PsychSim: Agent-based modeling of social interactions and influence

Stacy C. Marsella and David V. Pynadath
Information Sciences Institute
University of Southern California
4676 Admiralty Way, Marina del Rey, CA 90292 USA
{marsella,pynadath}@isi.edu

Stephen J. Read
Department of Psychology
University of Southern California
Los Angeles, CA 90089-1061 USA
read@usc.edu

Abstract

Agent-based modeling of human social behavior is an increasingly important research area. For example, it is critical to designing virtual humans, human-like autonomous agents that interact with people in virtual worlds. A key factor in human social interaction is our beliefs about others, in particular a theory of mind. Whether we believe a message depends not only on its content but also on our model of the communicator. The actions we take are influenced by how we believe others will react. However, theory of mind is usually ignored in computational models of social interaction. In this paper, we present PsychSim, an implemented multiagent-based simulation tool for modeling interactions and influence among groups or individuals. Each agent has its own decision-theoretic model of the world, including beliefs about its environment and recursive models of other agents. Having thus given the agents a theory of mind, PsychSim also provides them with a psychologically motivated mechanism for updating their beliefs in response to actions and messages of other agents. We discuss PsychSim and present an example of its operation.

Introduction

Human social interaction is complex. Rarely are our interactions independent, nor do they fall into narrow categories of full cooperation or competition. People may share some goals but not others, may cooperate at some times and compete at others. To navigate this complexity, we rely on forming beliefs about the goals and behavior of others, what is called a *theory of mind* [Whiten, 1991]. These beliefs inform not only our decisions about what to do, but also what to believe.

There are a range of applications where the computational modeling of such social interactions and influence in general [Prietula and Gasser, 1998, Liebrand and Hegselmann, 1998, Gilbert and Troitzsch, 1999], and theory of mind in particular, is important. Theory of mind, for example, plays a role in childhood aggression whereby one child's aggression may be rooted in misattribution of another child's intent or outcome expectancies on how people will react to the violence [Schwartz, 2000]. To develop a better understanding of the causes and remedies of school bullying, we could use agent models of the students that incorporate a theory of mind to simulate and study classroom social interactions. Models of social interaction have also been used to create social training environments where the learner explores high-stress social interactions in the safety of a virtual world (e.g., [Marsella et al., 2000, Swartout et al., 2001]). In particular, one might imagine an environment where a young teacher could train for handling unruly students by playing the role of a teacher in a virtual classroom populated with virtual students (e.g., see `www.victec.org`). The teacher could explore in the safety of a virtual world the same sorts of situations and dilemmas he is likely to face in the real world, with interaction between human and virtual characters dynamically unfolding based on choices they make in their various roles.

To facilitate such explorations and applications, we have developed a social simulation tool, called PsychSim, designed to explore how individuals and groups interact and how those interactions can be influenced. PsychSim allows an end-user to quickly construct a social scenario, where a diverse set of entities, either groups or individuals, interact and communicate among themselves. Each entity has its own goals, relationships (e.g., friendship, hostility, authority) with other entities, private beliefs and mental models about other entities. The simulation tool generates the behavior for these entities and provides explanations of the result in terms of each entity's goals and beliefs. A user can play different roles by specifying actions or messages for any entity to perform. Alternatively, the simulation itself can perturb the scenario to provide a range of possible behaviors that can identify critical sensitivities of the behavior to deviations (e.g., modified goals, relationships, or mental models).

A central aspect of the PsychSim design is that agents have fully specified decision-theoretic models of others. Such quantitative recursive models give PsychSim a powerful mechanism to model a range of factors in a principled way. For instance, we exploit this recursive modeling to allow agents to form complex attributions about others, enrich the messages between agents to include the beliefs and goals of other agents, model the impact such recursive models have on an agent's own behavior, model the influence observations of another's behavior have on the agent's model of that other, and enrich the explanations provided to the user. The decision-theoretic models in particular give our agents the ability to judge degree of credibility of a message in a subjective fashion that factors in a range of influences that sway such judgments in humans. In this paper, we present PsychSim, discuss key aspects of its approach to modeling social interaction, and illustrate its operation on a school bullying example.

