

Biased Games

Ioannis Caragiannis
University of Patras
caragian@ceid.upatras.gr

David Kurokawa
Carnegie Mellon University
dkurokaw@cs.cmu.edu

Ariel D. Procaccia
Carnegie Mellon University
arielpro@cs.cmu.edu

Abstract

We present a novel extension of normal form games that we call *biased games*. In these games, a player’s utility is influenced by the distance between his mixed strategy and a given *base strategy*. We argue that biased games capture important aspects of the interaction between software agents. Our main result is that biased games satisfying certain mild conditions always admit an equilibrium. We also tackle the computation of equilibria in biased games.

1 Introduction

Since the seminal work of Nash (1950), the notion of *Nash equilibrium* has been the cornerstone of game theory (and its interaction with AI, and computer science more broadly). Nash equilibrium is most commonly employed as a solution concept for *normal form games*, which are defined by a set of n players, their sets of *pure strategies*, and their *payoff functions* that map profiles of pure strategies (reflecting the choice of each player) to the player’s payoff. In Nash equilibrium, each player’s choice of strategy is a *best response* to other players’ strategies, in the sense that the player cannot secure a higher payoff by deviating to a different strategy. Nash’s key mathematical contribution was to show that if players have the option of playing *mixed strategies* — probability distributions over pure strategies — then a Nash equilibrium always exists.

More specifically, the payoffs under a mixed strategy profile are simply the expected payoffs. To compute this expectation, note that the probability of each possible pure strategy profile (s_1, \dots, s_n) is the product of probabilities of each player i playing strategy s_i , and the payoffs for this pure strategy profile are given by the players’ payoff functions. Importantly, even when players play mixed strategies, their utilities are completely determined by the payoff functions, which only take pure strategies as input.

In this paper, we are interested in a fundamentally different way in which a player’s mixed strategy can directly affect his utility. Specifically, the player may be biased towards (or away from) a specific *base strategy*, so his utility may also depend on the distance between his mixed strategy and the base strategy.

We believe that these issues have been largely overlooked, as normal form games are typically seen as *one-shot* interactions: the mixed strategy is only important insofar as it selects a pure strategy. However, the mixed strategy itself can play a more significant role in some settings, justifying the preceding notions of bias:

- In computational environments (e.g., networks), the mixed strategies of software agents can be encoded as programs that are submitted to a server, and therefore the mixed strategies themselves are visible to certain parties. Such game-theoretic settings were nicely motivated in the work of Rozenfeld and Tennenholtz (2007), and their justification is also implicit in the earlier work of Tennenholtz (2004). Once mixed strategies are visible, bias towards certain mixed strategies can arise due to *social norms* — agents are expected to play certain strategies, *mediation* (Monderer and Tennenholtz 2009; Rozenfeld and Tennenholtz 2007) — agents are *told* to play certain strategies, and *privacy* — certain mixed strategies reveal more about the agent’s preferences than others.
- In other settings, mixed strategies may be instantiated multiple times before the agents actually interact. For example, *security games* (Tambe 2012) are 2-player games played by a defender and an attacker. The defender’s strategy specifies a random allocation of security resources to potential targets, and the attacker’s strategy pinpoints a target that will be attacked. It is typically assumed that the defender would play a mixed strategy for a period of time before the attacker makes his move. The crux of this example is that redeploying security resources (such as boats, in the case of the US Coast Guard) is costly, and different instantiations of a mixed strategy lead to different deployments. This can bias the defender, say, towards pure strategies, or away from high-entropy strategies. While security games are often viewed as Stackelberg games (where the defender moves first), they have also been studied as normal form games that are solved using Nash equilibrium (Korzhyk et al. 2011).
- Rather than representing an agent’s actual utility, a bias towards a specific strategy may be hard-coded in order to lead the agent closer to certain behaviors that are known to be beneficial. For example, state-of-the-

art computer poker agents rely on computing an approximation of one equilibrium strategy (Sandholm 2010; Billings et al. 2003). One can bias agents towards strategies that, e.g., maximize opponent exploitation or minimize exploitability, and directly find an equilibrium of the biased game (as we will show). This approach is somewhat reminiscent of the work of Johanson et al. (2007), but is fundamentally different from most existing approaches for opponent exploitation, which create a model for the opponent’s strategy that is biased towards a pre-computed equilibrium strategy (of the original game), such that it is consistent with the agent’s observations of its play thus far (Ganzfried and Sandholm 2011); or perform the bias on the agent’s own strategy, requiring its exploitative strategy to be close to an equilibrium strategy of the original game (Ganzfried and Sandholm 2012).

