Simulating Human Behaviors in Agent Societies

Paper ID Number: 690

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 [20] that explores the motivation for 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 this knowledge about 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 draw understanding from human societies. In designing multiagent systems, computer scientists use understanding of rational agent behavior to create and cultivate agent societies. The primary distinction between these two fields is in how each defines its contributing agent, one organic and the other computational. Yet, as interactions between agents and people begin to define societies of their own, this boundary between sociology and computer science begins to blur. In this work, we explore the idea of using the results of sociological research to model human behavior in a multiagent society. We aspire to exploit this knowledge about human societies in order to observe its ramifications in an agent world, and to motivate the development of human-agent societies.

We have based our experimentation on the sociological research performed by Simpson and Willer [20] that explores the motivation for cooperative prosocial behavior in people, a conceivably non-rational process. The simulation presented here models agents based on the statistical findings of [20]. This research seeks to understand how these agents contend against purely rational agent models in order to provide insight into how humans fair when pitted against agents in real-world scenarios. Modeling human behaviors as we have done will present a means of exploring and understanding motivations, consequences, and resolutions to human-agent interactions.

Computer simulation has lately provided a significant and useful tool in sociology research. Scientists have begun using computer simulations to garner new insights into developmental trends in certain societies. For example, Bryson's work [5] has been acclaimed for its examination of the emergence of despotic versus egalitarian social structures of primate groups through multiagent simulations [11]. In the vein of this work, 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. Although there are pitfalls to which humans are vulnerable, there exist niches for human prosperity in the rational agent world. In addition, our results can provide a guide for developing mixed societies of agents and humans.

In section 2 we orient this work with respect to existing work in multiagent systems and other relevant disciplines. This is followed by a brief description of the sociological research that was used to model the human behavior herein [20]. In section 4 we describe the rational agent types that we have selected as competitors for our human modeled agents. An explanation of the design and execution of the experiments we conducted is reported in section 5, which also includes experimental results. We then provide an analysis of the implications and limitations of our experiments and results in section 6. The conclusions realized from this work and the expectations we have of its future potential

Padgham, Parkes, Mueller and Parsons (eds.): Proc. of AAMAS 2008, Estoril, May, 12-16, 2008, Portugal, pp. XXX-XXX

are presented in sections 7 and 8, respectively. The paper ends with acknowledgements and references.

2. BACKGROUND

The work presented in this paper takes advantage of and extends existing work spanning agent reciprocity, computational models of human behavior, aspects of human society, and simulations of agent societies. In this section, we describe the contributions made by the existing work and related it to our research.

Sen [10][16][17][18][19] provides significant and thorough work in reciprocity and motivations for cooperation in games between a variety of agent types. For our simulations described in section 4, we selected agents analogous to those used in [18] as representative of wide range of possible rational agents, including *philanthropic, reciprocative, and selfish*. Sen's *individual* agent type was not used, as there was no clear translation of this agent to our domain.

Bond [4] developed a model of an agent and discussed the relationship between it and humans. He hypothesized two important properties for the relationship: (1) agents and humans must act voluntarily and autonomously, but (2) they must also exert some control over each other. The properties were verified anecdotally, not computationally.

Our work is aligned with that of Rauchier [13], who uses observations of human social intelligence to substantiate her claim that both the designed agent and Artificial Life approaches to building agents do not endow them with the flexible communicative ability they will need for meaningful social behavior. Similarly, Chattoe [9] recommends basing agent system design on information about human societies, as done in our approach.

We model prosocial behavior in humans to build a similar society of prosocial agents. There have been other attempts to build prosocial computational agents, although the attempts have not involved mimicking humans faithfully. Castelfranchi et al. [8] have explored the use of simulating norms like those in human societies to motivate and manage agent behavior. Bryson et al. [5] base their models of non-human animal behavior on various artificial intelligence techniques, and then compare the simulated behaviors to those of real animal society. Bryson's research has focused primarily on primate behavior.

