Mastering multi-player games

Yossi Azar[*]
Blavatnik School of Computer Science
Tel-Aviv University
Tel-Aviv, Israel
azar@post.tau.ac.il

Uriel Feige[†]
Department of Computer Science and Applied Mathematics
the Weizmann Institute
Rehovot 76100, Israel
uriel.feige@weizmann.ac.il

Michal Feldman[‡]
Hebrew University
and Harvard University
Jerusalem, 91905, Israel
mfeldman@huji.ac.il

Moshe Tennenholtz
Microsoft Research and Technion
13 Shenkar street
Herzliyah, Israel
moshet@microsoft.com

ABSTRACT

We consider multi-player games, and the guarantees that a master player that plays on behalf of a set of players can offer them, without making any assumptions on the rationality of the other players. Our model consists of an $(n+1)$-player game, with m strategies per player, in which a *master* player M forms a coalition with nontransferable utilities among n players, and the remaining player is called the *independent* player. Existentially, it is shown that every game admits a *product-minimax-safe* strategy for M — a strategy that guarantees for every player in M's coalition an expected value of at least her *product minimax value* (which is at least as high as her minimax value and is often higher). Algorithmically, for any given vector of values for the players, one can decide in polytime whether it can be ensured by M, and if so, compute a mixed strategy that guarantees it. In symmetric games, a product minimax strategy for M can be computed efficiently, even without being given the safety vector. We also consider the performance guarantees that M can offer his players in repeated settings. Our main result here is the extension of the oblivious setting of Feldman, Kalai and Tennenholtz [ICS 2010], showing that in every symmetric game, a master player who never observes a single payoff can guarantee for each of its players a *similar* performance to that of the independent player, even if the latter gets to choose the payoff matrix after the fact.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multiagent Systems*; J.4 [**Computer Applications**]: Social and Behavioral Sciences—*Economics*

General Terms

Algorithms, Economics, Theory

Keywords

Game Theory, minmax, repeated games

1. INTRODUCTION

In many situations, a *master* is playing on behalf of multiple players, and wishes to offer them some guarantees. For example, a representative plaintiff in a class action represents a group of complainants, and wishes to guarantee the individual complainants a certain level of compensation. In many of these settings, the payoffs of the players under the master's control are nontransferable, due to the nature of the good or some legal or ethical concerns. Presumable, every individual player has some safety-level payoff, which is the payoff she can guarantee for herself, when playing independently. An individual player will typically not delegate her action to a third party unless he can guarantee for her a value of at least the safety level she can guarantee for herself. The focus of this paper is the performance guarantees that a master can provide to the members of the coalition he forms. This involves two aspects. One is *existential*, studying which guarantees are possible. The other is *algorithmic*, presenting efficient algorithms that find strategies for M.

As is well known [1], if the master can play on behalf of all the players in the game, he can offer them a correlated equilibrium. However, in this manuscript the master plays only on behalf of some of the players, and not all. Hence he forms a coalition among the players that he controls. The strength of coalitions has been studied in the past [2, 8], but mostly

[*]Supported in part by the Israel Science Foundation (grant No. 1404/10).

[†]This research was conducted while at Microsoft Research at Herzliyah.

[‡]This research was partially conducted while at Microsoft Research at Herzliyah. This work was partially supported by the Israel Science Foundation (grant number 1219/09), by the Leon Recanati Fund of the Jerusalem School of Business Administration, and by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement number 274919.

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), June, 4–8, 2012, Valencia, Spain.

in the context of incentives they may have to deviate from a Nash equilibrium. In this manuscript we assume nothing about what other players play, and ask for guarantees to the coalition players regardless of what non-coalition players do. As we assume nothing about the strategies played by non-coalition players, we group them into one player.

To study this problem, we use the following model. Let G be an arbitrary $(n+1)$-player game with m strategies per player. Consider a situation in which a *master* player M controls players $P_1, \dots P_n$, and the remaining player P_0 is called the *independent* player I. This induces a new two player game between the master player and the independent player. This new game is referred to as the *Master Game* of G, and is denoted G_2. The payoffs for the independent player in G_2 are the same as in G. For the master player, there is a vector of payoffs, one for each player P_i. The goal of M is to provide performance guarantees to the players under his control. Crucially, throughout the paper, the notion of performance *guarantee* refers to the *expected* value that can be guaranteed.

What types of guarantees can a player expect to get? There are various safety notions that correspond to different guarantees a player can guarantee for herself when playing as an individual. Clearly, the master can guarantee each player the minimum possible payoff in the payoff matrix for that player. But in general, can the master offer better guarantees? Two natural candidates that do not work are the following:

1. For every $i \in [n]$, guarantee player P_i the minimum payoff for P_i in any Nash equilibrium. This does not work because I is not forced to play an equilibrium strategy. We wish our guarantees to hold even if I plays arbitrarily.

2. For an $(n+1)$ player game G and a choice x of $n-1$ pure strategies, let G_x^i be the marginal game that remains between player P_i and P_0 when the remaining players play the strategies as in x, and the payoff for P_0 is the negative of the payoff for P_i. The suggested guarantee is to guarantee every player P_i the minimum over x of the value of the games G_x^i. To see that this cannot be guaranteed, consider for example a three player game in which P_1 and P_2 are playing matching pennies with each other with payoffs ± 1, and P_0 only observes (is irrelevant to the payoffs). Hence given what P_2 plays (this is x for P_1), P_1 can get a payoff of 1 ("against P_0") and vice versa, but clearly there is no strategy for M controlling P_1 and P_2 that ensures expected payoff of at least 1 for both players simultaneously.