1.1 Overview of Our Model and Results

Our definition of a *biased game* starts from a normal form game, but lets the utility of player i for a strategy profile $p = (p_1, \dots, p_n)$ be his (expected) payoff in the normal form game, minus a *bias term* of the form $f_i(\|p_i - \bar{p}_i\|)$. Here, $f_i(\cdot)$ is some real-valued function (possibly 0 everywhere), $\|\cdot\|$ is a norm, and \bar{p}_i is the *base strategy* of i . Given a biased game that includes all these components, we can define an equilibrium in the same way that the Nash equilibrium is defined (each strategy is a best response to other strategies).

We remark that a biased game and its equilibria cannot be captured as a normal form game and its Nash equilibria due to the potential nonlinearity of the utilities. From a mathematical viewpoint, we believe that biased games are a very natural (strict) generalization of normal form games. And from a conceptual viewpoint, as we have argued above, they provide a new perspective on important, AI-related problems.

In §3 we prove our main result: the existence of equilibria in biased games. Specifically, we show that an equilibrium always exists if each f_i is a non-decreasing continuous convex function, and the norms are all L_p norms. We also construct a biased game that satisfies these conditions, except for having decreasing f_i functions, in which equilibria do not exist.

In §4 we make some progress on the computation of equilibria in biased games, by (significantly) generalizing a basic algorithm for Nash equilibrium computation.

1.2 Related Work

A line of work that is somewhat related to ours includes the classic paper on *psychological games* (Geanakoplos, Pearce, and Stacchetti 1989), and influential papers that extended these ideas (see, e.g., (Battigalli and Dufwenberg 2009)). Psychological games are extensions of normal form games, where each player’s utility depends on: (i) the mixed strategies played by all players (as usual), and (ii) his *beliefs*. A psychological game also includes, for each player, a *coherent response* mapping, which maps the beliefs held by others and strategies played by others to the player’s belief. A *psychological equilibrium* consists of a vector of beliefs, and a

vector of strategies, such that the beliefs are coherent, and given these beliefs, the strategies form a Nash equilibrium.

Of course, in order to obtain technical results (such as equilibrium existence), one must restrict the structure of these various components of the game. One may wonder, though, whether our bias term can be encoded in a player’s beliefs. This is not the case, because in biased games the utility functions are defined in a way that, when the beliefs are fixed, we obtain a normal form game. That is, the beliefs only play a role to the extent that they determine the payoffs in the induced normal form game. Therefore, the psychological games framework cannot model our novel notion of bias.

2 Our Model

Before we begin, it will be helpful to establish some notation which we will use throughout the paper.

- $\forall m : [m] = \{1, 2, \dots, m\}$.
- $\mathbb{P}^n = \{x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 1, \forall i : x_i \geq 0\}$ is the space of allowed probability vectors in n dimensions.
- Given a vector $x \in \mathbb{R}^n$, we define (y, x_{-i}) as $(x_1, x_2, \dots, x_{i-1}, y, x_{i+1}, x_{i+2}, \dots, x_n)$.

We begin with the definition of the centerpiece of study in game theory.

Definition 1. A *normal form game* with n players is given by:

- An n -tuple of finite strategy spaces (S_1, S_2, \dots, S_n) .
- An n -tuple of multilinear payoff functions (T_1, T_2, \dots, T_n) where

$$T_i : \mathbb{P}^{|S_1|} \times \mathbb{P}^{|S_2|} \times \dots \times \mathbb{P}^{|S_n|} \rightarrow \mathbb{R}$$

is linear in each of its n parameters for all i .

In such a game, player i plays a *mixed strategy*, which is a probability distribution $p_i \in \mathbb{P}^{|S_i|}$ over his finite strategy space S_i , and receives an (expected) utility of $T_i(p)$ where $p = (p_1, p_2, \dots, p_n)$.

Alternatively, one can think of the payoff functions as assigning a payoff to every profile of *pure* strategies in $S_1 \times \dots \times S_n$. These functions induce payoffs for a mixed strategy profile by calculating expected payoffs. We will find it more convenient, though, to think of the domain of the functions T_i as mixed strategy profiles.

Importantly, when using a normal form game to analyze or predict actions of agents it is sensible to assume they act rationally: they will maximize their utility to the best of their ability. As a result, the equilibrium notion where each agent i *best responds* to other’s strategies by choosing a distribution p_i such that $T_i(p) \geq \max_{p'_i} T_i(p'_i, p_{-i})$ (i.e., such that no other mixed strategy gives a higher payoff) is a natural focal point of study. This is known as a *Nash Equilibrium* of a normal form game, and is so named due to Nash’s proof of its existence in any such game (Nash 1950).

