WEAKLY DOMINATED STRATEGIES: A MYSTERY CRACKED

JOHN HILLAS AND DOV SAMET

Abstract. An informal argument shows that common knowledge of rationality implies the iterative elimination of strongly dominated strategies. Rationality here means that players do not play strategies that are strongly dominated relative to their knowledge. We formalize and prove this claim. When by rationality we mean that players do not play strategies that are weakly dominated relative to their knowledge, then common knowledge of rationality does not imply iterative elimination of weakly dominated strategies. We show that it does imply an iterative elimination of *flaws* of weakly dominated strategies. The iterative elimination of flaws of strongly dominated strategies coincides with the iterative elimination of strongly dominated strategies.

1. Introduction

Iterative elimination of strongly dominated strategies is usually justified by the assumption that it is common knowledge that players are *strong-dominance rational*, by which we mean that they do not play strategies that are strongly dominated given their knowledge about the opponents' strategies.[1] The argument for this in the literature is informal, as knowledge, and a fortiori strong-dominance rationality which is defined in terms of knowledge, are not fully formalized. We show that the argument can be made rigorous by formally modeling knowledge and strong-dominance rationality.

The picture changes when weakly dominated strategies are considered. Iterative elimination of weakly dominated strategies fails to capture common knowledge of *weak-dominance rationality*, which requires that players do not play strategies that are weakly dominated given their knowledge about the opponents' strategies. The problem does not lie in the assumption of weak-dominance rationality, but in the process of elimination. In a nutshell, the problem is that strategies that are eliminated in early stages of the process, given the knowledge of the players at these stages, may not be weakly dominated given the knowledge at the end of the process, which is the knowledge players have, according to this argument, under the assumption of common knowledge of weak-dominance rationality.

Despite the awareness of this problem, no suggestion has been made how to fix the process of iterated elimination of weakly dominated strategies in order to capture common knowledge of weak-dominance rationality, due to the lack of formalization of weak-dominance rationality. Here, we formalize this notion analogously to the notion of strong-dominance rationality, and show that it implies an iterative process of elimination of *flaws* of weakly dominated strategies. This process circumvents the problem of iterative elimination of weakly dominated strategies

Date: Draft of December 9, 2013.

[1]See, for example, Myerson (1991, 57–61) and Binmore (1991, 149-150).

alluded above. A flaw of a dominated strategy of a player is a combination of this strategy with strategies of her opponents, which has a strong inequality in the set of inequalities that describe the dominance relationship.

Iterative elimination of flaws is the process that expresses common knowledge of both strong-dominance and weak-dominance rationality. When flaws of strongly dominated strategies are eliminated iteratively, the result is the same as the iterative elimination of strongly dominated *strategies*. In contrast, iterative elimination of flaws of weakly dominated strategies is different from the iterative elimination of weakly dominated strategies, and unlike the latter, it reflects common knowledge of weak-dominance rationality.

Hillas and Samet (2013) study strong-dominance and weak-dominance rationality under the names weak rationality and strong rationality correspondingly, and show their relation to correlated equilibrium. Here we use these notions to clarify the nature of the age-old solutions of iterative elimination of dominated strategies. We propose the iterative elimination of flaws which coincides with the known solution for the strong dominance case, and provides an alternative, contradiction free, replacement of the solution for the weak dominance case.

2. Iterative elimination of dominated strategies

Let G be a game with a finite set of players I, and a finite set of strategies S_i for each player i. The set of strategy profiles is $S = \times_i S_i$, and the set of the profiles of i's opponents is $S_{-i} = \times_{j \neq i} S_j$. The set of mixed strategies of i is denoted by $\Delta(S_i)$. The payoff function for i is $h_i \colon S \to R$. It is extended in the usual way to $\times_i \Delta(S_i)$. In order to describe iterated elimination of dominated strategies we use the following terminology.

Definition 1. (relative domination) Let T_{-i} be a nonempty set of profiles of i's opponents. A mixed strategy $\sigma_i \in S_i$ *strongly dominates* s_i *relative to* T_{-i} if $h_i(\sigma_i, t_{-i}) > h_i(s_i, t_{-i})$ for all $t_{-i} \in T_{-i}$. We say in this case that s_i is *strongly dominated relative to* T_{-i}. The strategy σ_i *weakly dominates* s_i *relative to* T_{-i} if $h_i(\sigma_i, t_{-i}) \geq h_i(s_i, t_{-i})$ for all $t_{-i} \in T_{-i}$, and at least one of these inequalities is strict. We say in this case that s_i is *weakly dominated relative to* T_{-i}.

