Pride and Perjury: A Computational Characterization of Multiagent Games with Fallacious Rewards

Ariel D. Procaccia and Jeffrey S. Rosenschein

School of Engineering and Computer Science
Hebrew University, Jerusalem, Israel
{arielpro,jeff}@cs.huji.ac.il

Abstract. Agents engaged in noncooperative interplay often seek convergence to Nash equilibrium; this requires that agents be aware of others' rewards. Misinformation about rewards leads to a gap between the real interaction model — the *explicit* game — and the game which the agents perceive — the *implicit* game. We identify two possible sources of fallacious rewards: *pride* and *perjury*.
If estimation of rewards is based on modeling, a *proud* agent is likely to err. We define a *robust equilibrium*, which is impervious to slight perturbations, and prove that one can be efficiently pinpointed. We relax this concept by introducing *persistent equilibrium pairs* — pairs of equilibria of the explicit and implicit games with nearly identical rewards — and resolve associated complexity questions.
Supposing agents simply reveal their valuations for different outcomes of the game, *perjuring* agents may report false rewards in order to improve their payoff. We define the Game-Manipulation (GM) decision problem, and fully characterize the complexity of this problem and some variations.

1 Introduction

"It is very often nothing but our own vanity that deceives us."
– Jane Austen, *Pride and Prejudice*

Game theory has long been a standard tool in the analysis of multiagent interactions. Cooperative game theory, for instance, is widely regarded as the standard model in the context of coalition formation among autonomous agents. Noncooperative game theory is, however, the field in which a large portion of multiagent research is truly grounded, as one often considers settings where self-interested agents act alone.

Nash equilibrium is considered to be the prominent solution concept for noncooperative games; when players' strategies are in Nash equilibrium, no player stands to gain by unilaterally changing his strategy. The applicability of this concept to multiagent systems has motivated researchers to study the complexity of finding a Nash equilibrium, in terms of both the computation required [5, 7] and the communication involved [6].

Unfortunately, it seems impossible to efficiently find a Nash equilibrium, and, to make things worse, agents usually lack complete information about their peers and environment. This has led researchers to devote much effort to learning in games, and in particular learning a Nash equilibrium [9, 10, 2].

Whichever method is used in order to guarantee that agents converge to an equilibrium, it is usually assumed that agents can observe others' rewards for given combinations of actions. But what are the possible sources of these observations? We suggest three:

1. Direct observation: an agent can see another's reward. For example, if the agents are playing poker, an agent can observe how much money each agent gets paid for a given combination of actions.
2. Modeling: by modeling the other agent and the environment, it is possible to estimate the other's rewards through observation.
3. Preference revelation: each agent announces its valuation for different action profiles, or, in other words, for different outcomes of the game.

In the first case, it seems reasonable that agents might observe the *correct* rewards for other agents; in the last two cases, the rewards an agent assumes for others through observation might very well be incorrect. In the second case, an incorrect model may lead to an incorrect estimation; in the third case, an agent may reveal its preferences strategically (or to put it simply, lie) in order to improve the final outcome. Below we use the *implicit games* terminology proposed by Bowling and Veloso [3]: the explicit game is the real game, whereas the implicit game is the same game but with possibly different rewards.[1]

When rewards are obtained by modeling, mistakes may occur as a result of "overconfidence" on the part of the modeling agent; we call the source of these mistakes *pride*. In this context, we study equilibrium concepts that satisfy some notion of stability in the face of changing rewards.[2] Our idea of an ϵ-robust equilibrium is a strategy profile which remains an equilibrium in any implicit game in which the rewards are perturbed by at most ϵ. We characterize these equilibria and show they can be found in polynomial time. Since robust equilibria rarely exist, we produce a relaxation in the form of *persistent equilibrium pairs*, and study the complexity of deciding the existence of such pairs.

In the case where errors originate in strategic preference revelation, we say the source of mistakes is *perjury*. The subject of the complexity of manipulation, especially in voting, has been the source of much interest [1, 13, 4]. In our predicament, an agent may improve its payoff by smartly revealing rewards that incite the other agents to converge to an equilibrium of the induced implicit

[1] This framework was introduced as a model of learning with limitations (such as broken actuators or hardwired behavior), but was discarded in favor of a different model, which is more suitable with respect to modeling limitations. A simpler version of Bowling and Veloso's implicit games model is appropriate for our purposes.