The main focus of our paper is on an extension of normal form games that we call *biased games*. Informally, a biased game is a normal form game but where the utilities

of the players are determined by a multilinear payoff function (similarly to the standard setting) summed to a function of the distance of the player's strategy to some base strategy.

Definition 2. A *biased game* with n players includes an n -tuple of finite strategy spaces (S_1, S_2, \dots, S_n) . For all i , player i 's utility for a mixed strategy profile is given by

$$u_i(p) := T_i(p) - f_i(\|p_i - \bar{p}_i\|),$$

where:

- $p = (p_1, p_2, \dots, p_n)$.
- The T_i are multilinear payoff functions similar to the classic setting.
- The f_i are real-valued functions.
- The norms $\|\cdot\|$ are any valid norms (possibly different for different players).¹
- The \bar{p}_i describe *base strategies* in $\mathbb{R}^{|S_i|}$ — it is not necessary, but intuitive, that they be in $\mathbb{P}^{|S_i|}$.

For ease of exposition we define the *bias term* as $f_i(\|p_i - \bar{p}_i\|)$.

Notice that despite the name *biased game*, we allow that all $f_i = 0$ — which reduces immediately to the definition of a standard game.

Definition 3. An *equilibrium* of a biased game is a strategy profile p such that $u_i(p) \geq u_i(q_i, p_{-i})$ for all $i \in [n]$ and all $q_i \in \mathbb{P}^{|S_i|}$.

Below we illustrate these definitions using a simple example.

Example 1. Let us consider a security game with two players, the defender and the attacker, and two targets. The defender has one resource that can cover a single target. He receives a utility of 1 if he catches the attacker and 0 otherwise. Similarly, the attacker receives a utility of 1 if he attacks an undefended target and 0 otherwise. Now suppose the defender has a base strategy of $[3/4, 1/4]$ (for pedagogical purposes). The utilities may then be described by:

$$u_1(p_1, p_2) = p_1^\top \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} p_2 - 2 \left\| p_1 - \begin{bmatrix} 3/4 \\ 1/4 \end{bmatrix} \right\|_2^2$$

$$u_2(p_1, p_2) = p_1^\top \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} p_2$$

If we let

$$p_1 = \begin{bmatrix} x \\ 1-x \end{bmatrix}, \quad p_2 = \begin{bmatrix} y \\ 1-y \end{bmatrix},$$

then simple analysis shows that simultaneously maximizing u_1 w.r.t. x and u_2 w.r.t. y , is equivalent to:

$$x = \frac{2y+5}{8}$$

$$y = \begin{cases} 1 & \text{if } x < 1/2 \\ 0 & \text{if } x > 1/2 \\ \text{anything} & \text{otherwise} \end{cases}$$

We can then see that the only equilibrium is for $x = 5/8$ and $y = 0$.

¹We do not attach an index to the norm to avoid confusion, as we later use subscripts for L_p norms and superscripts for exponents.

3 Existence of Equilibria

In this section we examine the question of existence of equilibria in biased games, and prove our main result:

Theorem 1. Every biased game in which f_i is a non-decreasing continuous convex function for all $i \in [n]$, and the norms are all L_p norms, has an equilibrium.

Proof. Consider the function $h(p) = (q_1, q_2, \dots, q_n)$ where $p_i, q_i \in \mathbb{P}^{|S_i|}$ are probability distributions over i 's actions and $q_i = \arg \max_{p'_i} v_i(p'_i, p)$ with $v_i(p'_i, p) = u_i(p'_i, p_{-i}) - \|p'_i - p_i\|_2^2$. We first show that this function is a well-defined² continuous function and thus, because it acts upon a convex compact set, must have a fixed point by Brouwer's theorem. We then proceed to show that any fixed point of h must be an equilibrium.

Lemma 1. h is well-defined.

Proof. Let i be given. We show that q_i is well-defined.

