A Generalized Strategy Eliminability Criterion and Computational Methods for Applying It*

Vincent Conitzer and **Tuomas Sandholm**
Carnegie Mellon University
Computer Science Department
5000 Forbes Ave.
Pittsburgh, PA, 15213
{conitzer, sandholm}@cs.cmu.edu

Abstract

We define a generalized strategy eliminability criterion for bimatrix games that considers whether a given strategy is eliminable relative to given dominator & eliminee subsets of the players' strategies. We show that this definition spans a spectrum of eliminability criteria from strict dominance (when the sets are as small as possible) to Nash equilibrium (when the sets are as large as possible). We show that checking whether a strategy is eliminable according to this criterion is coNP-complete (both when all the sets are as large as possible and when the dominator sets each have size 1). We then give an alternative definition of the eliminability criterion and show that it is equivalent using the Minimax Theorem. We show how this alternative definition can be translated into a mixed integer program of polynomial size with a number of (binary) integer variables equal to the sum of the sizes of the eliminee sets, implying that checking whether a strategy is eliminable according to the criterion can be done in polynomial time, given that the eliminee sets are small. Finally, we study using the criterion for iterated elimination of strategies.

Introduction

Solving general-sum games is a topic of growing interest in AI. To solve such games, the concept of (iterated) dominance is often too strong: it cannot eliminate enough strategies. But, if possible, we would like a stronger argument for eliminating a strategy than (mixed-strategy) Nash equilibrium. Hence, it is desirable to have eliminability criteria that are *between* these two concepts in strength. In this paper, we will introduce such a criterion. The criterion we introduce considers whether a given strategy is eliminable relative to given dominator & eliminee subsets of the players' strategies. The criterion spans an entire *spectrum* of strength between Nash equilibrium and strict dominance (in terms of which strategies it can eliminate), and in the extremes can be made to coincide with either of these two concepts, depending on how the dominator & eliminee sets are set. It can also be used for iterated elimination of strategies.

An important question to ask of any solution concept is how efficiently a solution can be *computed*. In particular, the question of how hard it is to compute a Nash equilibrium is still open and has been called one of the two most important concrete open questions on the boundary of P today (Papadimitriou 2001). (In contrast, *approximate* Nash equilibria can be found in quasi-polynomial time (Lipton, Markakis, & Mehta 2003). Also, Nash equilibria can be found in polynomial time for average-payoff repeated games (Littman & Stone 2003).) The best-known algorithm for finding a Nash equilibrium, the *Lemke-Howson* algorithm (Lemke & Howson 1964), has a worst-case exponential running time (Savani & von Stengel 2004), and methods based on exhaustively searching through the space of the mixed strategies' supports fare comparatively well for many classes of games (Porter, Nudelman, & Shoham 2004). It is known that finding Nash equilibria *with certain additional properties* (for example, the social-welfare maximizing Nash equilibrium) is NP-complete (Gilboa & Zemel 1989; Conitzer & Sandholm 2003). The computational complexity of dominance and iterated dominance has also been studied (Knuth, Papadimitriou, & Tsitsiklis 1988; Gilboa, Kalai, & Zemel 1993; Conitzer & Sandholm 2005). In this paper, we will study the computational complexity of applying the new eliminability criterion, and provide a mixed integer programming approach for it.

Throughout, we focus on two-player games only. The eliminability criterion itself can be generalized to more players, but the computational tools we introduce do not straightforwardly generalize to more players. Moreover, we focus only on normal-form games (rather than make use of structured representations of games (Kearns, Littman, & Singh 2001; Leyton-Brown & Tennenholtz 2003; Blum, Shelton, & Koller 2003; Gottlob, Greco, & Scarcello 2003)).

One of the benefits of the new criterion is that when a strategy cannot be eliminated by dominance (but it can be eliminated by the Nash equilibrium concept), the new criterion may provide a stronger argument than Nash equilibrium for eliminating the strategy, by using dominator & eliminee sets smaller than the entire strategy set. To get the strongest possible argument for eliminating a strategy, the dominator & eliminee sets should be chosen to be as small as possible while still having the strategy be eliminable relative to these sets.[1] Iterated elimination of strategies using the new crite-

*This material is based upon work supported by the National Science Foundation under ITR grants IIS-0121678 and IIS-0427858, and a Sloan Fellowship.