Using this terminology we define processes of elimination of dominated strategies.

Definition 2. A process of *iterated elimination of strongly dominated strategies* consists of sequences of strategy profile sets $(S^0, S^1, \ldots, S^m)$, where $S^0 = S$, and for $k \geq 1$, $S^k = \times_i S_i^k$, where S_i^k is obtained from S_i^{k-1} by eliminating some strategies in the latter set which are strongly dominated relative to S_{-i}^{k-1}. In the sets S_i^m there are no strongly dominated strategies relative to S_{-i}^m. A process of *iterated elimination of weakly dominated strategies* is similarly defined, where in each stage, weakly dominated strategies are eliminated.[2]

[2]A strategy is strongly dominated by a mixed strategy if and only if it is not a best response against any probability distribution on the opponents profiles. Thus, it is possible to describe the iterated elimination of strongly dominated strategies as the iterative elimination of strategies that are not a best response (see Myerson, 1991, 88–89). Bernheim (1984) and Pearce (1984) suggested an iterative process that terminates with *rationalizable* profiles. In their process, strategies are eliminated when they are not a best response against probabilistic distribution over the opponents profiles, where player's strategies are drawn independently.

Of great importance is the following property of monotonicity of relative strong dominance.

Claim 1. (strong-dominance monotonicity) *If a strategy of i is strongly dominated relative to $T_{-i} \subseteq S_{-i}$ then it is also strongly dominated relative to $T'_{-i} \subseteq T_{-i}$.*

In subsection 4.1 we define a family of processes of iterative elimination of profiles, which subsumes the family of processes of iterative elimination of strongly dominated strategies. This larger family has a similar monotonicity property. We show that this monotonicity property guarantees that all the processes terminate in the same set of profiles. In contrast, relative weak dominance does not have this property, and different processes of iterative elimination of weakly dominated strategies may end in different sets of strategy profiles.

3. Strong dominance

3.1. **Informal justification of the iterated process.** The process of iterated elimination of strongly dominated strategies, described above, can be justified by assuming common knowledge of rationality. We consider this justification, which is described below, informal, since the term knowledge is not formalized.

Rationality, here, means that players do not play strongly dominated strategies. Common knowledge of rationality is tantamount to saying that all players are rational, they all know it, they all know that they all know it and so on. We can now justify the process of the strategy elimination in steps.

(1) All players are rational, and thus the strategy profile they play must be in S^1.
(2) Moreover, all players know that all players are rational, and thus they all know that the profile played is in S^1. Being rational, the strategy profile they play must be in S^2.
(3) Moreover, all players know that all players know that all players are rational, and thus ... and so on.

Note, that in step k each player i knows that her opponents are playing a profile in S^{k-1}_{-i}, and this is why dominance is considered only relative to this set. According to this explanation of the process, a player is rational if she does not play a strongly dominated strategy relative to the set of her opponents' profiles that her knowledge does not exclude. We call this kind of rationality, *strong-dominance rationality*, and define it formally later.

This informal argument for justifying the iterative elimination of strongly dominated strategies raises the following problem. By this argument, common knowledge of rationality implies that each player i knows that her opponents are playing a profile in S^m_{-i}. But, in the process we eliminate strongly dominated strategies given that player i knows less than that. That is, we assume along the way that player i knows that the other players play S^0_{-i} and then S^1_{-i} and so on. This seeming contradiction is spurious because of the monotonicity of relative strong dominance in Claim 1: if a strategy is eliminated in stage k when player i knows that her opponents' profile is in S^{k-1}_i, then it should be eliminated also when she knows more than that, that is, when she knows that her opponents' profile is in S^m_{-i}.

