Cooperative Equilibrium

ABSTRACT

Nash equilibrium (NE) assumes that players always make a best response. However, this is not always true; sometimes people cooperate even it is not a best response to do so. For example, in the Prisoner's Dilemma, people often cooperate. Are there rules underlying cooperative behavior? In an effort to answer this question, we propose a new equilibrium concept: *perfect cooperative equilibrium* (PCE). PCE may help explain players' behavior in games where cooperation is observed in practice. A player's payoff in a PCE is at least as high as in any NE. However, a PCE does not always exist. We thus consider α-PCE, where α takes into account the degree of cooperation; a PCE is a 0-PCE. Every game has a Pareto-optimal *max-perfect cooperative equilibrium (M-PCE)*; that is, an α-PCE for a maximum α. We show that M-PCE does well at predicting behavior in quite a few games of interest. We provide further insight into M-PCE, at least in two-player games, by considering another generalization of PCE called *cooperative equilibrium* (CE), which takes the possibility of punishment into account. We show that a Pareto-optimal M-PCE must be a CE.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multiagent systems*

General Terms

Economics, Theory

Keywords

Cooperative equilibrium, Nash equilibrium, game theory, cooperation, punishment

1. INTRODUCTION

Nash Equilibrium (NE) assumes that players always make a best response to what other players are doing. However, this assumption does not always hold. Consider the Prisoner's Dilemma, in which two prisoners can choose either to defect or cooperate with payoffs as shown in the following table:

	Cooperate	Defect
Cooperate	(3,3)	(0,5)
Defect	(5,0)	(1,1)

Although the only best response here is to play Defect no matter what the other player does, people often do play (Cooperate, Cooperate). There are a number of other games in which Nash equilibrium does not predict actual behavior well.

In the Traveler's Dilemma [1, 2], two travelers have identical luggage, for which they both paid the same price. Their luggage is damaged (in an identical way) by an airline. The airline offers to recompense them for their luggage. They may ask for any dollar amount between \$2 and \$100. There is only one catch. If they ask for the same amount, then that is what they will both receive. However, if they ask for different amounts—say one asks for \$$m$ and the other for \$$m'$, with $m < m'$—then whoever asks for \$$m$ (the lower amount) will get \$$(m + 2)$, while the other traveler will get \$$(m - 2)$. A little calculation shows that the only NE in the Traveler's Dilemma is $(2, 2)$. (Indeed, $(2, 2)$ is the only strategy that survives iterated deletion of weakly dominated strategies and is the only rationalizable strategy; see [8] for a discussion of these solution concepts.) Nevertheless, in practice, people (even game theorists!) do not play (2,2). Indeed, when Becker, Carter, and Naeve [3] asked members of the Game Theory Society to submit strategies for the game, 37 out of 51 people submitted a strategy of 90 or higher. The strategy that was submitted most often (by 10 people) was 100. The winning strategy (in pairwise matchups against all submitted strategies) was 97. Only 3 of 51 people submitted the "recommended" strategy 2. In this case, NE is neither predictive nor normative; it is neither the behavior that was submitted most often (it was in fact submitted quite rarely) nor the strategy that does best (indeed, it did essentially the worst among all strategies submitted).

Note that in both Prisoner's Dilemma and Traveler's Dilemma, people display what might be called "cooperative" behavior. This cannot be explained by the best-response assumption of NE. Are there rules underlying cooperative behavior?

In this paper, we propose a new equilibrium concept, *perfect cooperative equilibrium* (PCE), in an attempt to characterize cooperative behavior. Intuitively, in a two-player game, a strategy profile (i.e., a strategy for each player) is a PCE if each player does at least as well as she would if the other player were best-responding. In Prisoner's Dilemma, both (Cooperate, Cooperate) and (Defect, Defect) are PCE.

Cite as: Title, Author(s), *Proc. of 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, van der Hoek, Kaminka, Luck and Sen (eds.), May, 10–14, 2010, Toronto, Canada, pp. XXX-XXX.

Copyright © 2010, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

To see why, suppose that the players are Amy and Bob. Consider the game from Amy's point of view. She gets a payoff of 3 from (Cooperate, Cooperate). No matter what she does, Bob's best response is Defect, that gives Amy a payoff of either 0 or 1 (depending on whether she cooperates or defects). Thus, her payoff with (Cooperate, Cooperate) is better than her payoff with any strategy, if Bob best-responds. The same is true for Bob. Thus, (Cooperate, Cooperate) is a PCE. The same argument shows that (Defect, Defect) is also a PCE.

This game already shows that some PCE are not NE. In Traveler's Dilemma, any strategy profile that gives each player a payoff above 99 is a PCE (see Section 2 for details). Thus, both (99, 99) and (100, 100) are PCE. Moreover, the unique NE is not a PCE. Thus, in general, PCE and NE are quite different. Note that in both examples, PCE explains the cooperative behaviors which NE does not. We can in fact show that, if a PCE exists, the payoff for each player is at least as good as it is in any NE. This makes PCE an attractive notion, especially for mechanism design.

This leads to some obvious questions. First, *why* should or do players play (their part of) a PCE? Second, does a PCE always exist? Finally, how do players choose among multiple PCE, when more than one exists?

With regard to the first question, first consider (one of) the intuitions for NE. The assumption is that players have played repeatedly, and thus have learned other players' strategies. They thus best respond to what they have learned. A NE is a stable point of this process: every players' strategy is already a best response to what the other players are doing. This intuition focuses on what players have done in the past; with PCE, we also consider the *future*. In a PCE such as (Cooperate, Cooperate) in Prisoner's Dilemma, players realize that if they deviate from the PCE, then the other player may start to best respond; after a while, they may well end up in some NE, and thus have a payoff that is guaranteed to be no better than (and is often worse than) that of the PCE. Although cooperation here (and in other games) gives a solution concept that is arguably more "fragile" than NE, players may still want to play a PCE because it gives a better payoff. Of course, we are considering one-shot games, not repeated games, so there is no future (or past); nevertheless, these intuitions may help explain why players actually play a PCE. (See Section 7 for a comparison of PCE and NE in repeated games.)

It is easy to see that a PCE does not always exist. Consider the Nash bargaining game [7]. Each of two players suggests a number of cents between 0 and 100. If their total demand is no more than a dollar, then they each get what they asked for; otherwise, they both get nothing. Each pair (x, y) with $x+y = 100$ is a NE, so there is clearly no strategy profile that gives both players a higher payoff than they get in every NE.

We define a notion of *α-PCE*, where s is an α-PCE if, playing s, each player can do at least α better than she would if the other player were best-responding (note that α may be negative). Thus, if a strategy is an α-PCE, then it is an α'-PCE for all $\alpha' \leq \alpha$. A strategy is a PCE iff it is a 0-PCE. We are most interested in *max-perfect cooperative equilibrium (M-PCE)*. A strategy is an M-PCE if it is an α-PCE, and no strategy is an α'-PCE for some $\alpha' > \alpha$. We show that every game has an M-PCE; in fact, it has a *Pareto-optimal* M-PCE (so that there is no other strategy profile where all players do at least as well and at least one does better). We show that M-PCE does well at predicting behavior in quite a few games of interest. For example, in Prisoner's Dilemma, (Cooperate, Cooperate) is the unique M-PCE; and in the Nash bargaining game, $(50, 50)$ is the unique M-PCE. As the latter example suggests, the notion of an M-PCE embodies a certain sense of fairness. In cases where there are several PCE, M-PCE gives a way of choosing among them.

Further insight into M-PCE, at least in two-player games, is provided by considering another generalization of PCE, called *cooperative equilibrium* (CE), which takes punishment into account. It is well-known that people are willing to punish non-cooperators, even at a cost to themselves (see, for example, [6, 11, 9] and the references therein). CE is defined only for two-player games. Intuitively, a strategy profile s in a two-player game is a CE if, for each player i and each possible deviation s'_i for i, either (1) i does at least as well with s as she would do if the other player j were best-responding to s'_i; or (2) all of j's best responses to s'_i result in j being worse off than he is with s, so he "punishes" i by playing a strategy s''_j that results in i being worse off. (Note that by punishing i, j may himself be worse off.)

It is almost immediate that every PCE is a CE. More interestingly, for our purposes, we show that every Pareto-optimal M-PCE is a CE. Thus, every two-player game has a CE. While CE does seem to capture reasoning often done by people, there are games where it does not have much predictive power. For example, in the Nash bargaining game, CE and NE coincide; all strategy profiles (x, y) where $x+y = 100$ are CE. CE also has little predictive power in the *Ultimatum* game [4], a well-known variant of the Nash bargaining game where player 1 moves first and proposes a division, which player 2 can either accept or reject; again, all offers give a CE. In practice, "unfair" divisions (typically, where player 2 gets less than, say, 30% of the pot, although the notion of unfairness depends in part of cultural norms) are rejected; player 2 punishes player 1 although he is worse off.

This type of punishment is not captured by CE, but can be understood in terms of M-PCE. For example, a strategy in the ultimatum game might be considered acceptable if it is close to an M-PCE; that is, if an M-PCE is an α-PCE, then a strategy might be considered acceptable if it is an α'-PCE, where $\alpha - \alpha'$ is smaller than some (possibly culturally-determined) threshold. Punishment is applied if the opponent's strategy precludes an acceptable strategy being played.

To summarize, M-PCE is a solution concept that is well-founded, has good predictive power, and may help explain when players are willing to apply punishment in games.

The rest of the paper is organized as follows. In Section 2, we introduce PCE, prove a number of properties of PCE, and give some examples. In Sections 3 and 4, we do the same for M-PCE and CE, respectively. In Section 5 we prove that, in two-player games, both a PCE and an M-PCE can be found in polynomial time, using bilinear programming. (We can also determine in polynomial time whether a PCE exists.) This is a contrast to Nash equilibrium, which is PPAD complete even in two-player games [**?**].

In Section 6, we compare M-PCE to the *coco value* (cooperative-competitive value), a notion introduced by Kalai and Kalai [**?**] that is also intended to cpature cooperative behavior in two-player games. Although the definition of the coco value is very different from that of M-PCE, it turns out that they are closely related if we consider games with *trasnferable utility*, that is, games where players can make side payments to other players. When defining the coco value, Kalai and Kalai implicitly assume that side payments are allowed. To enable us to do the comparison, we provide a technique to convert an arbitrary game to one that allows side payments that may be of independent interest, and characterize the *M-PCE value* in such games (that is, the payoff obtained in an M-PCE), both axiomatically and by providing a formula for the M-PCE value. We also provide a formula for the coco value, while Kalai and Kalai [?] provide an axiomatic characterization for it. These characterizations make it clear how closely related the two notions are.

We discuss further related work in Section 7, and conclude in Section ?? with some discussion of the notion of cooperation in games and some open problems.

2. PERFECT COOPERATIVE EQUILIBRIUM

In this section, we introduce PCE. For ease of exposition, we focus here on finite *normal-form* games $G = (N, A, u)$, where $N = \{1, \ldots, n\}$ is a finite set of players, $A = A_1 \times \ldots \times A_n$, A_i is a finite set of possible actions for player i, $u = (u_1, \ldots, u_n)$, and u_i is player i's utility function, that is, $u_i(a_1, \ldots, a_n)$ is player i's utility or payoff if the action profile $a = (a_1, \ldots, a_n)$ is played. Players are allowed to randomize. A strategy of player i is thus a distribution over actions in A_i; let S_i represent the set of player i's strategies. Let $U_i(s_1, \ldots, s_n)$ denote player i's expected utility if the strategy profile $s = (s_1, \ldots, s_n)$ is played. Given a profile $x = (x_1, \ldots, x_n)$, let x_{-i} denote the tuple consisting of all values x_j for $j \neq i$.

Definition 1. Given a game G, a strategy s_i for player i in G is a *best response* to a strategy s_{-i} for the players in $N - \{i\}$ if s_i maximizes player i's expected utility given that the other players are playing s_{-i}, that is, $U_i(s_i, s_{-i}) = \sup_{s'_i \in S_i} U_i(s'_i, s_{-i})$. Let $BR_i^G(s_{-i})$ be the set of best responses to s_{-i} in game G. We omit the superscript G if the game is clear from context.

We first define PCE for two-player games.

Definition 2. Given a two-player game G, let BU_i^G denote the best utility that player i can obtain if the other player j best responds; that is,

$$BU_i^G = \sup_{\{s_i \in S_i, s_j \in BR^G(s_i)\}} U_i(s).$$

(As usual, we omit the superscript G if it is clear from context.)

Definition 3. A strategy profile s is a *perfect cooperative equilibrium (PCE)* in a two-player game G if for all $i \in \{1, 2\}$, we have

$$U_i(s) \geq BU_i^G.$$

It is easy to show that every player does at least as well in a PCE as in a NE.

