Int. Journal of Game Theory, Vol. 7, Issue 2, page 73–80. Physica-Verlag, Vienna.

Refinements of the Nash Equilibrium Concept

By *R.B. Myerson*, Evanston[1][2]

Abstract. Selten's concept of perfect equilibrium for normal form games is reviewed, and a new concept of proper equilibrium is defined. It is shown that the proper equilibria form a nonempty subset of the perfect equilibria, which in turn form a subset of the Nash equilibria. An example is given to show that these inclusions may be strict.

1. Introduction

The concept of equilibrium, as defined by *Nash* [1951], is one of the most important and elegant ideas in game theory. Unfortunately, a game can have many Nash equilibria, and some of these equilibria may be inconsistent with our intuitive notions about what should be the outcome of a game. To reduce this ambiguity and to eliminate some of these counterintuitive equilibria, *Selten* [1975] introduced the concept of a *perfect equilibrium*. In this paper, we shall define the notion of a *proper equilibrium,* to further refine the equilibrium concept. We will show that, for any game, the proper equilibria form a nonempty subset of *Selten's* perfect equilibria, which are themselves a subset of the Nash equilibria.

To see how these counterintuitive equilibria can arise, consider the game in Figure 1.

Γ_1:

		Player 2: β_1	β_2
Player 1	α_1	1, 1	0, 0
	α_2	0, 0	0, 0

Fig. 1

[1]) The author acknowledges helpful conversations with Ehud Kalai and David Schmeidler.

[2]) Prof. *R.B. Myerson,* Graduate School of Management, Northwestern University, Nathaniel Leverone Hall, Evanston, Illinois 60201, USA.

There are two Nash equilibria in this game, (α_1, β_1) and (α_2, β_2), because in each case neither player can improve his payoff by unilaterally changing his strategy. But it would be unreasonable to predict (α_2, β_2) as the outcome of this game. If player 1 thought that there was any chance of player 2 using β_1, then player 1 would certainly prefer α_1. Thus (α_2, β_2) qualifies as an equilibrium only because Nash's definition presumes that a player will ignore all parts of the payoff matrix corresponding to opponents strategies which are given zero probability. The essential idea behind *Selten*'s perfect equilibria is that no strategy should ever be given zero probability, since there is always a small chance that any strategy might be chosen, if only by mistake. So, in our example, α_1 and β_1 always must get at least an infinitesimal probability weight, which will eliminate (α_2, β_2) from the class of perfect (and proper) equilibria.

2. Normal Form Games and Nash Equilibria

Although *Selten* [1975] initially developed his theory of perfect equilibria for extensive from games, we will limit our attention in this paper to normal form games. Γ is an n-person game in *normal form* if

$$\Gamma = (S_1, \ldots, S_n; U_1, \ldots, U_n) \tag{1}$$

where each S_i is a nonempty finite set, and each U_i is a real-valued function defined on the domain $S_1 \times S_2 \times \ldots \times S_n$. We interpret $\{1, 2, \ldots, n\}$ as the set of *players* in the game. For each player i, S_i is the set of *pure strategies* which are available to player i. Each U_i is the utility function for player i, so that $U_i(s_1, \ldots, s_n)$ would be the payoff to player i (measured in some von Neumann-Morgenstern utility scale) if $(s_1, \ldots, s_n)$ were the combination of strategies chosen by the players.

For any finite set M, let $\Delta(M)$ be the set of all probability distributions over M. Thus:

$$\Delta(S_i) = \{\sigma_i \in R^{S_i} \mid \sum_{s_i \in S_i} \sigma_i(s_i) = 1,\ \sigma_i(s_i') \geq 0 \quad \forall s_i' \in S_i\}. \tag{2}$$

So $\Delta(S_i)$ is the set of *randomized* or *mixed* strategies which player i could choose in Γ.

