Coordinate to cooperate or compete: Abstract goals and joint intentions in social interaction

Max Kleiman-Weiner[1] (maxkw@mit.edu), Mark K. Ho (mark_ho@brown.edu)[2],
Joseph L. Austerweil[2] (joseph_austerweil@brown.edu), Michael L. Littman[3] (mlittman@cs.brown.edu),
Joshua B. Tenenbaum[1] (jbt@mit.edu)
[1]Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139
[2]Cognitive, Linguistic, and Psychological Sciences / [3]Computer Science, Brown University, Providence, RI 02912

Abstract

Successfully navigating the social world requires reasoning about both high-level strategic goals, such as whether to cooperate or compete, as well as the low-level actions needed to achieve those goals. While previous work in experimental game theory has examined the former and work on multi-agent systems has examined the later, there has been little work investigating behavior in environments that require simultaneous planning and inference across both levels. We develop a hierarchical model of social agency that infers the intentions of other agents, strategically decides whether to cooperate or compete with them, and then executes either a cooperative or competitive planning program. Learning occurs across both high-level strategic decisions and low-level actions leading to the emergence of social norms. We test predictions of this model in multi-agent behavioral experiments using rich video-game like environments. By grounding strategic behavior in a formal model of planning, we develop abstract notions of both cooperation and competition and shed light on the computational nature of joint intentionality.

Keywords: joint intention, cooperation, coordination, reinforcement learning, teams

Introduction

Our most important relationships involve understanding when to cooperate and when to compete. From siblings to coworkers, humans rely on both planning and context to know which situations they should cooperate in and which they should compete in (Galinsky & Schweitzer, 2015; Rand & Nowak, 2013). And yet in real life, unlike a behavior economics experiment, cooperation and competition are abstract with respect to a given situation. A cooperative or competitive interaction unfolds over time – there isn't a single moment where competition or cooperation "happens". Even if the decision to cooperate or compete has been made, efficiently implementing those strategies can be difficult. A person determined to cooperate and knowing what the other person wants will have to develop a detailed plan of action to realize that cooperative intention. Likewise for a person intent on competing. In this work we aim to bridge high-level strategic decision making over abstract social goals such as cooperation and competition with low-level planning over actions to actually realize those goals.

The ability to form these hierarchical joint intentions is a key component of social behavior. The motivated instinct to both infer and evaluate complex social plans emerges in early childhood (Warneken & Tomasello, 2006; Hamann, Warneken, Greenberg, & Tomasello, 2011). Young children not only rapidly infer the goals of other agents, but spontaneously execute complex plans to cooperate with others. For instance, a cooperative intention might generalize to include not just the low-level details of a joint task but also tell how to share the spoils. The ability to infer the intentions of oth-

Matrix-Form Games

		Yellow	
		Cooperate	Compete
Blue	Cooperate	7,7	-1,8
	Compete	8, -1	4,4

Figure 1: A social dilemma written as a normal-form game. The numbers in each square specify the payoff in terms of utility to the blue and yellow player respectively for choosing the action corresponding to that square's row and column. If both agents choose cooperate they will collectively be better well off than if they both choose compete. However in any single interaction, either agent would be materially better off by choosing to compete.

