

From: Dasgupta, Partha, Douglas Gale, Oliver Hart, and Eric Maskin (editors). *Economic Analysis of Markets and Games*. The MIT Press (Cambridge, MA), 1992.

## Lexicographic Probabilities and Iterated Admissibility

### Adam Brandenburger

The implications of common knowledge of rationality in normal-form games have been explored in recent years by Aumann (1987), Bernheim (1984, 1986), Brandenburger and Dekel (1987, 1989), Pearce (1984), Tan and Werlang (1988), and others. This line of research is concerned with providing what Aumann has termed "analytic" foundations for game-theoretic solution concepts.

The most conservative approach is due to Bernheim and Pearce. Bernheim and Pearce's notion of rationalizability assumes that it is common knowledge that each player is a subjective expected utility maximizer (Savage 1954; Anscombe and Aumann 1963), but that the precise subjective probability distribution held by each player is not necessarily common knowledge. Bernheim and Pearce show that this assumption implies that players choose iteratively undominated actions.<sup>1</sup>

The hypothesis of subjective expected utility maximization does not preclude players from choosing actions that are weakly dominated (inadmissible). Within decision theory, admissibility has long been advanced as an important criterion for choice (see, for example, Luce and Raiffa 1957, chapter 13). On the premise that the criterion is equally reasonable in a game-theoretic context, Kohlberg and Mertens (1986, p. 1014) view "admissibility of the players' strategies as a basic requirement" of a satisfactory solution concept. Moreover, they argue in favor of an iterated admissibility requirement.<sup>2</sup>

This leads naturally to the following question: What are the implications of assuming, in addition to common knowledge of rationality, that it is common knowledge that players choose only admissible actions? Samuelson (1989) and Börgers (1990) have addressed this and other questions concerning the interplay between the concepts of common knowledge and admissibility. The present paper was inspired by Börgers' work in particular, and uses lexicographic probabilities to derive a result (proposition 2 below) which is analogous to a result of Börgers. Both results address the question posed at the beginning of this paragraph.

By analogy with the result linking common knowledge of subjective expected utility maximization and iterated deletion of dominated actions, one might conjecture that assuming it to be common knowledge that players choose *admissible* actions implies that they choose *iteratively admissible*

actions. However, as Samuelson has demonstrated, establishing such a result seems problematic. Börgers shows that a weaker result is achievable. He shows that *approximate* common knowledge of admissibility implies that players choose actions that survive one round of deletion of weakly dominated actions followed by iterated deletion of strongly dominated actions. This latter procedure was proposed by Dekel and Fudenberg (1990), who showed it to be equivalent to iterated admissibility when players may be uncertain about other players' payoffs.

Börgers is led to assume approximate common knowledge since, in order to incorporate admissibility, he supposes that a player's subjective probability distribution assigns strictly positive weight to *all* choices of the other players. If there are any dominance relationships in the game, this condition will violate common knowledge of rationality since it requires players to place positive probability on dominated choices by other players. In order to resolve this conflict, Börgers replaces common knowledge with a notion of approximate common knowledge that builds on the definitions of Monderer and Samet (1989) and Stinchcombe (1988).

The objective of the present paper is to show how lexicographic probabilities can be used to prove a result analogous to that of Börgers.<sup>3</sup> Thus, instead of approximate common knowledge, I use a lexicographic analogue, which I call common first-order knowledge. To define common first-order knowledge of admissibility, I employ the non-Archimedean version of subjective expected utility theory developed by Blume, Brandenburger, and Dekel (1991a). According to this theory, a decision maker possesses, in addition to a utility function on outcomes, a *hierarchy* of probability distributions on uncertain events that is used lexicographically in ranking actions. (In conventional subjective expected utility theory, a decision maker possesses a *single* probability distribution.) Admissibility of the decision maker's choices is ensured by requiring that every event receive positive measure under some probability distribution in the hierarchy. As for common first-order knowledge, one begins by saying that the decision maker "first-order knows an event" if that event is assigned probability 1 by the first-order probability distribution in the decision maker's hierarchy. An event is then common first-order knowledge if everyone first-order knows it, everyone first-order knows that everyone first-order knows it, and so on.

