CIRJE-F-395

Repeated Games

Entry in *The New Palgrave Dictionary of Economics*, 2nd Edition

Michihiro Kandori
University of Tokyo

January 2006

Repeated Games

Entry in *The New Palgrave Dictionary of Economics,* 2nd Edition

KANDORI, Michihiro

Faculty of Economics, University of Tokyo

January 8, 2006

Abstract

This entry shows why self-interested agents manage to cooperate in a long-term relationship. When agents interact only once, they often have an incentive to deviate from cooperation. In a repeated interaction, however, any mutually beneficial outcome can be sustained in an equilibrium. This fact, known as the folk theorem, is explained under various information structures. This entry also compares repeated games with other means to achieve efficiency and briefly discuss the scope for potential applications.

JEL classification: C72, C73, D43, L13

Repeated games

Repeated games provide a formal and quite general framework to examine why self-interested agents manage to cooperate in a long term relationship.

Formally, repeated games refer to a class of models where the same set of agents repeatedly play the same game, called the 'stage game', over a long (typically, infinite) time horizon. In contrast to the situation where agents interact only once, *any* mutually beneficial outcome can be sustained as an equilibrium when agents interact repeatedly and frequently. A formal statement of this fact is known as the folk theorem.

Repeated games and the general theories of efficiency

Thanks to the developments in the last three decades, economics now recognizes three general ways to achieve efficiency:

(1) Competition

(2) Contracts

(3) Long term relationships.

For standardized goods and services, with a large number of potential buyers and sellers, promoting market competition is an effective way to achieve efficiency. This is formulated as the classic First and Second Welfare Theorems in general equilibrium theory. There are, however, other important resource allocation problems which do not involve standardized goods and services. Resource allocation within a firm or an organization is a prime example, as pointed out by Ronald Coase, and examples abound in social and political interactions. In such a case, aligning individual incentives with social goals is essential for efficiency, and this can be achieved by means of *incentive schemes* (penalties or rewards). The incentive schemes, in turn, can be provided in two distinct ways: by a formal contract or by a long term relationship. The penalties and rewards specified by a formal contract are enforced by the court, while in a long term relationship, the value of future interaction serves as the rewards and penalties to discipline the agents' current behavior. The theory of contracts and mechanism design concerns the former case, and the theory of repeated games deals with the latter. These theories provide general methods to achieve efficiency, and they have become important building blocks of modern economic theory.

An example: collusion of gas stations and the trigger strategy

Consider two gas stations located right next to each other. They have identical and constant marginal cost c (the wholesale price of gasoline) and compete by publicly posting their prices. Suppose their joint profit is maximized when they both charge p=10, where each receives a large profit π. Although this is the best outcome for them, they have an incentive to deviate. By slightly undercutting the price, each of them can steal all the customers from the opponent, and its profit (almost) doubles. The only price free from such profitable deviation is p=c, where their profit is equal to zero. In other words, the only Nash equilibrium in the price competition game is an *inefficient* (for the gas stations) outcome where both charge p=c. This situation is the rule rather than the exception: The Nash equilibrium in the stage game, the only outcome that agents can credibly achieve in a one-shot interaction, is quite often inefficient for them. This is because agents only seek their private benefits, ignoring the benefits or costs of their actions for the opponents.

In reality, however, gas stations enjoy positive profits, even when there is another station nearby. An important reason may well be that their interaction is not one-shot. Formally, the situation is captured by a *repeated game*, where the two gas stations play the price competition game (the stage game) over an infinite time horizon t=0,1,2,…. Consider the following repeated game strategy:

1) Start with the optimal price p=10.
2) Stick to p=10 as long as no player (including oneself) has ever deviated from p=10.
3) Once anyone (including oneself) deviated, charge p=c forever.

This can be interpreted as an explicit or implicit agreement of the gas stations: charge the monopoly price p=10, and any deviation triggers cut-throat price competition (p=c with zero profit). Let us now check if each player has any incentive to deviate from this strategy. Note that, if no one deviates, each station enjoys profit π every day. As we saw above, a player can (almost) double its stage payoff by slightly undercutting the agreed price p=10. Hence the short term gain from deviation is at most π. If one deviates, however, her future payoff is reduced from π to zero in each and every period in the future. Now assume that the players discount future profits by the *discount factor* $\delta \in (0,1)$. The number δ measures the value of a dollar in the next period. The discounted future loss is $\delta\pi + \delta^2\pi + \Lambda = \frac{\delta}{1-\delta}\pi.$ If this is larger than the short-term gain from defection (π), no one wants to deviate from the collusive price p=10. The condition is $\pi \le \delta/(1-\delta)\pi$, or equivalently, $1/2 \le \delta$.