It is straightforward to extend the utility functions to the mixed strategies, using the formula:

$$U_j(\sigma_1, \ldots, \sigma_n) = \sum_{(s_1, \ldots, s_n) \in S_1 \times \ldots \times S_n} \left(\prod_{i=1}^{n} \sigma_i(s_i)\right) U_j(s_1, \ldots, s_n). \tag{3}$$

That is, $U_j(\sigma_1, \ldots, \sigma_n)$ is the expected utility which player j would get if each player i planned to independently randomize his strategy according to σ_i.

Suppose that $(\sigma_1, \ldots, \sigma_n) \in \mathop{\times}_{i=1}^{n} \Delta(S_i)$ is the combination of mixed strategies which the players are expected to use, but suppose that player j is considering whether to switch to one of his pure strategies $s_j^* \in S_j$. Let $V_j(s_j^* \mid \sigma_1, \ldots, \sigma_n)$ be the expected utility for j if he makes this switch and all others remain with their σ_i mixed strategies, so that:

$$V_j(s_j^* \mid \sigma_1, \ldots, \sigma_n) = U_j(\sigma_1, \ldots, \sigma_{j-1}, \sigma_j^*, \sigma_{j+1}, \ldots, \sigma_n) \tag{4}$$

$$\text{where } \sigma_j^*(s_j) = \begin{cases} 1 & \text{if } s_j = s_j^* \\ 0 & \text{if } s_j \neq s_j^*. \end{cases}$$

We say that s_j is a *best response* (in pure strategies) to $(\sigma_1, \ldots, \sigma_n)$ for player j iff

$$V_j(s_j \mid \sigma_1, \ldots, \sigma_n) = \max_{s_j' \in S_j} V_j(s_j' \mid \sigma_1, \ldots, \sigma_n).$$

A combination of mixed strategies $(\sigma_1, \ldots, \sigma_n)$ is a *Nash equilibrium* if no player can gain by unilaterally switching to any other mixed strategy. That is,

$(\sigma_1, \ldots, \sigma_n) \in \underset{i=1}{\overset{n}{\times}} \Delta(S_i)$ is a Nash equilibrium iff

$$U_j(\sigma_1, \ldots, \sigma_n) \geqslant U_j(\sigma_1, \ldots, \sigma_j', \ldots, \sigma_n) \quad \forall j, \forall \sigma_j' \in \Delta(S_j). \tag{5}$$

It is well known that a combination of mixed strategies forms a Nash equilibrium if and only if every player gives positive probability only to his pure strategies which are best responses for him. The following proposition states this fact in terms which will be most convenient for our purposes.

Proposition 1: $(\sigma_1, \ldots, \sigma_n)$ is a Nash equilibrium iff:
$(\sigma_1, \ldots, \sigma_n) \in \Delta(S_1) \times \ldots \times \Delta(S_n)$, and if $V_j(s_j \mid \sigma_1, \ldots, \sigma_n) < V_j(s_j' \mid \sigma_1, \ldots, \sigma_n)$ then $\sigma_j(s_j) = 0$, $\forall j, \forall s_j \in S_j, \forall s_j' \in S_j$.

To check this proposition, observe that

$$U_j(\sigma_1, \ldots, \sigma_n) - U_j(\sigma_1, \ldots, \sigma_j', \ldots, \sigma_n)$$

$$= \sum_{s_j \in S_j} \sum_{s_j' \in S_j} \sigma_j(s_j)\, \sigma_j'(s_j')\, (V_j(s_j \mid \sigma_1, \ldots, \sigma_n) - V_j(s_j' \mid \sigma_1, \ldots, \sigma_n)).$$

The proposition then follows easily from the definition of a Nash equilibrium.

3. Perfect Equilibria

In this section we review Selten's concept of *perfect equilbrium*, using a new approach which will be more convenient for our purposes.

For any finite set M, let Δ^0 (M) be the set of all probability distributions on M which give positive probability weight to all members of M. So, for any player i,

$$\Delta^0 (S_i) = \{\sigma_i \in R^{S_i} \mid \sum_{s_i \in S_i} \sigma_i (s_i) = 1, \sigma_i (s_i') > 0 \quad \forall s_i' \in S_i\}. \tag{6}$$

The members of Δ^0 (S_i) are called *totally mixed* strategies for player i.