Additional natural safety notions for a player P_i are the *minimax* and *maximin* values. The minimax value is the expected value that P_i can guarantee if the other players announce a (possibly correlated) strategy first and P_i responds, and the *maximin* value is the expected value that P_i can guarantee if she announces a strategy first and the other players respond. It is also interesting to consider the *product minimax* and *product maximin* values, which are the respective versions of minimax and maximin, where the other players are restricted to play a product mixed strategy.

One may observe that the definitions of maximin value and product maximin value are in fact identical. Once a player announces her mixed strategy, the most harmful response is a pure strategy on behalf of each of the remaining players, and hence is captured by a product strategy. The minimax value and maximin value are also identical, by the minimax theorem for two player games (here the player under consideration is one player and a coalition of all remaining players serves as the other player). Hence the maximin, minimax and product maximin notions are equivalent to each other. However, the product minimax value might be different; it is at least as high, and often higher.

In simultaneous play, a player may guarantee for herself the minimax value (which equals the maximin value and the product maximin value). However, a player cannot guarantee the product minimax value, at least not without making assumptions on the rationality of other players. We ask whether a master who plays on behalf of n players can guarantee for each one of them her product minimax value.

Our existential result asserts that for every game G, there exists a mixed strategy for M in G_2 that guarantees for every player P_i under his control her product minimax value. Such a strategy for M is called a *product-minimax-safe* strategy. Moreover, we show that given any vector $\bar{v} = (v_1, \dots, v_n)$, one can find in polynomial time whether there exists a mixed strategy for M that guarantees for every player P_i a value of v_i, and if so, compute it in polynomial time. Consequently, given a vector of the players' product minimax values, a product-minimax-safe strategy can be computed in polynomial time. One should, however, not misinterpret the last result to suggest that the master can always compute a product-minimax-safe strategy in polynomial time. Indeed, it is shown in [5] that computing the product minimax value of a player is NP-hard, even for symmetric games. This makes it difficult for M to deduce what are the product minimax values that he needs to attain for his players. It remains open whether a product minimax safe strategy can be found without the safety vector being given.

Much of the intuition of what can and cannot be achieved is given through the analysis of *symmetric games*. As a special case of our existential result, it follows that in every symmetric game, M can guarantee each one of his players an expected payoff that equals at least that of I. Moreover, in the symmetric case, the master can also compute in polynomial time (via a linear program) a strategy that guarantees each one of his players at least her product minimax value. Interestingly, there exist symmetric games in which M can guarantee each one of his players an expected payoff that is strictly greater than the payoff of I. The last result is particularly significant in competitive settings, where players care less about their absolute payoff, rather they wish to perform well relative to others. For example, when a group of potential employees compete over a limited set of positions, they are mostly concerned with their ranking within the group.

In addition to the *vector version* (where the payoff of M is a vector of his players' payoffs), we consider the *minimum version*, where the payoff for M is the minimum of the payoffs to all his players. This version models situations that exhibit the "weakest-link" characteristic — where the weakest player determines the outcome for all involved parties. A classical example, given by [9], is the level of flood protection of an island, which is determined by the lowest dike along the coastline. Other prominent examples include airport security, the spread of infectious diseases, and more. Unfortunately, the guarantees that can be offered in the vector version do not extend to the minimum version. Indeed, there exist symmetric games in which for every mixed strategy of M, there exists a player P_i whose expected payoff is strictly lower than I's payoff.

In the second part of the paper we extend the model to repeated settings under various information structures. In all settings, a two-player repeated game between M and I is considered, where the payoff for M is a vector of his players' payoffs, and I is assumed to have full information of the payoff matrix and the history of play. The various settings differ by the information available to M, and the nature of the desired safety vector depends on the specific setting. We say that M approaches a safety vector $\bar{v} = (v_1, \ldots, v_n)$ if regardless of the strategy of I, the expected sum of payoffs over T rounds, for every player P_i controlled by M, is $Tv_i - o(T)$.

We first consider a full information setting. In a 2-player game, where the payoff of player 1 is a vector, Blackwell's approachability theorem [4] implies that a vector is *vector-minimax achievable* (in a one-shot game) if and only if it is also approachable (in a repeated setting). This implication can be viewed as a generalization of the minimax theorem for vector payoffs. In our setting, the payoff of M is the vector of its players' payoffs; thus it is essentially a 2-players game between M and I, with M having a vector payoff. In the one-shot game, the payoff vectors that are *safe* for M are precisely those that are *vector-maximin achievable*. For example, by our existential result for the one-shot game, the class of vectors that consist of respective product-minimax values of the players are vector-maximin achievable in the one-shot game. Vectors that are minimax achievable in the one shot game are not necessarily safe for M in the one shot game. Nevertheless, based on the implication of Blackwell's theorem stated above, these vectors are approachable in the repeated settings. Thus, the set of safe vectors for M in the repeated setting is larger than the set of safe vectors in the one-shot game.