Next let us check if the players have an incentive to carry out the threat (the cut-throat price competition p=c). Since p=c is the Nash equilibrium of the stage game, charging p=c in each period is a best reply if the opponent always does so. Hence, the players are choosing mutual best replies. In this sense, the threat of p=c is credible or self-enforcing.

In summary, under the strategy defined above, players are choosing mutual best replies *after any history*, as long as $1/2 \leq \delta$. In other words, the strategy constitutes a *subgame perfect equilibrium* in the repeated game. Similarly, in a general game, any outcome which Pareto dominates the Nash equilibrium can be sustained by a strategy which reverts to the Nash equilibrium after a deviation. Such a strategy is called a *trigger strategy*.

Three remarks: multiple equilibria, credibility of threat and renegotiation, and finite versus infinite horizon

A couple of remarks are in order about the example. First, the trigger strategy profile is not the only equilibrium of the repeated game. The repetition of the stage game Nash equilibrium (p=c forever) is also a subgame perfect equilibrium. Are there any other equilibria? Can we characterize *all* equilibria in a repeated game? The latter question appears to be formidable at first sight, because there are an infinite number of repeated game strategies, and they can potentially be quite complex. We do have, however, some complete characterizations of all equilibria of a repeated game, such as folk theorems and self-generation conditions as will be discussed subsequently.

Second, one may question the credibility of the threat (p=c forever). In the above example, credibility was formalized as the subgame perfect equilibrium condition. According to this criterion, the threat p=c is credible because a *unilateral* deviation by a *single* player is never profitable. The threat p=c, however, may be upset by *renegotiation*. When players are called upon to carry out this grim threat after a deviation, they may well get together and agree to "let bygones be bygones." After all, when there is a better equilibrium in the repeated game (for example, the trigger strategy equilibrium), why do we expect the players to stick to the inefficient one (p=c)? This is the problem of *renegotiation proofness* in repeated games. The problem is trickier than it appears, however, and economists have not yet agreed on what is the right notion of renegotiation proofness for repeated games. The reader may get a sense of difficulty from the following observation. Suppose the players have successfully renegotiated away p=c to play the trigger strategy equilibrium again. This is self-defeating, however, because the players now have an incentive to deviate, as they may well anticipate that the threat p=c will be again subject to renegotiation and will not be carried out. For a comprehensive discussion of this topic (and also of a number of major technical results on repeated games), see an excellent survey by D.

Pearce (1990).

Third, let me comment on the assumption of an *infinite* time horizon. Suppose that the gas stations are to be closed by the end of next year (due to a new zoning plan, for example). This situation can be formulated as a *finitely* repeated game. On the last day of their business, the gas stations just play the stage game, and therefore they have no other choice but to play the stage game equilibrium p=c. In the penultimate day, they rationally anticipate that they will play p=c *irrespective of* their current action. Hence they are effectively playing the stage game in the penultimate day, and again they choose p=c. By induction, the *only* equilibrium of the finitely repeated price competition is to charge p=c in *every* period. The impossibility of cooperation holds no matter how long the time horizon is, and it is in sharp contrast to the infinite horizon case.

Although one may argue that players do not really live infinitely long (so that the finite horizon case is more realistic), there are some good reasons to consider the infinite horizon models. First, even though the time horizon is finite, if players do not know in advance exactly *when* the game ends, the situation can be formulated as an infinitely repeated game. Suppose that, with probability $r > 0$, the game ends at the end of any given period. This implies that, *with probability one*, the game ends in a finite horizon. Note, however, that the expected discounted profit is equal to $\pi(0) + (1 - r)\delta\pi(1) + (1 - r)^2\delta^2\pi(2) + \ldots$, where $\pi(t)$ is the stage payoff in period t. This is identical to the payoff in an infinitely repeated game with discount factor $\delta' = (1 - r)\delta$. Second, the drastic "discontinuity" between the finite and infinite horizon cases in the price competition example hinges on the uniqueness of equilibrium in the stage game. Benoit and Krishna (1985) show that, if each player has multiple equilibrium payoffs in the stage game, the long but finite horizon case enjoys the same scope for cooperation as the infinite horizon case (the folk theorem, discussed below, approximately holds for T-period repeated game, when $T \to \infty$).