Theorem 1. *If s is a PCE and s^* is a NE in a two-player game G, then for all $i \in \{1, 2\}$, we have $U_i(s) \geq U_i(s^*)$.*

PROOF. Suppose that s is a PCE and s^* is a NE. Then, by the definition of NE, $s^*_{3-i} \in BR(s^*_i)$, so by the definition of PCE, $U_i(s) \geq U_i(s^*)$. ∎

It is immediate from Theorem 1 that a PCE does not always exist. For example, in the Nash bargaining game, a PCE would have to give each player a payoff of 100, and there is no strategy profile that has this property. Nevertheless, we continue in this section to investigate the properties of PCE; in the following two sections, we consider generalizations of PCE that are guaranteed to exist.

A strategy profile s *Pareto dominates* strategy profile s' if $U_i(s) \geq U_i(s')$ for all players i, strategy s *strongly Pareto dominates* s' if s Pareto dominates s' and $U_j(s) > U_j(s')$ for some player j; strategy s is *Pareto optimal* if no strategy profile strongly Pareto dominates s; s is a *dominant strategy profile* if it Pareto dominates all other strategy profiles.

A dominant strategy profile is easily seen to be a NE; it is also a PCE.

Theorem 2. *If s is a dominant strategy profile in a two-player game G, then s is a PCE.*

PROOF. Suppose that s is a dominant strategy profile in G. Then for all $i \in \{1, 2\}$, all $s'_i \in S_i$ and all $s'_{3-i} \in BR_{3-i}(s'_i)$, we have that $U_i(s) \geq U_i(s')$. Thus, $U_i(s) \geq BU_i$ for all i, so s is a PCE. ∎

The next result shows that a strategy profile that Pareto dominates a PCE is also a PCE. Thus, if s is a PCE, and s' makes everyone at least as well off, then s' is also a PCE. Note that this property does not hold for NE. For example, in Prisoner's Dilemma, (Cooperate, Cooperate) is not an NE, although it strongly Pareto dominates (Defect, Defect), which is an NE.

Theorem 3. *In a two-player game, a strategy profile that Pareto dominates a PCE must itself be a PCE.*

PROOF. Suppose that s is a PCE and s^* Pareto dominates s. Thus, for all $i \in N$, we have

$$U_i(s^*) \geq U_i(s) \geq BU_i.$$

Thus, s^* is a PCE. ∎

Corollary 4. *If there is a PCE in a two-player game G, there is a Pareto-optimal PCE in G (i.e., a PCE that is Pareto optimal among all strategy profiles).*

PROOF. Given a PCE s, let S^* be the set of strategy profiles that Pareto dominate s. This is a closed set, and hence compact. Let $f(s) = U_1(s) + U_2(s)$. Clearly f is a continuous function, so f takes on its maximum in S^*; that is, there is some strategy $s^* \in S^*$ such that $f(s^*) \geq f(s')$ for all $s' \in S^*$. Clearly s^* must be Pareto optimal, and since s^* Pareto dominates s, it must be a PCE, by Theorem 3. ∎

We now want to define PCE for n-player games, where $n > 2$. The problem is that "best response" is not well defined. For example, in a 3-player game, it is not clear what it would mean for players 2 and 3 to make a best response to a strategy of player 1, since what might be best for player 2 might not be best for player 3. We nevertheless want to keep the intuition that player 1 considers, for each of her possible strategies s_1, the likely outcome if she plays s_1. If there is only one other player, then it seems reasonable to expect that that player will play a best response to s_1. There are a number of ways we could define an analogue if there are more than two players; we choose an approach that both seems natural and leads to a straightforward generalization of all our results. Given an n-player game G and a strategy s_i for player i, let G_{s_i} be the $(n-1)$-player game among the players in $N - \{i\}$ that results when player i plays s_i. We assume that the players in $N - \{i\}$ respond to s_i by playing some NE in G_{s_i}. Let $NE^G(s_i)$ denote the NE of G_{s_i}. Again, we omit the superscript G if it is clear from context. We now extend the definition of PCE to n-player games for $n > 2$ by replacing $BR(s_i)$ by $NE(s_i)$. Note that if $|N| = 2$, then $NE(s_i) = BR(s_i)$, so this gives a generalization of what we did in the two-player case. As a first step, we extend the definition of BU_i^G to the multi-player case by using $NE^G(s_i)$ instead of $BR^G(s_i)$; that is,

$$BU_i^G = \sup_{\{s \in S_i, s_{-i} \in NE_i^G(s_i)\}} U_i(s).$$

(As usual, we omit the superscript G it if it is clear from context.)

Definition 4. A strategy profile s is a *perfect cooperative equilibrium (PCE)* in a game G if for all $i \in N$, we have

$$U_i(s) \geq BU_i^G.$$

With this definition, we get immediate analogues of Theorems 1, 2, 3, and Corollary 4, with almost identical proofs. Therefore, we state the results here and omit the proofs.

Theorem 5. *If s is a PCE and s^* is a NE in a game G, then for all $i \in N$, we have $U_i(s) \geq U_i(s^*)$.*

Theorem 6. *If s is a dominant strategy profile in a game G, then s is a PCE.*

Theorem 7. *A strategy profile that Pareto dominates a PCE must itself be a PCE.*

Corollary 8. *If there is a PCE in a game G, there is a Pareto-optimal PCE in G.*

We now give some examples of PCE in games of interest.

Example 1. *A coordination game:* A coordination game has payoffs as shown in the following table:

	a	b
a	(k_1, k_2)	$(0,0)$
b	$(0,0)$	$(1,1)$

It is well known that if k_1 and k_2 are both positive, then (a, a) and (b, b) are NE (there is also a NE that uses mixed strategies). On the other hand, if $k_1 > 1$ and $k_2 > 1$, then (a, a) is the only PCE; if $k_1 < 1$ and $k_2 < 1$, then (b, b) is the only PCE; and if $k_1 > 1$ and $k_2 < 1$, then there are no PCE (since, by Theorem 1, a PCE would simultaneously have to give player 1 a payoff of at least k_1 and player 2 a payoff of at least 1).

Example 2. *Prisoner's Dilemma:* Note that, in Prisoner's Dilemma, $BU_1 = BU_2 = 1$, since the best response is always to defect. Thus, a strategy profile s is a PCE iff $\min(U_1(s), U_2(s)) \geq 1$. It is immediate that (Cooperate, Cooperate) and (Defect, Defect) are PCE, and are the only PCE in pure strategies, but there are other PCE in mixed strategies. For example, ($\frac{1}{2}$Cooperate+$\frac{1}{2}$Defect, Cooperate) and ($\frac{1}{2}$Cooperate+$\frac{1}{2}$Defect, $\frac{1}{2}$Cooperate+$\frac{1}{2}$Defect) are PCE (where αCooperate + $(1-\alpha)$Defect denotes the mixed strategy where Cooperate is played with probability α and Defect is played with probability $1 - \alpha$).

Example 3. *Traveler's Dilemma:* To compute the PCE for Traveler's Dilemma, we first need to compute BU_1 and BU_2. By symmetry, $BU_1 = BU_2$. We now show that BU_1 is between $98\frac{1}{6}$ and 99. If player 1 plays $\frac{1}{2}100 + \frac{1}{6}99 + \frac{1}{6}98 + \frac{1}{6}97$, then it is easy to see that player 2's best responses are 99 and 98 (both give player 2 an expected payoff of $98\frac{5}{6}$); player 1's expected payoff if player 2 plays 99 is $98\frac{1}{6}$. Thus, $BU_1 \geq 98\frac{1}{6}$. To see that BU_1 is at most 99, suppose by way of contradiction that it is greater than 99. Then there must be strategies $s_1 = p_{100}100 + p_{99}99 + \cdots + p_2 2 \in S_1$ and $s_2 \in BR_2(s_1)$ such that $U_1(s_1, s_2) > 99$. It cannot be the case that s_2 gives positive probability to 100 (for then s_2 would not be a best response). Suppose that s_2 gives positive probability to 99. Then 99 must itself be a best response. Thus, $U_2(s_1, 99) \geq U_2(s_1, 98)$, so $101p_{100} + 99p_{99} + 96p_{98} \geq 100(p_{100} + p_{99}) + 98p_{98}$, so $p_{100} \geq p_{99} + 2p_{98}$. Since a best response by player 2 cannot put positive weight on 100, the highest utility that player 1 can get if player 2 plays a best response is if player 2 plays 99; then $U_1(s_1, 99) \leq 97p_{100} + 99p_{99} + 100p_{98} + 99(1 - p_{100} - p_{99} - p_{98})$. Since $U_1(s_1, 99) > 99$, it follows that $p_{98} > p_{100}$. This gives a contradiction. Thus, s_2 cannot give positive probability to 99. This means that s_1 does not give positive probability to either 100 or 99. But then $U_1(s_1, s_2) \leq U_1(s_1, 98) \leq 99$, a contradiction.

Since s is a PCE if $U_i(s) \geq BU_i(s)$, for $i = 1, 2$, it follows that the only PCE in pure strategies are $(100, 100)$ and $(99, 99)$. There are also PCE in mixed strategies, such as $(\frac{1}{2}100 + \frac{1}{2}99, \frac{1}{2}100 + \frac{1}{2}99)$ and $(100, \frac{2}{3}100 + \frac{1}{3}99)$.

Example 4. *Centipede game:* In the Centipede game [10], players take turns moving, with player 1 moving at odd-numbered turns and player 2 moving at even-numbered turns. There is a known upper bound on the number of turns, say 20. At each turn $t < 20$, the player whose move it is can either stop the game or continue. At turn 20, the game ends if it has not ended before then. If the game ends after an odd-numbered turn t, then the payoffs are $(2^t + 1, 2^{t-1})$; if the game ends after an even-numbered turn t, then the payoffs are $(2^{t-1}, 2^t + 1)$. Thus, if player 1 stops at round 1, player 1 gets 3 and player 2 gets 1; if player 2 stops at round 4, then player 1 gets 8 and player 2 gets 17; if player 1 stops at round 5, then player 1 gets 33 and player 2 gets 16. If the game stops at round 20, both players get over 500,000. The key point here is that it is always better for the player who moves at step t to end the game than it is to go on for one more step and let the other player end the game. Using this observation, a straightforward backward induction shows the best response for a player if he is called upon to move at step t is to end the game. Not surprisingly, the only Nash equilibrium has player 1 ending the game right away. But, in practice, people continue the game for quite a while.

To compute the PCE for the game, we need to first compute BU_1 and BU_2. If player 1 continues to the end of the game, then player 2's best response is to also continue to the end of the game, giving player 1 a payoff of 2^{19} (and player 2 a payoff of $2^{20}+1$). If we take $q_{i,j}$ to be the strategy where player i quits at turn j and $q_{i,C}$ to be the strategy where player i continues to the end of the game, then a straightforward computation shows that $q_{2,C}$ continues to be a best response to $\alpha q_{1,19} + (1 - \alpha)q_{1,C}$ as long as $\alpha \geq \frac{3 \times 2^{18}}{3 \times 2^{18}+1}$. If we take $\alpha = \frac{3 \times 2^{18}}{3 \times 2^{18}+1}$ and player 2 best responds by playing $q_{2,C}$, then player 1's utility is $2^{19} + \frac{3\times 2^{18}}{3\times 2^{18}+1}$. It is then straightforward to show that this is in fact BU_1. A similar argument shows that, if player 1 is best responding, then the best player 2 can do is to play $\beta q_{2,18} + (1-\beta)q_{2,C}$, where $\beta = \frac{3\times 2^{17}}{3\times 2^{17}+1}$. With this choice, player 1's best response is $q_{1,19}$. using this strategy for player 2, we get that $BU_2 = 2^{18} + \frac{3\times 2^{17}}{3\times 2^{17}+1}$.

It is easy to see that there is no pure strategy profile s such that $U_1(s) \geq BU_1$ and $U_2(s) \geq BU_2$. However, there are many mixed PCE. For example, every strategy profile $(q_{1,C}, s_2)$ where $s_2 = \beta q_{2,18} + (1-\beta)q_{2,C}$ and $\beta \in [1 - \frac{3\times 2^{17}}{(3\times 2^{17}+1)(3\times 2^{18}+1)}, \frac{3\times 2^{18}}{3\times 2^{18}+1}]$ is a PCE.

3. α-PERFECT COOPERATIVE EQUILIBRIUM

In this section, we consider a more quantitative version of PCE called α-PCE, which intuitively takes into account the degree of cooperation exhibited by a strategy profile.