Stochastic Games

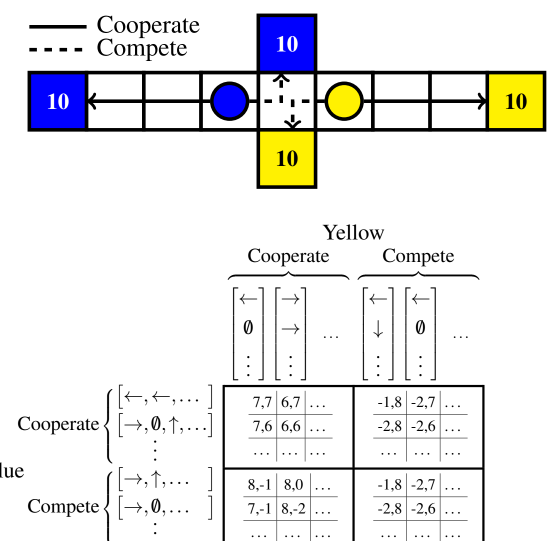

Figure 2: Two-player stochastic games. (top) Grid form representation of the stochastic game. The arrows show example strategies that can be used to realize both cooperative and competitive outcomes. (bottom) Matrix representation of the strategy space, with low-level strategies sorted by a high-level goal. The arrows correspond to moving in a specific direction and the ∅ corresponds to waiting. Note that the action space is effectively unbounded but the strategies naturally cluster into a small number of high-level goals. If both agents go to the sides then they will both score the reward but if they fight for the middle in hopes of using less moves they will collide and only one will get any reward.

ers and participate in a dynamic joint endeavor (sometimes called the "we-mode") is thought to be a key building block of large scale collaborative culture (Tomasello, Carpenter, Call, Behne, & Moll, 2005).

Naturalistic Games

Game-theoretic investigations of social behavior often represent strategic interactions as *matrix-form* games like the one shown in Figure 1. In these games, the rows and columns correspond to the action space of the two players and the cells describe the payoffs to each agent that would result from those actions. While useful as a succinct representation of a social decision, these games lack the ecological validity of real social decisions which require planning across space and time. When presented to participants, it can be difficult to extract the right information and even after significant training, many people don't even look at the payoffs most relevant for strategic reasoning (Costa-Gomes, Crawford, & Broseta, 2001). When the number of decisions grows beyond two decisions per agent, these problems are exacerbated.

Instead we use a paradigm commonly deployed in multi-agent systems research which has not been explored behaviorally (De Cote & Littman, 2008). In this paradigm, strategic interactions are represented as naturalistic spatial environments that people play intuitively like video-games. Figure 2 shows an example of one of these multi-agent planning environments that is conceptually related to the social dilemma shown in Figure 1. Unlike the matrix-form game, these environments also require low-level planning over spatial actions to realize a strategic goal. The action space of these games is much larger than those typically studied in matrix-form games but the strategies are still intuitive.

Each player controls the movement of one of the colored circles. On each turn players choose to either move their circle into an adjacent square (not including diagonal moves) or to remain in the same position. Attempting to move is costly resulted in the loss of one point. Choosing to remain in the same position did not incur any cost. Both players select an action during the same turn and their positions are updated simultaneously. Each square can only be occupied by one player at a time so if both players try to move to the same square, one of the players chosen by chance will enter the contested square while the other remains in place. However both pay the cost for attempting to move. If one player stays in the same position and the other player tries to move into their square, no movement occurs. Finally, players cannot move through each other and switch places.

The colored squares are the goals. When either player reaches a square with the same color as their avatar, that player receives ten points and the round ends. Thus the only way for both players to receive points is if they both enter squares that match their avatar's color on the same turn. These dynamics were chosen to be identical to those in (De Cote & Littman, 2008) so that our data can also be compared to the models of that work. Because each interaction generates data about both the action plan and the payouts, we can use these games to start to investigate the mechanisms people use to coordinate on cooperative and competitive outcomes. Furthermore, they allow us to study how humans innovate to find these strategies out of such a large possible space of action plans.

Model

Hierarchical Social Planning

We develop a hierarchical model of strategic planning that unifies low-level action planning with high-level strategic reasoning and allows for learning across both levels. In brief, agents have two "modes" of low-level planning: a cooperative mode and a competitive mode. These two modes are connected through a high-level strategic planner that determines which mode should be deployed based on previous interactions. After each round, agents use Bayesian theory-of-mind to determine whether or not the other agent's low-level actions are consistent with the cooperative planning mode vs. the competitive planning mode. The agent can then condition their own next actions on the inferred high-level intentions of the other agent realizing a sophisticated strategic response.

