FRANÇOISE FORGES

FIVE LEGITIMATE DEFINITIONS OF CORRELATED EQUILIBRIUM IN GAMES WITH INCOMPLETE INFORMATION

ABSTRACT. Aumann's (1987) theorem shows that correlated equilibrium is an expression of Bayesian rationality. We extend this result to games with incomplete information.

First, we rely on Harsanyi's (1967) model and represent the underlying multiperson decision problem as a fixed game with imperfect information. We survey four definitions of correlated equilibrium which have appeared in the literature. We show that these definitions are not equivalent to each other. We prove that one of them fits Aumann's framework; the 'agents normal form correlated equilibrium' is an expression of Bayesian rationality in games with incomplete information.

We also follow a 'universal Bayesian approach' based on Mertens and Zamir's (1985) construction of the 'universal beliefs space'. Hierarchies of beliefs over independent variables (states of nature) and dependent variables (actions) are then constructed simultaneously. We establish that the universal set of Bayesian solutions satisfies another extension of Aumann's theorem.

We get the following corollary: once the types of the players are not fixed by the model, the various definitions of correlated equilibrium previously considered are equivalent.

Keywords: correlated equilibrium, Bayesian rationality, incomplete information games

1. INTRODUCTION

The *correlated equilibrium* is a solution concept in strategic form games. It was introduced by Aumann (1974). The correlated equilibria of a given game are usually defined as Nash equilibria of some extension of this game, where the players receive private signals before the beginning of the original game. In a second article devoted to the subject, Aumann (1987) has shown that the correlated equilibrium could be defined without any reference to Nash equilibrium or even to the game-theoretical apparatus.

Aumann's (1987) study provides a bridge between game theory and (individual) decision theory. In a game, every player faces a decision

Theory and Decision **35**: 277–310, 1993.

problem where the actions of his opponents appear as states of nature. Obviously, every player realizes that the others proceed as himself and form beliefs about his behavior. This generates infinite hierarchies of beliefs, which constitute the 'space of all states of the world'. According to Mertens and Zamir (1985), this is a well-defined mathematical object (see below). Armbruster and Böge (1979) and Böge and Eisele (1979) also pioneered in this area. More recently, articles have been devoted to the study of the (Bayesian) foundations of game theory (see e.g. Bernheim, 1986; Tan and Werlang, 1988, and the survey of Brandenburger and Dekel, 1990).

Aumann (1987) formulates three assumptions. First, the players share a common prior over the space of all states of the world (this is referred to as 'consistency'). Second, every player knows his own action. Third, every player is Bayes rational; namely, he maximizes his expected utility given his information. Aumann (1987) proves that under these assumptions, the players' actions follow a correlated equilibrium distribution.

Aumann's (1987) framework is *complete information*. As we have already pointed out, in this context, the uncertainty of every player only concerns the other players' actions and the subsequent hierarchies of beliefs. Formally, the space of all states of the world is a 'consistent beliefs subspace' of the 'universal beliefs space' which can be constructed on the parameters space of players' *actions* (see Mertens and Zamir, 1985, and Mertens, Sorin and Zamir, forthcoming, Ch. 3). Mertens and Zamir's (1985) construction was motivated by an apparently unrelated problem, namely the formalization of Harsanyi's (1967) notion of 'type'. Harsanyi (1967) introduced this concept as well as those of 'Bayesian players' and 'consistency' in the context of *incomplete information*.[1]

The main goal of the present paper is to extend Aumann's (1987) theorem in games with incomplete information, as they were formalized by Harsanyi (1967–1968) and Mertens and Zamir (1985). We face two difficulties. First, the literature contains several different definitions of correlated equilibrium in the model usually referred to as 'games with incomplete information'. We shall survey four such legitimate definitions and check whether they fit Aumann's (1987) framework. Another difficulty is that these definitions cannot be immediately reconciled with a simultaneous construction of beliefs hierarchies over states of nature (as in Harsanyi, 1967) and over players' actions (as in Aumann, 1987).

The expression 'games with incomplete information' is often misleading. *Stricto sensu*, a game with incomplete information is *not a game*. However, as Harsanyi (1967) has shown, such a situation is 'equivalent' to a game with *imperfect information*, which can be analyzed by the standard methods of game theory. This approach is much guided by the solution concept to be applied, namely Nash equilibrium. In particular, the normal form, the extensive form and the agents normal form of the game with imperfect information modeling the multiperson decision problem with incomplete information are all equivalent. It is well-known that these models may not yield the same 'refined' Nash equilibria (see e.g. Fudenberg and Tirole, 1992, Myerson, 1991). However, in the case of correlated equilibrium, it is even not clear that any of these games is an adequate representation of the decision problem.

The study of correlated equilibrium in 'games with incomplete information' will show that this model is fragile, because it allows various legitimate interpretations. Typically, player's types may result from different scenarios. The way in which the players' beliefs are generated may matter for correlated equilibrium, although it is irrelevant for Nash equilibrium. Let us illustrate this by a trivial example. If at some precise date, a referee performs a lottery and secretly transmits information to the players, this point in time can be taken as a reference. The players can receive additional ('extrinsic') private signals *before* or *after* this event. On the contrary, the players' beliefs may result from a complicated process of introspection so that no timing can be made precise. In this case, one can just formulate timeless axioms such as the consistency of beliefs.

In Section 2 we recall the standard game theoretical-approach to incomplete information. In Section 3 we present Aumann's (1974, 1987) results on correlated equilibrium in games with complete information. Proposition 2 states that correlated equilibrium is an expression of Bayesian rationality (Aumann's (1987) theorem).

In Section 4 we fix a game G with *imperfect information* in order to represent the underlying multiperson decision with incomplete information. In G, a virtual chance move selects the types of the players. G is also called a 'Bayesian game' (see e.g. Myerson, 1991). We analyze four solution concepts for G, all of which have been called 'correlated equilibrium' in the literature. Obviously, they all coincide with Aumann's concept in the particular case of complete information.

First, Aumann's (1974, 1987) definition can be applied to the *strategic form* of G. This solution concept was considered (e.g.) in Cotter (1991) and Forges (1985, 1986b). Second, Samuelson and Zhang (1989) and Forges (1986a) use Aumann's definition in the *agents normal form* of G. The resulting solution concept may be hard to justify intuitively (see Myerson, 1991, pp. 261–262). However, it is the appropriate one to extend Aumann's (1987) theorem. More precisely, if we impose the three assumptions recalled above together with natural coherence requirements (connecting the space of all states of the world with the given game G), we get exactly the agents normal form correlated equilibria as solutions. This result is stated as Proposition 3.

Myerson (1982, 1991) claims that Aumann's correlated equilibrium is adequately generalized by the so-called *communication equilibrium* in games with incomplete information. This solution concept has been extensively used (see e.g. Samuelson and Zhang, 1989; Einy and Peleg, 1991; Forges, 1985, 1986b, 1990).

The fourth approach consists of imposing Aumann's assumptions in a naive way, which amounts to using the players' beliefs over the state of nature, which are specified by G, as priors. These probability distributions may be revised as a function of the players' intended actions. The corresponding solutions can be interpreted as the result of recommendations from an omniscient mediator. They are closely related to the 'jointly coherent outcomes' studied in Nau (1992) (see also Nau and McCardle, 1990).

Section 5 contains examples which illustrate that the four solutions concepts are not equivalent, although payoffs associated with some of them can automatically be achieved through the others.

We have seen that Aumann's (1987) Bayesian analysis of games with complete information was formally similar to Harsanyi's (1967) model of incomplete information. In both cases, hierarchies of beliefs are constructed, consistency is required, private information is identified for every player. It seems thus natural to generalize Aumann's (1987) result by constructing a single set of hierarchies of beliefs for each player, without distinguishing dependent and independent variables. More precisely, in Section 6 we start with a multiperson decision problem with incomplete information, rather than with a game. The main difference between the two models is that the beliefs of the players and hence their types are not specified in the former one.

Thus let K be the set of states of nature, which parametrizes the possible games, and let A be the actions space (without loss of generality, A does not depend on the state of nature). We construct the universal beliefs space (in the sense of Mertens and Zamir, 1985) over $K \times A$. This space is then used to define the space of all states of the world. As in Aumann (1987), we impose conditions of consistency, knowledge and rationality. This yields the *universal set of Bayesian solutions* to the decision problem with incomplete information. We establish that every such Bayesian solution corresponds to a correlated equilibrium (or even a Nash equilibrium) of some game with incomplete information consistent with the basic decision problem. This statement holds for *any* of the four definitions of correlated equilibrium considered in Section 4. The converse is also true: any correlated equilibrium of any game with incomplete information consistent with the basic decision problem induces a Bayesian solution in the universal set. This result is stated as Proposition 4. It contains Aumann's (1987) theorem in the particular case of complete information.

