Chapter 24

COMMUNICATION, CORRELATED EQUILIBRIA AND INCENTIVE COMPATIBILITY

ROGER B. MYERSON

Northwestern University

Contents

Handbook of Game Theory, Volume 2, Edited by R.J. Aumann and S. Hart

1. Correlated equilibria of strategic-form games

It has been argued [at least since von Neumann and Morgenstern (1944)] that there is no loss of generality in assuming that, in a strategic-form game, all players choose their strategies simultaneously and independently. In principle, anything that a player can do to communicate and coordinate with other players could be described by moves in an extensive-form game, so that planning these communication moves would become part of his strategy choice itself.

Although this perspective may be fully general in principle, it is not necessarily the most fruitful way to think about all games. There are many situations where the possibilities for communication are so rich that to follow this modeling program rigorously would require us to consider enormously complicated games. For example, to model player 1's opportunity to say just one word to player 2, player 1 must have a move and player 2 must have an information state for every word in the dictionary! In such situations, it may be more useful to leave communication and coordination possibilities out of the explicit model. If we do so, then we must instead use solution concepts that express an assumption that players have implicit communication opportunities, in addition to the strategic options explicitly described in the game model. We consider here such solution concepts and show that they can indeed offer important analytical insights into many situations.

So let us say that a game is *with communication* if, in addition to the strategy options explicitly specified in the structure of the game, the players have a very wide range of implicit options to communicate with each other. We do not assume here that they have any ability to sign contracts; they can only talk. Aumann (1974) showed that the solutions for games with communication may have remarkable properties, even without contracts.

Consider the two-player game with payoffs as shown in table 1, where each player i must choose x_i or y_i (for $i=1,2$). Without communication, there are three equilibria of this game: (x_1, x_2) which gives the payoff allocation $(5,1)$; (y_1, y_2) which gives the payoff allocation $(1,5)$; and a randomized equilibrium which gives the expected payoff allocation $(2.5, 2.5)$. The best symmetric payoff allocation $(4,4)$ cannot be achieved by the players without contracts, because (y_1, x_2) is not an equilibrium. However, even without binding contracts, communication may allow the players to achieve an expected payoff allocation that is better for both than

Table 1

	x_2	y_2
x_1	5,1	0,0
y_1	4,4	1,5

(2.5, 2.5). Specifically, the players may plan to toss a coin and choose (x_1, x_2) if heads occurs or (y_1, y_2) if tails occurs. Even though the coin has no binding force on the players, such a plan is *self-enforcing*, in the sense that neither player could gain by unilaterally deviating from this plan.

With the help of a *mediator* (that is, a person or machine that can help the players communicate and share information), there is a self-enforcing plan that generates the even better expected payoff allocation $(3\frac{1}{3}, 3\frac{1}{3})$. To be specific, suppose that a mediator randomly recommends strategies to the two players in such a way that each of the pairs (x_1, x_2), (y_1, y_2), and (y_1, x_2) may be recommended with probability $\frac{1}{3}$. Suppose also that each player learns only his own recommended strategy from the mediator. Then, even though the mediator's recommendation has no binding force, there is a Nash equilibrium of the transformed game with mediated communication in which both players plan always to obey the mediator's recommendations. If player 1 heard the recommendation "y_1" from the mediator, then he would think that player 2 may have been told to do x_2 or y_2 with equal probability, in which case his expected payoff from y_1 would be as good as from x_1 (2.5 from either strategy). If player 1 heard a recommendation "x_1" from the mediator then he would know that player 2 was told to do x_2, to which his best response is x_1. So player 1 would always be willing to obey the mediator if he expected player 2 to obey the mediator, and a similar argument applies to player 2. That is, the players can reach a self-enforcing understanding that each obey the mediator's recommendation when he plans to randomize in this way. Randomizing between (x_1, x_2), (y_1, y_2), and (y_1, x_2) with equal probability gives the expected payoff allocation

$$\tfrac{1}{3}(5,1) + \tfrac{1}{3}(4,4) + \tfrac{1}{3}(1,5) = (3\tfrac{1}{3} 3\tfrac{1}{3}).$$

Notice that the implementation of this correlated strategy $(\frac{1}{3}[x_1, x_2] + \frac{1}{3}[y_1, y_2] + \frac{1}{3}[y_1, x_2]$ without contracts required that each player get different partial information about the outcome of the mediator's randomization. If player 1 knew when player 2 was told to choose x_2, then player 1 would be unwilling to choose y_1 when it was also recommended to him. So this correlated strategy could not be implemented without some kind of mediation or noisy communication. With only direct unmediated communication in which all players observe anyone's statements or the outcomes of any randomization, the only self-enforcing plans that the players could implement without contracts would be randomizations among the Nash equilibria of the original game (without communication), like the correlated strategy $0.5[x_1, x_2] + 0.5[y_1, y_2]$ that we discussed above. However, Barany (1987) and Forges (1990) have shown that, in any strategic-form and Bayesian game with four or more players, a system of direct unmediated communication between pairs of players can simulate any centralized communication system with a mediator, provided that the communication between any pair of players is not directly observable by the other players. (One of the essential ideas behind this result is that, when there are four or more players, each pair of players can use two other players as parallel mediators. The messages that two mediating players carry can be suitably encoded so that neither of the mediating players can, by himself, gain by corrupting his message or learn anything useful from it.)

Consider now what the players could do if they had a bent coin for which player 1 thought that the probability of heads was 0.9 while player 2 thought that the probability of heads was 0.1, and these assessments were common knowledge. With this coin, it would be possible to give each player an expected payoff of 4.6 by the following self-enforcing plan: toss the coin, and then implement the (x_1, x_2) equilibrium if heads occurs, and implement the (y_1, y_2) equilibrium otherwise.

However, the players beliefs about this coin would be *inconsistent*, in the sense of Harsanyi (1967, 1968). That is because there is no way to define a prior probability distribution for the outcome of the coin toss and two other random variables such that player 1's beliefs are derived by Bayes's formula from the prior and his observation of one of the random variables, player 2's beliefs are derived by Bayes's formula from the prior and her observation of the other random variable, and it is common knowledge that they assign different probabilities to the event that the outcome of the coin toss will be heads. (See Aumann (1976) for a general proof of this fact).

The existence of such a coin, about which the players have inconsistent beliefs, would be very remarkable and extraordinary. With such a coin, the players could make bets with each other that would have arbitrarily large positive expected monetary value to both! Thus, as a pragmatic convention, let us insist that the existence of any random variables about which the players may have such inconsistent beliefs should be explicitly listed in the structure of the game, and should not be swept into the implicit meaning of the phrase "game with communication." (These random variables could be explicitly modelled either by a Bayesian game with beliefs that are not consistent with any common prior, or by a game in generalized extensive form, where a distinct subjective probability distribution for each player could be assigned to the set of alternatives at each chance node.) When we say that a particular game is played "with communication" we mean only that the players can communicate with each other and with outside mediators, and that players and mediators have implicit opportunities to observe random variables that have objective probability distributions about which everyone agrees.