[2] The literature already contains some examples of robust equilibria (see, for example, [14]). Nevertheless, such notions of robustness which build upon Bayesian Nash equilibria [11] are unsuitable for our purposes.

game, but in fact plays a strategy that maximizes its payoff in the explicit game. In more detail, the other gullible agents, observing the implicit game, believe that the manipulator would play a certain strategy, which is a part of a Nash equilibrium strategy profile in the implicit game; they think this strategy is a best response for the manipulator, and hence would be played. However, the manipulator, which is aware of its true rewards in the explicit game, plays a different strategy, which improves its outcome. The final outcome the manipulator receives is often better than the one it gets for a strategy profile which is an equilibrium in the explicit game. We formalize the above discussion in the form of a decision problem — GAME-MANIPULATION (GM) — and prove that the problem is in $\mathcal{P}$ for 2 players, but $\mathcal{NP}$-complete for at least 3 players. We also examine three variations on GM.

The paper proceeds as follows. In Section 2, we outline some basic ideas in game theory, and state relevant results concerning the complexity of Nash equilibria. In Section 3, we present our results regarding mistakes which are due to pride. In Section 4, we prove results about mistakes whose source is perjury. In Section 5 we conclude and suggest directions for future research.

2 Preliminaries

We start with a short introduction to game theory, and a formulation of some relevant results. The reader is urged to consult [11] for more details.

Definition 1. *A* game in strategic form *(hereinafter, simply* game*) G is a tuple $(n, A_{1,\dots,n}, R_{1,\dots,n})$. $N = \{1, \dots, n\}$ is the set of players, A_i is the finite action set of player i, and $R_i : A \to \mathbb{R}$ is the reward function of player i, where $A = A_1 \times \cdots \times A_n$.*

Many of our examples will rely on 2-player games were each player has two actions; we refer to the actions of player 1 as U (up) and D (down), and to the actions of player 2 as L (left) and R (right). If one of the players has 3 actions, we refer to the additional action as M (middle).

Definition 2. *A* strategy *for player i is $\pi_i \in \Delta(A_i)$; it specifies the probability of player i playing each action in its action set. A* pure *strategy gives probability 1 to some action. A* mixed *strategy is a strategy which is not pure. A* strategy profile *is a tuple containing one strategy for each player.*

For a strategy profile π, we denote by π_{-i} the tuple containing the strategies of all players other than i, and by π_C the tuple of strategies of the players in $C \subset N$.

A special case of games is *zero-sum* games. In such games, there are two players with diametrically opposed interests, i.e., $R_1 \equiv -R_2$. When dealing with zero-sum games, we write R instead of R_1; the goal of player 2 is to minimize the reward. It is known that in zero-sum games:

$$\max_{\pi_1 \in \Delta(A_1)} \min_{\pi_2 \in \Delta(A_2)} R(\pi_1, \pi_2) = \min_{\pi_2 \in \Delta(A_2)} \max_{\pi_1 \in \Delta(A_1)} R(\pi_1, \pi_2) = v(G).$$

$v(G)$ is called the *value* of the game G. Strategies that guarantee that a player receive at least payoff $v(G)$ (for player 1) or at most payoff $v(G)$ (for player 2) are called *optimal strategies*.

Returning to general-sum games, we discuss the concept of equilibrium. Given a strategy profile $\pi = (\pi_1, \ldots, \pi_n)$, we say that π_i is a *best response* (BR) to π_{-i} iff for all $\pi'_i \in \Delta(A_i)$:

$$R_i(\pi) \geq R_i(\pi'_i, \pi_{-i}). \tag{1}$$

Definition 3. *A strategy profile π is a* Nash equilibrium *(NE) iff for all i, π_i is a best response to π_{-i}.*

It is a well-known fact that every game has a Nash equilibrium.

Remark 1. A pair of optimal strategies in a zero-sum game is a Nash equilibrium.

Definition 4. *A Nash equilibrium π is a* coordination equilibrium *iff all players achieve their highest possible value:*

$$R_i(\pi) = \max_{a \in A} R_i(a) \tag{2}$$

for all $i \in N$.

It is still unclear what the complexity of finding a Nash equilibrium is. Nevertheless, some related questions have been resolved.

Theorem 1. *[5] Even in symmetric 2-player games, it is $\mathcal{NP}$-hard to decide the following problems:*

1. *Whether there is a NE where player 1 sometimes plays $a_1 \in A_1$.*
2. *Whether there is a NE where all players have expected utility greater than k.*
3. *Whether there is a NE where the social welfare is greater than k.*
4. *Whether there is more than one Nash equilibrium.*

3 Pride: Modeling Inaccurate Rewards

In this section, we discuss scenarios where players converge to equilibria based on erroneous models. In such settings, an agent estimates the payoffs other agents receive for action profiles, relying on knowledge of its peers. It is therefore likely that an agent might inaccurately estimate other agents' rewards. A more extraordinary situation is the one where an agent fails to correctly determine its own rewards; this may happen when an agent learns the game using an inaccurate model of the environment. The mistakes are often assumed to be bounded by a (relatively small) number ϵ.