We also consider an *oblivious* version, in which payoffs are never observed by M over the course of the game, extending the setting of [7] from two players to multiple players. This version models situations where players are engaged in games but the payoffs are determined only in later stages. This is the situtaion, for example, in cases where the actual behavior of players is evaluated by other parties only in retrospect or when the actual payoffs are determined based on an event whose outcome is unknown in the time actions are taken. A motivating example would be a reporter who writes a daily column about political issues. In some countries reporters may be recruited after the elections to represent government officials. Needless to say that the content of the daily columns by the different reporters may affect this selection. However, only after the government is selected the reporters will be able to determine their payoffs for their expressed opinions.

If players have similar capabilities and preferences, then these games are also symmetric. In the latter case an attractive objective for a player is to guarantee a payoff which is close to the one of his competitor despite the a-prior lack of information; indeed, for the case of two players this problem has been addressed by [7]. More generally, however, we ask whether a group of players can guarantee each one of them a payoff close to the one of an independent player, even if that player is all capable and able to predict the actual payoff function.

To model such situations, we consider a repeated game between M and I, where the game is first played T times, and after all T repetitions are completed I gets to choose the payoff matrix. The notion of safety for M that we use here is that as T tends to infinity, the regret is $o(T)$. We show that for every symmetric game G, M is safe in the oblivious repeated version. Moreover, a safe strategy for M can be found in time polynomial in m^n and T.

Related work.

Our work relates to the body of work on mediators in AI (see e.g. [10, 12]), which extends on classical literature on non-transferable utility coalition games, originated in [3], which relies on the fundamental work introduced in [2] (leading to the concepts of β-core and c-acceptable strategies, respectively). In particular, given a non-cooperative game, a c-acceptable strategy can be viewed as a recommended strategy profile to the group of all agents, augmented with a punishing correlated strategy for each strict subset of the agents. The idea is that the recommended strategy will be accepted if there is no subset of the players that can all gain with respect to this strategy payoffs, using a correlated strategy without side-payments when playing against the punishing correlated strategy of the remaining agents. A subset of deviating agents can be viewed as a master player, and the question is whether this master player can guarantee a vector of payoffs that exceeds the corresponding vector of payoffs of the recommended strategy profile. Some of our results complement this literature, by showing that if the recommended strategy payoffs are dominated by the product minimax outcomes associated with a particular master-player then it can not be accepted. We are not aware of a similar result.

2. SOME SAFETY NOTIONS

Given a multiplayer game G with n players, let us review some safety notions that a player may wish to achieve, and in some cases may indeed achieve. We use the following notation. S_i denotes a mixed strategy for player P_i. S^{-i} denotes a mixed strategy for all players except for P_i, when we do not assume that this is a product strategy. That is, all remaining players may collaborate to agree on a joint correlated mixed strategy. When the players each individually plays a mixed strategy, then $S^{-i\times}$ denotes the resulting product strategy (excluding player P_i). We use $v_i(S_i, S^{-i})$ to denote the expected payoff of player P_i when it plays the mixed strategy S_i and other players play the mixed strategy S^{-i}.

We consider here four different versions of safety notions, three of which are equivalent to each other.

- Minimax value (a best response notion). The other players announce their most harmful mixed strategy (min), and in response P_i chooses his best strategy (max). $\text{minimax}_i(G) = \max_{S_i} \min_{S^{-i}} v_i(S_i, S^{-i})$. Without loss of generality, S_i can be a pure strategy here.

- Maximin value (a best bid notion). P_i announces a mixed strategy that is deemed best (max), and in response the other players choose their most harmful strategy (min). $\text{maximin}_i(G) =$

$\min_{S^{-i}} \max_{S_i} v_i(S_i, S^{-i})$. Without loss of generality, S^{-i} can be a pure strategy here.

- Product minimax value. The other players announce their most harmful product mixed strategy, and in response P_i chooses his best strategy. $\text{minimax}_i^{\times}(G) = \max_{S_i} \min_{S^{-i\times}} v_i(S_i, S^{-i\times})$. Without loss of generality, S_i can be a pure strategy here.

- Product maximin value. P_i announces a mixed strategy that is deemed best, and in response the other players choose their most harmful product strategy. $\text{maximin}_i^{\times}(G) = \min_{S^{-i\times}} \max_{S_i} v_i(S_i, S^{-i\times})$. Without loss of generality, S^{-i} can be a pure strategy here.

As mentioned in the Introduction, the minimax, maximin and product maximin values are identical, whereas the minimax value is at least as high and may be higher. The superiority of the product minimax value is demonstrated in the following example.

Example. Suppose that P_1 has four pure strategies whereas P_2 and P_3 each has two pure strategies, and hence four combinations of pairs of pure strategies. If P_1 plays strategy 1 against combination 1 its payoff is 2. For $j > 1$, if P_1 plays strategy j against combination j its payoff is 1. In all other cases the payoff of P_1 is 0. The minimax value for P_1 in this game is 2/7. The unique S^{-1} that limits P_1 to this value is the mixed strategy in which the other players play combination 1 with probability 1/7, and each other combination with probability 2/7. But this S^{-1} cannot be represented as a product strategy, and hence the product minimax value is strictly larger than 2/7.

In simultaneous play, a player may guarantee for herself the minimax value, but not the product minimax value. In this context, it is interesting to note that in every (mixed) Nash equilibrium for the multiplayer game G, every player gets at least his product minimax value. A player not getting this value (in expectation) necessarily can gain by deviating from his current mixed strategy, contradicting the assumptions of a Nash equilibrium.

