Human Behavior in Mixed Human-Agent Societies

(Extended Abstract)

Alicia Ruvinsky
University of South Carolina
Computer Science and Engineering
Columbia, SC 29208 USA
ruvinsky@cec.sc.edu 2nd. author

Michael N. Huhns
University of South Carolina
Computer Science and Engineering
Columbia, SC 29208 USA
huhns@sc.edu

ABSTRACT

Agents are increasingly being implemented on the many processors embedded in our environment, resulting in more frequent interactions between agents and humans and the emergence of mixed human-agent societies. We are attempting to cultivate an interdisciplinary understanding of productive human-agent societies by investigating characteristics of human cooperation toward software agents. The contribution of this work is an exploration characterizing direct and indirect behavior of humans toward agents by way of experimentation on human subjects. Preliminary investigation into human-agent interactions reveals that in a context involving direct and indirect behavior, humans will exhibit prosocial behaviors toward agents on par with those they exhibit toward humans. These results seem to support Reeves and Nass's media equation describing human interactions with media entities as equivalent to human interactions with humans.

Categories and Subject Descriptors

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

General Terms

Experimentation, Human Factors

Keywords

Experimental/Empirical, Agent-Human Interactions, Societal Aspects of MAS, Artificial Social Systems, People-Agent Societies, Models of Social Behavior

1. INTRODUCTION

Current trends in agent theory, with respect to work flow, negotiation, and cognitive modeling to name a few, promise increasingly innovative and embedded uses for agents. Coupling agent-based technologies with increasing numbers of processors pervading the human environment, it is inevitable for interactions between agents and humans to occur, resulting in the emergence of mixed human-agent societies. Presently, it is unknown how a person might fare in mixed societies, or how a person's behavior might be altered when participating in such societies.

In considering how humans interact with computer technology, Mick and Fournier state that though essential for contemporary life, the effects of technology are not imperceptible [2]. The primary discrepancies of technology are the anxiety producing paradoxes it introduces into the lives of its users based on the fundamental social dialectic by which humans perceive technology: as enabling freedom, control, and efficiency, while degrading the environment, undermining human competence, and propagating human dependence [2].

Ficici and Pfeffer show that humans take their opponents into account during decision making processes [1]. As such, the opponent's agency, either human or artificial, is likely to play a role in human decision making as well.

Anthropomorphism is the process of attributing human qualities and characteristics to non-human entities. In [3], Nowak and Biocca's experimental results suggest that anthropomorphizing technology enables a social interpretation comparable to that of a human interaction.

When considering how people interact with media, Reeves and Nass devise the media equation for this human-technology interaction where media = real life [4]. The equation proposes that people exhibit an unconscious response to digital media as if they were interacting with humans.

Nowak and Biocca explore whether people can actually distinguish between an artificial versus a human entity behind the media. Though they hypothesized that people would feel more of a presence behind the media when interacting with humans versus agents, experimentation did not support these claims. The results suggest that participants interact with computers in a social manner comparable to interactions with other humans [3].

Our work presents an investigation characterizing human behavior toward agents designed to answer the following questions. Do people give resources to an agent as they do to a person? Do people give resources to another entity when being observed by an agent as they do when observed by a person? Do people indirectly reciprocate to an agent as they do to a person? Do people indirectly reciprocate to an entity when in response to an action toward an agent as they do in response to an action toward a person?

2. HUMAN SUBJECT EXPERIMENTS

Our study consists of two games, the dictator game and the indirect reciprocator game, and is based on experiments conducted by Simpson and Willer [5]. In the dictator game,

Cite as: Human Behavior in Mixed Human-Agent Societies, (Extended Abstract), Alicia Ruvinsky, Michael N. Huhns, *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009),* Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. 1165–1166

the human participant plays the role of the dictator who decides how many points of a resource pool to give to a receiver, a simulated human or agent. This interaction is either made in public where the observer is either a simulated human or agent, or in private.

In the indirect reciprocator game, the human participant plays the role of the indirect reciprocator deciding how many points of a slightly larger resource pool to give to the dictator that had previously given points to a receiver, where the dictator and receiver are played by simulated humans or agents. The reciprocation is in response the dictator's previous behavior that was enacted either in public or in private.

