the player’s available interactions are also outlined, as well as how the player can influence them.

This outline is going to serve as a hugely important resource for the development of my simulation. Once the semester begins, I’m going to spend some amount of time learning how to use Little-JIL (despite having read through the language report, I haven’t actually had any practice writing with language). After achieving some level of proficiency, I’ll begin creating my simulation in Little-JIL! CICS’s Heather Conboy is going to spend some time creating a platform in Little-JIL for these simulations over the summer, so we ought to jump into this part of the process fairly quickly.

When I’ve built a working simulation prototype, I’ll move into an evaluation phase. As of right now, my current plan is to present the simulation to a group of volunteer students (most likely recruited from Professor Osterweil’s 320H class; the potential for this has been mentioned a couple of times by Professor Osterweil). After running through the simulation, said students would fill out a survey with some of their thoughts on their experience. The survey would ask questions like, “were there any parts of the simulation that you felt were unrealistic?”, and, “what did you think about the pacing of the simulation?”

In tandem with the survey, I’d also look towards a number of other factors for the evaluating the sim’s effectiveness. As was previously mentioned, there are “reflection points” interspersed throughout the experience; we could look towards the volunteers’ responses to those prompts to get an impression of how players are thinking about the ethical dilemmas. Ideally, when the simulation is fully completed and integrated into courses, these reflection responses will be used by the professors to assess the impact of the simulation; depending on how the volunteers respond, we can tweak the simulation or the reflection prompt to elicit more of a response.

As there’s no real way to measure if players have become more ethical decision-makers, our research group has settled on another, more measurable metric to test effectiveness: player engagement. Engagement can be measured in a number of ways. The first that comes to mind is the amount that the player explores the simulation - if one player chooses to interact with twice as many simulated co-workers than another player, then the first player is probably more engaged with the narrative than the other one. If we’re able to implement a “rewind” feature, the amount of times that a player chooses to “rewind” to see other conclusions could also be a great indicator of interest.

We’ve spent some time researching more methods to check for engagement, but we’re still forming a consensus on which methods will be more effective: one of the undergraduate students, Jon, is focusing his research efforts on evaluating which metrics will serve as a good measure of engagement. Rest assured, though, by the beginning of the next semester, we’ll have hypothesized about these metrics and how to factor them into our interpretation of the sim evaluation, so that our evaluation will be able to proceed smoothly. The results of this evaluation will inform the iteration process, where I’ll be able to change the simulation according to the feedback I received; ideally, this process of evaluation and iteration will continue to happen even after my time in 499T, eventually culminating in a very polished version of the simulation that’s ready to be deployed in the classroom.

Along with the working prototype, I’ll have written my thesis about the players’ use of the rewind functionality. Through this thesis (and the other ones that will be written by Jon and Kyle), future developers of the simulation will have definitive pieces of knowledge about the effectiveness of different aspects of the simulation.

### Section 5: Evaluation

As was described in the previous section, the 499T semester will be broken into a couple of “phases”: prototype creation, prototype evaluation, and finally, prototype iteration. Another final phase, the thesis presentation, will happen near the end of the course - this will be a presentation of my work to my research group and others from the Honors college.

As follows is a more detailed description of what I’ll be evaluated on during each phase.

**Prototype Creation -** By the end of this phase, I ought to have some sort of a working prototype of the simulation. While it may not be as extensive in scale as we hope the final simulation will be, it ought to provide many of the functionalities we’ve talked about, including:

* A branching narrative that can end in a couple of different ways
* Underlying parameters that’ll affect how avatar interactions are handled
* Presentation that’s more than just textual (pictures, sounds, or even video)

**Prototype Evaluation -** During this phase, I’ll need to first complete some sort of evaluation plan; this plan will include the survey questions that’ll be given to the volunteers, as well as a guide for how I’ll plan on interpreting the results of my evaluation. (This moreso applies to some of the underlying engagement metrics we’ll be measuring, such as “percentage of the simulation explored”; these metrics will be quantitative, as opposed to the qualitative survey responses, so I’ll need to define how to interpret them.)

Actually *administering* the simulation to the volunteers is another important component of this phase; while I’m sure I’ll receive some level of help from the professors in my research group, I expect to take charge on setting up these volunteer sessions and making sure they run according to plan. Finally, once I’ve received all of the feedback from these playtests, I’ll be expected to write a report on what I’ve learned, and how I ought to use this newfound information to change the simulation.