Since $v_i(p'_i, p)$ — as a function of p'_i — is a continuous function on a compact space it must achieve its maximum. It therefore suffices to show that there exists a unique maximizer to v_i . Suppose for the purposes of contradiction there exist two such q_i , denoted x and y . Then let $\alpha \in (0, 1)$ and $z = \alpha x + (1 - \alpha)y$. Now consider the value v_i would achieve at z :

$$v_i(z, p) = u_i(z, p_{-i}) - \|z - p_i\|_2^2$$

$$= T_i(z, p_{-i}) - f_i(\|z - \bar{p}_i\|) - \|z - p_i\|_2^2.$$

Specifically, let us consider each term separately:

- $T_i(z, p_{-i}) = T_i(\alpha x + (1 - \alpha)y, p_{-i})$
 $= \alpha T_i(x, p_{-i}) + (1 - \alpha)T_i(y, p_{-i})$.
- $f_i(\|z - \bar{p}_i\|)$
 $= f_i(\|\alpha x + (1 - \alpha)y - \alpha \bar{p}_i - (1 - \alpha)\bar{p}_i\|)$
 $= f(\|\alpha(x - \bar{p}_i) + (1 - \alpha)(y - \bar{p}_i)\|)$
 $\leq f(\alpha\|x - \bar{p}_i\| + (1 - \alpha)\|y - \bar{p}_i\|)$
 $\leq \alpha f(\|x - \bar{p}_i\|) + (1 - \alpha)f(\|y - \bar{p}_i\|)$.

where we have used the triangle inequality and the definition of convexity.

- $\|z - p_i\|_2^2 = \|\alpha x + (1 - \alpha)y - p_i\|_2^2$
 $= \|\alpha(x - p_i) + (1 - \alpha)(y - p_i)\|_2^2$
 $\leq (\alpha\|x - p_i\|_2 + (1 - \alpha)\|y - p_i\|_2)^2$
 $\leq \alpha\|x - p_i\|_2^2 + (1 - \alpha)\|y - p_i\|_2^2$

where we have again used the triangle inequality and the definition of convexity. Importantly, there are two critical differences between this term and the previous one.

First, in this case the first inequality (created due to the use of the triangle inequality of $\|\cdot\|_2$) achieves equality if and

²We must prove that the arg max produces exactly one probability distribution.

only if $x - p_i = \beta(y - p_i)$ for some $\beta \geq 0$. This is a well-known fact of L_p norms³ and intuitively the condition is equivalent to the two vectors facing the same direction.

Second, in this case the second equality (created due to the use of convexity) is over a strictly convex function. We thus have equality if and only if $\|x - p_i\|_2 = \|y - p_i\|_2$. That is, the two vectors have the same magnitude.

From these two observations, we see that we have equality if and only if $x - p_i = y - p_i$ or more explicitly, $x = y$. As we are assuming this is not the case, we have strict inequality.

With this analysis of each term in separation, we get the following:

$$\begin{aligned} v_i(z, p) &= T_i(z, p_{-i}) - f_i(\|z - \bar{p}_i\|) - \|z - p_i\|_2^2 \\ &> \alpha T_i(x, p_{-i}) + (1 - \alpha) T_i(y, p_{-i}) \\ &\quad - \alpha f(\|x - \bar{p}_i\|) - (1 - \alpha) f(\|y - \bar{p}_i\|) \\ &\quad - \alpha \|x - p_i\|_2^2 - (1 - \alpha) \|y - p_i\|_2^2 \\ &= \alpha v_i(x, p) + (1 - \alpha) v_i(y, p) \\ &= \alpha v_i(x, p) + (1 - \alpha) v_i(x, p) \\ &= v_i(x, p). \end{aligned}$$

Thus, z is strictly better than the assumed maximizers x and y ! This is a clear contradiction. \square (Lem 1)

Lemma 2. h is continuous.

Proof. We first establish some notation which we will use throughout the proof.

- P will be the set of values that p can be drawn from.
- $\|\cdot\|$ will be the L_2 norm for vectors in \mathbb{R}^n and for $p \in P$, $\|p\|^2 = \sum_{i=1}^n \|p_i\|^2$. The induced metrics of these norms will allow us to formally discuss continuity.
- $p \in P$ will denote the arbitrary input point to h for which we wish to show continuity.

Now it suffices to show that q_i is continuous at p for arbitrary i . Suppose $\varepsilon > 0$ is given. We wish to show that there exists a δ such that $\|p - p'\| < \delta$ implies $\|q_i - q'_i\| < \varepsilon$ where $q_i = \arg \max_x v_i(x, p)$ and $q'_i = \arg \max_x v_i(x, p')$.

