Simulating Human Behaviors in Agent Societies (Short Paper)

Alicia Ruvinsky
University of South Carolina
Computer Science & Engineering
Columbia, SC 29208 USA
+1-803-777-3768
ruvinsky@engr.sc.edu

Michael N. Huhns
University of South Carolina
Computer Science & Engineering
Columbia, SC 29208 USA
+1-803-777-5921
huhns@sc.edu

ABSTRACT
As increasing numbers of processors and agents pervade the human environment, societies comprising both humans and agents will emerge. Presently, it is unknown how a person might fair in such mixed societies. For the societies to operate effectively and efficiently, it is important for the humans and agents to recognize and understand each other's behavior. This paper provides an initial step in that understanding via two contributions: (1) we provide models, within a limited domain, for agents that behave like humans and (2) we present the results of simulated interactions between the human-like agents and a variety of purely rational agents. Our models for the behaviors of people are based on recent sociological research by Simpson and Willer [10] that explores humans' cooperative prosocial behavior, a conceivably non-rational process. Modeling human behaviors presents a means of exploring and understanding motivations, consequences, and resolutions to human-agent interactions. We aspire to exploit measured human behavior in order to observe its ramifications in an agent world, and to motivate development of human-agent societies. Our results show that, although there are pitfalls to which humans are vulnerable, there exist niches for human prosperity in a rational agent world.

Categories and Subject Descriptors
J.4 [**Social and Behavioral Sciences**]: *Sociology*

General Terms
Experimentation, Human Factors

Keywords
Societal Aspects of MAS, Artificial Social Systems, People-agent Societies, Models of Social Behavior

1. INTRODUCTION
A sociologist's job is to observe and understand human societies. In designing multiagent systems, computer scientists use understanding of rational agent behavior to motivate agent societies. The primary distinction between these fields is in how each defines its contributing agent, one organic and the other computational. As interactions between agents and people begin to define a new mixed society, the boundary between sociology and computer science blurs. This work explores the idea of using the results of sociological research to model human behavior in a multiagent society. We aspire to exploit this knowledge to motivate the development of human-agent societies.

We have based our experimentation on the sociological research performed by Simpson and Willer [10] that explores the motivation for cooperative prosocial behavior in people, a conceivably non-rational process. Our simulation models agents based on the statistical findings of [10] in order to understand how these agents contend against purely rational agent models.

Scientists use computer simulations to garner new insights into developmental trends in societies. Our research simulates human behavior through probabilistic modeling based on results of sociological investigations into human interactions in an attempt to understand the developmental trends in a society comprised of rational agents and non-rational humans. The hypothesis is that humans will be at a disadvantage when interacting/competing with rational agents in a computational environment. However, our results reveal that this is not always the case. There are pitfalls to which humans are vulnerable, yet there exist niches for human prosperity in the rational agent world.

Section 2 orients this work in existing work in multiagent systems and other relevant disciplines. Proceeding this is a brief description of the sociological research used to model the human behavior herein [10]. Section 4 describes the rational agent types competing against our human modeled agents. An explanation of the design and execution of our experiments is reported in section 5, including experimental results. The implications and limitations of our experiments and results are found in section 6. Conclusions and future potential of this work are presented in sections 7 and 8, respectively. The paper ends with acknowledgements and references.

2. BACKGROUND
The work herein is motivated by existing work spanning agent reciprocity, computational models of human behavior, aspects of human society, and simulations of agent societies.

Sen [9] provides significant work in reciprocity among varieties of agent types. Our simulated agents described in section 4 were strongly motivated by agent types described in [9].

In [5], Chattoe recommends basing agent system design on information about human societies. Attempts to build socially inspired agents, though not mimicking humans faithfully, include Castelfranchi et al. [4] in which human-like norms were simulated to motivate and manage agent behavior.

