Proposal for
CS499 for FALL’2006
Intelligent Agents and Science Fiction

Milind Tambe, Emma Bowring, Jon Pearce, Zvi Topol
Computer Science Dept
University of Southern California

Sept 19, 2005

Mission: Introduce fundamental concepts of agents and multiagent systems to get undergraduate students interested in Artificial Intelligence in general and agent and multiagent systems in particular, using science fiction as our vehicle to motivate key concepts.

Essentially, we propose to introduce a technical computer science course with an emphasis on fundamental concepts in AI that are in heavy use in agents and multiagent systems, e.g. Nash equilibrium. This is not a discussion course, nor a course on sci-fi literature. In fact, as can be seen in the detailed syllabus provided, there are technical papers from the agents/multiagent systems arena that serve as readings for this course. By carefully selecting particular sections of these papers for emphasis, and with detailed notes, we eventually wish to match a 300-level course in CS. In other words, in this course, science fiction is the ‘spice’ for the main dish of agents and multiagent systems.

So, why have science fiction? Science fiction provides three key benefits in this course. First, it is a great motivator and it provides context, generating excitement about AI topics in general, and agents and multiagent systems in particular. This first point is the main role of Sci-Fi in this course. Second, science fiction also helps provide a perspective on how far we have come in our research, as well as current limitations, and future research challenges. Third, science fiction literature is a great vehicle for understanding the impact on society if agent-based computing truly succeeds.

Thus, at the end of the course, students: (i) most importantly will have been introduced to fundamental technical concepts in building agents and multiagent systems, (ii) will have broadly understood the challenges and limitations of agents and multiagent systems today, (iii) exposed to some cutting edge research/social issues in introducing agents and multiagent systems in society.

In the rest of the document, we cover the pre-req, our lecture plan motivation and description, detailed syllabus, grading, relationship to CS460, and other key issues.

Pre-req: CS101 or equivalent: Our target student population: Students who have some knowledge of computer programming. If this course were to be approved into a regular course sometime in the future, the intended level for this course would be a 300-level course in computer science.
Lecture plan motivation and description: There are two key new concepts in this course:

1. Each lecture or group of lectures in this course builds on a science-fiction specific story or a sci-fi TV episode. The sci-fi literature chosen exhibits the traits: agent modeling, emotions, teamwork, agent interactions under uncertainty, etc. After reading about one or more of these traits, our lecture will draw out what is needed to make these agent programs exhibit (or in some cases avoid) this behavior. For example, the story LIAR (by Isaac Asimov, introduced in Week 3) describes a “mind-reading” robot. This mind-reading ability essentially brings out the notion of agent modeling in the agents and multiagent systems literature: how might one model the beliefs, desires and intentions of humans or other agents? This also motivates discussion around agents that do mind-reading vs those that are mind-blind. The story RUNAROUND (also by Isaac Asimov) brings up a race condition, which in effect brings up the notion of intention reconsideration and conflicting commitments. We may explain the behavior in terms of Belief-desire-intention (BDI) concepts, and understand how to avoid such conflicts.

There are other courses offered in other universities that use science fiction as a way of introducing science in general, of which in general AI might be a very small part, and students there have given excellent feedback. Prof. Barry Luokkala of Carnegie Mellon University, Teaching Professor in the Physics department, uses science fiction for introducing science. His students have provided him excellent feedback. Prof. Luokkala also gives lectures on Commander Data from Star Trek, and he might be a guest lecturer. Prof. Maja Mataric pointed out a similar course offered at Brandeis University by Prof. Sekular.

Prof. Stacy Marsella has pointed to us the similarity of this course to NIH's program called "Science in the Cinema" [http://science.education.nih.gov/cinema](http://science.education.nih.gov/cinema) and the American Chemical Society's similar program: [http://www.scalacs.org/ScienceCinema/](http://www.scalacs.org/ScienceCinema/). In particular, those programs also use film and world leading scientists commenting on the films to educate students and the general public.

2. We also propose to invite key outside lecturers, possibly professors who specialize in science fiction, or science fiction writers or creators of science fiction TV episodes. These lectures may serve the purpose of (i) motivation, context and excitement of agents and multiagent systems; (ii) elaboration on the impact of science fiction on our science, or (iii) the cutting edge issues in impact of agents and multiagent systems on society in general (which may include general ethical issues, social issues). For example, we have had discussions with Prof. Paul Alkon of USC who teaches science fiction in the English department. Similarly, Prof. Barry Luokkala at CMU might be another candidate. Mike Zyda has agreed to get us in touch with key sci-fi figures, such as Alex Singer of Star Trek. However, these social issues are not the main thrust of this class.
Lecture plan (15 weeks):

Note:
- Students are NOT required to read all the suggested readings. We recognize that in some cases this may be asking too much from students to read.
- However, the suggested readings do provide strong technical content and hence a way of changing the precise level of technical content in the course. In some cases, we may promote some selected sections of such suggested readings to required readings. However, in all cases, we will have detailed lecture notes based on these readings.
- When we list movies, we don’t intend to show entire movies, but just small clips

1. Week 1
   a. Introduction to the course, lecture plans
   b. Introduction to agents and multiagent systems
   c. Introduction to basic concepts, e.g. agent state, agent plans
   d. Impact of Science fiction on agents and multiagent systems
   e. Suggested reading: Stan Franklin and Art Graesser “Is it an an agent or just a program: A taxonomy for autonomous agents”, Agents, Theories, Architectures and Languages (ATAL)1996.

