Nash Equilibrium

DAVID M. KREPS

The concept of a *Nash equilibrium* plays a central role in noncooperative game theory. Due in its current formalization to John Nash (1950, 1951), it goes back at least to Cournot (1838). This entry begins with the formal definition of a Nash equilibrium and with some of the mathematical properties of equilibria. Then we ask: To what question is 'Nash equilibrium' the answer? The answer that we suggest motivates further questions of *equilibrium selection*, which we consider in two veins: the informal notions, such as Schelling's (1960) *focal points*; and the formal theories for *refining* or *perfecting* Nash equilibria, due largely to Selten (1965, 1975). We conclude with a brief discussion of two related issues: Harsanyi's (1967–8) notion of a *game of incomplete information* and Aumann's (1973) *correlated equilibria.*

I. DEFINITION AND SIMPLE MATHEMATICAL PROPERTIES. We give the definition in the simple setting of a finite player and action game in normal form. There are I players, indexed by $i = 1, \ldots, I$. Player i chooses from N_i (pure) strategies; we write S_i for this set of strategies, and s_i for a typical member of S_i. A *strategy profile*, written $s = (s_1, \ldots, s_I)$, is a vector of strategies for the individual players – we write S for $\Pi_{i=1}^{I} S_i$, the set of all strategy profiles. For a strategy profile $s = (s_1, \ldots, s_I) \in S$ and a strategy $s'_i \in S_i$ for player i, we write $s|s'_i$ for the strategy profile $(s_1, \ldots, s_{i-1}, s'_{i+1}, \ldots, s_I)$ or s with the part of i changed from s_i to s'_i. For each player i and strategy profile $s, u_i(s)$ denotes i's expected utility or payoff if players employ strategy profile s.

Definition. A *Nash equilibrium* (in pure strategies) is a strategy profile s such that for each i and $s'_i \in S_i, u_i(s) \geqslant u_i(s|s'_i)$. In words, no single player, by changing his own part of s, can obtain higher utility if the others stick to their parts.

This basic definition is often extended to independently mixed strategy profiles, as follows. Given S_i, write Σ_i for the set of mixed strategies for player i; that is,

J. Eatwell et al. (eds.), *Game Theory*

all probability distributions over S_i. Write Σ for $\Pi_{i=1}^{I}\Sigma_i, \sigma=(\sigma_1,\ldots,\sigma_I), \cdot\sigma|\sigma_i'$, and so on, as before. Extend the utility functions u_i from domain S to domain Σ by letting $u_i(\sigma)$ be player i's expected utility:

$$u_i(\sigma)=\sum_{S_1}\ldots\sum_{S_I}u_i(s_1,\ldots,s_I)\sigma_1(s_1)\ldots\sigma_I(s_I).$$

Then define a Nash equilibrium in mixed strategies just as above, with σ in place of s and σ_i in place of s_i. Equivalently, player i puts positive weight on pure strategy s_i only if s_i is among the pure strategies that give him the greatest expected utility.

This formal concept is due to John Nash (1950, 1951). Luce and Raiffa (1957) provided an important and influential early commentary. Nash also proved that in a finite player and finite action game, there always exist at least one Nash equilibrium, albeit existence can only be guaranteed if we look at mixed strategies – standard examples (such as matching pennies) show that there are games with no pure strategy equilibria. The proof that a Nash equilibrium always exists is an application of Brouwer's fixed point theorem. The concept of a Nash equilibrium is extended in natural fashion to games with infinitely many players and/or pure strategies, although in such cases existence can be problematic; we do not discuss these matters further here.

II. THE PHILOSOPHY OF NASH EQUILIBRIUM. To what question is 'Nash equilibrium' the answer? This has been and continues to be the subject of much discussion and debate. Most authors take a position that is a variation on the following.

Suppose that, in a particular game, players *by some means unspecified at the moment* arrive at an 'agreement' as to how each will play the game. This 'agreement' specifies a particular strategy choice by each player, and each player is aware of the strategies chosen by each of his fellow players, although players may not resort to enforcement mechanisms except for those given as part of the formal specification of the game. One would not consider this agreement *self-enforcing* (or strategically stable) if some one of the players, hypothesizing that others will keep to their parts of the agreement, would prefer to deviate and choose some strategy other than that specified in the agreement. Thus, to be self-enforcing in this sense, it is *necessary* that the agreement form a Nash equilibrium. (If players could perform a public randomization as part of the agreement, we would get convex combinations of Nash equilibria as candidate self-enforcing agreements. See section VI for what can be done with partially private randomizations.)

This does not say that every Nash equilibrium is a self-enforcing agreement. For example, in the context being modelled, it might be appropriate to consider multi-player defections (and the concept of a *strong equilibrium*, a strategy assignment in which no coalition can profitably deviate, then comes into play). It does not say how this agreement comes about, nor what will transpire if there is no agreement. Indeed, in the latter case the concept of a Nash equilbrium has no particular claim upon us.