[1]There may be multiple minimal vectors of dominator & elim-

rion is also possible, and again, to get the strongest possible argument for eliminating a strategy, the sequence of eliminations leading up to it should use dominator & eliminee sets that are as small as possible.[2]

As another benefit, the algorithm that we provide for checking whether a strategy is eliminable according to the new criterion can also be used as a subroutine in the computation of Nash equilibria. Specifically, any strategy that is eliminable (even using iterated elimination) according to the criterion is guaranteed not to occur in any Nash equilibrium. Current state-of-the-art algorithms for computing Nash equilibria already use a subroutine that eliminates (conditionally) dominated strategies (Porter, Nudelman, & Shoham 2004). Because the new criterion can eliminate more strategies than dominance, the algorithm we provide may speed up the computation of Nash equilibria. (For purposes of speed, it is probably desirable to only apply special cases of the criterion that can be computed fast—in particular, as we will show, eliminability according to the criterion can be computed fast when the eliminee sets are small. Even these special cases are more powerful than dominance.)

A motivating example

Because the definition of the new eliminability criterion is complex, we will first illustrate it with an example. Consider the following (partially specified) game.

	σ_c^1	σ_c^2	σ_c^3	σ_c^4
σ_r^1	?, ?	?, 2	?, 0	?, 0
σ_r^2	2, ?	2, 2	2, 0	2, 0
σ_r^3	0, ?	0, 2	3, 0	0, 3
σ_r^4	0, ?	0, 2	0, 3	3, 0

A quick look at this game reveals that strategies σ_r^3 and σ_r^4 are both *almost* dominated by σ_r^2—but they perform better than σ_r^2 against σ_c^3 and σ_c^4, respectively. Similarly, strategies σ_c^3 and σ_c^4 are both almost dominated by σ_c^2—but they perform better than σ_c^2 against σ_r^4 and σ_r^3, respectively. So we are unable to eliminate any strategies using (even weak) dominance.

Now consider the following reasoning. In order for it to be worthwhile for the row player to ever play σ_r^3 rather than σ_r^2, the column player should play σ_c^3 at least $\frac{2}{3}$ of the time. (If it is exactly $\frac{2}{3}$, then switching from σ_r^2 to σ_r^3 will cost the row player 2 exactly $\frac{1}{3}$ of the time, but the row player will gain 1 exactly $\frac{2}{3}$ of the time, so the expected benefit is 0.) But, similarly, in order for it to be worthwhile for the column player to ever play σ_c^3, the row player should play σ_r^4 at least $\frac{2}{3}$ of the time. But again, in order for it to be worthwhile for the row player to ever play σ_r^4, the column player should play σ_c^4 at least $\frac{2}{3}$ of the time. Thus, if both the row and the column player accurately assess the probabilities that the other places on these strategies, and their strategies are rational with respect to these assessments (as would be the case in a Nash equilibrium), then, if the row player puts positive probability on σ_r^3, by the previous reasoning, the column player should be playing σ_c^3 at least $\frac{2}{3}$ of the time, and σ_c^4 at least $\frac{2}{3}$ of the time. Of course, this is impossible; so, in a sense, the row player should not play σ_r^3.

It may appear that all we have shown is that σ_r^3 is not played in any Nash equilibrium. But, to some extent, our argument for not playing σ_r^3 did not make use of the full elimination power of the Nash equilibrium concept. Most notably, we only reasoned about a small part of the game: we never mentioned strategies σ_r^1 and σ_c^1, and we did not even specify most of the utilities for these strategies. (It is easy to extend this example so that the argument only uses an arbitrarily small fraction of the strategies and of the utilities in the matrix, for instance by adding many copies of σ_r^1 and σ_c^1.) The locality of the reasoning that we did is more akin to the notion of dominance, which is perhaps the extreme case of local reasoning about eliminability—only two strategies are mentioned in it. So, in this sense, the argument for eliminating σ_r^3 is somewhere between dominance and Nash equilibrium in strength.

Definition of the eliminability criterion

We are now ready to give the formal definition of the generalized eliminability criterion. To make the definition a bit simpler, we define its negation—when a strategy is *not* eliminable relative to certain sets of strategies. Also, we only define when one of the *row player's* strategies is eliminable, but of course the definition is analogous for the column player.

The definition, which considers when a strategy e_r^* is eliminable relative to subsets D_r, E_r of the row player's pure strategies (with $e_r^* \in E_r$) and subsets D_c, E_c of the column player's pure strategies, can be stated informally as follows. To protect e_r^* from elimination, we should be able to specify the probabilities that the players' mixed strategies place on the E_i sets in such a way that 1) e_r^* receives nonzero probability, and 2) for every pure strategy e_i that receives nonzero probability, for every mixed strategy d_i using only strategies in D_i, it is conceivable that player $-i$'s mixed strategy[3] is completed so that e_i is no worse than d_i.[4] The formal definition follows.