Proposition 4 can obviously be interpreted as an *equivalence* theorem. As soon as the beliefs of the players are not fixed by a given game with incomplete information, all definitions of correlated equilibrium are equivalent. This suggests that the problem is not to identify the 'correct' definition of correlated equilibrium in games with incomplete information but rather to identify the 'correct' game with incomplete information, namely the 'correct' types which result from hierarchies of beliefs over states of nature *and actions*.

Harsanyi (1967) solved the problem of modeling games with incomplete information by distinguishing dependent and independent variables, which is fully justified in the game-theoretical paradigm based on the normative Nash equilibrium. Aumann (1987, point (a) of the Discussion) argued that this distinction did not make sense in a

Bayesian decision theoretical approach to games. We made a further step by pursuing Aumann's (1987) work in the context of incomplete information. At first sight, several different definitions of correlated equilibrium seem legitimate in this framework. But a more careful analysis shows that the main issue concerns the model of incomplete information itself. If a player can form beliefs over his opponents' information, he can also form beliefs over their actions and these beliefs can certainly interfere. The two-step model where types are fully clarified before decisions are made seems inadequate if games are to be played by Bayesian players.

2. FRAMEWORK

A *two person decision problem with incomplete information* is defined by

- a basic parameters space K,
- actions sets A_1, A_2 (for player 1 and player 2, respectively),
- utility functions u_i: $K \times A \rightarrow \mathbb{R}$ $(i = 1, 2)$ (where $A = A_1 \times A_2$).

For simplicity, we assume that K, A_1 and A_2 are finite sets. Our analysis should hold in more general models; the assumption on the number of players is only made for convenience of notation. We refer to the above decision problem as Γ_K.

If K is a singleton, Γ_K can be denoted as Γ and is just a *game in strategic form*. Otherwise, Γ_K does not correspond to any standard game-theoretical model, since it does not describe the information of the players, not their strategies. The variable $k \in K$ parametrizes the games that the two individuals under consideration may play; k is *independent* of the players' choices. The set A contains the *dependent* variables. As pointed out by Harsanyi (1967), there is no loss of generality in assuming that A is not indexed by $k \in K$ (see also Mertens, Sorin and Zamir, forthcoming, and Myerson, 1991).

An information scheme $\mathcal{S}_K$ on K consists of finite sets S_1 and S_2 together with a probability distribution P over $K \times S_1 \times S_2$. We write $\mathcal{S}_K = (S_1, S_2, P)$.

S_i will be interpreted as a set of signals to player i $(i = 1, 2)$. More precisely, Γ_K and $\mathscr{S}_K$ generate a game with a private signals described as follows:

stage 1: (k, s_1, s_2) is selected in $K \times S_1 \times S_2$ according to P; s_i is secretly transmitted to player i $(i = 1, 2)$.
stage 2: the players simultaneously choose an action $(a_i \in A_i, i = 1, 2)$.

The payoffs are described as in Γ_K, namely by $u_i(k, a_1, a_2)$. This game will be denoted as $G(\Gamma_K, \mathscr{S}_K)$. A pure strategy of player i in $G(\Gamma_K, \mathscr{S}_K)$ is a mapping

$$\sigma_i : S_i \to A_i \quad (i = 1, 2)$$

The *reduced form* of $G(\Gamma_K, \mathscr{S}_K)$ is obtained by 'forgetting' K in the above description. More precisely, (s_1, s_2) is selected according to the marginal distribution of P over $S_1 \times S_2$ and *expected* utility functions

$$v_i : S \times A \to \mathbb{R} \quad (i = 1, 2)$$

are considered (we set $S = S_1 \times S_2$):

$$\begin{aligned} v_i(s, a) &= E[u_i(k, a) \mid s] \\ &= \sum_k P(k \mid s) u_i(k, a) \end{aligned}$$

The game described by S_1, S_2, A_1, A_2, v_1 and v_2 has exactly the form of a *consistent Bayesian* game (see e.g. Myerson, 1991). S_i is then referred to as player i's set of *types*.

Before analyzing the relationships between games with private signals and Bayesian games, let us consider the particular case where K is a singleton. Any information scheme $\mathscr{S} = (S_1, S_2, P)$ is then independent of Γ and the private signals in S_1 and S_2 can be called 'extraneous'. $G(\Gamma, \mathscr{S})$ is an 'extension' of the original game Γ.

Assume now that K has at least two elements. We shall briefly recall Mertens and Zamir's (1985) construction, which formalizes ideas introduced by Harsanyi (1967) (see also Mertens, Sorin and Zamir, forthcoming, and Myerson, 1991). Given the basic parameters space K, each player i forms beliefs not only over K but also over the other player's beliefs over K, over the other player's beliefs over his own beliefs, and so on, which generates an infinite hierarchy of beliefs. Let Θ_i denote the space of all the beliefs hierarchies of player i ($i = 1, 2$). Mertens and Zamir (1985) have proved that Θ_i is a compact space and is homeomorphic to $\Delta(K \times \Theta_j)$ ($i, j = 1, 2, i \neq j$), where $\Delta(X)$ denotes the space of all probability distribution over a space X ($\Delta(X)$ is endowed with the weak* topology). We write this as

$$\Theta_i \simeq \Delta(K \times \Theta_j) .$$

$\Omega = K \times \Theta_1 \times \Theta_2$ is the universal beliefs space constructed from K and Θ_i is the universal types space of player i. The above result states that as emphasized by Harsanyi (1967), player i's type can be viewed as a belief over the basic parameter and the other player's type. Once the universal space is constructed, hierarchies of beliefs are 'closed'. Mertens and Zamir (1985) have established that any space with the same fundamental properties as Ω must be homeomorphic to Ω. They have also introduced the notion of a *beliefs subspace* X of Ω. This can be defined as a subspace X of Ω such that, at every point $x \in X$, every player assigns probability 1 to X (according to his type at x).

In practice, game theorists focus on beliefs subspaces. They often impose a further condition: *consistency* (also known as Harsanyi's 'common prior assumption'). Under this assumption, there exists a single probability distribution P over the underlying (say, finite) beliefs subspace $X = K \times S_1 \times S_2$, such that players' beliefs coincide with the ones which they derive from the information scheme $\mathcal{S}_K = (S_1, S_2, P)$ and the game $G(\Gamma_K, \mathcal{S}_K)$ can be used as a representation of the decision problem Γ_K completed by the description of beliefs. Observe in particular that the common knowledge of Γ_K and the construction of a beliefs subspace guarantee the common knowledge of the rules of $G(\Gamma_K, \mathcal{S}_K)$.

We have thus seen that a finite consistent beliefs subspace of the universal beliefs space over K typically has the form of an information scheme $\mathcal{S}_K$. Conversely, any information scheme induces a consistent beliefs subspace of the universal beliefs space (see Mertens, Sorin and Zamir). But obviously, an information scheme may involve signals which appear as irrelevant to the basic parameter (in K) once the scheme is projected on the universal beliefs space (relative to K, namely $\Omega = K \times \Theta_1 \times \Theta_2$). The particular case where K is a singleton provides a trivial example. Signals from the information scheme appear as extraneous in this context. The same phenomenon may also arise when K has more than one element. Suppose for instance that a single signal is selected in a finite set, independently of $k \in K$, and is transmitted to both players. The expected utility functions $v_i(s, a)$ do not depend on s. Such an information scheme cannot be generated by hierarchies of beliefs over K. More precisely, the marginal beliefs on S do not have a projection on the universal space *constructed from* K. The basic parameters space seems to be crucial. Nevertheless, in the applications, K is often enlarged without specific caution and one does not make any distinction between games with incomplete information and games with private signals, even if the latter may involve extraneous signals. This may have consequences on solution concepts as we shall see later.

3. CORRELATED EQUILIBRIUM IN STRATEGIC FORM GAMES

Throughout this section we fix a two person game in strategic form Γ. As in the previous section, Γ is described by sets of actions A_i and utility functions u_i $(i = 1, 2)$. We summarize the main properties of correlated equilibria; they have been obtained by Aumann (1974, 1987) (see also Myerson, 1991, and Forges, 1986b).

3.1. DEFINITION. The set of all *correlated equilibrium payoffs* of Γ is defined by

$$C(\Gamma) = \bigcup_{\mathscr{S}} N(G(\Gamma, \mathscr{S}))$$

where $\mathscr{S}$ varies over all (finite) information schemes, $G(\Gamma, \mathscr{S})$ is the game with private (extraneous) signals generated by Γ and $\mathscr{S}$ (see Section 2) and $N(G)$ denotes the set of *Nash* equilibrium payoffs of a game G.

