Overcoming Obstacles in Religion-and-Science Dialogue with an Agent-Based Computer Simulation Tool.

Ted Metzler, Amanda Beyers, and John Goulden

Abstract

This paper introduces a novel application of agent-based computer simulation that promises to improve dialogue among a number of scientific and religious communities. Specifically, it concerns a software tool that we believe will help overcome some obstacles to communication among those who are seeking to understand the important phenomenon of altruistic behavior. We currently are prototyping a proof-of-concept version of the tool at Oklahoma City University. To be named THAIST (Theological Artificial Intelligence Simulation Tool), this new *type* of resource for dialogue is expected to be posted as downloadable freeware on our LSI society's website. Other LSI societies, scientists, and theologians will be encouraged to explore—and, hopefully, to expand—its utility as a medium for thinking and communicating across disciplinary boundaries.

Keywords: agent-based simulation, altruistic behavior, artificial intelligence, dialogue, Java 2, user programmable

Biography

Ted Metzler's *B.A. degree in mathematics (philosophy minor) was followed by an M.S. in computer science and an M.A. and Ph.D. in philosophy. The uniting motivation in this educational path was his fascination with relations between traditional notions of the human person and new alternatives introduced by artificial intelligence (AI) and robotics. Appropriately, most of Ted's work experience in computer software development has involved AI applications. Several years ago—sensing a personal calling to help mediate dialogue between the communities of AI and religion—he returned to school and earned an M.A. in theology. Ted currently is an adjunct professor in religion at Oklahoma City University and serves as program coordinator for its new LSI program.*

Amanda Beyers *appears also to have embarked upon a multidisciplinary educational path. Although her B.A. degree from Oklahoma City University combined a French major with a philosophy minor, she chose the university's computer science department for her graduate program. Currently pursuing the M.S. degree in that discipline, Amanda is focusing upon software engineering involving games. The project described in our paper already is consistent with this focus, and it ultimately could become pertinent to Amanda's computer-game hobby.*

John Goulden *has earned degrees in varied disciplines, as well, holding a B.A. in music, B.S. in physics, M.A.s in physics and computer science, and a Ph.D. in physics. He has taught in all three areas and is currently chair of the Department of Computer Science at Oklahoma City University. His interests include computability, algorithm design, programming, and any issues that span science and religion.*

Introduction

Theologian Wesley J. Wildman, in recent observations concerning the multidisciplinary Divine Action Project[1] (DAP), reports "Most people involved in the DAP had been involved in science-religion dialogues enough to know how perilous any work plan would be" (33). He proceeds immediately to identify a number of "challenges" that probably all of us can recognize as characteristic obstacles in such dialogues: "theologians typically do not know enough science," "scientists typically do not appreciate the nuances of theology well enough," "scientists typically are expert in only one area of science," etc. Indeed, in an age that notably rewards specialization we are likely to encounter difficulty with nearly any dialogue that attempts to link distinct "areas," whether these represent religion and science, separate sciences, or even just contrasting theologies. Each "area," after all, understandably tends to grow its own distinctive vocabulary of concepts and terms, as well as its own web of theory that connects those concepts with a limited set of relations. Worse yet, distinct sciences and theologies—working with their diverse vocabularies and theories—occasionally find themselves studying the same phenomenon. One important phenomenon of this kind appears to be that of altruistic behavior. Although dialogue of a sort has arisen among the specialists in this case, it generally has lacked a "common meeting ground." That is to say, the diverse parties in the dialogue encounter the (not uncommon) obstacle of lacking a sharable and readily accessible medium that fosters and supports both expression and experimental comparison of their different vocabularies and theories.

With regard to this altruism dialogue, we invite the reader to consider a few *what ifs*. What if a technological resource that has emerged from computer science and artificial intelligence in recent times (viz., agent-based simulation) could be adapted to offer a flexible medium for *expressing* the diverse vocabularies of concepts and terms that we find in this dialogue? What if the same technology enabled instructive experiments with the dynamic *interactions* of these elements by means of "social simulations" involving artificially intelligent software agents? What if software incorporating this technology were already a relatively familiar tool within some of the sciences that have addressed altruistic behavior? What if extensions of such tools plausibly could comprise a shared medium for precise communication, allowing *rapid worldwide comparisons* of concepts and theories representing diverse perspectives on altruistic behavior?