In general, consider any finite strategic-form game $\Gamma = (N, (C_i)_{i\in N}, (u_i)_{i\in N})$, where N is the set of players, C_i is the set of pure strategies for player i, and $u_i\colon C \rightarrow \mathbb{R}$ is the utility payoff function for player i. We use here the notation

$$C = \mathop{\times}_{i\in N} C_i.$$

A mediator who was trying to help coordinate the players's actions would have (at least) to tell each player i which strategy in C_i was recommended for him. Assuming that the mediator can communicate separately and confidentially with each player, no player needs to be told the recommendations for any other players.

Without contracts, player i would then be free to choose any strategy in C_i after hearing the mediator's recommendation. So in the game with mediated communication, each player i would actually have an enlarged set of communication strategies that would include all mappings from C_i into C_i. each of which represents a possible rule for choosing an element of C_i to implement as a function of the mediator's recommendation in C_i.

Now, suppose that it is common knowledge that the mediator will determine his recommendations according to the probability distribution μ in $\Delta(C)$, so that $\mu(c)$ denotes the probability that any given pure strategy profile $c = (c_i)_{i\in N}$ would be recommended by the mediator. (For any finite set X, we let $\Delta(X)$ denote the set of probability distributions over X.) The expected payoff to player i under this correlated strategy μ, if everyone obeys the recommendations, is

$$U_i(\mu) = \sum_{c\in C} \mu(c)\, u_i(c).$$

Then it would be an equilibrium for all players to obey the mediator's recommendations iff

$$U_i(\mu) \geqslant \sum_{c\in C} \mu(c)\, u_i(c_{-i}, \delta_i(c_i)), \quad \forall i\in N, \quad \forall \delta_i{:}C_i \to C_i, \tag{1}$$

where $U_i(\mu)$ is as defined in (6.1). [Here $(c_{-i}, \delta_i(c))$ denotes the pure strategy profile that differs from c only in that the strategy for player i is changed to $\delta_i(c_i)$.] Following Aumann (1974, 1987), we say that μ is a *correlated equilibrium* of Γ iff $\mu\in\Delta(C)$ and μ satisfies condition (1). That is, a correlated equilibrium is any correlated strategy for the players in Γ that could be self-enforcingly implemented with the help of a mediator who can make nonbinding confidential recommendations to each player. The existence of correlated equilibria can be derived from the general existence of Nash equilibria for finite games, but elegant direct proofs of the existence of correlated equilibria have also been given by Hart and Scheidler (1989) and Nau and McCardle (1990).

It can be shown that condition (1) is equivalent to the following system of inequalities:

$$\sum_{c_{-i}\in C_{-i}} \mu(c)[u_i(c) - u_i(c_{-i}, e_i)] \geqslant 0, \forall i\in N, \quad \forall c_i\in C_i, \quad \forall e_i\in C_i. \tag{2}$$

[Here $C_{-i} = \times_{j\in N-i} C_j$, and $c = (c_{-i}, c_i)$.] To interpret this inequality, notice that, given any c_i, dividing the left-hand side by

$$\sum_{c_{-i}\in C_{-i}} \mu(c),$$

would make it equal to the difference between player i's conditionally expected payoff from obeying the mediator's recommendation and his conditionally expected payoff from using the action e_i, given that the mediator has recommended c_i. Thus, (2) asserts that no player i could expect to increase his expected payoff by using some disobedient action e_i after getting any recommendation c_i from the mediator. These inequalities (1) and (2) may be called *strategic incentive constraints*, because they represent the mathematical inequalities that a correlated strategy must satisfy to guarantee that all players could rationally obey the mediator's recommendations.

The set of correlated equilibria is a compact and convex set, for any finite game in strategic form. Furthermore, it can be characterized by a finite collection of linear inequalities, because a vector μ in $\mathbb{R}^N$ is a correlated equilibrium iff it satisfies the strategic incentive constraints (2) and the following *probability constraints*:

$$\sum_{e \in C} \mu(e) = 1 \text{ and } \mu(c) \geqslant 0, \quad \forall c \in C. \tag{3}$$

Thus, for example, if we want to find the correlated equilibrium that maximizes the sum of the players' expected payoffs in Γ, we have a problem of maximizing a linear objective $[\sum_{i \in N} U_i(\mu)]$ subject to linear constraints. This is a linear programming problem, which can be solved by any one of many widely-available computer programs. [See also the general conditions for optimality subject to incentive constraints developed by Myerson (1985).]

For the game in table 1, the correlated equilibrium μ that maximizes the expected sum of the players' payoffs is

$$\mu(x_1, x_2) = \mu(y_1, x_2) = \mu(y_1, y_2) = \tfrac{1}{3} \text{ and } \mu(x_1, y_2) = 0.$$

That is, $\mu = \frac{1}{3}[x_1, x_2] + \frac{1}{3}[y_1, y_2] + \frac{1}{3}[y_1, x_2]$ maximizes the sum of the players' expected payoffs $U_1(\mu) + U_2(\mu)$ subject to the strategic incentive constraints (2) and the probability constraints (3). So the strategic incentive constraints imply that the players' expected sum of payoffs cannot be higher than $3\frac{1}{3} + 3\frac{1}{3} = 6\frac{2}{3}$.

It may be natural to ask why we have been focusing attentions on mediated communication systems in which it is rational for all players to obey the mediator. The reason is that such communication systems can simulate any equilibrium of any game that can be generated from any given strategic-form game by adding any communication system. To see why, let us try to formalize a general framework for describing communication systems that might be added to a given strategic-form game Γ as above. Given a communication system, let R_i denote the set of all strategies that player i could use to determine the reports that he sends out, into the communication system, and let M_i denote the set of all messages that player i could receive from the communication system. For any $r = (r_i)_{i \in N}$ in $R = \times_{i \in N} R_i$ and any $m = (m_i)_{i \in N}$ in $M = \times_{i \in N} M_i$, let $\nu(m|r)$ denote the conditional probability that m would be the messages received by the various players if each player i were sending reports according to r_i. This function $\nu: R \to \Delta(M)$ is our basic mathematical characterization of the communication system. (If all communication is directly between players, without noise or mediation, then every player's message would be composed directly of other players' reports to him, and so $\nu(\cdot|r)$ would always put probability 1 on some vector m; but noisy communication or randomized mediation allows $0 < \nu(m|r) < 1$.)