Many game theoretical abstractions to human behaviors have also been investigated in the quest to garner further insight into the fundamental question motivating much research, including our own, as to the evolution of cooperation and cooperative social strategies such as altruism. This type of work is grounded in Axelrod's prisoner's dilemma competition, as described in [1]. Along these lines, Bazzan et al. [2] investigated the effects of altruism among agents playing the Iterated Prisoner's Dilemma.

In [7], Castelfranchi argues that deceitful agents will invariably exist in any society, including human and agent mixed societies. The underhandedness of these agents might not necessarily be for malicious purposes. There might be potentially constructive reasons for this seemingly negative behavior, such as social control. This is ultimately promising, as our work shows that human-like agents are as susceptible to anti-social rational agents as any other type of agent that encounters them. The ulterior reason for having deceitful agents may give reason to believe that this susceptibility may have beneficial consequences to the human-like agents.

3. MODELING HUMAN BEHAVIOR

The models on which the human-like agents in this work are based derive from the statistical results of the work by Simpson and Willer [20]. In their research, they presume a heterogeneous society in which humans are characterized not only by their social preference of altruism or egoism, but also by the situation under which the prosocial behavior is motivated, either public or private. For example, an egoist acting publicly is more motivated to behave prosocially compared to an egoist acting privately, due to the social exposure of the public action. Thus, there are four human characterizations distinguished in [20], namely altruists who act privately, altruists who act publicly, egoists who act privately, and egoists who act publicly.

In our simulations, we similarly model four human-like agent behaviors characterized in this same way. The distributions according to which the human-like agents behave are derived from the results obtained in experiments conducted in [20] on actual human participants. These experiments and the results used to develop our models are discussed in this section.

Preceding this discussion, the motivation for selecting these particular results and experiments as models for human behavior must be explained. The experiments performed in [20] were explicitly designed to expose the distinct prosocial tendencies of specific groups of human participants within a heterogeneous human society. The novelty of this approach is that it characterized human prosocial intentions as more than simply altruist and egoist by incorporating other motivating factors, such as taking into account social pressure from a public action. This provided us with a wider range of human behavioral possibilities without leaving the domain of cooperation.

The experiments in [20] were devised to allow a participant to establish a preference in behavior based on the particular controlled situation. This enabled, for example, an altruist to exhibit degrees of prosociality based on whether the action would be conducted in public or not. For our research, such scenarios that expose the differences in human behaviors are ideal for providing distinguishable observations describing how various human-like agents will perform against different rational agents. For this reason, our experiments for comparing human-like agent behaviors with rational agents, specifically the dictator game and the indirect reciprocator game, were strongly based on the same experimental methods executed in [20].

3.1 The Dictator Game

The dictator game is a two-player game consisting of a dictator and a receiver. In this game, the dictator agent is given a set of resources for which it must choose an amount to donate to a passive receiver agent. It decides how much to give to the receiver and keeps the remainder for itself. This constitutes a potential increase in resources for both the dictator and the receiver.

In [20], the dictator game is performed as an experiment on actual human participants. The data gathered is reported as the mean proportion of resources donated by participants who are characterized by the two parameters of social preference and situational context. The results are as follows:

- Altruist participants in a private situation donated a mean of 40%;
- Altruist participants in a public situation donated a mean of 51%;
- Egoist participants in a private situation donated a mean of 22%; and
- Egoist participants in a public situation donated a mean of 46%.

The conjecture proposed in [20] and supported by the above data is that situational context affects prosocial behavior. More specifically, each social preference of altruism or egoism will exhibit a greater prosocial tendency in public situations than in private situations. The difference between prosocial contributions made in public vs. in private is more significant for egoists, with a difference of 24%, than in altruists, with a difference of 11%. The conclusion is that egoists are more susceptible to social pressures than altruists, whose behaviors are more consistent despite social changes.

The results of the real-world experiment described above were used to generate statistical distributions for simulating human-like behaviors that could be characterized in the same way as the actual human participants' behaviors. The graphs of the generated distribution are shown in Figure 1. We then used the distributions in our NetLogo models of these four types of agents.

