Towards a Deeper Understanding of Cooperative Equilibrium: Characterization and Complexity

Nan Rong Joseph Y. Halpern
Department of Computer Science
Cornell University
Ithaca, NY 14853
{rongnan,halpern}@cs.cornell.edu

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. We consider two solution concepts that were introduced recently that try to capture such cooperation in two-player games: *perfect cooperative equilibrium* (PCE) (and an extension called *maximum PCE (M-PCE)*) [8] and the *coco value* [11]. We show that, despite their definitions being quite different, these notions are closely related, both in terms of axiomatization and algebraic characterization. We also consider the problem of computing how well players do when they cooperate according to these solution concepts, and show that in both cases in polynomial time. In the case of the coco value, this follows easily from the definition; in the case of the corresponding *M-PCE* value, it follows from a theorem showing that bilinear programming for a class of 2×2 matrices is in constant time, a result that may be of independent interest.

Categories and Subject Descriptors

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

General Terms

Economics, Theory

Keywords

Cooperative equilibrium, PCE, M-PCE, coco value, game theory, cooperation

1. INTRODUCTION

Standard solution concepts in game theory such as Nash equilibrium assume that players are self-interested, and are out to get the highest possible payoffs for themselves. However, they are not very good at predicting behavior in games with an element of cooperation such as Prisoner's Dilemma, Traveler's Dilemma [2, 3], or the Centipede Game [14]. In this paper, we consider two solution concepts that seem to provide better predictions in such games: *perfect cooperative equilibrium* (PCE) (and an extension called *maximum PCE (M-PCE)*) [8] and the *coco* (cooperative-competitive) value, introduced by Kalai and Kalai [11]. Although the definitions of the two notions are quite different on the surface, they often make similar predictions. In this paper, we explain this phenomenon by providing an axiomatic characterization of the M-PCE value of a game (the payoff that players obtain when they play an M-PCE), and show that it is closely related to the axiomatic characterization of the coco value given by Kalai and Kalai [11]. We also provide an algebraic characterization of the the M-PCE value and coco value in terms of some standard parameters of the game (specifically, the maximum social welfare obtainable in the game and the minimax value of the game), again showing how closely related they are.

There is a subtlety involved in making this comparison: the definition of coco value implicitly assumes that utility transfers are possible in the game; that is, a player can make a side payments to the other player, and players value these payments the same way (that is, they are implicitly equating utility with money). Side payments are not assumed when defining M-PCE. In order to make for a better comparison, we provide a general technique, which may be of independent interest, for converting a 2-player game without side payments into one with side payments. We also show that such games have a unique M-PCE value: that is, there is a unique payoff that players get when playing an M-PCE.

Finally, we turn to complexity-theoretic considerations. It follows easily from the characterizations that computing the coco value can be done in time polynomial in the description of the game (specifically, in time polynomial in the number of strategies available in the game). This is also true for the M-PCE value in games of transferable utility. Computing the strategies that give the coco value is also easy. Essentially, this is because the coco value is obtained by playing the strategies that give the maximum social welfare and then making side payments. Computing the strategies that give the M-PCE value in a game with side payments is easy as well. However, computing the strategies that give a PCE (if one exists) or an M-PCE in a general game is not so straightforward. We show that these can be computed in polynomial time in 2-player games by using ideas from bilinear programming [1, 15]. A key ingredient of our argument is a proof that bilinear programming for a class of 2×2 matrices is solvable in constant time, a result that may be of independent interest.

The rest of the paper is organized as follows. In Section 2,

Appears in: *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013),* Ito, Jonker, Gini, and Shehory (eds.), May 6–10, 2013, Saint Paul, Minnesota, USA.

we review PCE, M-PCE, and the coco value, and compare them in a number of games of interest. In Section 3, we compare M-PCE with the coco value, using both algebraic and axiomatic characterizations. In Section 4, we prove that both a PCE and an M-PCE can be computed in polynomial time in two-player games using bilinear programming. We conclude in Section 5.