Let $X = \{x \mid \|x - q_i\| < \varepsilon\}$ and assume that X^c is non-empty (if this is not the case, decrease the size of ε until this is the case). Furthermore, as X^c is a closed subset of a compact set, it is compact as well. Now note that $v_i(y, p)$ is a continuous function of y and therefore achieves its maximum on the compact set X^c , denoted M_ε . Importantly, Lemma 1 shows that $M \neq M_\varepsilon$ where

$$M = \max_x v_i(x, p) = v_i(q_i, p).$$

We now further observe that $v_i(x, q)$ is a continuous function of q on a compact set and thus, by the Heine-Cantor theorem is uniformly continuous. We therefore have that there exists some $\delta > 0$ such that if $\|p - p'\| < \delta$ then

$$|v_i(x, p) - v_i(x, p')| < \frac{M - M_\varepsilon}{2}$$

³This is due to the biconditional conditions for equality in Minkowski's inequality.

for all x . Two important consequences of this are:

- $v_i(q_i, p') = v_i(q_i, p) - v_i(q_i, p) + v_i(q_i, p')$
 $\geq v_i(q_i, p) - |v_i(q_i, p) - v_i(q_i, p')|$
 $> M - \frac{M - M_\varepsilon}{2} = \frac{M + M_\varepsilon}{2}$
- For all $y \in X^c$: $v_i(y, p') = v_i(y, p) - v_i(y, p) + v_i(y, p')$
 $\leq v_i(y, p) + |v_i(y, p) - v_i(y, p')|$
 $< M_\varepsilon + \frac{M - M_\varepsilon}{2} = \frac{M + M_\varepsilon}{2}$

Together, these imply that $v_i(q_i, p') > v_i(x, p')$ and so $\arg \max_x v_i(x, p') \in X$. Thus, $\|q_i - q'_i\| < \varepsilon$. \square (Lem 2)

We have shown that h is indeed well-defined and continuous. By Brouwer's fixed point theorem, it must therefore have a fixed point. Our next lemma demonstrates that such fixed points are equilibria.

Lemma 3. Any fixed point of h represents an equilibrium in the biased game.

Proof. Suppose for the purposes of contradiction we have a fixed point of h that is not an equilibrium. That is, suppose $h(p) = p$ and there exists some i such that player i gains by the unilateral deviation from p_i to p'_i . Now let $\alpha \in [0, 1]$ and $z = \alpha p'_i + (1 - \alpha) p_i$. Then we have:

$$\begin{aligned} v_i(z, p) &= u_i(z, p_{-i}) - \|z - p_i\|_2^2 \\ &= T_i(z, p_{-i}) - f_i(\|z - \bar{p}_i\|) - \|z - p_i\|_2^2. \end{aligned}$$

Similarly to the proof of Lemma 1 we find that

$$T_i(z, p_{-i}) = \alpha T_i(p'_i, p_{-i}) + (1 - \alpha) T_i(p_i, p_{-i})$$

and

$$f_i(\|z - \bar{p}_i\|) \leq \alpha f(\|p'_i - \bar{p}_i\|) + (1 - \alpha) f(\|p_i - \bar{p}_i\|).$$

Moreover,

$$\begin{aligned} \|z - p_i\|_2^2 &= \|\alpha p'_i + (1 - \alpha) p_i - p_i\|_2^2 \\ &= \|\alpha(p'_i - p_i)\|_2^2 \\ &= \alpha^2 \|p'_i - p_i\|_2^2. \end{aligned}$$

We therefore find that:

$$\begin{aligned} v_i(z, p) &= T_i(z, p_{-i}) - f_i(\|z - \bar{p}_i\|) - \|z - p_i\|_2^2 \\ &\geq \alpha T_i(p'_i, p_{-i}) + (1 - \alpha) T_i(p_i, p_{-i}) \\ &\quad - \alpha f(\|p'_i - \bar{p}_i\|) - (1 - \alpha) f(\|p_i - \bar{p}_i\|) \\ &\quad - \alpha^2 \|p'_i - p_i\|_2^2 \\ &= \alpha u_i(p'_i, p_{-i}) + (1 - \alpha) u_i(p_i, p_{-i}) - \alpha^2 \|p'_i - p_i\|_2^2. \end{aligned}$$

Now let us consider this lower bound, call LB . For $\alpha = 0$ we have $LB = u_i(p_i, p_{-i})$ — which is tight (as $z = p_i$ in this case). That is, we have equality. Moreover, we have

$$\frac{\partial}{\partial \alpha} LB|_{\alpha=0} = u_i(p'_i, p_{-i}) - u_i(p_i, p_{-i}).$$

4 Computation of Equilibria

Importantly, we assumed that it was advantageous for player i to deviate from p_i to p'_i and thus we have that $u_i(p'_i, p_{-i}) - u_i(p_i, p_{-i}) > 0$. Therefore, there exists a small α such that

$$\begin{aligned} v_i(z, p) &> u_i(p_i, p_{-i}) \\ &= u_i(p_i, p_{-i}) - \|p_i - p_i\|_2^2 \\ &= v_i(p_i, p). \end{aligned}$$

This is a clear contradiction. \square (Lem 3)

To conclude, we have argued that h has a fixed point, and every fixed point is an equilibrium of the biased game. This completes the proof of the theorem. \square (Thm 1)

Let us now examine the assumptions that Theorem 1 makes. The assumption that the f_i are continuous and convex, and that the norms are L_p norms, can be seen as mild technical assumptions. But the assumption that the f_i are non-decreasing has more conceptual bite.