Heuristically, a perfect equilibrium could be described as a combination of totally mixed strategies $(\sigma_1, \ldots, \sigma_n)$ such that, for any player j and any pure strategy $s_j \in S_j$, if s_j is not a best response to $(\sigma_1, \ldots, \sigma_n)$ for j then $\sigma_j (s_j)$ should be infinitesimally small. Thus, in a perfect equilibrium a player cannot ignore any point in $S_1 \times \ldots \times S_n$ as "impossible" when he computes his best responses, since every one of his opponents' pure strategies has a positive probability; and yet no player wants to make any substantial change in his strategy, since each player is already putting almost all his probability weight on his best responses.

To make these ideas about "infinitesimally small" probabilites precise, let ϵ be any small positive number. Then we define an *ϵ-perfect equilibrium* to be any combination of totally mixed strategies $(\sigma_1, \ldots, \sigma_n) \in \Delta^0$ $(S_1) \times \ldots \times \Delta^0$ (S_n) such that:

$$\begin{aligned} &\text{if } V_j (s_j \mid \sigma_1, \ldots, \sigma_n) < V_j (s_j' \mid \sigma_1, \ldots, \sigma_n) \text{ then } \sigma_j (s_j) \leqslant \epsilon, \\ &\forall j, \forall s_j \in S_j, \forall s_j' \in S_j. \end{aligned} \tag{7}$$

So an ϵ-perfect equilibrium is a combination of mixed strategies such that every pure strategies gets a positive probability, but only best-response strategies get more than ϵ probability.

A perfect equilibrium is then defined to be any limit of ϵ-perfect equilibria. That is, $(\sigma_1, \ldots, \sigma_n) \in \Delta (S_1) \times \ldots \times \Delta (S_n)$ is a *perfect equilibrium* iff there exist some sequences $\{\epsilon_k\}_{k=1}^{\infty}$ and $\{(\sigma_1^k, \ldots, \sigma_n^k)\}_{k=1}^{\infty}$ such that:

$$\text{each } \epsilon_k > 0 \text{ and } \lim_{k \to \infty} \epsilon_k = 0, \tag{8a}$$

$$\text{each } (\sigma_1^k, \ldots, \sigma_n^k) \text{ is an } \epsilon_k\text{-perfect equilibrium, and} \tag{8b}$$

$$\lim_{k \to \infty} \sigma_i^k (s_i) = \sigma_i (s_i), \text{ for all } i \text{ and all } s_i \in S_i. \tag{8c}$$

(Notice that every σ_i^k will have to be in Δ^0 (S_i), but σ_i need not be, since the closure of Δ^0 (S_i) is Δ (S_i).)

Selten has shown that a perfect equilibrium must be a Nash equilibrium [see Lemma 9 in *Selten*]. The proof follows easily from the fact that $V_j (s_j \mid \sigma_1, \ldots, \sigma_n)$ is continuous in $(\sigma_1, \ldots, \sigma_n)$. So a perfect equilibrium must satisfy the conditions of our Proposition 1, because it is the limit of ϵ-perfect equilbria which satisfy (7).

All perfect equilibria are Nash equilibria, but the converse does not hold. For example, consider the game Γ_1 defined in Figure 1. The only ϵ-perfect equilibria are those pairs of totally mixed strategies in which σ_1 $(\alpha_2) \leqslant \epsilon$ and σ_2 $(\beta_2) \leqslant \epsilon$. Thus the only perfect equilibrium is (α_1, β_1) – or, more precisely, it is (σ_1^*, σ_2^*), where $\sigma_1^*(\alpha_1) = 1$ and $\sigma_2^*(\beta_1) = 1$. (α_2, β_2) was a Nash equilibrium, but it is not a perfect equilibrium.

4. Proper Equilibria

Consider now the following simple example:

Γ_2 Player 2

(U_1, U_2)		β_1	β_2	β_3
	α_1	1, 1	0, 0	–9, –9
Player 1	α_2	0, 0	0, 0	–7, –7
	α_3	–9,–9	–7,–7	–7, –7

Fig. 2

As in our first example, (α_1, β_1) would seem like the obvious outcome for this game. There are three Nash equilibria, and all are in pure strategies (or, more precisely, in mixed strategies which assign all weight to one pair of pure strategies); these equilibria are (α_1, β_1), (α_2, β_2), and (α_3, β_3).

