

PsychSim: Agent-based modeling of social interactions and influence

Stacy C. Marsell 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
University of Southern California

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 is a range of applications where rich models of human social behavior are important. We can use them to study social and psychological interactions. To develop a better understanding of the causes and remedies of school bullying, for example, we could use agent models of the students to create simulations of classroom social interactions. We could also use such models as a backend to interactive drama applications, modeling the behavior of the characters in the drama. In particular, one might imagine an interactive pedagogical drama [Marsella et al., 2000] 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.

We have tackled such issues in the context of creating 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 unique aspect of the PsychSim design is that agents have fully specified models of others. Such 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, give our agents the ability to reason about message credibility in a human-like fashion, model the influence 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. 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

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? Were there unintended side effects? Such exploration is ideally part of a three-part brainstorming cycle: (1) hypothesize an approach, (2) test it, (3) and then refine that hypothesis. The cycle facilitates both creative and critical thinking, reveals unexpected phenomena, suggests alternative approaches to explore, and avoids facile answers to problems that rarely have simple (or predictable) answers. Figure 1 shows a screenshot of PsychSim's graphical user interface.

The user interacts with PsychSim the exploratory simulation tool in three overlapping phases. First, there is a design phase where the user populates the simulation scenario with

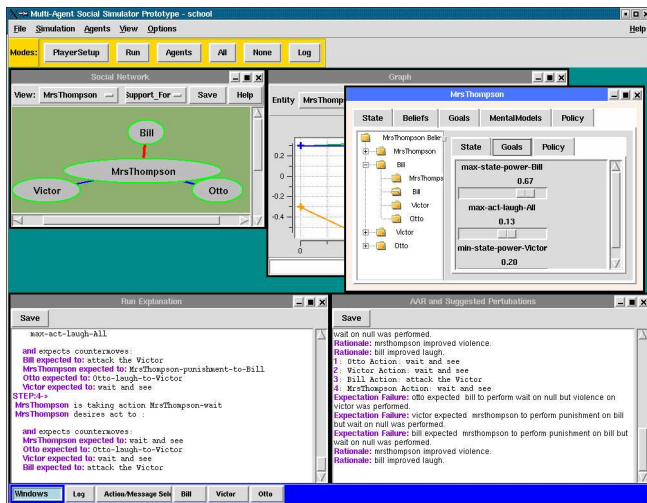


Figure 1: Screenshot of PsychSim interface

autonomous agents that represent the key groups or individuals, selected from a range of provided models, and then specializes those models as needed. Second, the execution phase runs the simulation. Third, the Analysis/Perturbation phase analyzes the simulation run, presents those analyses and suggests perturbations to the models for use in subsequent runs.

In the design phase, the user sets up the simulation by selecting generic agent models that will play the roles of the various groups that will be simulated and specializing those models as needed. A key feature here is the degree of automated assistance and fitting incorporated into the design phase. 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 is a significant improvement in simulation design and promises to greatly simplify simulation setup.

The execution phase allows the user to control the simulation. It 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, the Analysis/Perturbation phase 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. This is an important distinction, because some of these features 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. We used `hardship(agent)` to represent an aggregation of violence suffered by the agent. `trust(truster, trustee)` represents the degree of trust that the agent `truster` has in another agent `trustee's` messages. It is implemented as an average history of whether an agent has accepted (believed) prior messages from another agent. `support(supporter, supportee)` is the strength of support that an agent `supporter` has for another agent `supportee`. Currently, it is an averaged history of whether there is an overlap of shared goals.

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 `punish(teacher, bully)` represents punishment that the teacher gives to the bully. Orders of a superior to a subordinate (such as the principal ordering the teacher to punish the bully) are also treated as acts. Actions affect the state of the world, via a transition function that modifies state features.

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. Each agent's behavior is then motivated by a desire to derive as much benefit as possible. We divide the possible goals into two classes: *features* and *actions*. A goal of **Minimize/maximize** `feature(agent)` corresponds to a negative/positive reward proportional to the given agent's value of the state feature of interest. For example, an agent can have the goal of minimizing another agent's hardship or 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 the bully teases the victim or, more generally, the number of times *any* student teases any other.