In other words, a correlated equilibrium of Γ is a Nash equilibrium of an extension $G(\Gamma, \mathscr{S})$ of Γ where the players get private (possibly correlated) signals before playing Γ.

We focus on equilibrium payoffs rather than equilibrium strategies because the dimension of the strategies sets depends on the underlying information scheme. By contrast, for every $\mathscr{S}$, the set $N(G(\Gamma, \mathscr{S}))$ is included in $\mathbb{R}^2$.

3.2. *Canonical Representation*

The next result is well known; it is analogous to the 'revelation principle' of the mechanisms literature.

An information scheme $\mathscr{S} = (S_1, S_2, Q)$ is *canonical* for Γ if $S_i = A_i$ $(i = 1, 2)$. Given a canonical information scheme $\mathscr{S}$, the identity mapping on A_i is called the *canonical strategy* of player i $(i = 1, 2)$. A *canonical correlated equilibrium payoff* of Γ is achieved by means of canonical strategies in an extension $G(\Gamma, \mathscr{S})$ where $\mathscr{S}$ is canonical. The proposition below states that one can restrict to canonical correlated equilibrium payoffs without loss of generality.

PROPOSITION 1. *$C(\Gamma)$ coincides with the set of canonical correlated equilibrium payoffs of Γ.*

$C(\Gamma)$ is thus easily characterized, namely by all probability distributions Q over $A_1 \times A_2$ satisfying the following property: if (a_1, a_2) is selected according to Q and a_i is secretly recommended to player i $(i = 1, 2)$ before Γ, then playing the recommendation is a Nash equilibrium of the corresponding extension of Γ.

The canonical representation leads us to the usual interpretation of the correlated equilibrium as a *mediated equilibrium*. A mediator is assumed to perform the lottery Q over $A_1 \times A_2$ and to make recommendations to the players accordingly. Q induces a correlated equilibrium if none of the players can gain by unilaterally deviating from the recommendation. This is not the only interpretation of correlated equilibria, as we shall see in the next subsection.

3.3. *Correlated Equilibrium as an Expression of Bayesian Rationality*

As in Aumann (1987), let us complete the description of Γ by the following elements:

Y: the space of 'all states of the world', which contains everything which may be uncertain to the players. According to Aumann, we may assume that Y is finite. We may also view Y as a beliefs subspace of the universal beliefs space Ω' constructed on the parameters space $A_1 \times A_2$ (this is also suggested in Aumann, 1987, under point c of the Discussion).

Π: probability distribution over Y; Π is the *common prior* of the players. This corresponds to a *first assumption*: *consistency*.

$\mathcal{S}_i$: player i's information partition of Y $(i = 1, 2)$.

α_i: A_i-valued random variable on Y $(\alpha_i : Y \to A_i)$ describing player i's action. A *second assumption* expresses that player i knows the action he chooses: α_i is measurable with respect to $\mathcal{S}_i$ $(i = 1, 2)$.

Finally, let us formulate a *third assumption*: each player is Bayes rational at every state of the world, namely maximizes his expected utility given his information. For player 1 the condition is

$$E[u_1(\alpha_1, \alpha_2) \mid \mathcal{S}_1] \geqslant E[u_1(a_1, \alpha_2) \mid \mathcal{S}_1] \quad \forall a_1 \in A_1 \tag{3.1}$$

where E is the expectation with respect to Π. A similar condition can be written for player 2.

Let us define $B(\Gamma)$ as the *set of all payoffs to Bayes rational solutions* of Γ, namely all payoffs that are achieved by taking account of the space of all states of the world, under the three above assumptions. Observe that we did not refer to any game-theoretic solution concept here. We did not even define strategies. In (3.1), player 1 does not make any conjecture on player 2's behavior. However, player 2's action α_2 is a random variable on the space of all states of the world, so that player 1 has a distribution over it, given his information. The space of all states of the world contains in particular the players' actions.

The next proposition states that correlated equilibrium can be viewed as an expression of Bayesian rationality. The proof is so short that we recall it here.

PROPOSITION 2.

$$B(\Gamma) = C(\Gamma)$$

Proof. Let us take the conditional expectation of (3.1) given α_1:

$$E[u_1(\alpha_1, \alpha_2) \mid \alpha_1] \geq E[u_1(a_1, \alpha_2) \mid \alpha_1] \quad \forall a_1 \in A_1$$

because α_1 is $\mathscr{S}_1$-measurable. A similar inequality obviously holds for player 2. This expresses that the distribution of (α_1, α_2) (over $A_1 \times A_2$) corresponds to a canonical correlated equilibrium. ■

In this section we have recalled three equivalent approaches to correlated equilibrium in strategic form games. In the next section, we shall see that the three views are no longer equivalent in the context of incomplete information.

4. CORRELATED EQUILIBRIUM IN GAMES WITH INCOMPLETE INFORMATION

In this section we fix a game with private signals $G(\Gamma_K, \mathscr{S}_K)$. We only need to consider the reduced form of the game, which we simply denote by G. Let us set $\mathscr{S}_K = (T_1, T_2, P)$; G is described as follows:

stage 1: (t_1, t_2) is chosen in $T_1 \times T_2$ according to P; t_i is secretly transmitted to player i $(i = 1, 2)$.
stage 2: the players simultaneously choose an action (in A_1 and A_2, respectively).

The payoff function of player i is

$$v_i : T \times A \to \mathbb{R} \quad (i = 1, 2)$$

where $T = T_1 \times T_2$ and $A = A_1 \times A_2$. G corresponds to the standard description of a 'game with incomplete information'.

We shall survey four definitions of correlated equilibrium in G, which have all been used in the literature.

4.1. Strategic Form Approach

A natural approach is to consider the strategic form of G and to apply the solution concept of the previous section. The set of pure strategies of player i in G is

$$\Sigma_i = A_i^{T_i} \quad (i = 1, 2)$$

and payoffs can be expressed as a function of pure strategies, in the usual way.

Let $C(G)$ be the set of all *strategic form correlated payoffs* of G. In the context of games with incomplete information, we might consider vector payoffs, namely parametrize payoffs by types. $C(G)$ would then be a subset of $\mathbb{R}^{T_1} \times \mathbb{R}^{T_2}$. We rather deal with expected payoffs, in view of the results of Section 6. However, all results of this section hold for sets of vector payoffs.

Proposition 1 obviously applies to $C(G)$. In a canonical strategic form correlated equilibrium, the mediator selects (σ_1, σ_2) in $\Sigma_1 \times \Sigma_2$ according to some probability distribution Q and transmits σ_i to player i $(i = 1, 2)$. σ_i is a vector of recommendations; player i of type t_i should play $\sigma_i(t_i) \in A_i$. In this representation, the mediator is fully independent of the game and he cannot distinguish the different types of the players. Hence he transmits to player i a recommendation for every type t_i.

The given probability P over $T_1 \times T_2$ and the canonical correlated equilibrium Q over $\Sigma_1 \times \Sigma_2$ induce a probability distribution Π over $T_1 \times T_2 \times A_1 \times A_2$.

$$\Pi(t_1, t_2, a_1, a_2) = P(t_1, t_2) Q(\sigma_1(t_1) = a_1, \sigma_2(t_2) = a_2)$$

The marginal probability distribution of Π over $T_1 \times T_2$ is obviously P; Π satisfies the following

CONDITIONAL INDEPENDENCE PROPERTY: player 2's type (t_2) is conditionally independent of player 1's action (a_1), given player 1's type (t_1). Similarly, player 1's type (t_1) is conditionally independent of player 2's action (a_2), given player 2's type (t_2).

As Proposition 1, Proposition 2 holds here. However, this result may be difficult to interpret because in the present formulation, types are not part of the state of the world. We shall come back to this in Subsection 4.4 and in Section 6.

The strategic form correlated equilibrium has been used to solve games with incomplete information by Cotter (1991) and Forges (e.g.

1985, 1990). The advantage of this solution concept is that the correlation mechanism does not interfere with the game, so that the game can be played according to its original rules. However, the rules of G are somewhat artificial because the first stage of G is only a representation of the players' beliefs. One could argue that this stage is just a 'historical node' (Myerson, 1991), which does not really occur.