PsychSim Overview

PsychSim is shown in Figure 1. The user sets up a simulation by selecting generic agent models that will play the roles of the various groups that will be simulated and spe-

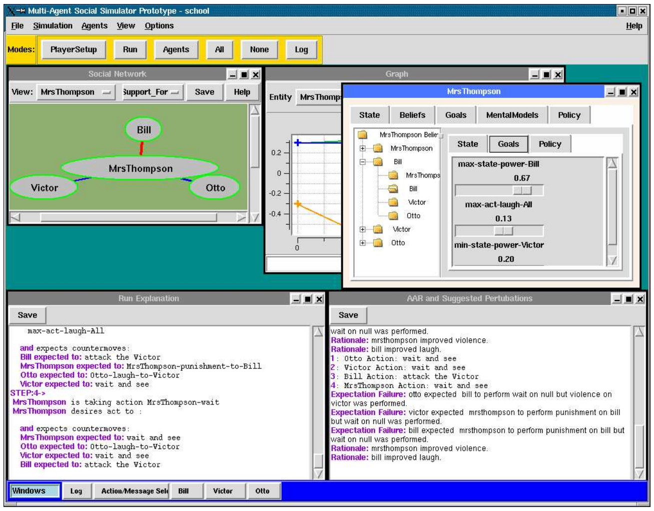

Figure 1: Screenshot of PsychSim interface

cializing those models as needed. To facilitate setup, PsychSim uses an automated fitting algorithm. For example, if the user wants the bully to initially attack a victim and wants the teacher to threaten the bully with punishment, then the user specifies those behaviors and the model parameters are fitted accordingly [Pynadath and Marsella, 2004]. This degree of automation has the potential to significantly simplify simulation setup.

Execution of the simulation incorporates a variety of visualization techniques along with analyses of the simulations. PsychSim allows one to explore multiple tactics for dealing with a social issue and to see potential consequences of those tactics. How might a bully respond to admonishments, appeals to kindness or punishment? How might other groups react in turn? What are the predictions or unintended side-effects?

Finally, there is an analysis/perturbation capability that supports the iterative refinement of the simulation. The intermediate results of the simulation (e.g., the reasoning of the agents in their decision-making, their expectations about other agents) are all placed into a database. Inference rules analyze this database to explain the results to the user in terms of the agents' motivations, including how their beliefs and expectations about other agents influenced their own behavior and whether those expectations were violated. Based on this analysis, the system also reports sensitivities in the results, as well as potentially interesting perturbations to the scenario.

The rest of this paper describes PsychSim's underlying architecture in more detail, using a school bully scenario for illustration. The agents represent different people and groups in the school setting. The user can analyze the simulated behavior of the students to explore the causes and cures for school violence. One agent represents a bully, and another represents the student who is the target of the bully's violence (for young boys, the norm would be physical violence, while young girls tend to employ verbal abuse and ostracizing). A third agent represents the group of onlookers, who encourage the bully's exploits by, for example, laughing at the victim as he is beaten up. A final agent represents the class's teacher trying to maintain control of the classroom, for example by doling out punishment in response to the violence.

The Agent Models

We embed PsychSim's agents within a decision-theoretic framework for quantitative modeling of multiple agents. Each agent maintains its independent beliefs about the world, has its own goals and it owns policies for achieving those goals. The PsychSim framework is an extension to the Com-MTDP model [Pynadath and Tambe, 2002].

