Incorporating Emotion Perception into Opponent Modeling for Social Dilemmas

Rens Hoegen, Giota Stratou and Jonathan Gratch
University of Southern California, Institute for Creative Technologies
12015 Waterfront Dr., Playa Vista, CA 90094
{rhoegen, stratou, gratch}@ict.usc.edu

ABSTRACT

Many everyday decisions involve a social dilemma: cooperation can enhance joint gains, but also make one vulnerable to exploitation. Emotion and emotional signaling is an important element of how people resolve these dilemmas. With the rise of affective computing, emotion is also an important element of how people resolve these dilemmas with machines. In this article, we learn a predictive model of how people make decisions in an iterative social dilemma. We further show that model accuracy improves by incorporating a player's emotional displays as input to this model, and provide some insight into which emotions influence social decisions. Finally, we show how this model can be used to perform "social planning": i.e., to generate a sequence of actions and expressions that achieve social goals (such as maximizing individual rewards). These techniques can be used to enhance machine-understanding of human behavior, as social decision-aids, or to drive the actions of virtual and robotic agents.

CCS Concepts

•**Computing methodologies** → **Machine learning; Modeling and simulation;** *Agent / discrete models;* Cognitive science;

Keywords

agent-based analysis of human interactions, affective behavior, virtual agents for social dilemmas, player modeling, game theory, emotion in social dilemmas

1. INTRODUCTION

Social dilemmas are a big part of our everyday lives. Competing nations deciding to either invest in weapons or reduce defense spending (disarm) [19], competing economies curbing CO_2 emissions to avoid global impact on climate change, athletes opting to use illegal performance enhancing drugs; all of the above are examples of competitive social dilemmas where joint cooperation has the most benefits long-terms but defecting can give one competitor the clear advantage. In those cases, it is important to be able to predict cooperative intentions so that we can opt to work together towards joint gains or issue corrective (punitive) actions when those are necessary.

However, predicting people's decisions in social dilemmas is a complex and multi-faceted problem. Even in simple two-person dilemmas, people don't usually make rational decisions [18, 17] and are influenced among other factors by affect, communication with their partners, and elements of reciprocity [14, 26, 10]. We are particularly interested in the role that emotion plays in how people solve competitive social dilemmas. Prior research has shown 1) that emotional expressions can telegraph a person's cooperative tendencies and 2) that observers use these expressions to inform their own decision-making [10, 15, 7, 29, 27, 24, 4]. For example, in the prisoner's dilemma, Brosig and colleagues found that observers could predict if a person would cooperate by observing their expressions [4] and Stratou and colleagues used automatic expression analysis to show that smiles predicted these cooperative tendencies [27]. Other work has examined how expressions shape observers decisions. For instance, de Melo and colleagues had participants play the iterated prisoner's dilemma with an expressive agent [7]. They showed that agents that displays guilt after exploiting a participant elicits greater cooperation from participants than an agent that smiles. Krumhuber and colleagues showed that temporal dynamics of a smile can also shape decisions in a two-player trust game [15]. Even the lack of emotion can be an important signal, as Schug and colleagues showed that inexpressive opponents are viewed as untrustworthy [24]. Finally, Stratou and colleagues [27] showed that untrustworthy players can take advantage of their opponents by attending to their expressions. For instance, they showed that smiling players were more likely to be exploited in an iterated prisoner's dilemma, because positive emotions signal cooperation.

Despite this extensive research, there are still limitations in our understanding of these signaling processes. Many of the aforementioned findings gloss over the moment-to-moment patterns of expressions. For example, Stratou and colleagues [27] treat a multi-round prisoner's dilemma game as a single unit of analysis and examine how the frequency of expressions over this entire interaction is correlated with cooperation rate. Research that focuses on the moment-to-moment dynamics has done so using artificial emotional displays and scripted interactions (e.g.,[7]), which may not be suitable for modeling emotional exchanges in real human-human interactions. In general, there is limited work on computational models that actually predict a human oppo-

Appears in: *Proc. of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), S. Das, E. Durfee, K. Larson, M. Winikoff (eds.), May 8–12, 2017, São Paulo, Brazil.*

nent's decision in real-time based on their immediate reactions (as one of the few exceptions, Park and colleagues [21] trained a model that could predict whether an opponent would accept or reject offers in a multi-issue bargaining task).

In this paper, we develop a real-time prediction model that predicts a player's actions in a social dilemma, based not only on their opponent's prior actions, but also on their opponent's emotional expressions. This model can be used to drive the actions of virtual and robotic agents that can perceive or even participate in the exchange of strategic emotional displays, simulating the way humans interact in these scenarios [6, 8]. The model could also be used to provide real-time decision support to a decision-maker by analyzing the behavior of their opponent or to help teach students how to attend to emotions in social interactions (e.g [3, 13]).