**Prototype Iteration -** Throughout this phase, I’m going to be working on implementing some of the changes I’ve outlined in my evaluation report. I’m predicting that some of these changes will be easier than others to implement - for example, making a specific interaction’s dialogue more realistic will be much simpler than totally rehauling the pacing of the simulation. For this reason, I don’t expect that the second iteration of the prototype will totally address *all* of the feedback given during the evaluation phase; instead, I’ll make a concerted effort to cover as much feedback as possible.

By the end of this phase, I should have a second iteration of the prototype, along with a report of all of the changes that have been made to the simulation. I’ll use this changelog report as a resource for writing my thesis, showing what’s worked and what hasn’t during the development of the simulation.

**Thesis Presentation -** During this phase, I’ll write a thesis on my progress throughout the entire semester. Within the thesis, I’ll address what I’ve learned about the process of building the simulation, as well as how it was received by the volunteers who played it. (This’ll read like an overview -of the work done in the previous three phases!) There will be a focus placed on the rewind mechanic, and an analysis about whether it’s worthwhile for future developers to retain this functionality in their simulations. Furthermore, as the development of these simulations are meant to continue after my 499T semester concludes, I’ll outline what I think the “next steps” are, and things future developers ought to take into consideration. Along with this thesis, I’ll deliver a presentation of this report to the professors in my research group (and any other professors on the honors board).

My work for the semester will be evaluated based off of the timely completion of each of these phases, along with the thoroughness with which each of them are completed. A full mark in the course, in my prediction, would be given for satisfactory results (judged by the professors in my research group) produced during each phase.

My research group plans on continuing the same weekly meeting schedule we’d set up throughout the 499Y semester; during these meetings, we’ll show off our progress on each of our milestones and receive helpful feedback from the professors.

Once the end of the semester is over, I’ll save any source code and executables to a USB drive, and pass that in with my 499T thesis!

### Section 6: Communication

Like was mentioned in the previous section, I expect to meet once a week with all of the members of my research group. During this meeting, we’ll spend some time talking about the progress of our work, and ask any relevant questions. We’ve had a lot of great discussions about the project (regarding the simulation’s purpose, how to best deliver the simulation’s message, how to make the simulation realistic, etc.) throughout this past semester in 499Y, and I expect these sorts of discussions to continue into this semester. We also might devote some time to speak on current events within the world of computing, and how they might be relevant to our work as a whole.

Additionally, a couple of times throughout the semester, I expect to meet one-on-one with Professor Osterweil. (This was *also* done during the 499Y semester!) During these one-on-one meetings, Professor Osterweil will ask some more targeted questions about the progress of the research (regarding whether development is going how I might have initially imagined it, about any issues I might have with the process, about specific topics I’m looking to research, etc).

Generally, between these group meetings, the one-on-one meetings, and individual time spent working on the project, I expect to be devoting roughly 10 hours per week towards the development of the simulation!

### Section 7: Timeline

As follows is a timeline for the completion of the phases laid out in Section 5:

**Sept. 1st - Oct. 1st:** Development of the first prototype of the game

**Oct. 1st - Nov. 1st:** Evaluation of the first prototype

**Nov. 1st - Dec. 1st:** Creation of second prototype w/ feedback from the evaluation

**Dec. 1st - Dec. 15th:** Writing the thesis

### Section 8: References

As follows is a list of references used for this report; this list uses ACM reference formatting:

Alexander Wise. 2006. *Little-JIL 1.5 - Language Report*. Department of Computer Science.

University of Massachusetts, Amherst, MA.

ACM Code 2018 Task Force. 2018. *ACM Code of Ethics*. (June 2018). Retrieved May 19, 2019

from https://www.acm.org/code-of-ethics

Cathy ONeil. 2017. *Weapons of Math Destruction*, New York, NY: Broadway Books.

Center for Devices and Radiological Health. 2012. *Content of Investigational Device Exemption*

*and Premarket Approval Applications for Artificial Pancreas Device Systems*, FDA.

Center for Devices and Radiological Health. 2018. *Content of Premarket Submissions for*

*Management of Cybersecurity in Medical Devices*, FDA.

Karen Schrier. 2014. *Designing and using games to teach ethics and ethical thinking*. In

Learning, Education and Games, Karen Schrier (Ed.). ETC Press, Pittsburgh, PA, USA

141-158.