State: Each agent model includes several features representing its "true" state. This state consists of objective facts about the world, some of which may be hidden from the agent itself. For our example bully domain, we included such state features as `power(agent)`, to represent the strength of an agent, though the agent may have its own subjective view of its own power. It is impacted by acts of violence, conditional on the relative powers of the interactants. `trust(truster,trustee)` represents the degree of trust that the agent `truster` has in another agent `trustee`'s messages. `support(supporter,supportee)` is the strength of support that an agent `supporter` has for another agent `supportee`. We represent the state as a vector, $\vec{s}^t$, where each component corresponds to one of these state features and has a value in the range $[-1, 1]$.

Actions: Agents have a set of actions that they can choose to perform in order to change the world. An action consists of an action type (e.g., `punish`), an agent performing the action (i.e., the actor), and possibly another agent who is the object of the action. For example, the action `laugh(onlooker, victim)` represents the laughter of the onlooker directed at the victim. Actions affect the state of the world, via a transition function, $T(\vec{s}^t, a)$, that returns a modified state vector, $\vec{s}^{t+1}$, that captures the effects of the performed action, a.

Goals: An agent's goals represent its incentives (and disincentives) for behavior. In PsychSim's decision-theoretic framework, we represent goals as a reward function that maps the current state of the world into a real-valued evaluation of benefit for the agent. A goal of **Minimize/maximize** `feature(agent)` corresponds to a negative/positive reward proportional to the value of the given state feature. For example, an agent can have the goal of maximizing its own power. A goal of **Minimize/maximize** `action(actor, object)` corresponds to a negative/positive reward proportional to the number of matching actions performed. For example, the teacher may have the goal of minimizing the number of times any student teases any other.

We can represent the overall goals of an agent, as well as the relative priority among them, as a vector of weights, $\vec{g}$, so that the product, $\vec{g} \cdot \vec{s}^t$, quantifies the degree of satisfaction that the agent receives from the world, as represented by the state vector, $\vec{s}^t$. For example, in the school violence simulation, the bully's reward function consists of goals of maximizing `power(bully)`, minimizing `power(victim)`, minimizing `power(teacher)`, and maximizing `laugh(onlookers, victim)`. We can model a sadistic bully with a high weight on the goal of minimizing `power(victim)` and an attention-seeking bully with a high weight on maximizing `laugh(onlookers, victim)`. In other words, by modifying the weights on the different goals, we can alter the motivation of the agent and, thus, its behavior in the simulation.

Beliefs: The simulation agents have only a *subjective* view of the world, where they form beliefs, denoted by the vector $\vec{b}^t$, about what they *think* is the state of the world, $\vec{s}^t$. The structure of an agent's beliefs is identical to the representation of the objective world state. In other words, each agent's beliefs consist of models of all of the agents (including itself), representing their state, beliefs, goals, and policy of behavior. For example, an agent's beliefs may include its subjective view on states of the world: "The bully believes that the teacher is weak", "The onlookers believe that the teacher supports the victim", or "The bully believes that he/she is powerful." These beliefs may also include its subjective view on beliefs of other agents: "The teacher believes that the bully believes the teacher to be weak." An agent may also have a subjective view of the *goals* of other agents: "Teacher believes that the bully has a goal to increase his power."

Policies: Each agent's policy is a function, $\pi(\vec{b}, \vec{g})$, that represents the process by which it selects an action or message based on its beliefs and goals. An agent's policy allows us to model critical psychological distinctions such as reactive vs. deliberative behavior. PsychSim models this process as a table of "Condition⇒Action" rules. The left-hand side conditions may trigger on an *observation* of some action or a *belief* of some agent (e.g., such as the bully believing himself as powerful). The right-hand side takes on one of the following forms: action, wait, or bounded lookahead. An action on the right-hand side causes the agent to execute the specified action. A *wait* right-hand side means that the agent does nothing.