Specifically, we incorporate immediate reactions or emotional displays of a player as input into the decision model of the opponent in an iterated prisoner's dilemma task. We train this model on data from natural, unscripted interactions between pairs of human players and consider game decisions as well as emotional signaling by each player. In the next section we describe our goals and the related literature. In Section 3, we describe the data collection setup and how we extracted the behaviors. In Section 4 we present the performance of the opponent decision model and we show that by adding information to the model such as emotional expressions we can increase the predictive accuracy. Furthermore we will provide insight into the circumstances under which emotions have the potential to shape decisions. Then, in Section 5, we demonstrate the potential of this model to perform social planning (i.e., to generate a sequence of actions and expressions that achieve a goal like maximizing individual or group reward). Again, we show that incorporating emotional signals enhances the ability to generate these plans. Finally, we follow up with discussion in Section 6 and conclusions in Section 7.

2. MODELING DECISION-MAKING IN SOCIAL DILEMMAS

To examine the importance of emotional signals in social dilemmas we train predictive models on a large corpus of people playing an iterated prisoner's dilemma game [22]. The prisoner's dilemma is a game commonly analyzed in game theory and can serve as a means of studying cooperation between people. The game is interesting because it provides an incentive to cooperate (players will do well if they both cooperate), but a temptation to exploit (a player receives the highest payout if their opponent cooperates but they are non-cooperative). In the standard prisoner's dilemma, each player's must choose cooperation or defection (by convention, the non-cooperative choice is referred to as defecting) without knowing how their opponent will decide (the joint outcome is revealed after both players decide). In the iterated prisoner's dilemma, players repeat this decision over multiple rounds. In our corpus, players participated in a finite-horizon (10-round) iterated prisoner's dilemma where both players knew in advance that the game would terminate after ten rounds.

There are several approaches to predicting how players will act in this game. By applying game theory, it can be shown that mutual defection is the rational strategy (defection dominates the single-round game and thus, by backwards induction, it also dominates all decisions in a multi-round game). Thus, if all players were rational (in a game theoretical context), they should always make the non-cooperative choice, regardless of context. Interestingly, human players cooperate far more than this "rational" solution predicts, and thereby receive a far higher payout than players that adopt the game-theoretic solution [5]. Empirically, human players tend to cooperate between 60 and 80 percent of the time across a wide variety of studies. As a result, they perform far better than the game-theoretic solution. Thus, a model that always predicts cooperation would perform reasonably well and we use this as a naive baseline model to compare the performance of our learned model.

Both the rational solution (always defect) and our baseline model (always cooperate) are context-insensitive in that they predict the same decision regardless of the opponent's past decisions. Yet psychological research shows that most people engage in "conditional cooperation" [9]. This means that participants cooperate as long as their opponent does so as well. However, should their opponent fail to cooperate, they will choose to punish them for this by choosing to defect in the next round of the game. This behavior is captured in the popular "tit-for-tat" approach [1, 14]. This seems to model human interactions well, but humans have been observed to be more generous with each other (e.g., allowing some percentage of the opponent's defections to go unpunished) [12] and as such tit-for-tat might be too rigid as a predictive model. Thus we use tit-for-tat as a second baseline and seek to improve upon this.

In order to improve upon the tit-for-tat approach, we will use a machine learning approach, using a Naive Bayes classifier. Machine learning has previously already been explored as a possible approach to improving decision models in the prisoner's dilemma. Sandholm and Crites used an online approach in order to learn an optimal approach to playing an iterated prisoner's dilemma [23]. Using agents that implemented reinforcement learning and by having them play against an opponent using tit-for-tat, these agents would all learn to play optimally against tit-for-tat in an infinite-horizon game. Applying an algorithm such as this in a finitely repeated game might not be the optimal solution. Gal and colleagues used an offline approach to a decision-making model in a negotiation context [11], using a rule based mechanism and supplementing it with character traits on its negotiation partners it applied different negotiation strategies based on their nationality. As a result their model managed to apply different strategies in a finite-horizon negotiation on participants without using data on the participant's past behavior.

We aim to develop a model that predicts an opponent's decision for any given round of the iterative social dilemma, that takes into account both game behaviors (actions in previous rounds) and their emotional displays shown directly before the round. Some previous work has looked at emotional signals before as a means of predicting decisions, however this work was in different contexts than a prisoner's dilemma (e.g. Park and colleagues [21] in a negotiation context) and to the best of our knowledge this approach has not yet been applied to the iterated prisoner's dilemma. For our model we are specifically interested in the effect of a player using emotional displays as a means of signaling on the decision of the opponent.[1]. Based on prior work and the nature of the game, we hypothesize that:

- *Game acts in previous rounds can help predict the opponent's decision in the next round*
- *A player's emotional displays (reactions) in the previous round can influence the opponent's decision and a model can use this information to enhance predictive performance*

By learning a predictive model of the opponent's decision in the iterative social dilemma, we will be able to investigate the intricate dynamics in iterative social dilemmas. For example, using the opponent model in simulations, one can explore different strategies. In this case, we are incorporating emotional signaling in this model, and we are especially interested in emotional signaling strategies: how do the player's emotional displays affect the outcome of the game, and can one form strategies based on displayed emotion? We hypothesize that:

- *Emotional signaling in combination with game behavior can influence opponent game acts*

As far as the emotion perception is concerned we use methods that work in real-time so that we can implement our findings in a real-time agent system. Specifically we use computer vision to automatically detect and classify facial expressions in videos. This way, we capture natural expressions from real interactions and can incorporate realistic emotion displays in our model (as opposed to assigning theoretical states of emotion). Besides investigating the subtle dynamics of human facial display exchanges, we are able to point to realistic behaviors as part of strategic social planning for humans or virtual agent opponents.