Observing the formulation of Bowling and Veloso [3], we dub the "real" game — the game with the correct rewards — the *explicit* game, and the game with the same players and action sets, but perhaps different rewards, the *implicit* game. The explicit game is denoted by $G = (n, A_{1,\ldots,n}, R_{1,\ldots,n})$, while the implicit game is represented by $\widetilde{G} = (n, A_{1,\ldots,n}, \widetilde{R}_{1,\ldots,n})$.

Definition 5. *Let G be an explicit game, and $\widetilde{G}$ be an implicit game. The pair $\langle G, \widetilde{G} \rangle$ is an* ϵ-perturbed system *iff for all $a \in A$ and all $i \in N$: $|R_i(a) - \widetilde{R}(a)| \leq \epsilon$.*

Example 1. We briefly demonstrate how small changes in the rewards of players in the explicit game significantly impact players' utilities in the Nash equilibrium. In all examples, the explicit game will be specified using a normal matrix, while the implicit game will be described with a square matrix.

$$\begin{pmatrix} (-10,-10) & (9,-9) \\ (-9,9) & (10,10) \end{pmatrix}$$

The game has a single equilibrium at (D, R), with payoffs $(10, 10)$. Several small errors create the implicit game:

$$\begin{bmatrix} (-10,-10) & (9,-11) \\ (-11,9) & (8,8) \end{bmatrix}$$

This game has a single equilibrium at (U, L) with payoffs $(-10, -10)$.

3.1 Robust Equilibria

It is desirable that games be robust to small perturbations of the rewards, in the sense that there is a strategy profile which is an equilibrium in every perturbed implicit game. Should this property hold, mistakes would little affect the outcome of the game.

Definition 6.

1. *Let G be a game, and π a strategy profile. We say that π_i is an* ϵ-robust best response *to π_{-i} in G iff for all games $\widetilde{G}$ such that $\langle G, \widetilde{G} \rangle$ is an ϵ-perturbed system, π is a best-response to π_{-i} in $\widetilde{G}$.*
2. *Let G be a game. A strategy profile π is an* ϵ-robust equilibrium *iff for all i, π_i is an ϵ-BR to π_{-i} in G.*

Example 2. In the following explicit game, (U, L) is a 1-robust equilibrium, while (D, R) is an equilibrium which is not ϵ-robust for any $\epsilon > 0$.

$$\begin{pmatrix} (2,2) & (0,0) \\ (0,0) & (0,0) \end{pmatrix}$$

It is clearly often advantageous for agents to jointly converge to a robust equilibrium — if such an equilibrium exists. How hard is it to find one? We first address the existence problem.

Definition 7. *In the* Robust-NE *problem, we are given a game G, a strategy profile π, and $\epsilon > 0$, and are asked whether π is an ϵ-robust NE in G.*

Lemma 1. *Let G be a game, and π_i be an ϵ-robust BR to π_{-i} in G. Then π_i is a pure strategy.*

Proof. Omitted due to lack of space. □

Lemma 2. *$\widehat{a}_i \in A_i$ is an ϵ-robust BR to π_{-i} iff for all $a_i \neq \widehat{a}_i$,*

$$R_i(\widehat{a}_i, \pi_{-i}) \geq R_i(a_i, \pi_{-i}) + 2\epsilon. \tag{3}$$

Proof. Omitted due to lack of space. □

Lemma 3. Robust-NE *is in $\mathcal{P}$.*

Proof. It is sufficient to show that it is possible to determine whether a strategy is an ϵ-robust BR in polynomial time. By Lemma 1, we can assume the given strategy π_i is a pure strategy $\widehat{a}_i$. By Lemma 2, we only have to check if Equation (3) holds for all other pure strategies in A_i — and this can be accomplished in polynomial time. □

The theorem follows directly from the above lemmas.