Definition 1 *Given a two-player game in normal form, subsets D_r, E_r of the row player's pure strategies Σ_r, subsets*

inee sets relative to which the strategy is eliminable; in this paper, we will not attempt to settle which of these minimal vectors, if any, constitutes the most powerful argument for eliminating the strategy.

[2]Here, there may also be a tradeoff with the length of the elimination path. For example, there may be a path of several eliminations using dominator & eliminee sets that are small, as well as a single elimination using dominator & eliminee sets that are large, both of which eliminate a given strategy. (In fact, we will *always* be confronted with this situation, as Corollary 3 will show.) Again, in this paper, we will not attempt to settle which argument for eliminating the strategy is stronger.

[3]As is common in the game theory literature, $-i$ denotes "the player other than i."

[4]This description may sound similar to the concept of *rationalizability*. However, in two-player games (the subject of this paper), rationalizability is known to coincide with iterated strict dominance (Pearce 1984).

D_c, E_c *of the column player's pure strategies* Σ_c*, and a distinguished strategy* $e_r^* \in E_r$*, we say that* e_r^* *is* not eliminable relative to D_r, E_r, D_c, E_c*, if there exist functions (partial mixed strategies)* $p_r : E_r \rightarrow [0,1]$ *and* $p_c : E_c \rightarrow [0,1]$ *with* $p_r(e_r^*) > 0$, $\sum_{e_r \in E_r} p_r(e_r) \leq 1$, *and* $\sum_{e_c \in E_c} p_c(e_c) \leq 1$, *such that the following holds. For both* $i \in \{r, c\}$*, for any* $e_i \in E_i$ *with* $p_i(e_i) > 0$*, for any mixed strategy* d_i *placing positive probability only on strategies in* D_i*, there is some pure strategy* $\sigma_{-i} \in \Sigma_{-i} - E_{-i}$ *such that (letting* $p_{-i} \diamond \sigma_{-i}$ *denote the mixed strategy that results from placing the remaining probability* $1 - \sum_{e_{-i} \in E_{-i}} p_{-i}(e_{-i})$ *that is not used by the partial mixed strategy* p_{-i} *on* σ_{-i}*), we have:* $u_i(e_i, p_{-i} \diamond \sigma_{-i}) \geq u_i(d_i, p_{-i} \diamond \sigma_{-i})$*. (If* p_{-i} *already uses up all the probability, we simply have* $u_i(e_i, p_{-i}) \geq u_i(d_i, p_{-i})$*—no* σ_{-i} *needs to be chosen.)*[5]

In the example from the previous subsection, we can set $D_r = \{\sigma_r^2\}$, $D_c = \{\sigma_c^2\}$, $E_r = \{\sigma_r^3, \sigma_r^4\}$, $E_c = \{\sigma_c^3, \sigma_c^4\}$, and $e_r^* = \sigma_r^3$. Then, by the reasoning that we did, it is impossible to set p_r and p_c so that the conditions are satisfied, and hence σ_r^3 is eliminable relative to these sets.

The spectrum of strength

In this section we show that the generalized eliminability criterion we defined in in the previous section spans a spectrum of strength all the way from Nash equilibrium (when the sets D_r, E_r, D_c, E_c are chosen as large as possible), to strict dominance (when the sets are chosen as small as possible). First, we show that the criterion is monotonically increasing, in the sense that the larger we make the sets D_r, E_r, D_c, E_c, the more strategies are eliminable. (We omit many of the proofs of this paper due to space constraint.)

Theorem 1 *If* e_r^* *is eliminable relative to* $D_r^1, E_r^1, D_c^1, E_c^1$*, and* $D_r^1 \subseteq D_r^2, E_r^1 \subseteq E_r^2, D_c^1 \subseteq D_c^2, E_c^1 \subseteq E_c^2$*, then* e_r^* *is eliminable relative to* $D_r^2, E_r^2, D_c^2, E_c^2$*.*

Next, we show that the Nash equilibrium concept is weaker[6] than our generalized eliminability criterion—in the sense that the generalized criterion can never eliminate a strategy that is in some Nash equilibrium. So, if a strategy can be eliminated by the generalized criterion, it can be eliminated by the Nash equilibrium concept.

