# Dominance and Iterated Dominance

Adam Brandenburger*

Version 01/04/07

## 1 Introduction

Our goal in these notes is to answer the question: Suppose that the players in a game are rational, each player believes that the other players are rational, each player believes that the other players believe this, and so on ad infinitum? Where does this assumption lead? What strategies might the players choose when it holds?

**Definition 1** *Say that an event is* ***common belief*** *if all players believe it, all players believe that all players believe it, all players believe that all players believe that all players believe it, and so on ad infinitum.*

With this terminology, we can rephrase our question as: What strategies can players choose, when they are rational and their rationality is common belief?

To answer this question, we introduce some formal apparatus.

## 2 Strategic-Form Games

We consider a *two-player strategic-form game* $\Gamma = \langle S^a, S^b, \pi^a, \pi^b \rangle$, where $S^a$ (resp. $S^b$) is Ann's (resp. Bob's) *strategy set* and $\pi^a : S^a \times S^b \to \Re$ (resp. $\pi^b : S^b \times S^a \to \Re$) is Ann's (resp. Bob's) *payoff function.* A piece of notation: For any finite set $X$, let $\Delta(X)$ denote the set of probability measures (distributions) over $X$.

We shall assume, unless otherwise noted, that the strategy sets $S^a$ and $S^b$ are finite.

**Definition 2** *A* ***mixed strategy*** *of Ann is a probability measure $\sigma^a \in \Delta(S^a)$. For each $s^a \in S^a$, write $\sigma^a(s^a)$ for the weight assigned to $s^a$ by $\sigma^a$.*[1]

*With the assistance of Amanda Friedenberg. Do not copy or circulate these notes without the permission of the author. dominance-01-04-07

[1] This and subsequent definitions have obvious counterparts for Bob.

(An element $s^a \in S^a$ may be called a *pure strategy* of Ann.) If Ann uses the mixed strategy $\sigma^a$ and Bob uses the mixed strategy $\sigma^b$, then we shall denote the resulting expected payoff to Ann by $v^a(\sigma^a, \sigma^b)$, where

$$v^a(\sigma^a, \sigma^b) = \sum_{(s^a, s^b) \in S^a \times S^b} \sigma^a(s^a)\sigma^b(s^b)\pi^a(s^a, s^b).$$

The expected payoff to Bob, denoted by $v^b(\sigma^b, \sigma^a)$, is defined similarly. We won't distinguish between the pure strategy $s^a \in S^a$ and the mixed strategy $\sigma^a \in \Delta(S^a)$ that assigns probability one to $s^a$. Accordingly, we shall feel free to write $v^a(s^a, \sigma^b)$ for Ann's expected payoff when Ann uses the strategy just described and Bob uses the mixed strategy $\sigma^b$; and to use other, similar expressions. (In particular, then, we have $v^a(s^a, s^b) = \pi^a(s^a, s^b)$.)

**Definition 3** *A strategy $s^a \in S^a$ is **(strongly) dominated** if there is a $\sigma^a \in \Delta(S^a)$ such that $v^a(\sigma^a, s^b) > v^a(s^a, s^b)$ for all $s^b \in S^b$.*

Note that in this definition it is not the same to say "...if there is an $r^a \in S^a$ such that $v^a(r^a, s^b) > v^a(s^a, s^b)$ for all $s^b \in S^b$." Consider Figure 1. The strategy $B$ is dominated by the mixed strategy that puts equal weight on $T$ and $M$, but does not satisfy the condition of the preceding sentence.

| | *L* | *R* |
|---|---|---|
| *T* | 3 | 0 |
| *M* | 0 | 3 |
| *B* | 1 | 1 |

**Figure 1**

[One might also wonder whether the following definition would be different: A strategy $s^a \in S^a$ is dominated if and only if there is a $\sigma^a \in \Delta(S^a)$ such that $v^a(\sigma^a, \sigma^b) > v^a(s^a, \sigma^b)$ for all $\sigma^b \in \Delta(S^b)$. This is, in fact, equivalent to Definition 3. Clearly, it is at least as stringent, i.e. if a strategy is dominated under this definition then it is certainly dominated under Definition 3. So, let us suppose that $s^a$ is dominated by $\sigma^a \in \Delta(S^a)$ under Definition 3, and fix any $\sigma^b \in \Delta(S^b)$. Then

$$\begin{aligned} v^a(\sigma^a, \sigma^b) &= \sum_{s^b \in S^b} \sigma^b(s^b) v^a(\sigma^a, s^b) \\ &> \sum_{s^b \in S^b} \sigma^b(s^b) v^a(s^a, s^b) \\ &= v^a(s^a, \sigma^b), \end{aligned}$$

establishing that $s^a$ is dominated under the second definition.]