## 1 The Result

Consider an  $n$ -person normal-form game

$$\Gamma = \langle A^1, \dots, A^n; u^1, \dots, u^n \rangle$$

where, for each  $i = 1, \dots, n$ ,  $A^i$  is player  $i$ 's finite set of pure strategies (henceforth actions) and  $u^i: A^1 \times \dots \times A^n \rightarrow \mathbb{R}$  is player  $i$ 's von Neumann-Morgenstern utility function. For any  $i = 1, \dots, n$ , the set  $A^1 \times \dots \times A^{i-1} \times A^{i+1} \times \dots \times A^n$  will be denoted by  $A^{-i}$ .

In developing the theory of rationalizability, Bernheim and Pearce augment the orthodox description of a normal-form game by specifying for each player not only a utility function over outcomes but also a subjective probability distribution over the uncertainty that player faces. By contrast, we are going to assume that players adhere to the non-Archimedean version of subjective expected utility theory developed by Blume, Brandenburger, and Dekel (1991a). It is therefore supposed that each player  $i$  possesses, in addition to a utility function  $u^i$ , a vector  $\rho^i = (p_1^i, \dots, p_{K^i}^i)$ , for some integer  $K^i$ , of probability measures on  $A^{-i}$ , where  $p_1^i$  is player  $i$ 's first-order belief about the other players' choices of action,  $p_2^i$  is player  $i$ 's second-order belief, and so on. The component probability measures are used lexicographically in ranking actions. Thus, player  $i$  prefers an action  $a^i \in A^i$  over another action  $b^i \in A^i$  if and only if<sup>4</sup>

$$\left[ \sum_{a^{-i} \in A^{-i}} p_k^i(a^{-i}) u^i(a^i, a^{-i}) \right]_{k=1}^{K^i} \geq_L \left[ \sum_{a^{-i} \in A^{-i}} p_k^i(a^{-i}) u^i(b^i, a^{-i}) \right]_{k=1}^{K^i}.$$

Following the terminology of Blume, Brandenburger, and Dekel (1991a), I will refer to a vector  $\rho^i = (p_1^i, \dots, p_{K^i}^i)$  of probability measures as a *lexicographic probability system* (LPS).

An important consequence of assuming that player  $i$  adheres to the non-Archimedean version of subjective expected utility theory just described is that admissibility of  $i$ 's choices can be guaranteed without at the same time ruling out the possibility that  $i$  assigns probability 0 (in a certain sense) to some of the other players' choices.

**DEFINITION 1** An LPS  $\rho^i = (p_1^i, \dots, p_{K^i}^i)$  of player  $i$  has *full support* if, for each  $a^{-i} \in A^{-i}$ ,  $p_k^i(a^{-i}) > 0$  for some  $k = 1, \dots, K^i$ .

**PROPOSITION 1** Suppose player  $i$ 's LPS  $\rho^i = (p_1^i, \dots, p_{K^i}^i)$  has full support. Then any action  $a^i \in A^i$  that satisfies

$$\left[ \sum_{a^{-i} \in A^{-i}} p_k^i(a^{-i}) u^i(a^i, a^{-i}) \right]_{k=1}^{K^i} \geq_L \left[ \sum_{a^{-i} \in A^{-i}} p_k^i(a^{-i}) u^i(b^i, a^{-i}) \right]_{k=1}^{K^i}$$

for all  $b^i \in A^i$  is admissible.

Proposition 1 will not be proved here, since it is very closely related to theorem 4.2 in Blume, Brandenburger, and Dekel 1991a and proposition 1 in Blume, Brandenburger, and Dekel 1991b. Proposition 1 says that if player  $i$ 's LPS has full support, then  $i$ 's optimal choices will be admissible. Of course, the same effect could have been achieved in the case in which player  $i$  possessed a *single* probability distribution on  $A^{-i}$ , by simply requiring that distribution to have full support in the conventional sense. The advantage of the non-Archimedean formulation is that in guaranteeing admissibility of player  $i$ 's choices, one does not at the same time have to forgo the possibility that certain action profiles of the other players are deemed infinitely unlikely by player  $i$ . Specifically, these are the action profiles  $a^{-i} \in A^{-i}$  that are assigned first-order probability 0 by player  $i$  ( $p_1^i(a^{-i}) = 0$ ). of course, any such action profile  $a^{-i}$  is not viewed as *completely* impossible by  $i$ , since it must be that  $p_k^i(a^{-i}) > 0$  for some  $k = 2, \dots, K^i$ .