4.2. *Agents Normal Form Approach*

The agents normal form of G is the game where every type t_i of player i is represented by a different agent indexed by (i, t_i). We denote this game as G_a; the number of players in G_a is $|T_1| + |T_2|$. Agent (i, t_i) is only active when t_i is chosen; in this case, he gets the payoff of player i of type t_i. Payoffs are thus conditional on activity. This representation is known as the 'Selten game' (see Harsanyi, 1967–1968), and Myerson, 1991). Nash equilibrium payoffs satisfy the following property:

$$N(G) = N(G_a)$$

Such an equivalence does not hold for correlated equilibria. This can be understood from the canonical representation (see below). An example will also be given in Section 5. Let us define the set $C_a(G)$ of *agents normal form correlated equilibrium payoffs* of G as

$$C_a(G) = C(G_a)$$

A *canonical* correlated equilibrium of G_a is represented by a probability distribution Q over $\Sigma_1 \times \Sigma_2$, exactly as in the previous subsection, but now agent (i, t_i) only receives the recommendation $\sigma_i(t_i)$; player i of type $t'_i \neq t_i$ does not know this recommendation. It is clear that

$$C(G) \subseteq C_a(G)$$

since the players' information decreases in the agents normal form representation. But there is no hope to get the converse (see Example 3 in Section 5).

Samuelson and Zhang (1989) have proposed the agents normal form correlated equilibrium as a legitimate generalization of Aumann's solution concept in games with incomplete information. Agents normal form correlated equilibria have also been used in Forges (1986a). Myerson (1991, p. 262) argues that

> these equilibria have no clear interpretation in terms of the given Bayesian game with communication, unless one makes the unnatural assumption that a mediator can send each player a message that depends on his actual type while the players can send no reports to the mediator at all!

Nevertheless, agents normal form correlated equilibria seem appropriate as soon as types are *verifiable* by a mediator, even if they are unverifiable by the participants in the game. Another justification can be elaborated. Even if an omniscient mediator coordinates the players, it may seem reasonable that his recommendations *do not modify the players' beliefs over each other's types*, namely that the conditional probability of t_2 (say) given t_1 and the recommendation a_1 is $P(t_2 \mid t_1)$. This amounts to requiring that the induced probability distribution over $T_1 \times T_2 \times A_1 \times A_2$ satisfies the *conditional independence property* stated in Subsection 4.1. This condition is obviously fulfilled in every agents normal form correlated equilibrium since the associated distribution over $T_1 \times T_2 \times A_1 \times A_2$ is the same as in a strategic form correlated equilibrium. The two concepts differ only in the signals received by the players.

In Subsection 4.4, we shall show that agents normal form correlated equilibria are not as artificial as they may appear at first sight. They can be interpreted as *Bayesian solutions* of the original game (with only two players) and thus as the proper generalization of correlated equilibrium in the framework of Subsection 3.3.

4.3. *Communication Equilibrium*

The communication equilibrium is likely the best known generalization of Aumann's concept in games with incomplete information. It has been used in many applications (see e.g. Myerson, 1982, 1991; Einy and Peleg, 1991; Forges, 1985, 1986b, 1990). Samuelson and Zhang (1989) compare the communication equilibrium with the agents normal form correlated equilibrium; they consider these solution concepts as two legitimate extensions of the correlated equilibrium in games with incomplete information.

The communication equilibrium can be interpreted as a *mediated* *equilibrium*. The role of the mediator is indeed very important in the scenario traditionally associated with this solution concept. As in the mechanisms literature, the players send messages to the mediator after having learnt their types. As in the case of games with complete information, the mediator sends then a private signal to every player. A generalized revelation principle applies, and every communication equilibrium payoff can be achieved in a canonical way.

A *canonical communication device* is described by a system q of conditional probability distributions $q(\cdot \mid t_1, t_2)$ over $A_1 \times A_2$, for every $(t_1, t_2) \in T_1 \times T_2$. $q(a_1, a_2 \mid t_1, t_2)$ is interpreted as the probability that the mediator *recommends* a_i to player i $(i = 1, 2)$ if (t_1, t_2) is the vector of types *reported* by the players. In order to define a canonical communication equilibrium, q must satisfy linear inequalities which express that no player can gain by lying about his type and/or deviating from the recommended action, as long as the other player is honest and obedient.

We shall denote by $M(G)$ the set of all communication equilibrium payoffs of G.

The difference between the present scenario and the previous ones (in Subsections 4.1 and 4.2) is that here, there is a *two-way* communication between the players and the mediator. It is clear that $C(G) \subseteq M(G)$ (see e.g. Forges, 1986b); hence

$$C(G) \subseteq C_a(G) \cap M(G)$$

This inclusion may be strict as Example 3 of Section 5 shows. Examples 1 and 2 also illustrate that there is no inclusion relationship between $C_a(G)$ and $M(G)$. On the one hand, the equilibrium payoffs in $C_a(G)$ do not result from any information transmission by the players, which is made possible in a communication equilibrium. In the latter case, the players can for instance reveal their type to each other even if G does not allow them to communicate. Moreover, the probability distribution Π over $T_1 \times T_2 \times A_1 \times A_2$ induced by the original probability distribution P over $T_1 \times T_2$ and a canonical communication device q is defined by

$$\Pi(t_1, t_2, a_1, a_2) = P(t_1, t_2) q(a_1, a_2 \mid t_1, t_2)$$

It does not satisfy the conditional independence property, which holds in the case of $C_a(G)$. On the other hand, the agents normal form correlated equilibrium allows type-dependent recommendations which could not be realized with a communication device because some player could have incentives to lie.

4.4. *Partial Bayesian Approach*

Subsections 4.1 and 4.2 were inspired by the first definition of correlated equilibrium in games with complete information (Subsection 3.1). Similarly, the communication equilibrium can be viewed as a natural generalization of Subsection 3.2. Here we extend the Bayesian setup expounded in Subsection 3.3. Our analysis will be partial in the sense that we shall not question the basic types in $T_1 \times T_2$ nor the beliefs over them. The game G is kept fixed.

Let Y, Π, $\mathcal{S}_i$ and α_i $(i = 1, 2)$ be defined as in Subsection 3.3. Since Y is the space of all states of the world, it contains the *players' types*; we thus add T_i-valued random variables τ_i $(i = 1, 2)$ to the previous model in order to describe the types $(\tau_i : Y \to T_i)$.

We make the natural assumption that each player knows his type; τ_i is thus $\mathcal{S}_i$-measurable $(i = 1, 2)$. We also make a *minimal coherence requirement*: the distribution of (τ_1, τ_2), which is derived from the probability distribution Π over the space of all states of the world, must coincide with the probability P over $T_1 \times T_2$ which is part of the rules of the games G. In other words,

$$\Pi(\tau_1 = t_1, \tau_2 = t_2) = P(t_1, t_2) \tag{4.1}$$

for every $(t_1, t_2) \in T_1 \times T_2$.

We can extend the Bayes rationality condition for player 1 (see (3.1)) as follows

$$E[v_1(\tau_1, \tau_2, \alpha_1, \alpha_2) \mid \mathcal{S}_1] \geq E[v_1(\tau_1, \tau_2, a_1, \alpha_2) \mid \mathcal{S}_1] \quad \forall a_1 \in A_1 \tag{4.2}$$

where E is the expectation with respect to Π. A similar condition holds for player 2. We define $B(G)$ as the set of all *Bayesian solutions* to G, namely the set of all payoffs which can be achieved as above, for some Y, Π, $\mathcal{S}_i$, α_i, τ_i $(i = 1, 2)$ satisfying (4.1), (4.2) and the analogue of (4.2) for player 2.

$B(G)$ has a *canonical representation*, exactly as the previous sets of equilibrium payoffs. Every payoff in $B(G)$ is characterized by a probability distribution over $T_1 \times T_2 \times A_1 \times A_2$, which can be denoted by Π without any risk of confusion. The marginal distribution of Π over $T_1 \times T_2$ coincides with P. Π also satisfies rationality conditions:

$$\sum_{t_2, a_2} \Pi(t_2, a_2 \mid t_1, a_1)[v_1(t_1, t_2, a_1, a_2) - v_1(t_1, t_2, a_1', a_2)] \geqslant 0$$
$$\forall t_1 \in T_1, a_1 \in A_1 : \Pi(t_1, a_1) > 0, \quad \forall a_1' \in A_1 \tag{4.3}$$

for player 1, and similarly for player 2. These inequalities are simply obtained by taking the conditional expectation of (4.2) given the $\mathcal{S}_1$-measurable random vector (τ_1, α_1).

The canonical representation of the solutions in $B(G)$ suggests that they are achieved with the help of an *omniscient* mediator. He chooses a pair of actions conditionally on the actual types of the players although the players do not transmit any information to him.

The solutions in $B(G)$ are akin to the *jointly coherent outcomes* defined in Nau (1992) (see also Nau and McCardle, 1990). More precisely, the vector of types and actions which have positive probability Π in a Bayesian solution (such that Π satisfies a suitable property[2]) coincide with jointly coherent outcomes of the game G.

Nau's approach is axiomatic and guided by de Finetti's no arbitrage condition. His formalization of types and beliefs is intermediate between the model of this section (where beliefs over types are fixed i.e. are described by P) and the next one (where beliefs will be constructed, so that the sets of types themselves will vary). In Nau (1992), the set of types T_1 and T_2 are given but the beliefs are derived from no-arbitrage conditions. The model does not seem to have the same basic features as Harsanyi's one.