We remark that although our focus here is on PCE and the coco value, there have been a number of attempts to model cooperation; see [4, 5, 6, 9, 10, 12, 16, 18]. Moreover, other solution concepts, such as *iterated regret minimization* [7], while not intended to model cooperation, often produce results similar to PCE. We discuss this work in the full paper, which can be found at www.cs.cornell.edu/home/halpern/papers/CoopEq.pdf. Omitted proofs can also be found there.

2. PCE, M-PCE, AND COCO VALUE: A REVIEW

In this section, we review PCE, α-PCE, and M-PCE [8] and the coco value [11]. Kalai and Kalai define the coco value only for 2-player Bayesian games. We defined PCE and its variants for arbitrary normal-form games, but the definition is best motivated for 2-player games. Thus, in this paper, we consider only 2-player normal-form games, and review the definition only for that case.

Formally, a two player game has the form (A, u), where $A = A_1 \times A_2$, A_i is a finite set of possible actions for player i, $u = (u_1, u_2)$, and u_i is player i's utility function, that is, $u_i(a_1, a_2)$ is player i's utility or payoff if the action profile $a = (a_1, a_2)$ 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 $U_i(s_1, s_2)$ denote player i's expected utility if the strategy profile $s = (s_1, s_2)$ is played. Given a profile $x = (x_1, x_2)$, let x_{-i} denote the strategy of the player other than i.

2.1 PCE and M-PCE

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 other player if $U_i(s_i, s_{-i}) = \sup_{s'_i \in S_i} U_i(s'_i, s_{-i})$. 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.

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.)

We defined PCE for two-player games as follows [8].

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

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

We showed [8] that PCE has a number of attractive properties. For example, in the Prisoner's Dilemma, cooperation is a PCE, as is any mixed strategy that guarantees both players a payoff of at least what they would get if they defected. More generally, players do at least as well in a PCE as they do in any NE of the game. Unfortunately, it easily follows from this that a PCE does not always exist (see Example 1 below).

Motivated in part by this observation, we considered α-PCE, a more quantitative version of PCE that intuitively takes into account the degree of cooperation exhibited by a strategy profile.

Definition 4. A strategy profile s is an *α-PCE* in a game G if $U_i(s) \geq \alpha + BU_i^G$ for all $i \in N$. [8]

Definition 5. 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.[8]

It is easy to see that every game is guaranteed to have an M-PCE. Moreover, the M-PCE does well at predicting behavior in quite a few games of interest [8]. For example, in the Prisoner's Dilemma, cooperation gives the unique M-PCE. We consider other example shortly.

2.2 The coco value

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 [11] 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.

The *minimax value of game G for player i*, denoted $mm_i(G)$, is the payoff player i gets when the opponent is minimizing i's maximum payoff; formally,

$$mm_1(G) = \min_{s_2 \in S_2} \max_{s_1 \in S_1} U_1(s_1, s_2);$$

$mm_2(G)$ is defined similarly, interchanging 1 and 2.

We are now ready to define the coco value. Given a game G, let a 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[1] 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 [11] argue that it is also fair in an appropriate sense.

[1]We use minimax rather than maximin (recall that $\max\min_i(G) = \max_{s_i \in S_i} \min_{s_j \in S_j} U_i(s)$) although they are equal in zero-sum games, because minimax characterizes the M-PCE value even in non-zero-sum games, while maximin does not.

2.3 Examples

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

Example 1. *The Nash bargaining game [13]:* In the Nash bargaining game, 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. There is clearly no strategy profile that gives both players a higher It easily follows that the Nash bargaining game has no PCE.

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. Thus, the coco value of the Nash bargaining game is $(50, 50)$; it is also the unique M-PCE value [8]. To see this, observe that in the Nash bargaining game, $BU_1 = BU_2 = 100$, so the best α-PCE that can be obtained is a 50-PCE, which is obtained when both players ask for 50. ▮

Example 2. *Prisoner's Dilemma:* In the Prisoner's Dilemma, 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 NE here is (Defect, Defect), people often do play (Cooperate, Cooperate).

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 $(3, 3)$. Again, this is the unique M-PCE value [8]: it is easy to see that for Prisoner's Dilemma with these payoffs, $BU_1 = BU_2 = 1$, so by both cooperating, the players have a 2-PCE, which is clearly also an M-PCE. ▮

Example 3. *Traveler's Dilemma:* In the Traveler's Dilemma [2, 3], 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 $(2, 2)$.