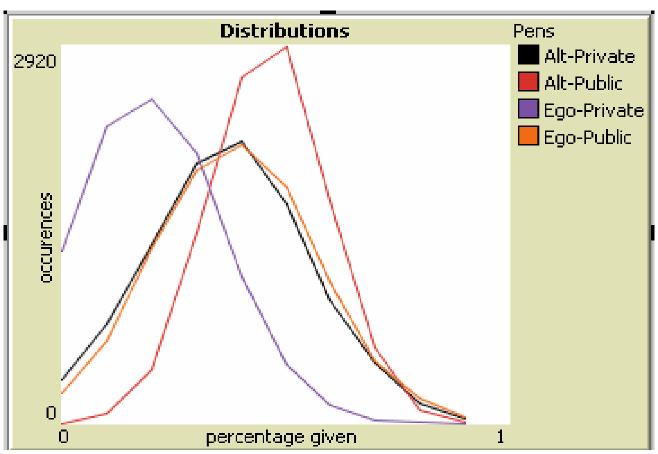

Figure 1. The distribution of donations given by different types of prosocial humans under different circumstances. These distributions were used for our simulated-human agents in the experiments we conducted with NetLogo.

In order to describe the accuracy of the simulated distribution with respect to the actual experimental results reported in [20], the standard deviation of the average percent donated by the simulated human-like agents were compared to the mean proportions donated by the real human participants. If the calculated standard deviation is large, then the simulated distribution does not accurately represent the results generated by the actual human experiments. Small differences in standard deviation signify correspondence between simulation and real world results. Table 1 shows the standard deviations between the average percent donated by each human-like agent type and the mean proportion donated by their human counterparts. The negligible differences may be attributed to simulation error, and our simulated humans are thus an accurate model of the real humans for this domain. These results are presented visually in Figure 2.

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

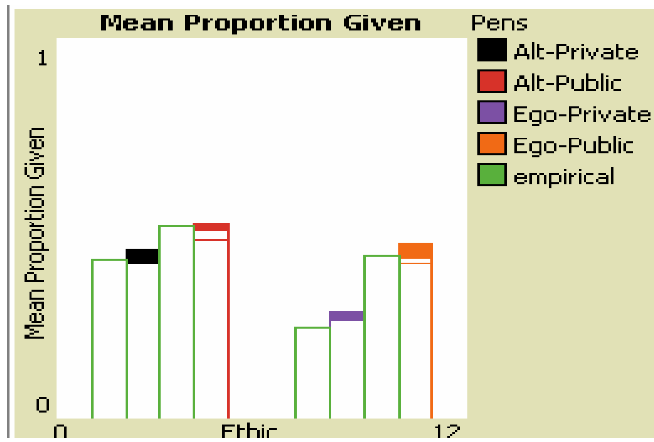

Figure 2: A visual representation of the miniscule amount of error exhibited in when comparing our simulated results with results from the sociological experiment from which the simulation is modeled. This error may be attributed to simulation error. The simulation results are identified in the legend. Each simulation result is located to the right of its corresponding sociological result.

3.2 The Indirect Reciprocity Game

The indirect reciprocity game is also a two-player game consisting of a dictator and an indirect reciprocator. The game begins with the premise that a dictator game has already occurred, although this initial dictator game is fabricated. An independent member of the society is then asked to indirectly reciprocate the original dictator's behavior from the dictator game. This independent member is the indirect reciprocator. The indirect reciprocator is given a set of resource units. This player is then told the portion of resources that the original dictator donated to the receiver in the dictator game, as well as whether this donation was made in public or private. The indirect reciprocator then decides how much of the resources to give to the dictator and keeps the remainder for itself. This is a socially charged game in which the ramifications of an agent's actions come from peers with which the agent did not interact directly.