Theorem 2 *If there is some Nash equilibrium that places positive probability on pure strategy* σ_r^**, then* σ_r^* *is not eliminable relative to any* D_r, E_r, D_c, E_c*.*

We next show that by choosing the sets D_r, E_r, D_c, E_c as large as possible, we can make the generalized eliminability criterion coincide with the Nash equilibrium concept.[7]

Theorem 3 *Let* $D_r = E_r = \Sigma_r$ *and* $D_c = E_c = \Sigma_c$*. Then* e_r^* *is eliminable relative to these sets if and only if there is no Nash equilibrium that places positive probability on* e_r^**.*

Moving to the other side of the spectrum, we now show that the concept of strict dominance is stronger than the generalized eliminability criterion—in the sense that the generalized eliminability criterion can always eliminate a strictly dominated strategy (as long as the dominating strategy is in D_r).

Theorem 4 *If pure strategy* σ_r^* *is strictly dominated by some mixed strategy* d_r*, then* σ_r^* *is eliminable relative to any* D_r, E_r, D_c, E_c *such that 1)* $\sigma_r^* \in E_r$*, and 2) all the pure strategies on which* d_r *places positive probability are in* D_r*.*

Finally, we show that by choosing the sets E_r, E_c as small as possible, we can make the generalized eliminability criterion coincide with the strict dominance concept.

Theorem 5 *Let* $E_c = \{\}$ *and* $E_r = \{e_r\}$*. Then* e_r *is eliminable relative to* D_r, E_r, D_c, E_c *if and only if it is strictly dominated by some mixed strategy that places positive probability only on elements of* D_r*.*

We are now ready to turn to computational aspects of the new eliminability criterion.

Applying the new eliminability criterion can be computationally hard

In this section, we demonstrate that applying the eliminability criterion can be computationally hard, in the sense of worst-case complexity.[8] We show that applying the eliminability criterion is coNP-complete in two key special cases (subclasses of the problem). The first case is the one in which the D_r, E_r, D_c, E_c sets are set to be as large as possible. Here, the hardness follows directly from Theorem 3 and a known hardness result on computing Nash equilibria (Gilboa & Zemel 1989; Conitzer & Sandholm 2003).

Theorem 6 *Deciding whether a given strategy is eliminable relative to* $D_r = E_r = \Sigma_r$ *and* $D_c = E_c = \Sigma_c$ *is coNP-complete, even when the game is symmetric.*

While this shows that the eliminability criterion is, in general, computationally hard to apply, we may wonder if there are special cases in which it is computationally easy to apply. Natural special cases to look at include those in which some of the sets D_r, E_r, D_c, E_c are small. The next theorem shows that applying the eliminability criterion remains coNP-complete even when $|D_r| = |D_c| = 1$.

Theorem 7 *Deciding whether a given strategy is eliminable relative to given* D_r, E_r, D_c, E_c *is coNP-complete, even when* $|D_r| = |D_c| = 1$*.*

[5] We need to make this case explicit for the case $E_{-i} = \Sigma_{-i}$.

[6] When discussing elimination of strategies, it is tempting to say that the stronger criterion is the one that can eliminate more strategies. However, when discussing solution concepts, the convention is that the stronger concept is the one that implies the other. Therefore, the criterion that can eliminate fewer strategies is actually the stronger one. For example, strict dominance is stronger than weak dominance, even though weak dominance can eliminate more strategies.

[7] Unlike Nash equilibrium, the generalized eliminability criterion does not discuss what probabilities should be placed on strategies that are not eliminated, so it only "coincides" with Nash equilibrium in terms of what it can eliminate.

[8] Because we only show hardness in the worst case, it is possible that many (or even most) instances are in fact easy to solve.

However, we will show later that the eliminability criterion can be applied in polynomial time if the E_i sets are small (regardless of the size of the D_i sets). To do so, we first need to introduce an alternative version of the definition.

An alternative, equivalent definition of the eliminability criterion

In this section, we will give an alternative definition of eliminability, and we will show it is equivalent to the one presented in Definition 1. While the alternative definition is slightly less intuitve than the original one, it is easier to work with computationally, as we will show in the next section. Informally, the alternative definition differs from the original one as follows: in the alternative definition, the completion of player $-i$'s mixed strategy has to be chosen *before* player i's strategy d_i is chosen (but after player i's strategy e_i with $p_i(e_i) > 0$ is chosen). The formal definition follows.