Definition 5. A strategy profile s is an *α-PCE* in a game G if $U_i(s) \geq \alpha + BU_i^G$ for all $i \in N$.

Clearly, if s is an α-PCE, then s is an α'-PCE for $\alpha' \leq \alpha$, and s is a PCE iff s is a 0-PCE. Note that an α-PCE imposes some "fairness" requirements. Each player must get at least α more (where α can be negative) than her best possible outcome if the other players best respond.

We again get analogues of Theorems 1 and 3, and Corollary 4, with similar proofs.

Theorem 9. *If s is an α-PCE and s^* is a NE in a game G, then for all $i \in N$, we have $U_i(s) \geq \alpha + U_i(s^*)$.*

Theorem 10. *A strategy profile that Pareto dominates an α-PCE must itself be an α-PCE.*

Corollary 11. *If there is an α-PCE in a game G, there is a Pareto-optimal α-PCE in G.*

Of course, we are interested in α-PCE with the maximum possible value of α.

Definition 6. The strategy profile s is an *maximum-PCE (M-PCE)* in a game G if s is an α-PCE and for all $\alpha' > \alpha$, there is no α'-PCE in G.

A priori, an M-PCE may not exist in a game G. For example, it may be the case that there is an α-PCE for all $\alpha < 1$ without there being a 1-PCE. The next theorem, which uses the fact that the strategy space is compact, shows that this cannot be the case.

Theorem 12. *Every game G has a Pareto-optimal M-PCE.*

PROOF. Let $f(s) = \min_{i\in N}(U_i(s) - BU_i^G)$. Clearly f is a continuous function; moreover, if $f(s) = \alpha$, then s is an α-PCE. Since the domain consists of the set of strategy profiles, which can be viewed as a closed subset of $[0,1]^{|A|\times N}$, the domain is compact. Hence f takes on its maximum at some strategy profile s^*. Then it is immediate from the definition that s^* is an M-PCE. The argument that there is a Pareto-optimal M-PCE is essentially the same as that given in Corollary 4 showing that there is a Pareto-optimal PCE; we leave details to the reader. ∎

The following examples show that M-PCE gives some very reasonable outcomes.

Example 5. *The Nash bargaining game, continued:* Clearly $BU_1 = BU_2 = 100$; $(50,50)$ is a (-50)-PCE and is the unique M-PCE.

Example 6. *A coordination game, continued:* If $k_1 > 1$ and $k_2 > 1$, then (a,a) is the unique M-PCE; if $k_1 < 1$ and $k_2 < 1$, then (b,b) is the unique M-PCE. In both cases, $\alpha = 0$. If $k_1 > 1$ and $k_2 < 1$, then the M-PCE depends on the exact values of k_1 and k_2. If $k_1 - 1 > 1 - k_2$, then (a,a) is the unique M-PCE; if $k_1 - 1 = 1 - k_2$, then both (a,a) and (b,b) are M-PCE; otherwise, (b,b) is the unique M-PCE. In all three cases, $\alpha = -\min(k_1 - 1, 1 - k_2) < 0$.

Example 7. *Prisoner's Dilemma, continued:* Clearly (Cooperate, Cooperate) is a 2-PCE and (Defect, Defect) is a 0-PCE; (Cooperate, Cooperate) is the unique M-PCE.

Example 8. *The Traveler's Dilemma, continued:* $(100,100)$ is easily seen to be the unique M-PCE; since there is no strategy profile that guarantees both players greater than 100 (since for any pair of pure strategies, the total payoff to the players is at most 200, and the total payoff from a mixed strategy profile is a convex combination of the payoff of pure strategy profiles).

Example 9. *The centipede game, continued:* A straightforward computation shows that the M-PCE in this game is unique, and is the strategy profile s^* of the form $(\alpha q_{1,C} + (1-\alpha)q_{1,19}, q_{2,C})$, where α is chosen so as to maximize $\min(U_1(s^*) - BU_1, U_2(s^*) - BU_2)$. This can be done by taking $\alpha = \frac{1}{3\times 2^{18}+2} - \frac{3\times 2^{17}}{(3\times 2^{18}+2)(3\times 2^{18}+1)(3\times 2^{17}+1)}$.

4. COOPERATIVE EQUILIBRIUM

We can gain further insight into M-PCE (and into what people actually do in a game) by considering a notion that we call *cooperative equilibrium*, which generalizes PCE by allowing for the possibility of punishment. We define CE for two-player games. (As we discuss below, it is not clear how to extend the definition to n-player games for $n > 2$.)

Definition 7. A strategy profile s is a *cooperative equilibrium (CE)* in a two-player game if, for all $i \in \{1,2\}$ and all strategies $s_i' \in S_i$, one of the following conditions holds:

(a) for all $s_{3-i}' \in BR_{3-i}(s_i')$, we have $U_i(s) \geq U_i(s')$;

(b) for all $s_{3-i}' \in BR_{3-i}(s_i')$, we have $U_{3-i}(s') < U_{3-i}(s)$, and, for some $s_{3-i}' \in S_{3-i}$, we have $U_i(s) \geq U_i(s')$.

If we considered only the first condition, then the definition would be identical to PCE. The second condition is where punishment comes in. Suppose that there is no response that the other player can make to s_i' that makes the other player better off than he is with s. Then, intuitively, the other player becomes unhappy, and will seek to punish i. If there is some way to punish i that leads to i being no better off than he is with s, then i will not deviate to s_i'.

We are not sure how to generalize CE to arbitrary games. We could, of course, replace $BR_{3-i}(s_i')$ by $NE_{-i}(s_i')$ in the first clause. The question is what to do in the second clause. We could say that if each player in $N - \{i\}$ is worse off in every Nash equilibrium in the game G_{s_i}, they punish player i. But punishment may require a coordination of strategies, and it is not clear how the players achieve such coordination, at least in a one-shot game. Not surprisingly, the examples in the literature where players punish others are two-player games like the Ultimatum game. In general, the intuition of punishment seems most compelling in two-player games.

Our main interest in CE is motivated by the following result, which shows that every Pareto-optimal M-PCE is a CE.

Theorem 13. *Every Pareto-optimal M-PCE is a CE.*

PROOF. Suppose that s is a Pareto-optimal M-PCE. To see that s is a CE, consider the maximum α such that s is an α-PCE. If $\alpha \geq 0$, then s is a PCE, and hence clearly a CE. So suppose that $\alpha < 0$ and, by way of contradiction, that s is not a CE. Then one of the players, say 1, has a deviation to a strategy s'_1 such that either (1) player 2 has a best response s'_2 to s'_1 such that $U_1(s') > U_1(s)$ and $U_2(s') \geq U_2(s)$ or (2) for all $s'_2 \in S_2$, it must be the case that $U_2(s') < U_2(s)$ and $U_1(s) < U_1(s')$; that is, player 2 does worse than $U_2(s)$ no matter what he does, and cannot punish player 1. In case (1), it is immediate that s is not a Pareto-optimal M-PCE. So suppose that case (2) applies.

By definition, $U_i(s) \geq \alpha + BU_i$, for $i = 1, 2$. By compactness, there must be a strategy profile s^* such that $s^*_1 \in BR_1(s^*_2)$ and $U_2(s^*_2) = BU_2$. We claim that s^* is an α'-PCE for some $\alpha' > \alpha$. Since $s^*_1 \in BR_1(s^*_2)$, we must have $U_1(s^*) \geq U_1(s'_1, s^*_2) > U_1(s)$. Since $U_1(s) \geq \alpha + BU_1$, there must be some $\alpha' > \alpha$ such that $U_2(s) \geq \alpha' + BU_1$. By definition, $U_2(s^*) = BU_2$. Thus, s^* is an α'-PCE. This completes the proof. ∎

We can also prove the following analogues of Theorem 3 and Corollary 4. Since the proofs are quite similar to proofs of Theorem 3 and Corollary 4, we omit them here.

Theorem 14. *A strategy profile that Pareto dominates a CE must itself be a CE.*

Corollary 15. *There is a Paret-optimal CE in every game.*

We now consider how CE works in the examples considered earlier.

Example 10. *The Nash bargaining game:* Recall that the Nash bargaining game does not have a PCE, and that every profile of the form $(a, 100 - a)$ is a NE. We now show that these profiles are all CE as well. To see this, first observe that $U_1(s) + U_2(s) \leq 100$ for any strategy profile s. (This is clearly true for pure strategy profiles, and the total expected utility for a mixed strategy is just a convex combination of expected utilities from pure strategies.) Now suppose that player 1 deviates from $(a, 100 - a)$ to some strategy s_1, and that player 2's expected utility from a best response s'_2 to s_1 is b. If $b \geq 100 - a$, then $U_1(a, s'_2) \leq a$, and the first condition of CE applies. If $b < 100 - a$, then player 2 will punish player 1 by playing 100 (i.e., the second condition of CE applies). The same considerations apply to a deviation by player 2 from $100 - a$. Thus, $(a, 100 - a)$ is a CE. Of course, only one of these CE is an M-PCE: $(50, 50)$.

There are also Nash equilibria in mixed strategies; for example, $(\frac{1}{3}25 + \frac{2}{3}75, \frac{1}{3}25 + \frac{2}{3}75)$ is a NE. However, it is not hard to show that there are no nontrivial CE in mixed strategies. For suppose that s is a CE where either s_1 or s_2 are nontrivial mixed strategies. It is easy to see that $U_1(s) + U_2(s) < 100$. That means that there is pair $(a, 100 - a)$ such that $a > U_1(s)$ and $100 - a > U_2(s)$. So if player 1 deviates to a and player 2 deviates to $100 - a$, neither of the two conditions that characterize CE hold.

Example 11. *A coordination game, continued:* If $k_1 > 1$ and $k_2 > 1$, then (a, a) is the only CE; if $k_1 < 1$ and $k_2 < 1$, then (b, b) is the only CE; if $k_1 > 1$ and $k_2 < 1$, then the two NE, (a, a) and (b, b), are both CE (although neither is a PCE). There is one other NE s in mixed strategies; s is not a CE. To see this, note that in s both players have to put positive probability on each pure strategy. It easily follows that $U_2(s) = U_2(s_1, b) < 1$ (since s_1 puts positive probability on a); similarly, $U_1(s) < 1$. Hence, if player 1 plays b instead of s_1, player 2 has a unique best response of b, which strictly increases both players' payoffs. Thus, s is not a CE.

Example 12. *Prisoner's Dilemma, continued:* Clearly all the PCE in Prisoner's Dilemma are CE. We now prove that there are no other CE. Suppose, by way of contradiction, that s is a CE that is not a PCE. Then some player must get a payoff with s that is strictly less than 1. Without loss of generality, we can suppose that this player is player 1. Suppose that $U_1(s) = r_1 < 1$. But then if player 1 plays Defect, he is guaranteed a better payoff—at least 1—no matter what player 2 does, so s cannot be a CE.

Example 13. *The Traveler's Dilemma, continued:* Of course, all the PCE in Traveler's Dilemma are also CE. There are other CE as well. For example, $(100, 99)$ is a CE but not a PCE. To see this, note that with $(100, 99)$, player 1 gets a payoff of 97 and player 2 gets 101, the maximum possible payoff. Suppose that there exists some strategy s_1 that gives player 1 a payoff strictly greater than 97 when player 2 best responds. This strictly decreases player 2's payoff. However, player 2 can punish player 1 by playing 2, so that player 1 gets at most 2, strictly less than what he gets originally. It easily follows that (100, 99) is a CE. A similar argument shows that all other Pareto-optimal strategy profiles are also CE.

Recall that (100, 100) is the unique M-PCE of this game. Intuitively, M-PCE has fairness requirements that CE does not have.

Example 14. *The centipede game, continued:* Again, all the PCE are CE. In addition, all Pareto-optimal strategies are CE. Thus, for example, the strategy profile where both players continue to the end of the game is a CE (although it is not a PCE), as is the profile where player 2 continues at all his moves, but player 1 ends the game at his last turn. To see that Pareto-optimal strategies are CE, let s be a Pareto-optimal strategy profile and, by way of contradiction, suppose that s is not a CE. Then there must be a strategy s'_i for some player i such that either (1) there is a best response s'_{3-i} to s'_i such that $U_i(s) > U_i(s')$ and $U_{3-i}(s') \geq U_{3-i}(s)$ or (2) for all $s'_{3-i} \in S_{3-i}$, it must be the case that $U_{3-i}(s') < U_{3-i}(s)$ and $U_i(s) < U_i(s')$; that is, player $3 - i$ does worse than $U_{3-i}(s)$ no matter what he does, and cannot punish player i. In case (1), it is immediate that s is not Pareto optimal; and case (2) cannot hold, since player $3 - i$ can always punish player i by exiting at his first turn.