Both modes include forms of model-based learning which allows for learning to generalize across environments as well as model-free reinforcement of actions. In this work we focus specifically on the high-level goals of cooperation and competition but other high-level goals such as teaching, punishing or communication are also relevant in these games and will be investigated in future work. The challenge of hierarchical planning is to link these high-level goals to a lower-level plan of action.

Our work builds on and is inspired by classical formalisms of intention and joint planning from the AI literature (Levesque, Cohen, & Nunes, 1990; Grosz & Kraus, 1996) as well as more modern formulations for planning under uncertainty such as DEC-POMDPs and I-POMDPs (Gmytrasiewicz & Doshi, 2005; Gal & Pfeffer, 2008; De Cote & Littman, 2008). However the earlier models do not handle uncertainty in a probabilistic way and hence struggle with quantitative predictions about behavior while the later are often intractable over long planning horizons and don't explicitly represent abstract social goals.

We briefly introduce stochastic games following the notation of De Cote and Littman (2008) and then discuss repeated stochastic games. A two-player stochastic game is: $\langle S, s_0, A_1, A_2, T, U_1, U_2, \gamma \rangle$ where S is the set of all possible states with $s_0 \in S$ the starting state. Each agent can choose from a set of actions A_1 and A_2 which together form a joint action space $A_1 \times A_2$. The state-transition function, $T(s, a_1, a_2) = P(s'|s, a_1, a_2)$ maps a state and joint action to a distribution over new states. The utility functions of the two agents $U(s', s, a_1, a_2) = R$ describe the agent's goals in terms of quantitative costs and rewards. Finally $0 \leq \gamma_{\texttt{game}} \leq 1$ is the discount rate of reward. In repeated stochastic games, a series of stochastic games are played one after another in succession between the same pair of players. We now discuss the cooperative and competitive modes of planning in detail.

Cooperative Planning

Since there is no specific action that corresponds to cooperation in these stochastic games (all actions are spatial movements), we develop an abstract notion of cooperation which generalizes across contexts. We postulate that a cooperative action is one that is good for the group i.e., efficiently maximizes the utility of all agents. Since under this assumption, the goal of cooperation is to rationally achieve a group goal, we consider a *group-agent* that optimizes a utility function composed of the utility of all agents (Sugden, 1993, 2003; De Cote & Littman, 2008).

Computationally, we represent this group utility function as a linear weighting of the utility of the two agents: $U^G = (w)U_1 + (1-w)U_2$ where $w \in [0,1]$ controls how the two agents are relatively valued by the group-agent. For example when $w = 0.5$ the group-agent impartially weighs the utility of both agents equally. We are not implying that this group-agent actually exists but rather that each player can simulate the same group-agent by taking an objective view of the planning environment outside and separate of their own personal goals (Nagel, 1986). We note that this utility function can include other social preference such as inequality aversion or merit based allocations.

Since the group-agent can directly control the actions of both players (like a "we" agent), it can treat the stochastic game as a single-agent MDP. Rational planning over joint actions (a_1, a_2) is achieved through value-iteration:

$$P(a_1,a_2|s) = \pi^G(s,a_1,a_2) \propto e^{(\beta Q^G(s,a_1,a_2))}$$
$$Q^G(s,a_1,a_2) = \sum_{s'} P(s'|s,a_1,a_2)[U^G(s',s,a_1,a_2) + \gamma \max_{(a'_1,a'_2)} Q^G(s',a'_1,a'_2)]$$

where the group-agent policy, $\pi^G(s)$, is to choose actions with probability proportional to their future expected utility. A high value of β means that the group-agent is more likely to select the action with the highest Q-value and a low value of β means that the group-agent is more likely to select suboptimal-actions. In all experiments we used a relatively high value of $\beta = 4$. We note that π^G is not only a policy for action, but also includes the future-oriented intentions of what the two agents *should* do once they get to a new state. These intentions include how to recover from failed coordination attempts. We used a discount rate of $\gamma = 0.9$ in all the models presented here.