Theorem 2. *It is possible to find an ϵ-robust equilibrium, or determine that one does not exist, in polynomial time.*

Proof. Given a game G, by Lemma 1, in order to determine whether there exists a robust equilibrium it is sufficient to test each pure $a \in A$ for this property. By Lemma 3, the test can be executed in polynomial time. □

Remark 2. It is possible to define the *robustness level* ϵ^* of a given game as the maximal ϵ such that there exists an ϵ-robust equilibrium in the game. The above results show that the robustness level can also be calculated in polynomial time, using the formula:

$$\epsilon^* = \max_{\widehat{a} \in A} \min_{i \in N} \min_{a_i \neq \widehat{a}_i} \frac{R_i(\widehat{a}) - R_i(a_i, \widehat{a}_{-i})}{2}.$$

Indeed, given $\widehat{a} \in A$, by Lemma 2 the value

$$\min_{i \in N} \min_{a_i \neq \widehat{a}_i} \frac{R_i(\widehat{a}) - R_i(a_i, \widehat{a}_{-i})}{2}$$

is a non-negative number ϵ iff $\widehat{a}$ is an ϵ-robust equilibrium, and is not ϵ'-robust for all $\epsilon' > \epsilon$. By Lemma 1, it is sufficient to maximize over all pure strategy profiles in order to find the robustness level.

3.2 Persistent Equilibrium Pairs

Robust equilibria are characterized by Lemmas 1 and 2; in particular, a robust equilibrium must be a pure strategy profile. It seems that, although efficiently solvable, robust equilibrium is too strong a solution concept. We wish to relax our notion of robustness.

Suppose the players have errors in their game models, and converge to an equilibrium in the implicit game. Even if the new equilibrium is not an equilibrium in the explicit game, we would like to know how "far" (in terms of rewards) this equilibrium is from an equilibrium in the explicit game. If there is a pair of equilibria in the explicit and implicit games that are "close", we think of this pair as satisfying some notion of stability. Of course, the distance between the equilibria depends on the magnitude of the mistakes. Formally:

Definition 8. *Let $\langle G, \widetilde{G} \rangle$ be a perturbed system. An ordered pair $\langle \pi, \rho \rangle$ where π is a Nash equilibrium in G and ρ is a Nash equilibrium in $\widetilde{G}$ is an* ϵ-persistent equilibrium pair *of $\langle G, \widetilde{G} \rangle$ iff:*

$$\forall i,\ |R_i(\pi) - \widetilde{R}_i(\rho)| \leq \epsilon.$$

Remark 3. If $\langle \pi, \rho \rangle$ is an ϵ-persistent equilibrium pair in the ϵ-perturbed system $\langle G, \widetilde{G} \rangle$, then:

$$\forall i,\ |R_i(\pi) - R_i(\rho)| \leq 2\epsilon.$$

A persistent equilibrium pair differs from a robust equilibrium in two respects: it is relevant only to a specific perturbed system (instead of every ϵ-perturbed system), and it can hold that $\pi \neq \rho$, as long as the rewards are close.

Before proving some existence theorems for such pairs in special games, we deal with the complexity of determining whether a given perturbed system has a persistent equilibrium pair.

Definition 9. *In the* Persistent-Pair *problem, we are given an explicit game G, an implicit game $\widetilde{G}$, and a number $\epsilon \geq 0$. We are asked whether $\langle G, \widetilde{G} \rangle$ has an ϵ-persistent equilibrium pair.*

Theorem 3. Persistent-Pair *is $\mathcal{NP}$-complete, even for two players.*

Proof. The problem is in $\mathcal{NP}$. Indeed, a specific choice of π and ρ can serve as a witness; it can be determined in polynomial time whether indeed π is a NE in G, ρ a NE in $\widetilde{G}$, and whether their rewards differ by at most ϵ.

We prove that the problem is $\mathcal{NP}$ hard via a reduction from the problem of determining whether a given game has a NE where all players have expected payoff at least k. Given an instance $\langle G, k \rangle$ of the former problem, the reduction creates an instance of Persistent-Pair: define $r_{max} = \max_{i \in N, a \in A} R_i(a)$; G is identical to the given game, $\widetilde{G}$ is the game where all rewards are r_{max}, and $\epsilon = r_{max} - k$.[3] Notice that all equilibria in $\widetilde{G}$ have a payoff of r_{max} to all players. Assume that the given instance is a "yes" instance; thus, there is an equilibrium π with payoff at least k to all players, but clearly with payoff at most r_{max}. Picking some equilibrium ρ in $\widetilde{G}$, we have for all i:

$$|R_i(\pi) - \widetilde{R}_i(\rho)| \leq r_{max} - k = \epsilon.$$

[3] It is safe to assume that $r_{max} \geq k$, otherwise the given instance is clearly a "no" instance.