What would a biased game with decreasing f_i look like? For example, recall our discussion of security games, where a mixed strategy equates to a potentially costly redeployment of resources. One of the possible interpretations is to set the base strategy to be the uniform distribution (that is, a uniformly random assignment of resources), and let the defender be biased *away* from this strategy, that is, the f_i functions are negative and strictly decreasing.

Unfortunately, it turns out that Theorem 1 is simply not true for decreasing f_i functions: a biased game may not admit equilibria even when all the other assumptions hold, i.e., the f_i are continuous convex functions and the norms are L_p norms. Below we construct such an example.

Example 2. Suppose we have a two-player game with:

$$T_1 = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}, \quad T_2 = \begin{bmatrix} 0 & 4 \\ 0 & 0 \end{bmatrix},$$

$$B_1 \left(\begin{bmatrix} x \\ 1-x \end{bmatrix} \right) = - \left\| \begin{bmatrix} x \\ 1-x \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\|_2^2,$$

and

$$B_2 \left(\begin{bmatrix} y \\ 1-y \end{bmatrix} \right) = - \left\| \begin{bmatrix} y \\ 1-y \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\|_2^2,$$

where the B_i describe the bias terms. Then the utility of the (row) player 1 as a function of x and y is: $u_1(x, y) = 2(x^2 + 2(y-1)x + 1)$. Similarly, the utility of the (column) player 2 is: $u_2(x, y) = 2(y^2 - 2xy + 2x)$. Now note that as u_1 is an upward-facing parabola in x , its maximum over the set $x \in [0, 1]$ is reached at one of the endpoints (i.e. $x \in \{0, 1\}$). So let us consider these two cases.

Suppose first that $x = 0$. Then $u_2(x, y) = u_2(0, y) = 2y^2$ and so u_2 is maximized for $y \in [0, 1]$ when $y = 1$. However, this implies that $u_1(x, y) = u_1(x, 1) = 2x^2 + 2$ and thus u_1 is maximized when $x = 1$ — a contradiction.

Now suppose instead that $x = 1$. Then $u_2(x, y) = u_2(1, y) = 2(y^2 - 2y + 2)$ and so u_2 is maximized for $y \in [0, 1]$ when $y = 0$. However, this implies that $u_1(x, y) = u_1(x, 0) = 2(x^2 - 2x + 1)$ and thus u_1 is maximized when $x = 0$ — a contradiction.

In this section we investigate the computation of equilibria in biased games. From the outset our expectations are quite low, as even in normal form games, computing a Nash equilibrium is computationally hard (Daskalakis, Goldberg, and Papadimitriou 2009; Conitzer and Sandholm 2008). However, there are various equilibrium computation algorithms that work well on average or in special cases, such as the famous Lemke-Howson (1964) Algorithm.

A major challenge in our (much more general) setting is that algorithms for computing Nash equilibria in normal form games typically rely on the property that if p is a Nash equilibrium and s_i is a pure strategy in the support of p_i , then s_i is itself a best response to p_{-i} . This is clearly not the case when it comes to biased games and their equilibria. This can be seen in Example 1, where pure strategies in the support of p_1 are never best responses to p_2 .

To further illustrate the difficulties that one encounters in our setting, we note that in normal form games with two players and rational payoffs, there exist Nash equilibria defined by rational probabilities (Nash 1950), and this is, of course, true if there is only one player. In contrast, below we give an example of a biased game with *one player* and rational parameters, but only irrational equilibrium. Note that this also demonstrates the strict generality of this setting to the standard setting.

Example 3. Consider the game where the sole player has two strategies with payoffs 1 and 0, and the bias term is $\|p - [1/2, 1/2]^T\|_2^4$. A simple analysis then yields that the sole equilibrium is to play the first strategy with probability $\frac{1}{2} + \frac{1}{2\sqrt{3}}$.

Due to these difficulties, we focus on certain subsets of biased games (which, in particular, circumvent Example 3). Specifically, we consider the two-player (and later, more generally the n -player) setting with a bias term of the form $c\|\cdot\|_1$ or $c\|\cdot\|_2^2$ where $c \geq 0$ is some constant. Crucially, this still generalizes the classic setting. Our goal is to generalize the (extremely simple) *support enumeration* algorithm for computing Nash equilibria (see, e.g., (von Stengel 2007, Algorithm 3.4)).