An agent's overall reward function is a combination of its set of goals. We model this combination as a weighted sum, with the weight on a particular goal representing its priority to the agent. 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 about what they *think* is the state of the world. Therefore, 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. However, while the simulation-level model of the agents represent the true values for these components, an agent's beliefs consist of that agent's subjective view of the values of these components. 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: Some bullies are reactively responding after an attribution of hostility toward them. Other bullies pursue aggressive behavior as part of a more cold-blooded calculus, believing that aggression will achieve some goal. A key component of the computational model that allows us to model such distinctions is the agent's policy, which represents the process by which it selects an action or message based on its beliefs and goals. PsychSim models this process as a table of "Condition \Rightarrow Action" rules. When executing its policy, the agent tests the pairs until it finds one whose left-hand side matches, in which case it performs the action on the right-hand side. The left-hand side conditions may trigger on an *observation* of some action over some specified preceding time interval or a *belief* of some agent (e.g., such as the bully believing himself as powerful). The conditions may also be more complicated combinations of these basic triggers (e.g., a *conjunction* of conditions that matches when each and every individual condition matches). There is also a special *default* condition that always matches if no other rule applies. 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 typically 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. However, the agents *believe* that the other agents have policies with many direct action rules in them. Thus, the agents do not model each other as doing much lookahead, even though 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. The fact that agents have mental models of others gives a rich, principled basis for message content, because messages can now 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." It may make a claim about the beliefs of the message subject: "the principal believes that he is powerful." It may make a claim about the goals of the message subject: "the bully wants to increase his power." It may make a claim about the policy about the policy table entries of message subject: "if the bully thinks the victim is weak, he will attack 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. We have already 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. A mental model can have supporters, in the form of others that are arguing that this particular mental model is the correct one. For example, the bully could claim that the teacher is seeking dominance, and he could try to convince other students as well, while the teacher could instead counter-claim that she is altruistic.

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 can inform the modeling of such influence, from a computational perspective, they typically suffer from two shortcomings. First, they identify factors that effect 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. For example, there have been many attempts to study factors that lead to influence [Cialdini, 2001], but they leaves open the question of how to operationalize or combine these factors. To provide a sufficient basis for computational models, our approach has been to distill key psychological factors and map those factors into the simulation. We started with some common themes that are apparent in the social psychology literature:

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: People are self-interested. Self-interest impacts how information influences us in numerous ways. It impacts how we interpret appeals to one's self-interest, emotions or 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 argue that these factors can in fact be modeled by a few simple mechanisms in the simulation. The factors are reduced to evaluating *consistency*, *self-interest*, and *bias*, which we can render computationally operational as follows.

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. The more matches a set of beliefs generates, the higher its 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 through a projection into the future within the context of the agent's beliefs. The agent then evaluates the reward it expects to re-

ceive in that forward projection. The set of beliefs that generates the higher reward for the agent has a higher degree of self-interest. 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.

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). We see the computation of these factors as a toolkit for the user to explore the system's behavior under different assumptions or models of how these factors interact. One can imagine variations consistent with specific theoretical models.

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 can illustrate how one might use PsychSim to explore psychological theories by demonstrating how PsychSim can represent the possible entities in a simulation of school violence. Because the agents in PsychSim have a theory of mind, they are capable of modeling both *attributional style* and *outcome expectancies*. The user can then manipulate each factor in a variety of ways, thus generating 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. For example, agents with a hostile attributional style could mentally model another student as having the goal of hurting them (i.e., minimizing their power) and a policy of aggression. A student's attributional style can also manifest itself in the beliefs that s/he has about the other students and their beliefs about him/her. For example, a student may believe that the other students view him/her as weak. Our PsychSim agents can represent such a belief within our recursive agent models of `believes(bully, believes(victim, power(bully)))`.

Our agents already compute the second factor of interest, outcome expectancies, as part of their bounded lookahead. Thus, when considering possible aggression, the agents consider the immediate effect of an act of violence, as well as the possible consequences. The agents evaluate these effects to determine whether an act of aggression will improve themselves with respect to their goals. An agent considers the immediate effect of an act of aggression according to its beliefs about the dynamics of the world. In particular, it considers the change in the state of the world (e.g., the relative power values of itself and its victim), as well as the change in the beliefs of the other agents. For example, a bully may have incentive to perform an act of aggression because his classmates will then believe him to be strong. Our bully agent can model such an incentive as a goal of maximizing `believes(classmates, power(bully))`, as well as a belief that an act of aggression will increase `believes(classmates, power(bully))`. Alternatively, the bully may be motivated by a desire for popularity, which we would represent as a goal of maximizing `support(classmates, bully)`. The bully may believe that his classmates dislike the victimized student or, more precisely, that they have a goal of minimizing `power(victim)`. Following PsychSim's model of the dynamics of support relationships, the bully would believe that an act of aggression against the victim would decrease `power(victim)`, which helps achieve his classmates' goals, which, in turn, would lead them to his popularity.