To summarize, in this paper, we describe the creation of a real-time decision model based on real human interactions that can be used for robotic or virtual agents in live systems. This model factors in both game decisions and naturalistic emotion displays that will be captured live from the player. This emotion-aware decision model can support research efforts by allowing investigation of the dynamics of an interaction on both those levels (game acts and emotion displays) and can also produce virtual opponents with emotion perception for training systems, where one can test different social planning strategies.

3. IPD CORPUS

In this section we describe the Iterated Prisoner's Dilemma (IPD) dataset. In subsection 3.1, we describe our data acquisition process and in subsection 3.2, we describe extraction process of our behavior descriptors and ground truth.

3.1 Data Collection

This dataset includes a large corpus of people playing an iterated prisoner's dilemma. The data was collected using a computer-mediated version of the game based on the British TV show *Golden Balls*. This framing makes the structure of

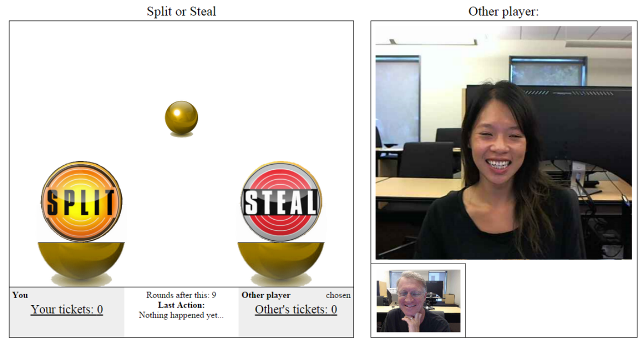

Figure 1: The interface of the game. On the left side, the game is displayed to the participants where they make their choice and can see game info (score, rounds left etc.) On the right side, participants see themselves and their opponent via a videochat-like interface.

the game very clear to players and we felt this would reduce any variance in the data resulting from player confusion. In the TV show, players interact face-to-face and are given two golden balls that represent the two choices in the game. The cooperative choice is labeled Split (meaning that the player intends to evenly split a set of resources) and the non-cooperative choice is labeled steal (meaning that the player intends to steal all the resources). In our version, players interact through a Skype-like interface (see fig. 1) where they can see each other but cannot speak to each other. Players engage in a 10-round game. In each round players received a payout described in Table 1. For those familiar with the TV show, this payoff matrix is slightly different as the TV show doesn't strictly follow the constraints of the prisoner's dilemma (in the TV show, mutual non-cooperation yields zero points for each player).

Data was collected in a large room with computers. Each player was seated individually in front of a computer surrounded by a barrier. Before playing the game, players filled out a demographic questionnaire and some personality scales. They then read instructions about the game and had the opportunity to ask questions before the game commenced. The game interface contains a panel displaying the graphical representation of the game on one side of the screen. Below this some additional information was displayed, namely the current score, the number of rounds remaining and a description of the last game event that occurred. On the other side of the screen a panel displayed the real time webcam video feed of their opponent on top. Below this a smaller video feed of their own webcam was displayed, so participants could ensure they were visible to the camera and to reinforce that their opponent could see them. After the game they filled out additional questions about their impressions.

Each player was paid a $30 for their participation; additionally they were incentivized to maximize their own self-interest by having a chance to receive an additional $100 through a group lottery. In each of the 10 rounds, both players were offered 10 lottery tickets to divide amongst themselves. Based on the decisions whether or not to cooperate, player would divide the tickets (see table 1 for these round

[1]Emotion displays also show felt affect which has been shown to influence a player's decisions as well [14], but given our goal of creating a virtual opponent such displays would be the output of the system rather than input.

Table 1: Round payoffs in the iterated prisoner's dilemma

	Other cooperate	Other defect
Player cooperate	5,5	0,10
Player defect	10,0	1,1

Table 2: Overview of extracted features used for modeling

Category	Feature	Description\encoding	Source
Game Behavior	Self last pick	C or D	IPD event database
	Opp. last pick	C or D	
	Prev. state	CC,CD,DC,DD	
	Curr. round	2,...,10	
	Relative score	$\frac{playerscore}{jointscore}$	
	Coop. rate	% cooperation so far	
	CC rate	% of joint cooperation so far	
	CD rate	% of being exploited so far	
	DC rate	% of exploiting opponent so far	
	DD rate	% of mutual defection so far	
Facial Displays	AU6	Cheek raiser	FACET[16]
	AU10	Upper lip raiser	
	AU12	Lip corner puller	
	AU14	Dimpler	
	AU17	Chin raiser	
	AU24	Lip pressor	

payoffs). By gathering more tickets, players could increase their chances at winning the lottery.

In total, we collected data on 370 players. Besides the synchronized videos of each player, the corpus consists of decision on each round (split or steal), including the time of this decision and the time of the start of a new round. The webcam video feeds were stored on a server for further analysis.