As one may expect, $B(G)$ contains all the sets of equilibrium payoffs considered previously:

$$C(G) \subseteq C_a(G) \cup M(G) \subseteq B(G) \tag{4.4}$$

In a communication equilibrium, player 1 faces condition (4.3) when he reveals truthfully his type and gets the recommendation a_1. In this context, *incentive compatibility conditions* must also be satisfied. In an agents normal form correlated equilibrium, (4.3) clearly holds (with

$a_1 = \sigma_1(t_1)$), but in this case Π satisfies the *conditional independence property*. It is thus not surprising that the (second) inclusion in (4.4) may be strict. This is illustrated by Example 4 in Section 5.

We are going to prove that $C_a(G)$ coincides with the subset of solutions in $B(G)$ which satisfy a stronger coherence requirement than (4.1). Let us consider a world Y, Π, $\mathcal{S}_i$, α_i, τ_i $(i = 1, 2)$ as above and let us strengthen the connections between this world and the game G by adding the following conditions

$$\Pi(\tau_j = t_j \mid \mathcal{S}_i) = \Pi(\tau_j = t_j \mid \tau_i) \quad i, j = 1, 2 \quad i \neq j \tag{4.5}$$

This means that the beliefs of player i over the other player's type cannot be altered by his knowledge of the world. In other words, the description of G remains 'correct' when G is embedded in the space of all states of the world.

PROPOSITION 3. *The set of all payoffs in* $B(G)$ *which can be achieved in a world* $(Y, \Pi, \mathcal{S}_i, \alpha_i, \tau_i, i = 1, 2)$ *satisfying* (4.1), (4.2) *and* (4.5) *is exactly* $C_a(G)$.

Proof. It is clear that $C_a(G)$ is included in the subset of $B(G)$ described in the statement. Indeed, the probability distribution Π over $T_1 \times T_2 \times A_1 \times A_2$ associated with a solution in $C_a(G)$ satisfies the conditional independence property and hence (4.5).

Conversely, let us start with a solution in $B(G)$ as in the above statement. The conditional expectation of (4.5) given (α_i, τ_i) yields that α_i is independent of τ_j given τ_i $(i, j = 1, 2,\ i \neq j)$. Hence, the conditional independence property holds. In order to show that the solution in $B(G)$ belongs to $C_a(G)$, we construct a probability distribution Q over $\Sigma_1 \times \Sigma_2$ from Π. We first define the marginal distributions over $A_1 \times A_2$ by

$$Q(\sigma_1(t_1) = a_1, \sigma_2(t_2) = a_2) = \Pi(\alpha_1 = a_1, \alpha_2 = a_2 \mid t_1, t_2)$$

for every (t_1, t_2) (the right-hand side member is defined arbitrarily if $P(t_1, t_2) = 0$). The conditional independence condition guarantees that this definition is meaningful: the marginal distribution over A_1 deduced from the right-hand side is independent of t_2. We complete the construction of Q by defining it as the product distribution of the above probability distribution over $A_1 \times A_2$:

$$Q(\sigma_1, \sigma_2) = Q((\sigma_1(t_1))_{t_1 \in T_1}, (\sigma_2(t_2))_{t_2 \in T_2})$$
$$= \prod_{(t_1,t_2)} \Pi(\alpha_1 = a_1, \alpha_2 = a_2 \mid t_1, t_2)$$

The rationality condition (4.3) is exactly the non-deviation condition of player 1 of type t_1 in the agent normal form correlated equilibrium described by Q. Indeed, the construction of Q implies that

$$\Pi(t_2, a_2 \mid t_1, a_1) = P(t_2 \mid t_1) Q(\sigma_2(t_2) = a_2 \mid \sigma_1(t_1) = a_1) \qquad \blacksquare$$

An analogous result does not hold for strategic form correlated equilibria nor for communication equilibria. This follows from the timeless structure of Aumann's world. Strategic form correlated equilibria require a virtual stage where players do not know their types yet. And communication equilibria require to model explicitly a process of information transmission.

4.5. Remark

We have presented four definitions of correlated equilibrium in (the reduced form of) $G(\Gamma_K, \mathcal{S}_K)$. Many other variants are conceivable. For instance, we could imagine a solution described by a probability distribution Π over $K \times T_1 \times T_2 \times A_1 \times A_2$, whose marginal distribution over $K_1 \times T_1 \times T_2$ would be P (determined by $G(\Gamma_K, \mathcal{S}_K)$). We could require that Π satisfies a further conditional independence property:

$$\Pi(k, t_j \mid t_i, a_i) = P(k, t_j \mid t_i) \quad i, j = 1, 2 \quad i \neq j .$$

We would otherwise impose the same conditions as in Subsection 4.4: each player knows his type and his action and maximizes his expected utility given his information.

The solutions corresponding to the previous approach may not belong to $B(G(\Gamma_K, \mathcal{S}_K))$, but they are in the set $U(\Gamma_K)$ which will be introduced in Section 6. In particular, the restriction to the reduced form of $G(\Gamma_K, \mathcal{S}_K)$, which is so common in the applications, is not innocuous in the present context. In order to illustrate this, let K contain two states of nature, let T_1 and T_2 be singletons and suppose that each player has two actions. The game is essentially a game with complete information. But the pair of actions need not be independent of k here.

A probability distribution Π with the following conditional distributions over $A_1 \times A_2$:

$$\begin{pmatrix} 1/8 & 3/8 \\ 3/8 & 1/8 \end{pmatrix} \quad \text{if } k=1$$

$$\begin{pmatrix} 3/8 & 1/8 \\ 1/8 & 3/8 \end{pmatrix} \quad \text{if } k=2$$

satisfies the above requirement. Knowledge of a_1 (resp. a_2) does not reveal anything on the state of nature to player 1 (resp. 2). However, $\Pi(k, a_j \mid a_i)$ depends on k and it should not be difficult to construct payoffs such that all rationality conditions are satisfied but the corresponding payoff is not a correlated equilibrium payoff of the reduced form of the game.

5. EXAMPLES

The following examples clarify the relationships between the set of solutions considered above. We observed that for any Bayesian game G,

$$C(G) \subseteq C_a(G) \cap M(G)$$

$$C_a(G) \cup M(G) \subseteq B(G)$$

Example 1 consists of a game G such that $M(G)$ is not included in $C_a(G)$. Similarly, Example 2 shows that $C_a(G)$ may not be included in $M(G)$. Games with the same properties were also exhibited in Samuelson and Zhang (1989). Example 3 (resp. 4) illustrates that the first (resp. second) inclusion above may be strict.

Example 1. $T_1 = \{t_1, t_1'\}$, T_2 is a singleton, $P(t_1) = P(t_1') = \frac{1}{2}$, A_1 is a singleton, $A_2 = \{a_2, a_2'\}$, v_1 and v_2 are described by the following payoff matrices

$$\begin{array}{ccc} & a_2 & a_2' \\ t_1 & (1,2 & 0,0) \end{array}$$

$$t_1' \begin{matrix} a_2 & a_2' \\ (0,0 & 1,4) \end{matrix}$$

In $C_a(G)$ – *a fortiori* in $C(G)$ – players 2's payoff cannot be larger than $2 = \frac{1}{2}(0+4)$. But $(1,3)$ belongs to $M(G)$ since

$$q(a_2 \mid t_1) = q(a_2' \mid t_1') = 1$$

induces a communication equilibrium.

Example 2. $T_1 = \{t_1, t_1'\}$; T_2 is a singleton; $P(t_1) = P(t_1') = \frac{1}{2}$; $A_1 = \{1, 2, \ldots, n\}$ $(n \geqslant 2)$; $A_2 = \{a_2, a_2'\}^n$.

These data may be interpreted as follows: after having learnt his type, player 1 sends a message in A_1 to player 2; the latter chooses then an action in $\{a_2, a_2'\}$ as a function of this message. The message is costless; payoffs only depend on player 1's type and player 2's action; they are described by

$$t_1 \begin{matrix} a_2 & a_2' \\ (1,2 & 0,0) \end{matrix}$$

$$t_1' \begin{matrix} a_2 & a_2' \\ (1,0 & 0,4) \end{matrix}$$

In order to construct a payoff in $C_a(G) \backslash M(G)$, consider a correlation device which selects a_1^* uniformly in $\{1, \ldots, n\}$ and transmits a_1^* to agent $(1, t_1)$ and player 2 (but not to agent $(1, t_1')$). The corresponding strategy of agent $(1, t_1)$ consists of sending a_1^* to player 2; the strategy of player 2 is to play a_2 if he receives a_1^*, a_2' otherwise. (Recall that in the agents normal form, player 2 does not know whether his opponent is agent $(1, t_1)$ or agent $(1, t_1')$. The previous strategy consists of choosing a_2 if the signal of the correlation device and the message from player 1 coincide.) If agent $(1, t_1')$ sends some message a_1 to player 2, he induces a_2 with probability $1/n$ (namely if $a_1 = a_1^*$). Any strategy of agent $(1, t_1')$ is a best reply to the previous scenario and the expected payoff of this agent is $1/n$. Agent $(1, t_1)$ should send the message a_1^* recommended by the correlation device; his expected payoff is 1.