We now establish an equivalent definition of dominance that indicates why a rational player won't play a dominated strategy.

**Definition 4** *A strategy $s^a \in S^a$ is **rational** if there is a $\sigma^b \in \Delta(S^b)$ such that $v^a(s^a, \sigma^b) \geq v^a(r^a, \sigma^b)$ for all $r^a \in S^a$.*

Note that we state this definition in terms of one of Bob's mixed strategies $\sigma^b$. But in the interpretation, we are going to think of this mixed strategy as really being a probability distribution held by Ann about Bob's (definite) choice of strategy. We can do this, of course, because formally these are the same object.

Note also that in this definition it is not the same to say "...if there is an $s^b \in S^b$ such that $v^a(s^a, s^b) \geq v^a(r^a, s^b)$ for all $r^a \in S^a$." Consider Figure 2. The strategy $B$ maximizes expected payoff with respect to the mixed strategy that puts equal weight on $L$ and $R$, but does not satisfy the condition of the preceding sentence.

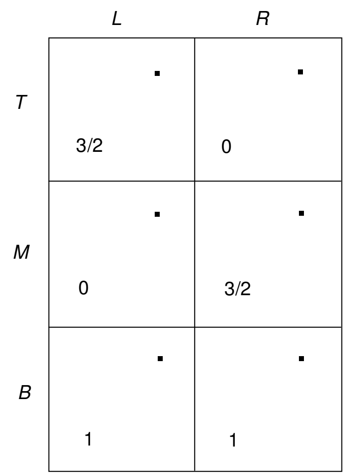


**Figure 2**

**Proposition 1** *A strategy $s^a \in S^a$ is rational if and only if it is undominated.*

**Sketch of Proof.** We give a complete proof of the "Only If" direction, and sketch the "If" direction. So, suppose that $s^a$ is rational with respect to some $\sigma^b \in \Delta(S^b)$. Then for any $\sigma^a \in \Delta(S^a)$, we have

$$\begin{aligned} v^a(\sigma^a, \sigma^b) &= \sum_{r^a \in S^a} \sigma^a(r^a) v^a(r^a, \sigma^b) \\ &\leq \sum_{r^a \in S^a} \sigma^a(r^a) v^a(s^a, \sigma^b) \\ &= v^a(s^a, \sigma^b). \end{aligned}$$

It follows, using the parenthetical remark after Definition 3, that $s^a$ is undominated.

Next, suppose that $s^a$ is undominated. We sketch the proof that $s^a$ is then rational, for the case that $S^b$ contains just two strategies, say $r^b$ and $s^b$. Refer to Figure 3, where the strategies $r^a \in S^a$ are plotted as the points $(v^a(r^a, r^b), v^a(r^a, s^b))$.

It is clear geometrically that we can find positive numbers $\lambda$ and $\mu$ such that

$$\lambda v^a(s^a, r^b) + \mu v^a(s^a, s^b) \geq \lambda v^a(r^a, r^b) + \mu v^a(r^a, s^b)$$

for all $r^a \in S^a$. Setting $\sigma^b(r^b) = \lambda/(\lambda+\mu)$ and $\sigma^b(s^b) = \mu/(\lambda+\mu)$ completes the argument. (The precise, and general, argument uses the Supporting Hyperplane Theorem from convex analysis.) ■

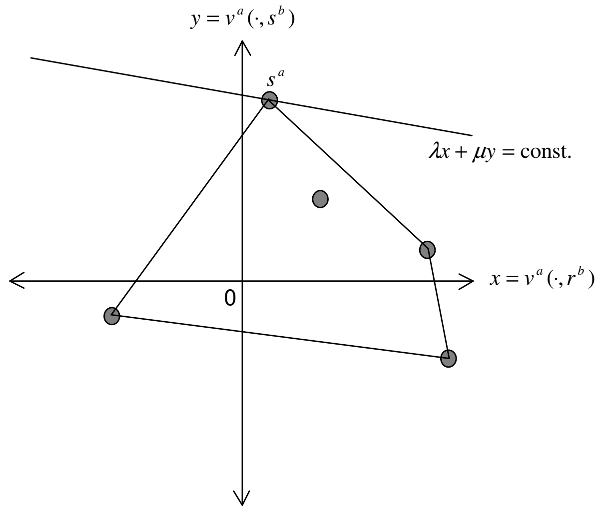


**Figure 3**

## 3 Iterated Dominance

**Example 1** *Consider the matrix in Figure 4. Neither of Row's strategies is dominated. Both L and M are undominated for Column, but R is dominated (by a 50:50 mixture of L and M). Eliminating R, we obtain a submatrix, in which D now is dominated. Eliminating D yields a further submatrix, in which M becomes dominated. We are left with U for Row and L for Column.*

| | *L* | *M* | *R* |
|---|---|---|---|
| *U* | 10<br>4 | 0<br>3 | 3<br>1 |
| *D* | 0<br>0 | 10<br>2 | 3<br>10 |

**Figure 4**

This example illustrates a general procedure (called *iterated dominance*) which we will first define verbally and then formally. Verbally, the procedure is:

Step 1: For each player, delete all strategies of that player that are dominated in the game. This defines a 'reduced' game.