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 $(100, 100)$. Again, this is also the unique M-PCE value [8]. ▮

As the next example shows, there are games in which the coco value and M-PCE value differ.

Example 4. *The centipede game:* In the Centipede game [14], 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 with a very low payoff profile $(1, 3)$. But, in practice, people continue the game for quite a while.

It can be computed that, in this game, $BU_1 = 2^{19} + \frac{3\times 2^{18}}{3\times 2^{18}+1}$ and $BU_2 = 2^{18} + \frac{3\times 2^{17}}{3\times 2^{17}+1}$. Thus, any strategy profile s where $U_1(s) \geq BU_1$ and $U_2(s) \geq BU_2$ is a PCE. A straightforward computation shows that the unique M-PCE s^* in this game is one where player 1 plays through round 20 with probability β, and stops at round 19 with probability $1 - \beta$, while player 2 plays through round 20, where $\beta = \frac{1}{3\times 2^{18}+2} - \frac{3\times 2^{17}}{(3\times 2^{18}+2)(3\times 2^{18}+1)(3\times 2^{17}+1)}$. This gives the M-PCE value $(2^{19}+1-\beta, 2^{18}+(3\times 2^{18}+1)\beta) \approx (2^{19}+1, 2^{18}+1)$.

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 $(\frac{2^{19}+2^{20}+1}{2} + 1, \frac{2^{19}+2^{20}+1}{2} - 1) = (\frac{2^{19}+2^{20}+3}{2}, \frac{2^{19}+2^{20}-1}{2})$. 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 the largest payoff obtainable in G_t 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 in the next section.

3. CHARACTERIZING THE M-PCE AND COCO VALUES

In this section, we characterize M-PCE and coco values using two approaches: first algebraically, then axiomatically. The characterizations help clarify the relationship between the two notions.

As we have observed, the coco value makes sense only if players can make side payments. 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).[2] The ability to make side payments is not explicitly modeled in the description of the games considered by Kalai and Kalai [11]. 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.

3.1 Two-player games with side payments

In this subsection, we describe 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 transferred. If the players propose the same deal, then the suggested strategy profile is played, and the money is transferred. Otherwise, a "backup" action 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* G, where A^* and u^* are defined as follows. A^* extends A by adding a collection of actions that we call *deal actions*. A deal action for player i is a triple of the form $(a, r, a'_i) \in A \times I\!\!R \times A_i$. Intuitively, this action proposes that the players play the action profile a and that player 1 should transfer r to player 2; if the deal is not accepted, then player i plays a'_i. Given this intuition, it should be clear how u^* extends u. For action profiles $a \in A$, $u^*(a) = u(a)$. For profiles actions $a \in (A_1^* - A_1) \times (A_2^* - A_2)$, the players agree on a deal if they both propose a deal strategy with the same first two components (a, r). In this case they play a and r is transferred. Otherwise, players just play the backup action. More precisely, for $a, a' \in A$, $b_i \in A_i$, and $r, r' \in I\!\!R$:

- $u^*(a) = u(a)$;
- $u_1^*((a, r, b_1), (a, r, b_2)) = u_1(a) - r$;
  $u_2^*((a, r, b_1), (a, r, b_2)) = u_2(a) + r$;
- $u^*((a, r, b_1), (a', r', b_2)) = u(b_1, b_2)$ if $(a, r) \neq (a', r')$;
- $u^*((a, r, b_1), b_2) = u^*(b_1, (a', r', b_2)) = u(b_1, b_2)$.

As usual, players are allowed to randomize, and a strategy of player i in G^* is a distribution over actions in A_i^*; let S_i^* represent the set of player i's strategies. Let $U_i^*(s)$ denote player i's expected utility if the strategy profile $s \in S^*$ is played. We call G^* the game with side payments *extending* G, and call G the game *underlying* G^*.

Intuitively, when both players play deal actions, we can think of them as giving their actions to a trusted third party. If they both propose the same deal, the third party ensures that the deal action is carried out and the transfer is made. Otherwise, the appropriate backup actions are played.