Given the signal a_1^* of the correlation device and the message a_1 of his opponent, player 2's beliefs are

$$\pi(t_1 \mid a_1) = \begin{cases} \dfrac{n}{n+1} & \text{if } a_1 = a_1^* \\ 0 & \text{otherwise} \end{cases}$$

so that playing a_2 on a_1^*, a_2' on $a_1 \neq a_1^*$ is an optimal response. Player 2's expected payoff is $(3n-1)/n$. We have thus shown that

$$\left(\frac{n+1}{2n}, \frac{3n-1}{n}\right) \text{ belongs to } C_a(G)\,.$$

By the same reasoning as in the revelation principle, $M(G)$ is not larger than the set of communication equilibrium payoffs of the game G' where player 1 does not send any message to player 2. More precisely, G' has the same form as the game in example 1. A_1 is a singleton and $A_2 = \{a_2, a_2'\}$. It is easily checked that the only payoff in $M(G')$ is $(0, 2)$. The incentive conditions guaranteeing that player 1 reveals some information are not satisfied.

The previous examples have already appeared in a different context. Example 1 is borrowed from Forges (1986b) and example 2 is inspired from Forges (1986a). The game of Example 3 has been used in Forges (1985).

Example 3. $T_1 = \{t_1, t_1'\}$, T_2 is a singleton, $P(t_1) = P(t_1') = \frac{1}{2}$, $A_1 = \{a_1, a_1'\}$, $A_2 = \{a_2, a_2', a_2''\}^{A_1}$.

These data can be interpreted as in the previous example; the payoffs are described in the same way.

$$\begin{array}{c} \\ t_1 \end{array} \begin{array}{ccc} a_2 & a_2' & a_2'' \\ (6,0 & 0,4 & 4,6) \end{array}$$

$$\begin{array}{c} \\ t_1' \end{array} \begin{array}{ccc} a_2 & a_2' & a_2'' \\ (0,6 & 6,4 & 4,0) \end{array}$$

First, the communication device described by

$$q(a_2' \mid t_1) = q(a_2'' \mid t_1) = \tfrac{1}{2}$$

$$q(a_2' \mid t_1') = 1$$

induces an equilibrium. Hence, $(4, 4.5) \in M(G)$. Observe that, as above, player 1's action becomes superfluous once a communication device is added to the game.

In Forges (1985), it is proved that $(4, 4.5) \notin C(G)$. It remains to show that this payoff belongs to $C_a(G)$. Let us consider a correlation device which selects a_1 and a_1' with equal probability (1/2) for agent $(1, t_1)$, independently of the signals to agent $(1, t_1')$ and player 2. The recommendations to these two players are correlated as follows: with probability 1/2, the device recommends a_1 to agent $(1, t_1')$ and the strategy $\alpha_2 : A_1 \to \{a_2, a_2', a_2''\}$ defined by $\alpha_2(a_1) = a_2'$, $\alpha_2(a_1') = a_2''$ to player 2; similarly, with probability 1/2, the device recommends a_1' to agent $(1, t_1')$ and the strategy $\alpha_2(a_1) = a_2''$, $\alpha_2(a_1') = a_2'$ to player 2.

Agent $(1, t_1)$ is indifferent between a_1 and a_1' whatever the recommendation of the device. He can thus follow this recommendation. Agent $(1, t_1')$ deduces player 2's effective strategy α_2 from the recommendation of the device. This leads to the conclusion that he should be obedient too. Finally, assume that player 2 receives the advice '$\alpha_2(a_1) = a_2'$, $\alpha_2(a_1') = a_2''$.' He deduces that $(1, t_1')$ will send him message a_1 and that $(1, t_1)$ will send a_1 and a_1' with equal probability 1/2. He is thus exactly in the same situation as in the communication equilibrium above.

Example 4. Let us replace the payoffs of example 1 by the ones of example 2, while keeping the simple structure of the first example. In particular, player 1 cannot transmit any information to player 2. The payoff $(1/2, 3)$ belongs to $B(G)$ but it is not in $C_a(G)$ nor in $M(G)$. These two sets coincide here and contain only the payoff $(0, 2)$. Remark that the payoff $(1/2, 0)$ is feasible but is not in $B(G)$, because it does not satisfy player 2's rationality condition.

6. UNIVERSAL BAYESIAN APPROACH

In the context of complete information, Aumann's space of all states of the world Y can be thought of as a beliefs subspace of the universal beliefs space Ω' constructed over the actions space $A_1 \times A_2$. In the last section, we fixed a game with incomplete information, in particular types sets T_1, T_2, consistent beliefs induced by a probability distribution P and actions spaces A_1, A_2. According to Harsanyi (1967) and Mertens and Zamir (1985) (see also Mertens, Sorin and Zamir, forthcoming), the types sets are part of a beliefs subspace. More precisely, let K denote the basic parameters space (as in Section 2); $K \times T_1 \times T_2$ is a consistent beliefs subspace of the universal beliefs space Ω constructed over K. If Aumann's space of states of the world has to be viewed as above, the approach of Subsection 4.4 does not seem appropriate, since two hierarchies of beliefs are constructed in a sequential way (a first one for the beliefs over the state of nature and a second one for the beliefs over actions). In order to respect the spirit of Aumann's theorem (Proposition 2), it seems that a single construction should be performed, namely that beliefs over the basic parameter (in K) and over the actions (in $A_1 \times A_2$) should be generated at the same time.

Before detailing such a model, let us point out that this approach contradicts Harsanyi's principles. He insisted on distinguishing independent and dependent variables (types and actions) and would not have incorporated dependent variables in the states space. But Aumann did and it seems natural to proceed in the same way in the extension of Proposition 2. Armbruster and Böge (1979) and Böge and Eisele (1979) considered hierarchies of beliefs over basic parameters as well as over actions, but they imposed rationality restrictions at each step of the construction. As Aumann, we only impose such restrictions once the beliefs subspace is well-defined (see also Mertens, Sorin and Zamir, forthcoming).

Let us start with a two-person decision problem Γ_K, as in Section 2. K is the basic parameters space. $A_1 \times A_2$ is the actions space. Let us apply Mertens and Zamir's construction to $K \times A_1 \times A_2$. The universal beliefs space has the form

$$\Omega^* = K \times A_1 \times A_2 \times \Theta_1^* \times \Theta_2^*$$

and satisfies the homeomorphism property

$$\Theta_i^* \simeq \Delta(K \times A_1 \times A_2 \times \Theta_j^*) \quad i, j = 1, 2 \quad i \neq j$$

An element of Θ_i^* describes a *full* type of player i, namely specifies his beliefs over the basic parameter, the actions and the type of the other player; the latter can obviously be interpreted in the same way.

Let us formulate Aumann's assumptions in this framework. The first one (*finiteness* and *consistency* of the space of states of the world) leads us to restrict ourselves to a finite, consistent beliefs subspace

$$Y^* = K \times A_1 \times A_2 \times S_1^* \times S_2^*$$

of Ω^*. Let P^* be the corresponding probability distribution over Y^*.

The second assumption states that each player i *knows his action*. Player i's knowledge is described by his type $s_i \in S_i^*$ or equivalently by the conditional probability distribution $P^*(\cdot \mid s_i)$ over $K \times A_1 \times A_2 \times S_j^*$. Player i's knowledge of his own action means that $P^*(\cdot \mid s_i)$ assigns probability one to some element of A_i, which we denote as $\gamma_i(s_i)$ $(i = 1, 2)$. If we combine the knowledge restrictions of both players with consistency, we get that P^* satisfies

$$P^*(k, a_1, a_2, s_1, s_2) = P^*(k, s_1, s_2) I(a_1 = \gamma_1(s_1), a_2 = \gamma_2(s_2)) \tag{6.1}$$

for every element of Y^* (I denotes the indicator function). This equality just says that the players' types determine their actions.