Therefore, $\langle \pi, \rho \rangle$ is an ϵ-persistent equilibrium pair. In the other direction, assume that the given instance is a "no" instance. Any equilibrium π in G has a player i such that $R_i(\pi) < k$. Hence, for any equilibrium ρ in $\widetilde{G}$ it holds that:

$$\widetilde{R}_i(\rho) - R_i(\pi) > r_{max} - k = \epsilon.$$

□

A setting where a persistent equilibrium pair is guaranteed to exist is the one where some of the equilibria of the explicit and implicit games satisfy special properties. One such example is requiring that both games have coordination equilibria.[4]

Proposition 1. *Let $\langle G, \widetilde{G} \rangle$ be an ϵ-perturbed system, and let π and ρ be coordination equilibria in G and $\widetilde{G}$, respectively. Then $\langle \pi, \rho \rangle$ is an ϵ-persistent equilibrium pair.*

Proof. Omitted due to lack of space. □

Another important case where persistent equilibrium pairs always exist is when G and $\widetilde{G}$ are both zero-sum games.

Definition 10. *An ϵ-perturbed system $\langle G, \widetilde{G} \rangle$ is* zero-sum *iff both G and $\widetilde{G}$ are (2 player) zero-sum games.*

Remark 4. In zero-sum perturbed systems, the errors are implicitly assumed to be consistent, in a way: an error e in the reward of one of the players entails an error $-e$ in the reward of the other.

Proposition 2. *Let $\langle G, \widetilde{G} \rangle$ be a zero-sum ϵ-perturbed system, and let π and ρ be optimal strategy profiles in G and $\widetilde{G}$, respectively. Then $\langle \pi, \rho \rangle$ is an ϵ-persistent equilibrium pair.*

Proof. Omitted due to lack of space. □

4 Perjury: Revealing False Rewards

In the previous section we examined settings where agents converge to a false equilibrium, as a result of erroneous estimation of rewards. In this section our study takes a sinister turn: we look at settings where agents reveal their preferences to other agents, by reporting their rewards for different action profiles. In such cases, an agent may improve its payoff by reporting false rewards. When this happens, we say the lying agent reveals its preferences *strategically*, and further slander it by referring to it as a *manipulator*.

[4] Our interest in coordination equilibria stems from Littman's results: Littman provided a Q-learning algorithm that converges to Nash equilibrium provided the (stochastic) game has a coordination equilibrium.

Example 3. The following example presents a successful manipulation by player 1. Observe the explicit game:

$$\begin{pmatrix} (0,0) & (1,1) \\ (1,1) & (2,0) \end{pmatrix}$$

The unique NE is is (D, L), with payoff 1 to both players. We would like to know if player 1 can do better. By changing player 1's rewards, we obtain the implicit game:

$$\begin{bmatrix} (0,0) & (1,1) \\ (-1,1) & (-1,0) \end{bmatrix}$$

Observing the implicit game, player 2 converges to the only NE: (U, R), i.e., plays strategy R. Player 2 believes (U, R) is a Nash equilibrium, and thus player 1 can do no better than play U. However, being aware of the real rewards and knowing that player 2 would play R, player 1 counters with D, receiving a payoff of 2 (and leaving player 2 with a sucker's payoff of 0).

The next definition formalizes the above discussion.

Definition 11. *In the* Game-Manipulation *(GM) problem, we are given an explicit n-player game G, a player $\underline{i} \in N$, and an integer $k \in \mathbb{Z}$. We are asked whether there is an implicit game $\widetilde{G}$ where only the rewards of $\underline{i}$ are changed, a (possibly mixed) strategy profile ρ and $\pi_{\underline{i}} \in \Delta(A_{\underline{i}})$, such that ρ is a Nash equilibrium in $\widetilde{G}$, and*

$$R_{\underline{i}}(\pi_{\underline{i}}, \rho_{-\underline{i}}) > k.$$

4.1 Complexity of GM

Our goal in this subsection is to prove:

Theorem 4 (Dichotomy of GM). *GM with at least 3 players is $\mathcal{NP}$-complete, while GM with 2 players is in $\mathcal{P}$.*

The proof is naturally decomposed into two parts.

Proposition 3. *GM with at least 3 players is $\mathcal{NP}$-complete.*

Proof. To show that GM is in $\mathcal{NP}$, observe that for a given instance of the problem, a specific example of $\widetilde{G}$, ρ and $\pi_{\underline{i}}$ is a witness; it can be ascertained in polynomial time whether indeed ρ is a Nash equilibrium in $\widetilde{G}$ and

$$R_{\underline{i}}(\pi_{\underline{i}}, \rho_{-\underline{i}}) > k.$$

For the $\mathcal{NP}$-hardness, we prove that the problem of determining whether there exists a NE in a 2-person game where player 1 sometimes plays $\widehat{a}_1 \in A_1$ reduces to GM with 3 players. Given an instance of the former problem, construct as an instance of the latter a three-player game G, where player 3 has only one action; the rewards for players 1 and 2 for any action profile are the same as in the given game (when the action of player 3 is disregarded), and the payoff for player 3 is 1 for any action profile where player 1 plays $\widehat{a}_1$, and 0 otherwise. We also set $k = 0$, and the manipulator to be player 3 (i=3).