This indirect reciprocator game is also performed as an experiment on actual 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 performed a donation in private would give the same proportion of its resources as the dictator gave to the receiver. If the dictator gave 50% of his resources, for example, then the indirect reciprocator would give 50% of his resources.
- An altruist indirectly reciprocating to a dictator that performed the donation in public would give 90% of the proportion of its resources that the dictator gave to the receiver. For example, if the dictator gave 50%, then the indirect reciprocator would give 45% of its resources to the dictator.
- An egoist indirectly reciprocating to a dictator that performed the donation in private would give 86% of the proportion of its resources that the dictator gave to the receiver. If the dictator gave 50%, then the indirect reciprocator would give 43%.
- An egoist indirectly reciprocating to a dictator that performed the donation in public would give 64% of the proportion of its resources that the dictator gave to the receiver. If the dictator gave 50%, then the indirect reciprocator would give 32%.

These results show that dictators that had given in private are considered more valuable to the indirect reciprocators than those who gave in public. These prosocial dictators are also more valuable to egoist indirect reciprocators who offer 11% more resources to private actions than to altruist reciprocators that only offer 5 % more resources.

The results of this experiment were used to model analogous indirect reciprocation behaviors in agents.

4. RATIONAL AGENTS

The rational agents devised for this simulation were inspired by those created by Sen [18] for proving that reciprocity promotes cooperation. Sen defined four agents, a philanthropic agent, a selfish agent, a reciprocative agent, and an individual agent. We adopted three of these agent types for the dictator and indirect reciprocator domains: philanthropic agents, selfish agents, and reciprocative agents. The fourth agent, individual agent, could not be translated appropriately into our domain.

The philanthropic agent is characterized as a perpetually cooperative agent. We have defined the philanthropic agent such that when it assumes the role of dictator in the dictator game, it will always behave prosocially by donating 50% of its resources to the receiver. Similarly, in the indirect reciprocator game, the philanthropic agent will also donate 50% of its resources to the dictator, independent of the original donation of the dictator or the situation in which the donation occurred (public or private).

The selfish agent accepts the prosocial kindness of others, but never itself behaves prosocially. We define this agent as one that will willingly accept any donations made by either dictator or indirect reciprocators, but it will never itself donate anything to any other agent when it is in the role of dictator or indirect reciprocator.

The reciprocative agent assesses its indebtedness to another agent in its consideration of how much to donate to that agent. The reciprocative agent assumes both roles of dictator and the indirect reciprocator by behaving in this way. In an attempt to inspire reciprocative behavior, this agent will periodically contribute to an agent to which it is not indebted.

5. EXPERIMENTATION AND RESULTS

5.1 Experiment

As described above, there are 4 human-like agent types and 3 rational agent types. The human-like agents are human altruists in a private context (HAPr), human altruists in a public context (HAPu), human egoists in a private context (HEPr), and human egoists in a public context (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

For the dictator game, agents are paired such that opponents are of different agent types. One of the agents will then randomly decide which of the pair is to be the dictator. One round of the dictator game is played in which the dictator is given 8 resource units. The dictator decides how much to give to the receiver based on the dictator's particular agent type. The receiver's funds are incremented by the amount given, and the dictator's funds are incremented by the amount remaining. Then the agents swap roles so that the receiver now becomes the dictator, and the game is played again. This ensures an equal representation of dictator agents from both agent types throughout the execution of the simulation. Once the two games are played, the average amount of resources acquired by each agent type in this round is calculated. The average amount acquired per round is accrued over repeated iterations of the dictator game in order to observe the rate of change in the amount of resources each agent type is able to accumulate while contending with another agent type.

5.1.2 Implementing the Indirect Reciprocator Game

In the indirect reciprocator game, agents are again paired so that agents of opposing agent type are competing against each other. One agent randomly selects which of the two agents becomes the indirect reciprocator. The remaining agent becomes the dictator. The dictator now fabricates a round of the dictator game. This fabrication is used to produce the amount of resources that the dictator would give to the receiver, but no resources are actually disseminated in this step to either a receiver or the dictator. The indirect reciprocator then identifies an amount to give to the dictator as reciprocity for its behavior in the fabricated dictator game. The indirect reciprocator gives this amount to the dictator and keeps the remainder for itself. The indirect reciprocator and the dictator now swap roles, and the game is played again. After these two games are completed, the average amount of resources that each agent type has accumulated in this round is calculated. This value is also accrued over repeated iterations of the indirect reciprocator game in order to observe the rate of change in the amount of resources that each competing agent type accumulates.

