Science Fiction as an Introduction to AI Research

Judy Goldsmith and Nicholas Mattei
Department of Computer Science
University of Kentucky
Lexington, KY 40506

Abstract
The undergraduate computer science curriculum is generally focused on skills and tools; most students are not exposed to much research in the field, and do not learn how to navigate the research literature. We describe how science fiction reviews were used as a gateway to research reviews. Students learn a little about current or recent research on a topic that stirs their imagination, and learn how to search for, read critically, and compare technical papers on a topic related their chosen science fiction book, movie, or TV show.

Introduction
Science fiction has inspired generations of would-be computer scientists and engineers. Some draw direct lines from particular works to subsequent inventions: From Heinlein’s “Waldo, Inc.” to modern automated assembly lines (Heinlein 1950); from Neal Stephenson’s Snow Crash to Second Life (Stephenson 1992). Many computer science students are avid science fiction readers. This paper addresses the question of how to harness their enthusiasm to propel those students past a state of passively fulfilling course obligations, and into the world of artificial intelligence research.

Too many undergraduate computer science students finish their education with no idea what computer science research is. Many of the students are in school and the major to get a well-paying job and do not pursue graduate study (Computing Research Association 2009). Many of these students have no motivation to explore the research literature, no clue about what constitutes research, peer review, or publication.

It is better for the field of computer science if practitioners have some connection to research in the field. Technology and knowledge are changing rapidly; practitioners must be able to find more information about things they read about in Wired or Slashdot, and must be able to evaluate the sources of information. It’s important that our students be exposed to the research literature, and have at least a rudimentary understanding of the peer review process. Some may end up discovering a taste for, or passion for, research. They may contribute to others’ research by offering technical challenges, or possibly funding. And they should be able to access and take advantage of the work that has been done.

The goal of the exercise described here is for students to start from their own interests, and explore research related to those interests. The context of this discovery is a senior-level computer science elective, Introduction to Artificial Intelligence. The exercise is student reviews of science fiction books, movies, or other media that contain significant AI content.

For the last several years, I’ve walked in to the first Artificial Intelligence class of the semester and asked the students’ permission to not give exams. After a while, someone in the class usually recovers from the shock and asks what would replace the exams. I tell them that there are many options, and each will have to choose one option for their midterm and a different one for the final. We discuss the options, and they vote.

Approximately 2% of the students to whom I’ve offered these choices have requested an exam.

The options on offer are that they may review a book, movie, or game with significant AI content; write a survey paper on some AI topic; implement an algorithm from the research literature; give an in-class talk on a research paper or papers; or write a short story with AI content based on actual AI research. All of the options have been chosen, but the book and movie reviews are by far the most popular choice. This paper will discuss the advantages of that option, and what we all bring to the process.

Background and Related Work
The use of AI as a hook for CS participation is not a new one. AAAI had a recent symposium on “Using AI to motivate greater participation in Computer Science” (Sahami

1The same exercise has also been used in graduate courses on AI, machine learning, and cognitive sciences, with similarly enthusiastic responses. This paper focuses on the undergraduate experience.

2The first author is the primary instructor for this course. Where appropriate, the first person is used to relay personal anecdotes and preferences. The second author has served as a sounding board, pedagogical expert, and occasional lecturer for the classes over the last several years.

3This section owes a debt to the AAAI “AI Topics” site (AAAI 2011).
Similarly, the idea of AI inspiring science, and computer science in particular, is very popular (see, for example, (Sawyer 2002; Watson 2003)).

This course is not the only one using science fiction to inspire students or to drive interest in artificial intelligence. Bates uses science fiction as motivation to talk about AI in a general education (nonmajors) course, and as an entry point for talking about ethics (Bates 2011). Bowring and Tambe describe their use of science fiction and games as drivers of courses—offered to audiences as varied as students’ parents; pre-freshmen; and computer science majors—on multi-agent systems. They use short stories and short video clips, TV shows and movies, to present issues such as agents’ models of other agents, risk averseness, and even the use of emotions (Bowring and Tambe 2009).

The assignment described here is different from Bowring and Tambe’s because they present the science fiction as class readings or in-class video, rather than letting the students choose their own readings. Furthermore, they lead the students’ explorations of the relevant technical issues, rather than pushing them to discover and explore the research literature. By being directive, Bowring and Tambe are able to focus attention on specific topics, such as distributed Partially Observable Markov Decision Processes, more deeply than our students usually do. However, our approach allows students a higher level of autonomy, thus allowing them to pursue their own interests in science fiction and AI, and giving them a flavor of self-driven research.