Assume the given instance has a NE π where player 1 plays $\widehat{a}_1$ with probability $r > 0$. Construct the instance of GM as above, and choose $\widetilde{G} \equiv G$. Clearly, when players 1 and 2 use π_1 and π_2 in G (and player 3 plays his single action), this is a Nash equilibrium with the property that player 3 has payoff r.

On the other hand, it is clear that if (π_1, π_2) is not a Nash equilibrium in the given instance, then (π_1, π_2, a_3) (where a_3 is the single action in A_3) is not a Nash equilibrium in any $\widetilde{G}$ where only the rewards of player 3 have changed. Therefore, if the given instance has no NE where player 1 sometimes plays $\widehat{a}_1$, every NE in every admissible $\widetilde{G}$ has the property that player 1 plays $\widehat{a}_1$ with probability 0, and thus the utility of player 3 is 0. □

In the second part of the theorem's proof we require the following lemma. For ease of exposition, in the lemma and subsequent proposition we consider player 1 to be the manipulator, but this is of course an arbitrary choice.

Lemma 4. *A given instance of GM with two players where player 1 is the manipulator is a "yes" instance iff there exists a pure strategy profile* $(\widehat{a}_1, \widehat{a}_2)$ *such that* $R_1(\widehat{a}_1, \widehat{a}_2) > k$*, and a strategy* $\rho_1 \in \Delta(A_1)$ *such that* $\widehat{a}_2$ *is a best-response to* ρ_1 *(in the explicit game).*

Proof. Omitted due to lack of space. □

Proposition 4. *GM with 2 players is in* $\mathcal{P}$.

Proof. We present an algorithm for GM, and prove that it always terminates in polynomial time.

1: **for** each $(\widehat{a}_1, \widehat{a}_2) \in A$ s.t. $R_1(\widehat{a}_1, \widehat{a}_2) > k$ **do**
2:     **if** $\exists \rho_1 \in \Delta(A_1)$ s.t. $\widehat{a}_2$ is a BR to ρ_1 **then**
3:         **return true**
4:     **end if**
5: **end for**
6: **return false**

It follows straightforwardly from Lemma 4 that a given instance of GM is a "yes" instance iff the above algorithm accepts. Therefore, we only need to show that the algorithm can be implemented efficiently.

The algorithm performs the test in line 2 a polynomial number of repetitions; this test can also be performed in polynomial time. Indeed, there is ρ_1 such that $\widehat{a}_2$ is a best response to ρ_1 iff there is a feasible solution to the linear program with the following constraints:

$$\textbf{find } \rho_1(a_1) \textbf{ such that}$$

$$\forall a_2 \in A_2,\ \sum_{a_1 \in A_1} \rho_1(a_1)[R_2(a_1, \widehat{a}_2) - R_2(a_1, a_2)] \geq 0$$

$$\sum_{a_1 \in A_1} \rho_1(a_1) = 1$$

$$\forall a_1 \in A_1,\ 0 \leq \rho_1(a_1) \leq 1$$

It is well-known that it is possible to determine whether a linear program has a feasible solution in polynomial time. This completes the proof of the proposition, as well as the proof of Theorem 4. □

4.2 Variations on GM

In this subsection, we are concerned with variations on the GM problem.

GM is naturally generalized to a setting where there are several manipulators. The manipulators are all interested in securing a payoff greater than k. However, the manipulators are still selfish, and do not hesitate to improve their own payoff at the expense of other manipulators. Hence, we require that the manipulators' strategies in the explicit game also be in equilibrium.

Definition 12. *In the* COALITIONAL-GAME-MANIPULATION *(CGM) problem, we are given an explicit n-player game G, a subset of players $C \subseteq N$, and an integer $k \in \mathbb{Z}$. We are asked whether there is an implicit game $\widetilde{G}$ where only the rewards of the players in C are changed, a (possibly mixed) strategy profile ρ and $\pi_C \in \prod_{i \in C} A_i$, such that ρ is a Nash equilibrium in $\widetilde{G}$, π_i is a BR to (π_{C-i}, ρ_{-C}) in G for all $i \in C$, and*

$$\forall i \in C, R_i(\pi_C, \rho_{-C}) > k.$$

Example 4. We clarify the definition by showing a coalitional manipulation in a three player game, in which players 1 and 2 are manipulators. The game is described by two matrices: the left is associated with an action of player 3 which we dub B (backward), while the right matrix is associated with F (forward).