It remains to express that *each player maximizes his expected payoff given his information*. Player 1's information is described by $s_1 \in S_1^*$; as we have just seen, the corresponding action is $\gamma_1(s_1)$; the rationality condition of player 1 is thus that $\gamma_1(s_1)$ maximizes

$$E^*[u_1(k, a_1, a_2) \mid s_1]$$

over A_1, where E^* is the expectation with respect to P^*. As above (in Subsections 3.3 and 4.4), strategies are not explicitly defined but the model specifies player 1's probability distribution over the variables k, a_2 which are relevant to his utility. Player 1's rationality condition can be precisely written as

$$a_1 = \gamma_1(s_1) \text{ maximizes } \sum_{k, a_2} P^*(k, a_2 \mid s_1) u_1(k, a_1, a_2) \tag{6.2}$$

A similar condition obviously expresses player 2's Bayesian rationality.

Observe that in the setup adopted here, we do not have to assume that the world is consistent with the underlying game nor that the players know their types. No game is fixed in advance and types are included in the states of the world by construction. As in Mertens and Zamir (1985), the fact that each agent knows his own type is implicit.

Let us denote as $U(\Gamma_K)$ the set of all payoffs corresponding to universal Bayesian solutions of Γ_K (namely, the solutions achieved in some beliefs subspace Y^*, P^* under the assumptions above). With Aumann's approach in mind, we can regard $U(\Gamma_K)$ as the set of correlated equilibrium payoffs of Γ_K. We are thus confronted with a fifth definition. The natural questions which arise now concern the relationship between $U(\Gamma_K)$ and the four sets defined previously. An important difference between $U(\Gamma_K)$ and the other sets is that $U(\Gamma_K)$ is defined directly on the decision problem Γ_K while the others describe the solutions of a Bayesian game G in reduced form. In particular, the beliefs are partially specified by G in the latter case.

In Section 2 we associated games with private signals $G(\Gamma_K, \mathcal{S}_K)$ with Γ_K by adding information schemes $\mathcal{S}_K = (S_1, S_2, P)$ to Γ_K. The solution concepts considered in Section 4 apply to the (reduced form of the) game $G(\Gamma_K, \mathcal{S}_K)$. The next proposition states that $U(\Gamma_K)$ coincides with the set of *all* solutions (according to *any* of the four definitions of Section 4) which can be achieved in $G(\Gamma_K, \mathcal{S}_K)$ for some $\mathcal{S}_K$. In particular, the four solution concepts are *equivalent* in that sense. Furthermore, the same property holds for the Nash equilibrium solution concept.

PROPOSITION 4

$$U(\Gamma_K) = \bigcup_{\mathcal{S}_K} Z(G(\Gamma_K, \mathcal{S}_K))$$

where $\mathcal{S}_K$ varies over all information schemes on K and Z denotes the set of all payoffs corresponding to one of the previous solution concepts, namely $Z = N$, C, C_a, M or B.

Proof. Given formula (4.4) (and the fact that Nash equilibrium payoffs are particular correlated equilibrium payoffs), we only need to show that

$$U(\Gamma_K) \subseteq \bigcup_{\mathcal{S}_K} N(G(\Gamma_K, \mathcal{S}_K))$$

and that

$$\bigcup_{\mathcal{S}_K} B(G(\Gamma_K, \mathcal{S}_K)) \subseteq U(\Gamma_K)$$

In order to establish the first part, let us fix a payoff z in $U(\Gamma_K)$. We must construct an information scheme $\mathcal{S}_K$ such that z is a Nash equilibrium payoff of the game $G(\Gamma_K, \mathcal{S}_K)$. z is associated with some types spaces S_1^*, S_2^* and a probability distribution P^* as above, which satisfy (6.1), (6.2) and its analogue for player 2. Given (6.1), (6.2) is equivalent to $\gamma_1(s_1)$ maximizes $\Sigma_{k,s_2} P^*(k, s_2 \mid s_1) u_1(k, a_1, \gamma_2(s_2))$ which can be rewritten as $\gamma_1(s_1)$ maximizes $\Sigma_{s_2} P^*(s_2 \mid s_1) v_1(s_1, s_2, a_1, \gamma_2(s_2))$ be defining v_1 exactly as in Section 2. We can proceed in the same way for player 2.

Let us consider the information scheme $\mathcal{S}_K$ defined by S_1^*, S_2^* and the marginal distribution of P^* over $K \times S_1^* \times S_2^*$. The previous analysis shows that the strategies $\gamma_i : S_i^* \to A_i$ form a Nash equilibrium of $G(\Gamma_K, \mathcal{S}_K)$. The payoff is z by construction.

To prove the second part of the proposition, let us start from an information scheme $\mathcal{S}_K = (T_1, T_2, P)$ and a payoff $z \in B(G(\Gamma_K, \mathcal{S}_K))$ as in Subsection 4.4. z is thus associated with a probability distribution Π over the states of the world and random variables τ_i, α_i $(i = 1, 2)$ satisfying the assumptions of Subsection 4.4. We shall only need to consider the distribution of $(\tau_1, \tau_2, \alpha_1, \alpha_2)$, over $T_1 \times T_2 \times A_1 \times A_2$, which we also write Π. We must construct S_1^*, S_2^*, P^* such that $z \in U(\Gamma_K)$.

Let us set $S_i^* = T_i \times A_i$ $(i = 1, 2)$. In this expression, A_i is not the set of actual actions of player i, but a copy of this set. In order to avoid confusion, we denote as a_i' an element of the actual actions set A_i and as (t_i, a_i) an element of S_i^*. We define the probability distribution P^* over $Y^* = K \times A_1 \times A_2 \times S_1^* \times S_2^*$ by

$$\begin{aligned} P^*(k, a_1', a_2', s_1, s_2) &= P^*(k, a_1', a_2', (t_1, a_1), (t_2, a_2)) \\ &= P(k \mid t_1, t_2)\Pi(t_1, t_2, a_1, a_2) \\ &\quad \times I(a_1' = a_1, a_2' = a_2) \end{aligned}$$

$$= P(k, t_1, t_2)\Pi(a_1, a_2 \mid t_1, t_2) \\ \times I(a_1' = a_1, a_2' = a_2)$$

The last equality follows from (4.1). The definition of P^* implies that Y^* is indeed a beliefs subspace and that the first two assumptions of the model of $U(\Gamma_K)$ are satisfied. In particular, (6.1) holds with $\gamma_i(s_i) = \gamma_i(t_i, a_i) = a_i$ (γ_i is just the canonical projection from S_i to A_i). There remains to check the rationality conditions, namely (6.2) and its analogue for player 2.

Given the above definitions, (6.2) can be rewritten as: $a_1' = a_1$ maximizes

$$\sum_{k,t_2,a_2} P(k \mid t_1, t_2)\Pi(t_2, a_2 \mid t_1, a_1)u_1(k, a_1', a_2)$$

This is exactly (4.3) if v_1 is defined as in Section 2. The same holds for player 2. ∎

A first immediate interpretation of Proposition 4 is that as soon as the beliefs of the players over the states of nature are not fixed by the model and may even be correlated with the beliefs over actions, all definitions of correlated equilibrium considered in Section 4 are *equivalent*. For every payoff z in $U(\Gamma_K)$, there is a game with private signals $G(\Gamma_K, \mathcal{S}_K)$ such that z is the payoff of a solution of this game. This holds for all the solution concepts studied here, from Nash equilibrium[3] to Bayesian solution, including all variants of correlated equilibrium. Obviously, the underlying game $G(\Gamma_K, \mathcal{S}_K)$ may vary with the solution concept. The payoff z may be a Nash equilibrium payoff in a game $G(\Gamma_K, \mathcal{S}_K)$, a communication equilibrium payoff in a (possibility different) game $G(\Gamma_K, \mathcal{S}_K')$ and the payoff of a Bayesian solution in $G(\Gamma_K, \mathcal{S}_K'')$.

To some extent, Proposition 4 depends on the fact that the information schemes $\mathcal{S}_K$ may involve extraneous signals (namely, signals which do not affect directly the payoffs). As we pointed out in Section 2, such signals are not precluded in the standard model of games with incomplete information: the (reduced form of the) games with private signals $G(\Gamma_K, \mathcal{S}_K)$ formally correspond to consistent

Bayesian games. Nevertheless, let us consider the particular case of a singleton set K. As in Section 2, we can write $\Gamma_K = \Gamma$; $\mathcal{S}_K = \mathcal{S}$ varies over *extraneous* information schemes. All signals are extraneous here, so that this notion is not ambiguous. In this context, Proposition 4 first states that

$$U(\Gamma) = \bigcup_{\mathcal{S}} N(G(\Gamma, \mathcal{S}))$$

which can be rewritten as

$$U(\Gamma) = C(\Gamma)$$

by the definition of Subsection 3.1. We find another version of Proposition 2. Indeed, as we already noticed in Subsection 3.3, the space of all states of the world which sustains the construction of $B(\Gamma)$ may be viewed as a beliefs subspace (of the universal beliefs space over $A_1 \times A_2$). Hence

$$B(\Gamma) = U(\Gamma)$$

Proposition 4 can thus certainly be interpreted as a generalization of Proposition 2 to games with incomplete information.