5. THE COMPLEXITY OF COMPUTING A PCE AND M-PCE

In this section, we show that in two-player games, both a PCE and an M-PCE can be found in polynomial time. The first step in the argument involves showing that in two-player games, for all strategy profile s, there is a strategy profile $s' = (s'_1, s'_2)$ that Pareto dominates s such that both

s'_1 and s'_2 have support at most two pure strategies (i.e., they give positive probability to at most two pure strategies). We then show that both the problem of computing a PCE and an M-PCE can be reduced to solving a polynomial number of "small" bilinear programs, each of which can be solved in constant time. This gives us the desired polynomial time algorithm.

Notation: For a matrix $\mathbf{A}$, let $\mathbf{A}^T$ denote $\mathbf{A}$ transpose, let $\mathbf{A}[i,\cdot]$ denote the ith row of $\mathbf{A}$, let $\mathbf{A}[\cdot,j]$ denote the jth column of $\mathbf{A}$, and let $\mathbf{A}[i,j]$ be the entry in the ith row, jth column of $\mathbf{A}$. We say that a vector x is *nonnegative*, denoted $x \geq 0$, if its all of its entries are nonnegative.

We start by proving the first claim above. In this discussion, it is convenient to identify a strategy for player 1 with a column vector in $\mathbb{R}^n$, and a strategy for player 2 with a column vector in $\mathbb{R}^m$. The strategy has a support of size at most two if the vector has at most two nonzero entries.

Lemma 16. *In a two-player game, for all strategy profiles s^*, there exists a strategy profile $s' = (s'_1, s'_2)$ that Pareto dominates s^* such that both s'_1 and s'_2 have support of size at most two.*

PROOF. Let $\mathbf{A}$ and $\mathbf{B}$ be the payoff matrices (of size $n \times m$) for player 1 and player 2 respectively. Given a strategy profile $s^* = (s^*_1, s^*_2)$, let $U_1(s^*) = r^*_1$ and $U_2(s^*) = r^*_2$. We first show that there exists a strategy s'_2 for player 2 with support of size at most two such that (s^*_1, s'_2) Pareto dominates s^*. We then show that there exists a strategy s'_1 for player 1 with support of size at most two such that (s'_1, s'_2) Pareto dominates (s^*_1, s'_2), and hence s^*.

Consider the following linear program P_1, where y is a column vector in $\mathbb{R}^m$:

$$\begin{array}{ll} \text{maximize} & (s^*_1)\mathbf{A}y \\ \text{subject to} & (s^*_1)^T\mathbf{B}y = r^*_2 \\ & \sum_{i=1}^m y[i] = 1 \\ & y \geq 0. \end{array}$$

As usual, an *optimal solution* of P_1 is a vector y that maximizes the objective function ($(s^*_1)\mathbf{A}y$) and satisfies the three constraints; a *feasible solution* of P_1 is one that satisfies the constraints; finally, an *optimal value* of P_1 is the value of the objective function for the optimal solution y (if it exists). We show that P_1 has an optimal solution y^* with at most two nonzero entries.

Since all constraints in P_1 are equality constraints except for the nonnegativity constraint, P_1 is a *standard-form linear program* [?]. We can rewrite the equality constraints in P_1 as

$$\mathbf{D}y = \begin{bmatrix} r^*_2 \\ 1 \end{bmatrix},$$

where $\mathbf{D}$ is an $(m \times 2)$ matrix whose first row is $(s^*_1)^T\mathbf{B}$ and whose second row has all entries equal to 1. In geometric terms, the region represented by the constraints in P_1 is a convex polytope. Since P_1 is a standard-form linear program, it is well-known that y is a vertex of the polytope (i.e., an extreme point of the polytope) iff all columns i in $\mathbf{D}$ where $y[i] \neq 0$ are linearly independent [?]. Since the columns of $\mathbf{D}$ are vectors in $\mathbb{R}^2$, at most two of them can be linearly independent. Thus, a vertex y of the polytope can have at most two nonzero entries.

Clearly s^*_2 is a feasible solution of P_1. Since $(s^*_1)\mathbf{A}s^*_2 = r^*_1$, by assumption, the optimal value of P_1 is at least r^*_1. Moreover, since the objective function of P_1 is linear, $y \geq 0$, and $\sum_{i=1}^m y[i] = 1$, the optimal value is bounded. Therefore, the linear program has an optimal solution. By *the fundamental theorem of linear programming*, if a linear program has an optimal solution, then it has an optimal solution at a vertex of the polytope defined by its constraints [?]. Let s'_2 be the strategy defined by an optimal solution at the vertex of the polytope. As we observed above, s'_2 has at most two nonzero entries. It is immediate that $U_1((s^*_1, s'_2)) \geq r^*_1$ and $U_2((s^*_1, s'_2)) \geq r^*_2$.

This completes the first step of the proof.

The second step of the proof essentially repeats the first step. Suppose that $U_1((s^*_1, s'_2)) = r_1$ and $U_2((s^*_1, s'_2)) = r_2$. Consider the following linear program P_2, where x is column vector in $\mathbb{R}^n$:

$$\begin{array}{ll} \text{maximize} & x^T\mathbf{B}s'_2 \\ \text{subject to} & x^T\mathbf{A}s'_2 = r_1 \\ & \sum_{i=1}^n x[i] = 1 \\ & x \geq 0. \end{array}$$

Since s^*_1 is a feasible solution of P_2 and $(s^*_1)^T\mathbf{B}s'_2 \geq r^*_2$, the optimal value of P_2 is at least r^*_2. As above, if we take s'_2 to be an optimal solution of P_2 that is a vertex of the polytope defined by the constraints, then s'_2 has support of size at most two, and (s'_1, s'_2) Pareto dominates s^*. ∎

The rest of the section makes use of *bilinear* programs. There are a number of slightly different variants of bilinear programs. For our purposes, we use the following definition.

Definition 8. A *bilinear program* P (of size $n \times m$) is a quadratic program of the form

$$\begin{array}{ll} \text{maximize} & x^T\mathbf{A}y + x^Tc + y^Tc' \\ \text{subject to} & x^T\mathbf{B}_1 y \geq d_1 \\ & \mathbf{B}_2 x = d_2 \\ & \mathbf{B}_3 y = d_3 \\ & x \geq 0 \\ & y \geq 0, \end{array}$$

where $\mathbf{A}$ and $\mathbf{B}_1$ are $n \times m$ matrices, $x, c \in \mathbb{R}^n$, $y, c' \in \mathbb{R}^m$, $\mathbf{B}_2$ is a $k \times n$ matrix for some k, and $\mathbf{B}_3$ is a $k' \times m$ matrix for some k'. P is *simple* if $\mathbf{B}_2$ and $\mathbf{B}_3$ each has one row, which consists of all 1's. (Thus, in a simple bilinear program, we have a bilinear constraint $x^T\mathbf{B}_1 y \geq d_1$, non-negativity constraints on x and y, and constraints on the sum of the components of the vectors x and y; that is, constraints of the form $\sum_{i=1}^n x[i] = d'$ and $\sum_{j=1}^m y[j] = d''$.) ∎

Lemma 17. *A simple bilinear program of size 2×2 can be solved in constant time.*

PROOF. See Appendix. ∎

We can now give our algorithm for finding a PCE. The idea is to first find BU_1 and BU_2, which can be done in polynomial time. We then use Lemma 16 to reduce the problem to $\binom{n}{2}\binom{m}{2} = O(n^2m^2)$ smaller prolems, each of a which is a simple bilinear program of size 2×2. By Lemma 17, each of these smaller problems can be solved in constant time, giving us a polynomial-time algorithm.

Theorem 18. *Given a two-player $G = (\{1,2\}, A, u)$, we can compute in polynomial time whether G has a PCE and, if so, we can compute a PCE in polynomial time.*

PROOF. Suppose that $G = (\{1,2\}, A, u)$, where $A = A_1 \times A_2$, $|A_1| = n$, $|A_2| = m$, u_1 is characterized by the payoff matrix $\mathbf{A}$, and u_2 is characterized by the payoff matrix $\mathbf{B}$.

In order to compute a PCE for the game, we need the values of BU_1 and BU_2. These can be computed in polynomial time, as follows. For BU_1, for each $i \in \{1, \dots, m\}$, we solve the following linear program P_i:

$$\begin{array}{ll} \text{maximize} & s_1^T(\mathbf{A}[\cdot, i]) \\ \text{subject to} & s_1^T(\mathbf{B}[\cdot, i]) \geq s_1^T(\mathbf{B}[\cdot, j]) \ \text{ for all } j \in \{1, \dots m\} \\ & \sum_{l=1}^n s_1[l] = 1 \\ & s_1 \geq 0. \end{array}$$

Suppose that r_i is the optimal value of P_i. Since P_i is a linear program, r_i can be computed in polynomial time. Intuitively, r_i is the maximum reward player 1 gets when action b_i is a best response for player 2. (The first constraint ensures that, given s_1, action b_i is a best response for player 2.) $BU_1 = \max_{i=1}^m r_i$, so can be computed in polynomial time. BU_2 can be similarly computed.

After computing BU_1 and BU_2, we can compute a PCE. Recall that a strategy profile s is a PCE iff $U_1(s) \geq BU_1$ and $U_2(s) \geq BU_2$. Suppose that game G has a PCE s^*. By Lemma 16, there must exist a strategy profile $s' = (s_1', s_2')$ that Pareto dominates s^*, where both s_1' and s_2' have support of size at most two. By Theorem 7, s' is also a PCE. We call such a PCE a (2×2)-*PCE*. Our arguments above show that G has a PCE iff it has a (2×2)-PCE. Thus, in order to check whether G has a PCE, it suffice to check whether it has a (2×2)-PCE.

We do this exhaustively. For all $i_1, i_2 \in \{1, 2, \dots, n\}$ with $i_1 \neq i_2$ and all $j_1, j_2 \in \{1, 2, \dots, m\}$ with $j_1 \neq j_2$, we check whether G has a (2×2)-PCE in which player 1 places positive probability only on strategies i_1 and i_2, and player 2 places positive probability only on strategies j_1 and j_2. For each choice of i_1, i_2, j_1, j_2, this question can be expressed as the following 2×2 simple bilinear programming problem P_{i_1,i_2,j_1,j_2}, where $\mathbf{A}_{i_1,i_2,j_1,j_2}$ is the 2×2 matrix $\begin{bmatrix} \mathbf{A}[i_1, j_1] & \mathbf{A}[i_1, j_2] \\ \mathbf{A}[i_2, j_1] & \mathbf{A}[i_2, j_2] \end{bmatrix}$, and $\mathbf{B}_{i_1,i_2,j_1,j_2}$ is the 2×2 matrix $\begin{bmatrix} \mathbf{B}[i_1, j_1] & \mathbf{B}[i_1, j_2] \\ \mathbf{B}[i_2, j_1] & \mathbf{B}[i_2, j_2] \end{bmatrix}$:

$$\begin{array}{ll} \text{maximize} & [x_1 \ x_2] \ \mathbf{A}_{i_1,i_2,j_1,j_2} \ [y_1 \ y_2]^T \\ \text{subject to} & [x_1 \ x_2] \ \mathbf{B}_{i_1,i_2,j_1,j_2} \ [y_1 \ y_2]^T \geq BU_2 \\ & x_1 + x_2 = 1 \\ & y_1 + y_2 = 1 \\ & x \geq 0, \ y \geq 0. \end{array}$$

The first constraint ensures that player 2's reward is at least BU_2, the remaining constraints ensure that player 1 puts positive probability only on strategies i_1 and i_2, while player 2 puts positive probability only on j_1 and j_2. If the optimal value of P_{i_1,i_2,j_1,j_2} for some choice of of (i_1, i_2, j_1, j_2) is at least BU_1, then the corresponding optimal solution (x, y) is a PCE of G. (Recall that a strategy profile s is a PCE if $U_1(s) \geq BU_1$, and $U_2(s) \geq BU_2$.) On the other hand, if the optimal value for each P_{i_1,i_2,j_1,j_2} is strictly less than BU_1, then G does not have a (2×2)-PCE and so, by the arguments above, G does not have a PCE.

The algorithm above must solve $\binom{n}{2} \times \binom{m}{2}$ simple 2 bilinear programs. By Lemma 17, each can be solved in constant time. Thus, the algorithm runs in polynomial time, as desired. ∎

The argument that an M-PCE can be found in polynomial time is very similar.

Theorem 19. *Given a two-player $G = (\{1,2\}, A, u)$, we can compute an M-PCE in polynomial time.*

PROOF. We start by compute BU_1 and BU_2, as in Theorem 18. Again, this takes polynomial time.