Although each agent might consider the policy of the group-agent, the individual agents can only control their own actions. To transform this group-agent policy into an individual policy, individual agents marginalize out the actions of the other player from the joint policy: $\pi_1^G(s,a_1) = \sum_{a_2} \pi^G(s,a_1,a_2)$ and $\pi_2^G(s,a_2) = \sum_{a_1} \pi^G(s,a_1,a_2)$. These policies contain intertwined intentions, not only an *intention to* take a specific action but also the *intention that* the other agent reach certain states. This "meshing" of plans between the two agents has been called a key component of joint and shared intentionality (Bratman, 1993, 2014). Unlike social preference based accounts of cooperative behavior where each agent individually plans to maximize joint utility, in this account, cooperation is a built in cognitive feature of planning itself – agents *plan together*.

When there is a single unambiguous action for both players that maximizes joint utility, coordination is readily achieved. However in the environments we investigate, there are often multiple actions that can generate optimal rewards for the group-agent. We now discuss two mechanisms for learning social norms that can break these symmetries and lead to robust coordination on a single jointly optimal plan.

We first consider the case where two different actions are equally good from the perspective of a group-agent that weighs the utility of the two agents equally but the rewards will be allocated unequally. For example, consider game (C) in Figure 3 where one agent needs to go around the other. Because moving costs 1 point, the agent who goes around the other will only earn 7 points while the agent who waits will earn 9 points. From the perspective of the group-agent with $w = 0.5$, it doesn't matter who goes around since the joint utility is equal. However if one agent was favored over the other ($w \neq 0.5$) this symmetry would be broken and the disfavored agent would take the long route. Thus prior knowledge about asymmetries in how the group should operate can lead to more robust coordination although potentially at the cost of less fair cooperation.

The two agents may start with a different prior on the value of w and thus when simulating the group-agent will fail to coordinate. Consider the case where both agents think they should be valued more than the other and hence expect the other player to go around them. We propose a mechanism based on "virtual bargaining" accounts of social choice that lead to each agent's w to converge over time to the same value without any explicit communication (Binmore, 1998; Misyak, Melkonyan, Zeitoun, & Chater, 2014). After each interaction, agents can infer the w that best explains the joint behavior of their previous interaction: $P(w|H) \propto P(H|w)P(w)$ where H are the data from previous interactions and the likelihood of those interactions is defined by the marginalized joint policies generated from planning with a specific w: π_1^G and π_2^G. In our analysis, each agent starts out with a prior of $w = 0.5$ and updates it after each round based on the inferred w of the previous interaction. Thus over time w will converge and as predicted by the theory of virtual bargaining, more patient agents who insist on the advantage will gain a greater share of the joint reward in future coordinated interactions where an equitable split isn't possible. For example, if in a previous interaction agent 1 took a more costly route, then in the next round agent 1 will be more likely to take the costly route again generating a social norm for cooperative coordination. Since w is an input to the planning process itself, it allows for generalizing these norms to new environments.

Finally, in some environments, there are multiple plans that are equally good for both agents, creating a different type of symmetry which cannot be broken by w. For example, the decision to go clockwise or counterclockwise in game (A) of Figure 3 is equally good for both players as long as they both go in the same direction. To capture the intuition that once agents successfully coordinate, they should continue to coordinate in that way e.g., after luckily choosing to go clockwise in game (A), they will go clockwise again on the next round, agents learn a function $N^G(s, a_1, a_2)$ based on the frequency of previous joint actions which is added to the state-action Q^G-value used by the group-agent. This norm based reinforcement affects the policies of the individual agents through marginalization. The norms reinforced by this mechanism do not generalize across environments although feature based norms can generalize when there are features in common between two environments e.g., see Ho et al. in this years proceedings.