5.2 Results

In our simulation, the amount of resources acquired during each iteration is averaged for each competing agent type. This value is then accrued over many iterations of the games. Figure 3 shows an example of a graph of this accumulation over time for two competing agent types, namely human altruists in a private context versus philanthropic.

The rates at which an agent type accumulates resources as compared to its competitor are calculated. The differences between the rates of change of competing agent types will serve as our metric for characterizing the possible 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 game and the indirect reciprocator game.

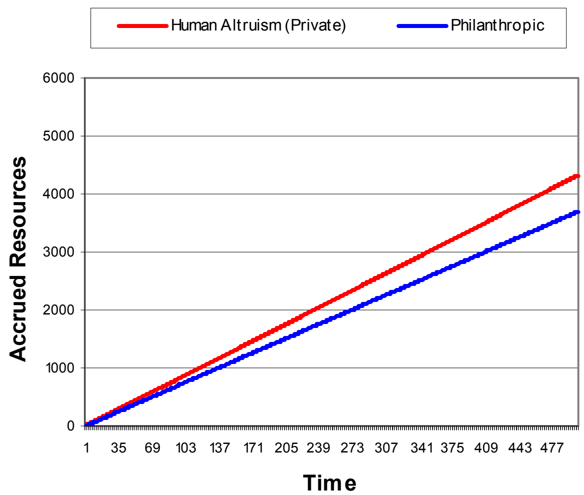

Figure 3. An example of the results obtained from our NetLogo simulation when a simulated human altruist acting privately competes against a rational philanthropic agent. The upper curve shows the wealth accumulated over time by the simulated human, while the lower curve shows the wealth accumulated by the rational agent.

Table 2 depicts the dictator game results of differences in rates at which acquired resources accumulate for contending agents. Table 3 shows these results for the indirect reciprocator game. Table values are with respect to the agent type specified by the row label, such that positive values mean the row agent type accrues resources faster than the column agent type.

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 states of interaction indicated by the difference in rates of accumulated wealth.

(1) Difference is zero. This means that both agent types are gaining and losing resources at the same rate. Thus, neither agent type is exhibiting an apparent benefit over the other. This occurs, for example, when an agent type is competing against itself, against which the agent cannot develop an advantage in accumulating wealth. This can be observed by noting the diagonals of both game tables.

(2) Difference is not zero. This means that the playing field is not equal between these two agent types. In such a case, one of the agents will have a rate of change lower than its opponent. This is due to the agent making larger

donations to its opponent than it is receiving from its opponent. Such an agent is exhibiting a greater prosocial tendency. The opponent, on the other hand, is exploiting the agent's prosociality. An example of a prosocial agent is the human altruist in a public context in the dictator game, whose values are all negative except for its competition against itself. An example of an exploitative agent is the selfish agent in the dictator game.

6.1 General Observations

An obvious aspect of both result tables is their inverse symmetry. This implies consistent performance of the competing social agent types and indicates correctness of the simulation results.

Another observation is that the diagonal of both tables is nearly zero. This is due to the fact that the diagonal represents simulations in which an agent type is competing against itself, thereby gaining and losing approximately the same amount of resources. In these situations, the agent type is unable to develop a prosocial or exploitative advantage over its opponent, namely its own agent type.

A fundamental difference between our human-like agents and rational agents is that the human-like agents, just as their human counterparts, expect their opponents to be susceptible to social pressures as they themselves are. For example, the human-like agents seem to sanction the selfish agents in the indirect reciprocator game in an attempt to motivate the selfish agents' to behave more prosocially in the dictator game. The selfish agent gave nothing in the dictator game, so it gets nothing in the indirect reciprocator game. This is meant to be a kind of lesson to the selfish agent. The human-like agent is attempting to change the opponent's behavior to be more complementary to its own behavior. In this sense, the fear of social ramifications that incites the human-like agent to behave as it does in the dictator game becomes the agent's right to donate nothing to the selfish agent in the indirect reciprocator game. The human-like agent has the right to fear the selfish agent, and thereby sanction it [15]. This sanctioning becomes a means of communicating preferences between agents. The interplay between the agent applying the social pressure and the one receiving this pressure creates the potential for a social learning interaction [2].