Given such a communication system, the set of pure communication strategies that player i can use for determining the reports that he sends and the action in C_i that he ultimately implements (as a function of the messages that he receives) is

$$B_i = \{(r_i, \delta_i) | r_i \in R_i, \quad \delta_i : M_i \to C_i\}.$$

Player i's expected payoff depends on the communication strategies of all players according to the function $\bar{u}_i$, where

$$\bar{u}_i((r_j, \delta_j)_{j \in N}) = \sum_{m \in M} v(m|r)\, u_i((\delta_j(m_j))_{j \in N}).$$

Thus, the communication system $v: R \to \Delta(M)$ generates a *communication game* Γ_v, where

$$\Gamma_v = (N, (B_i)_{i \in N}, (\bar{u}_i)_{i \in N}).$$

This game Γ_v is the appropriate game in strategic form to describe the structure of decision-making and payoffs when the game Γ has been transformed by allowing the players to communicate through the communication system v before choosing their ultimate payoff-relevant actions. To characterize rational behavior by the players in the game with communication, we should look among the equilibria of Γ_v.

However, any equilibrium of Γ_v is equivalent to a correlated equilibrium of Γ as defined by the strategic incentive constraints (2). To see why, let $\sigma = (\sigma_i)_{i \in N}$ be any equilibrium in randomized strategies of this game Γ_v.

Let μ be the correlated strategy in $\Delta(C)$ defined by

$$\mu(c) = \sum_{(r,\delta) \in B} \sum_{m \in \delta^{-1}(c)} \left(\prod_{i \in N} \sigma_i(r_i, \delta_i) \right) v(m|r), \quad \forall c \in C,$$

where we use the notation:

$$B = \times_{i \in N} B_i, \quad (r, \delta) = ((r_i, \delta_i)_{i \in N}),$$

$$\delta^{-1}(c) = \{m \in M | \delta_i(m_i) = c_i, \quad \forall i \in N\}.$$

That is, the probability of any outcome c in C under the correlated strategy μ is just the probability that the players would ultimately choose this outcome after participating in the communication system v, when early player determines his plan for sending reports and choosing actions according to σ. So μ effectively simulates the outcome that results from the equilibrium σ in the communication game Γ_v. Because μ is just simulating the outcomes from using strategies σ in Γ_v, if some player i could have gained by disobeying the mediator's recommendations under μ, when all other players are expected to obey, then he could have also gained by similarly disobeying the recommendations of his own strategy σ_i when applied against σ_{-i} in Γ_v. More precisely, if (1) were violated for some i and δ_i, then player i could gain by switching from σ_i to $\hat{\sigma}_i$ against σ_{-i} in Γ_v, where

$$\hat{\sigma}_i(r_i, \gamma_i) = \sum_{\zeta_i \in Z(\delta_i, \gamma_i)} \sigma_i(r_i, \zeta_i), \quad \forall (r_i, \gamma_i) \in B_i, \quad \text{and}$$

$$Z(\delta_i, \gamma_i) = \{\zeta_i \,|\, \delta_i(\zeta_i(m_i)) = \gamma_i(m_i), \quad \forall m_i \in M_i\}.$$

This conclusion would violate the assumption that σ is an equilibrium. So μ must satisfy the strategic incentive constraints (1), or else σ could not be an equilibrium of Γ_v.

Thus, any equilibrium of any communication game that can be generated from a strategic-form game Γ by adding a system for preplay communication must be equivalent to a correlated equilibrium satisfying the strategic incentive constraints (1) or (2). This fact is known as the *revelation principle* for strategic-form games. [See Myerson (1982).]

For any communication system v, there may be many equilibria of the communication game Γ_v, and these equilibria may be equivalent to different correlated equilibria. In particular, for any equilibrium $\bar{\sigma}$ of the original game Γ, there are equilibria of the communication game Γ_v in which every player i chooses a strategy in C_i according to $\bar{\sigma}_i$, independently of the reports that he sends or the messages that he receives. (One such equilibrium σ of Γ_v could be defined so that

$$\text{if } \delta_i(m_i) = c_i, \quad \forall m_i \in M_i, \text{ then } \sigma_i(r_i, \delta_i) = \bar{\sigma}_i(c_i)/|R_i|,$$

and

$$\text{if } \exists \{m_i, \hat{m}_i\} \subseteq M_i \text{ such that } \delta_i(m_i) \neq \delta_i(\hat{m}_i) \text{ then } \sigma_i(r_i, \delta_i) = 0.)$$

That is, adding a communication system does not eliminate any of the equilibria of the original game, because there are always equilibria of the communication game in which reports and messages are treated as having no meaning and hence are ignored by all players. Such equilibria of the communication game are called *babbling equilibria.*

The set of correlated equilibria of a strategic-form game Γ has a simple and tractable mathematical structure, because it is closed by convex and is characterized by a finite system of linear inequalities. On the other hand, the set of Nash equilibria of Γ, or of any specific communication game that can be generated from Γ, does not generally have any such simplicity of structure. So the set of correlated equilibria, which characterizes the union of the sets of equilibria of all communication games that can be generated from Γ, may be easier to analyze than the set of equilibria of any one of these games. This observation demonstrates the analytical power of the revelation principle. That is, the general conceptual approach of accounting for communication possibilities in the solution concept, rather than in the explicit game model, not only simplifies our game models but also generates solutions that are much easier to analyze.

To emphasize the fact that the set of correlated equilibria may be strictly larger

Table 2

	x_2	y_2	z_2
x_1	0.0	5.4	4.5
y_1	4.5	0.0	5.4
z_1	5.4	4.5	0.0

than the convex hull of the set of Nash equilibria, it may be helpful to consider the game in table 2, which was studied by Moulin and Vial (1978).

This game has only one Nash equilibrium,

$$(\tfrac{1}{3}[x_1]+\tfrac{1}{3}[y_1]+\tfrac{1}{3}[z_1],\ \ \tfrac{1}{3}[x_2]+\tfrac{1}{3}[y_2]+\tfrac{1}{3}[z_2]),$$

which gives expected payoff allocation (3,3). However, there are many correlated equilibria, including

$$((\tfrac{1}{6}[(x_1,y_2)]+\tfrac{1}{6}[(x_1,z_2)]+\tfrac{1}{6}[(y_1,x_2)]+\tfrac{1}{6}[(y_1,z_2)]+\tfrac{1}{6}[(z_1,x_2)]+\tfrac{1}{6}[(z_1,y_2)]]),$$

which gives expected payoff allocation (4.5,4.5).

2. Incentive-compatible mechanisms for Bayesian games

The revelation principle for strategic-form games asserted that any equilibrium of any communication system can be simulated by a communication system in which the only communication is form a central mediator to the players, without any communication from the players to the mediator. The one-way nature of this communication should not be surprising, because the players have no private information to tell the mediator about, within the structure of the strategic-form game. More generally, however, players in Bayesian game [as defined by Harsanyi (1967, 1968)] may have private information about their types, and two-way communication would then allow the players' actions to depend on each others' types, as well as on extraneous random variables like coin tosses. Thus, in Bayesian games with communication, there may be a need for players to talk as well as to listen in mediated communication systems. [See Forges (1986).]