Let us first consider the L_2 case: player i has a bias term of the form $c_i\|\cdot\|_2^2$ where $c_i \geq 0$. Recall that for each player i , if the strategy of the other player is fixed, then i simply wishes to maximize his utility $u_i(p_i)$. That is, for every player i , we wish to have that $u_i(p_i)$ is maximized subject to the constraints that the entries of p_i are nonnegative and sum to one. The Karush-Kuhn-Tucker (KKT) conditions on a player's utility then give necessary and sufficient conditions for maximization — sufficiency is due to the concavity of the objective and the affine nature of the constraints. Thus, equilibrium computation is equivalent to solving the

following system. For all i and pure strategies j of i :

$$\begin{aligned} p_{i,j} &\geq 0 \\ \mu_{i,j} &\geq 0 \\ \mu_{i,j} p_{i,j} &= 0 \\ p_i^\top \vec{1} &= 1 \\ \text{STANDARD}(i,j) - 2 \text{BIAS}(i,j) + \lambda_i + \mu_{i,j} &= 0. \end{aligned}$$

where

$$\text{STANDARD}(i,j) = \frac{d}{dp_{i,j}} T_i(p_1, p_2),$$

and

$$\text{BIAS}(i,j) = c_i (p_{i,j} - \bar{p}_{i,j}).$$

Crucially, aside from the $\mu_{i,j} p_{i,j} = 0$ conditions, the complete characterization is then a linear feasibility program. We can thus consider the $2^{|S_1|+|S_2|}$ possibilities (recall that $|S_i|$ is the number of pure strategies of player i) of which one of $\mu_{i,j}$ and $p_{i,j}$ are zero to find the equilibria. That is, for every player i and strategy j of i we set one of $\mu_{i,j}$ and $p_{i,j}$ to zero and solve the resulting linear program. This computes an equilibrium exactly (albeit in exponential time).

Dealing with bias terms of the form $c_i \|\cdot\|_1$ where $c_i \geq 0$ is largely analogous. The important difference appears due to the discontinuity of the derivative of the L_1 norm. Via a simple case analysis which we omit here, we see that for all i and pure strategies j of i :

$$\begin{aligned} p_{i,j} &\geq 0 \\ \mu_{i,j} &\geq 0 \\ \mu_{i,j} p_{i,j} &= 0 \\ p_i^\top \vec{1} &= 1. \end{aligned}$$

and at least one of two discontinuity-caused requirements must be satisfied: either

$$\begin{aligned} \text{STANDARD}(i,j) - \text{BIAS}(i,j) + \lambda_i + \mu_{i,j} &= 0 \\ \text{BIAS}(i,j) &\geq 0 \end{aligned}$$

or

$$\begin{aligned} \text{STANDARD}(i,j) + \text{BIAS}(i,j) + \lambda_i + \mu_{i,j} &= 0 \\ \text{BIAS}(i,j) &\leq 0. \end{aligned}$$

Therefore, to compute an equilibrium when all players have $c \|\cdot\|_1$ bias terms we can consider the $4^{|S_1|+|S_2|}$ linear programs that arise. That is, for every player i who has a $c \|\cdot\|_1$ bias term, and strategy j of i , we set one of $\mu_{i,j}$ and $p_{i,j}$ to zero (as before) and also determine which of the two discontinuity-caused requirements we will choose to include in our linear program — the players who have $c \|\cdot\|_2^2$ bias terms are dealt with as before.

The next statement summarizes the discussion above.

Theorem 2. *Given a biased game with two players and bias terms of the form $c_i \|p_i - \bar{p}_i\|_2^2$ or $c_i \|p_i - \bar{p}_i\|_1$, an equilibrium can be computed by solving at most $4^{|S_1|+|S_2|}$ linear programs.*

Although we have discussed the algorithm only for the two-player setting, this approach can be extended to multiple players at the cost of solving multilinear programs instead of linear ones. The discrepancy is caused due to the $\text{STANDARD}(i,j)$ term: the derivative in its expression is multilinear for $n \geq 3$. This is a generalization of the sequential approach discussed by Widger and Grosu (2009) for normal form games.