Under these rhetorically hypothetical conditions, would it not be helpful actually to *build* a user programmable agent-based simulation tool to accommodate the multiple scientific and religious communities who are interested in examining altruistic behavior? We think so. Accordingly, we shall now present a fuller account of our rationale for undertaking an experimental development of just that sort, and report our current progress with the project that is doing it.

The Case for this Application of Agent-Based Computer Simulation

As the sub-field of computer science known as *artificial intelligence* (AI) has increased its ability to mimic intelligent behavior of biological creatures (especially, humans) it has generated a special area of research called *distributed* artificial intelligence (DAI).

Within DAI research, investigation of *multi-agent systems* (MAS) has supported development of computer simulation tools that "distribute" capability for intelligent behavior among populations of relatively autonomous, but interactive software *agents*.

In broad description, *agents* appearing in *social simulations* simply are software models of intelligent creatures that interact with simulated environments containing active models of other intelligent creatures. The complexity of the agents that one finds in different social simulation systems can vary substantially. In the simpler cases, one encounters *fine grained agents* which, as DAI researcher Nicholas Avouris observes, are agents "with reactive behaviour, i.e. [,] agents with no complex reasoning capabilities, not owning representations of themselves, other agents or the environment in which they exist" (145). Agents of this type have been used, for example, to simulate behavior within populations of rather simple creatures such as ants (Avouris 146). In other systems, though, one may find so-called *coarse grained* agents with considerably more impressive capabilities. Petra Funk and Jurgen Lind, in their contribution to "Socially Intelligent Agents," a Technical Report of the American Association for Artificial Intelligence (AAAI), offer definitions of coarse grained agents that they regard as "more appropriate agents in a multi-agent system" (47). The agents they have in mind can learn, deliberate, behave proactively and reactively, possess knowledge about other agents, cooperate with other agents, act autonomously in behalf of other agents, communicate with other agents, and incorporate some form of self awareness (47). Workers in the MAS field often treat the incorporation of "self awareness" as an agent's own possession and/or construction of what is called a *self model*. Similarly, the agent's knowledge about other agents is said to be represented in the *acquaintance* *models* that it builds as it becomes “acquainted” with other agents during a simulation (Jennings 32).

Not surprisingly, the capability to set up and execute computer simulations of behavior in complex “societies” of agents has attracted interest in a number of different research areas. A representative online bibliography of work in this category includes entries, for example, from computer science, biology, politics, law, economics, psychology, sociology, and anthropology (Axelrod and D’Ambrosio). Of course, any simulated *social* behavior among artificially intelligent agents must also presume some specification of interactions that the agents may have with each other. One reasonably standardized model of agent interaction that has been of special interest to researchers concerned with altruistic behavior was popularized by Robert Axelrod’s *The Evolution of Cooperation* about two decades ago. This model has come to be known generically as the “Prisoner’s Dilemma” game, reflecting the originally imagined social situation that inspired it. We shall now briefly sketch an illustration of that situation to explain the dynamics of this “game.”

Two current prisoners at the local jail—who have been apprehended by police as suspected partners in commission of a recent bank robbery—actually *were* partners in the robbery. To avoid accusations of sexist presumptions, let us name them Bonnie and Clyde. In advance of the robbery, Bonnie and Clyde agreed that if either were apprehended he or she would not give authorities any information that might harm the other. The District Attorney, it seems, has enough evidence to convict either prisoner on a “weapons” charge (indeed, each was caught carrying a virtual arsenal!). Alas, it is also the case that the D.A.'s only hope of getting a *bank robbery* conviction is the familiar strategy of persuading one of the prisoners to turn state's evidence in return for promise of a lighter sentence. In fact, the D.A. offers a suspended sentence (no jail time at all!) to each prisoner—if he or she will selfishly abandon the alleged "partner" (who will then get *ten years* in the slammer). In the case that both prisoners accept this deal, they will both be rewarded for their cooperation with a somewhat lighter sentence (five years). Surprisingly, though, if Bonnie and Clyde both stay true to their original agreement, they will each get only a sentence of *one* year (we surmise that not all of their arsenal was illegal). The dynamic for Bonnie and Clyde's "Prisoner's Dilemma" is summarized in the following grid (wherein "selfish" abbreviates "turning state's evidence," "cooperative" means "upholding the prior agreement among the thieves," and resulting time in jail for each prisoner is shown—in years—for each possible combination of their choices):

Bonnie = cooperative	*Bonnie = selfish*	Years of jail time shown for each combination of choices
Bonnie: 10, Clyde: 0	Bonnie: 5, Clyde: 5	*Clyde = selfish*
Bonnie: 1, Clyde: 1	Bonnie: 0, Clyde: 10	*Clyde = cooperative*

Note that each "partner" can minimize his (or her) *own* jail time by choosing the selfish option, while their best choice *as a team* is mutual cooperation.