Let $\Gamma^b=(N,\ (C_i)_{i\in N},\ (T_i)_{i\in N},\ (p_i)_{i\in N},\ (u_i)_{i\in N})$, be a finite Bayesian game with incomplete information. Here N is the set of players, C_i is the set of possible actions for player i, and T_i is the set of possible types (or private information states) for player i. For any $t=(t_j)_{j\in N}$ in the set $T=\times_{i\in N}T_i$, $p_i(t_{-i}|t_i)$ denotes the probability that player i would assign to the event that $t_{-i}=(t_j)_{j\in N-i}$ is the profile of types for the players other than i if t_i were player i's type. For any t in T and any $c=(c_j)_{j\in N}$ in the set $C=\times_{j\in N}C_j$, $u_i(c,t)$ denotes the utility payoff that player i would get if c were the profile of actions chosen by the players and t were the profile of their actual types. Let us suppose now that Γ^b is a game with communication, so that the players have wide opportunities to communicate, after each player i learns his type in T_i but before he chooses his action in C_i.

Consider mediated communication systems of the following form: first, each player is asked to report his type confidentially to the mediator; then, after getting these reports, the mediator confidentially recommends an action to each player. The mediator's recommendations may depend on the players' reports in a deterministic or random fashion. For any c in C and any t and T, let $\mu(c|t)$ denote the conditional probability that the mediator would recommend to each player i that he should use action c_i, if each player j reported his type to be t_j. Obviously, these numbers $\mu(c|t)$ must satisfy the following *probability constraints*

$$\sum_{c\in C} \mu(c|t) = 1 \text{ and } \mu(d|t) \geqslant 0, \quad \forall d\in C, \quad \forall t\in T. \tag{4}$$

In general, any such function $\mu: T \to \Delta(C)$ may be called a *mediation plan* or *mechanism* for the game Γ^b with communication.

If every player reports his type honestly to the mediator and obeys the recommendations of the mediator, then the expected utility for type t_i of player i from the plan μ would be

$$U_i(\mu|t_i) = \sum_{t_{-i}\in T_{-i}} \sum_{c\in C} p_i(t_{-i}|t_i)\,\mu(c|t)\,u_i(c,t),$$

where $T_{-i} = \times_{j\in N-i} T_j$ and $t = (t_{-i}, t_i)$.

We must allow, however, that each player could lie about his type or disobey the mediator's recommendation. That is, we assume here that the players' types are not verifiable by the mediator, and the choice of an action in C_i can be controlled only by player i. Thus, a mediation plan μ induces a communication game Γ^b_μ in which each player must select his type report and his plan for choosing an action in C_i as a function of the mediator's recommendation. Formally Γ^b_μ is itself a Bayesian game, of the form

$$\Gamma^b_\mu = (N, (B_i)_{i\in N}, (T_i)_{i\in N}, (p_i)_{i\in N}, (\bar{u}_i)_{i\in N}),$$

where, for each player i,

$$B_i = \{(s_i, \delta_i) | s_i \in T_i, \quad \delta_i: C_i \to C_i\},$$

and $\bar{u}_i: (\times_{i\in N} B_j) \times T \to \mathbb{R}$ is defined by the equation

$$\bar{u}_i((s_j, \delta_j)_{j\in N}, t) = \sum_{c\in C} \mu(c|(s_j)_{j\in N})\, u_i((\delta_j(c_j))_{j\in N}, t).$$

A strategy (s_i, δ_i) in B_i represents a plan for player i to report s_i to the mediator, and then to choose his action in C_i as a function of the mediator's recommendation according to δ_i, so that he would choose $\delta_i(c_i)$ if the mediator recommended c_i. The action that player i chooses cannot depend on the type-reports or recommended actions of any other player, because each player communicates with the mediator separately and confidentially.

Suppose, for example, that the true type of player i were t_i, but that he used the strategy (s_i, δ_i) in the communication game Γ^b_μ. If all other players were honest and obedient to the mediator, then i's expected utility payoff would be

$$U_i^*(\mu, \delta_i, s_i | t_i) = \sum_{t_{-i} \in T_{-i}} \sum_{c \in C} p_i(t_{-i} | t_i)\, \mu(c | t_{-i}, s_i)\, u_i((c_{-i}, \delta_i(c_i)), t).$$

[Here (t_{-i}, s_i) is the vector in T that differs from $t = (t_{-i}, t_i)$ only in that the i-component is s_i instead of t_i.]

Bayesian equilibrium [as defined in Harsanyi (1967–68)] is still an appropriate solution concept for a Bayesian game with communication, except that we must now consider the Bayesian equilbria of the induced communication game Γ^b_μ, rather than just the Bayesian equilibria of Γ. We say that a mediation plan μ is *incentive compatible* iff it is a Bayesian equilibrium for all players to report their types honestly and to obey the mediator's recommendations when he uses the mediation plan μ. Thus, μ is incentive compatible iff it satisfies the following general *incentive constraints*:

$$U_i(\mu | t_i) \geqslant U_i^*(\mu, \delta_i, s_i | t_i), \quad \forall i \in N, \quad \forall t_i \in T_i, \quad \forall s_i \in T_i, \quad \forall \delta_i : C_i \to C_i. \tag{5}$$

If the mediator uses an incentive-compatible mediation plan and each player communicates independently and confidentially with the mediator, then no player could gain by being the only one to lie to the mediator or disobey his recommendations. Conversely, we cannot expect rational and intelligent players all to participate honestly and obediently in a mediation plan unless it is incentive compatible.

In general, there may be many different Bayesian equilibria of a communication game Γ^b_μ, even if μ is incentive compatible. Furthermore, as in the preceding section, we could consider more general communication systems, in which the reports that player i can send and the messages that player i may receive are respectively in some arbitrary sets R_i and M_i, not necessarily T_i and C_i. However, given any general communication system and any Bayesian equilibrium of the induced communication game, there exists an equivalent incentive-compatible mediation plan, in which every type of every player gets the same expected utility as in the given Bayesian equilibrium of the induced communication game. In this sense, there is no loss of generality in assuming that the players communicate with each other through a mediator who first asks each player to reveal all of his private information and who then reveals to each player only the minimum information needed to guide his action, in such a way that no player has any incentive to lie or disobey. This result is the *revelation principle* for general Bayesian games.