The repeated game model

Now let me present a general formulation of a repeated game. Consider an infinitely repeated game, where players i=1,2,…,N repeatedly play the same stage game over an infinite time horizon t=0, 1, 2, …. In each period, player i takes some action $a_i \in A_i$, and her payoff in that period is given by a stage game payoff function $g_i(a)$, where $a = (a_1, \ldots, a_N)$ is the action profile in that period. The repeated game payoff is given by

$$\Pi_i = \sum_{t=0}^{\infty} g_i(a(t))\delta^t ,$$

where a(t) denotes the action profile in period t and $\delta \in (0,1)$ is the discount factor. It is often quite useful to look at the *average payoff* of the repeated game, which is defined to be $(1-\delta)\Pi_i$. Note that, if one receives the same payoff x in each period, the repeated game payoff is $\Pi_i = x + \delta x + \delta^2 x + \ldots = x/(1-\delta)$. This example helps to understand the definition of average payoff: in this case $(1-\delta)\Pi_i$ is indeed equal to x, the payoff per period.

A *history* up to time t is the sequence of realized action profiles before t: $h^t = (a(0), a(1), \ldots, a(t-1))$. A *repeated game strategy* for player i, denoted by s_i, is a complete contingent action plan, which specifies a current action after any history: $a_i(t) = s_i(h^t)$ (a minor note: to determine $a_i(0)$, we introduce a dummy history h^0 such that $a_i(0) = s_i(h^0)$). A repeated game strategy profile $s = (s_1, \ldots, s_N)$ is a *subgame perfect equilibrium* if it specifies mutual best replies after any history.

The folk theorem

Despite the fact that a repeated game has an infinite number of strategies, which can be arbitrarily complicated, we do have a *complete* characterization of equilibrium payoffs. The folk theorem shows exactly which payoff points can be achieved in a repeated game.

Before stating the theorem, we need to introduce a couple of concepts. First, let us determine the set of physically achievable average payoffs in a repeated game. Note that, by alternating between two pure strategy outcomes, say u and v, one may achieve any point between u and v as the average payoff profile. Hence, an average payoff profile can be a weighted average (in other words, a convex combination) of pure strategy payoff profiles in the stage game. Let us denote the set of all such points by V. Formally, the set of *feasible average payoff profiles* V is the smallest convex set that contains the pure strategy payoff profiles of the stage game.

Second, let us determine the points in V that cannot possibly be an equilibrium outcome. For example, if a player has an option to stay out to enjoy zero profit in each period, it is *a priori* clear that her equilibrium average payoff cannot be less than zero. In general, there is a payoff level that a player can guarantee herself in any equilibrium, and this is formulated as the *minimax* payoff. Formally, the minimax payoff for player i is defined as

$$\underline{v}_i = \min_{\alpha_{-i}} \max_{\alpha_i} g_i(\alpha),$$

where $\alpha = (\alpha_1, \ldots, \alpha_N)$ is a mixed action profile (α_i is a probability distribution over player i's pure actions) and $g_i(\alpha)$ is the associated expected payoff. To understand why min and max are taken in that particular order, consider the situation where player i always *correctly anticipate what others do.* If player i knows that others choose α_{-i} = (α_1,…α_{i-1}, α_{i+1},…,α_N), he can play a best reply against α_{-i} to obtain $\max_{\alpha_i} g_i(\alpha)$. Note well that $\max_{\alpha_i} g_i(\alpha)$ is a function of α_{-i}. In the worst case, where others take the most damaging actions α_{-i}, player i obtains the minimax payoff (this is exactly what the definition says). From this definition it is clear that, in any equilibrium of the repeated game, *the average payoff to each player is at least her minimax payoff.* In any equilibrium, each player correctly anticipates what others do, and simply by playing the stage game best reply in each period, any player can make sure that her average payoff is more than her minimax payoff. (A comment: we consider mixed strategies in the definition of the minimax payoff because in many games the minimax payoff is smaller when we consider mixed strategies.)