Courses in other fields use science fiction for motivation. Bowring and Tambe list two, including an intriguing case that used science fiction as a tool for teaching children to think about the future (Dils 1987).

Others have used games as a motivation and platform for exploring issues in AI (McGovern and Fager 2007). In McGovern and Fager’s course, they use a gaming platform for assignments. Our students can also choose to investigate AI in games instead of science fiction. Some of them implement searching or pathfinding strategies to improve game-generated characters. However, the complexity of the programming challenges tend to limit the depth to which they explore the AI literature. This is balanced by the students’ enthusiasm for the projects, when they choose those projects. They are required to choose different forms of projects/papers/presentations for the midterm and final, so they are exposed to the research literature in at least one of those instances.

In computer science, as with many high consensus fields, there is a tendency to teach from a “facts and principles” standpoint (Colbeck 1998). While many feel that this is a necessary evil in order to establish a firm foundation of core knowledge within students, there is no doubt that it reinforces “absolute truth” views of knowledge (Haworth and Conrad 1995). Students may graduate with gaps in their ability to think and reason in situations where there may be more than one answer. While the day to day programming tasks of the majority of our graduates may be routine, there are still critical moments in students’ careers where they will need to judge two or more seemingly similar technologies. We must properly train them to cut through marketing and PR material and recognize that one, the other, or both technologies may be appropriate for the task at hand. Without addressing and fostering our students’ ability to think critically, we may forever leave them unable to judge multiple solutions in a fair and discerning way. Engaging multiplicity and other forms of critical thinking through exposure to research and writing will provide our students with examples of thinking that move beyond dualism and other didactic modes of reasoning (Davis 2009; Perry 1980). We believe that exposure to multiplicity and critical thinking will better equip our students to engage in debates about technology, applications, and knowledge from multiple viewpoints.

The Assignment

Book or movie reports will cover a work that uses AI as an integral part of the plot. You will describe the plot and the role played by AI; analyze the author(s)’ concept of AI, and conclude by discussing the feasibility of this concept. Discussions of social ramifications are welcome but not required. You will be graded on the thoughtful analysis of AI in this work; readability, prose structure, and technical details of writing (spelling, punctuation, etc.) will be a non-negligible part of the grade (Goldsmith 2010b).

AI is taught as a senior-level elective. Most of the students who take it are CS majors who are out of the habit of writing, and until recently, none had taken a technical writing class. Left to their own devices, many would begin writing a day or two before the assignment was due. Therefore, proposal deadlines are set several weeks ahead. Helping students schedule and plan their writing activities is essential to ensuring success in their writing endeavor (Davis 2009).

You will list the book/movie that you will be reviewing, cite its dominant AI theme(s), and outline the review (Goldsmith 2010a).

Proposals are submitted on paper or emailed. Email is preferred, to enable quick responses and dialogue. The goal of the discussion is to find the “hook”: the AI challenge that interests them. If a students proposal is too weak or there is not significant AI content then I work with the student extensively to find an appropriate choice. While I have opinions on the suitability of many novels, the student has the opportunity to argue for their choice.

Once the student has chosen a work of fiction and an AI challenge, the next email interchange helps them to build a bibliography. This is an intensive process, both because it is different for each student and because students often bring up topics with which the instructor is not familiar. Ideally, the instructor should be able to point them to authors who write accessibly on the topic of interest, or at least, conferences and journals related to the topic. The goal for an undergraduate is that they actually support their conclusions about feasible and possible technologies by investigating ongoing or recent research on their chosen subject. Students are expected to read at least two peer-reviewed technical papers or chapters of advanced textbooks or monographs.
The goal is not to denigrate Wikipedia, but to convince them to learn more about graduate school. That leads to a discussion of theses and dissertations, and graduate school in general. This provides an opportunity to extend the knowledge of the field. The scholarship of integration, application, and teaching. The scholarship of discovery, which involves research into effective teaching methodologies, is not covered in this assignment. The scholarship of integration is taking ideas from one field and applying them to another field or body of work; students learn that applying existing ideas within the field to new domains in order to extend the knowledge of the field. The scholarship of teaching, which involves research into effective teaching methodologies, is not covered in this assignment. The scholarship of integration is taking ideas from one field and applying them to another field or body of work; students learn that the act of reading and writing a review article is "research," specifically, the scholarship of integration.