6.2 Trends in Individual Human-Like Agents

Based on the observed data, how does a human-like agent fair against rational agents? In the following three subsections, the ramifications of human-like behaviors are analyzed with respect to each rational agent type.

6.2.1 Human-Like Agents vs. the Philanthropic Rational Agent

All human agents fare best competing against the philanthropic rational agent in the dictator game as compared to the other rational agents. This is due to the philanthropic agent's persistent philanthropy.

The most similar agents in the dictator game are the public human altruistic agent and the philanthropic rational agent with an average difference of 0.12. They lose this similarity in the indirect reciprocity game, because the public human altruistic agent's donation is now dependent on the original dictator contribution, which will most likely be less than 50%, whereas the philanthropic agent's donations remain stable at 50%. This is particularly apparent in competing against the selfish agent type, where the public human altruist absolutely does not reciprocate to the selfish agent, whereas the philanthropic agent does reciprocate significantly.

6.2.2 Human-Like Agents vs. the Selfish Rational Agent

In the dictator game, the selfish agent exploits both the human and rational agents alike, though the private human egoistic agent puts up the best fight of all the agent types. Observations show that selfish agents perform very well when competing against the human-like agents. This is shown by the relatively large negative values garnered by the human-like agents when contending against selfish agents. The negative values mean that the human-like agents are accruing resources at a much slower rate than the selfish agents when these agent types play the iterated dictator game. This exposes the selfish agents' exploitation of the prosocial concerns of the human-like agents. The human-like agents' behavior is a result of human temperance of their prosocial decision making process, a condition proven effectively by [20]. A human does not want to behave too prosocially for fear of impairing him/herself, yet a human also 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. This is why the selfish agents win out against the human-like agents in the dictator game; the human agents exhibit some degree of fear of social ramifications if they behave uncaringly towards the selfish agents. For both the philanthropic and reciprocative agent types, the human-like agents' risk in making a donation pays off because these agents make complementary donations.

The fear of social ramification of the human-like agents is well-founded, though, as can be observed in the results to the indirect reciprocator game. In this game, the selfish agent is unable to exploit any of the human agents. This inability to exploit an opponent is a direct result of the selfish agents' most extreme uncaring behavior in the dictator game as expressed by its absolute denial of resources to receiver agents. Since the selfish agent never makes any donation when playing the dictator game, then it receives no indirect reciprocity from any of the human-like agents. An agent inspires indirect reciprocity by its prosocial behavior, of which the selfish agent exhibits none.

6.2.3 Human-Like Agents vs. the Reciprocative Rational Agent

The reciprocative rational agent produces nearly balanced resource distribution for all agent types, both human and rational, in both the dictator and indirect reciprocity games. The human-like agents we have developed fare relatively well when competing against reciprocative rational agents. The nearly zero values for all competitions of both game types show that there is no significant exploitation of the reciprocative rational agent type over the human-like agent types or vise versa. The reciprocative rational agent type creates a fair environment in which the human agents may compete. Reciprocity is known to be an equalizer of rational agents, but here it exhibits flexibility in also accommodating human-like agents as well. [1]

7. CONCLUSIONS

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. We aspire to exploit this knowledge about human societies in order to observe its ramifications in an agent world, and to motivate 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 very recent sociological research characterizing human actions based on person X situation relationships. The relationships identified in [20] and duplicated in this work are human altruism in a private context, human altruism in a public context, human egoism in a private context, and human egoism in a public context.

These agent types based on 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. The human-like agents use reciprocative type behavior to attempt to communicate and inspire this same fear in their opponents.