In addition to this immediate effect of aggression, our agents' bounded lookahead also considers the actions that the other agents may take in response. The agents' theory of mind is crucial here, because it allows them to predict these responses, albeit limited by the subjective view of the predicting agent. 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.

It is just as easy for PsychSim's agents to model reactive aggression. In the reactive case, the agent simply retaliates for any perceived injury, without computing an outcome ex-

pectancy. In other words, the agent's policy includes a simple rule, whose trigger is an observed action by another agent which hurts it (i.e., negatively impacts its goals), and which specifies an act of aggression as a response. This rule short-circuits the bounded lookahead, while still allowing the agent to reason about outcome expectancies under those circumstances where there has been no triggering victimization.

Given PsychSim's ability to represent factors like attributional style and outcome expectancies within its theory of mind, we can define classes of entities using different settings for those factors. We can create an agent model for a nonaggressive victim by using a hostile attributional style and a negative outcome expectancy for aggression. By removing the negative outcome expectancy and adding a policy for reactive aggression, we change our agent model into that of an aggressive victim. By giving the agent a positive outcome expectancy for aggression and removing the hostile attributional style, we create a model of a nonvictimized aggressor. We can thus quickly instantiate models covering a wide range of possible student behaviors.

We can construct diverse social scenarios by using different combinations of these instantiated models to populate our simulation environment. Thus, we can systematically explore the space of possible class makeups to understand the social behavior that arises out of different configurations of student psychologies. Through this exploration, we can potentially find the configuration that best matches a real-world class dynamic, allowing us to find an underlying psychological explanation for a specific instance of behavior. Furthermore, we can try out different interventions in the simulation to understand their impact under varying student models. Each alternate scenario 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. Multi-agent based social simulation can be roughly divided into two classes. On one hand, there is a large body of work in modeling teams or organizations, such as the STEAM teamwork model [Tambe, 1997], where each agent has an explicit model of the team plan, its distinct role in the plan, commitments to communicate, and sophisticated problem-solving capabilities. Although STEAM has been applied to simulate a platoon of human operated helicopters, it has not been concerned with modeling a wide range of human social interaction.

In contrast, there are relatively massive simulations of 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 so-

cial training environments. Our current goals are to explore and evaluate the application of PsychSim to these 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

- [Cialdini, 2001] Cialdini, R. (2001). *Influence: Science and Practice*. Allyn and Bacon, Boston, MA.
- [Dodge and Coie, 1987] Dodge, K. and Coie, J. (1987). Social information processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology*, 53:389–409.
- [Gmytrasiewicz and Durfee, 1995] Gmytrasiewicz, P. and Durfee, E. (1995). A rigorous, operational formalization of recursive modeling. In *Proceedings of the First International Conference on Multiagent Systems, ICMAS95*, pages 125–132.
- [M. 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.
- [Marsella et al., 2000] Marsella, S. C., Johnson, W. L., and LaBore, C. (2000). Interactive pedagogical drama. In *Proceedings of the Fourth International Conference on Autonomous Agents*, pages 301–308, New York. ACM Press.
- [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.
- [Pynadath and Marsella, 2004] Pynadath, D. V. and Marsella, S. C. (2004). Fitting and compilation of multiagent models through piecewise linear functions. Submitted to the International Conference on Autonomous Agents and Multiagent Systems.
- [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 First International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 873–880, New York. ACM Press.
- [Schwartz, 2000] Schwartz, D. (2000). Subtypes of victims and aggressors in children's peer groups. *Journal of Abnormal Child Psychology*, 28:181–192.
- [Tambe, 1997] Tambe, M. (1997). Towards flexible teamwork. *Journal of Artificial Intelligence Research*, 7:83–124.
- [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.