We model each agent's real policy table as including a bounded lookahead policy rule that seeks to best achieve the agent's goals given its beliefs. To do so, the policy considers all of the possible actions/messages it has to choose from and measures the results by simulating the behavior of the other agents and the dynamics of the world in response to the selected action/message. They compute a quantitative value, $V_a(\vec{b}^t)$, of each possible action, a, given their beliefs, $\vec{b}^t$.

$$V_a(\vec{b}^t) = \vec{g} \cdot T(\vec{b}^t, a) + V(T(\vec{b}^t, a)) \quad (1)$$

$$V(\vec{b}^{t+1}) = \sum_{\tau=1}^{N} \vec{g} \cdot \vec{b}^{t+\tau} \quad (2)$$

Thus, an agent first uses the transition function, T, to project the immediate effect of the action, a, and then projects another N steps into the future, weighing each state along the path against its goals, $\vec{g}$. Thus, the agent is seeking to maximize the expected value of its behavior, along the lines of decision policies in decision theory and decision theory. However, PsychSim's agents are only boundedly rational, given that they are constrained, both by the finite horizon, N, of their lookahead and the possible error in their belief state, $\vec{b}$.

In addition, we give PsychSim's agents erroneous beliefs about the policies of other agents. In particular, the agents *believe* that the other agents do not do much lookahead, even though, in reality, they use lookahead exclusively when choosing their own actions. This achieves two desirable results. First, from a human modeling perspective, the agents perform a shallower reasoning when thinking about other agents, which provides an accurate model of the real-world entities they represent. Second, from a computational perspective, the direct action rules are cheap to execute, so the agents gain significant efficiency in their reasoning by avoiding expensive lookahead.

Messages: Messages are attempts by one agent (or the user) to influence the beliefs of recipients. Messages have five components: a source, recipients, a message subject, content and overhearers. For example, the teacher (source) could send a message to the bully (recipient) that the principal (subject of the message) will punish acts of violence by the bully (content). Finally, overhearers are agents who hear the message even though they are not one of the intended recipients. Messages can refer to beliefs, goals, policies, or any other aspect of other agents. Thus, a message may make a claim about a state feature of the message subject ("the principal is powerful"), the beliefs of the message subject ("the principal believes that he is powerful"), the goals of the message subject ("the bully wants to increase his power"), or the policy of the message subject ("if the bully thinks the victim is weak, he will pick on him").

Mental Models: An agent's beliefs about another agent are realized as a fully specified agent model of the other agent, including goals, beliefs and policies. To simplify the setup of the system, these mental models are realized as stereotypes. For our simulation model of a bullying scenario, we have implemented mental models corresponding to *selfishness*, *altruism*, *dominance-seeking*, etc. For example, a model of a selfish agent specifies a goal of increasing self-wealth as paramount, while a model of an altruistic agent specifies a goal of helping the weak. Similarly, a model of an agent seeking dominance specifies a goal of having more power than its competitors.

Modeling Influence and Belief Change

A challenge in creating a social simulation is addressing how groups or individuals influence each other, how they update their beliefs and alter behavior based on observations of, as well as messages from, others. Although many psychological results and theories must inform the modeling of such influence (e.g., [Cialdini, 2001, Abelson et al., 1968, Petty and Cacioppo, 1986]), they often suffer from two shortcomings from a computational perspective. First, they identify factors that affect influence but do not operationalize those factors. Second, they are rarely comprehensive and do not address the details of how various factors relate to each other or can be composed. To provide a sufficient basis for our computational models, our approach has been to distill key psychological factors and map those factors into our simulation framework. Here, our decision-theoretic models are helpful in quantifying the impact of factors and in such a way that they can be composed.

Specifically, a survey of the social psychology literature identified the following key factors:

Consistency: People expect, prefer and are driven to maintain consistency, and avoid cognitive dissonance, between beliefs and behaviors. This includes consistency between their old and new information, between beliefs and behavior, as well as consistency with the norms of their social group.