Recall that an M-PCE is an α-PCE such that for all $\alpha' > \alpha$, there is no α'-PCE in G. Clearly, a strategy that Pareto dominates an α-PCE must itself be an α-PCE. Thus, using Lemma 16, it easily follows that there must be an M-PCE for G such that the support of both strategies involved is of size at most 2. Call such an M-PCE a (2×2)-M-PCE.

Thus, to compute an M-PCE, for each tuple (i_1, i_2, j_1, j_2), we compute the optimal α for which we can get an α-PCE when player 1 is restricted to putting positive probability on actions i_1 and i_2, while player 2 is restricted to putting positive probability in j_1 and j_2. Using the notation of Theorem 18, we want to solve the following problem Q_{i_1,i_2,j_1,j_2}, where $d_1(x_1, x_2, y_1, y_2) = [x_1 \ x_2] \ \mathbf{A}_{i_1,i_2,j_1,j_2} \ [y_1 \ y_2]^T - BU_1$ and $d_2(x_1, x_2, y_1, y_2) = [x_1 \ x_2] \ \mathbf{B}_{i_1,i_2,j_1,j_2} \ [y_1 \ y_2]^T - BU_2$:

$$\begin{array}{ll} \text{maximize} & \min(d_1(x_1, x_2, y_1, y_2), d_2(x_1, x_2, y_1, y_2)) \\ \text{subject to} & x_1 + x_2 = 1 \\ & y_1 + y_2 = 1 \\ & x \geq 0, \ y \geq 0. \end{array}$$

The objective function maximizes the α for which the strategy profile determined by $[x_{i_1}, x_{i_2}]$ and $[y_{i_1}, y_{i_2}]$ is an α-PCE (recall that s is an α-PCE if $\alpha = \min(U_1(s) - BU_1, U_2(s) - BU_2)$). The problem here is that since the objective function involves a min, this is not a bilinear program. However, we can solve this problem by solving two simple bilinear programs of size 2×2, depending on which of $[x_{i_1} x_{i_2}]\mathbf{A}_{i_1,i_2,j_1,j_2}$ $[y_{i_1} y_{i_2}]^T - BU_1$ and $[x_{i_1} x_{i_2}]\mathbf{A}_{i_1,i_2,j_1,j_2}[y_{i_1} y_{i_2}]^T - BU_2$ is smaller.

Let Q'_{i_1,i_2,j_1,j_2} be the following simple bilinear program,

$$\begin{array}{ll} \text{maximize} & d_1(x_1, x_2, y_1, y_2) \\ \text{subject to} & d_1(x_1, x_2, y_1, y_2) \leq d_2(x_1, x_2, y_1, y_2) \\ & x_1 + x_2 = 1 \\ & y_1 + y_2 = 1 \\ & x \geq 0, \ y \geq 0. \end{array}$$

Let Q''_{i_1,i_2,j_1,j_2} be the same bilinear program with the roles of d_1 and d_2 reversed. It is easy to see that the larger of the solutions to Q'_{i_1,i_2,j_1,j_2} and Q''_{i_1,i_2,j_1,j_2} is the solution to Q_{i_1,i_2,j_1,j_2}. It thus follows that an M-PCE can be computed in polynomial time. ∎

6. M-PCE AND COCO VALUE

Kalai and Kalai [**?**] have introduced a solution concept for two player games that they call the *cooperative-competitive (coco) value.* As the name suggests, it also attempts to capture some of the cooperative behavior in games. As we show by example, the coco value is an M-PCE value (i.e., the payoffs that the players get in an M-PCE) in many games of interest. This motivates us to look more carefully at the relationship between the coco value and the M-PCE value. We start with a review of the coco value. The coco value is defined only for two-player games where side payments are possible. Intuitively, it is best to think of the outcome of the game being expressed in dollars, assume that money can be transferred between the two players, and that each player values money the same way (so if player 1 transfers \$5 to player 2, then player 1's utility decreases by 5, while player 2's increases by 5).[1] The coco value is viewed as a fair and efficient outcome in such games.

In the rest of this section, we make use of the following definitions. In a two-player game G, we say that (r_1, r_2) is an *M-PCE value of G* if there is an M-PCE s such that $U_1(s) = r_1$ and $U_2(s) = r_2$. Let $MSW(G)$ be the maximum social welfare of G; formally, $MSW(G) = \max_{s \in S}(U_1(s) + U_2(s))$. Finally, the *minimax value of game G for player i*, denoted $mm_i(G)$, is the best payoff that i can guarantee himself; formally,

$$mm_i(G) = \min_{s_{3-i} \in S_{3-i}} \max_{s_i \in S_i} U_1(s_1, s_2).$$

The coco value is computed by decomposing a game into two components, which can be viewed as a purely cooperative component and a purely competitive component. The cooperative component is a *team game*, a game where both players have identical utility matrices, so that both players get identical payoffs, no matter what strategy profile is played. The competitive component is a *zero-sum* game, that is, one where if player 1's payoff matrix is A, then player 2's payoff matrix is $-A$.

As Kalai and Kalai [?] observe, every game G can be uniquely decomposed into a team game G_t and a zero-sum game G_z, where if $(\mathbf{A}, \mathbf{B})$, $(\mathbf{C}, \mathbf{C})$, and $(\mathbf{D}, -\mathbf{D})$ are the utility matrices for G, G_t, and G_z, respectively, then $\mathbf{A} = \mathbf{C} + \mathbf{D}$ and $\mathbf{B} = \mathbf{C} - \mathbf{D}$. Indeed, we can take $\mathbf{C} = (\mathbf{A} + \mathbf{B})/2$ and $\mathbf{D} = (\mathbf{A} - \mathbf{B})/2$. We call G_t *the team game of G* and call G_z *the zero-sum game of G*.

Using this decomposition, we can define the coco value. Given a game G, let c be the largest value obtainable in the team game G_t (i.e., the largest value in the utility matrix for G_t), and let z be the minimax value for player 1 in the zero-sum game G_z. Then the *coco value of G*, denoted *coco*(G), is $(a+z, a-z)$. Note that the coco value is attainable if utilities are transferable: the players simply play the strategy profile that gives the value c in G_t; then player 2 transfers z to player 1 (z may be negative, so that 1 is actually transferring money to 2). Clearly this outcome maximizes social welfare. Kalai and Kalai [?] argue that it is also fair in an appropriate sense.

The coco value and M-PCE value are closely related in a number of games of interest, as the following examples show.

Example 15. *The Nash bargaining game, continued:* Clearly, the largest payoff obtainable in the team game corresponding to the Nash Bargaining game is $(50, 50)$. Since the game is symmetric, the minimax value of each player in the zero-sum game is 0. in the zero-sum game, so $mm_1(G_z) = mm_2(G_z) = 0$. Thus, the coco value of the Nash bargaining game is $(50, 50)$, which is also the unique M-PCE value.

Example 16. *Prisoner's Dilemma, continued:* Clearly, the largest payoff obtainable in the team game corresponding to Prisoner's Dilemma (given the payoffs shown in the Introduction) is $(3, 3)$. Since the game is symmetric, again, the minimax value in the corresponding zero-sum game is 0. Thus, the coco value is $(3, 3)$, which is also the unique M-PCE value.

Example 17. *The Traveler's Dilemma, continued:* Clearly, the largest payoff obtainable in the team game corresponding to the Traveler's Dilemma is $(100, 100)$. And again, since the game is symmetric, the minimax value for each player in the zero-sum game is 0. Thus, the coco value is $(100, 100)$, which is also the unique M-PCE value.

On the other hand, in some games, the coco value and M-PCE value differ.

Example 18. *The centipede game, continued:* It is easy to see that the largest payoff obtainable in the team game corresponding to the centipede game is $(\frac{2^{19}+2^{20}+1}{2}, \frac{2^{19}+2^{20}+1}{2})$: both players play to the end of the game and split the total payoff. It is also easy to compute that, in the zero-sum game corresponding to the centipede game, player 1's minimax value is 1, while player 2's minimax value is -1, obtained when both players quit immediately. Thus, the coco value is $(\frac{2^{19}+2^{20}+1}{2} + 1, \frac{2^{19}+2^{20}+1}{2} - 1)$. This value is not achievable without side payments, and is higher than the M-PCE value.

Although, as the centipede game shows, the coco value and the M-PCE value may differ, it is worth noting that the coco value of a game is the sum of the M-PCE values of its decomposed games. Clearly c is the unique M-PCE value of G_t, since it is the unique Pareto-optimal payoff; moreover, the unique M-PCE value of a zero-sum game can easily be shown to be the payoffs in NE, which are given by the minimax values. But we can say more. Part of the problem in the centipede game is that the computation of the coco value assumes that side payments are possible. The M-PCE value does not take into account the possibility of side payments. Indeed, once we extend the centipede game to allow side payments in an appropriate sense, it turns out that the coco value and the M-PCE value are the same. To do a fairer comparison of the M-PCE and coco values, we consider games with side payments, which we define next.

6.1 Two-player games with side payments

As we have observed, the coco value makes sense only if players can make side payments. The ability to make side payments is not explicitly modeled in the description of the games considered by Kalai and Kalai [?]. Since the M-PCE value calculation does not assume side payments are possible, we do need to explicitly model this possibility if we want to do a reasonable comparison of the M-PCE value and coco value.

In this subsection, we decribe how an arbitrary two-player game without payments can be transformed into a game with side payments. There is more than one way of doing this—we focus on one, and briefly discuss a second alternative. Our procedure may be of interest beyond the specific application to coco and M-PCE. We implicitly assume throughout that outcomes can be expressed in dollars and that players value the dollars the same way. The idea is to add strategies to the game that allow players to propose "deals", which amount to a description of what strategy profiles should be played and how much money should be trasferred. If the players propose the same deal, then the suggested strategy profile is played, and the money is transferred. Otherwise, a "backup" strategy is played.

Given a two-player game $G = (\{1, 2\}, A, U)$, let $G^* = (\{1, 2\}, A^*, U^*)$ be the *game with side payments extending*

[1] Without the assumption that players value money the same way, the intuition behind the coco value breaks down.

G, where S^* and U^* are defined as follows. S^* extends S by adding a collection of strategies that we call *deal strategies*. A deal strategy for player i is a triple of the form $(s, r, s'_i) \in S \times I\!R \times S_i$. Intuitively, this strategy proposes that the players play the strategy profile s and that player 1 should transfer r to player 2; if the deal is not accepted, then player i plays s'_i. Given this intuition, it should be clear how U^* extends U. For strategy profiles $s \in S$, $U^*(s) = U(s)$. The players agree on a deal if they both propose a deal strategy with the same first two components (s, r). In this case they play s and r is transferred. Otherwise, players just play the backup strategy. That is, for $s, s' \in S$, $t_i \in S_i$, and $r, r' \in I\!R$

- $U^*(s) = U(s)$;
- $U_1^*((s, r, t_1), (s, r, t_2)) = U_1(s) - r$;
  $U_2^*((s, r, t_1), (s, r, t_2)) = U_2(s) + r$;
- $U^*((s, r, t_1), (s', r', t_2)) = U(t_1, t_2)$ if $(s, r) \neq (s', r')$;
- $U^*((s, r, t_1), t_2) = U^*(t_1, (s', r', t_2)) = U(t_1, t_2)$.

We call G^* the game with side payments *extending* G, and call G the game *underlying* G^*.

Intutively, we can think of both players as giving their strategies to a trusted third party. If they both propose the same deal strategy, the third party ensures that it is carried out and the transfer is made. Otherwise, the appropriate backup strategies are played.

In our approach, we have allowed players to propose arbitrary backup strategies in case their deal offers are not accepted. We also considered an alternative approach, where if a deal is proposed by one of the parties but not accepted, then the players get a fixed default payoff (e.g., they could both get 0, or a default strategy could be played, and the players get their payoff according to the default strategy). Essentially the same results as those we prove hold for this approach as well; see the end of Section 6.2.

6.2 Characterizing the coco value and the M-PCE value

At first glance, the coco value and the M-PCE value seem quite different, although both are trying to get at the notion of fairness. However, we show below that both have quite similar characterizations. We show this in two ways. In this section, we characterize the two notions using two similar formulas involving the maximum social welfare and the minimax value. In the next section, we compare axiomatic characterizations of the notions.

Before proving our results, we first show that, although they are different games, G and G^* agree on the relevant parameters (recall that G^* is the game with side payments extending G).