This fairly simple game affords rich possibilities for illuminating agent-based social simulations because (1) externally observed *behavior* of a live *human* agent does not reliably identify his or her internally experienced *intentions*, but (2) coarse grained *artificially intelligent* agents, in social simulations, allow the experimenter to model both externally observed behavior *and* internally experienced intentions. Indeed, modeling Bonnie and Clyde with such agents can permit exploration of interactions between their objective choices and their hypothesized cognitive structures and motivations. Suppose, for example, we have equipped both of the software agents representing Bonnie and Clyde with self models, acquaintance models (of each other, at least), and with representations of goals, strategies, and values *such that* their individual choices will be forever dominated by (hypothesized) *self-interest*. Let us further suppose that the simulation will permit *repeated* arrests of the same two agents, and that the agents also can *learn* in the course of a simulation (i.e., conditions for what is commonly called the "*Iterated* Prisoner's Dilemma" game). Under these set-up conditions for a simulation experiment it is at least conceivable that the agents might both come, as the simulation progresses, to display a growing preference for the "cooperation" choice. It is important to notice, however, that the same *behavioral outcome* of the simulation might conceivably result *as well* if we have set up the experiment with substantially *different* hypotheses regarding the agents. We might hypothesize the presence, say, of influence from a "divine" agent, who progressively leads both human agents to find the "cooperation" choice more *rewarding—regardless* of the respective "payoffs," in terms of years in jail. With provision of suitable "graphical user interface" resources, the tool we are proposing could be made sufficiently "user programmable" to set up simulations representing either of the foregoing approaches.

If the reader is familiar with agent-based simulation research in such disciplines as economics and sociobiology, the first experiment just described most likely will sound familiar. However, the second approach we have sketched probably strikes most readers as somewhat bizarre. This is not the case for any *scientific* reason—it simply happens to be an artifact of the tacit *theories* about human beings and their place in the cosmos that have generally characterized the research communities *using* such computer simulation methods … so far.

Theoretical and metaphysical assumptions that accompany the use of computer simulation should always be acknowledged, but they are *particularly* significant for the class of simulations that we are addressing in this paper—i.e., in the case of social simulations employing artificially intelligent agents. When a computer simulation applies equations from classical mechanics and/or relativity theory to study of gravitational interactions among celestial bodies it is drawing upon mathematical laws whose margins of predictive error have been fully explored. A computer simulation that tests theories regarding the decisions of stock traders by modeling their hypothesized cognitive strategies with artificially intelligent agents is a much different *kind* of simulation. The latter kind of simulation is less a *predictive calculating machine* than it is a *heuristic aid to thinking*—a medium for rapid exploration of complex interactions that the researcher believes *might* be occurring. Accordingly, if a sociobiologist feels comfortable about using simulations of this type to express and explore his or her hypothesis that cooperative behavior could emerge among totally self-interested human agents, then theologians who wish to express and explore, say, Whiteheadian metaphysics and the effects of "initial aims" upon human agents *need not feel uncomfortable about using the same kinds of tools*. The vocabularies and theories in the two cases obviously will differ—but neither thinker has exclusive rights to draw upon the technology of agent-based simulation as an aid for expressing and exploring theoretical understandings of observable phenomena.

Similarly, we recognize no need for thinking people in the religious *or* scientific communities to abandon the word “altruism” to those research specialties who wish persuasively to redefine it as a form of self-interest—and specialties of this kind do exist. Theologian Colin Grant, for example, reminds us that it was artificial intelligence pioneer Herbert Simon who asserted "activities that aid others are so foreign to the foundational predilection to self-interest that they can only be attributed to docility and stupidity” (72). Again, philosopher Joseph Des Jardins characterizes the basic view of human agency in economics by saying “human beings act, primarily if not solely, on the basis of self-interest” (53). Within the scientific community itself, biologist Jeffrey Schloss complains that “sociobiology remains committed to seeing altruism as self-interest by another name” (248). Commitments of this sort, wherever they appear, have every right to draw upon agent-based simulation for their expression; we argue only that *other* theoretical commitments may legitimately do the same.