Of these three Nash equilibria, (α_3, β_3) is not perfect, but (α_1, β_1) and (α_2, β_2) are both perfect equilibria. To check that (α_2, β_2) is a perfect equilibrium, define σ_1^ϵ and σ_2^ϵ by

$$\sigma_1^\epsilon(\alpha_1) = \epsilon, \sigma_1^\epsilon(\alpha_2) = 1 - 2\epsilon, \sigma_1^\epsilon(\alpha_3) = \epsilon$$

$$\sigma_2^\epsilon(\beta_1) = \epsilon, \sigma_2^\epsilon(\beta_2) = 1 - 2\epsilon, \sigma_2^\epsilon(\beta_3) = \epsilon,$$

and observe that $(\sigma_1^\epsilon, \sigma_2^\epsilon)$ forms an ϵ-perfect equilibrium. (For example $V_1(\alpha_1 \mid \sigma_1^\epsilon, \sigma_2^\epsilon) = -8\epsilon$, $V_1(\alpha_2 \mid \sigma_1^\epsilon, \sigma_2^\epsilon) = -7\epsilon$ and $V_1(\alpha_3 \mid \sigma_1^\epsilon, \sigma_2^\epsilon) = -7 - 2\epsilon$, so α_2 is best response. As required, $\sigma_1^\epsilon(\alpha_1) \leqslant \epsilon$ and $\sigma_1^\epsilon(\alpha_3) \leqslant \epsilon$.) Then, as $\epsilon \to 0$, these σ_1^ϵ and σ_2^ϵ converge to the strategies which select α_2 and β_2 with probability 1. In effect, adding the α_3 row and β_3 column has converted (α_2, β_2) into a perfect equilibrium even though α_3 and β_3 are obviously dominated strategies.

To discriminate between (α_1, β_1) and (α_2, β_2) in this example, we need a new refinement of the equilibrium concept: the proper equilibrium.

For our general normal form game, we define an *ϵ-proper equilibrium* to be any combination of totally mixed strategies $(\sigma_1, \ldots, \sigma_n) \in \Delta^0(S_1) \times \ldots \times \Delta^0(S_n)$ such that:

$$\text{if} \quad V_j(s_j \mid \sigma_1, \ldots, \sigma_n) < V_j(s_j' \mid \sigma_1, \ldots, \sigma_n) \text{ then } \sigma_j(s_j) \leqslant \epsilon \cdot \sigma_j(s_j'),$$
$$\forall j, \forall s_j \in S_j, \forall s_j' \in S_j. \tag{9}$$

So an ϵ-proper equilibrium is a combination of totally mixed strategies in which every player is giving his better responses much more probability weight than his worse responses (by a factor $1/\epsilon$), whether or not those "better" responses are "best".

It is easy to check that an ϵ-proper equilibrium must be ϵ-perfect (since $\epsilon \cdot \sigma_j(s_j') \leqslant \epsilon$ in (9)), but the converse does not hold. In the example above, $(\sigma_1^\epsilon, \sigma_2^\epsilon)$ is not ϵ-proper because, for $0 < \epsilon < 1$, $V_1(\alpha_3 \mid \sigma_1^\epsilon, \sigma_2^\epsilon) < V_1(\alpha_1 \mid \sigma_1^\epsilon, \sigma_2^\epsilon)$ but $\sigma_1(\alpha_3) > \epsilon \cdot \sigma_1(\alpha_1)$.