Self-interest: Self-interest impacts how information influences us in numerous ways. It impacts how we interpret appeals to one's self-interest, values and promises of reward or punishment. The inferences we draw are biased by self-interest (e.g., motivated inference) and how deeply we analyze information in general is biased by self-interest. Self-interest may be in respect to satisfying specific goals like "making money" or more abstract goals such as psychological reactance, the tendency for people to react to potential restrictions on freedom such as their freedom of choice (e.g., the child who is sleepy but refuses to go to bed when ordered by a parent.)

Speaker's Self-interest: If the sender of a message benefits greatly if the recipient believes it, there is often a tendency to be more critical and for influence to fail.

Trust, Likability, Affinity: The relation to the source of the message, whether we trust, like or have some group affinity for him, all impact whether we are influenced by the message.

Providing each agent with a model of itself and, more importantly, fully specified models of other agents gave us a powerful mechanism to model this range of factors in a principled way. We can model these factors by a few simple mechanisms in the simulation: *consistency*, *self-interest*, and *bias*. We can render each as a quantitative function on beliefs that allows an agent to compare alternate candidate belief states (e.g., an agent's original $\vec{b}$ vs. the $\vec{b}'$ implied by a message).

Consistency is an evaluation of whether the content of a message or an observation was consistent with prior observations. In effect, the agent asks itself, "If this message is true, would it better explain the past better than my current beliefs?". An agent assesses the quality of the competing explanations by a re-simulation of the past history. In other words, it starts at time 0 with the two worlds implied by the two candidate sets of beliefs, projects each world forward up to the current point of time, and compares the projected behavior against the behavior it actually observed. In particular, the consistency of a sequence of observed actions, $\omega^0, \omega^1, \ldots$, with a given belief state, $\vec{b}$, corresponds to:

$$\text{consistency}(\vec{b}^t, [\omega^0, \omega^1, \ldots, \omega^{t-1}]) = \sum_{\tau=0}^{t-1} V_{\omega^\tau}(\vec{b}^\tau) \quad (3)$$

Note that the value function, V, computed is with respect to the agent performing the action at time τ. In other words, we are summing the value of the observed action to the acting agent, given the set of beliefs under consideration. The higher the value, the more likely that agent is to have chosen the observed action, and, thus, the higher the degree of consistency.

Self-interest is similar to consistency, in that the agent compares two sets of beliefs, one which accepts the message and one which rejects it. However, while consistency requires evaluation of the past, we compute self-interest by evaluating the future using Equation 2. In other words, if $V(\vec{b}) > V(\vec{b}')$ (all else equal), then an agent prefers believing $\vec{b}$. An agent can perform an analogous computation using its beliefs about the sender's reward to compute the sender's self-interest in sending the message.

Bias factors act as tie-breakers when consistency and self-interest fail to decide acceptance/rejection. We treat support (or affinity) and trust as such a bias on message acceptance. Agents compute their support and trust levels as a running history of their past interactions. In particular, one agent increases (decreases) its trust in another, when the second sends a message that the first decides to accept (reject). Similarly, an agent increases (decreases) its support for another, when the second selects an action that has a high (low) reward, with respect to the goals of the first. In other words, if an agent selects an action a, then the other agents modify their support level for that agent by a value proportional to $\vec{g} \cdot T(\vec{b}, a)$, where $\vec{g}$ corresponds to the goals of the agent modifying its support.

Upon receiving any information (whether message or observation), an agent must consider all of these various factors in deciding whether to accept it and how to alter its beliefs (including its mental models of the other agents). For a message, the agent determines acceptance using a weighted sum of the five components: consistency, self-interest, speaker self-interest, trust and support.

We see the computation of these factors as a toolkit for the user to explore the system's behavior under existing theories, which we can encode in PsychSim. For example, the elaboration likelihood model (ELM)[Petty and Cacioppo, 1986] is a two-process model that argues that the way messages are processed differs according to the relevance of the message to the receiver. High relevance or importance would lead to a deeper assessment of the message, which is consistent with the self-interest calculations our model performs. For less relevant messages, more peripheral processing of perceptual cues such as "liking for" the speaker would dominate. PsychSim's linear combination of factors is roughly in keeping with ELM because self-interest values of high magnitude would tend to dominate.