5 Discussion

In our model of biased games, a player's bias term is $f_i(\|p_i - \bar{p}_i\|)$. Consider a more general definition of the bias term, $f_i(\|p_i - (\bar{p}_i + \sum_{j \in [n]} \gamma_{i,j} p_j)\|)$, where the $\gamma_{i,j}$ are real weights such that $\gamma_{i,j} = 0$ if players i and j do not have the same number of pure strategies. The new term $\sum_{j \in [n]} \gamma_{i,j} p_j$ captures a different type of bias, towards playing a mixed strategy that is similar to, or different from, those played by other players. This extension can help capture well-known social phenomena, such as homophily (the propensity of individuals to bond with others that are similar to themselves). Interestingly, all our results (existence and computation) also hold for the more general definition of biased games. However, it seems poorly motivated in the context of the settings discussed in Section 1, which is why we chose to simplify the presentation by focusing on the more restricted definition of biased games.

We feel that our existence result (Theorem 1) is rather general, but the computation of equilibria in biased games poses difficult problems even for zero-sum games. Still, we have been able to identify some rather interesting properties of two-player *zero-sum* biased games with bias terms of the form $c \|p_i - d \bar{p}_i\|_2^2$ for some constants $c > 0$ and d , which lead us to conjecture that, under these conditions, there exists a unique equilibrium. Of course, the conditions of this conjecture are quite restrictive. In order to tackle equilibrium computation in biased games under assumptions as general as those of Theorem 1, fundamentally new ideas are required.

Acknowledgements

We thank Sam Ganzfried and the anonymous AAAI'14 reviewers for helpful comments. Kurokawa and Procaccia were supported in part by the NSF under grants NSF IIS-1350598 and NSF CCF-1215883. Caragiannis was supported in part by the European Social Fund and Greek national funds through the research funding program Thales on "Algorithmic Game Theory" and by a Caratheodory research grant from the University of Patras.

References

- Battigalli, P., and Dufwenberg, M. 2009. Dynamic psychological games. *Journal of Economic Theory* 144(1):1–35.
- Billings, D.; Burch, N.; Davidson, A.; Holte, R.; Schaeffer, J.; Schauenberg, T.; and Szafron, D. 2003. Approximating game-theoretic optimal strategies for full-scale poker. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI)*, 661–668.

- Conitzer, V., and Sandholm, T. 2008. Complexity results about Nash equilibria. *Games and Economic Behavior* 63(2):621–641.
- Daskalakis, C.; Goldberg, P. W.; and Papadimitriou, C. H. 2009. The complexity of computing a Nash equilibrium. *SIAM Journal on Computing* 39(1):195–259.
- Ganzfried, S., and Sandholm, T. 2011. Game theory-based opponent modeling in large imperfect-information games. In *Proceedings of the 10th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 533–540.
- Ganzfried, S., and Sandholm, T. 2012. Safe opponent exploitation. In *Proceedings of the 13th ACM Conference on Electronic Commerce (EC)*, 587–604.
- Geanakoplos, J.; Pearce, D.; and Stacchetti, E. 1989. Psychological games and sequential rationality. *Games and Economic Behavior* 1:60–79.
- Johanson, M.; Zinkevich, M.; and Bowling, M. 2007. Computing robust counter-strategies. In *Proceedings of the 21st Annual Conference on Neural Information Processing Systems (NIPS)*.
- Korzhyk, D.; Yin, Z.; Kiekintveld, C.; Conitzer, V.; and Tambe, M. 2011. Stackelberg vs. Nash in security games: An extended investigation of interchangeability, equivalence, and uniqueness. *Journal of Artificial Intelligence Research* 41:297–327.
- Lemke, C. E., and Howson, J. T. 1964. Equilibrium points of bimatrix games. *SIAM Journal on Applied Mathematics* 12(2):413–423.
- Monderer, D., and Tennenholtz, M. 2009. Strong mediated equilibrium. *Artificial Intelligence* 173(1):180–195.
- Nash, J. F. 1950. Equilibrium points in N -person games. *Proceedings of the National Academy of Sciences of the United States of America* 36:48–49.
- Rozenfeld, O., and Tennenholtz, M. 2007. Routing mediators. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI)*, 1488–1493.
- Sandholm, T. 2010. The state of solving large incomplete-information games, and application to poker. *AI Magazine* 31(4):13–32.
- Tambe, M. 2012. *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned*. Cambridge University Press.
- Tennenholtz, M. 2004. Program equilibrium. *Games and Economic Behavior* 49(2):363–373.
- von Stengel, B. 2007. Equilibrium computation for two-player games in strategic and extensive form. In Nisan, N.; Roughgarden, T.; Tardos, E.; and Vazirani, V., eds., *Algorithmic Game Theory*. Cambridge University Press. chapter 9.
- Widger, J., and Grosu, D. 2009. Parallel computation of Nash equilibria in N -player games. In *Proceedings of the 12th International Conference on Computational Science and Engineering (CSE)*, 209–215.