The formal proof of the revelation principle for Bayesian games is almost the same as for strategic-form games. Given a general communication system $v: R \to \Delta(M)$ and communication strategy sets $(B_i)_{i \in N}$ as in Section 1 above, a Bayesian equilibrium of the induced communication game would then be a vector σ that specifies, for each i in N, each (r_i, δ_i) in B_i, and each t_i in T_i, a number $\sigma_i(r_i, \delta_i | t_i)$ that represents the probability that i would report r_i and choose his final action according to δ_i (as a function of the message that he receives) if his actual type were t_i. If σ is such a Bayesian equilibrium of the communication game Γ^b_ν induced by the communication system ν, then we can construct an equivalent incentive-compatible mediation plan μ by letting

$$\mu(c|t) = \sum_{(r,d)\in B} \sum_{m\in\delta^{-1}(c)} \left(\prod_{i\in N} \sigma_i(r_i, \delta_i|t_i)\right)\nu(m|r), \quad \forall c\in C, \quad \forall t\in T,$$

where $\delta^{-1}(c) = \{m\in M|\delta_i(m_i) = c_i, \quad \forall i\in N\}$.

This construction can be described more intuitively as follows. The mediator first asks each player (simultaneously and confidentially) to reveal his type. Next the mediator computes (or simulates) the reports that would have been sent by the players, with these revealed types, under the given equilibrium. Then he computes the recommendations or messages that would have been received by the players, as a function of these reports, under the given communication system or mechanism. Then he computes the actions that would have been chosen by the players, as a function of these messages and the revealed types in the given equilibrium. Finally, the mediator tells each player to do the action computed for him at the last step. Thus, the constructed mediation plan simulates the given equilibrium of the given communication system. To check that this constructed mediation plan is incentive compatible, notice that any type of any player who could gain by lying to the mediator or disobeying his recommendations under the constructed mediation plan (when everyone else is honest and obedient) could also gain by similarly lying to himself before implementing his equilibrium strategy or disobeying his own recommendations to himself after implementing his equilibrium strategy in the given communication game, which is impossible (by definition of a Bayesian equilibrium).

If each player's type set consists trivially of only one possible type, so that the Bayesian game is essentially equivalent to a strategic-form game, then an incentive-compatible mechanism is a correlated equilibrium. So incentive-compatible mechanisms are a generalization of correlated equilibria to the case of games with incomplete information. Thus, we may synonymously use the term *communication equilibrium* (or *generalized correlated equilibrium*) to refer to any incentive-compatible mediation plan of a Bayesian game.

Like the set of correlated equilibria, the set of incentive-compatible mediation plans is a closed convex set, characterized by a finite system of inequalities [(4) and (5)] that are linear in μ. On the other hand, it is generally a difficult problem to characterize the set of Bayesian equilibria of any given Bayesian game. Thus, by the revelation principle, it may be easier to characterize the set of all equilibria of all games that can be induced from Γ^b with communication than it is to compute the set of equilibria of Γ^b or of any one communication game induced from Γ^b.

For a simple, two-player example, suppose that $C_1 = \{x_1, y_1\}$, $C_2 = \{x_2, y_2\}$, $T_1 = \{1.0\}$ (so that player 1 has only one possible type and no private information),

Table 3

	$t_2 = 2.1$			$t_2 = 2.2$	
	x_2	y_2		x_2	y_2
x_1	1,2	0,1	x_1	1,3	0,4
y_1	0,4	1,3	y_1	0,1	1,2

$T_2 = \{2.1, 2.2\}$, $p_1(2.1 \mid 1.0) = 0.6$, $p_1(2.2 \mid 1.0) = 0.4$, and the utility payoffs (u_1, u_2) depend on the actions and player 2's type as in table 3.

In this game, y_2 is a strongly dominated action for type 2.1, and x_2 is a strongly dominated action for type 2.2, so 2.1 must choose x_2 and 2.2 must choose y_2 in a Bayesian equilibrium. Player 1 wants to get either (x_1, x_2) or (y_1, y_2) to be the outcome of the game, and he thinks that 2.1 is more likely than 2.2. Thus the unique Bayesian equilibrium of this game is

$$\sigma_1(\cdot \mid 1.0) = [x_1], \quad \sigma_2(\cdot \mid 2.1) = [x_2], \quad \sigma_2(\cdot \mid 2.2) = [y_2].$$

This example illustrates the danger of analyzing each matrix separately, as if it were a game with complete information. If it were common knowledge that player 2's type was 2.1, then the players would be in the matrix on the left in table 3, in which the unique equilibirum is (x_1, x_2). If it were common knowledge that player 2's type was 2.2, then the players would be in the matrix on the right in table 3, in which the unique equilibrium is (y_1, y_2). Thus, if we looked only at the full-information Nash equilibria of these two matrices, then we might make the prediction "the outcome of the game will be (x_1, x_2) if 2's type is 2.1 and will be (y_1, y_2) if 2's type is 2.2."

This prediction would be absurd, however, for the actual Bayesian game in which player 1 does not initially know player 2's type. Notice first that this prediction ascribes two different actions to player 1, depending on 2's type (x_1 if 2.1, and y_1 if 2.2). So player 1 could not behave as predicted unless he got some information from player 2. That is, this prediction would be impossible to fulfill unless some kind of communication between the players is added to the structure of the game. Now notice that player 2 prefers (y_1, y_2) over (x_1, x_2) if her type is 2.1, and she prefers (x_1, x_2) over (y_1, y_2) if her type is 2.2. Thus, even if communication between the players were allowed, player 2 would not be willing to communicate the information that is necessary to fulfill this prediction, because it would always give her the outcome that she prefers less. She would prefer to manipulate her communications to get the outcomes (y_1, y_2) if 2.1 and (x_1, x_2) if 2.2.

Suppose that the two players can communicate, either directly or through some mediator, or via some tatonnement process, before they choose their actions in C_1 and C_2. In the induced communication game, could there ever be a Bayesian equilibrium giving the outcomes (x_1, x_2) if player 2 is type 2.1 and (y_1, y_2) if player 2 is type 2.2, as naive analysis of the two matrices in table 3 would suggest? The answer is No, by the revelation principle. If there were such a communication game, then there would be an incentive-compatible mediation plan achieving the same outcomes. But this would be the plan satisfying

$$\mu(x_1, x_2 | 1.0, 2.1) = 1, \quad \mu(y_1, y_2 | 1.0, 2.2) = 1.$$

which is not incentive compatible, because player 2 could gain by lying about her type. In fact, there is only one incentive-compatible mediation plan for this example, and it is $\bar{\mu}$, defined by

$$\bar{\mu}(x_1, x_2 | 1.0, 2.1) = 1, \quad \bar{\mu}(x_1, y_2 | 1.0, 2.2) = 1.$$

This is, this game has a unique communication equilibrium, which is equivalent to the unique Bayesian equilibrium of the game without communication.

Notice this analysis assumes that player 2 cannot choose her action and show it verifiably to player 1 before he chooses his action. She can say whatever she likes to player 1 about her intended action before they actually choose, but there is nothing to prevent her from choosing an action different from the one she promised if she has an incentive to do so.

In the insurance industry, the inability to get individuals to reveal unfavorable information about their chances of loss is known as adverse selection, and the inability to get fully insured individuals to exert efforts against their insured losses is known as moral hazard. This terminology can be naturally extended to more general game-theoretic models. The need to give players an incentive to report their information honestly may be called *adverse selection.* The need to give players an incentive to implement their recommended actions may be called *moral hazard.* In this sense, we may say that the incentive constraints (5) are a general mathematical characterization of the effect of adverse selection and moral hazard in Bayesian games.