Game theoretical abstractions to human behaviors have been investigated in exploring the evolution of cooperation. This work

Cite as: Simulating Human Behaviors in Agent Societies (Short Paper), Alicia Ruvinsky, Michael N. Huhns, *Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, Padgham, Parkes, Müller and Parsons (eds.), May, 12-16., 2008, Estoril, Portugal, pp. 1513-1516.

is grounded in Axelrod's prisoner's dilemma competition [1]. Along these lines, Bazzan et al. [2] investigated the effects of altruism among agents playing the Iterated Prisoner's Dilemma.

3. MODELING HUMAN BEHAVIOR

The models for the human-like agents in this work are based on the statistical results of Simpson and Willer [10]. In their research, they presume a heterogeneous society in which humans are characterized by their social preference of altruism or egoism and by the situation motivating the prosocial behavior, either public or private. There are four human characterizations distinguished in [10], namely altruists acting privately, altruists acting publicly, egoists acting privately, and egoists acting publicly.

We similarly model four human-like agent behaviors characterized in this same way. The human-like agents behaviors are derived from the results of experiments conducted in [10] on human participants. The experiments and results used to develop our models are discussed in this section.

3.1 The Dictator Game

The two-player dictator game consists of a dictator and a receiver. The dictator agent is given a set of resources for which it must choose an amount to donate to a passive receiver agent. It gives this amount to the receiver and keeps the remainder for itself.

In [10], the dictator game is played by human participants. The data gathered is reported as the mean proportion of resources donated by participants who are characterized by social preference and situational context. The results are as follows:

- Altruists in a private situation donated a mean of 40%;
- Altruists in a public situation donated a mean of 51%;
- Egoists in a private situation donated a mean of 22%;
- Egoists in a public situation donated a mean of 46%.

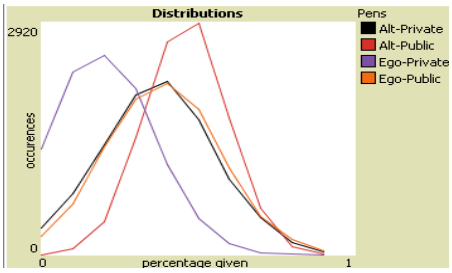

Figure 1. Distributions of donations given by various prosocial humans X situation relationships used in simulating human agents in our NetLogo model.

The results of the experiment were used to generate statistical distributions for simulating human-like behaviors characterized in the same way as the human participants' behaviors (see Figure 1).

The accuracy of our distributions in simulating the experimental results reported in [10], are expressed by comparing the standard deviation of the average percent donated by the simulated human-like agents compared to the mean proportions donated by the real human participants. Small differences in standard deviation signify correspondence between our simulation and the real world results (see table 1). Consequently, our simulated humans are an accurate model of the real humans for this domain.

Table 1: The standard deviation between the donation percentage amounts identified by the models of human-like behavior and the actual human behaviors they are mimicking.

Human Altruist Private	Human Altruist Public	Human Egoist Private	Human Egoist Public
0.02	0.007	0.05	0.02

3.2 The Indirect Reciprocity Game

The two-player indirect reciprocity game consists of a dictator and an indirect reciprocator. The game begins with the premise that a dictator game has already occurred. An independent member of the society (the indirect reciprocator) is then asked to indirectly reciprocate the original dictator's behavior from the dictator game. The indirect reciprocator is given a set of resource units, and then told the percentage of the original dictator's resources that were donated to the receiver in the dictator game, as well as situation of the donation (public or private). The indirect reciprocator decides how much of its resources to reciprocate to the dictator.

The indirect reciprocator game was performed as an experiment on human participants [20]. The data gathered is reported as the mean proportion of resources donated by the indirect reciprocator participants.