In our approach, we have allowed players to propose arbitrary backup actions 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 3.2.

3.2 Algebraic characterization

At first glance, the coco value and the M-PCE value seem quite different, although both are trying to get at the notion of cooperation. However, we show below that both have quite similar characterizations. In this section, we characterize the two notions algebraically, 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). Let $MSW(G)$ be the maximum social welfare of G; formally, $MSW(G) = \max_{a \in A}(u_1(a) + u_2(a))$.

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

This proof (and all further omitted proofs) can be found in the full paper.

We now characterize the coco value.

Theorem 2. *If G is a two-player game, then $coco(G) = \left(\frac{MSW(G)+mm_1(G_z)-mm_2(G_z)}{2}, \frac{MSW(G)-mm_1(G_z)+mm_2(G_z)}{2}\right)$.[3] Moreover, $coco(G) = coco(G^*)$.*

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

$$
\begin{aligned}
& coco(G) \\
= \; & \left(\tfrac{MSW(G)}{2}, \tfrac{MSW(G)}{2}\right) + (mm_1(G_z), mm_2(G_z)) \\
= \; & \left(\tfrac{MSW(G)+2mm_1(G_z)}{2}, \tfrac{MSW(G)+2mm_2(G_z)}{2}\right) \\
= \; & \left(\tfrac{MSW(G)+mm_1(G_z)-mm_2(G_z)}{2}, \tfrac{MSW(G)-mm_1(G_z)+mm_2(G_z)}{2}\right).
\end{aligned}
$$

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 1), and the fact that $(G_z)^* = (G^*)_z$, which we leave to the reader to check. ∎

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

[3] Note that $mm_1(G_z) = -mm_2(G_z)$ by von Neumann's minimax theorem [17] (which says that in every two-player zero-sum games, there is an equilibrium where both players play a minimax strategy). We write the expression in the form above to better compare it to the M-PCE value.

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 3. *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 a^* be an action profile in G that maximizes social welfare, that is, $U_1(a^*) + U_2(a^*) = MSW(G)$, and let (s_1', s_2') be a strategy 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 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 $s = (s_1, s_2)$ be such that, in s_1, player 1 plays deal action $(a^*, mm_2(G) - u_2(a^*), a_1')$ with the same probability that she plays a_1' in s_1' (where s_1' is as defined above) for all $a_1' \in A_1$; and $s_2 = (a^*, mm_2(G) - u_2(a^*), a_2)$ for some fixed $a_2 \in A_2$. Intuitively, s_1 does the following: if player 2 agrees to the deal in s_1, then a^* is carried out, and player 1 transfers $mm_2(G) - u_2(a^*)$ to player 2; otherwise player 1 plays the mixed strategy s_1'. s_2 is a deal action that agrees to s_1. Thus, $U_1^*(s) = u_1(a^*) - (mm_2(G) - u_2(a^*)) = U_1(a^*) + u_2(a^*) - mm_2(G) = MSW(G) - mm_2(G)$, and $U_2^*(s) = mm_2(G)$. On the other hand, if player 2 plays an action $a_2 \in A_2$, then

$$U_2^*(s_1, s_2) = U_2(s_1', a_2) \le U_2(s') = mm_2(G).$$

Thus, player 2 gets at most $mm_2(G)$ when player 1 plays s_1, 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 1, 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 a^* be the action profile in G defined above that maximizes social welfare, and let $a' \in A$. Let $s^+ = (s_1^+, s_2^+)$, where $s_1^+ = (a^*, u_1(a^*) - \frac{MSW(G)+mm_1(G)-mm_2(G)}{2}, a_1')$ and $s_2^+ = (a^*, u_1(a^*) - \frac{MSW(G)+mm_1(G)-mm_2(G)}{2}, a_2')$. It is also 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 also easily be checked that $U_i(s^+) = BU_i + \beta$ for $i = 1, 2$, so s^+ is indeed a 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 2 and 3 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 minimax 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 2 and 3 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 can 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, as the following example shows.

Example 5. 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 11, 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 transfers 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, then 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 actions 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

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

Thus, rather than 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.