We note also that there are multiplayer games for which in every Nash equilibrium, all players get expected payoff strictly larger than their respective product minimax values. For example, consider an n-player game in which each player has two strategies, and the payoffs to all players is the number of players playing strategy 1. This game has a unique Nash equilibrium, with value n to all players, but the product minimax value is only 1.

2.1 Computational complexity of safety notions

The minimax (and hence maximin and product maximin) value for a player P_i in a multiplayer game can be computed in polynomial time (time polynomial in the game matrix). This is done by viewing the setting as a 0-sum two player game, where P_i is one player, and a coalition of all players is the other player.

When the number of players is three or more, computing the product minimax value of a game is NP-hard. See [5], where the vector of product minimax values is called the *threat point*. This also holds when the 3-player game is symmetric.

3. SAFETY GUARANTEES IN ONE-SHOT GAMES

We consider multi-player games, and the guarantees that a master player that plays on behalf of a set of players can offer the players. Let G be an arbitrary $(n+1)$-player game with m strategies per player. Consider a situation in which a *master* player M controls players $P_1, \dots P_n$, and the remaining player P_0 is called the *independent* player I. Having only one independent player can be assumed without loss of generality: as we assume nothing about the strategies played by non-coalition players, we may as well view them as one player. (See Section 3.3 for a brief discussion on this point.) This induces a new two player game G_2 between the master player and the independent player. The payoffs for the independent player in G_2 are the same as in G. For the master player, there is a vector of payoffs, one for each player P_i. We use $[n]$ to denote the set $\{1, \dots, n\}$. We use the notion of *guarantee* to refer to guarantees on the *expected* payoff.

3.1 Algorithmic results

The $\bar{v}$-safe problem is defined as follows.

$\bar{v}$-safe problem. Given a multi-player game G and a vector of values $\bar{v} = \{v_1, \dots, v_n\}$, is there a mixed strategy for M that guarantees for every player P_i an expected payoff of at least v_i, and if so, present such a mixed strategy.

A key concept that we shall use here is the following.

Selection version. Given a multi-player game G we introduce a *selection version* G'. In G' players $P_1, \dots, P_n$ have the same set of strategies as in G. The set of strategies for I includes two coordinates, one being the set of strategies for P_0 and the other a choice of an index in $[n]$. Given the strategies played by the players, player P_i (where i is the value of the second coordinate chosen by I) gets the same payoff as in G, the payoff for I is the negative of the payoff for P_i, and all other players get payoff 0. The master version G'_2 of G' (where M controls $P_1, \dots, P_n$) is a 0-sum game, if the payoff for M is taken to be as the sum of payoffs of its players. This implies that linear programming can be used in order to find optimal strategies in G'_2 in time polynomial in the size of the normal form representation of G'_2, namely, polynomial in $(n+1)m^{n+1}$.

The concept of a selection version is used to show that the $\bar{v}$-safe problem if polynomial.

THEOREM 1. *There is a polynomial time algorithm for the $\bar{v}$-safe problem.*

PROOF. Given a game G and a vector $\bar{v}$, scale the payoff for each player P_i (for $i \in [n]$) by subtracting v_i, so that the $\bar{v}$-safe problem now asks for nonnegative expected payoff for every player controlled by M in G_2. We claim that this is equivalent to finding a strategy for M with nonnegative expected payoff in the corresponding selection game G'_2. If such a strategy exists for G_2, then it can be used in G'_2 and I is helpless in choosing $i \in [n]$. If such a strategy exists in G'_2, then the same strategy can be used in G_2, and every player controlled by M has expected nonnegative payoff, as otherwise there would be a strategy for I in G'_2 giving I expected positive payoff. Now polynomiality of G_2 follows from polynomiality of G'_2. □

3.2 Existential results

The product minimax value of a player is always at least as high as her minimax value (and hence also the maximin and product maximin values), and is often higher (see Section 2). While a player cannot guarantee this value for herself when playing independently, we show that a master who plays on behalf of a set of players can guarantee each one of them her product minimax value.

Definition 2. *We say that a mixed strategy for M is* product-minimax-safe *for G_2 if for each player P_i controlled by M it guarantees an expected payoff of at least her product minimax value. We say that G_2 is product-minimax-safe for M if M has a product-minimax-safe mixed strategy for G_2.*

Theorem 3. *For every multi-player game G, the corresponding game G_2 is product-minimax-safe for M. Moreover, given the players' product minimax values, a product-minimax-safe strategy for M can be found in time polynomial in m^n.*

Note that the algorithmic content of Theorem 3 is implied by Theorem 1, and hence the main new content is in the existential statement. Before proving Theorem 3, we consider the special case of symmetric games which gives much of the intuition of what can and cannot be achieved.

3.3 Symmetric games

A multi-player game is *symmetric* if the payoff matrix of each player is symmetric (namely, permuting the identities of the remaining players keeps the payoff matrix unchanged), and moreover, the payoff matrices of any two players are identical to each other.

In this section G is a symmetric game, or more accurately, the *difference version* $\hat{G}$ of a symmetric game G. The payoff for P_i in $\hat{G}$ is computed as the payoff in G to P_i minus the payoff in G to P_0. Based on the symmetry of G and the 0-sum nature of $\hat{G}$, it follows that the product minimax value of all the players is 0. Thus, product-minimax-safety in $\hat{G}$ means obtaining a nonnegative expected payoff for every player.