- An altruist indirectly reciprocating to a dictator that donated in private would reciprocate with an equal proportion of its resources. If the dictator gives 50% of his resources, then the indirect reciprocator gives 50% of his resources.
- An altruist indirectly reciprocating to a dictator that donated in public would match 90% of the percentage the dictator donated to the receiver. If the dictator gives 50%, then the indirect reciprocator gives 45%.
- An egoist indirectly reciprocating to a dictator that donated in private would match 86% of the percentage the dictator donated to the receiver. If the dictator gives 50%, then the indirect reciprocator gives 43%.
- An egoist indirectly reciprocating to a dictator that donated in public would match 64% of the percentage the dictator donated to the receiver. If the dictator gives 50%, then the indirect reciprocator gives 32%.

4. RATIONAL AGENTS

The rational agents in this simulation were inspired by those in [9] where Sen defines four agent types: philanthropic, selfish, reciprocative, and individual. We adopted three of these agent types for the dictator and indirect reciprocator domains: philanthropic, selfish, and reciprocative. The fourth agent type, individual, could not be translated appropriately into our domain.

The philanthropic agent is a perpetually cooperative agent. It will always donate 50% of its resources.

The selfish agent accepts any donations made by others, but never donating anything.

The reciprocative agent assesses its indebtedness to another agent in its consideration of how much to donate to that agent. This agent will periodically contribute to an agent to which it is not indebted. [9]

5. EXPERIMENTATION AND RESULTS

5.1 Experiment

There are 4 human-like agents and 3 rational agents. The human-like agents are human altruists acting privately (HAPr), human altruists acting publicly (HAPu), human egoists acting privately (HEPr), and human egoists acting publicly (HEPu). The rational agent types are philanthropic (P), selfish (S), and reciprocative (R). Each of these agent types are pitted against each other in both the dictator and indirect reciprocity games.

5.1.1 Implementing the Dictator Game

Agents of each contending agent type are paired. Dictatorship is then randomly assigned. The dictator is given 8 resource units of which it decides how much to give to the receiver. The transfer of resource units is made. Agents now swap roles so that the receiver becomes the dictator, and the game is played again, ensuring an equal representation of dictator agents from both agent types.

5.1.2 Implementing the Indirect Reciprocator Game

Agents of each contending agent type are paired. Indirect reciprocatorship is then randomly assigned. The other agent becomes the dictator. The dictator fabricates a round of the dictator game to produce the amount of resources that the dictator would give to the receiver, but no resources are actually disseminated in this step. The indirect reciprocator receives 9 resource units of which it decides how much to reciprocate to the dictator. The indirect reciprocator gives this amount to the dictator, keeping the remainder for itself. The indirect reciprocator and the dictator swap roles, and the game is played again.

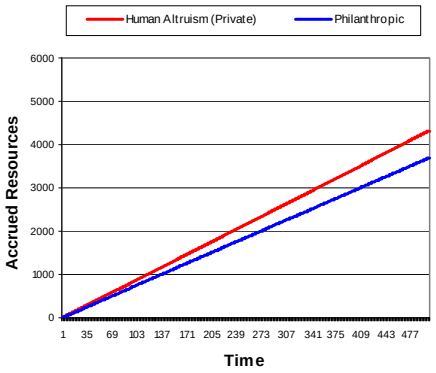

Figure 3. An example of the results obtained from our NetLogo simulation when a simulated human altruist acting privately competes against a rational philanthropic agent.

5.2 Results

The amount of resources acquired during each simulation iteration is averaged for each competing agent type and then accrued over many iterations (see figure 3). The rates at which an agent type accumulates resources as compared to its competitor are calculated. The differences between the rates of competing agent types will serve as a metric for characterizing the relative success of one agent type over another. The results for each simulated competition between agent types are shown as the difference between the accumulated resources of one agent type versus the accumulated resources of the contending agent type. These values are identified for all agent type pairs in both the dictator and indirect reciprocator games (see tables 2 and 3).

Table 2. Results for dictator game: Differences in rates at which acquired resources accumulate for contending agents. (Values calculated as rate of change of row agent's resources minus rate of change of column agent's resources.)