3.2. **The formal justification of the iterated process.** We formalize the argument of the previous subsection by using a state space in which knowledge is formally defined. Such a model enables us to directly capture common knowledge of rationality as an event in the state space, without the hierarchy of knowledge. We show that this modeling reaffirms the informal justification of the iterated process.[3]

Let Ω be a finite state space with a partition Π_i for each player i. At a state ω player i knows all the events that contain $\Pi_i(\omega)$, the element of i's partition that contains ω. For simplicity we assume that the meet of the partition consists of Ω, and thus, in each state, Ω is the only event that is common knowledge (see Aumann, 1976). In order to define knowledge about strategy profiles, we assume that each state determines the strategy profile played in the state . The strategies played in each state are given by functions $\mathbf{s}_i \colon \Omega \to S_i$, such that $\mathbf{s}_i(\omega)$ is the strategy i plays at state ω. We further assume that each player knows which strategy she plays. This means that $\mathbf{s}_i$ is measurable with respect to Π_i, or in other words, in each element of Π_i, i plays the same strategy in all the states in the element.

We can now define the event that a player is rational, in agreement with the informal definition of the previous subsection.

Definition 3. Player i is *strong-dominance rational* in state ω if the strategy she plays in ω is not strongly dominated relative to the set of her opponents' profiles which are not excluded by her knowledge in ω. That is, there is no strategy of hers that strongly dominates $\mathbf{s}_i(\omega)$ relative to the set $T_{-i}(\omega) = \{\mathbf{s}_{-i}(\omega') \mid \omega' \in \Pi_i(\omega)\}$.

The antecedent in the following proposition is that it is common knowledge that all players are strong-dominance rational. By our assumption that Ω is the only element of the meet, this is equivalent to saying that each of the players is strong-dominance rational in each state.

Proposition 1. *If it is common knowledge that the players are strong-dominance rational, then the strategy profiles played survive the iterated elimination of strongly dominated strategies.*

Proof. The proof, not surprisingly, resembles the iterated elimination of strongly dominated strategies, using the monotonicity property in Claim 1. Since strong-dominance rationality of all players is common knowledge, players are strong-dominance rational in each state. Therefore, for any state ω and player i, $\mathbf{s}_i(\omega)$ must be a strategy in S_i^1. Otherwise, it is strongly dominated in the game and therefore by monotonicity and the strong-dominance rationality of i, when i knows $T_{-i}(\omega)$ she cannot play it in ω. Thus, in all states the profiles played are in S^1. But then, for all ω and i, $T_{-i}(\omega) \subseteq S_{-i}^1$, and by the same argument as above, all the profiles played must be in S^2, and so on. $\square$

4. Weak dominance

The informal argument for iterated elimination of strongly dominated strategies can be stated verbatim for the case of weakly dominated strategies by interpreting rationality as *weak-dominance rationality.* That is, a player is weak-dominance rational if she does not play a weakly dominated strategy relative to the set of her opponent' profiles that she does not exclude. However, the argument that common

[3]As mentioned in footnote 1, eliminating dominated strategies is the same as eliminating strategies that are not a best response. Tan and Werlang (1988) formalize the iterative elimination of non-best response strategies in probabilistic belief spaces. Their characterization can be considered as dual to ours.

	L	C	R
T	1, 0	2, 0	3, 0
M	1, 2	2, 1	2, 1
B	1, 1	0, 3	4, 2

	L	C	R
T	1, 0	2, 0	3, 0
M	1, 2	3, 1	2, 1
B	1, 1	0, 3	

	L	C	R
T	1, 0	2, 0	3, 0
M	1, 2	3, 1	2, 1
B	1, 1		

	L	C	R
T	1, 0	2, 0	3, 0
M	1, 2		
B	1, 1		

FIGURE 1. Common knowledge of weak-dominance rationality

knowledge of rationality implies iterative elimination of weakly dominated strategies is flawed. Suppose that a process of iterative elimination of weakly dominated strategies ends with a set of profiles S^m. If, as the argument goes, common knowledge of weak-dominance rationality implies that the strategy profile must be in S^m, then players know it. Thus, weak-dominance rationality requires only that each player i does not play a strategy which is weakly dominated relative to S^m_{-i}. However, along the process we eliminated strategies of i relative to sets of profiles that reflect less knowledge. Since weak dominance does not have the monotonicity property in Claim 1, these eliminated strategies may not be weakly dominated relative to S^m_{-i} and thus their elimination is not justified in view of the player's knowledge.