The following proposition is a special case of Theorem 3.

Proposition 4. *For every symmetric game G, $\hat{G}_2$ is product-minimax-safe for M.*

Proof. For every entry in the payoff matrix G, add a constant to all payoffs to make the game 0-sum. It remains symmetric. Every symmetric game has a symmetric equilibrium [11]. Let M play this symmetric equilibrium. Then for I to play it as well is a mixed Nash, and the expected payoff for each player is 0, by symmetry. If I deviates from the mixed Nash, his expected payoff cannot increase, and hence the expected sum of payoffs of other player cannot decrease. By symmetry, no player controlled by M can then have negative expected payoff. This implies that the same strategy for M in $\hat{G}_2$ is product-minimax-safe. □

Let us end this section with a few comments on what happens when G is symmetric, but I represents several players rather than just one. For concreteness, suppose that I represents two players (but the same principles can be generalized to a larger number of players).

If n, the number of players controlled by M, is divisible by 2, then the results extend in the following sense. M can partition his players into pairs, resulting in a new symmetric game with half the players (where each pair of players in G is a single player in the new game). Hence all results transfer, but with guarantees to pairs of players rather than to individual players. This can further be changed to a guarantee for an individual player by randomly permuting the order of players in a pair, and hence guaranteeing for each player the average of the guarantee for the pair. (In symmetric games, coalitions with nontransferable utilities can implement a transfer of utility to achieve fairness within coalition, by randomizing over the members).

If n is not divisible by 2, the guarantees become weaker. (Theorem 3 is still true as stated, but does not form an incentive for players to join the coalition formed by M.) Consider for example a 5-player symmetric game in which each player announces a number in $[m]$ (where m is large), and a player gets payoff of 1 if exactly one other player announced the same number as he did, and a payoff of 0 otherwise. In this game, a coalition M of three players cannot guarantee to its members an average expected payoff equal to the average of the remaining two players (that if controlled by I can coordinate to both report the same number in $[m]$, chosen at random).

3.4 Proof of safety for arbitrary games

Here we prove the existential part of Theorem 3. (The algorithmic part then follows from Theorem 1.)

Proof. Let v_i be the product minimax value of player P_i. As in the proof of Theorem 1, scale the payoff for each player P_i (for $i \in [n]$) by subtracting v_i, and let $\tilde{G}$ denote the obtained game. Showing that G_2 is product-minimax-safe for M is equivalent to showing that M has a strategy for $\tilde{G}_2$ that guarantees for each player P_i a nonnegative expected value.

Consider an arbitrary mixed Nash for $\tilde{G}'$ (the selection version of $\tilde{G}$). We claim that in this mixed Nash the expected payoff for every player P_i (for $i \in [n]$) is nonnegative. Equivalently, in this mixed Nash the expected payoff of every player P_i in G' is at least v_i. Assume otherwise, that there exists a player P_i ($i \in [n]$), whose expected payoff in G' is lower than v_i. Let x denote the mixed strategies for P_j ($j \in [n]$, $j \neq i$) as in the mixed Nash. Then P_i is *almost* playing the marginal game that remains between herself and P_0 when the remaining players play according to x, which has value at least v_i. The only difference is that P_i gets payoff only if P_0 chooses i. However, the choice of i is independent of the choice of x, and the strategy of P_0 on its first coordinate conditioned on having i on the second coordinate is a strategy for P_0 in the original game G. Hence the lower bound guarantee on expected payoff in G is transferred also to G' (otherwise P_i would not be in a Nash equilibrium in G' – he will have an incentive to change his mixed strategy).

We next claim that this implies that regardless of the strategy played by P_0, the expected sum of payoffs for P_i ($i \in [n]$) when playing according to this mixed Nash is nonnegative. Assume otherwise, that P_0 has a strategy for which the expected sum of payoffs for P_i (when playing according to the mixed Nash equilibrium) is negative. Then, the payoff for P_0 would be positive, which means that P_0 can improve its expected payoff from the mixed Nash equilibrium, in contradiction. This means that in the 0-sum game $\tilde{G}'_2$ (which is similar to $\tilde{G}'$, except that M controls players P_i for $i \in [n]$ and gets as payoff their sum of payoffs from $\tilde{G}'$), the strategy given above ensures M a nonnegative expected payoff.

Finally, we claim that using the same strategy in $\tilde{G}_2$ ensures a nonnegative payoff for every player P_i controlled by M. If this was not the case, say for a particular P_i, then I would have a strategy in $\tilde{G}'_2$ picking this particular i and achieving a positive payoff, implying a negative payoff for M in $\tilde{G}'_2$, reaching a contradiction.

□

3.5 A note on complexity

In [5] it is shown that computing the threat value (which is precisely the product minimax value) is NP-hard, even for symmetric games. This makes it difficult for M to deduce what are the product minimax values that it needs to attain for the players that it represents. Nevertheless, if these values are given to M as input, it can attain it via a polynomial time strategy, by Theorem 1.