Competitive Planning

As before, in these stochastic games there is no action that directly corresponds to "compete". Instead, we ground competitive planning as each agent attempting to maximize their individual utility under the assumption that the other agent is doing the same. To tractably realize this game-theoretic best-response, we extend the cognitive hierarchy / level-K formalism used in behavioral game theory to temporally extended polices instead of just actions (Camerer, Ho, & Chong, 2004). In brief, a level-K agent best responds to a level-$(K-1)$ agent which grounds out in the level-0 agent. Specification of the level-0 agent is sufficient to specify the full hierarchy.

In this work we use a level-0 agent that doesn't consider the existence of the other player and tries to efficiently reach her goal without taking any strategic consideration of how the other player might affect her progress. This level-0 agent is more naturalistic than randomly acting agents which are commonly used in behavioral modeling (Camerer et al., 2004; Yoshida, Dolan, & Friston, 2008). A level-0 agent of this type only makes sense in these naturalistic environments since one can easily imagine acting alone unlike in matrix-form games. The level-0 agent for player i is:

$$P(a_i|s, k=0) = \pi_i^0(s) \propto e^{\beta Q_i^0(s, a_i)}$$
$$Q_i^0(s, a_i) = \sum_{s'} P(s'|s, a_i)(U_i(s, a_i, s') + \gamma \max_{a_i'} Q_i^0(s', a_i'))$$

where $P(s'|s, a_i)$ represents transition dynamics that do not depend on the other player. Having defined the level-0 player we can recursively define all of the other levels in the hierarchy in terms of lower levels:

$$P(a_i|s, k) = \pi_i^k(s) \propto e^{\beta Q_i^k(s, a_i)}$$
$$Q_i^k(s, a_i) = \sum_{s'} P(s'|s, a_i)(U(s, a_i, s') + \gamma \max_{a_i'} Q_i^k(s', a_i')))$$

Since the other agent is treated as a knowable stochastic part of the environment, the dynamics of the other player are encapsulated in $P(s'|s, a_i)$ which are marginalized out using the $k-1$ player: $P(s'|s, a_i) = \sum_{a_{-i}} P(s'|s, a_i, a_{-i}) P(a_{-i}|s, k = k-1)$ where $-i$ is a shorthand to refer to the "other" player. Because of the maximization operator, a level-K agent implements a best response to a level-$K-1$ agent. Thus zeroth-order agents have their own goals but ignore the other player, first-order agents act on their own goals but assume that the other agent is ignoring their existence and so on. In our experiments we used $K = 1$ although results were similar with higher values of K.

Even when competitively planning, agents can still improve their behavior through learning and can even develop certain conventions when they serve mutual self-interest such as symmetry breaking in coordination games. Again we consider two mechanisms. The first mechanism improves agent i's model of agent $-i$ by using the frequency of i's previously successful behavior to modify the state-action Q-values of $-i$ such that previously successful action are more likely to occur again. This model-based mechanism, improves agent i's policy since she will best-respond to a more accurate model of agent $-i$. The second mechanism is model-free reinforcement of player i's state-action Q-values when player i herself successfully reaches a goal. Neither of these norms trivially generalize across different planning environments that don't share states.

Coordinating Cooperation and Competition

Finally, we describe how agents can use both the cooperative and competitive modes of planning to decide whether to cooperate or compete. Since these modes of planning abstract away the details of cooperation and competition, high-level strategic planning can use these low-level planners without considering their details. Agents first use these planning modes to infer the high-level intention I of the other player (i.e., their planning mode) using Bayesian theory-of-mind: $P(I|D) \propto P(D|I)P(I)$ where $P(D|I)$ are just the cooperative or competitive policies. This probabilistic approach is justified because intentions can be ambiguous. For instance, when both agents reach the goal in a coordination game it could just be because of luck so the behavior isn't very diagnostic of the intention. Yet in social dilemma only the cooperative intention is consistent with behavior where both reach the goal. Using these inferred strategic intentions, a high-level planner can take a simple and intuitive form such as reciprocal cooperation (e.g., tit-for-tat) or reinforcement learning at the level of strategy rather than actions (Fudenberg & Levine, 1998).