	HAPr	HAPu	HEPr	HEPu	P	S	R
HAPr	-0.05	1.24	-2.49	0.14	1.24	-6.61	-0.39
HAPu	-1.41	0.01	-3.83	-1.13	-0.09	-8.01	-0.73
HEPr	2.46	3.81	-0.05	2.65	3.85	-4.19	-0.09
HEPu	-0.19	1.10	-2.69	-0.05	1.13	-6.91	-0.43
P	-1.25	0.11	-3.79	-0.97	0	-7.99	-0.49
S	6.69	8.11	4.23	6.94	7.99	0	0.69
R	0.41	0.70	0.09	0.43	0.49	-0.71	0

Table 3. Results for indirect reciprocity game: Differences in rates at which acquired resources accumulate for contending agents. (Values are with respect to the agent type specified by the row label.)

	HAPr	HAPu	HEPr	HEPu	P	S	R
HAPr	0.03	-0.89	1.73	-2.59	0	0	0.03
HAPu	1.95	0.05	3.28	-1.15	0.80	0	0.01
HEPr	-1.57	-3.21	-0.03	-3.35	1.11	0	0.01
HEPu	2.62	1.09	3.34	0	2.87	0	0.11
P	0	-0.80	-1.11	-2.87	0	-7.99	-0.49
S	0	0	0	0	7.99	0	1.18
R	-0.03	-0.01	-0.03	-0.11	0.49	-1.17	-0.01

6. ANALYSIS

The difference in the rates at which resources are accumulated between competing agent types is used to assess how the two agent types interact. There are two possible interactions indicated by differences in rates of accumulated wealth.

(1) Difference is zero. This means that both agent types are gaining and losing resources at the same rate. Neither agent type is benefitting over the other.

(2) Difference is not zero. The playing field is not equal between these two agent types. One of the agents is making larger donations to its opponent than it is receiving from its opponent. Such an agent has a greater prosocial tendency. The opponent, on the other hand, is exploiting the agent's prosociality.

6.1 General Observations

Both result tables 2 and 3 exhibit inverse symmetry. This implies consistent performance of the competing social agent types and indicates correctness of the simulation results.

The diagonal of both tables is nearly zero. The diagonal represents simulations in which an agent type is competing against itself, and is unable to develop a prosocial or exploitative advantage.

The fear of social ramifications that incites the human-like agents to behave as they do in the dictator game becomes the agents'

right to reciprocate in the indirect reciprocator game in a manner that cultivates fear of social ramifications [8]. The interplay between the agent applying the social pressure and the one receiving this pressure has potential for a social learning interaction [2].

6.2 Trends in Human-Like Agents

Based on the observed data, how does a human-like agent fare against rational agents? Here we analyze the ramifications of human-like behaviors with respect to each rational agent type.

6.2.1 Human-Like Agents vs. Philanthropic Agent

The rational agent against which all human agents fare best in the dictator game is the philanthropic agent, due to its persistent philanthropy. The most similar agents in the dictator game are the HAPu agent and the philanthropic rational agent with an average rate difference of 0.12. They lose this similarity in the indirect reciprocity game, because the HAPu agent's donation is now dependent on the original dictator contribution, whereas the philanthropic agent's donations remain stable at 50%.

6.2.2 Human-Like Agents vs. Selfish Agent

Observations show that selfish agents perform very well when competing against the human-like agents. As in humans, these agents do not want to behave too prosocially for fear of impairing themselves, yet they does not want to appear too uncaring for fear of social ramifications. The human-like agents manifest this temperament by making donations (varying in amount by agent type) to selfish agents, despite the steadfast abstinence of the selfish agents to make any donation to the human agents. The fear of social ramification of the human-like agents is well-founded, as observed in the indirect reciprocator game where the selfish agent is unable to exploit any of the human-like agents. Since the selfish agent never donates in the dictator game, then it receives no indirect reciprocity from any of the human-like agents.