Example Scenario in Operation

The research literature on childhood bullying and aggression provides interesting insight into the role that theory of mind plays in human behavior. Although a number of factors are related to bullying, two social cognitive variables have been shown to play a central role. One variable discussed is a hostile attributional style [Nasby et al., 1979], wherein typical playground behaviors are interpreted as having a hostile intent. Children who tend to see other children as intending to hurt them are more likely to display angry, retaliatory aggression. A second variable is outcome expectancies for the effectiveness of aggression. Children develop outcome expectancies for the effectiveness of aggression depending on whether in the past they have been rewarded for its use or found it to be ineffective or punished for it.

Investigations of bullying and victimization [Schwartz, 2000] have identified four distinct types of children: those who characteristically display reactive aggression (aggressive victims), those who display proactive aggression (nonvictimized aggressors), those who are typically victimized (nonaggressive victims), and normal children. Nonaggressive victims display a hostile attributional style and have negative outcome expectancies for aggression. Aggressive victims tend to have a hostile attributional style, but neither positive nor negative outcome expectancies for aggression. Nonvictimized aggressors have positive outcome expectancies for aggression, but do not have a hostile attributional style.

We have begun to use use PsychSim to explore psychological theories by demonstrating how PsychSim can represent both *attributional style* and *outcome expectancies* in a simulation of school violence. The user can manipulate each factor to generate a space of possible student behaviors for use in simulation and experimentation. For example, an agent's attributional style corresponds to the way in which it updates its beliefs about others to explain their behavior. A hostile attributional style corresponds to an agent who tends to adopt negative mental models of other agents. In our example scenario, agents with a hostile attributional style mentally model another student as having the goal of hurting them (i.e., minimizing their power).

Our agents already compute the second factor of interest, outcome expectancies, as the expected value of actions (V_a from Equation 1). Thus, when considering possible aggression, the agents consider the immediate effect of an act of violence, as well as the possible consequences, including the change in the beliefs of the other agents. In our example scenario, a bully has two incentives to perform an act of aggression: (1) to change the power dynamic in the class (i.e., weaken his victim and make himself stronger), and (2) to earn the approval of his peers (as demonstrated by their response of laughter at the victim). Our bully agent models the first incentive as a goal of maximizing `power(bully)` and minimizing `power(victim)`, as well as a belief that an act of aggression will increase the former and decrease the latter. The second incentive must also consider the actions that the other agents may take in response. The agents' theory of mind is crucial here, because it allows our bully agent to predict these responses, albeit limited by its subjective view.

For example, a bully motivated by the approval of his classmates would use his mental model of them to predict whether they would enjoy his act of aggression and laugh along with him. Similarly, the bully would use his mental model of the teacher to predict whether he will be punished or not. The agent will weigh the effect of these subjective predictions along with the immediate effect of the act of aggression itself to determine an overall expected outcome. Thus, the agents' ability to perform bounded lookahead easily supports a model for proactive aggression.

We explored the impact of different types of proactive aggression by varying the priority of the two goals (increasing power and gaining popularity) within our bully agent. When we ran PsychSim using an agent model where the bully cares about each goal equally, then the threat of punishment is insufficient to change the bully's behavior, because he expects to still come out ahead in terms of his popularity with his peers. On the other hand, a threat against the whole class in response to the bully's violence is effective, because the bully then believes that an act of violence will *decrease* his popularity among his peers. If we instead use an agent model where the bully favors the first goal, then even this threat against the whole class is ineffective, because the bully no longer cares about his popularity in the class.