Lemma 20. *For all two-player games G, $MSW(G) = MSW(G^*)$ and $mm_i(G^*) = mm_i(G)$ for $i = 1, 2$.*

PROOF. To see that $MSW(G) = MSW(G^*)$, observe To see that $MSW(G) = MSW(G^*)$, observe that, by the definition of U^*, for all strategies $s^* \in S^*$, there exists a strategy $s \in S$ and $r \in I\!R$ such that $U^*(s^*) = (U_1(s) + r, U_2(s) - r)$, so $U_1^*(s^*) + U_2^*(s^*) = U_1(s) + U_2(s)$.

To see that $mm_1(G^*) = mm_1(G)$, observe that for all $t \in S_2$, we have that $U_1^*((s, r, s'_1), t) = U_1(s'_1, t)$, so

$$\max_{s'_1 \in S_1^*} U_1^*(s'_1, t) = \max_{s'_1 \in S} U_1(s'_1, t).$$

Thus,

$$\min_{t \in S_2} \max_{s'_1 \in S_1} U_1(s'_1, t) = \min_{t \in S_2} \max_{s'_1 \in S_1^*} U_1^*(s'_1, t).$$

Therefore,

$$\begin{array}{ll} mm_1(G^*) & = \min_{t \in S_2^*} \max_{s'_1 \in S_1^*} U_1^*(s'_1, t) \\ & \leq \min_{t \in S_2} \max_{s'_1 \in S_1^*} U_1^*(s'_1, t) \quad \text{[since } S_2^* \supset S_2\text{]} \\ & = \min_{t \in S_2} \max_{s'_1 \in S_1} U_1(s'_1, t) \\ & = mm_1(G). \end{array}$$

Thus, $mm_1(G^*) \leq mm_1(G)$. Similarly, for $s'_1 \in S_1$, we have $\min_{t \in S_2^*} U_1^*(s'_1, t) = \min_{t \in S_2} U_1(s'_1, t)$. Thus,

$$\min_{t \in S_2^*} \max_{s'_1 \in S_1} U_1^*(s'_1, t) = \min_{t \in S_2} \max_{s'_1 \in S_1} U_1(s'_1, t).$$

It follows that

$$\begin{array}{lll} mm_1(G^*) & = & \min_{t \in S_2^*} \max_{s'_1 \in S_1^*} U_1^*(s'_1, t) \\ & \geq & \min_{t \in S_2^*} \max_{s'_1 \in S_1} U_1^*(s'_1, t) \\ & = & \min_{t \in S_2} \max_{s'_1 \in S_1} U_1(s'_1, t) \\ & = & mm_1(G). \end{array}$$

Thus, $mm_1(G^*) = mm_1(G)$. A similar argument shows that $mm_2(G^*) = mm_2(G)$. ∎

We now characterize the coco value.

Theorem 21. *If G is a two-player game, then $coco(G) = (\frac{MSW(G)+mm_1(G_z)-mm_2(G_z)}{2}, \frac{MSW(G)-mm_1(G_z)+mm_2(G_z)}{2})$.[2] Moreover, $coco(G) = coco(G^*)$.*

PROOF. It is easy to see that the Pareto-optimal payoff profile in G_t is $(\frac{MSW(G)}{2}, \frac{MSW(G)}{2})$. Thus, by definition,

$$\begin{array}{ll} & coco(G) \\ = & (\frac{MSW(G)}{2}, \frac{MSW(G)}{2}) + (mm_1(G_z), mm_2(G_z)) \\ = & (\frac{MSW(G)+2mm_1(G_z)}{2}, \frac{MSW(G)+2mm_2(G_z)}{2}) \\ = & (\frac{MSW(G)+mm_1(G_z)-mm_2(G_z)}{2}, \frac{MSW(G)-mm_1(G_z)+mm_2(G_z)}{2}). \end{array}$$

The last equation follows since G_z is a zero-sum game, so $mm_1(G_z) = -mm_2(G_z)$.

The fact that $coco(G) = coco(G^*)$ follows from the characterization of $coco(G)$ above, the fact that $MSW(G) = MSW(G^*)$ (Lemma 20), and the fact that $(G_z)^* = (G^*)_z$, which we leave to the reader to check. ∎

The next theorem provides an analogous characterization of the M-PCE value in two-player games with side payments. It shows that in such games the M-PCE value is unique and has the same form as the coco value. Indeed, the only difference is that we replace $mm_i(G_z)$ by $mm_i(G)$.

Theorem 22. *If G is a two-player game, then the unique M-PCE value of the game G^* with side payments extending G is $(\frac{MSW(G)+mm_1(G)-mm_2(G)}{2}, \frac{MSW(G)-mm_1(G)+mm_2(G)}{2})$.*

PROOF. We first show that $BU_1^{G^*} = MSW(G) - mm_2(G)$ and $BU_2^{G^*} = MSW(G) - mm_1(G)$. For $BU_1^{G^*}$, let s^* be a strategy profile in G that maximizes social welfare, that is, $U_1(s^*) + U_2(s^*) = MSW(G)$, and let (s'_1, s'_2) be a strateg profile in G such that $s'_2 \in BR^G(s'_1)$ and $U_2(s'_1, s'_2) = mm_2(G)$. (Thus, by playing s'_1, player 1 ensures that player 2 can get

[2] Of course, $mm_1(G_z) = -mm_2(G_z)$, since G_z is a zero-sum game; however, we write the expression in this form to make the comparison to the M-PCE value easier.

no more utility than $mm_2(G)$, and by playing s_2', player 2 ensures that she does get utility $mm_2(G)$ when player 1 plays s_1'.

Let $s = (s_1, s_2)$, where $s_1 = (s^*, mm_2(G) - U_2(s^*), s_1')$ and $s_2 = (s^*, mm_2(G) - U_2(s^*), s_2')$. By definition, if s is played, the players agree to the deal, so s^* is played in G and player 1 transfers $mm_2(G) - U_2(s^*)$ to player 2. Thus, $U_1^*(s^*) = U_1(s^*) - (mm_2(G) - U_2(s^*)) = U_1(s^*) + U_2(s^*) - mm_2(G) = MSW(G) - mm_2(G)$, and $U_2^*(s^*) = mm_2(G)$. Player 2 gets the same payoff if she plays any strategy of the form $(s^*, mm_2(G) - U_2(s^*), t)$. On the other hand, if player 2 plays a strategy t' not of this form, then $U_2^*(s_1, t') = U_2(s_1', t') \le mm_2(G)$. Thus, if player 1 plays s_1, then player 2 gets a utility of at most $mm_2(G)$ no matter what she plays, so $s_2 \in BR_2^{G^*}(s_1)$. This shows that $BU_1^{G^*} \ge MSW(G) - mm_2(G)$.

To see that $BU_1^{G^*} \le MSW(G) - mm_2(G)$, consider a strategy profile $s'' = (s_1'', s_2'') \in S^*$ with $s_2'' \in BR_2^{G^*}(s_1'')$. Since $mm_2(G^*) = mm_2(G)$, it follows that $U_2^*(s'') \ge mm_2(G)$. Since $MSW(G^*) = MSW(G)$ by Lemma 20, it follows that $U_1^*(s'') + U_2^*(s'') \le MSW(G)$. Thus, $U_1^*(s'') \le MSW(G) - mm_2(G)$, so $BU_1^{G^*} \le MSW(G) - mm_2(G)$. Thus, $BU_1^{G^*} = MSW(G) - mm_2(G)$, as desired.

The argument that $BU_2^{G^*} = MSW(G) - mm_1(G)$ is similar.

Now suppose that we have a strategy $s^+ \in S^*$ such that $U_1(s^+) \ge BU_1^{G^*} + \alpha$ and $U_2^*(s^+) \ge BU_2^{G^*} + \alpha$. Since $MSW(G^*) = MSW(G)$, it follows that $BU_1(G^*) + BU_2(G^*) + 2\alpha \le MSW(G)$. Plugging in our characterizations of $BU_1(G^*)$ and $BU_2(G^*)$, we get that $\alpha \le \frac{-MSW(G)+mm_1(G)+mm_2(G)}{2}$. Taking $\beta = \frac{-MSW(G)+mm_1(G)+mm_2(G)}{2}$, we now show that we can find a β-PCE. It follows that this must be an M-PCE.

Let s^* and s' be the strategy profiles in S as defined above; in particular, s^* maximizes social welfare. (The choice of s' is not so relevant; the role of s' below could be played by any strategy profile in S.) Let $s^+ = (s_1^+, s_2^+)$, where $s_1^+ = (s^*, U_1(s^*) - \frac{MSW(G)+mm_1(G)-mm_2(G)}{2}, s_1')$ and $s_2^+ = (s^*, U_1(s^*) - \frac{MSW(G)+mm_1(G)-mm_2(G)}{2}, s_2')$. Thus, s^+ is a variant of s above, where the transfers are modified. It is easy to check that $U_1(s^+) = \frac{MSW(G)+mm_1(G)-mm_2(G)}{2}$, and $U_2(s^+) = \frac{MSW(G)-mm_1(G)+mm_2(G)}{2}$.

It can easily be checked that $U_i(s^+) = BU_i + \beta$ for $i = 1, 2$, so s^+ is indeed a β-PCE. Moreover, s^+ must Pareto dominate any other β-PCE. Therefore, s^+ is an M-PCE, and its value is an M-PCE value, as desired. Since $U_1(s^+) + U_2(s^+) = MSW(G)$, it follows that the M-PCE value is unique. ∎

As Theorems 21 and 22 show, in a two-player game G^* with side payments, the coco value and M-PCE value are characterized by very similar equations, making use of $MSW(G^*)$ and minmax values. The only difference is that coco value uses the minimax value of the zero-sum game G_z, while the M-PCE value uses minimax value of G. It immediately follows from Theorem 21 and 22 that the coco value and the M-PCE value coincide in all games where

$$mm_1(G_z) - mm_2(G_z) = mm_1(G) - mm_2(G).$$

Such games include team games, *equal-sum games* (games with a payoff matrices (A, B) such that $A + B$ is a constant matrix, all of whose entries are identical), *symmetric games* (games where the strategy space is the same for both players, that is, $S_1 = S_2$, and $U_1(s_1, s_2) = U_2(s_2, s_1)$ for all $s_1, s_2 \in S_1$), and many others. We could also use these theorems to show that the M-PCE value and the coco value can differ, even in a game where side payments are allowed.

Example 19. Let G be the two-player game described by the payoff matrix below, and let G^* be the game with side payments extending G.

	a	b
c	(3,2)	(1,0)

Let player 1 be the row player, and player 2 be the column player. It is easy to check that $MSW(G) = 5$, $mm_1(G) = 1$, and $mm_2(G) = 2$, Thus, by Theorem 19, the M-PCE value of G^* is $(\frac{5+1-2}{2}, \frac{5-1+2}{2}) = (2, 3)$. On the other hand, it is easy to check that $coco(G) = coco(G^*) = (3, 2)$.

It seems somewhat surprising that the M-PCE here should be $(2, 3)$, since player 1 gets a higher payoff than player 2 no matter which strategy profile in G is played. Moreover, $BU_1^G = 3$ and $BU_2^G = 2$. But things change when transferred are allowed. It is easy to check that it is still the case that $BU_1^{G^*} = 3$; if player 1 plays c, then player 2's best response is to play a. But $BU_2^{G^*} = 4$; if player 2 plays $((c, a), 2, b)$, offering to play (c, a), provided that player 1 transfers an additional 2, player 1's best response is to agree (for otherwise player 2 plays b), giving player 2 a payoff of 4. The possibility that player 2 can "threaten" player 1 in this way (even though the moves are made simultaneously, so no actual threat is involved) is why $mm_2(G) \ge mm_1(G)$. ∎

We conclude this subsection by considering what happens if a default strategy profile is used instead of backup strategies when defining games with side payments. Let the default payoffs be (d_1, d_2). Then a similar argument to that above shows that the M-PCE value becomes

$$(\frac{MSW(G) + d_1 - d_2}{2}, \frac{MSW(G) - d_1 + d_2}{2}).$$

Thus, rather then using the minimax payoffs in the formula, we now use the default payoffs. Note that if the default payoffs are $(0, 0)$, then the M-PCE amounts to the players splitting the maximum social welfare. We leave the details to the reader.

6.3 Axiomatic comparison

In this section, we compare M-PCE value to coco value axiomatically. Before jumping into the axioms, we first explain the term "axiomatize". Given a function $f : A \to B$, we say a set AX of axioms *axiomatizes f in A*, if f is the unique function mapping A to B that satisifies all axioms in AX. Recall that every two-player normal form game has a unique coco value. We can thus view the coco value as a function from two-player normal form games to $\mathbb{R}^2$. Therefore, a set AX of axioms axiomatizes the coco value if the coco value is the unique function that maps from the set to $\mathbb{R}^2$ that satisfies all the axioms in AX.