For each other player  $j$ , where  $j \neq i$ , there is a set of  $j$ 's actions that player  $i$  considers infinitely unlikely. This set consists of those actions  $a^j \in A^j$  assigned probability 0 by (the marginal on  $A^j$  of)  $p_1^i$ . Similarly, there is a set of  $j$ 's actions that  $i$  is almost sure contains  $j$ 's choice, namely the set of actions  $a^j \in A^j$  assigned positive probability by  $p_1^i$ . I will say that player  $i$  first-order knows that player  $j$ 's choice lies in this latter set.

So far, no mention has been made of what player  $i$  might know about each other player  $j$ 's rationality. I am going to assume that  $i$  first-order knows that  $j$  chooses an admissible action in the same manner as  $i$  does. That is, for each action  $a^j \in A^j$  to which  $i$  assigns positive first-order probability, there must be an LPS  $\rho^j = (p_1^j, \dots, p_{K^j}^j)$ , where each  $p_k^j$  ( $k = 1, \dots, K^j$ ) is a probability measure on  $A^{-j}$ , that has full support in a sense analogous to definition 1 (for each  $a^{-j} \in A^{-j}$ ,  $p_k^j(a^{-j}) > 0$  for some  $k = 1, \dots, K^j$ ) and is such that

$$\left[ \sum_{a^{-j} \in A^{-j}} p_k^j(a^{-j}) u^j(a^j, a^{-j}) \right]_{k=1}^{K^j} \geq_L \left[ \sum_{a^{-j} \in A^{-j}} p_k^j(a^{-j}) u^j(b^j, a^{-j}) \right]_{k=1}^{K^j}$$

for all  $b^j \in A^j$ . But we also want to assume that player  $i$  first-order knows

that player  $j$  first-order knows that each player chooses an admissible action, and so on. That is, we want it to be *common first-order knowledge* that the players choose admissible actions. In the following I say that an LPS  $\rho^i = (p_1^i, \dots, p_{k^i}^i)$  justifies an action  $a^i$  of player  $i$  if

$$\left[ \sum_{a^{-i} \in A^{-i}} p_k^i(a^{-i}) u^i(a^i, a^{-i}) \right]_{k=1}^{K^i} \geq_L \left[ \sum_{a^{-i} \in A^{-i}} p_k^i(a^{-i}) u^i(b^i, a^{-i}) \right]_{k=1}^{K^i}$$

for all  $b^i \in A^i$ .

**DEFINITION 2** An action  $a^i$  of player  $i$  is *permissible* if there is an LPS  $\rho^i$  of player  $i$  that has full support and justifies  $a^i$ , if for each  $j \neq i$  and every  $a^j \in A^j$  assigned positive first-order probability by  $\rho^i$  there is an LPS  $\rho^j$  of player  $j$  that has full support and justifies  $a^j$ , and so on.

For each  $i = 1, \dots, n$ , let  $P^i \subset A^i$  be the set of permissible actions of player  $i$ . Then, if it is common first-order knowledge that the players choose admissible actions, each player  $i$  selects an action from  $P^i$ .

The main result of this paper relates the sets  $P^i$  to the sets of actions that survive the following elimination procedure. To define the procedure, a shorthand will be useful. Given the original game  $\Gamma = \langle A^1, \dots, A^n, u^1, \dots, u^n \rangle$  and subsets  $S^i \subset A^i$  ( $i = 1, \dots, n$ ), the reduced game,

$$\langle S^1, \dots, S^n; v^1, \dots, v^n \rangle,$$

where, for each  $i$ ,  $v^i$  is the restriction to  $S^1 \times \dots \times S^n$  of  $u^i$ , will be referred to simply as "the game  $S^1 \times \dots \times S^n$ ." Now, for each  $i = 1, \dots, n$ , define sets

$$W^i = \{a^i \in A^i : a^i \text{ is not weakly dominated in the game } A^1 \times \dots \times A^n\}$$

and

$$S_1^i = \{a^i \in W^i : a^i \text{ is not strongly dominated in the game } W^1 \times \dots \times W^n\},$$

and, for  $m = 2, 3, \dots$ , define inductively

$$S_m^i = \{a^i \in S_{m-1}^i : a^i \text{ is not strongly dominated in the game}$$

$$S_{m-1}^1 \times \dots \times S_{m-1}^n\}.$$

Finiteness of the original sets of actions  $A^i$  immediately implies the existence of an  $M$  such that, for all  $i = 1, \dots, n$ ,