3.2 Automatic Behavior Extraction

For the purpose of studying the IPD interactions we create both game behavior and displayed behavior (expression) features that describe these actions and expressions. Table 2 summarizes those features by category.

When discussing the game behavior of players, we will be using the term *player* to refer to the player whose facial displays we are looking at and who makes that prediction, and *opponent* to refer to the player's opponent whose next move we are trying to predict.

Game Behavior: Player decisions are presumably influenced by actions taken on previous rounds. Thus, one feature is the individual choice made by each player on the preceding round: cooperate (C) or defect (D). A second feature is the game state which is decided by the joint decision made by the dyad on the preceding round: *joint cooperation* (CC), the *player is exploited* by the opponent (CD), the *player exploits* the opponent (DC) or *joint defect* (DD).

Additional features include the current round number, the ratio of the participant's score of the combined participants' score, the cooperation rate of both participants over preceding rounds, and rates for of how often the four game states occurred throughout the game (e.g, how often did joint cooperation occur in the preceding rounds).

Facial Displays: Player decisions may well be influenced by how their opponent emotionally reacts to the previous outcome (for example, de Melo et al. showed that people are more likely to cooperate if their opponent looks guilty after exploiting them). To examine this, we extracted facial expressions using a tool called FACET [16]. FACET analyzes each frame of a video individually for evidence of 19 facial action units (these are components of a facial expression: e.g., action unit 12, or AU12 for short, denotes upturned corners of the mouth). The evidence values describe the likelihood on a logarithmic scale of a specific action unit being active, with positive values indicating that FACET classifies the action unit as active and negative values indicating inactivity. These values where then averaged for each round using the mean evidence values between specific event timings, for example, while participants are deciding what to do and when the joint decision is revealed. Since the strongest facial expressions occurred following this reveal, we used the average evidence in the first six seconds after this event. To consider only facial action units that have reasonable occurrence rate in the data we screened the AUs by activation ratio, and kept only the action units with an overall activation of more than 25% of the data frames. This process left us with six AU features in total: AU6 (cheek raiser), AU10 (upper lip raiser), AU12 (lip corner puller), AU14 (dimpler), AU17 (chin raiser) and AU24 (lip pressor). We only considered emotional reactions to the immediately preceding reveal; although in future work we might consider features that summarize facial activity over all preceding rounds of the game.

Moreover, participant data was screened using a threshold value based on the number of frames the software output had confidence in its output. Dyads with detection rates of less than 75% of the video were discarded from this work, leaving us with 296 participants total. Finally, the first round of the IPD was excluded from our modeling process due to its special circumstances, since participants at that point do not have prior game actions or expression displays to base their decision on.

All our features presented in Table 2 were extracted automatically and are available online in real-time for evaluation of our model.

4. OPPONENT MODELING

In this section, we will give details on our model of opponent decisions in our corpus. We look at the available information in each independent round, such as the previous events that occurred in the game and the player's facial displays, to predict the game decision of the opponent in the next round. Specifically, the model uses the joint game decisions and their results (e.g. current score and game state) and the player's facial displays during the reveal of the results of the previous round as inputs. In Subsection 4.1, we describe our methodology and in Subsection 4.2, we describe the performance of the model and in Subsection 4.3, what we learned.

4.1 Method

Although a variety of learning techniques are possible, we used a Naive Bayes classifier using a kernel density estimate function in order to construct our model. This classifier was chosen since the learned model is easily interpretable (so that we might gain insight into how emotion impacts decision-making) and because it readily facilitates the social planning that we will discuss in Section 5. To examine the contribution of different modalities, we learn several models using different grouping of features. We learn a model using all features (both actions and opponent expressions), but then contrast this with models trained on actions alone or with expressions alone. These learned models are further compared with our two baselines described in Section 2. Recall that the naive baseline always chooses the most common action (i.e., cooperate), whereas tit-for-tat cooperates following the opponent's cooperation and defects following the opponent's defection.

Models were trained with leave-one-participant-out training and testing method. Within each fold, we first performed feature selection. Using the Lasso regression analysis method, we tested several different feature sets based on specific degrees of freedom; 3, 5, 7, 10, 15 and 20. After selecting these different feature sets we used a Naive Bayes algorithm to construct the model.

4.2 Model Performance

Table 3 summarizes the performance of each model in terms of F1 score and overall accuracy, as well as the degrees of freedom that gave us the best results.

As shown in the table, a model trained on actions alone can outperform our naive baseline and the tit-for-tat model, confirming our hypothesis that actions factor in the decision in the next round, and people are more complex than the baseline and tit-for-tat strategy can capture. A model based on expressions alone does not manage to outperform the tit-for-tat strategy but still gives more balanced performance (as measured by F1 score) than the naive baseline. The best performing model used both actions and facial expressions, confirming our hypothesis that opponent expressions enhance predictive accuracy. This model yielded correct decisions almost 75 percent of the time.

The difference in performance between the models using both actions and expressions and the model using actions alone is small, however using logistic regression we demonstrate that the action and expression features both add an independent contribution to our prediction. Using this approach both the actions-only model ($p<0.001$, coeff=-2.269) and the expressions-only model ($p=0.008$, coeff=-0.280) add a significant unique contribution to the prediction.