Kalai and Kalai [**?**] show that the following collection of axioms axiomatizes the coco value. We describe the axioms in terms of an arbitrary function f. If $f(G) = (a_1, a_2)$, then we take $f_i(G) = a_i$, for $i = 1, 2$.

1. **Maximum social welfare.** f maximizes social welfare: $f_1(G) + f_2(G) = MSW(G)$.

2. **Shift invariance.** Shifting payoffs by constants leads to a corresponding shift in the value. That is, if $a = (a_1, a_2)$, $G = (\{1,2\}, S, U)$ and $G^a = (\{1,2\}, S, U^a)$, where $U_i^a(s) = U_i(s) + a_i$ for all $s \in S$, then $f(G^a) = (f_1(G) + a_1, f_2(G) + a_2)$.

3. **Monotonicity in actions.** Removing a pure strategy of a player cannot increase her value. That is, if $G = (\{1,2\}, S_1 \times S_2, U)$, a_1 is a pure strategy of player 1, and $G' = (\{1,2\}, S_1' \times S_2, U|_{S_1' \times S_2})$, where $S_1' = S_1 - \{a_1\}$, then $f_1(G') \le f_1(G)$, and similarly if we replace S_2 by $S_2' \subseteq S_2$.

4. **Payoff dominance.** If, for all pure strategy profile $s \in S$, a player's expected payoff is strictly larger than her opponent's, then her value should be at least as large as the opponent's. That is, if $U_i(s) \ge U_{3-i}(s)$ for all $s \in S$, then $f_i(G) \ge f_{3-i}(G)$.

5. **Invariance to replicated strategies.** Adding a mixed strategy of player 1 as a new pure strategy for her does not change the value of the game; similarly for player 2. That is, if $G = (\{1,2\}, S_1 \times S_2, U)$, $s_1, s_1' \in A_1$ (recall that A_1 is the set of pure actions of player 1) , $t \notin A_1$, and $\alpha \in [0,1]$, let $G' = (\{1,2\}, S_1' \times S_2, U')$, such that $A_1' = A_1 \cup \{t\}$, $U'(t, s_2) = \alpha U(s_1, s_2) + (1-\alpha) U(s_1', s_2)$ for all $s_2 \in S_2$, and $U'(s) = U(s)$ for all $s \in S$ (so that G' extends G by adding to S_1 one new strategy, that is a convex combination of two strategies in S_1). Then $f(G) = f(G')$. The same holds if we add a replicated strategy to S_2.

Theorem 23. *[?] Axioms 1-5 characterize the coco value in two-player normal-formal games.*[3]

PROOF. See [?]. ∎

We prove that the M-PCE value in two-player games with side payments satisfies all axioms 1-5 except payoff dominance. Now we apply the function f only to games of the form G^*. We state the assumptions in terms of G, but could equally well state them in terms of G^*. Thus, for Axiom 1, it is irrelevant whether we require $f_1(G^*) + f_2(G^*)$ to be $MSW(G)$ or $MSW(G^*)$, since they are the same. Similarly , for shift invariance, it is easy to check $(G^a)^* = (G^*)^a$, so it does not matter whether we apply the shift before or after transforming the game to one that allows side payments.

The fact that the M-PCE value does not satisfy payoff dominance was already observed in Example 19. The following result shows that it satisfies all the other axioms.

Theorem 24. *The function mapping 2-player games with side payments to their (unique) M-PCE value satisfies maximum social welfare, shift invariance, monotonicity in actions, and invariance in replicated strategies.*

PROOF. We consider each property in turn:

- The fact that the function satisfies maximum social welfare is immediate from the characterization in Theorem 22.
- It is easy to see that $MSW(G^a) = MSW(G) + a_1 + a_2$, $mm_1(G^a) = mm_1(G) + a_1$ and $mm_2(G^a) = mm_2(G) + a_2$. It then follows from Theorem 22 that the M-PCE value of $(G^a)^*$ is the result of adding a to the M-PCE value of G^*.
- Let G' be as in the description of Axiom 3 (Monotonicity in actions). It is almost immediate from the definitions that $MSW(G') \le MSW(G)$, $mm_1(G') \le mm_1(G)$, and $mm_2(G') \ge mm_2(G)$. The result now follows from Theorem 22.
- Let G' be the result of adding a replicated action to S_1, as described in the statement of Axiom 4 (Invariance of replicated actions). Clearly $MSW(G') = MSW(G)$, $mm_1(G') = mm_1(G)$, and $mm_2(G') = mm_2(G)$. (For the latter two equalities, note that if $t = \alpha s_1' + (1-\alpha) s_1''$, then the expected utility if player 1 plays t is always between the expected utility if player 1 plays s_1' and if player 1 plays s_1'', so adding t does not change the minimax value of the game.) Again, the result now follows from Theorem 22.

∎

Our goal now is to axiomatize the M-PCE value in games with side payments. Since the M-PCE value and the coco value are different in general, there must be a difference in their axiomatizations. Interestingly, we can capture the difference by replacing payoff dominance by another simple axiom:

6. **Minimax dominance.** If a player's minimax value is no less than her opponent's minimax value, then her value is no less than her opponent's. That is, if $mm_i(G) \ge mm_{3-i}(G)$, then $f_i(G^*) \ge f_{3-i}(G^*)$.

It is immediate from Theorem 22 that the M-PCE value satisfies minimax dominance; Example 19 shows that the coco value does not satisfy it. We now prove that the M-PCE value is characterized by axioms 1, 2, and 6. (Although axioms 3 and 5 also hold for the M-PCE value, we do not need them for the axiomatizaion.)

Theorem 25. *Axioms 1, 2, and 6 characterize the M-PCE value in two-player games with side payments.*

PROOF. Theorem 24 shows that the M-PCE value satisfies axioms 1 and 2. As we observed, the fact that the M-PCE value satisfies axiom 6 is immediate from Theorem 22.

To see that the M-PCE value is the unique mapping that satisfies axioms 1, 2, and 6, suppose that f is a mapping that satisfies these axioms. We want to show that $f(G^*)$ is the M-PCE value for all games G. So consider an arbitrary game G such that the M-PCE value of G^* is $v = (v_1, v_2)$. By shift invariance, the M-PCE value of $(G^{-v})^*$ is $(0,0)$. By axiom 1, $MSW(G) = v_1 + v_2$. and $MSW(G^{-v}) = 0$. Note that it follows from Theorem 22 that $0 = MSW(G^{-v}) + mm_1(G^{-v}) - mm_2(G^{-v})$. Since $MSW(G^{-v}) = 0$, it follows that $mm_1(G^{-v}) = mm_2(G^{-v})$. Suppose that $f((G^{-v})^*) = (v_1', v_2')$. By axiom 1, we must have $v_1' + v_2' = 0$. By axiom 6, since $mm_1(G^{-v}) = mm_2(G^{-v})$, we must have $v_1' = v_2'$. Thus, $f((G^{-v})^*) = (0,0)$. By shift invariance, $f(G^*) = f((G^{-v})^*) + v = (v_1, v_2)$, as desired. ∎

[3] Kalai and Kalai actually consider Bayesian games in their characterization, and have an additional axiom that they call *monotonicity in information.* This axiom trivializes in normal-form games (which can be viewed as the special case of Bayesian games where players have exactly one possible type). It is easy to see that their proof shows that Axioms 1-5 characrerizes the coco value in normal-forml games.

Again, we end this subsection by considering what happens if a default payoff is used instead of backup strategies when defining games with side payments. It is still the case that the M-PCE value satisfies axioms 1, 2, 3, and 5 and does not satisfy axiom 4 (payoff dominance). To get an axiomatization of the M-PCE value in such games with side payments, we simply need to change the minimax dominance axiom to a default value dominance axiom: if the default value of a player is no less than the default value of the opponent, then the player's value is no less than the opponent's value. Thus, variations in the notion of games with side payments lead to straightforward variations in the characterizaiton of the M-PCE value.

7. RELATED WORK

In this section, we compare PCE to other solution concepts in the literature.

We have already seen the PCE is incomparable to NE. There are games (like Traveler's Dilemma) where the NE is not a PCE, and no PCE is a NE. Of course, the same will be true for refinements of NE. *Rationalizability* [8] is a solution concept that generalizes NE; every NE is rationalizable, but the converse is not necessarily true. Intuitively, a strategy of player i is rationalizable if it is a best response to some beliefs that player i may have about the strategies that other players are following, assuming that these strategies are themselves best responses to beliefs that the other players have about strategies that other players are following, and so on. Again, the Traveler's Dilemma shows that the notion rationalizability is incomparable to PCE—the only rationalizable strategy profile in Traveler's Dilemma is $(2, 2)$. Similarly, in Prisoner's Dilemma (Cooperate, Cooperate) is not rationalizable since Cooperate is not a best response to any action. Thus, rationalizability is not getting at the notion of cooperation in the way the PCE is.

Although PCE is meant to apply to one-shot games, our motivation for it involved repeated games. It is thus interesting to compare Cooperative Equilibrium to solutions of repeated games. The well-known *Folk Theorem* [8] says that any payoff profile that gives each player at least his minimax utility is the payoff profile of some NE in the repeated game. Moreover, the proof of the Folk Theorem shows that if s is a strategy in the underlying normal-form game where each player's utility is higher than the minimax utility in the repeated game, then there is a NE in the repeated game where s is played at each round. Thus, playing cooperatively repeatedly in the repeated game will typically be an outcome of a NE. However, so will many other behaviors. Because so many behaviors are consistent with the Folk Theorem, it has very little predictive power. For example, in repeated Traveler's Dilemma, a player can ensure a payoff of at least 2 per iteration simply by always playing 2. It follows from the Folk Theorem that for any strategy profile s in the one-shot game where each player gets at least 2, there is a NE in the repeated game where each player i plays s_i in each round. By way of contrast, as we have seen, in a PCE of the single-shot game, each player gets more than 98. More generally, we can show that, for each PCE s in a normal-form game, there is a NE of the repeated game where s is played repeatedly.

Perhaps the solution concept that gives results closest to PCE is the recently-introduced notion of *iterated regret minimization* (IRM) [5]. As its name suggests, IRM iteratively deletes strategies that do not minimize regret. Although it based on a quite different philosophy than PCE or its variants, IRM leads to quite similar predictions as PCE in a surprising number of games. For example, in Traveler's Dilemma, (97, 97) is the unique profile that survives IRM. In the Nash bargaining game, (50, 50) is the unique profile that survives IRM and is also the unique M-PCE of the game. There are a number of other games of interest where PCE and IRM either coincide or are close.

There are also games in which they behave differently. For example, consider a variant of Prisoner's Dilemma with the following payoff matrix:

	Cooperate	Defect
Cooperate	(10000,10000)	(0,10001)
Defect	(10001,0)	(1,1)

It can be shown that, in general, if there are dominant actions in a game, then these are the only actions that survive IRM. Since defecting is the only dominant action in this game, it follows that (Defect, Defect) is the only strategy profile that survives IRM, giving a payoff (1, 1). On the other hand, the unique M-PCE is (Cooperate, Cooperate) with payoffs (10000, 10000) (although (Defect, Defect) is also a PCE). In this game, M-PCE seems to do a better job of explaining behavior than PCE.

Nevertheless, the fact that PCE and IRM lead to similar answers in so many games of interest suggests that there may be some deep connection between them. We leave the problem of explaining this connection to future work.

8. CONCLUSION

APPENDIX

LEMMA 17. *A simple bilinear program of size* 2×2 *can be solved in constant time.*

PROOF. Let P be the following simple bilinear program, where $x = [x_1\ x_2]^T$, $y = [y_1\ y_2]^T$:

$$\begin{array}{ll} \text{maximize} & x^T \mathbf{A} y + x^T c + y^T c' \\ \text{subject to} & x^T \mathbf{B} y \geq d_1 \\ & x_1 + x_2 = d_2 \\ & y_1 + y_2 = d_3 \\ & x \geq 0 \\ & y \geq 0, \end{array}$$

where $\mathbf{A}$ and $\mathbf{B}$ are 2×2 matrices.

We show that P can be solved in constant time. That is, we either find an optimal solution of P, or find that P has no optimal solution in constant time. The idea is to show that P can be reduced into eight simpler problems, each of which can more obviously be solved in constant time.

Suppose that $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$.