Behavioral Experiments

We developed client/server software that allows for real-time interactions between two participants randomly matched through mTurk. All participants went through a short single player tutorial that familiarized them with the controls of the games, the dynamics of the game environment, the costs of movement and value of the goals. After the tutorial, pairs of

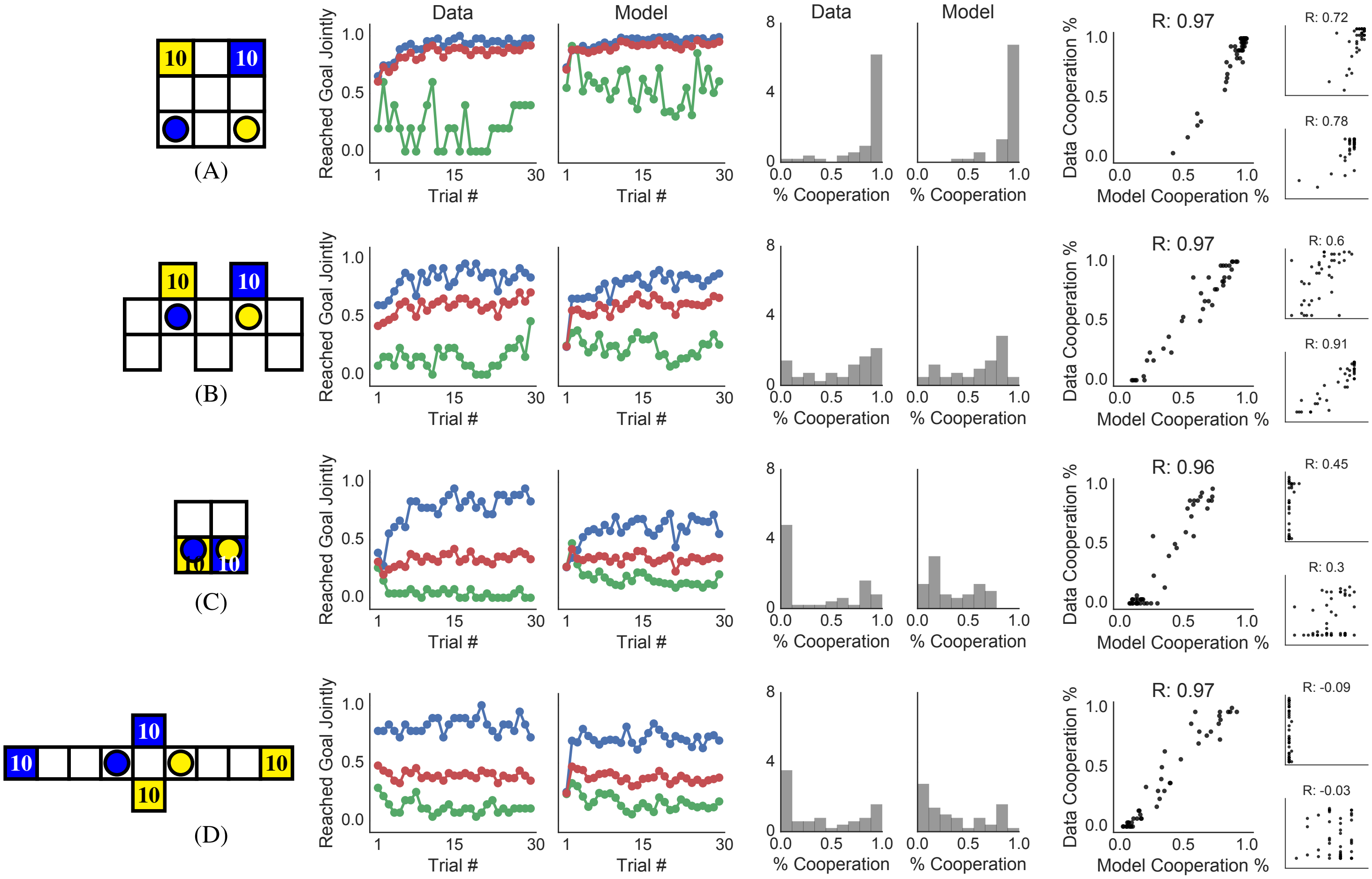

Figure 3: Participant data and model predictions for four environments. Each row shows data and model predictions for the environment in column 1 which was repeated 30 times. Rows 1 and 2 are coordination games and rows 3 and 4 are social dilemmas. Column 2 shows the average rate of cooperation for each round of play averaged over the high-cooperating cluster of participants (blue), low-cooperating cluster of participants (green) and all participants (red). Column 3 are histograms of the proportion of cooperation for all pairs of participants. Column 4 quantifies the model predictions where each point represents the frequency of cooperation for a given dyad observed in the data and as predicted by the model. The inset shows correlations of the two lesioned models with the same human data: (top) only compete (bottom) only cooperate.

participants were matched together and played 30 rounds of the same game with the same partner. Subjects were not told the exact number of rounds they would play together in order to prevent horizon effects from backward induction. Once both participants submitted moves, the game state and score were updated and the process continued until the end of the round. Participants had 30 second for each move and the game ended if a participant exceeded their 30 second time bank two moves in a row. We only analyzed data from complete interactions where the pair of participants completed all 30 rounds of the game together. All experiments were incentivized with bonuses proportional to the number of points accumulated.

To compare model predictions with human behavior, we first focused on analyzing whether or not both players reached a goal on a given round, a behavioral signature of cooperation in these games. For each pair of participants, the model observes the interaction in the previous rounds, performs inference on the latent high-level goal and social norms, and samples a prediction for the behavior of the pair in the next round. We compare this sampled prediction with actual human behavior to assess model performance. The same model parameters were used for all pairs of participants.