Of course, this example illustrates one outcome, where we do not change any of the other variables (e.g., bully's power relative to victim, teacher's credibility of threats). PsychSim's full range of variables provide a rich space of possible class makeups that we can systematically explore to understand the social behavior that arises out of different configurations of student psychologies. We have also begun developing algorithms that are capable of finding the configuration that best matches a real-world class dynamic, allowing us to find an underlying psychological explanation for a specific instance of behavior [Pynadath and Marsella, 2004]. Furthermore, as illustrated, we can try out different interventions in the simulation to understand their impact under varying student models. As we have seen, alternate scenarios will have different results, but by systematically varying the scenario, we can draw general conclusions about the effectiveness of these different intervention methods. Finally, although this section uses a specific taxonomy of student behavior to illustrate PsychSim's operation, the methodology itself is general enough to support the exploration of many such taxonomies.

Related Work

PsychSim agents model each other's beliefs, goals, policies, etc. and are able to reason with it. This is essentially a form of recursive agent modeling [Gmytrasiewicz and Durfee, 1995], but specifically organized to model psychological factors that play a role in influence, human communication about theory of mind and human social interaction in general. Other agent theories have also addressed the modeling of social interaction, including many based on the Beliefs-Desires-Intentions (BDI) framework [Rao and Georgeff, 1995]. Existing BDI models are based in first-order logic, which can often lead to ambiguous conclusions (e.g., multiple beliefs are consistent with the observations). PsychSim's decision-theoretic foundation supports the quantitative evaluation of beliefs (e.g., computing *degrees* of consistency), as well as the development of novel algorithms for automatic model construction [Pynadath and Marsella, 2004].

Typical work on multi-agent based social simulation has used often thousands or more agents used to study complex systems (e.g. [Terna, 1998]). However, the focus of such work is often the study of the emergent properties of the interactions of simpler agents. For example, Mosler and Tobias use a simulation of 10,000 agents to study the emergence of collective action. Although the agents employed in the work are rather sophisticated by the standards of such simulations, they are nevertheless far simpler than the agents in PsychSim. The agents' decision-making is parametrized by exogenous variables, rather than controlled by deep recursive modeling of others as in PsychSim. Of course, in a simulation with 10,000 agents exploring aggregate properties that emerge from simple behaviors, it would probably not be desirable to give the agents theories of mind and complex lookahead reasoning. Thus, the complexity is in the system, not the individual agents.

Directions and Conclusion

We have presented PsychSim, an environment for multi-agent simulation of human social interaction that employs a formal decision-theoretic approach using recursive models. This approach allows us to model phenomena rarely if at all addressed in simulated worlds. For example, our agents can reason about the behavior and beliefs of other agents. In psychological terms, they have a theory of mind. This allows agents to communicate beliefs about other agent's beliefs, goals and intentions and be motivated to use communication to influence other agents' beliefs about agents. Within PsychSim, we have developed a range of technology to simplify the task of setting up the models, exploring the simulation and analyzing results. This includes new algorithms for fitting multi-agent simulations. Within PsychSim, there is also an ontology for modeling communications about theory of mind. We have exploited the recursive models to provide a psychologically motivated computational model of how agents influence each other's beliefs.

The rich belief structure of PsychSim's agents offers avenues for future exploration. For example, our current reward structures can represent goals on the true state of the world, but it is straightforward to extend it to encompass goals regarding *beliefs* about the world. For example, the bully may have a goal of maximizing the perception that he is powerful, rather than his actual power. We can easily implement such a goal as being proportional to the power that the victim believes the bully to have. We can define many more such goals by applying reward functions to any aspect of the agents' beliefs. These beliefs also include the agents' subjective view of the world dynamics. In our current implementation, the agents all have correct beliefs about the world dynamics. In other words, the bully's beliefs about how beating up the victim affects his power is identical to the true effect. We could easily open up such beliefs about dynamics to the same errors and biases that are possible within the other beliefs. We can also enrich the message space. Certainly, as we expand the belief space to more possibilities for deviation from reality, the more possibility there is for interesting communication between entities with differing perspectives. Furthermore, it is reasonable to enrich the types of performatives supported. Our current messages are intended as statements of fact, but we could easily imagine request messages as well. For example, an agent who has uncertainty in some areas of belief could ask a trusted second agent about its belief in those areas. We believe PsychSim has a range of innovative applications, including computational social science and the model of social training environments. Our current goals are to expand the exploration already begun in the school violence scenario and begin evaluating the application of PsychSim there and in these other areas.