Some of those alternative commitments might even restore luster to the meaning assigned to the term “altruism” by pioneer sociologist August Comte when he *coined* it in the nineteenth century. The 1971 Unabridged Edition of *The Random House Dictionary of the English Language* defines “altruism” as “the principle or practice of unselfish concern for or devotion to the welfare of others.” This seems not to have strayed far from the meaning with which the word was coined—according to psychologists Samuel and Pearl Oliner, “Comte conceived of altruism as devotion to the welfare of others, based in selflessness" (4). Normally, the word is contrasted with its opposite notion, "egoism," which the economist Henry Hazlitt has defined concisely as "the pursuit of personal ends at the cost of those of others" (par. 4).

Upon reflection, these understandings of "altruism" and "egoism" seem to us to mark out reasonably well the basic behavioral choices facing Bonnie and Clyde. Nor are we alone in this semantic judgment. Contemporary computer scientists at the University of Massachusetts and University College London, for instance, report results of social simulations that involve groups of "*altruistic agents*" and "*egoistic agents*" who interact in terms of the Iterated Prisoner's Dilemma (Bazzan, et al. 5). Moreover, it should be possible (and perhaps fruitful) even for *theologians* to use the Iterated Prisoner's Dilemma game in social simulations—although they *might not always choose* self-interested practical rewards to *individuals* as "payoff functions" that any "rational" human agents should seek to maximize at all times.

In fairness, for that matter, we should acknowledge that even members of the *scientific* community have begun to consider alternatives to Egoistic-Tooth-and-Claw-Survival as viable strategies for agent-based simulations. Consider, for example, the following Abstract comments by L. M. Hogg and N. R. Jennings in a 1997 article titled (perspicuously, we believe) "*Socially* Rational Agents" [emphasis added]:

> Autonomous agents are designed to carry out problem solving actions in order to achieve given, or self generated, goals. A central aspect of this design is the agent's decision making function which determines the right actions to perform in a given situation to best achieve the agent's objectives. Traditionally, this function has been solipsistic in nature and based upon the principle of individual

> utility maximisation. However, we believe that when designing multi-agent systems this may not be the most appropriate choice. Rather we advocate a more social view of rationality which strikes a balance between the needs of the individual and those of the overall system. (61)

Does it not seem that work of this kind strongly invites dialogue with theologians? Even if theologians were not motivated to devise experiments of their own with the tool we have suggested, they could arguably find that familiarity with its methods would make potentially interesting work of the kind being reported by Hogg and Jennings more accessible. Moreover, we live in an age in which the proposed tool could be made globally portable on the World Wide Web, offering opportunities for rapid exchange of simulation ideas and results among *all* of its users. Let us, then, turn to an account of our initial steps at Oklahoma City University (OCU) toward actually *building* the proposed "user programmable agent-based simulation tool."

Current Progress with the Project

The first milestone in our current prototyping project at OCU will be completion of a proof-of-concept version of the new agent-based simulation tool for which our foregoing remarks have "argued the case." Recognizing that innovations tend to die from neglect if they cannot be associated from the beginning with a simple name, we christen the intended tool "THAIST." Abbreviating "Theological Artificial Intelligence Simulation Tool," the name deliberately puns on "theist" to help make it memorable—and to represent our hope that the tool can eventually become a vehicle for improved communication between communities of religion and science.

Our Local Societies Initiative group at OCU (GOOD STAR, Mark Y. A. Davies, Chairman) is not financially supporting the THAIST prototyping, although it has sound reasons for interest in the project's outcome. First, we currently are planning a GOOD STAR event that will be focused upon the *topic* of altruistic behavior. Biologist Jeffrey Schloss has been invited to keynote this event in October of 2004, and we should obviously be pleased if THAIST has reached a stage of development that will allow it to be demonstrated at that time. Second, we hope to take advantage of the fact that the THAIST tool is being written in the Java 2 programming language. We intend to exploit this feature by making THAIST available for downloading as freeware from our GOOD STAR website, and shall certainly encourage widespread experimentation with the new tool among other LSI groups, scientists, and theologians.