For many students, the idea that "research" can be incremental is both startling and reassuring. This sometimes leads to a discussion of theses and dissertations, and graduate school in general. This provides an opportunity to encourage them to learn more about graduate school. That conversation often occurs late in the semester, if and when they have become excited about a current research area.

The next topic is quality control. I try to have recent instances of nonsensical news items or bad Wikipedia entries. The goal is not to denigrate Wikipedia, but to convince them that there is value in stable, peer-reviewed research presentations. I divide technical writing into several categories.

1. Peer-reviewed journals and conferences;
2. Textbooks, handbooks, and monographs;
3. Tech reports and web pages;

Within the first category, discussion centers on the reviewing process and criteria, and the social structure of reviewing. This is an opportunity to talk about a professor’s job, and the expectations that we will be on conference program committees and editorial boards. This further allows us to discuss researchers’ responsibilities with respect to others’ research—a more engaging angle to the oft-repeated discussion of plagiarism, as well as what is often their first discussion on co-authorship.

Students ask about the reviewing process for textbooks, and often complain about books they’ve used. This is a springboard to remind them of the time our colleagues have put into writing textbooks, and suggest that there is social value to buying books, whether paper or electronic.

The discussion of web pages begins with a claim that anyone can post anything. I provide several examples; once students themselves uncover examples of nonsensical web pages, they are more invested in the idea of peer review.

Next, we discuss how to find relevant articles. One goal of this exercise is to make students more information literate: they should know not only how to navigate different sources of literature but also how to extract relevant information from what they find. They are introduced to backward referencing (reading the papers referenced in the current paper) and forward referencing (reading papers that cite the current paper). We start with Citeseer (Lawrence et al. 2005) and Google Scholar, but they are encouraged to look elsewhere as well.

Because students are so used to search engines, they tend to focus on individual papers. Once they’ve chosen their AI topics, we share links to relevant conferences and journals. They are encouraged to investigate multiple issues/instances of the journal or conference.

We discuss the value of citations: to give pointers to details and related work, and to give credit. This leads to a discussion of how to cite, how to quote, and how to choose between them. That leads to a discussion of opinions. Many students will read an opinion, stated factually, and will report that opinion as if it were their own. For instance, “Modern AI research has focused on reasoning under uncertainty,” or some other broad generalization for which the students do not have sufficient information or experience in the field. Students are encouraged to take ownership of their own opinions, and distinguish others’ opinions from facts.

Discussing Writing

In the course of the semester, the students generally read four or more technical papers. If the AI topic that they choose is outside of my own interests, I cannot necessarily guide them to the best-written papers. Student complaints about poorly written papers are teachable moments: I ask them to analyze what makes the paper hard to read, and discuss the importance of clarity and organization.
Since some of the students are, themselves, writing technical papers (the survey option appeals to students who already have an interest in an AI topic, or who dislike genre fiction and presentations), I spend time discussing what makes a good technical paper.

What problem is being solved? Good papers should address a problem and make it clear both what the problem is and why it is important.

What have others done to solve the problem? Papers should include a survey of what others have done in order to legitimize the problem itself and to frame the new result within the existing literature.

What is new here? If the paper is about a new idea, method, or problem. It should be clear what the paper’s new contribution is, and how that contribution was evaluated.

Why is it better than other solutions (or why not)? If a new idea or method is to be adopted by the greater community it must be made clear how the new method improves over old methods. Though just as valuable, but much more rare in the literature, methods that are not improvements need to be explained as well. By explaining clearly what went wrong and why the scientific community has learned what approaches do not work.

One of the most difficult aspects of teaching survey writing is to move students from the annotated bibliography format (one paragraph or section per paper read) to a more analytic framework. While instilling analytical writing is not the focus of the assignment, discussions on organizing technical materials pays off in the students reviews.

Examples

The quality of writing has varied. When I first gave this option, I did not insist on a bibliography of research sources. The students tended to pull conclusions from their imaginations, heavily influenced by other works of fiction and popular magazines such as Wired. At first I was afraid that their initial encounters with potentially impenetrable technical prose would scare them away from AI. However, I have not seen that happen. As the requirements have grown more rigorous, the papers have improved significantly, as has the students’ satisfaction with the exercise.