Then we can write P as the following quadratic program Q:

$$\begin{aligned}\text{maximize } & a_{11}x_1y_1 + a_{12}x_1y_2 + a_{21}x_2y_1 + a_{22}x_2y_2 + c[1]x_1 + \\ & c[2]x_2 + c'[1]y_1 + c'[2]y_2 \\ \text{subject to } & b_{11}x_1y_1 + b_{12}x_1y_2 + b_{21}x_2y_1 + b_{22}x_2y_2 - d_1 \geq 0 \\ & x_1 + x_2 = d_2 \\ & y_1 + y_2 = d_3 \\ & x_1, x_2, y_1, y_2 \geq 0.\end{aligned}$$

After replacing x_2 with $(d_2 - x_1)$ and y_2 with $(d_3 - y_1)$, then rearranging terms, the objective function of Q becomes

$$(a_{11} - a_{12} - a_{21} + a_{22})x_1y_1 + (a_{12}d_3 - a_{22}d_3 + c[1] - c[2])x_1 + (a_{21}d_2 - a_{22}d_2 + c'[1] - c'[2])y_1 + (a_{22}d_2d_3 + c[2]d_2 + c'[2]d_3),$$

and the first constraint becomes

$$(b_{11} - b_{12} - b_{21} + b_{22})x_1y_1 + (b_{12}d_3 - b_{22}d_3)x_1 + (b_{21}d_2 - b_{22}d_2)y_1 + (b_{22}d_2d_3 - d_1).$$

We can get an equivalent problem by removing the constant terms $a_{22}d_2d_3 + c[2]d_2 + c'[2]d_3$ from the objective function, since adding or removing additive constants from a function that we want to maximize does not affect its optimal solutions (e.g., "maximize x" has the same optimal solutions as "maximize $(x+1)$").

Thus, Q is equivalent to the following quadratic program Q':

$$\begin{aligned}\text{maximize } & \gamma_1x_1y_1 + \gamma_2x_1 + \gamma_3y_1 \\ \text{subject to } & \gamma_4x_1y_1 + \gamma_5x_1 + \gamma_6y_1 + \gamma_7 \geq 0 \\ & x_1 \in [0, d_2] \\ & y_1 \in [0, d_3],\end{aligned}$$

where

$$\begin{aligned}\gamma_1 &= a_{11} - a_{12} - a_{21} + a_{22} \\ \gamma_2 &= a_{12}d_3 - a_{22}d_3 + c[1] - c[2] \\ \gamma_3 &= a_{21}d_2 - a_{22}d_2 + c'[1] - c'[2] \\ \gamma_4 &= b_{11} - b_{12} - b_{21} + b_{22} \\ \gamma_5 &= b_{12}d_3 - b_{22}d_3 \\ \gamma_6 &= b_{21}d_2 - b_{22}d_2 \\ \gamma_7 &= b_{22}d_2d_3 - d_1.\end{aligned}$$

(Note that γ_1–γ_7 are all constants.)

The first step in solving Q' involves expressing the values of y_1 that make (x_1, y_1) a feasible solution, that is, one that satisfies the constraint

$$\gamma_4x_1y_1 + \gamma_5x_1 + \gamma_6y_1 = (\gamma_4y_1 + \gamma_5)x_1 + \gamma_5x_1 + \gamma_6y_1 \geq 0.$$

For each $y_1 \in [0, d_3]$, let $\Psi_1(y_1)$ be the set of x_1 such that (x_1, y_1) a feasible solution of Q'. The characterization of $\Psi_1(y_1)$ depends on the sign of $\gamma_4y_1 + \gamma_5$. Specifically:

$$\begin{cases}\Psi_1(y_1) = \left[\frac{-\gamma_6y_1-\gamma_7}{\gamma_4y_1+\gamma_5}, d_2\right] \cap [0, d_2], \\ \quad \text{if } \gamma_4y_1 + \gamma_5 > 0, \frac{-\gamma_6y_1-\gamma_7}{\gamma_4y_1+\gamma_5} \leq d_2 \\ \Psi_1(y_1) = \left[0, \frac{-\gamma_6y_1-\gamma_7}{\gamma_4y_1+\gamma_5}\right] \cap [0, d_2], \\ \quad \text{if } \gamma_4y_1 + \gamma_5 < 0, \frac{-\gamma_6y_1-\gamma_7}{\gamma_4y_1+\gamma_5} \geq 0 \\ \Psi_1(y_1) = [0, d_2], \text{ if } \gamma_4y_1 + \gamma_5 = 0, \gamma_6y_1 + \gamma_7 \geq 0 \\ \Psi_1(y_1) = \emptyset, \quad \text{if } \gamma_4y_1 + \gamma_5 = 0, \gamma_6y_1 + \gamma_7 < 0.\end{cases} \tag{1}$$

Note that the first three regions are intervals.

Let $f(x_1, y_1) = \gamma_1x_1y_1 + \gamma_2x_1 + \gamma_3y_1$ be the objective function of Q'. We want to maximize f over all feasible pairs (x_1, y_1). Taking the derivative of f with respect to x_1, we get

$$\frac{\partial f(x_1, y_1)}{\partial x_1} = \gamma_1y_1 + \gamma_2,$$

which is a linear function of y_1. Because the derivative is linear, for each fixed value of y_1, the value that maximizes $f(x_1, y_1)$ must lie at an endpoint of the interval appropriate for that value of y_1. Whether it is the left endpoint or the right endpoint depends on whether the derviative is negative or positive. For example, if y_1 satisifes the constraints corresonding to the first interval in (1) (i.e., if $\gamma_4y_1 + \gamma_5 > 0$ and $\frac{-\gamma_6y_1-\gamma_7}{\gamma_4y_1+\gamma_5} \leq d_2$) and $\gamma_1y_1 + \gamma_2 > 0$, then $x_1 = d_2$ (i.e., the right endpoint of the interval of $Psi_1(y_1)$) maximizes $f(x_1, y_1)$; the problem of maximizing $f(x_1, y_1)$ reduces to that of maximizing $f(d_2, y_1)$ (see Q_1 below). On the other hand, if $\gamma_1y_1 + \gamma_2 > 0$, then maximizing $f(x_1, y_1)$ reduces to maximizing $f(0, y_1)$ or $f(\frac{\gamma_6-\gamma_7}{\gamma_4y_1+\gamma_5}, y_1)$, depending on whether $\frac{-gamma_6y_1-\gamma_7}{\gamma_4y_1+\gamma_5}$ is negative (see Q_5 and Q_6 below.

These considerations show that to find that value (x_1, y_1) that maximizes $f(x_1, y_1)$, it suffices to find the value of y_1 that maximizes each of the expressions below, and take the one that is best among these:

Q_1 : maximize $f(d_2, y_1)$, subject to
$\gamma_4y_1 + \gamma_5 > 0, \frac{-\gamma_6y_1-\gamma_7}{\gamma_4y_1+\gamma_5} \leq d_2, \gamma_1y_1 + \gamma_2 \geq 0, y_1 \in [0, d_3]$
Q_2 : maximize $f(d_2, y_1)$, subject to
$\gamma_4y_1 + \gamma_5 < 0, \frac{-\gamma_6y_1-\gamma_7}{\gamma_4y_1+\gamma_5} \geq d_2, \gamma_1y_1 + \gamma_2 \geq 0, y_1 \in [0, d_3]$
Q_3 : maximize $f(\frac{-\gamma_6y_1-\gamma_7}{\gamma_4y_1+\gamma_5}, y_1)$, subject to
$\gamma_4y_1 + \gamma_5 < 0, 0 \leq \frac{-\gamma_6y_1-\gamma_7}{\gamma_4y_1+\gamma_5} \leq d_2, \gamma_1y_1 + \gamma_2 \geq 0, y_1 \in [0, d_3]$
Q_4 : maximize $f(d_2, y_1)$, subject to
$\gamma_4y_1 + \gamma_5 = 0, \gamma_6y_1 + \gamma_7 \geq 0, \gamma_1y_1 + \gamma_2 \geq 0, y_1 \in [0, d_3]$
Q_5 : maximize $f(0, y_1)$, subject to
$\gamma_4y_1 + \gamma_5 > 0, \frac{-\gamma_6y_1-\gamma_7}{\gamma_4y_1+\gamma_5} \leq 0, \gamma_1y_1 + \gamma_2 < 0, y_1 \in [0, d_3]$
Q_6 : maximize $f(\frac{-\gamma_6y_1-\gamma_7}{\gamma_4y_1+\gamma_5}, y_1)$, subject to
$\gamma_4y_1 + \gamma_5 > 0, 0 \leq \frac{-\gamma_6y_1-\gamma_7}{\gamma_4y_1+\gamma_5} \leq d_2, \gamma_1y_1 + \gamma_2 < 0, y_1 \in [0, d_3]$
Q_7 : maximize $f(0, y_1)$, subject to
$\gamma_4y_1 + \gamma_5 < 0, \frac{-\gamma_6y_1-\gamma_7}{\gamma_4y_1+\gamma_5} \geq 0, \gamma_1y_1 + \gamma_2 < 0, y_1 \in [0, d_3]$
Q_8 : maximize $f(0, y_1)$, subject to
$\gamma_4y_1 + \gamma_5 = 0, \gamma_6y_1 + \gamma_7 \geq 0, \gamma_1y_1 + \gamma_2 < 0, y_1 \in [0, d_3]$.

Note that Q_1, Q_2, and Q_3 describe the possibilities for the first interval in (1), Q_4, Q_5, and Q_6 are the possibilities for the second subinterval, and Q_7 and Q_8 are the possibilities for the third subinterval.

Each of Q_1, Q_2, Q_4, Q_5, Q_7, and Q_8 can be easily rewritten as linear programs of a single variable (y_1), so can be solved in constant time. With a little more effort, we can show Q_3 and Q_6 can also be solved in constant time. We explain how this can be done for Q_3. The argument for Q_6 is similar and left to the reader. All the constraints in Q_3 can again be viewed as linear constraints; the set of feasible values of y_1 is thus an interval, whose endpoints can clearly be computd in constant time. Now the objective function is

$$f(\frac{-\gamma_6y_1-\gamma_7}{\gamma_4y_1+\gamma_5}, y_1) = \frac{\gamma_1(-\gamma_6y_1-\gamma_7)y_1}{\gamma_4y_1+\gamma_5} + \frac{\gamma_2(-\gamma_6y_1-\gamma_7)}{\gamma_4y_1+\gamma_5} + \gamma_3y_1.$$

To find the maximum value of the objective function among the feasible values, we need to take its derivative (with respect to y_1). A straightforward calculation shows that this derivative is

$$\frac{(-2\gamma_1\gamma_6y_1 - \gamma_1\gamma_7 - \gamma_2\gamma_6)(\gamma_4y_1+\gamma_5) - \gamma_4(\gamma_6y_1+\gamma_7)(\gamma_1y_1+\gamma_2)}{(\gamma_4y_1+\gamma_5)^2} + \gamma_3.$$

This derivative is 0 when its numerator is 0 (since the constraints in Q_3 guarantee that the denominator is positive). The numerator is a quadratic, so can be solved in constant time.

Thus, to find the optimal value for Q_3, we must just check f at the endpoints of the interval defined by the constraints (which, as we observed above, can be computed in constant time) and at the point where the derivative is 0 (which can also be computed in constant time). Thus, Q_3 can be solved in constant time.

This completes the argument that Q can be solved in constant time. ∎

A. REFERENCES

[1] K. Basu. The traveler's dilemma: paradoxes of rationality in game theory. *American Economic Review*, 84(2):391–395, 1994.

[2] K. Basu. The traveler's dilemma. *Scientific American*, June:90–95, 2007.

[3] T. Becker, M. Carter, and J. Naeve. Experts playing the Traveler's Dilemma. Discussion paper 252/2005, Universität Hohenheim, 2005.

[4] W. Güth, R. Schmittberger, and B. Schwarze. An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, 3(4):367–388, 1982.

[5] J. Y. Halpern and R. Pass. Iterated regret minimization: A new solution concept. In *Proc. Twenty-First International Joint Conference on Artificial Intelligence (IJCAI '09)*, pages 153–158, 2009.

[6] C. Hauert, A. Traulsen, H. Brandt, M. Nowak, and K. Sigmund. Via freedom to coercion: the emergence of costly punishment. *Science*, 316(5833):1905–1907, 2007.

[7] J. Nash. The barganing problem. *Econometrica*, 18:155–162, 1950.

[8] M. J. Osborne and A. Rubinstein. *A Course in Game Theory*. MIT Press, Cambridge, Mass., 1994.

[9] D. Quervain, U. Fischbacher, V. Treyer, M. Schellhammer, U. Schnyder, A. Buck, and E. Fehr. The neural basis of altruistic punishment. *Science*, 305:1254–1258, August 2004.

[10] R. W. Rosenthal. Games of perfect information, predatory pricing, and the chain store paradox. *Journal of Economic Theory*, 25:92–100, 1982.

[11] K. Sigmund. Punish or perish? retaliation and collaboration among humans. 22(11):593–600, 2007.