Each participant played a public and private dictator game, randomly ordered, followed by a public and private indirect reciprocator game, randomly ordered. Public actions are intended to impose social pressure on the human participants by creating reputational incentive. Contrarily, private actions are design to elicit the human participant's true social nature by eliminating social pressure.

The intention of public and private contexts in the dictator game was to observe participant's behavior with and without social pressure. In the indirect reciprocator game, the contexts were intended to show how the participants reward publicly versus privately made transactions.

3. DICTATOR GAME RESULTS

We are interested in whether the agency of the receiver and observer in the dictator game affects a participant's first-order direct contributing behavior.

Do people feel less social pressure to give resources to an agent rather than to a person? We hypothesize that participants will feel less social pressure when giving to an agent than a person, despite whether they are being observed by a person or an agent. Social pressures that exist between people do not transcend to artificial agents.

The observed data showed that the amount a person contributes to a person versus to an agent does not differ significantly regardless of whether the action is observed by a person or an agent.

Do people feel less social pressure to give resources to a receiver (human or artificial) when they are observed by an agent rather than a person? The social pressures that exist between people do not transcend to agents. We hypothesize that a person will feel less social pressure to give resources to a receiver when they are observed by an agent than when they are observed by a person.

The observed data showed that the contributing behavior of people towards a receiver, regardless of human or agent, remains consistent when being observed by a person versus an agent.

4. INDIRECT RECIPROCATOR RESULTS

We are interested in whether the agency of the dictator and receiver in the indirect reciprocator game affects a participant's reciprocating behavior.

Do people indirectly reciprocate to an agent as they do to a person? We predict that participants will indirectly reciprocate to an agent less than they will indirectly reciprocate to a person. The norm of indirect reciprocity which informs interactions between people will not transcend to agents. People will presume that software agents do not subscribe to the norm of indirect reciprocity.

The data showed that people's indirect reciprocity towards agents is similar to their indirect reciprocity towards other people. The agency of the dictator in the indirect reciprocator game did not affect the behavior of the participants.

Do people indirectly reciprocate differently when in response to an action towards an agent as they do in response to an action towards a person? We predict that participant will value an interaction with an agent less than an interaction with a person, so the participant will indirectly reciprocate less in response to an interaction with an agent than an interaction with a person.

The data showed that people's indirect reciprocation behavior is consistent when responding to an action towards an agent versus when responding to an action towards a person. The agency of the receiver in the indirect reciprocator game did not affect the behavior of the participants in the role of indirect reciprocator.

5. CONCLUSIONS

Our investigation into human-agent interactions reveals that in a context involving first-order direct contributing behavior and second-order indirect reciprocation behavior, there is no support for an affect of agency in prosocial behavior of humans towards agents versus towards humans. Specifically, we have shown probable support for the following statements. The amount a person contributes to a person versus to an agent does not differ significantly. The contributing behavior of people remains consistent when being observed by a person versus an agent. Indirect reciprocity of people towards agents is similar to the indirect reciprocity of people towards other people. The indirect reciprocation behavior of people remains consistent when responding to an action towards an agent versus when responding to an action towards a person. These results seem to support Reeves and Nass's media equation describing human interactions with media entities as equivalent to human interactions with humans.

6. REFERENCES

[1] S. G. Ficici and A. Pfeffer. Modeling how humans reason about others with partial information. In *AAMAS '08: Proc. of 7th Int. Conf. on Autonomous agents and multiagent systems*, pages 315–322, 2008.

[2] D. G. Mick and S. Fournier. Paradoxes of technology: Consumer cognizance, emotions, and coping strategies. *Journal of Consumer Research*, 25(2):123–43, September 1998.

[3] K. L. Nowak and F. Biocca. The effect of the agency and anthropomorphism of users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoper. Virtual Environ.*, 12(5):481–494, 2003.

[4] B. Reeves and C. Nass. *The media equation: how people treat computers, television, and new media like real people and places*. Cambridge University Press, 1998.

[5] B. Simpson and R. Willer. Altruism and indirect reciprocity: The interaction of person and situation in prosocial behavior. *Social Psychology Quarterly*, 71(1):37–52, March 2008.