The Bad

In the first few iterations of this assignment, I regularly got submissions on 2001 A Space Odyssey (Kubrik and Clarke 1968) and I, Robot (the original stories, usually) (Asimov 1950). Students found 2001 more obscure than they expected. They tended to describe Asimov’s ethical conundrums without much link to modern scholarship. I began to discourage students from choosing these books. Given the renewed interest in “robot ethics” (See (Anderson and Anderson 2007)); I might be willing to see that collection revisited.

Despite my best efforts, some students do not absorb the lessons about the quality of unrefereed websites, for instance. One of the earliest papers I received on I, Robot insisted that positronic brains were, indeed, feasible, “because Dana [sic], the robot in Star Wars [sic], has one.” The reference was to what appeared to be an undergraduate paper at another university.

The Good

As with any undergraduate writing assignment, some instances are earnest and awkward. On the other hand, there are absolute gems. Recently, a student submitted a paper that began with a description of Sigourney Weaver manipulating a walking forklift (Aliens (Cameron 1986)) using her feet. The challenge there was to control a mechanism with many degrees of freedom, although the controller had only three degrees of freedom. The student then looked into applications of AI to understanding and applying complex, multi-dimensional motion in response to lower-dimensional controls. The AI comes in matching the sense inputs to a goal, and the goal to a set of motions or a control policy.

Another student started with Babel-17 (Delaney 1966) and wrote about the Sapiir-Worrff hypothesis and natural language processing.

Many students ask for suggestions for novels or movies. I have suggested Hellspark (Kagan 1988) for the explicit discussion of intelligence, and received a paper on intelligent gaze control (for the robot whose intelligence is investigated). I suggested the movie, Minority Report, for the swarm, and received a paper that investigated the engineering and AI aspects of iris recognition, as portrayed in the movie. The student investigated processing speed in visual pattern recognition, as well as focal lengths for then-current cameras. He concluded that the technology was not yet in place to pull sharp enough images out of video from cameras as far away as the scanners were in the movie. The bottleneck was not, he said, the software, but the hardware.

Measures of Success

A common measure for course success is students’ evaluations of the course. In recent years, the evaluations for the first author’s Introduction to Artificial Intelligence course have gone up. The comments from students were uniformly enthusiastic this year, including, “Best class I took at UK [the University of Kentucky]. I really enjoyed doing projects over exams. I feel I learned much more that way.”

One of the goals in using science fiction and requiring writing is that students will continue to be scholars long after the last class of the semester. On teacher course evaluations (TCEs) there is a specific question, “The course stimulated me to read farther in the area.” Table 1 shows the average TCE score for this question before and after implementing the writing assignment described in this paper. All TCE question responses from students are integers between 1 and 4. Table 2 shows the student’s responses to the question “Rate the overall value of the course,” both before and after the writing assignment was implemented. The results of a single factor ANOVA test before and after the writing assignment was implemented shows a statistically significant (p < 0.0011) increase in student interest in continued reading. In addition to a continued interest in reading, there
Table 1: Student responses to “The course stimulated me to read farther in the area.” There is a statistically significant increase in the student ratings before and after the implementation of book reviews, $F(1, 5) = 43.639, p < 0.0011$.

<table>
<thead>
<tr>
<th>Semester</th>
<th>Mean</th>
<th>Semester</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2001</td>
<td>3.00</td>
<td>Fall 2007</td>
<td>3.60</td>
</tr>
<tr>
<td>Fall 2003</td>
<td>2.90</td>
<td>Fall 2009</td>
<td>3.70</td>
</tr>
<tr>
<td>Fall 2004</td>
<td>2.80</td>
<td>Fall 2010</td>
<td>4.00</td>
</tr>
<tr>
<td>Fall 2005</td>
<td>2.60</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Student responses to “Overall value of the course.” There is a statistically significant increase in the student ratings before and after the implementation of book reviews, $F(1, 5) = 8.653, p < 0.05$.

<table>
<thead>
<tr>
<th>Semester</th>
<th>Mean</th>
<th>Semester</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2001</td>
<td>3.40</td>
<td>Fall 2007</td>
<td>3.40</td>
</tr>
<tr>
<td>Fall 2003</td>
<td>3.20</td>
<td>Fall 2009</td>
<td>3.80</td>
</tr>
<tr>
<td>Fall 2004</td>
<td>2.80</td>
<td>Fall 2010</td>
<td>3.90</td>
</tr>
<tr>
<td>Fall 2005</td>
<td>3.10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

is a statistically significant increase in the student perception of the overall value of the course ($p < 0.05$). The enrollment of the courses has been roughly constant over the years.