3. Sender–receiver games

A sender–receiver game is a two-player Bayesian game with communication in which player 1 (the sender) has private information but no choice of actions, and player 2 (the receiver) has a choice of actions but no private information. Thus, sender–receiver games provide a particularly simple class of examples in which both moral hazard and adverse selection are involved. [See Crawford and Sobel (1982).]

A general *sender–receiver game* can be characterized by specifying (T_1, C_2, p, u_1, u_2), where T_1 is the set of player 1's possible types, C_2 is the set of player 2's possible actions, p is a probability distribution over T_1 that represents player 2's beliefs about player 1's type, and $u_1: C_2 \times T_1 \to \mathbb{R}$ and $u_2: C_2 \times T_1 \to \mathbb{R}$ are utility functions for player 1 and player 2 respectively. A sender–receiver game is finite iff T_1 and C_2 are both finite sets.

A *mediation plan* or *mechanism* for the sender–receiver game as above is any function $\mu: T_1 \to \Delta(C_2)$. If such a plan μ were implemented honestly and obediently by the players, the expected payoff to player 2 would be

$$U_2(\mu) = \sum_{t_1 \in T_1} \sum_{c_2 \in C_2} p(t_1)\,\mu(c_2|t_1)\,u_2(c_2, t_1)$$

and the conditionally expected payoff to player 1 if he knew that his type was t_1 would be

$$U_1(\mu|t_1) = \sum_{c_2 \in C_2} \mu(c_2|t_1) u_1(c_2, t_1).$$

The general incentive constraints (5) can be simplified in sender–receiver games. Because player 1 controls no actions, the incentive constraints on player 1 reduce to purely informational incentive constraints. On the other hand, because player 2 has no private information, the incentive constraints on player 2 reduce to purely strategic incentive constraints, as in (1) or (2). Thus, a mediation plan μ is *incentive compatible* for the sender–receiver game if any $\mu: T_1 \to \Delta(C_2)$ such that

$$\sum_{c_2 \in C_2} \mu(c_2|t_1) u_1(c_2, t_1) \geqslant \sum_{c_2 \in C_2} \mu(c_2|s_1) u_1(c_2, t_1), \forall t_1 \in T_1, \quad \forall s_1 \in T_1, \tag{6}$$

and

$$\sum_{t_1 \in T_1} p(t_1)[u_2(c_2, t_1) - u_2(e_2, t_1)]\,\mu(c_2|t_1) \geqslant 0, \forall c_2 \in C_2, \quad \forall e_2 \in C_2. \tag{7}$$

The informational incentive constraints (6) assert that player 1 should not expect to gain claiming that his type is s_1 when it is actually t_1, if he expects player 2 to obey the mediator's recommendations. The strategic incentive constraints (7) assert that player 2 should not expect to gain by choosing action e_2 when the mediator recommends c_2 to her, if she believes that player 1 was honest to the mediator.

For example, consider a sender–receiver game [due to Farrell (1993)] with $C_2 = \{x_2, y_2, z_2\}$ and $T_1 = \{1.a, 1.b\}$, $p(1.a) = 0.5 = p(1.b)$, and utility payoffs (u_1, u_2) that depend on player 1's type and player 2's action as in table 4.

Suppose first there is no mediation, but that player 1 can send player 2 any message drawn from some large alphabet or vocabulary, and that player 2 will be sure to observe player 1's message without any error or noise. Then, as Farrell (1993) has shown, in every equilibrium of the induced communication game, player

Table 4

	x_2	y_2	z_2
1.*a*	2,3	0,2	−1,0
1.*b*	1,0	2,2	0,3

2 will choose action y_2 for sure, after any message that player 1 might send with positive probability. To see why, notice that player 2 is indifferent between choosing x_2 and z_2 only if she assesses a probability of 1.*a* of exactly 0.5, but with this assessment she prefers y_2. Thus, there is no message that can generate beliefs that would make player 2 willing to randomize between x_2 and z_2. For each message that player 1 could send, depending on what player 2 would infer from receiving this message, player 2 might respond either by choosing x_2 for sure, by randomizing between x_2 and y_2, by choosing y_2 for sure, by randomizing between y_2 and z_2, or by choosing z_2 for sure. Notice that, when player 1's type is 1.*a*, he is not indifferent between any two different responses among these possibilities, because he strictly prefers x_2 over y_2 and y_2 over z_2. Thus, in an equilibrium of the induced communication game, if player 1 had at least two messages (call them "α" and "β") that are sent with positive probability and to which player 2 would respond differently, then type 1.*a* would be willing to send only one of these messages (say, "α"), and so the other message ("β") would be sent with positive probability only by type 1.*b*. But then, player 2's best response to this other message ("β") would be z_2, which is the worst outcome for type 1.*b* of player 1, so type 1.*b* would not send it with positive probability either. This contradiction implies that player 2 must use the same response to every message that player 1 sends with positive probability. Furthermore, this response must be y_2, because y_2 is player 2's unique best action given here beliefs before she receives any message. (This argument is specific to this example. However, Forges (1994) has shown more generally how to characterize the information that can be transmitted in sender–receiver games without noise.)

Thus, as long as the players are restricted to perfectly reliable noiseless communication channels, no substantive communication can occur between players 1 and 2 in any equilibrium of this game. However, substantive communication can occur when noisy communication channels are used. For example, suppose player 1 has a carrier pigeon that he could send to player 2, but, if sent, if would only arrive with probability $\frac{1}{2}$. Then there is an equilibrium of the induced communication game in which player 2 chooses x_2 if the pigeon arrives, player 2 chooses y_2 if the pigeon does not arrive, player 1 sends the pigeon if his type is 1.*a*, and player 1 does not send the pigeon if his type is 1.*b*. Because of the noise in the communication channel (the possibility of the pigeon getting lost), if player 2 got the message "no pigeon arrives," then she would assign a $\frac{1}{3}$ probability to the event that player 1's type was 1.*a* (and he sent a pigeon that got lost), and so she would be willing to choose y_2, which is better than x_2 for type 1.*b* of player 1. (See Forges (1985), for a seminal treatment of this result in related examples.)