In the following example we demonstrate, at this point still informally, how common knowledge of weak-dominance rationality implies a process of elimination different from iterative elimination of weakly dominated strategies.

Example 1. Consider the game on the left side of Figure 1, and assume that it is common knowledge that the players are weak-dominance rational. There is a unique dominated strategy in this game, strategy R, which is weakly dominated by C. However, eliminating strategy R may be unjustified, as was pointed out in explaining why the informal argument for iterative elimination of weakly dominated strategies is flawed; It is possible that player 2 is weak-dominance rational and nevertheless plays R if she knows that player 1 does not play B. Since we have not discovered at this stage what player 2 knows about player 1's strategy when there is common knowledge of weak-dominance rationality, we cannot eliminate the possibility that she plays R. But, obviously, if player 2 is weak-dominance rational and she *does* play R it must be because she knows that player 1 *does not* play B. Thus, the only thing we can infer from player 2's weak-dominance rationality is that the profile (B,R) cannot possibly be played. We refer to this profile as a *flaw* of the weakly dominated strategy R. We eliminate this profile and get the set of profiles with non-blank payoffs in the second table from the left, which we denote by T^1. We conclude that when players are weak-dominance rational, the profile played should be from this set.

Now, since the players know that they are weak-dominance rational, they know that the profile played is in T^1. Thus, if player 1 plays B, she excludes the possibility that player 2 plays R, and she knows that player 2 plays either L or C. Relative to this knowledge, strategy B is weakly dominated by M. However, we cannot eliminate strategy B since it is possible that player 1 knows that player 2 plays L in which case B is not weakly dominated by M. But we can conclude by player 1's weak-dominance rationality, that if she does play B it must be the case that she knows that the profile (B,C), the flaw of B relative to player 1's knowledge, is not played. Hence, we eliminate this profile and the set T^2 of the remaining profiles is depicted in the third table from the left. We conclude that when players know that they are weak-dominance rational then the profile they play must be in T^2.

When players know that they know that they are weak-dominance rational, then they know that the profile played is in T^2. If player 1 plays either C or R then, being weak-dominance rational, it must be the case that she excludes the possibility that player 1's strategy is M, because in that case L is weakly dominant, given her knowledge, which contradicts her weak-dominance rationality. Thus, we eliminate (M,C) and (M,R), the flaws of M, and the remaining set of profiles, T^3, is presented in the last table.

At this stage we cannot eliminate any profile from T^3 and we conclude that if there is common knowledge of weak-dominance rationality, then the strategy profile played should be in T^3.

Note that there exists a unique process of iterated elimination of weakly dominated strategies, in which the order of elimination is R, B, C. The remaining profiles are (T,L) and (M,L). Thus this process results in a smaller set of profiles than the process described above. However, common knowledge of weak-dominance rationality does not imply that only one of these two profiles is played. Indeed, if this were the case, then under the assumption of common knowledge of weak-dominance rationality player 1 knows that player 2 is playing L. But, then B is not weakly dominated given player 1's knowledge. The elimination of B is justified only if C is considered possible by player 1 when she plays B.

4.1. **Eliminating flaws.** We generalize the process described in Example 1; it is not the weakly dominated strategy which is the culprit that has to be eliminated, but rather the profiles in which this strategy is played in which the player is strictly worse off in comparison to the dominating strategy. We define it formally.

Definition 4. The profile (s_i, t_{-i}) is a *flaw of s_i relative to $T_{-i} \subseteq S_{-i}$*, if $t_{-i} \in T_{-i}$ and there exists some mixed strategy σ_i of i that dominates (strongly or weakly) s_i relative to T_{-i} such that $h_i(\sigma_i, t_{-i}) > h_i(s_i, t_{-i})$.

It is straightforward to show that flaws relative to some set have a property, stated next, which is similar to the monotonicity of strongly dominated strategies.