3.3 Axiomatic comparison

In this section, we provide an axiomatization of the M-PCE value and compare it to the axiomatization of the coco value given by Kalai and Kalai [11]. Before jumping into the axioms, we first explain the term "axiomatize" in this context. Given a function $f : A \rightarrow B$, we say a set AX of axioms *axiomatizes f in A*, if f is the unique function mapping A to B that satisfies 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 [11] 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 $c = (c_1, c_2) \in \mathbb{R}^2$, $G = (\{1, 2\}, A, u)$ and $G^c = (\{1, 2\}, A, u^c)$, where $u_i^c(a) = u_i(a) + c_i$ for all $a \in A$, then $f(G^c) = (f_1(G) + c_1, f_2(G) + c_2)$.

3. **Monotonicity in actions.** Removing an action of a player cannot increase her value. That is, if $G = (\{1, 2\}, A_1 \times A_2, u)$, and $G' = (\{1, 2\}, A_1' \times A_2, u|_{A_1' \times A_2})$, where $A_1' \subseteq A_1$, then $f_1(G') \leq f_1(G)$, and similarly if we replace A_2 by $A_2' \subseteq A_2$.

4. **Payoff dominance.** If, for all action profiles $a \in A$, 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 $u_i(a) \geq u_j(a)$ for all $a \in A$, then $f_i(G) \geq f_j(G)$.

5. **Invariance to replicated strategies.** Adding a mixed strategy of player 1 as a new action for her does not change the value of the game; similarly for player 2. That is, if $G = (\{1, 2\}, A_1 \times A_2, u)$, $t \in S_1$, and $G' = (\{1, 2\}, A_1' \times A_2, u')$, where $A_1' = A_1 \cup \{t\}$, $u'(t, a_2) = U(t, a_2)$ for all $a_2 \in A_2$, and $u'(a) = u(a)$ for all $a \in A$ (so that G' extends G by adding to A_1 one new action, which can be identified with a mixed strategy in S_1). Then $f(G) = f(G')$. The same holds if we add a strategy to A_2.

Theorem 4. *[11] Axioms 1-5 characterize the coco value in two-player normal-formal games.*[4]

PROOF. See [11]. ∎

Note that, following Kalai and Kalai [11], we have stated the axioms for the coco value in terms of the underlying game G. Since, as we have argued, Kalai and Kalai are assuming there are side payments, we might consider stating the axioms in terms of G^*. We could certainly replace all occurrences of $f_i(G)$ by $f_i(G^*)$; nothing would change if we did this, since, by Theorem 2, $coco(G) = coco(G^*)$. But we could go further, replacing G, A, and u uniformly by G^*, A^*, and u^*. For example, Axiom 1 would say $f_1(G^*) + f_2(G^*) = MSW(G^*)$; Axiom 2 would say that $f((G^*)^c) = (f_1(G^*) + c_1, f_2(G^*) + c_2)$. It is not hard to check that the resulting axioms are still sound. Moreover, for all axioms but Axiom 4 (payoff dominance), the resulting axiom is essentially equivalent to the original axiom. (In the case of shift invariance, this is because $(G^*)^c = (G^c)^*$.) However, the version of Axiom 4 for G^* is vacuous. No matter what the payoffs are in G, it cannot be the case that a player's expected payoff is larger than his opponent's for all actions in G^*, since players can always agree to a deal action that results in the opponent getting a large transfer. Thus, we must express payoff dominance in terms of G in order to prove Theorem 4.

We now characterize the M-PCE value axiomatically. The M-PCE value of G is not equal to that of G^* in general. Since we want to compare the M-PCE value and coco value, it is most appropriate to consider games with side payments. Thus, in the axioms, we write $f_i(G^*)$ rather $f_i(G)$. We start by considering the extent to which the M-PCE value satisfies the axioms above for coco value (with $f_i(G)$ replaced by $f_i(G^*)$, a change which, as we noted, has no impact for coco value). Example 5 shows that the M-PCE value does not satisfy payoff dominance. The following result shows that it satisfies all the remaining axioms.

Theorem 5. *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 3.