Let us point out that for symmetric games, neither Proposition 4 nor its proof ensures for M that its vector of expected payoffs attains or exceeds the product minimax safety vector. Consider a 3-player symmetric game with two strategies, "cooperate" and "defect". If all players cooperate then all payoffs are 3. If one player defects his payoff is 0 and other payoffs are 1. If two or more players defect all payoffs are 2. The product minimax value for each player is strictly above 1. (Here is a simple way of achieving such a value. If one other player plays the pure strategy defect and the other plays the pure strategy cooperate, then the given player plays defect. In any other case, the given player cooperates.) However, the proof of Proposition 4 may (though also may not, as it has nondeterministic components) produce for M the strategy of always cooperating, which might (if I defects) lead to a payoff of only 1 to each of its players.

Nevertheless, despite the fact that it is NP-hard to compute the product minimax value of the players even in symmetric games, for the case of symmetric games, we have the following algorithmic results.

Proposition 5. *There is a polytime strategy for M in symmetric games that ensures at least the product minimax value to each of its players.*

Proof. By symmetry, the product minimax value for each player is the same. We denote this value by v. Consider now the 0-sum two player game between M and I is which M attempts to maximize the expected sum of payoffs for its players, and I tries to minimize it. The value of this game for M is at least nv (since we have shown that existentially M can guarantee each of its players an expected payoff of v), and perhaps larger. Let $V \geq nv$ denote this value. A mixed strategy for M that achieves value V can be computed in polynomial time (as this is a 0-sum game). Thereafter, this strategy can be further randomized by permuting the players controlled by M at random, ensuring (by symmetry) each one of the players expected payoff at least $V/n \geq v$. □

3.6 Additional safety notions in symmetric games

Recall that $\hat{G}$ denotes the difference version of a symmetric game G (where the payoff for P_i is computed as the payoff in G to P_i minus the payoff in G to P_0). The product minimax value of all the players in $\hat{G}$ is 0; thus, product-minimax-safety in $\hat{G}$ means obtaining a nonnegative expected payoff for every player. We can extend the notion of *product-minimax-safe*, and say that M *product-minimax-wins* a game if it can guarantee every player an expected value that is strictly greater than her product minimax value. Thus, M product-minimax-wins $\hat{G}_2$ if M has a mixed strategy that guarantees positive expected payoff for each one of its players.

Up to this point we considered the *vector version* of a master game, where the payoff of M is a vector of its players' payoffs. It will be instructive for us to consider also the *minimum version* of $\hat{G}_2$ in which the payoff for M is the minimum of the payoffs to all its players. (Note that expectation of minimum is never higher than minimum of expectation, and hence the minimum version is more difficult for M than the vector version.)

Proposition 6. *There are symmetric games G for which M product-minimax-wins the corresponding game $\hat{G}_2$ even in the minimum version, and there are symmetric games G for which M product-minimax-wins the corresponding game $\hat{G}_2$ in the vector version, but not in the minimum version.*

Proof. Consider G^+ with three players and two strategies (0 or 1) per player, where the payoff to a player is the number of other players playing his strategy. Then in the respective $\hat{G}_2^+$, M can play $(0, 0)$ with probability 1/2 and $(1, 1)$ with probability 1/2, achieving payoff 1/2 to each of its players. Here M achieves expected payoff of 1/2 even in the minimum version, thus product-minimax-wins even in the minimum version.

Alternatively, consider G^- where the payoff to a player is the number of other players not playing his strategy. Then in the respective $\hat{G}_2^-$, M can play $(0, 1)$ with probability 1/2 and $(1, 0)$ with probability 1/2. Here M product-minimax-wins in the vector version and the selection version, but not in the minimum version (where he can guarantee only 0). □

Some comments on Proposition 6 are in order. It shows that for some games (e.g., G^+) it is good for M to play identically for all its players, and for others (e.g., G^-) it is a mistake to do so. Also, observe that for G^+ and G^-, playing independently randomly and uniformly is product-minimax-safe for M (but not winning) in the vector version, but not product-minimax-safe in the minimum version.

The following proposition shows that the offers that can be guaranteed in the vector version (asserted in Proposition 4) do not extend to the minimum version.

Proposition 7. *There are games for which M is not product-minimax-safe in the minimum version of the corresponding $\hat{G}_2$.*

Proof. Let G be the following three player game (for large enough m). If two (or three) players play the same strategy they get payoff 0. For players that do not collide their payoff is computed as follows. Add the numbers (in $[m]$) reported by the players and consider the sum modulo 3. This determines which of the three player wins a payoff (the lowest report if the result is 0, second lowest if 1, highest if 2). Other players get no payoff. One of M's player cannot get payoff. Hence the minimum payoff for M is always 0. If I plays randomly (and m is large enough), it has probability roughly 1/3 to get a positive payoff. □

4. SAFETY GUARANTEES IN REPEATED GAMES

We consider here repeated games. We say that M approaches a safety vector $\bar{v}$ for its n players if regardless of the strategy of the independent player I, for every player i controlled by M, the expected sum of payoffs for player i over T rounds is $T\bar{v}_i - o(T)$. The nature of the safety vector $\bar{v}$ will be dependent on the context. We describe several settings of repeated games, and they differ by the information available to M.

4.1 Full information

Blackwell's approachability theorem [4] deals with approachability issues in general. One of its consequences is the following generalization of the minimax theorem to vector payoffs.