Our prototyping team is not large, but we believe it is qualified for its tasks. Ted Metzler has formal education and work experience in the field of artificial intelligence, some of the latter involving use and development of simulation applications pertinent to THAIST (Heineken, et al; Ortiz and Metzler; Metzler and Ortiz; Metzler and Nordyke). He also formulated functional requirements for a tool of this type in his 2002 theology thesis. Amanda Beyers is currently pursuing her M.S. degree in computer science at OCU, and is focusing upon software engineering involving games. Clearly, our use of the classical Iterated Prisoner's Dilemma in THAIST already makes this an appropriate and interesting graduate project for Amanda. She is being advised on the project by the Chairman of the OCU Computer Science Department, John Goulden, who happens also to hold a professional degree in physics. John Starkey, from OCU's Wimberly

School of Religion, is our team's "resident theologian," who will satisfy the important developmental need for a "test user" with expertise in that area. Our methodology will make use of progressive build-test-revise cycles, and will emphasize user-centered design of THAIST's graphical user interface (GUI).

Prototyping has begun with the basic case of *human* agents interacting through Iterated Prisoner's Dilemma (IPD) encounters, although functional requirements for classes of agents representing *deity* and environmental systems of *nature* will be addressed later. These latter developments are expected to mark the extensions that will distinguish THAIST from social simulation resources that are already available and allow it to begin supporting dialog between religion and science.

Principal features of the tool's GUI, at this point, are illustrated in Figure 1 and Figure 2. We are initially working with fairly simple textual displays while we gain experience with various computational options for implementing IPD encounters among human agents; graphical displays will be added later. Figure 1 is focused upon specifications that the user of the tool may input to "set up" conditions for a simulation execution (hereafter, just "simulation"), while Figure 2 shows reported results of a simulation.

The "Applet Viewer" window displayed at the top of Figure 1 appears after the user has specified, in an opening window, the number of agents to be used in the simulation (in this case, three). The pull-down choices shown for Agent 1 determine the strategy by which that agent will play the IPD game during the simulation. The darkened choice, "Tit for Tat," refers to a *strategy*, often exercised in IPD simulations, by which the agent always cooperates when it encounters a new “acquaintance” – but repeats the choice *selected by the acquaintance* the next time that agent is encountered. Obviously, this strategy requires that our agents already are capable of building acquaintance models