- It is easy to see that $MSW(G^c) = MSW(G) + c_1 + c_2$, $mm_1(G^c) = mm_1(G) + c_1$ and $mm_2(G^c) = mm_2(G) + c_2$. It then follows from Theorem 3 that the M-PCE value of $(G^c)^*$ is the result of adding c 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 $MSW(G') \leq MSW(G)$, $mm_1(G') \leq mm_1(G)$, and $mm_2(G') \geq mm_2(G)$. The result now follows from Theorem 3.

- Let G' be the result of adding a replicated action to S_1, as described in the statement of Axiom 5 (Invariance to replicated strategies). Clearly $MSW(G') = MSW(G)$, $mm_1(G') = mm_1(G)$, and $mm_2(G') = mm_2(G)$. Again, the result now follows from Theorem 3. ∎

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:

[4]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 characterizes the coco value in normal-form games.

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 $mm_i(G) \geq mm_j(G)$, then $f_i(G^*) \geq f_j(G^*)$.

It is immediate from Theorem 3 that the M-PCE value satisfies minimax dominance; Example 5 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 axiomatization.) Interestingly, for all these axioms, we can replace G, A, and u by G^*, A^*, and u^* to get an equivalent axiom; it really does not matter if we state the axiom in terms of G or G^* (although the argument to f must be G^*).

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

Proof. Theorem 5 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 3.

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 3 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. ∎

Again, we conclude this subsection by considering what happens if a default payoff is used instead of backup actions 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 characterization of the M-PCE value.

4. THE COMPLEXITY OF COMPUTING A COOPERATIVE EQUILIBRIUM

In this section, we consider the complexity of computing the M-PCE value and the coco value, and the corresponding strategy profiles.

It follows easily from the characterization in Theorem 2 that in a two-player game G with (or without) side payments, the coco value is determined by $MSW(G)$, $mm_1(G_z)$, and $mm_2(G_z)$. G_z can clearly be determined from G in polynomial time (polynomial in the number of strategies), and $MSW(G)$ can be determined in polynomial time (simply by inspecting the payoff matrix for G). It can be proved that the minimax value of a game can be computed in polynomial time. (See the full paper for details.) Thus, the coco value can be computed in polynomial time. Moreover, if (c_1, c_2) is the coco value of G, and s^* is a pure strategy profile that obtains $MSW(G)$, the strategy profile that gives players the coco value is $((s^*, U_1(s^*) - c_1), (s^*, U_1(s^*) - c_1))$, which is simply the deal strategy profile in which both players agree to play s^*, and agree that player 1 pays player 2 $(U_1(s^*) - c_1)$ so that they eventually get (c_1, c_2).

Similarly, we can compute an M-PCE in a two-player game with side payments in polynomial time.

Theorem 7. *In a two-player game G^* with side payments, we can compute the M-PCE value and a strategy profile that obtains it in polynomial time.*

Proof. Let G be the game underlying G^*. By Theorem 3, the M-PCE value of G is determined by $MSW(G)$, $mm_1(G)$, and $mm_2(G)$. It can be proved that the minimax value of a two-player game can be computed in polynomial time using linear programming (see the full paper for details). It then follows that the M-PCE value can be computed in polynomial time.

Let the M-PCE value be (m_1, m_2), and let s^* be a pure strategy profile that obtains $MSW(G)$. Then the following must be an M-PCE: $((s^*, U_1(s^*) - m_1), (s^*, U_1(s^*) - m_1))$, which is simply the deal strategy profile in which both players agree to play s^*, and agree that player 1 pays player 2 $U_1(s^*) - m_1$ so that they eventually get payoffs (m_1, m_2) – their M-PCE value. ∎

Moving to arbitrary games (not necessarily ones with side payments), it is much less obvious how the M-PCE value or the strategies that achieve it can be computed efficiently. We now show that in two-player games (without side payments), a PCE can be found in polynomial time if one exists; moreover, determining whether one exists can also be done in polynomial time. Similarly, in two-player games, an M-PCE can always be found in polynomial time. The first step in the argument involves showing that in two-player games, for all strategy profiles 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 for PCE and M-PCE. We then use similar techniques to show that a Pareto-optimal M-PCE, and thus a CE, can be found in polynomial time,