Claim 2. (monotonicity of flaws) *If (s_i, t_{-i}) is a flaw of s_i relative to T_{-i}, and $t_{-i} \in T'_{-i} \subseteq T_{-i}$, then (s_i, t_{-i}) is also a flaw of s_i relative to T'_{-i}.*

A process of *iterative elimination of flaws of weakly dominated strategies* consists of sequence of strategy profile sets $S^0, S^1, \ldots, S^m$, where $S^0 = S$, and for each $k > 0$, S^k is obtained by eliminating from S^{k-1} some profiles (s_i, t_{-i}) for some players i, where s_i is weakly dominated relative to $\{t'_{-i} \mid (s_i, t'_{-i}) \in S^{k-1}\}$, and (s_i, t_{-i}) is a flaw of s_i relative to the same set. The set S^m has no profiles that can be eliminated.

A process of *iterative elimination of flaws of strongly dominated strategies* is similarly defined, by changing "weakly dominated" in the previous description to "strongly dominated".

Due to the monotonicty property in Claim 2, iterative elimination of flaws has the desired property that all processes end in the same set of profiles.

Proposition 2. *All processes of iterative elimination of flaws of weakly dominated strategies end at the same set of profiles, and all processes of iterative elimination of flaws of strongly dominated strategies end at the same set of profiles.*

Proof. For any subset A of profiles denote by $f(A)$ the subset of profiles in A that cannot be eliminated in the following sense. The set $f(A)$ consists of all profiles $s \in A$ such that there is no i for which s_i is weakly dominated relative to $\{t'_{-i} \mid (s_i, t'_{-i}) \in A\}$, and $s = (s_i, s_{-i})$ is a flaw of s_i relative to the same set.

We show that f is monotonic, that is, if $B \subseteq A$, then $f(B) \subseteq f(A)$. Indeed, suppose s is not in $f(A)$. We show that s is not in $f(B)$. If s is not in B then obviously it is not in $f(B)$. Assume that $s \in B$. Then $s \in A \setminus f(A)$, and therefore for some i, s_i is weakly dominated relative to $\{t'_{-i} \mid (s_i, t'_{-i}) \in A\}$, and $s = (s_i, t_{-i})$ is a flaw of s_i, relative to the same set. By Claim 2, s is a flaw of s_i relative to $\{t'_{-i} \mid (s_i, t'_{-i}) \in B\}$, and therefore it is not in $f(B)$.

A process of elimination of flaws of weakly dominated strategies is a sequence $S^0, S^1, \ldots, S^m$, where $S^0 = S$, $f(S^m) = S^m$, and for each $k \geq 0$, $f(S^k) \subseteq S^{k+1} \subseteq S^k$. Suppose that $\hat{S}$ is the last set in another such process. We show by induction on k that $\hat{S} \subseteq S^k$. Thus, in particular, $\hat{S} \subseteq S^m$ which shows that any two processes end with the same set. As $S^0 = S$, the claim for $k = 0$ is obvious. Suppose that $\hat{S} \subseteq S^k$, for $k < m$. Then, by the monotonicity of f and the definition of the iterative process, $\hat{S} = f(\hat{S}) \subseteq f(S^k) \subseteq S^{k+1}$.[4]

The proof for elimination of flaws of strongly dominated strategies is the same. □

Note, that if s_i is strongly dominated relative to T_{-i}, then for *all* $t_{-i} \in T_{-i}$, (s_i, t_{-i}) is a flaw of s_i relative to T_{-i}. Thus, a process of iterative elimination of strongly dominated strategies is in particular a process of iterative elimination of flaws of strongly dominated strategies, in which all the profiles that contain a certain strongly dominated strategy are eliminated in one round. Thus, in view of Proposition 2 we conclude:

Corollary 1. *The set of profiles that survive iterative elimination of strongly dominated strategies coincides with the set of profiles that survive iterative elimination of flaws of strongly dominated strategies.*

However, as demonstrated in Example 1, the set of profiles that survive iterative elimination of flaws of weakly dominated strategies may differ from any set of profiles that end a process of iterative elimination of weakly dominated strategies. It is the first process that captures common knowledge of weak-dominance rationality, as we show it in the next subsection.

4.2. **Common knowledge of weak dominance rationality formalized.** We first define weak-dominance rationality analogously to strong-dominance rationality in Definition 3, and then state our main result.