We now define a proper equilibrium to be any limit of ϵ-proper equilbria. That is $(\sigma_1, \ldots, \sigma_n) \in \Delta(S_1) \times \ldots \times \Delta(S_n)$ is a *proper equilibrium* iff there exist some sequences $\{\epsilon_k\}_{k=0}^{\infty}$ and $\{(\sigma_1^k, \ldots, \sigma_n^k)\}_{k=0}^{\infty}$ such that:

$$\text{each } \epsilon_k > 0 \text{ and } \lim_{k\to\infty} \epsilon_k = 0 \tag{10a}$$

$$\text{each } (\sigma_1^k, \ldots, \sigma_n^k) \text{ is an } \epsilon_k\text{-proper equilibrium, and} \tag{10b}$$

$$\lim_{k\to\infty} \sigma_i^k(s_i) = \sigma_i(s_i), \text{ for all } i \text{ and all } s_i \in S_i. \tag{10c}$$

As with perfect equilibria, a proper equilibrium need not be totally mixed; it must only be the limit of totally mixed ϵ-proper equilibria.

Proposition 2: For any game in normal form, the proper equilibria form a subset of the perfect equilbria, which in turn form a subset of the Nash equilibria. These inclusions may both be strict inclusions.

Proof: We already remarked that a perfect equilibrium must be a Nash equilibrium, and that an ϵ-proper equilibrum must be an ϵ-perfect equilibrium. So a proper equilibrium, as a limit of ϵ-proper equilibria, is also a limit of ϵ-perfect equilibria, and is therefore perfect.

The example in Figure 2 shows that these inclusions may be strict, since this game has three Nash equilibria $((\alpha_1, \beta_1), (\alpha_2, \beta_2)$, and $(\alpha_3, \beta_3))$, but only two perfect equilibria $((\alpha_1, \beta_1)$ and $(\alpha_2, \beta_2))$, and only one proper equilibrium, (α_1, β_1). To verify that (α_1, β_1) is the only proper equilibrium, suppose $0 < \epsilon < 1$, and let (σ_1, σ_2) be an ϵ-proper equilibrium. Since α_2 dominates α_3, and σ_2 is totally mixed, we have $V_1(\alpha_3 \mid \sigma_1, \sigma_2) < V_1(\alpha_2 \mid \sigma_1, \sigma_2)$, and so $\sigma_1(\alpha_3) \leqslant \epsilon \cdot \sigma_1(\alpha_2)$. This implies that $V_2(\beta_3 \mid \sigma_1, \sigma_2) < V_2(\beta_1 \mid \sigma_1, \sigma_2)$, so $\sigma_2(\beta_3) \leqslant \epsilon \cdot \sigma_2(\beta_1)$. This in turn implies that $V_1(\alpha_2 \mid \sigma_1, \sigma_2) < V_1(\alpha_1 \mid \sigma_1, \sigma_2)$, so $\sigma_1(\alpha_2) \leqslant \epsilon \cdot \sigma_1(\alpha_1)$. So $\sigma_1(\alpha_2) \leqslant \epsilon \cdot \sigma_1(\alpha_1) \leqslant \epsilon$ and $\sigma_1(\alpha_3) \leqslant \epsilon \cdot \sigma_1(\alpha_2) \leqslant \epsilon^2$. A similar argument shows that $\sigma_2(\beta_2) \leqslant \epsilon$ and $\sigma_2(\beta_3) \leqslant \epsilon^2$. Since the probabilities must sum to 1, $\sigma_1(\alpha_1) \geqslant 1 - \epsilon - \epsilon^2$ and $\sigma_2(\beta_1) \geqslant 1 - \epsilon - \epsilon^2$. As $\epsilon \to 0$, our ϵ-proper equilibria must converge to the mixed strategies which select α_1 and β_1 with probability 1. Thus, although (α_2, β_2) is perfect for this game, it is not proper.

5. Existence of Proper Equilibria

To be useful, a refinement of the Nash equilibrium concept should generate a nonempty set of Nash equilibria for any game. We have already shown that our proper eqr:ilibria do form a subset fo the Nash equilibria. The following theorem assures us that the set of proper equilibria is also nonempty.

Theorem: For any normal form game Γ (as in (1)), there exists at least one proper equilibrium.

Proof: We show first that there exists an ϵ-proper equilibrium, for any ϵ, $0 < \epsilon < 1$.