Thus, using this noisy communication channel, there is an equilibrium in which player 2 and type 1.*a* of player 1 get better expected payoffs than they can get in equilibrium with direct noiseless communication. By analyzing the incentive constraints (6) and (7), we can find other mediation plans $\mu: T_1 \rightarrow \Delta(C_2)$ in which they both do even better. The informational incentive constraints (6) on player 1 are

$$2\,\mu(x_2|1.a)-\mu(z_2|1.a)\geqslant 2\,\mu(x_2|1.b)-\mu(z_2|1.b),$$

$$\mu(x_2|1.b)+2\,\mu(y_2|1.b)\geqslant \mu(x_2|1.a)+2\,\mu(y_2|1.a),$$

and the strategic incentive constraints (7) on player 2 are

$$0.5\,\mu(x_2|1.a)-\mu(x_2|1.b)\geqslant 0,$$

$$1.5\,\mu(x_2|1.a)-1.5\,\mu(x_2|1.b)\geqslant 0,$$

$$-0.5\,\mu(y_2|1.a)+\mu(y_2|1.b)\geqslant 0,$$

$$\mu(y_2|1.a)-0.5\,\mu(y_2|1.b)\geqslant 0,$$

$$-1.5\,\mu(z_2|1.a)+1.5\,\mu(z_2|1.b)\geqslant 0,$$

$$-\mu(z_2|-1.a)+0.5\,\mu(z_2|1.\mathrm{b})\geqslant 0.$$

(The last of these constraints, for example, asserts that player 2 should not expect to gain by choosing y_2 when z_2 is recommended.) To be a mediation plan, μ must also satisfy the probability constraints

$$\mu(x_2|1.a)+\mu(y_2|1.a)+\mu(z_2|1.a)=1,$$

$$\mu(x_2|1.b)+\mu(y_2|1.b)+\mu(z_2|1.b)=1,$$

and all $\mu(c_2|t_1)\geqslant 0$.

If, for example, we maximize the expected payoff to type $1.a$ of player 1

$$U_1(\mu|1.a)=2\,\mu(x_2|1.a)-\mu(z_2|1.a)$$

subject to these constraints, then we get the mediation plan

$$\mu(x_2|1.a)=0.8,\quad \mu(y_2|1.a)=0.2,\quad \mu(z_2|1.a)=0,$$

$$\mu(x_2|1.b)=0.4,\quad \mu(y_2|1.b)=0.4,\quad \mu(z_2|1.b)=0.2.$$

Honest reporting by player 1 and obedient action by player 2 is an equilibrium when a noisy communication channel or mediator generates recommended-action messages for player 2 as a random function of the type-reports sent by player 1 according to this plan μ. Furthermore, no equilibrium of any communication game induced by any communication channel could give a higher expected payoff to type $1.a$ of player 1 than the expected payoff of $U_1(\mu|1.a)=1.6$ that he gets from this plan.

On the other hand, the mechanism that maximizes player 2's expected payoff is

$$\mu(x_2|1.a)=\tfrac{2}{3},\quad \mu(y_2|1.a)=\tfrac{1}{3},\quad \mu(z_2|1.a)=0,$$

$$\mu(x_2|1.b)=0,\quad \mu(y_2|1.b)=\tfrac{2}{3},\quad \mu(z_2|1.b)=\tfrac{1}{3}.$$

This gives expected payoffs

$$U_1(\mu \mid 1.a) = 1.333, \quad U_1(\mu \mid 1.b) = 1.333, \quad U_2(\mu) = 2.5.$$

Once we have a complete characterization of the set of all incentive-compatible mediation plans, the next natural question is: which mediation plans or mechanisms should we actually expect to be selected and used by the players? That is, if one or more of the players has the power choose among all incentive-compatible mechanisms, which mechanisms should we expect to observe?

To avoid questions of interpersonal equity in bargaining, which belong to cooperative game theory, let us here consider only cases where the power to select the mediator or design the communication mechanism belongs to just one of the players. To begin with suppose that player 2 can select the mediation plan. To be more specific, suppose that player 2 will first select a mediator and direct him to implement some incentive-compatible mediation plan, and then player 1 can either accept this mediator and communicate with 2 thereafter only through him, or 1 can reject this mediator and thereafter communicate with 2 only face-to-face.

It is natural to expect that player 2 will use her power to select a mediator who will implement the incentive-compatible mediation plan that is best for 2. This plan is worse than y_2 for 1 if his type is 1.*b*, so one might think that, if 1's type is 1.*b*, then he should reject player 2's proposed mediator and insist on communicating face-to-face. However, there is an equilibrium of this mediator-selection game in which player 1 always accepts always player 2's proposal, no matter what his type is. In this equilibrium, if 1 rejected 2's mediator, then 2 might reasonably infer that 1's type was 1.*b*, in which case 2's rational choice would be z_2 instead of y_2, and z_2 is the worst possible outcome for both of 1's types.

Unfortunately, there is another sequential equilibrium of this mediator-selection game in which player 1 always rejects player 2's mediator, no matter what mediation plan she selects. In this equilibrium, player 2 infers nothing about 1 if he rejects the mediator and so does y_2, but if he accepted the mediator then she would infer (in this zero-probability event) that player 1 is type 1.*b* and so she would choose z_2.

Now consider the mediator-selection game in which the informed player 1 can select the mediator and choose the mediation plan that will be implemented, with the only restriction that player 1 must make the selection after he already knows his own type, and player 2 must know what mediation plan has been selected by player 1. For any incentive-compatible mediation plan μ, there is an equilibrium in which 1 chooses μ for sure, no matter what his type is, and they thereafter play honestly and obediently when μ is implemented. In this equilibrium, if any mediation plan other than μ were selected then 2 would infer from 1's surprising selection that his type was 1.*b* (she might think "only 1.*b* would deviate from μ"), and therefore she would choose z_2 no matter what the mediator might subsequently recommend. Thus, concepts like sequential equilibrium cannot determine the outcome of such a mediator-selection game beyond what we already knew from the revelation principle.

To get a more definite prediction about what mediation plans or mechanisms are likely to be selected by the players, we need to make some assumptions that go beyond traditional noncooperative game theory. Concepts of inscrutable intertype compromise and credibility of negotiation statements need to be formalized. Formal approaches to these questions have been offered by Farrell (1993), Grossman and Perry (1986), Maskin and Tirole (1990), and Myerson (1983, 1989).

4. Communication in multistage games

Consider the following two-stage two-player game. At stage 1, player 1 must choose either a_1 or b_1. If he chooses a_1 then the game ends and the payoffs to players 1 and 2 are (3,3). If he chooses b_1 then there is a second stage of the game in which each player i must choose either x_1 or y_i, and the payoffs to players 1 and 2 depend on their second-stage moves as follows:

Table 5

	x_2	y_2
x_1	7,1	0,0
y_1	0,0	1,7

The normal representation of this game in strategic form may be written as follows:

Table 6

	x_2	y_2
a_1x_1	3,3	3,3
a_1y_1	3,3	3,3
b_1x_1	7,1	0,0
b_1y_1	0,0	1,7

In this strategic-form game, the strategy b_1y_1 is strongly dominated for player 1. So it may seem that any theory of rational behavior should imply that there is zero probability of player 1 choosing b_1 at the first stage and y_1 at the second stage.