This work shows promise of the potential for a symbiotic prosocial relationship between agents and people existing together in a mixed society.

8. FUTURE WORK

There are three avenues of investigation arising from the conclusions derived in this work. First, we propose to challenge and relax some of the assumptions made by this work. Primarily, this work presumes that a human will behave in the same way towards another human as he or she will towards an agent. This might not necessarily be true. There has been research in human-computer interactions showing that people feel an element of mistrust towards agents such as avatars or robots that appear too human [12]. On the other hand, there is also research supporting the view that human behaviors shown towards certain agents are equivalent to those shown towards fellow humans [14]. Such sources claim that humans must develop trust for agents through particular interactions designed to inspire trust similarly to developing trust in humans [6]. In order to explore the boundaries of assumptions about human behavior toward agents, a sociological experiment must be conducted with actual humans. These tests should duplicate the work of [20] with the alteration that the opponents of the participants be identified as computer agents. The awareness of the non-human quality of the participant's opponent may alleviate the perceived pressure of social ramifications, causing the human participant to behave prosocially to the same degree as when interacting with a human.

A second avenue of investigation is to consider the human-like agents developed in this research as feasible agents in other multiagent domains. The social pressures that the human-like behaviors elicit may be useful in socially significant problem spaces, such as negotiation, or in multiplayer domain spaces not explored in this work, such as auctions or task allocation.

Third, we are interested in exploring the idea of creating human-like agents to train learning agents to better understand and complement human-like interactions. It has been discussed that the human-like agents in this work generate certain behaviors designed to attempt to change the opponent's behavior to be more complementary to their own. This indirect communication of opponent's behavioral preferences creates the potential of a learning environment for this opponent. If agents are able to learn from human-like agents, essentially adopting a human's behaviors as a social "model" [2], then they can be designed to learn interaction preferences from humans themselves. In this case, the problem being investigated here may change from assessing how humans may contend with rational agents to developing rational agents that adapt to complement human preferences. In this way we may ensure that the agents pervading our environments already are, or can be made to be, complementary to our interests.

9. ACKNOWLEDGMENTS

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

10. 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., and 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] Bond, A. H. 2002. Modeling Social Relationship: An Agent Architecture for Voluntary Mutual Control. 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, 29-36.

[5] Bryson, J., Ando, Y., Lehmann, H. 2007. Agent-based Modeling as Scientific Method: a case study analyzing primate social behavior. (Spet. 2007), 1685-1698. DOI= 10.1098/rstb.2007.2061.

[6] Cassell, J., Bickmore, T. 2000. External Manifestations of Trustworthiness in the Interface. Communications of the ACM. (Dec. 2000), 50-56. DOI= http://doi.acm.org/10.1145/355112.355123.

[7] Castelfranchi, C. 2000. Artificial Liars: Why computers will (necessarily) deceive us and each other. Ethics and Information Technology. (June 2000), 113-119.

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

[9] 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.

[10] Dutta, P. Sen, S. 2003. Forming Stable Partnerships. Cognitive Systems Research. (Sept. 2003), 211-221. DOI= http://dx.doi.org/10.1016/S1389-0417(03)00005-6

[11] Fleming, N., Highfiefld, R. 2007. Computers Shed Light on Lives of Monkeys.

[12] 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= http://www.liebertonline.com/doi/abs/10.1089/cpb.2007.9977.

[13] Rauchier, J. 2002. Social Intelligence for Computers: Making Artificial Entities Creative in their Interactions. 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, 85-92.

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

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

[16] Saha, S., Sen, S., Dutta, P. Helping Based on Future Expectations. In Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems, (Melbourne, Australia, July 2003) ACM, 289-296.

[17] Saha, S., Sen, S. 2007. Reciprocal Negotiation over Shared Resources in Agent Societies. In Proceedings of the Autonomous Agent and Multiagent Based Systems Conference, (Honolulu, Hawai'I, USA, May 2007), 771-773.

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

[19] Sen, S. 2002. Believing others: Pros and cons. Artificial Intelligence. (Dec. 2002), 179-203.

[20] 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.