From the actions, the features most commonly selected were: *oppLastPick*, *selfCoopRate*,*CC%*,*DD%* (referencing Table 2). From the set of expressions, all AUs were selected uniformly in the different models. We show this using the prediction of a model using either game or expression features as input for a logistic regression to predict the ground truth.

Table 3: Performance of baselines and different models

Model	F1 score	Accuracy	DoF
Actions and expressions	**0.742**	**0.744**	5
Actions alone	0.739	0.742	5
Expressions alone	0.521	0.581	7
Tit-for-tat	0.691	0.708	-
Baseline	0.375	0.601	-

4.3 What Did it Learn?

We performed additional analysis to give insight into how the opponent's emotions are influencing a player's decisions. In particular, we wanted to see if specific expressions were more important at different points in the game. Recall that the majority of participants cooperated in the game. As a result, most of the predictions in our corpus are made when the previous round involved joint cooperation (CC). We had conjectured that emotion may be most valuable when players deviate from this most-common state. Therefore we partitioned our corpus into four subsets based on the four basic states that occur in the prisoners' dilemma game: joint cooperation (CC), joint defection (DD), the player exploits opponent (DC) and the player being exploited by opponent (CD). In other words, the player being exploited (CD) subset only contained decisions that followed the situation where the opponent has just exploited the player. We then trained models from these four subsets and examined how they performed. As before, we trained combined models (using both action- and expression-features), action-only models, expression-only models, and contrasted these with our two baselines.

As shown in Table 4, the action-only model performs best following the joint cooperation (CC) state, however the combined model performs best in the being exploited state (CD), player exploits opponent (DC) and joint defection (DD) states. Additionally, we can see that the action unit-only model does better when a player exploits or gets exploited (the CD and DC states) than it does in the joint cooperation state, indicating that these states in particular might benefit from using facial displays when modeling them. These results indicate that expressivity might be more important in these states compared to the much more common joint cooperation state. One final thing to note is that the performance of the state models appear lower than that of the overall models in table 3. However this is only the case because of the weighted nature of the F1-score, when combining the specific state models into a "hybrid" model its F1-score is similar or better to the full game models.

We next tried to visualize what the models learned. Although the model is nonlinear, we can display the way the Naive Bayes model use these descriptors as kernels by plotting them using the probability distribution. As an illustration of the effects of facial expressions, AU12 has a different effect when shown while the player is being exploited (CD) than when the player exploits (DC). Figure 2 shows the probability distribution used by two different models for this action unit. What this graph indicates is that when the

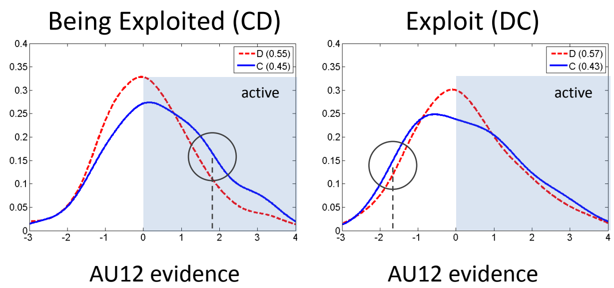

Figure 2: Probability distribution of AU12 used for the "player being exploited" (CD) and "player exploits" (DC) models. A circle notes the AU12 evidence value where $max\{P(C) - P(D)\}$ is observed.

Table 4: F1-score performance of state based models

	Model				
State	Combined	Action	AUs	Baseline	Tit-for-tat
CC	0.662	**0.670**	0.459	0.460	0.460
CD	**0.617**	0.594	0.527	0.308	0.308
DC	**0.615**	0.607	0.514	0.301	0.363
DD	**0.627**	0.598	0.504	0.227	0.414

player exploits the opponent (DC) it is better not to smile, a finding in line with previous research as smiling might indicate that you are more willing concede [7]. Additionally, by smiling after choosing to defect a participant can signal that they enjoy taking advantage of their opponent and are therefore more likely to be punished for this. Whereas, if it was the player being exploited (CD), it is beneficial for the player to display AU12 (a smile) in order to make their opponent return to cooperating, perhaps as a means of showing you forgive their transgression. Smiling appears to only be beneficial in the specific state when your opponent just betrayed you.

When looking at the specific action units used as descriptors by each model, differences can be found as well. The main action units used when the player is exploited (CD) are AU12 (lip corner puller) and AU24 (lip pressor), AU12 is often related to joy while AU24 can often be interpreted as negative. If AU12 is active it is more likely that the opponent will cooperate, whereas AU24 should be inactive in order to maximize cooperation likelihood. AU24, the lip pressor can be seen as a type of mouth control. Therefore if it is active at the same time as AU12 this could indicate that the smile is not as sincere as when AU24 is absent. Therefore this seems to indicate that after an opponent defects it would be best to smile sincerely in order to have them cooperate again. When the player exploits their opponent (DC) AU6 (cheek raiser) and AU17 (chin raiser) are used, again this is an action unit related to positivity (AU6) and one to negativity (AU17). Although AU6 being active only seems to slightly improve cooperation likelihood, however when it is inactive it will greatly increase the likelihood of the opponent defecting (and therefore punishing the participant).