Definition 2 *Given a two-player game in normal form, subsets D_r, E_r of the row player's pure strategies Σ_r, subsets D_c, E_c of the column player's pure strategies Σ_c, and a distinguished strategy $e^*_r \in E_r$, we say that e^*_r is* not eliminable relative to *D_r, E_r, D_c, E_c, if there exist functions (partial mixed strategies) $p_r : E_r \to [0, 1]$ and $p_c : E_c \to [0, 1]$ with $p_r(e^*_r) > 0$, $\sum_{e_r \in E_r} p_r(e_r) \leq 1$, and $\sum_{e_c \in E_c} p_c(e_c) \leq 1$, such that the following holds. For both $i \in \{r, c\}$, for any $e_i \in E_i$ with $p_i(e_i) > 0$, there exists some completion of the probability distribution over $-i$'s strategies, given by $p^{e_i}_{-i} : \Sigma_{-i} \to [0, 1]$ (with $p^{e_i}_{-i}(e_{-i}) = p_{-i}(e_{-i})$ for all $e_{-i} \in E_{-i}$, and $\sum_{\sigma_{-i} \in \Sigma_{-i}} p^{e_i}_{-i}(\sigma_{-i}) = 1$), such that for any pure strategy $d_i \in D_i$, we have $u_i(e_i, p^{e_i}_{-i}) \geq u_i(d_i, p^{e_i}_{-i})$.*

We now show that the two definitions are equivalent.

Theorem 8 *The notions of eliminability put forward in Definitions 1 and 2 are equivalent. That is, e^*_r is eliminable relative to D_r, E_r, D_c, E_c according to Definition 1 if and only if e^*_r is eliminable relative to (the same) D_r, E_r, D_c, E_c according to Definition 2.*

Proof: The definitions are identical up to the condition that each strategy with positive probability (each $e_r \in E_r$ with $p_r(e_r) > 0$ and each $e_c \in E_c$ with $p_c(e_c) > 0$) must satisfy. We will show that these conditions are equivalent across the two definitions, thereby showing that the definitions are equivalent.

To show that the conditions are equivalent, we introduce another, zero-sum game that is a function of the original game, the sets D_r, E_r, D_c, E_c, the chosen partial probability distributions p_r and p_c, and the strategy e_i for which we are checking whether the conditions are satisfied. (Without loss of generality, assume that we are checking it for some strategy $e_r \in E_r$ with $p_r(e_r) > 0$.)

The zero-sum game has two players, 1 and 2 (not to be confused with the row and column players of the original game). Player 1 chooses some $d_r \in D_r$, and player 2 chooses some $\sigma_c \in \Sigma_c - E_c$. The utility to player 1 is $u_r(d_r, p_c \diamond \sigma_c) - u_r(e_r, p_c \diamond \sigma_c)$ (and the utility to player 2 is the negative of this). (We assume without loss of generality that p_c does not already use up all the probability, because in this case the conditions are trivially equivalent across the two definitions.)

First, suppose that player 1 must declare her probability distribution (mixed strategy) over D_r first, after which player 2 best-responds. Then, letting $\Delta(X)$ denote the set of probability distributions over set X, player 1 will receive $\max_{\delta_r \in \Delta(D_r)} \min_{\sigma_c \in \Sigma_c - E_c} \sum_{d_r \in D_r} \delta_r(d_r)(u_r(d_r, p_c \diamond \sigma_c) - u_r(e_r, p_c \diamond \sigma_c)) = \max_{\delta_r \in \Delta(D_r)} \min_{\sigma_c \in \Sigma_c - E_c} u_r(\delta_r, p_c \diamond \sigma_c) - u_r(e_r, p_c \diamond \sigma_c)$. This expression is at most 0 if and only if the condition in Definition 1 is satisfied.

Second, suppose that player 2 must declare his probability distribution (mixed strategy) over $\Sigma_c - E_c$ first, after which player 1 best-responds. Then, player 1 will receive $\min_{\delta_c \in \Delta(\Sigma_c - E_c)} \max_{d_r \in D_r} \sum_{\sigma_c \in \Sigma_c - E_c} \delta_c(\sigma_c)(u_r(d_r, p_c \diamond \sigma_c) - u_r(e_r, p_c \diamond \sigma_c)) = \min_{\delta_c \in \Delta(\Sigma_c - E_c)} \max_{d_r \in D_r} \sum_{e_c \in E_c} p_c(e_c)(u_r(d_r, e_c) - u_r(e_r, e_c)) + \sum_{\sigma_c \in \Sigma_c - E_c} (1 - \sum_{e_c \in E_c} p_c(e_c))\delta_c(\sigma_c)(u_r(d_r, \sigma_c) - u_r(e_r, \sigma_c)) = \min_{\delta_c \in \Delta(\Sigma_c - E_c)} \max_{d_r \in D_r} u_r(d_r, p_c \diamond \delta_c) - u_r(e_r, p_c \diamond \delta_c)$. This expression is at most 0 if and only if the condition in Definition 2 is satisfied.