Acknowledgments We would like to thank Dr. Kelton Rhoads for his help in developing the critical factors in PsychSim's model of influence.

References

[Abelson et al., 1968] Abelson, R., Aronson, E., McGuire, W., Newcomb, T., Rosenberg, M., and Tannenbaum, P., editors (1968). *Theories of Cognitive Consistency: A Sourcebook*. Rand McNally, Chicago, IL.

[Cialdini, 2001] Cialdini, R. (2001). *Influence: Science and Practice*. Allyn and Bacon, Boston, MA.

[Gilbert and Troitzsch, 1999] Gilbert, N. and Troitzsch, K., editors (1999). *Simulation for the Social Scientist*. Open University, Buckingham, UK.

[Gmytrasiewicz and Durfee, 1995] Gmytrasiewicz, P. and Durfee, E. (1995). A rigorous, operational formalization of recursive modeling. In *Proceedings of the International Conference on Multiagent Systems*, pp. 125–132.

[Liebrand and Hegselmann, 1998] W. Liebrand, A. N. and Hegselmann, R., editors (1998). *Computer Modeling of Social Processes*. Sage, London, UK.

[Marsella et al., 2000] Marsella, S. C., Johnson, W. L., and LaBore, C. (2000). Interactive pedagogical drama. In *Proceedings of the International Conference on Autonomous Agents*, pp. 301–308.

[Nasby et al., 1979] Nasby, W., Hayden, B., and DePaulo, B. (1979). Attributional biases among aggressive boys to interpret unambiguous social stimuli as displays of hostility. *Journal of Abnormal Psychology*, 89:459–468.

[Petty and Cacioppo, 1986] Petty, R. and Cacioppo, J. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. Springer, New York, NY.

[Prietula and Gasser, 1998] M. Prietula, K, C. and Gasser, L., editors (1998). *Simulating Organizations: Computational Models of Institutions and Groups*. AAAI Press, Menlo Park, CA.

[Pynadath and Marsella, 2004] Pynadath, D. V. and Marsella, S. C. (2004). Fitting and compilation of multiagent models through piecewise linear functions. *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*, to appear.

[Pynadath and Tambe, 2002] Pynadath, D. V. and Tambe, M. (2002). Multiagent teamwork: Analyzing the optimality and complexity of key theories and models. In *Proceedings of the International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pp. 873–880.

[Rao and Georgeff, 1995] Rao, A. S. and Georgeff, M. P. (1995). BDI-agents: from theory to practice. In *Proceedings of the International Conference on Multiagent Systems*, pp. 312–319.

[Schwartz, 2000] Schwartz, D. (2000). Subtypes of victims and aggressors in children's peer groups. *Journal of Abnormal Child Psychology*, 28:181–192.

[Swartout et al., 2001] Swartout, W., Hill, R., Gratch, J., Johnson, W., Kyriakakis, C., LaBore, C., Lindheim, R., Marsella, S., Miraglia, D., Moore, B., Morie, J., Rickel, J., Thiébaux, M., Tuch, L., Whitney, R., and Douglas, J. (2001). Toward the holodeck: Integrating graphics, sound, character and story. In *Proceedings of the International Conference on Autonomous Agents*, pp. 409–416.

[Terna, 1998] Terna, P. (1998). Simulation tools for social scientists: Building agent based models with SWARM. *Journal of Artificial Societies and Social Simulation*, 1(2).

[Whiten, 1991] Whiten, A., editor (1991). *Natural Theories of Mind*. Basil Blackwell, Oxford, UK.