Let $m = \max_i |S_i|$. Given ϵ, let $\delta = \frac{1}{m} \cdot \epsilon^m$. For any player j, let

$$\Delta^* (S_j) = \{\sigma_j \in \Delta (S_j) \mid \sigma_j (s_j) \geqslant \delta, \forall\, s_j \in S_j\}.$$

Observe that $\Delta^* (S_j)$ is a nonempty compact subset of $\Delta^0 (S_j)$. We now define a point-to-set map $F_j : \mathop{\times}\limits_{i=1}^{n} \Delta^* (S_i) \Rightarrow \Delta^* (S_j)$ by:

$$F_j (\sigma_1, \ldots, \sigma_n) = \left\{ \sigma_j^* \in \Delta^* (S_j) \;\middle|\; \begin{array}{l} \text{if } V_j (s_j \mid \sigma_1, \ldots, \sigma_n) < V_j (s_j' \mid \sigma_1, \ldots, \sigma_n) \\ \text{then } \sigma_j^* (s_j) \leqslant \epsilon \cdot \sigma_j^* (s_j'), \\ \forall s_j \in S_j, A\, s_j' \in S_j . \end{array} \right\}$$

For any $(\sigma_1, \ldots, \sigma_n)$ the points in $F_j (\sigma_1, \ldots, \sigma_n)$ satisfy a finite collection of linear inequalities, so $F_j (\sigma_1, \ldots, \sigma_n)$ is a closed convex set. To check that $F_j (\sigma_1, \ldots, \sigma_n)$ is nonempty, let $\rho (s_j)$ be the number of pure strategies $s_j' \in S_j$ such that $V_j (s_j \mid \sigma_1, \ldots, \sigma_n) < V_j (s_j' \mid \sigma_1, \ldots, \sigma_n)$; then letting $\sigma_j^* (s_j) = \epsilon^{\rho(s_j)} / (\sum_{s_j' \in S_j} \epsilon^{\rho(s_j')})$ will give us $\sigma_j^* \in F_j\ (\sigma_1, \ldots, \sigma_n)$. (Observe that $\sigma_j^* (s_j) \geqslant \epsilon^m/m$, so $\sigma_j^* \in \Delta^* (S_j)$.) Finally, continuity of each $V_j (s_j \mid \cdot)$ function implies that $F_j (\cdot)$ must be upper-semicontinuous.

Let $F (\cdot) = \mathop{\times}\limits_{j=1}^{n} F_j (\cdot)$. Then $F: \mathop{\times}\limits_{i=1}^{n} \Delta^* (S_i) \Rightarrow \mathop{\times}\limits_{i=1}^{n} \Delta^* (S_i)$ satisfies all the conditions of the *Kakutani* Fixed Point Theorem [*Kakutani*], so there exists some $(\sigma_1^\epsilon, \ldots, \sigma_n^\epsilon) \in \mathop{\times}\limits_{i=1}^{n} \Delta^* (S_i)$ such that $(\sigma_1^\epsilon, \ldots, \sigma_n^\epsilon) \in F (\sigma_1^\epsilon, \ldots, \sigma_n^\epsilon)$. This $(\sigma_1^\epsilon, \ldots, \sigma_n^\epsilon)$ is clearly an ϵ-proper equilibrium.

So for any $0 < \epsilon < 1$ there exists an ϵ-proper equilibrium $(\sigma_1^\epsilon, \ldots, \sigma_n^\epsilon)$. Since $\mathop{\times}\limits_{i=1}^{n} \Delta (S_i)$ is a compact set there must exist a convergent subsequence and a proper equilibrium $(\sigma_1, \ldots, \sigma_n) = \lim_{\epsilon \to 0} (\sigma_1^\epsilon, \ldots, \sigma_n^\epsilon)$.

References

Kakutani, S.: A Generalization of Brouwer's Fixed Point Theorem. Duke Mathematical Journal 8, 1941, 457–459.

Nash, J.F.: Non-Cooperative Games. Annals of Mathematics 54, 1951, 286–295.

Selten, R.: Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games. International Journal of Game Theory 4, 1975, 25–55.

Received September, 1977