However, by the Minimax Theorem (von Neumann 1927), the two expressions must have the same value, and hence the two conditions are equivalent. ■

Informally, the reason that Definition 2 is easier to work with computationally is that all of the continuous variables (the values of the functions $p_r, p_c, p^{e_r}_c, p^{e_c}_r$) are set by the party that is trying to prove that the strategy is not eliminable; whereas in Definition 1, some of the continuous variables (the probabilities defining the mixed strategies d_r, d_c) are set by the party trying to refute the proof that the strategy is not eliminable. This will become more precise in the next section.

A mixed integer programming approach

In this section, we show how to translate Definition 2 into a mixed integer program that determines whether a given strategy e^*_r is eliminable relative to given sets D_r, E_r, D_c, E_c. The variables in the program, which are all restricted to be nonnegative, are the $p_i(e_i)$ for all $e_i \in E_i$; the $p^{e_{-i}}_i(\sigma_i)$ for all $e_{-i} \in E_{-i}$ and all $\sigma_i \in \Sigma_i - E_i$; and *binary* indicator variables $b_i(e_i)$ for all $e_i \in E_i$ which can be set to zero if and only if $p_i(e_i) = 0$. The program is the following:

maximize $p_r(e^*_r)$ **subject to**

(probability constraints): for both $i \in \{r, c\}$, for all $e_i \in E_i$, $\sum_{e_{-i} \in E_{-i}} p_{-i}(e_{-i}) + \sum_{\sigma_{-i} \in \Sigma_{-i} - E_{-i}} p^{e_i}_{-i}(\sigma_{-i}) = 1$

(binary constraints): for both $i \in \{r, c\}$, for all $e_i \in E_i$, $p_i(e_i) \leq b_i(e_i)$

(main constraints): for both $i \in \{r, c\}$, for all $e_i \in E_i$ and all $d_i \in D_i$, $\sum_{e_{-i} \in E_{-i}} p_{-i}(e_{-i})(u_i(e_i, e_{-i}) - u_i(d_i, e_{-i})) + \sum_{\sigma_{-i} \in \Sigma_{-i} - E_{-i}} p^{e_i}_{-i}(\sigma_{-i})(u_i(e_i, \sigma_{-i}) - u_i(d_i, \sigma_{-i})) \geq (b_i(e_i) - 1)U_i$

In this program, the constant U_i is the maximum difference between two different utilities that player i may receive in the game, that is, $U_i = \max_{\sigma_r, \sigma'_r \in \Sigma_r, \sigma_c, \sigma'_c \in \Sigma_c} u_i(\sigma_r, \sigma_c) - u_i(\sigma'_r, \sigma'_c)$.

Theorem 9 *The mixed integer program has a solution with objective value greater than zero if and only if e^*_r is not eliminable relative to D_r, E_r, D_c, E_c.*

We obtain the following corollaries:

Corollary 1 *Checking whether a given strategy can be eliminated relative to given D_r, E_r, D_c, E_c is in coNP.*

Corollary 2 *Using the mixed integer program above, the time required to check whether a given strategy can be eliminated relative to given D_r, E_r, D_c, E_c is exponential only in $|E_r| + |E_c|$ (and not in $|D_r|, |D_c|, |\Sigma_r|$, or $|\Sigma_c|$).*

Iterated elimination

In this section, we study what happens when we eliminate strategies *iteratively* using the new criterion. The criterion can be iteratively applied by removing an eliminated strategy from the game, and subsequently checking for new eliminabilities in the game with the strategy removed, *etc.* (as in the more elementary, conventional notion of iterated dominance). First, we show that this procedure is, in a sense, sound.

Theorem 10 *Iterated elimination according to the generalized criterion will never remove a strategy that is played with positive probability in some Nash equilibrium of the original game.*

Because (the single-round version of) the eliminability criterion extends all the way to Nash equilibrium by Theorem 3, we get the following corollary.