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 8. *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.*

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

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

where $\mathbf{A}$ and $\mathbf{B}_1$ are $n \times m$ matrices, $x, c \in I\!R^n$, $y, c' \in I\!R^m$, $\mathbf{B}_2$ 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 9. *A simple bilinear program of size* 2×2 *can be solved in constant time.*

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 8 to reduce the problem to $\binom{n}{2}\binom{m}{2} = O(n^2m^2)$ smaller problems, each of a which is a simple bilinear program of size 2×2. By Lemma 9, each of these smaller problems can be solved in constant time, giving us a polynomial-time algorithm.

Theorem 10. *Given a two-player game* $G = (\{1, 2\}, A, u)$*, we can compute in polynomial time whether* G *has a PCE and, if so, we can compute a PCE in polynomial time.*

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

Theorem 11. *Given a two-player game* $G = (\{1, 2\}, A, u)$*, we can compute an M-PCE in polynomial time.*

5. CONCLUSION

In this paper, we considered two solution concepts that are intended to capture cooperative behavior: PCE (and M-PCE) and the coco value. As we show, M-PCE value and the coco value coincide in many games of interest. We examined the two solution concepts by characterizing them both algebraically and axiomatically. Our characterizations shows that, despite the apparent differences in their definitions, the two notions are closely related. In the process, we define a technique for converting a 2-player normal-form game to a game with side payments. The fact that these two notions turn out to be so similar gives us hope that they are getting at deep intuitions regarding cooperation.

We also consider the complexity of the two notions. In both two-player games with and without side payments, we show that the coco value and its corresponding strategy profile can be computed in polynomial time. While the same holds for M-PCE in games with side payments, it is far from obvious in games without side payments. We used bilinear programming to show that in two-player games without side payments, both a PCE and an M-PCE can be computed in polynomial time. We also show that bilinear programming for a class of 2×2 matrices is solvable in constant time.

6. REFERENCES

[1] F. A. Al-Khayyal and J. E. Falk. Jointly constrained biconvex programming. *Mathematics of Operations Research*, 9:273–286, 1983.

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

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

[4] M. Chwe. Farsighted coalitional stability. *Journal of Economic Theory*, 63:299–325, 1994.

[5] E. Diamantoudi and L. Xue. Farsighted stability in hedonic games. *Social Choice and Welfare*, 21:39–61, 2003.

[6] J. Greenberg. *The theory of social situations: an alternative game-theoretic approach*. Cambridge University Press, 1990.

[7] J. Y. Halpern and R. Pass. Iterated regret minimization: a new solution concept. *Games and Economic Behavior*, 74(1):194–207, 2012.

[8] J. Y. Halpern and N. Rong. Cooperative equilibrium. In *Proc. 9th International Conference on AAMAS: volume 1*, pages 1465–1466, 2010.

[9] J. C. Harsanyi. An equilibrium-point interpretation of stable sets and a proposed alternative definition. *Management Science*, 20:1472–1495, 1974.

[10] W. Jamroga and M. Melissen. Doubtful deviations and farsighted play. In *15th Portuguese Conference on Artificial Intelligence (EPIA2011, LNAI7026)*, pages 506–520, 2011.

[11] A. Kalai and E. Kalai. Engineering cooperation in two-player games. 2009.

[12] N. Nakanishi. Purely noncooperative farsighted stable set in an n-player prisoners dilemma. *Technical Report*, 707, 2007.

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

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

[15] H. D. Sherali and A. Alameddine. A new reformulation-linearization technique for bilinear programming problems. *Journal of Global Optimization*, 2:379–410, 1992.

[16] A. Suzuki and S. Muto. Farsighted stability in an n-person prisoners dilemma. *International Journal of Game Theory*, 33:431–445, 2005.

[17] J. M. von Neumann. Zur theorie der gesellschaftsspiele (on the theory of games of strategy). *Mathematische Annalen*, 100:295–320, 1928.

[18] J. M. von Neumann. *Theory of games and economic behaviour*. Princeton University Press, 1944.