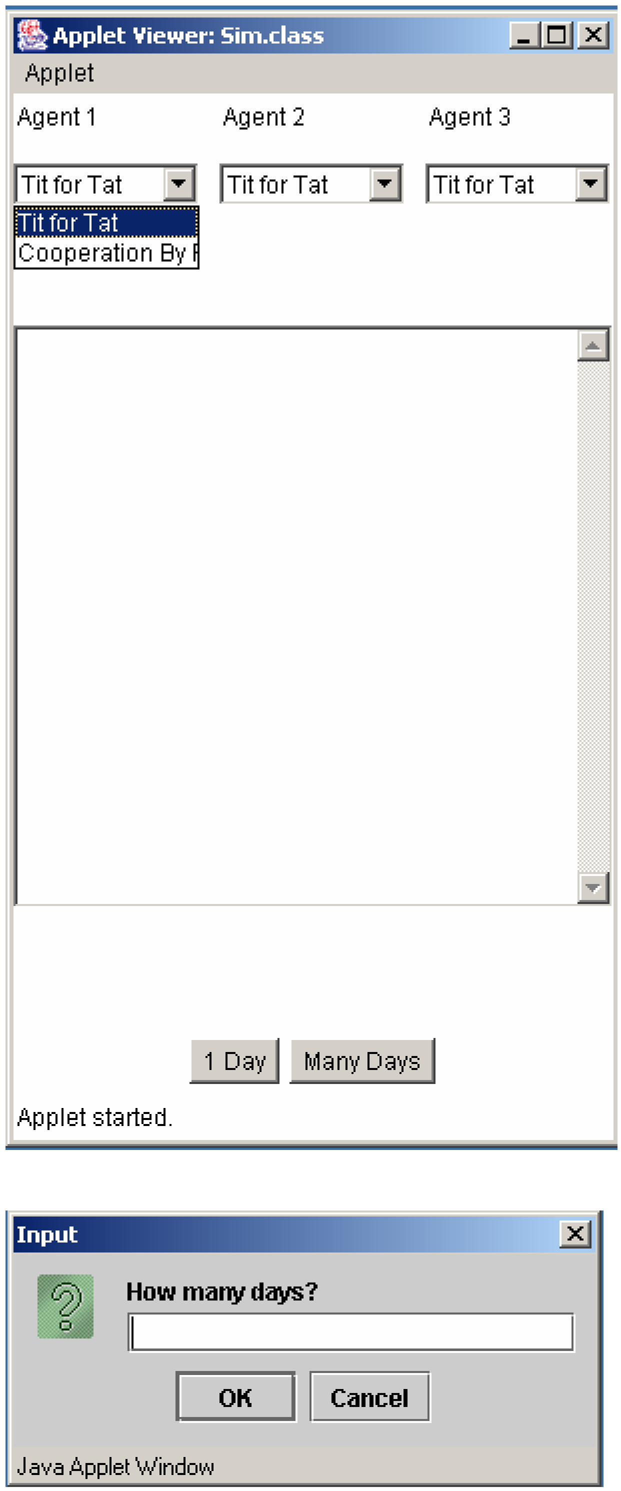

Figure 1.

of each other during simulations. The other strategy choice (partly occluded in the figure) is “Cooperation By Probability.” In this case, the user may assign an initial probability of cooperating to the agent, although the agent will dynamically adjust this probability as the simulation unfolds in response to choices made by other agents that it encounters. Finally, the user may choose either to step through simulations one-day-at-a-time or to select the “Many Days” button, in which case the bottom window shown in Figure 1 allows the user to specify the duration of the simulation.

In each simulation “day,” all mathematically possible combinations of (two) agents currently engage in the IPD game, as indicated by the results displayed in Figure 2. The “scores” shown in this figure for each agent currently follow a convention in the literature of reporting some measure of the individual agent’s practical *gain* from the IPD encounter. Accordingly, low numbers (of years of jail time) in our prior Bonnie-and-Clyde illustration are represented by high “scores” in the reports shown.

Although our simulation tool has not yet begun to exhibit the distinctive features that we eventually intend THAIST to have, it is already capable of facilitating the set up and execution of agent-based social simulations that use standard IPD interaction. To this extent, therefore, it is already likely to be recognizable and to “make sense” as a web-portable software tool within scientific communities who are familiar with its methods. Indeed, the novel challenge of extending it in ways that will also “make sense” to our team’s “test user” in the Wimberly School of Religion still lies ahead of us. We welcome the challenge, for we believe our experiment has sound potential for improving dialogue between religion and science.

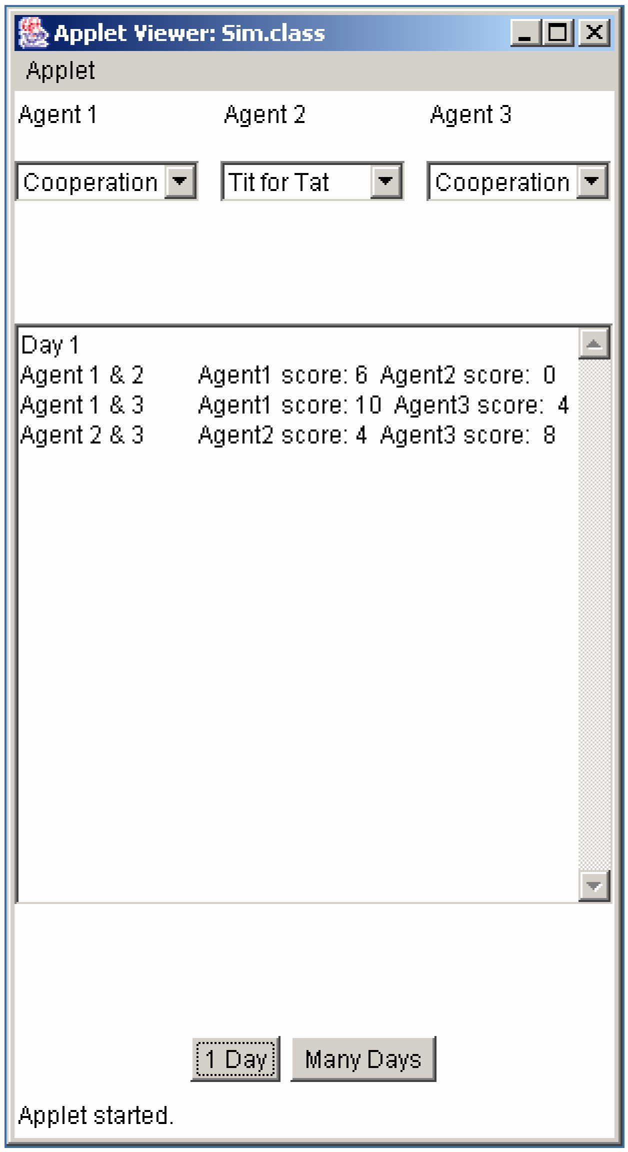

Figure 2.