Corollary 3 *Any strategy that can be eliminated using iterated elimination can also be eliminated in a single round (that is, without iterated application of the criterion).*

Interestingly, iterated elimination is in a sense incomplete:

Proposition 1 *Removing an eliminated strategy from a game sometimes decreases the set of strategies that can be eliminated.*

Proof: Consider the following game:

	L	M	R
U	2, 2	0, 1	0, 5
D	1, 0	1, 1	1, 0

The unique Nash equilibrium of this game is (D, M), for the following reasons. In order for it to be worthwhile for the row player to play U with positive probability, the column player should play L with probability at least $1/2$. But, in order for it to be worthwhile for the column player to play L with positive probability (rather than M), the row player should play U with probability at least $1/2$. However, if the row player plays U with probability at least $1/2$, then the column player's unique best response is to play R. Hence, the row player must play D in any Nash equilibrium, and the unique best response to D is M.

Thus, by Theorem 3, all strategies besides D and M can be eliminated. In particular, R can be eliminated. However, if we remove R from the game, the remaining game is:

	L	M
U	2, 2	0, 1
D	1, 0	1, 1

In this game, (U, L) is also a Nash equilibrium, and hence U and L can no longer be eliminated, by Theorem 2. ■

This example highlights an interesting issue with respect to using this eliminability criterion as a preprocessing step in the computation of Nash equilibria: it does not suffice to simply throw out eliminated strategies and compute a Nash equilibrium for the remaining game. Rather, we need to use the criterion more carefully: if we know that a strategy is eliminable according to the criterion we can restrict our attention to supports for the player that do not include this strategy.

The example also directly implies that iterated elimination according to the generalized criterion is path-dependent (the choice of which strategy to remove first affects which strategies can/will be removed later). The same phenomenon occurs with iterated weak dominance (one strategy weakly dominates another if the former always does at least as well as the latter, and in at least one case, strictly better). There is a sizeable literature on path (in)dependence for various notions of dominance (Gilboa, Kalai, & Zemel 1990; Borgers 1993; Osborne & Rubinstein 1994; Marx & Swinkels 1997; 2000; Apt 2004).

In light of these results, it may appear that there is not much reason to do iterated elimination using the new criterion, because it never increases and sometimes even decreases the set of strategies that we can eliminate. However, we need to keep in mind that Theorem 10, Corollary 3, and Proposition 1 do not pose any restrictions on the sets D_r, E_r, D_c, E_c, and therefore (by Theorems 2 and 3) are effectively results about iteratively removing strategies based on whether they are played in a Nash equilibrium. However, the new criterion is more informative and useful when there are restrictions on the sets D_r, E_r, D_c, E_c. Of particular interest is the restriction $|E_r| + |E_c| \leq k$, because by Corollary 2 this quantity determines the (worst-case) runtime of the mixed integer programming approach that we presented in the previous section. Under this restriction, it turns out that iterated elimination can eliminate strategies that single-round elimination cannot.

Proposition 2 *Under a restriction of the form $|E_r|+|E_c| \leq k$, iterated elimination can eliminate strategies that single-round elimination cannot (even when $k = 1$).*

Of course, even under this (or any other) restriction iterated elimination remains sound in the sense of Theorem 10. Therefore, one sensible approach to eliminating strategies is the following. Iteratively apply the eliminability criterion (with whatever restrictions are desired to increase the strength of the argument, or are necessary to make it computationally manageable, such as $|E_r| + |E_c| \leq k$), removing each eliminated strategy, until the process gets stuck. Then, start again with the original game, and take a different path of iterated elimination (which may eliminate strategies that could no longer be eliminated after the first path of elimination, as described in Proposition 1), until the process gets stuck—*etc.* In the end, any strategy that was eliminated in any one of the elimination paths can be considered "eliminated", and this is safe by Theorem 10.[9]

Interestingly, here the analogy with iterated weak dominance breaks down. Because there is no soundness theorem such as Theorem 10 for iterated weak dominance, considering all the strategies that are eliminated in some iterated weak dominance elimination path to be simultaneously "eliminated" can lead to senseless results. Consider for example the following game:

	L	M	R
U	1, 1	0, 0	1, 0
D	1, 1	1, 0	0, 0

U can be eliminated by removing R first, and D can be eliminated by removing M first—but these are the row player's only strategies, so considering both of them to be eliminated makes little sense.