5. SOCIAL PLANNING

Our analysis on the performance of the model indicated that the facial expressions of one participant influences their opponent's decisions. This raises the question of whether it is possible to strategically influence the opponent's decisions by choosing different facial expressions (we refer to this process as social planning). Second, even if we can determine which expression will influence the opponent, is this expression practically significant (i.e., would it make a meaningful difference in outcomes)? In this section, we provide preliminary evidence that the answer to both of these questions is yes.

In its broadest sense, social planning would allow a player to generate a sequence of actions and expressions that influence the opponent to achieve a goal, such as maximizing the player's individual reward. In describing how this is possible, consider the simple case of influencing the opponent's decision on the very next round.

More specifically, consider that player A has just been exploited by their opponent (player B) and would like to influence their opponent to start cooperating again.

This situation is illustrated by the probability distributions on the left-hand side of Figure 2. This graph indicates what player B will do after they just exploited player A and they observe player A showing AU12. If player A fails to show AU12 (i.e, evidence is zero or lower), then player B is most likely to defect, as because the probability of defection (the red line) is higher than the probability of cooperation (the blue line). In contrast, if player A clearly shows AU12 (i.e, evidence is greater than 1), then player B is most likely to cooperate a second time. Therefore, if player A wants to exploit player B again, they should smile. This concept can be extended to multi-step plans by using forward chaining.

Although the best test of this idea would be to run this social planning model against actual human opponents, here we provide a feasibility test of this idea by using the idea of a simulated opponent. In particular, we use the model we learned in Section 4 as a proxy for an actual human opponent (we leave a test against human opponents for future research). We consider if a player could increase their individual score by showing specific facial expressions and contrast this to the score that could be obtained by showing only a neutral expression. If judicious choice of expressions can increase a players score, then this provides evidence that such expressions are practically significant (in that they meaningfully shape outcomes).

Specifically, we attempt to derive an action and expression policy for player A that maximizes player A's score. Because of the nature of the payoff matrix for prisoner's dilemma, player A can maximize their score by coercing player B to cooperate as much as possible while, further, exploiting player B as much as player A can get away with. (By inducing player B to cooperate, player A will have the opportunity to score five points when cooperating and 10 points when exploiting, whereas a defecting player B will afford only zero or one points per round). We derive this policy through brute-force forward-simulation[2] with the simulated

[2]For this simulation we used a well performing model that uses the six action units with an activation rate over 25% as descriptors. Since our model does not simulate the first round, we assume that participants start in the CC state for this, as it is the most common game state in the first round in our dataset.

Table 5: Examples of optimal strategies when using a neutral expression and manipulating certain decisions

	Maximum Score	Example decision
Neutral	70	C-C-D-C-D-C-D-C-C-D C-C-C-D-D-C-D-C-C-D C-C-C-C-C-C-D-D-D-D
Manip. C	90	C-D-D-D-D-D-D-C-D-D C-C-D-D-D-D-D-D-D-D

opponent. At each round in the game we branch on player A's action (cooperate vs. defect) and identify the expression that maximizes B's probability of cooperation. We then compare this against an agent that shows only neutral expressions. For the no expression simulation we will assume none of the action units are active (using a FACET value of -1).

By plotting the probability distributions we were able to identify the optimal strategy for using our expression. This strategy involves displaying the action units that are often related to happiness, as well as the Duchenne smile; AU6, the cheek raiser, should be active (with a value of 1.76) and AU12, the lip corner puller, had an optimal value of 1.45. AU10, the upper lip raiser, can be detrimental to the result when active and has an optimal value is -1.89. AU10 can often be related to negative emotions such as anger. AU24 should also be inactive as well (-0.87), while the last two action units used in the model can generally have some low form of activation: AU14 has a value of 0.49 and AU17's optimal value is 0.65.

Although this model shows that it is possible to achieve a better result when being able to predict your opponent's behavior, it is likely that actual human opponents will change their strategy when they are confronted with an approach such as this. Additionally it is possible to deceive a model such as this, for instance, by presenting behavior that is completely different than the behavior that was used to learn the model. Therefore the results of the simulation should not be considered as an optimal strategy when playing against human opponents. Nonetheless the simulation illustrates that a better understanding of action and emotion can inform social planning, although the effectiveness of this approach will need to be demonstrated in subsequent research.

Table 5 summarizes the score obtained and the resulting policy that results from these social planning simulations. When using the neutral expression strategy, the maximum possible score you can achieve is 70 points (out of a maximum of 100). This score can be obtained by defecting four times throughout the game, with one of these defections occurring in the final round as this leaves the opponent unable to respond. This pattern is sufficient to induce the simulated opponent to always cooperate and shows one possible explanation why our model performs better than a tit-for-tat approach, as the tit-for-tat approach would defect at least three times when faced by four defections. In contrast, by using appropriate facial expressions, it is possible to defect up to eight times, resulting in a score of 90 points. Again, these results are suggestive and need to be verified against human opponents.