Definition 5. Player i is *weak-dominance rational* in state ω if the strategy she plays at ω is not weakly dominated relative to the set of her opponents' profiles

[4] The proof hinges on only two properties of the function f: it is a contraction, that is, for each A, $f(A) \subseteq A$, and it is monotonic. Thus, we proved that for any monotonic elimination function, all processes of iterative elimination converge to the same fixed point. Monotonicity alone guarantees, by Tarski's fixed point theorem or Kleene's fixed point theorem, that the sequence $f(S), f^2(S), \ldots$ converges to the largest fixed point of f. In particular, the set of profiles that survive iterative elimination of flaws of weakly (strongly) dominated strategies is the largest fixed point of this type of elimination, and hence it subsumes any set of profiles that do not have flaws of weakly (strongly) dominated strategies.

which are not excluded by her knowledge at ω. That is, there is no strategy of hers that weakly dominates $\mathbf{s}_i(\omega)$ relative to the set $T_{-i}(\omega) = \{\mathbf{s}_{-i}(\omega') \mid \omega' \in \Pi_i(\omega)\}$.

Proposition 3. *If it is common knowledge that the players are weak-dominance rational, then the strategy profiles played survive the iterative elimination of flaws of weakly dominated strategies.*

Proof. The proof, like that of Proposition 1, mimics the iterative process. Let $S^0, \ldots, S^m$ be a process of iterative elimination of flaws of weakly dominated strategies. Since weak-dominance rationality of all players is common knowledge, players are weak-dominance rational in each state. Therefore, for any state ω and player i, $\mathbf{s}(\omega)$ must be a strategy in S^1. Otherwise, for some ω and i, $\mathbf{s}(\omega) = (s_i, s_{-i})$ where s_i is weakly dominated relative to S_{-i}, and (s_i, s_{-i}) is a flaw of s_i relative to S_{-i}. But, if so, then by Claim 2, s_i is also a flaw of s_i relative to $T_{-i}(\omega) = \{\mathbf{s}_{-i}(\omega') \mid \omega' \in \Pi_i(\omega)\}$. This implies that s_i is weakly dominated with respect to this set, which contradicts our assumption.

Now, as $\mathbf{s}(\omega) \in S^1$ for each ω, $T_{-i}(\omega) = \{\mathbf{s}_{-i}(\omega') \mid \omega' \in \Pi_i(\omega)\} \subseteq \{s'_{-i} \mid (s_i, s'_{-i}) \in S^1\}$. Therefore, for each ω, $\mathbf{s}(\omega) \in S^2$, or else, for some i, $\mathbf{s}(\omega) = (s_i, s_{-i})$ is a flaw of s_i relative to $\{s'_{-i} \mid (s_i, s'_{-i}) \in S^1\}$ and hence, by monotonicity, also relative to $T_{-i}(\omega)$, which contradicts the weak-dominance rationality of i at ω. The argument for the next stages is similar. $\square$

In light of Corollary 1, Proposition 1 can be restated analogously to Proposition 3.

Proposition 4. *If it is common knowledge that the players are strong-dominance rational, then the strategy profiles played survive the iterative elimination of flaws of strongly dominated strategies.*

We conclude that iterative elimination of flaws captures both the case of common knowledge of strong-dominance rationality and common knowledge of weak dominance rationality. Iterative elimination of strongly dominated *strategies* coincides with iterative elimination of flaws of such strategies. This coincidence does not hold in the weak dominance case, and the iterative elimination of weakly dominated strategies fails to capture common knowledge of weak-dominance rationality.

References

Robert J. Aumann (1976): "Agreeing to disagree," *Ann. Statist.*, 4(6), 1236–1239.

B. Douglas Bernheim (1984): "Rationalizable strategic behavior," *Econometrica*, 52(4), 1007–1028.

Kenneth Binmore (1991): *Fun and Games: A Text on Game Theory*, D. C. Heath and Company.

John Hillas and Dov Samet (2013): "Non-probabilistic correlated equilibrium as an expression of non-Bayesian rationality," a manuscript.

Roger Myerson (1991): *Game Theory: Analysis of Conflict*, Harvard University Press, Cambridge, MA.

David G. Pearce (1984): "Rationalizable strategic behavior and the problem of perfection," *Econometrica*, 52(4), 1029–1050.

Tommy Chin-Chiu Tan and Sérgio Ribeiro da Costa Werlang (1988): "The Bayesian foundations of solution concepts of games," *Journal of Economic Theory*, 45(2), 370–391.