$$\begin{pmatrix} (2,2,0) & (1,1,0) \\ (1,1,1) & (0,0,1) \end{pmatrix} \qquad \begin{pmatrix} (1,1,2) & (0,0,0) \\ (0,0,0) & (0,0,0) \end{pmatrix}$$

In this game there is a Nash equilibrium at (U, L, F), with payoff $(1, 1, 2)$. Now observe the implicit game:

$$\begin{bmatrix} (2,2,0) & (1,1,0) \\ (1,1,1) & (2,2,1) \end{bmatrix} \qquad \begin{bmatrix} (1,1,2) & (0,0,0) \\ (2,2,0) & (0,0,0) \end{bmatrix}$$

(U, L, F) is no longer an equilibrium, but (D, R, B) is. Player 3 plays B, and the manipulators counter with (U, L); the utility of both manipulators is 2. The manipulators are not motivated to deviate, since U is a BR for player 1 against (L, B) in the explicit game, and L is a BR against (U, B).

Proposition 5. *CGM with at least 3 players is $\mathcal{NP}$-complete, while CGM with 2 players is in $\mathcal{P}$.*

Proof. Omitted due to lack of space. □

Remark 5. It is also possible to formulate the CGM problem for a cooperative setting, where the manipulators act as a coalition. In this case, we would only require that the total reward of the coalition be greater than k, and would drop the requirement that the manipulators' strategies be in equilibrium in the explicit game — but this is outside the scope of this paper.

In the GM problem, the manipulator was only concerned about its own reward. Nevertheless, some "Robin Hood" type manipulators wish to engage in deceit in order to improve the welfare of all players.

Example 5. Consider the explicit game:

$$\begin{pmatrix} (1,1) & (0,0) \\ (0,0) & (-1,10) \\ (0,10) & (9,9) \end{pmatrix}$$

The unique NE is (U, L), with payoff 1 to both players. In this case, player 1 can significantly increase the payoff to both players by reporting false rewards.

$$\begin{bmatrix} (0,1) & (0,0) \\ (1,0) & (1,10) \\ (0,10) & (0,9) \end{bmatrix}$$

In the implicit game, there is a single equilibrium in (M, R). Player 2 plays R, but player 1 counters with D. Consequently, both players get a payoff of 9. The social welfare increased from 2 to 18.

Definition 13. *In the* Benevolent-Game-Manipulation *(BGM) problem, we are given an explicit n-player game G, a player $\underline{i} \in N$, and an integer $k \in \mathbb{Z}$. We are asked whether there is an implicit game $\widetilde{G}$ where only the rewards of $\underline{i}$ are changed, a (possibly mixed) strategy profile ρ and $\pi_{\underline{i}} \in \Delta(A_{\underline{i}})$, such that ρ is a Nash equilibrium in $\widetilde{G}$, and*

$$\sum_{i \in N} R_{\underline{i}}(\pi_{\underline{i}}, \rho_{-\underline{i}}) > k.$$

Proposition 6. *BGM with at least 3 players is $\mathcal{NP}$-complete, while BGM with 2 players is in $\mathcal{P}$.*

Proof. Omitted due to lack of space. □

Returning to the selfish manipulator setting (the GM formulation), another issue requires attention. Even if the manipulator succeeds in finding $\widetilde{G}$, ρ, and $\pi_{\underline{i}}$ as before, it is not at all certain that the other players, observing the implicit game $\widetilde{G}$, would converge to ρ. This is only guaranteed if ρ is a unique NE in $\widetilde{G}$. We define a strong manipulation to be exactly as before, except that we now require that ρ be unique.

Definition 14. *In the* Strong-Game-Manipulation *(SGM) problem, we are given an explicit n-player game G, a player $\underline{i}$, and an integer k. We are asked whether there is an implicit game $\widetilde{G}$ where only the rewards of player $\underline{i}$ are changed, a (possibly mixed) strategy profile ρ and $\pi_{\underline{i}} \in \Delta(A_{\underline{i}})$, such that ρ is a* unique *Nash equilibrium in $\widetilde{G}$, and*

$$R_{\underline{i}}(\pi_{\underline{i}}, \rho_{-\underline{i}}) > k.$$

Remark 6. Example 3 is in fact an example of a strong manipulation.