Consider a two player game between players Q_1 and Q_2 where the payoff for Q_1 is a vector. A vector $\bar{v}$ is *vector-minimax achievable* if for every mixed strategy of Q_2, player Q_1 has a mixed strategy with expected payoff at least $\bar{v}$. Vector $\bar{v}$ is *vector-maximin achievable* if Q_1 has a mixed strategy such that no matter what Q_2 plays, the expected payoff for Q_1 is at least $\bar{v}$. For vector payoffs the well known minimax theorem does not apply. That is, a vector that is vector-minimax achievable need not be vector-maximin achievable. However, Blackwell's approachability theorem implies that if a vector $\bar{v}$ is vector-minimax achievable, then it is also approachable (in a repeated game, where one considers the average payoff vector over all rounds). Moreover, this is an if and only if relation.

The above general result about games with vector payoffs is applicable in our setting, with M serving as Q_1, and I serving as Q_2. The master player M controls n players, and hence his payoff can be viewed as an n-dimensional vector (one coordinate for each player that M represents). In the one-shot game, the payoff vectors that are safe for M are precisely those that are vector-maximin achievable for Q_1. Under this interpretation, Theorem 3 identifies a particular class of vectors that are vector-maximin achievable, namely, those vectors that in each coordinate i have the product-minimax value for player P_i. Proposition 6 provides examples showing vectors that are strictly higher than the vector of product-minimax values, but nevertheless are vector-maximin achievable.

Blackwell's theorem implies that there are vectors that are not safe for M in the one shot game but are nevertheless approachable in the repeated games settings. Here is an explicit example that illustrates this. Consider a 3-player game with players P_0, P_1 and P_2. Only actions of P_0 and P_1 determine the payoffs of the game, and only P_1 and P_2 can receive any payoff. Each of the two players P_0 and P_1 has two possible actions, where one action gives P_1 a payoff of 1 and the other action gives P_2 a payoff of 1. The master player M controls P_1 and P_2, and the independent player I is P_0. The vector $\bar{v} = (1, 1)$ is not safe for M in the one shot game: whatever mixed strategy M has, there is one player (P_1 or P_2) for which P_1 gives expected payoff less than 1, and then I can give this player an additional payoff of 0, preventing it from achieving an expected payoff of 1. However, $\bar{v} = (1, 1)$ is approachable by M in the repeated setting. One could prove this fact using Blackwell's theorem, but for this simple game, one can see this directly: M plays arbitrarily in round 1, and in every round after that M plays (for P_1) the opposite of what I played in the previous round. Hence at every round t, either both P_1 and P_2 have accumulated a payoff of exactly t, or one of them accumulated a payoff of $t-1$ and the other $t+1$. As t grows, the average payoff vector converges to $(1, 1)$.

4.2 Only realized payoffs are observable

Here we consider a version in which the payoff matrices are not known to M. Moreover, they cannot be inferred from repeated play because the actions of I are not observable. All that M observes is the actual realized payoff vector after each round of play.

For symmetric games, a natural goal for M is to maximize the sum (over all n players) of payoffs, because randomizing over players spreads this sum evenly among them. This task of M in a repeated game can be cast as a so called *bandit problem*, for which solutions are known [6].

For nonsymmetric games the situation becomes more complicated, but M can still attain sublinear regret (for an appropriate choice of safety vector) using a combination of bandit algorithms and Blackwell's approachability theorem. This turns out to be a special case of a more general result (concerning expert and bandit algorithms in vector settings) that will be presented in a companion paper (in preparation), and hence will not be discussed further in the current paper.

4.3 Only actions are observable

Here we consider an *oblivious* version in which M does not know the payoffs, and never observes them, extending a setting of [7] from two players to multiple players. The game G is initially defined only in terms of sets of strategies available to the players. In the one round version the players M and I announce their strategies, and only after the game is played I announces a payoff matrix of its choice. In the repeated version the game is first played T times, and after all T repetitions are completed I gets to choose the payoff matrix. To give M hope of providing meaningful performance guarantees, we limit the choice of payoff matrix to be symmetric and the entries to be bounded independently of T (say by 1).

We shall use here the notion of *differential safety*; that is, as T tends to infinity, the difference between the expected average payoff of every player i over T rounds and that of I is $o(T)$.

THEOREM 8. *For every symmetric game G, M is differential safe in the oblivious repeated version. Moreover, a safe strategy for M can be found in time polynomial in m^n and T.*

PROOF. We combine the proof technique of [7] (that was previously used for $n = 1$) together with our Theorem 3.

We provide a strategy for the master player that guarantees for each player P_i under his control an expected average difference (over T periods) of

$$\mathbf{E}\left[\left|\frac{1}{T}\sum_{t=1}^{T} v_i^t\right|\right] \leq \sqrt{\frac{m}{T}} m^{n/2},$$

where v_i^t is the difference between the payoff of player P_i and player I in period t. Note that since the expectation is taken over the absolute value, this bounds actually serves as a lower bound on the payoff of player i.

In round t consider $nm^{(n-1)}$ auxiliary matrices $A^i_x[t]$, where index $i \in [n]$ specifies a player controlled by M, and index x specifies a choice of strategies for the other players controlled by M. Entry $A^i_x(j,k)[t]$ shows the difference in number of times P_i played k and I played j versus a play of (j,k) on periods $1,\ldots,t-1$, conditioned on x. Each matrix $A^i_x[t]$ is antisymmetric. A mixed strategy for M also implicity defines a distribution over x, and hence a distribution over auxiliary matrices. Define the "pretend" payoff matrix of P_i as the weighted average of the antisymmetric matrices $A^i_x[t]$, induced by the distribution over x. As a weighted average of antisymmetric matrices, the pretend payoff matrix is still antisymmetric. Being antisymmetric and 0-sum, P_i has a strategy for this payoff matrix with nonnegative expected payoff.