While TCEs are not the most reliable instrument for gauging student involvement and course success, they are considered valid by many researchers (McKeachie 1997; Marsh and Roche 1997). Systematic biases in TCEs have been hypothesized, however, biasing factors are generally some combination of: instructor gender, grading leniency, course requiredness, course workload, class size, and class time of day (Fleming, Bazen, and Wetzstein 2005; Guerin and Michler 2011); all these factors are nearly constant across all course offerings considered here.

In 2009, over half of the 18 students who completed the class chose a book or movie review. (Records are incomplete, but that year there were at least 4 original short stories that included aspects of AI.) This past year, 7 of 8 students in the class chose book or movie reviews for their midterm project. There were two reviews of Blade Runner (Fancher and Peoples 1982), that looked at emotions, and ethics; one of Daughters of Elysium (Slonczewski 2009), that looked at smart house projects; The Hacker and the Ants (Rucker 2003), that looked at swarms; the movie Moon (Jones and Parker 2009), that looked at ethics; and one review that did not actually discuss the book at all, but looked at intelligent prostheses and mental control of insect-based cyborgs. The eighth student gave a talk about game AI.

One recent student, writing about the movie Screamers (O’Bannon and Tejada-Flores 1995), wrote,

My intention for the paper is twofold; To explain the fact vs. fiction behind the movie (what capabilities we are currently able to achieve) and to explain the social consequences, limitations, and potential benefits of the use of such weapons.

Many of the students go on to do a final project on a related subject (they are limited to at most one review.) The student who looked at swarm papers for his midterm went on to implement an insect colony algorithm for his final. The student who spoke about game AI then wrote a survey paper on chess algorithms.

The students express great satisfaction with being able to choose the topics and formats of these projects/papers. Several have come to the class with particular interests, and have explored aspects of their topics via fiction, surveys, implementations, and presentations. All discovered something new. In all of the reviews, the fictional presentation of some aspect of AI provided direction and motivation for their reading.

I have not yet convinced any students to extend their papers for publication, although some have been good enough. However, most years, I recruit at least one student from the class for a research project, and several of those have led to publications or presentations.

**Points To Relevant Science Fiction**

I maintain a list of science fiction novels with significant AI content. In addition, AAAI maintains a web page that discusses the influences of science fiction on AI research, and vice versa (AAAI 2011). Wikipedia has a very useful page, AI in Fiction (Wikipedia a), which includes print media, movies, and TV shows. There is also Fictional Computers page which includes these media, plus comics, graphic novels, computer and video games, board games, and role-playing games (Wikipedia b), and Fictional Robots and Androids page with a significant non-American set of entries, particularly under graphic novels (Wikipedia c). There are other sites, including a particularly nice site, Robots in Films (Dirks ). Also worth mentioning is the Cognitive Science Movie Index from Indiana University (Mott ), which explicitly states the cognitive science themes for each movie. (A student in the Intro to the Cognitive Sciences class brought this site to my attention, and several students in the class used it.)

**Conclusions**

Late in the semester, students are told that the purpose of the reviews is purely subversive: to change how they experience popular culture, particularly fictional (and popular media) portrayals of technology and AI. The goal is that they approach encounters with authors’ imaginations with both wonder and a desire to investigate. Even bad science fiction can inspire us to learn more about some facet of AI.

Good science fiction builds on the feasible and extrapolates to the possible. Good research does the same. What an author can imagine, perhaps our students can implement, now or later in their careers. A researcher’s currency is new ideas. The students learn that ideas can come from fiction as well as the research literature.

In addition, the freedom the students have to choose or discover an AI topic is often their first taste of self-guided
research. Many of them conclude the semester eager to continue reading about one of their chosen topics. Typically, at least one student from the class does a research project the next semester. Several have been coauthors on published papers, albeit unrelated to their class papers. However, the tracking in tracking down, reading, and making sense of research papers has served them well.

References


Asimov, I. 1950. I, Robot. Gnome Press. There are many other editions!


Guerin, J. T., and Michler, D. 2011. Analysis of undergraduate teaching evaluations in computer science. In SIGCSE.