Step 2: For each player, delete all strategies of that player that are dominated in the reduced game. This defines a further reduced game.

...

The procedure continues until no further deletion is possible, and a 'residual' game is produced. The strategies in this residual game are called the *iteratively undominated* strategies.

To give a formal definition, some notation is needed: Given subsets of strategies $X \subseteq S^a$ and $Y \subseteq S^b$, let $\pi^a|(X \times Y)$ denote the restriction of $\pi^a$ to $X \times Y$ and $\pi^b|(Y \times X)$ the restriction of $\pi^b$ to $Y \times X$. Now, let $S_0^a = S^a$ and $S_0^b = S^b$, and define $S_m^a$ and $S_m^b$ inductively by

$$S_{m+1}^a = \{s^a \in S_m^a : s^a \text{ is undominated in the game } \langle S_m^a, S_m^b, \pi^a|(S_m^a \times S_m^b), \pi^b|(S_m^b \times S_m^a)\rangle\},$$

$$S_{m+1}^b = \{s^b \in S_m^b : s^b \text{ is undominated in the game } \langle S_m^a, S_m^b, \pi^a|(S_m^a \times S_m^b), \pi^b|(S_m^b \times S_m^a)\rangle\},$$

Note that since $S^a$ and $S^b$ are finite, there is an $M$ such that $S_m^a = S_M^a$ and $S_m^b = S_M^b$ for all $m \geq M$. (Check why this is true!)

**Definition 5** *A strategy $s^a \in S_M^a$ or $s^b \in S_M^b$ is called* ***iteratively undominated****.*

Of course, our interest in the iteratively undominated strategies is that they would appear to be the strategies that can be played when both players are rational and there is common belief of rationality. We can see the idea in Example 1 above. We eliminated $R$ on the assumption that Column was rational. Presumably, we then eliminated $D$ on the assumption that Row was rational and believed Column to be rational. (This was why Row could effectively restrict attention to the submatrix without $R$.) Likewise, the elimination of $M$ would appear to follow from the assumption that Column was rational (again), believed Row to be rational, and believed that Row believed Column to be rational.

Our line of reasoning suggests the following theorem. (To get to the theorem, there is one step we have glossed over so far. Can you see what it is?)

**Theorem 1** *(Informal statement) Suppose that all players are rational and that their rationality is common belief. Then each player will choose an iteratively undominated strategy.*

Complete mathematical treatments of this theorem can be found in the game theory literature. See, for example, the references in "Knowledge and Equilibrium in Games," by Adam Brandenburger, *Journal of Economic Perspectives*, 6, 1992, 83-101.

**Example 2** [2] *Two women–let's call them Hannah and Sarah–come to King Solomon and ask him to adjudicate who is the true mother of a small baby. Solomon decides on the following procedure. Each woman has to say who she considers the true mother to be–i.e. she has to say either "Hannah" or "Sarah." Solomon also explains that if both women say "Hannah," then the baby will be awarded to Hannah; likewise, if both say "Sarah," then the baby will be awarded to Sarah. Finally, if the two women disagree, then the baby will be cut in half, and one half will be awarded to each woman.*

*To model this situation as a game, we assume (with no loss of generality) that the true mother is Hannah. A possible payoff matrix is then the one in Figure 5 below, where $H$ (resp. $S$) denotes the strategy of saying "Hannah" (resp. "Sarah"). Analyze this game using iterated dominance.*

| | | Sarah | |
|---|---|---|---|
| | | *H* | *S* |
| Hannah | *H* | 2, 0 | 0, 1 |
| | *S* | 0, 1 | 1, 2 |

**Figure 5**

[2] Taken from teaching material developed by Professor Ben Polak (Department of Economics, Yale University).

# 4 Caveat

It is important to recognize that there is nothing inevitable about the assumption of rationality and common belief of rationality. Indeed, it seems like a very stringent assumption in general. Perhaps both players are indeed rational. But can either really be sure that the other player is rational? Even if Ann is sure that Bob is rational, can Ann be absolutely sure that Bob thinks that she is rational? Etc. It is probably best to treat the assumption of rationality and common belief of rationality as a kind of 'baseline.' It describes an 'ideal' situation which is certainly of interest, but which may well differ from many real situations.