Using Theorem 3, M has a strategy that simultaneously gives expected nonnegative payoff to each of his players P_i in their respective pretend payoff matrices. Moreover, this strategy can be computed efficiently (as implied by Theorem 1). We claim that by playing such a strategy in every round (relative to the pretend payoff matrices of the respective round), M achieves the asserted bound for every player P_i.

Let $\phi_i = \sum_{t=1}^T v_i^t$. Then, $|\phi_i| \le \frac{1}{2}\sum_{x,j,k}\left|A^i_x(j,k)[T+1]\right|$. It holds that

$$\begin{aligned}(\mathbf{E}[|\phi_i|])^2 &\le \mathbf{E}[\phi_i^2]\\ &\le \mathbf{E}\left[\left(\sum_{x,j,k}\frac{1}{2}\left|A^i_x(j,k)[T+1]\right|\right)^2\right]\\ &\le \mathbf{E}\left[\frac{m^2m^{n-1}}{4}\sum_{x,j,k}\left(A^i_x(j,k)[T+1]\right)^2\right],\end{aligned}$$

where the last inequality follows by Cauchy-Schwartz.

Now let $\beta^i_t = \sum_{x,j,k}(A^i_x(j,k)[t])^2$. We claim that $\mathbf{E}\left[\beta^i_t\right] \le 2t$ for every t. To see this, let φ^i_t denote the expected "pretend" payoff of player P_i in period t, under the specified strategy of M. Simple algebra reveals that $\mathbf{E}\left[\beta^i_{t+1}-\beta^i_t\right] \le 2-2\varphi^i_t$. But since $\varphi^i_t \ge 0$, it follows that $\mathbf{E}\left[\beta^i_{t+1}-\beta^i_t\right] \le 2$. Combining the last inequality with $\beta^i_0 = 0$ implies that $\mathbf{E}\left[\beta^i_t\right] \le 2t$, as promised. It follows that $\mathbf{E}\left[|\phi_i|\right] \le \sqrt{mT}m^{n/2}$. Hence, the average number of deviations that I is expected to build against a single player over T periods is bounded by $\sqrt{\frac{m}{T}}m^{n/2}$. □

5. CONCLUSION

One of the major challenges in the "agent perspective" to multi-agent systems is to deal with guarantees an agent can obtain, without assuming rationality of the other agents. This work pushes the envelope of that fundamental attempt by showing several results on the effective use of a master agent (aka mediator). Our two major results are:

- Every game admits a product-minimax-safe strategy for M — a strategy that guarantees for every player in M's coalition an expected value of at least her product minimax value (which is at least as high as her minimax value and is often higher). A player cannot guarantee for herself this value, when playing independently.
- In repeated symmetric games a master player who never observes a single payoff can guarantee (in an asymptotic sense) for each of its players a similar performance to that of the independent player, even if the latter gets to choose the payoff matrix after the fact.

These results are augmented with corresponding algorithmic results. The results expand on early foundational work in game theory, and work on mediators in AI.

6. REFERENCES

[1] R. Aumann. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1:67–96, 1974.

[2] R. Aumann. Acceptable Points in General Cooperative n-Person Games. *Contribution to the Theory of Games, Vol. IV, Annals of Mathematics Studies, 40, editors A.W. Tucker and R.D. Luce*, 287–324, 1959.

[3] R. Aumann. The Core of cooperative Games without Payments. *Transactions of the American Mathematical Society*, 1961.

[4] D. Blackwell. An Analog of the Minimax Theorem for Vector Payoffs. *Pacific Journal of Mathematics*, 6: 1–8, 1956.

[5] C. Borgs, J. Chayes, N. Immorlica, A. Kalai, V. Mirrokni, and C. Papadimitriou. The Myth of the Folk Theorem, Games and Economic Behavior, 2009. Conference version appeared in STOC 2008.

[6] N. Cesa-Bianchi and G. Lugosi. Prediction, Learning, and Games. Cambridge University press, 2006.

[7] M. Feldman, A. Kalai and M. Tennenholtz. Playing games without observing payoffs. *Innovations in computer science*, 106–110, 2010.

[8] D.B. Gillies. Solutions to general non-zero-sum games. *In A. W. Tucker and R. D. Luce, editors, Contributions to the theory of games, volume IV, pp. 47–85.* Princeton University Press, 1959.

[9] J. Hirshleifer. From weakest-link to best-shot: The voluntary provision of public goods. Public Choice, Vol. 41, Number 3, 1983.

[10] D. Monderer and M. Tennenholtz. Strong mediated equilibrium. *Artificial Intelligence, Vol. 173, Number 1, pp. 180–195*, 2009.

[11] J. Nash. Non-cooperative Games. *The Annals of Mathematics*, 54(2):286–295, 1951.

[12] O. Rozenfeld and M. Tennenholtz. Routing Mediators. *Proceedings of IJCAI*, pp. 1488–1493, 2007.