2. Week 2: Introduction to “beliefs-desires-intentions”
   a. Asimov short story: “Runaround”
   b. Course notes based on suggested reading
   d. Suggested reading: Some selected sections of Martha Pollack “The uses of plans” Artificial Intelligence, 57(1):43-69

3. Week 3: Introduction to “Agent Modeling”
   a. Asimov short story: “LIAR”
   b. Course notes based on suggested readings
   c. Required reading: Chapter 2 and chapter 3 of the book “Mindblindness” by Simon Baron-Cohen

4. Week 4-5: Teamwork: Multiagent or agent-human
   a. Movies:
      i. Minority report (for Robot-agent-person teams)
      ii. Star trek: BORG
      iii. Star Wars: R2D2/C3PO
   b. Course notes based on suggested readings
c. Required readings (selected sections of):
   i. Searle “Collective Intentions” from “Intentions in communication” (edited by Perrault, Cohen et al) 1990;
   ii. Grosz/Sidner’s article “SharedPlans” from the same book

d. Suggested reading:
   i. Practical application: Schurr et al, IAAI’05 “The DEFACTO System: Training tool for Incident Commanders”
   ii. M Tambe, “Towards flexible teamwork” JAIR 1997, first few sections

e. Experiment with the RoboCup Rescue disaster rescue system

5. Week 7-8: Elementary decision theory and agent interaction under uncertainty
   a. Movie: Star Wars V (one chapter)
   b. Movie: Star Trek movie with interactions under uncertainty
   c. Single decisions under uncertainty; introduction to MDPs (may only be based on course notes)
   d. Course notes based on suggested readings
   e. Suggested reading:
      iii. Etzioni/Weld “First law of robotics” AAAI’94

6. Week 9: Emotions
   a. Movie: Star Trek V or IV (Data gets Emotions)
   b. Course notes based on suggested readings and/or practical applications
   c. Suggested readings:
      ii. Practical application: "Tears and Fears: Modeling emotions and emotional behaviors in synthetic agents," in Proceedings of the 5th International Conference on Autonomous Agents, Montreal, Canada, June 2001

7. Week 10: Social networks
   a. Star Trek: BORG ??
   b. Course notes based on suggested readings
c. Required readings: Chapter 2 or chapter 3 of the book “Linked” by Barabasi

d. Suggested readings: Matthew Gaston and Marie desJardin “Agent organized network for dynamic team formation”, AAMAS’2005

e. Experiment with agent networks (we will locate suitable software)

8. Week 11: Agents and Privacy
   a. Movie: The Final Cut
   b. Course notes based on suggested readings
   c. Suggested readings: ??

9. Week 12-13: Game theory, prisoner’s dilemma, Nash
   a. Course notes based on suggested readings
   b. Students will play prisoner’s dilemma in class
   c. Suggested readings: Kreps, D. "Game theory and economic modeling"
      Chapter 3

10. Week 14: Invited lectures (with help from Mike Zyda)
    a. Social impact of future agents
    b. Discussion

11. Week 15: Wrapup classes, discussions
    a. State of the art in agents and multiagents
    b. How far behind are we compared to the world depicted in sci-fi stories?
    c. What is possible in the next 5 to 10 years in agents and multiagent systems
    d. Suggested reading: Chapters from “Hal’s legacy: 2001’s legacy as dream and reality” David Stork, Arthur C. Clarke (subject to change)

Grading

- Class participation: 10%

- Assignments 1-4 may ask questions in two forms: (i) a direct question from a concept in class, such as finding the Nash equilibria in a payoff matrix or a question from results in Kaminka’s PhD thesis, asking for conditions where/why his failure detection theorem will fail; (ii) a question inspired by a science fiction story (e.g. Asimov’s story “Reason”) to be read and explained using concepts from class, such as BDI agent modeling concepts introduced in week 3. Assignment 5 will be a short paper (4 to 5 pages) focused on a key discussion issue (e.g. in introducing agents in society).
  - Assignment 1: 10%
  - Assignment 2: 10%
  - Assignment 3: 10%
  - Assignment 4: 10%
• Assignment 5: 10%

• Mid term: 20%
• Final exam: 30%

Discussion:

• **Relationship to CS460**: We envision this course to not conflict with CS460 (undergraduate AI class). There are three key points to consider. First, the focus of our course is on concepts that are typically used in agent-based systems, and with particular emphasis on multiagent systems, i.e. interactions of agents with humans, and agents with other agents. There is thus not a significant overlap in the content of our course with CS460, which does not focus on multiagent systems; and indeed, we expect this class to complement and not compete with CS460. Second, CS (CECS) majors are required to take a minimum of two (three) technical electives (TEs). So this course indeed could be a nice companion to CS460 and it does not automatically eject CS460 from a student’s schedule. Third, the goal of this course is to excite students about AI. We expect that students having taken this class will actually enroll in CS460 as their second TE.

• **Readings in Robotics, Game Theory**: We have yet to settle on the right readings for some lectures to introduce robotic applications/issues, and also in the game theory lectures on the right science-fiction short-story or movie to bring out the key concepts and the right technical papers to include.

**Academic Integrity**
The use of unauthorized material, communication with fellow students during an examination, attempting to benefit from the work of another student, and similar behavior that defeats the intent of an examination or other class work is unacceptable to the University. It is often difficult to distinguish between a culpable act and inadvertent behavior resulting from the nervous tension accompanying examinations. When the instructor determines a violation has occurred, appropriate action, as determined by the instructor, will be taken.

All students should read, understand and abide by the University Student Conduct Code [http://www.usc.edu/dept/publications/SCAMPUS/governance/gov03.html](http://www.usc.edu/dept/publications/SCAMPUS/governance/gov03.html)

**Students with Disabilities**
Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to me (or to your TA) as early in the semester as possible. DSP is located in STU 301 and is open 8:30 a.m. - 5:00 p.m., Monday through Friday. The phone number for DSP is (213) 740-0776.