6.2.3 Human-Like Agents vs. Reciprocative Agent

The reciprocative rational agent produces nearly balanced resource distribution for all agent types, human and rational, in the dictator and indirect reciprocity games. The nearly zero values for all competitions between human-like and reciprocative agents show that there is no significant exploitation by either agent type over the other.

7. CONCLUSIONS AND FUTURE WORK

As increasing numbers of processors and agents pervade the human environment, societies comprising both humans and agents will emerge. Presently, it is unknown how a person might fair in such mixed societies. To understand these heterogeneous societies, the boundary between sociology and computer science must be redefined. In this work, we explore the idea of using the results of sociological research to model human behavior in a multiagent society as a means of motivating the development of human-agent societies.

This study makes two important contributions to understanding the possible dynamics between agents and humans in a heterogeneous environment: (1) we provide models, within a limited domain, for agents that behave like humans, and (2) we present the results of simulated interactions between the human-like agents and a variety of purely rational agents.

The models for human behaviors are based on recent sociological research characterizing human actions based on person X situation relationships. Agent types based on these models of human behaviors are then pitted against rational agents in various two-player games. The results to these competitions provide evidence that human-like agents elicit prosocial inclinations that facilitate prosperity, for both human-like and rational competitors, when contending with mutually considerate agents. Yet, when human-like agents compete against antisocial agents designed to exploit prosocial motivations, the human-like agents are unable to effectively combat this parasitic behavior due to their own fears of social ramifications for behaving inconsiderately.

This work presumes that a human will behave in the same way towards another human as he or she will towards an agent. There is contradictory research claiming that people feel mistrust for agents and that people trust agents as they do humans [6][7]. To explore the boundaries of human behavior toward agents, we propose a sociological experiment duplicating the work of [10] where human participants are pitted against computer agents.

8. ACKNOWLEDGMENTS

Our thanks go to Dr. Brent Simpson for his advice and assistance in producing this work.

9. REFERENCES

[1] Axelrod, R. 1984. Evolution of Cooperation. Basic Books, New York.

[2] Bandura, A. 1976. Social Learning Theory. Prentice Hall.

[3] Bazzan, A. L. C., Bordini, R. H., Campbell, J. A. 2002. Evolution of Agents with Moral Sentiments in an IPD Exercise. In Game Theory and Decision Theory in Agent-Based Systems, S. Parsons, P. Gmytrasiewicz, and M. Wooldridge, Ed. Kluwer Academic Publishers, Norwell, MA, 43-64.

[4] Castelfranchi, C., Conte, R., Paolucci, M. 1998 Normative Reputation and the Costs of Compliance. J. Artificial Societies and Social Simulation. (June 1998).

[5] Chattoe, E. 2002. Building Empirically Plausible Multi-Agent Systems: A Case Study of Innovative Diffusion. In Socially Intelligent Agents: Creating Relationships with Computers and Robots, K. Dautenhahn, A. H. Bond, L. Canamero, and B. Edmonds, Ed. Kluwer Academic Publishers, Norwell, MA, 109-116.

[6] Oyedele, A., Hong, S., Minor, M. 2007. Contextual Factors in the Appearance of Consumer Robots: Exploratory Assessment of Perceived Anxiety Toward Humanlike Consumer Robots. CyberPsychology & Behavior. (Oct. 2007), 624-632. DOI= 10.1089/cpb.2007.9977.

[7] Reeves, B., Nass, C. 1996. The Media Equation. Cambridge University Press.

[8] Ruvinsky, A. 2007. Encyclopedia of Information Ethics and Security. IGI Global.

[9] Sen, S. 1996. Reciprocity: a foundational principle for promoting cooperative behavior among self-interested agents. In Proceedings of the 2nd International Conference on Multiagent Systems, AAAI Press, CA, 322-329.

[10] Simpson, B., Willer, R. 2007. Altruism and Indirect Reciprocity: The Interaction of Person and Situation in Prosocial Behavior. Social Psychology Quarterly, American Sociological Association, in press.