Figure 3 shows the results of the behavioral experiments and the model predictions for four environments (≈ 50 participant pairs per environment), two coordination games and two social dilemma. Since model predictions were made at the level of each pair of participants, averaging the behavior and model predictions across dyads obscures individual differences in the dynamics of cooperative and competitive learning. To investigate the model predictions in a more fine-grained way, we used unsupervised clustering to split the pairs of participants into two group. In short, for each pair of participants we construct a 30-dimensional binary vector where each dimension corresponds to one of the 30 rounds. Each element is set to one if both participants reached a goal in the round corresponding to that dimension and set to zero otherwise. We ran K-means clustering with $K = 2$ which split the data into a high-cooperating cluster and a low-cooperating cluster allowing for better visualization of the data and model prediction and gave some rough indication about the model ability to handle individual differences.

In all four environments, some of the pairs converged on a cooperative plan but the incentive structure of the game i.e., whether or not the game was a coordination game or social dilemma affected the likelihood that both participants jointly reached a goal. Overall, participants jointly reached the goal more frequently in coordination games than in the social dilemma. As shown in Figure 3 the model qualitatively captures the rate of cooperation and competition in both the high-cooperating cluster and the low-cooperating cluster as well as the average over all participants. Another coarse measure of behavior in these games is the distribution of the frequency of cooperative behavior across pairs of participants. In coordination games, the distribution was left-skewed and in social dilemma the distribution was right-skewed. These distributions were captured both qualitatively and quantitatively across these games by the model.

We compared the full model which included both modes of planning and strategic reasoning over those two modes with two lesioned models which just used one of the two planning modes. One lesioned model always used the competitive planning mode and the other lesioned model always used the cooperative planning mode. Overall, neither lesioned model could capture the rates of cooperation between the two clusters and qualitatively failed to explain the distribution of cooperative behavior in each game. Both lesioned models failed to predict the dynamics of strategic reasoning between cooperation and competition in social dilemma and had weaker correlation with participants' behavior in the coordination games.