Proposition 7. *SGM with at least 3 players is co$\mathcal{NP}$-hard.*

Proof. Omitted due to lack of space. □

5 Conclusions

We have demonstrated that when agents converge to a Nash equilibrium on the basis of fallacious rewards, their utilities may change substantially. For the setting where agents estimate rewards by relying on a model of other agents and the environment, we have defined the concept of ϵ-robust equilibrium, and have shown that if one exists, it can be found in polynomial time. However, our characterization of these equilibria implies that they rarely exist. Accordingly, we have relaxed the definition to obtain persistent equilibrium pairs. Although such pairs always exist when the explicit and implicit games are both zero-sum or both have coordination equilibria, in general deciding the existence of such a pair is $\mathcal{NP}$-hard. We wish to comment that the problem of determining whether some multiagent setting has a robust equilibrium or persistent pair is not associated with the agents (which are not aware of the explicit game), but rather with the system designer, who strives to create a stable system.

We have also considered mistakes which are grounded in false reports by manipulative agents. We have shown that the Game-Manipulation problem is in $\mathcal{P}$ for 2 players, and is $\mathcal{NP}$-complete for at least 3 players. We have additionally demonstrated that similar results hold for the "coalitional", "beneficial", and "strong" variations of GM. These results suggest that manipulation may not be a major concern when there are at least three agents, but it remains to determine how hard it is to decide these problems in the average-case [13]. The "beneficial" setting seems especially interesting: although, in general, one wishes to avoid manipulations, a well-intentioned lie can help agents avoid paying the price of anarchy [12].

It is important to note that the results in this paper are general, in the sense that they are independent of the specific algorithms the agents use to converge to an equilibrium; it suffices to assume that they all use such an algorithm (not necessarily the same one). Moreover, the results are stated with respect to games in normal form, commonly known as *stage games*; this makes them applicable in repeated games. Indeed, in several papers [2], the notion of convergence in repeated games is that the *stage game* strategies converge to a Nash equilibrium of the stage game; when played repeatedly, this is a Nash equilibrium of the repeated game. Nevertheless, multiagent interactions are often modeled in the wider framework of stochastic (Markov) games; a very interesting direction for future research is generalizing our results to such games.

Another way to extend our research is to derive results concerning other equilibrium concepts, such as correlated equilibria [11]. We are motivated to explore this solution concept specifically, since there has been work on learning correlated equilibria [8]. In the context of manipulation, the problems GM, BGM and SGM can be defined exactly as before, with the exception that the equilibria are correlated instead of Nash. Problems that are reformulated for correlated equilibria are expected to be easier, as correlated equilibria can be found in polynomial time.

6 Acknowledgment

This work was partially supported by grant #039-7582 from the Israel Science Foundation.

References

1. Y. Bachrach and J. S. Rosenschein. Achieving allocatively-efficient and strongly budget-balanced mechanisms in the network flow domain for bounded-rational agents. In *Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence*, pages 1653–1654, 2005.
2. M. Bowling and M. Veloso. Multiagent learning using a variable learning rate. *Artificial Intelligence*, 136:215–250, 2002.
3. M. Bowling and M. Veloso. Existence of multiagent equilibria with limited agents. *Journal of Artificial Intelligence Research*, 22:353–384, 2004.
4. V. Conitzer, J. Lang, and T. Sandholm. How many candidates are needed to make elections hard to manipulate? In *Proceedings of the International Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 201–214, 2003.
5. V. Conitzer and T. Sandholm. Complexity results about Nash equilibria. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 765–771, 2003.
6. V. Conitzer and T. Sandholm. Communication complexity as a lower bound for learning in games. In *Proceedings of the 21st International Conference on Machine Learning*, pages 185–192, 2004.
7. I. Gilboa and E. Zemel. Communication complexity as a lower bound for learning in games. *Games and Economic Behavior*, pages 80–93, 1989.
8. A. Greenwald and K. Hall. Correlated Q-learning. In *Proceedings of the 20th International Conference on Machine Learning*, pages 242–249, 2003.
9. J. Hu and M. Wellman. Nash Q-learning for general-sum stochastic games. *Journal of Machine Learning Research*, 4:1039–1069, 2004.
10. M. L. Littman. Friend-or-foe Q-learning in general-sum games. In *Proceedings of the 18th International Conference on Machine Learning*, pages 322–328, 2001.
11. M. J. Osborne and A. Rubinstein. *A Course in Game Theory*. MIT Press, 1994.
12. C. Papadimitriou. Algorithms, games, and the Internet. In *Proceedings of the 33rd Annual ACM Symposium on the Theory of Computing*, pages 749–753, 2001.
13. A. D. Procaccia and J. S. Rosenschein. Junta distributions and the average-case complexity of manipulating elections. In *Proceedings of the 5th International Joint Conference on Autonomous Agents and Multiagent Systems*, 2006. To appear.
14. T. Ui. Robust equilibria of potential games. *Econometrica*, 69:1373–1380, 2001.