6. DISCUSSION

We presented work on modeling opponent decisions in an iterative prisoner's dilemma task. We created a real-time model trained on real interactions between humans, which incorporates both game actions and the player's emotional displays as input and decides what the opponent will pick in the next round (cooperate or defect). We showed that previous game actions can be a strong input in deciding what the opponent will pick next (more so than deciding purely on emotion signals). A simple interpretation is that because of the iterative nature of this task previous actions speak "louder than words" (or in this case signaled intentions via emotional signaling), but still, combining both game acts and emotional displays is what yields the best performance. This finding validates previous research output that emotional signals influence game decisions in joint tasks and moreover, we provided a model, a mechanism, on how emotion contributes to the opponent's decision in a certain round.

To dissect the impact of emotion in more detail, we also modeled the opponent decision separately in the four possible states of the game (joint cooperation, joint defection, player exploits and player being exploited), as discussed in Section 4.3. By adding this context information we were able to highlight better the impact of emotion in the decision of the opponent by coupling it with specific game events (e.g., being exploited, or self exploiting on an opponent). This analysis revealed that emotion input has more impact in the opponent's decision following a state that involved defection (joint defection, player exploits or opponent exploits) as opposed to joint cooperation (Table 4). This observation perhaps makes sense considering that the majority of people picked cooperation (joint cooperation, makes up for ∼46% of the data) and with continuous cooperation being the default scenario there is perhaps less need to signal or communicate intentions for the next round. Defecting however, breaks the loop of default expectations (continuous cooperation) therefore triggering more emotional reactions and also creating more uncertainty, and thus need to plan the next move and that includes estimating the opponent's intentions by all available input (actions and emotions). This also suggests that the impact of emotion may in fact be underrepresented in the overall model, since most of the population in this scenario opted for continuous cooperation.

By employing the learned model in a simulated opponent, we also present a sophisticated tool to study human interactions in social dilemmas and perform social planning in an environment where emotional signals can also yield rewards. We would like to use this tool as a simulated human opponent and plan policies to achieve social goals, similar to reinforcement learning approaches[20, 25]. In this paper, we have shown that this is possible, based on a simple example of employing a combination of action and emotion signals towards maximizing one's own score. Although some of these action sequences seem unlikely while playing against a human opponent we provide them as an illustration of how understanding of action and emotion can inform this form of planning (which is also an important link to how this informs robots and virtual humans). Indeed, many of the human participants succeeded in exploiting their opponent multiple times including one that was exploited 10-times and quite a few up to 6-times (so while this behavior seems unusual it is not unprecedented in our data). Still, we would claim that the ultimate effectiveness of the approach can only be judged empirically in subsequent research against human opponents.

Another aspect of this is to create agents that display more "human-like" behavior. Based on Blascovich's theoretical model of social influence [2], agents that exhibit more agency and are perceived as more realistic will be treated more like humans than those who do not. Using this model we could have a virtual agent simulate human behavior more accurately and investigate whether these agents are treated differently from those that use different approaches, such as tit-for-tat.

Another interesting point of discussion is how we capture and handle emotion input in this model. We believe a novelty of this work to be the inclusion of naturalistic emotional displays from real interactions captured via automatic methods. As a result, we are using realistic (human-appropriate) emotion signals in our model and not theoretical emotional states. This translates to generating more meaningful and realistic policies in order to achieve social planning. For example, anger has been studied a lot in the context of negotiation and is commonly viewed as a means of signaling a tough position in order to avoid exploitation [28, 7], however in our data we observe limited occurrences of anger as it is traditionally defined. Perhaps the specific task does not trigger the specific expression as much, and we fail to observe its impact here, or perhaps, in real interactions people choose to show other more subtle signals (such as lip pressing- AU24) or derivative expressions of frustration/confusion rather than anger. This makes working with this data more challenging since we are dealing with subtle expressions, but it is more realistic since this is what people actually do.

As a summary, we saw that most facial expressions observed in this data span into the two loosely defined categories of smiling (AU6, AU12), and mouth controls (AU10, AU14, AU17, AU24), and this is an intuitive policy that we could either train people to employ for social planning, or implement in a virtual agent to test with.

7. CONCLUSION

We have shown that both joint game actions and the facial displays of a player have predictive power in a classification model for decision making behavior in a iterated prisoner's dilemma. Secondly we have shown how an individual or agent can use this model for social planning by employing not only strategic game decisions but also emotions.

We wish to further investigate the impact of a realistic model using emotional displays such as this in a real life setting and are therefore planning to implement these models into our framework. By using the webcam video feed as input, analyzing the facial expressions of the participant in the video and then using them as descriptors for our model in a real-time to study the effects of a virtual human using these models while playing against human participants.

Acknowledgments

We would like to thank the anonymous reviewers for their insightful comments and suggestions. This research is supported by the National Science Foundation under grant 14196221 and the US Army. The content does not necessarily reflect the position or the policy of any Government, and no official endorsement should be inferred.

REFERENCES

[1] R. Axelrod and W. D. Hamilton. The evolution of cooperation, 1981.

[2] J. Blascovich. A theoretical model of social influence for increasing the utility of collaborative virtual environments. In *Proceedings of the 4th international conference on Collaborative virtual environments*, pages 25–30. ACM, 2002.