Conclusions

We defined a generalized eliminability criterion for bimatrix games that considers whether a given strategy is eliminable relative to given dominator & eliminee subsets of the players' strategies. We showed that this definition spans a spectrum of eliminability criteria from strict dominance (when the sets are as small as possible) to Nash equilibrium (when the sets are as large as possible). Thus, eliminating a strategy relative to dominator & eliminee sets of intermediate size can provide a stronger argument for eliminating a strategy than Nash equilibrium, even when the strategy cannot be eliminated by (iterated) dominance. We showed that checking whether a strategy is eliminable according to this criterion is coNP-complete (both when all the sets are as large as possible and when the dominator sets each have size 1). We then gave an alternative definition of the eliminability criterion and showed that it is equivalent using the Minimax Theorem. We showed how this alternative definition can be translated into a mixed integer program of polynomial size with a number of (binary) integer variables equal to the sum of the sizes of the eliminee sets, implying that checking whether a strategy is eliminable according to the criterion can be done in polynomial time if the eliminee sets are small. Finally, we studied using the criterion for iterated elimination of strategies.

[9]This procedure is reminiscent of iterative sampling.

There are numerous avenues for future research. One is to use the new eliminability criterion and the computational tools we provided for it to speed up search-based techniques for computing Nash equilibria. Another avenue is to characterize the eliminability criterion at intermediate points of the spectrum. Yet another possibility is to try to find other special cases that can be computed in polynomial time. Finally, we can experimentally analyze the runtime of the mixed integer programming approach on random games (such as those generated by GAMUT (Nudelman *et al.* 2004)).

References

Apt, K. 2004. Uniform proofs of order independence for various strategy elimination procedures. *Contrib. Theoretical Econ.* 4(1).

Blum, B.; Shelton, C. R.; and Koller, D. 2003. A continuation method for Nash equilibria in structured games. In *IJCAI*.

Borgers, T. 1993. Pure strategy dominance. *Econometrica* 61(2):423–430.

Conitzer, V., and Sandholm, T. 2003. Complexity results about Nash equilibria. In *IJCAI*, 765–771.

Conitzer, V., and Sandholm, T. 2005. Complexity of (iterated) dominance. In *ACM Conf. on Electronic Commerce (ACM-EC)*.

Gilboa, I., and Zemel, E. 1989. Nash and correlated equilibria: Some complexity considerations. *Games & Economic Behavior*.

Gilboa, I.; Kalai, E.; and Zemel, E. 1990. On the order of eliminating dominated strategies. *OR Letters* 9:85–89.

Gilboa, I.; Kalai, E.; and Zemel, E. 1993. The complexity of eliminating dominated strategies. *Math. of OR* 18:553–565.

Gottlob, G.; Greco, G.; and Scarcello, F. 2003. Pure Nash equilibria: hard and easy games. In *TARK*, 215–230.

Kearns, M.; Littman, M.; and Singh, S. 2001. Graphical models for game theory. In *UAI*.

Knuth, D. E.; Papadimitriou, C. H.; and Tsitsiklis, J. N. 1988. A note on strategy elimination in bimatrix games. *OR Letters* 7(3):103–107.

Lemke, C., and Howson, J. 1964. Equilibrium points of bimatrix games. *J. Society of Industrial & Applied Math.* 12:413–423.

Leyton-Brown, K., and Tennenholtz, M. 2003. Local-effect games. In *IJCAI*.

Lipton, R.; Markakis, E.; and Mehta, A. 2003. Playing large games using simple strategies. In *ACM-EC*, 36–41.

Littman, M., and Stone, P. 2003. A polynomial-time Nash equilibrium algorithm for repeated games. In *ACM-EC*, 48–54.

Marx, L. M., and Swinkels, J. M. 1997. Order independence for iterated weak dominance. *Games & Econ. Behavior* 18:219–245.

Marx, L. M., and Swinkels, J. M. 2000. Corrigendum, order independence for iterated weak dominance. *Games & Economic Behavior* 31:324–329.

Nudelman, E.; Wortman, J.; Leyton-Brown, K.; and Shoham, Y. 2004. Run the GAMUT: A comprehensive approach to evaluating game-theoretic algorithms. In *AAMAS*.

Osborne, M. J., and Rubinstein, A. 1994. *A Course in Game Theory*. MIT Press.

Papadimitriou, C. 2001. Algorithms, games and the Internet. In *Annual Symposium on Theory of Computing (STOC)*, 749–753.

Pearce, D. G. 1984. Rationalizable strategic behavior and the problem of perfection. *Econometrica* 52:1029–1050.

Porter, R.; Nudelman, E.; and Shoham, Y. 2004. Simple search methods for finding a Nash equilibrium. In *AAAI*, 664–669.

Savani, R., and von Stengel, B. 2004. Exponentially many steps for finding a Nash equilibrium in a bimatrix game. In *FOCS*.

von Neumann, J. 1927. Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen* 100:295–320.