However, this conclusion does not hold if we consider the original two-stage game as a game with communication. Consider, for example, the following medation plan. At stage 1, the mediator recommends that player 1 should choose b_1. Then, at stage 2, with probability $\frac{1}{2}$ the mediator will recommend the moves x_1 and x_2, and with probability $\frac{1}{2}$ the mediator will recommend the moves y_1 and y_2. In either case, neither player will be able to gain by unilaterally disobeying the mediator at stage 2. At stage 1, disobeying the mediator would give player 1 a payoff of 3, which is less than the expected payoff of $\frac{1}{2} \times 7 + \frac{1}{2} \times 1 = 4$ that he gets from obedience. Thus, this mediation plan is incentive compatible, and it can lead player 1 to choose b_1 and then y_1 with probability $\frac{1}{2}$.

The key to this mediation plan is that player 1 must not learn whether x_1 or y_1 will be recommended to him at stage 2 until after it is too late to go back and choose a_1. That is, when we study a multistage game with communication, we should take account of the possibility of communication at every stage of the game. For multistage games, the revelation principle asserts that any equilibrium of any communication game that can be induced by adding a communication structure is equivalent to some mediation plan of the following form: at the beginning of each stage, the players confidentially report their new information to the mediator; then the mediator determines the recommended actions for the players at this stage, as a function of all reports received at this and all earlier stages, by applying some randomly selected feedback rule; then the mediator confidentially tells each player the action that is recommended for him at this stage; and (assuming that all players know the probability distribution that the mediator used to select his feedback rule) it is an equilibrium of the induced communication game for all players to always report their information honestly and choose their actions obediently as the mediator recommends. The probability distributions over feedback rules that satisfy this last incentive-compatibility condition can be characterized by a collection of linear incentive constraints, which assert that no player can expect to gain by switching to any manipulative strategy of lying and disobedience, when all other players are expected to be honest and obedient. For a formal statement of these ideas, see Myerson (1986a).

The inadequacy of the strategic-form game in table 6 for analysis of the incentive-compatible mediation in the above example shows that we may have to make a conceptual choice between the revelation principle and the generality of the strategic form, when we think about multistage games. If we want to allow communication opportunities to remain implicit at the modeling stage of our analysis, then we get a solution concept which is mathematically simpler than Nash equilibrium (because, by the revelation principle, communication equilibria can be characterized by linear incentive constraints) but which cannot necessarily be analyzed via the normal strategic-form representation. If we want in general to study any multistage game via the strategic form, then all communication opportunities must be made an explicit part of the extensive game model before we construct the normal representation in strategic form.

Sequential rationality and termbling-hand refinements of equilibrium for games with communication have been considered by Myerson (1986a, b). In Myerson (1986a), *acceptable correlated equilibria* are defined as a natural analogue of Selten's (1975) trembling-hand perfect equilibria for strategic-form games. The main result of Myerson (1986b) is that these acceptable correlated equilibria satisfy a kind of *strategy-elimination* property, which may be stated as follows: for any strategic-form game with communication, there exists a set of *codominated* strategies such that the set of acceptable correlated equilibria of the given game is exactly the set of correlated equilibria of the game that remains after eliminating all the codominated strategies. A similar strategy-elimination characterization of sequentially rational correlated equilibria for multistage games with communication is derived in Myerson (1986a).

References

Aumann, R.J. (1974) 'Subjectivity and Correlation in Randomized Strategies', *Journal of Mathematical Economics*, **1**: 67–96.

Aumann, R.J. (1976) 'Agreeing to Disagree', *Annals of Statistics*, **4**: 1236–1239.

Aumann, R.J. (1987) 'Correlated Equilibria as an Expression of Bayesian Rationality', *Econometrica*, **55** :1–18.

Barany, I. (1987) 'Fair Distribution Protocols or How the Players Replace Fortune', CORE Discussion Paper 8718, Université Catholique de Louvain. *Mathematics of Operations Research*, **17**: 327–340.

Crawford, V. and J. Sobel (1982) 'Strategic Information Transmission', *Econometrica*, **50**: 579–594.

Grossman, S. and M. Perry (1986) 'Perfect Sequential Equilibrium', *Journal of Economic Theory*, **39**: 97–119.

Farrell, J. (1993) 'Meaning and Credibility in Cheap-Talk Games', *Games and Economic Behavior*, **5**: 514–531.

Forges, F. (1985) 'Correlated Equilibria in a Class of Repeated Games with Incomplete Information', *International Journal of Game Theory*, **14**: 129–150.

Forges, F. (1986) 'An Approach to Communication Equilibrium', *Econometrica*, **54**: 1375–1385.

Forges, F. (1994) 'Non-zero Sum Repeated Games and Information Transmission', in: N. Megiddo, ed., *Essays in Game Theory*. Berlin: Springer, pp. 65–95.

Forges, F. (1990) 'Universal Mechanisms'. *Econometrica*, **58**: 1341–1364.

Harsanyi, J.C. (1967–8) 'Games with Incomplete Information Played by 'Bayesian' Players', *Management Science*, **14**: 159–182, 320–334, 486–502.

Hart, S. and D. Schmeidler (1989) 'Existence of Correlated Equilibria'. *Mathematics of Operations Research*, **14**: 18–25.

Maskin, E. and J. Tirole (1990) 'The Principal-Agent Relationship with an Informed Principal: the Case of Private Values', *Econometrica*, **58**: 379–409.

Moulin, H. and J.-P. Viral (1978) 'Strategically Zero-Sum Games: The Class Whose Completely Mixed Equilibria Cannot be Improved Upon', *International Journal of Game Theory*, **7**: 201–221.

Myerson, R.B. (1982) 'Optimal Coordination Mechanisms in Generalized Principal-Agent Problems'. *Journal of Mathematical Economics*, **10**: 67–81.

Myerson, R.B. (1983) 'Mechanism Design by an Informed Principal', *Econometrica*, **51**: 1767–1797.

Myerson, R.B. (1985) 'Bayesian Equilibrium and Incentive Compatibility', In: L. Hurwicz, D. Schmeidler and H. Sonnenschein, eds., *Social Goals and Social Organization*. Cambridge: Cambridge University Press, pp. 229–259.

Myerson, R.B. (1986a) 'Multistage Games with Communication', *Econometrica*, **54**: 323–358.

Myerson, R.B. (1986b) 'Acceptable and Predominant Correlated Equilibria', *International Journal of Game Theory*, **15**: 133–154.

Myerson, R.B. (1989) 'Credible Negotiation Statements and Coherent Plans', *Journal of Economic Theory*, **48**: 264–303.

Nau, R.F., and K.F. McCardle (1990) 'Coherent Behavior in Noncooperative Games', *Journal of Economic Theory*, **50**: 424–444.

Neumann, J. von and O. Morgenstern (1944) *Theory of Games and Economic Behavior*. Princeton: Princeton University Press. 2nd edn., 1947.

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