[3] J. Broekens, M. Harbers, W.-P. Brinkman, C. M. Jonker, K. Van den Bosch, and J.-J. Meyer. Virtual reality negotiation training increases negotiation knowledge and skill. In *Intelligent Virtual Agents*, pages 218–230. Springer, 2012.

[4] J. Brosig. Identifying cooperative behavior: some experimental results in a prisonerÓş dilemma game. *Journal of Economic Behavior & Organization*, 47(3):275–290, 2002.

[5] V. Capraro. A model of human cooperation in social dilemmas. *PLoS One*, 8(8):e72427, 2013.

[6] C. M. De Melo, P. Carnevale, and J. Gratch. The influence of emotions in embodied agents on human decision-making. In *IVA*, pages 357–370. Springer, 2010.

[7] C. M. de Melo, P. Carnevale, and J. Gratch. The effect of expression of anger and happiness in computer agents on negotiations with humans. In *The 10th International Conference on Autonomous Agents and Multiagent Systems - Volume 3*, AAMAS '11, pages 937–944, Richland, SC, 2011. International Foundation for Autonomous Agents and Multiagent Systems.

[8] C. M. de Melo, J. Gratch, and P. J. Carnevale. The importance of cognition and affect for artificially intelligent decision makers. In *Twenty-Eighth AAAI Conference on Artificial Intelligence*, 2014.

[9] U. Fischbacher, S. Gächter, and E. Fehr. Are people conditionally cooperative? evidence from a public goods experiment. *Economics letters*, 71(3):397–404, 2001.

[10] R. H. Frank, T. Gilovich, and D. T. Regan. The evolution of one-shot cooperation: An experiment. *Ethology and sociobiology*, 14(4):247–256, 1993.

[11] Y. Gal, S. Kraus, M. Gelfand, H. Khashan, and E. Salmon. An adaptive agent for negotiating with people in different cultures. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 3(1):8, 2011.

[12] R. A. Jianzhong Wu. How to cope with noise in the iterated prisoner's dilemma. *The Journal of Conflict Resolution*, 39(1):183–189, 1995.

[13] J. Kim, R. W. Hill, P. Durlach, H. C. Lane, E. Forbell, M. Core, S. C. Marsella, D. V. Pynadath, and J. Hart. BiLAT: A Game-Based Environment for Practicing Negotiation in a Cultural Context. *International Journal of Artificial Intelligence in Education*, 19(Issue on Ill-Defined Domains):289–308, 2009.

[14] P. Kollock. Social dilemmas: The anatomy of cooperation. *Annual review of sociology*, pages 183–214, 1998.

[15] E. Krumhuber, A. Manstead, D. Cosker, D. Marshall, P. L. Rosin, and A. Kappas. Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion*, 7(4):730–735, 2007.

[16] G. Littlewort, J. Whitehill, T. Wu, I. Fasel, M. Frank, J. Movellan, and M. Bartlett. The computer expression recognition toolbox (cert). In *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*, pages 298–305. IEEE, 2011.

[17] G. Loewenstein and J. S. Lerner. The role of affect in decision making. *Handbook of affective science*, 619(642):3, 2003.

[18] M. W. Macy and A. Flache. Learning dynamics in social dilemmas. *Proceedings of the National Academy of Sciences*, 99(suppl 3):7229–7236, 2002.

[19] S. J. Majeski. Arms races as iterated prisoner's dilemma games. *Mathematical Social Sciences*, 7(3):253 – 266, 1984.

[20] A. Papangelis and K. Georgila. Reinforcement learning of multi-issue negotiation dialogue policies. In *16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, page 154, 2015.

[21] S. Park, S. Scherer, J. Gratch, P. Carnevale, and L.-P. Morency. Mutual behaviors during dyadic negotiation: Automatic prediction of respondent reactions. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*, pages 423–428. IEEE, 2013.

[22] W. Poundstone. *Prisoner's Dilemma*. Doubleday, New York, NY, USA, 1st edition, 1993.

[23] T. W. Sandholm and R. H. Crites. Multiagent reinforcement learning in the iterated prisoner's dilemma. *Biosystems*, 37(1):147–166, 1996.

[24] J. Schug, D. Matsumoto, Y. Horitaa, T. Yamagishi, and K. Bonnet. Emotional expressivity as a signal of cooperation. *Evolution and Human Behavior*, 31:87–94, 2010.

[25] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.

[26] A. Sparks, T. Burleigh, and P. Barclay. We can see inside: Accurate prediction of prisoner's dilemma decisions in announced games following a face-to-face interaction. *Evolution and Human Behavior*, 37(3):210 – 216, 2016.

[27] G. Stratou, R. Hoegen, G. Lucas, and J. Gratch. Emotional Signaling in a Social Dilemma: an Automatic Analysis. In *Proceedings of ACII 2015*, Xi'an, China, Sept. 2015. IEEE.

[28] G. A. Van Kleef, C. K. De Dreu, and A. S. Manstead. The interpersonal effects of anger and happiness in negotiations. *Journal of personality and social psychology*, 86(1):57, 2004.

[29] G. A. Van Kleef, C. K. De Dreu, and A. S. Manstead. An interpersonal approach to emotion in social decision making: The emotions as social information model. *Advances in experimental social psychology*, 42:45–96, 2010.