# 5 Correlation

All of the preceding analysis, which was for two players, extends immediately to $n > 2$ players. Other than some more notation, nothing new is needed at the technical level. But a very important conceptual issue *does* arise, as we now show. Refer to the game in Figure 6.

| $A$ | $L$ | $R$ |
|---|---|---|
| $U$ | 1, 1, 1 | 1, 0, 1 |
| $D$ | 0, 1, 0 | 0, 0, 0 |

| $B$ | $L$ | $R$ |
|---|---|---|
| $U$ | 2, 2, 0.7 | 0, 0, 0 |
| $D$ | 0, 0, 0 | 2, 2, 0.7 |

| $C$ | $L$ | $R$ |
|---|---|---|
| $U$ | 1, 1, 0 | 1, 0, 0 |
| $D$ | 0, 1, 1 | 0, 0, 1 |

**Figure 6**

In this game, player 1 chooses the row $U$ or $D$, player 2 chooses the column $L$ or $R$, and player 3 chooses the matrix $A$, $B$, or $C$. (As always, each choice is made without knowledge of the other players' choices.)

Consider the following probability measure on the set $\{U, D\} \times \{L, R\}$, i.e. the product of the set of strategies of player 1 and the set of strategies of player 2: Probability 1/2 is assigned to the pair $(U, L)$ and probability 1/2 is assigned to the pair $(D, R)$. In words, player 3 assigns probability 1/2 to the event that player 1 chooses $U$ and player 2 chooses $L$, and probability 1/2 to the event that player 1 chooses $D$ and player 2 chooses $R$. It is easy to see that the strategy $B$ maximizes expected payoff with respect to this probability measure. It yields player 3 an expected payoff of 0.7 while either $A$ or $B$ yields $(1/2)\times 0+(1/2)\times 1 = 0.5$.

Note, however, that this probability measure exhibits *correlation*: It is not the product of two measures, one on $\{U, D\}$ and the other on $\{L, R\}$. The traditional view in game theory was to rule out such dependence, on the grounds that players 1 and 2 are assumed to choose their strategies 'independently.' But this view has been challenged by Aumann ("Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55, 1987, 1-18), who writes:

> In games with more than two players, correlation may express the fact that what 3, say, thinks 1 will do may depend on what he thinks 2 will do. This has no connection with any overt or even covert collusion between 1 and 2; they may be acting entirely independently. Thus it may be common knowledge that both 1 and 2 went to business school, or perhaps to the same business school; but 3 may not know what is taught there. In that case 3 would think it quite likely that they would take similar actions, without being able to guess what those actions might be.

If correlations are allowed, as Aumann argues they should be, then Proposition 1 extends to games with more than two players. That is, a strategy is rational if and only if it is undominated. In the present example, the strategy $B$ is rational, as we saw, and it is undominated. To see the latter, let $\sigma_3 = (p, 1-p-q, q)$ be any mixed strategy of player 3, where weight $p$ (resp. $q$) is put on the strategy $A$ (resp. $C$). Then, if players 1 and 2 choose $U$ and $L$ respectively, the expected payoff of $\sigma_3$ is

$$p \times 1 + (1-p-q) \times 0.7 + q \times 0,$$

while if players 1 and 2 choose $D$ and $R$ respectively, the expected payoff of $\sigma_3$ is

$$p \times 0 + (1-p-q) \times 0.7 + q \times 1.$$

For $B$ to be dominated, we would then require the two inequalities

$$\begin{aligned} p \times 1 + (1-p-q) \times 0.7 + q \times 0 &> 0.7, \\ p \times 0 + (1-p-q) \times 0.7 + q \times 1 &> 0.7, \end{aligned}$$

to hold, which is easily seen to be impossible.

But, if correlations are ruled out, we lose the equivalence between rational and undominated strategies. To see this, let $(\alpha, 1-\alpha)$ be a probability measure on player 1's strategy set $\{U, D\}$, and let $(\beta, 1-\beta)$ be a probability measure on player 2's strategy set $\{L, R\}$. Then, the expected payoff to player 3 of the strategy $A$ is $\alpha$, the expected payoff of $C$ is $(1-\alpha)$, while the expected payoff of $B$ is $\alpha\beta \times 0.7 + (1-\alpha)(1-\beta) \times 0.7$. It is not hard to see that

$$\alpha\beta \times 0.7 + (1-\alpha)(1-\beta) \times 0.7 < \max\{\alpha, 1-\alpha\},$$

so that $B$ is not rational, when we require independence. As far as the other direction is concerned, you can check that the "If" argument in the proof of

Proposition 1 continues to go through. So, we still have that if a strategy is rational, then it is undominated.

Summing up, the conclusion is that if correlations are allowed, the sets of rational strategies and undominated strategies coincide (just as in the two-player case). But, if independence is imposed, then any rational strategy is undominated, but not necessarily vice versa.