Discussion

In this work we developed a hierarchical model of social planning to understand how humans coordinate their low-level action plans to realize high-level strategic goals such as cooperation and competition. We formalize cooperation and competition as abstract planning procedures over low-level actions. Both model-based and model-free learning can create social norms which facilitate robust and stable coordination. One of our main contributions is formalizing how cooperative norms can make cooperation more robust across environments, a key step for long-lasting collaborative endeavors. While we only had space to show a subset of our full results, we are currently looking at how agents use these planning programs and the norms that they learn to generalize cooperation to completely new environments with the same partner. We will also use these models to study how observers attribute cooperative and competitive intentions to other agents.

One interesting feature of the model is how an asymmetric w in the cooperative planner can break symmetries making successful coordination more likely. In future work we'd like to explore how priors on this parameter in social hierarchies might enable more effective teamwork e.g., boss-employee relations (Galinsky & Schweitzer, 2015). Finally, in our current paradigm, the desires of all agents are common knowledge. Investigating environments that require jointly inferring the goals of others and the plan needed to help realize a cooperative outcome will be examined in future work. By grounding strategic social reasoning in a theory of planning we can begin to investigate the mechanisms of joint intentionality and how these joint intentions enable the scale and scope of human cooperative behavior (Tomasello, 2014).

Acknowledgement This work was supported by a Hertz Foundation Fellowship, NSF-GRFP, the Center for Brains, Minds and Machines (CBMM), NSF STC award CCF-1231216 and by an ONR grant N00014-13-1-0333. We thank Alejandro Vientós, Banti Gheneti, and Paul Masterson.

References

Binmore, K. G. (1998). *Game theory and the social contract: just playing* (Vol. 2). Mit Press.
Bratman, M. E. (1993). Shared intention. *Ethics*, 97–113.
Bratman, M. E. (2014). *Shared agency: A planning theory of acting together*. Oxford University Press.
Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 861–898.
Costa-Gomes, M., Crawford, V. P., & Broseta, B. (2001). Cognition and behavior in normal-form games: An experimental study. *Econometrica*, 1193–1235.
De Cote, E. M., & Littman, M. L. (2008). A polynomial-time nash equilibrium algorithm for repeated stochastic games. In *24th conference on uncertainty in artificial intelligence.*
Fudenberg, D., & Levine, D. K. (1998). *The theory of learning in games* (Vol. 2). MIT press.
Gal, Y., & Pfeffer, A. (2008). Networks of influence diagrams: A formalism for representing agents' beliefs and decision-making processes. *Journal of Artificial Intelligence Research*, *33*(1), 109–147.
Galinsky, A., & Schweitzer, M. (2015). *Friend and foe: When to cooperate, when to compete, and how to succeed at both*. Random House.
Gmytrasiewicz, P. J., & Doshi, P. (2005). A framework for sequential planning in multi-agent settings. *J. Artif. Intell. Res.(JAIR)*, *24*, 49–79.
Grosz, B. J., & Kraus, S. (1996). Collaborative plans for complex group action. *Artificial Intelligence*, *86*(2), 269–357.
Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature*, *476*(7360), 328–331.
Levesque, H. J., Cohen, P. R., & Nunes, J. H. (1990). On acting together. In *Aaai* (Vol. 90, pp. 94–99).
Misyak, J. B., Melkonyan, T., Zeitoun, H., & Chater, N. (2014). Unwritten rules: virtual bargaining underpins social interaction, culture, and society. *Trends in cognitive sciences*.
Nagel, T. (1986). *The view from nowhere*. Oxford University Press.
Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in cognitive sciences*, *17*(8), 413.
Sugden, R. (1993). Thinking as a team: Towards an explanation of nonselfish behavior. *Social philosophy and policy*, *10*(01), 69–89.
Sugden, R. (2003). The logic of team reasoning. *Philosophical explorations*, *6*(3), 165–181.
Tomasello, M. (2014). *A natural history of human thinking*. Harvard University Press.
Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and brain sciences*, *28*(05), 675–691.
Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*(5765), 1301–1303.
Yoshida, W., Dolan, R. J., & Friston, K. J. (2008). Game theory of mind. *PLoS Computational Biology*, *4*(12).