In the case of complete information, Proposition 4 further recalls us that the correlated equilibrium is the most general conceivable non-cooperative solution concept involving correlation and/or communication. For instance, the communication equilibria (*a fortiori*, the correlated equilibria) of extensions $G(\Gamma, \mathcal{S})$ of Γ yield the same outcomes as the correlated equilibria of Γ.

No extraneous signal is used in the construction of $B(\Gamma)$ (or $U(\Gamma)$). On the contrary, Aumann's approach is precisely a justification of correlated equilibria which does not appeal to explicit extraneous signals. A similar effect is achieved by the players' hierarchies of beliefs over actions. The same holds for $U(\Gamma_K)$ in the case of incomplete information. In this context, one may either fix a Bayesian game as in Subsection 4.4 or construct all beliefs at once as we did here. The consequence of proceeding in this way is that the construction does not yield a single well-defined Bayesian game but a family of games; and the same payoff may correspond to different solution concepts applied to different games. Let us illustrate this by a simple example (similar to Example 4 above).

$$K = \{1, 2\}, A_1 = \{a_1\}, A_2 = \{a_2, a_2'\}$$

and the utility functions u_i $(i = 1, 2)$ are described by the following payoff matrices.

$$k = 1 \qquad \begin{array}{cc} & \begin{array}{cc} a_2 & a_2' \end{array} \\ t_1 & (1, 2 \quad 0, 0) \end{array}$$

$$k = 2 \qquad \begin{array}{cc} & \begin{array}{cc} a_2 & a_2' \end{array} \\ t_1' & (1, 0 \quad 0, 4) \end{array}$$

As we showed in Section 5, the payoff $(1/2, 3)$ can be achieved by a *Bayesian solution* of the game where player 1 knows the state of nature, player 2 does not know it but assigns the same probability to each state of nature and these beliefs are common knowledge.

The same payoff $(1/2, 3)$ can also be achieved as a *Nash* equilibrium of another game with private signals, which is also consistent with the two-person decision problem described above. Indeed, it suffices to assume that player 2 *knows* the state of nature, so that he can choose a_2 if $k = 1$ and a_2' if $k = 2$.

In this example, the identification of the underlying game does not depend on the distinction of extrinsic and intrinsic signals. The problem rather comes from the lack of dynamics in Aumann's approach. The beliefs subspace Y^*, the probability distribution P^* and the rationality conditions summarize a possibly very rich information and many scenarios (namely, games and solutions) can be constructed which are consistent with the summary Y^*, P^*. As we already mentioned above, the model of a precise but arbitrary timing (as in Harsanyi's approach, for instance) appears as a false solution to the problem.

7. FINAL REMARKS

In this study we have respected Aumann's (1987) framework as closely as possible, because our main goal was to extend the results of this paper. In particular, we did not carefully formalize some notions, as *knowledge* and *common knowledge* (of the underlying model or of rationality). The references on this subject abound (see e.g. the surveys of Binmore and Brandenburger, 1990, and of Brandenburger and Dekel, 1990).

As Harsanyi (1967–1968) and Aumann (1987), we insisted on the *common prior assumption*. But both authors showed that their respective approaches could be extended to an inconsistent world. Aumann (1974, 1987) introduced *subjective correlated equilibria*. This solution concept was studied e.g. in Brandenburger and Dekel (1987), Brandenburger, Dekel and Geanokoplos (1992) and Nau (1991). Obviously, the universal beliefs space of Mertens and Zamir (1985) contains infinitely many inconsistent states of the world.

Finally, we did not exploit very far from intuitive consequences of the incomplete information model. The correlated equilibrium concept is based on *extrinsic* possibilities of coordination. However, intrinsic lacks and differences of information also offer correlation possibilities and it seems that *games with incomplete information should favor correlated behavior*. This is partly hidden in the equivalence result stated as Proposition 4. We also pointed out the difficulty of the distinction between extrinsic and intrinsic information. Furthermore, the players cannot usually coordinate on the latter without risk. Anyway, the use of incomplete information for coordination purposes may be a subject of future research.

NOTES

[1] The fact that the Bayesian approach was initiated in models of incomplete information and later exploited in the context of complete information is clearly acknowledged in Brandenburger and Dekel (1990).

[2] Π must be a *supporting* probability distribution for every player, namely must assign a positive expected value to every *belief gamble* a player may accept.

[3] As far as Nash equilibrium property is concerned, an analogous property was derived in Mertens, Sorin and Zamir (forthcoming, Ch. 3).

REFERENCES

Armbruster, W. and Böge W.: 1979, 'Bayesian game theory' in Moeschlin and Pallaschke (Eds.), *Game Theory and Related Topics*, North Holland, pp. 17–28.

Aumann, R.: 1974, 'Subjectivity and correlation in randomized strategies', *Journal of Mathematical Economics* **1**, 67–96.

Aumann, R.: 1987, 'Correlated equilibrium as an expression of Bayesian rationality', *Econometrica* **55**, 1–18.

Bernheim, D.: 1986, 'Axiomatic characterizations of rational choice in strategic environment', *Scandinavian Journal of Economics* **88**, 473–488.

Binmore, K. and Brandenburger, A.: 1990, 'Common knowledge and game theory' in K. Binmore, *Essays on the Foundations of Game Theory*, Basil Blackwell, Cambridge, Massachusetts.

Böge, W. and Eisele, T.H.: 1979, 'On solutions of Bayesian games', *International Journal of Game Theory* **8**, 193–215.

Brandenburger, A. and Dekel, E.: 1987, 'Rationalizability and correlated equilibria', *Econometrica* **55**, 1391–1402.

Brandenburger, A. and Dekel, E.: 1990, 'The role of common knowledge assumptions in game theory' in F. Hahn (Ed.), *The Economics of Missing Markets, Information and Games*, Oxford University Press, pp. 46–61.

Brandenburger, A., Dekel, E., and Geanokoplos, J.: 1992, 'Correlated equilibrium with generalized information structures', *Games and Economic Behavior* **4**, 182–201.

Cotter, K.: 1991, 'Correlated equilibrium in games with type-dependent strategies', *Journal of Economic Theory* **54**, 48–68.

Einy, E. and Peleg, B.: 1991, 'Coalition-proof communication equilibria', mimeo, Hebrew University of Jerusalem.

Forges, F.: 1985, 'Correlated equilibria in a class of repeated games with incomplete information', *International Journal of Game Theory* **14**, 129–150.

Forges, F.: 1986a, 'Correlated equilibria in repeated games with lack of information on one side: a model with verifiable types', *International Journal of Game Theory* **15**, 65–82.

Forges, F.: 1986b, 'An approach to communication equilibria', *Econometrica* **54**, 159–182.

Forges, F.: 1990, 'Universal mechanisms', *Econometrica* **58**, 1341–1364.

Fudenberg, D. and Tirole, J.: 1991, *Game Theory*, MIT Press, Cambridge, Massachusetts.

Harsanyi, J.: 1967–1968, 'Games with incomplete information played by Bayesian players', *Management Science* **14**, 159–182 (Part I); 320–334 (Part II); 486–502 (Part III).

Mertens, J-F. and Zamir, S.: 1985, 'Formulation of Bayesian analysis for games with incomplete information', *International Journal of Game Theory* **14**, 1–29.

Mertens, J-F., Sorin, S., and Zarmir, S. (forthcoming): 'Repeated games'.

Myerson, R.: 1982, 'Optimal coordination mechanisms in generalized principal-agent problems', *Journal of Mathematical Economics* **10**, 67–81.

Myerson, S.: 1991, *Game Theory: Analysis of Conflict*, Harvard University Press, Cambridge, Massachusetts.

Nau, R.: 1991, 'The relativity of subjective and objective correlated equilibria', Fuqua School of Business, Working Paper 9101, Duke University.

Nau, R.: 1992, 'Joint coherence in games with incomplete information', *Management Science* **38**, 374–387.

Nau, R. and McCardle, K.: 1990, 'Coherent behavior in noncooperative games', *Journal of Economic Theory* **50**, 424–444.

Samuelson, L. and Zhang, J.: 1989, 'Correlated equilibria and mediated equilibria in games with incomplete information', mimeo, Penn State University.
Tan, T.C.C. and Werlang, S.R.C.: 1988, 'The Bayesian foundations of solution concepts of games', *Journal of Economic Theory* **45**, 370–391.

CORE, Université Catholique de Louvain,
34 Voie du Roman Pays,
1348 Louvain-la-Neuve, Belgium.