Simulation Results Report.

Notes

[1] The Divine Action Project may be described as a sustained multi-year effort to formulate accounts of divine action in this world that both theologians and scientists can find intelligible.

Works Cited

"Altruism." Def. *The Random House Dictionary of the English Language*. The Unabridged Edition. 1971.

Avouris, N. M. "User Interface Design for DAI Applications: An Overview." *Distributed Artificial Intelligence: Theory and Praxis*. Ed. Nicholas M. Avouris and Les Gasser. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1992. 141-162.

Axelrod, Robert. *The Evolution of Cooperation*. New York: Basic Books, 1984.

Axelrod, Robert, and Lisa D'Ambrosio. *Annotated Bibliography on The Evolution of Cooperation*. 12 April 2004 <http://www.lifl.fr/IPD/references/from_others/Evol_of_Coop_Bibliography.html>

Bazzan, Ana L. C., Rafael H. Bordini, and John A. Campbell. "Agents with Moral Sentiments in an Iterated Prisoner's Dilemma Exercise." *Socially Intelligent Agents: Papers from the 1997 AAAI Fall Symposium. Technical Report FS-97-02*. Menlo Park, CA: AAAI Press, 1997. 4-6.

Des Jardins, Joseph R. *Environmental Ethics: An Introduction to Environmental Philosophy*. Belmont: Wadsworth/Thomson Learning, 2001.

Funk, Petra, and Jurgen Lind. "What is a Friendly Agent?" *Socially Intelligent Agents: Papers from the 1997 AAAI Fall Symposium. Technical Report FS-97-02*. Menlo Park, CA: AAAI Press, 1997. 46-48.

Grant, Colin. *Altruism and Christian Ethics*. Cambridge: Cambridge University Press, 2001.

Hazlitt, Henry. Chapter 13. *The Foundations of Morality*. Los Angeles: Nash

Publishing, 1972. 1 Oct. 2001 <http://www.hazlitt.org/e-texts/morality/>

Heineken, Carolyn, Theodore Metzler, Rafael Ortiz, and Ray Wonsewitz. "Tuning a Fuzzy Expert System for Simulations of Population Behavior." *Proceedings of the IASTED International Conference on Modeling and Simulation*. ACTA Press, 1999. 317-321.

Hogg, L. M., and N. R. Jennings. "Socially Rational Agents." *Socially Intelligent Agents: Papers from the 1997 AAAI Fall Symposium. Technical Report FS-97-02.* Menlo Park, CA: AAAI Press, 1997. 61-63.

Jennings, Nick. *Cooperation in Industrial Multi-Agent Systems*. Singapore: World Scientific Publishing Co. Pte. Ltd., 1994.

Metzler, Theodore A. "Experimental Theology, Using Artificially Intelligent Agents." Thesis. Andover Newton Theological School, 2002.

Metzler, Theodore, and John Nordyke. "Use of ModSAF in Development of an Automated Training Analysis and Feedback System." *Proceedings of the 6th Conference on Computer Generated Forces and Behavioral Representation.* Orlando: Institute for Simulation and Training, 1996. 151-155.

Metzler, Theodore, and Rafael Ortiz. "Modeling Population Behavior with Fuzzy Logic Agents." *Smart Engineering Systems: Neural Networks, Fuzzy Logic, Data Mining, and Evolutionary Programming.* Ed. Cihan H. Dagle, et al. New York: ASME Press, 1997. 273-278.

Ortiz, Rafael, and Theodore Metzler. "Planning US Border Patrol Staffing with a Modeling and Simulation Tool." *Modeling and Simulation: Proceedings of the IASTED International Conference*. Ed. H. M. Hamza. Anaheim: IASTED/ ACTA Press, 1998. 312-316.

Wildman, Wesley J. “The Divine Action Project, 1988-2003.” *Theology and Science* 2.1 (2004): 31-75.