From what we saw, now it is clear that the set of equilibrium average payoff profiles of a repeated game is *at most*

$$V^* = \{v \in V \mid \forall i \ v_i > \underline{v}_i\}.$$

(The points with $v_i = \underline{v}_i$ are excluded to avoid minor technical complications.) The set V* is called the *feasible and individually rational payoff set.* This is the set of physically achievable average payoff profiles in the repeated game where each player receives more than her minimax payoff. The folk theorem shows that any point in this "maximum possible region" can indeed be an equilibrium outcome of the repeated game. (Throughout this entry, I maintain a minor technical assumption that each player has a finite number of actions in the stage game.)

Folk theorem: In a N-player infinitely repeated game, any feasible and individually rational payoff profile v ∈ V* can be achieved as the average payoff profile of a subgame perfect equilibrium when the discount factor δ is close enough to 1, provided that either

(i) N = 2, or

(ii) N ≥ 3 and no two players have identical interests.

Formally, no two players have identical interests if there are *no* players i and j (i≠j) whose payoffs satisfy $g_i(a) = bg_j(a) + c$, $b > 0$ (i.e., no two players have the same preferences over the stage game outcomes). This is a "generic" condition that is almost always satisfied: the case where players have identical interests is very special in the sense that the equality $g_i(a) = bg_j(a) + c$ fails by even a slight change of the payoff functions. Hence, the folk theorem provides a general theory of efficiency: it shows that, for virtually any game, any mutually beneficial outcome can be achieved in a long term relationship, if the discount factor is close to 1. *Although game theoretic predictions quite often depend on the fine details of the model, this result is a notable exception for its generality.*

The crucial condition in the folk theorem is a high discount factor. The discount factor δ may measure the (subjective) patience of a player, or, it may be equal to $1/(1+r)$, where r is the interest rate per period. Although the discount factor may not be directly observable (in particular, in the former case), it should be high when one period is short. Hence, an empirically testable implication is that players who have daily interaction (such as the gas stations in our example) have a better scope for cooperation than those who interact only once a year. An important message of the folk theorem is that a high *frequency of interaction* is essential for the success of a long term relationship.

The name 'folk theorem' comes from the fact that game theorists had anticipated that something like it should be true long before it was precisely formulated and proved. In this sense, the assertion had been folklore in the game theorists' community. The proof is, however, by no means obvious, and there is a body of literature to prove the theorem in various degrees of generality. Early contributions include Aumann (1959), Friedman (1971) and Rubinstein (1979). The statement above is based on Fudenberg and Maskin (1986) and its generalization by Abreu, Dutta and Smith (1994). The proof is constructive: a clever strategy, which has a rather simple structure, is constructed to support any point in V^*.

Repeated games versus formal contracts

To discuss the scope of applications, I now compare a long-term relationship (repeated game) and a formal contract as a means to enforce efficient outcomes. As our gas station example shows, quite often an agent has an incentive to deviate from an efficient outcome, because it increases her private returns at the expense of the social benefit. Such a deviation can be deterred if we impose a sufficiently high penalty so that the *incentive constraint*

$$\text{gain from deviation} \leq \text{penalty}$$

is satisfied. This is the basic and common feature of repeated games and contracts. A formal contract explicitly specifies the penalty and it is enforced by the court. In repeated games, the penalty is indirectly imposed through future interaction. In this sense the theory of repeated games can be regarded as the theory of *informal or relational contracts*.

When is a long-term relationship a better way to achieve cooperation than a formal contract? First, a long-term relationship is useful when a formal contract is too costly or impractical. For example, it is often quite costly for a third party (the court) to verify if there was any deviation from an agreement, while defections may be directly observed by the players themselves. In practice, what constitutes 'cooperation' is often so fuzzy or complicated that it is hard to write it down explicitly, although the players have a common and good understanding about what it is. 'Pulling enough weight" in a joint research project may be a good example. In those situations, a long-term relationship is a more practical way to achieve cooperation than a formal contract. In fact, a classic study by Macaulay (1963) indicates that a vast majority of business transactions are executed without writing formal contracts. Second, there are some cases where a court powerful enough to enforce formal contracts simply does not exist. For example, in many problems in development economics and economic history, the legal system is highly imperfect. Even for developed countries in the modern age, there are no legal institutions which have enough binding power to enforce international agreements. Hence, repeated games provide a useful framework to address such problems as the organization of medieval trade, informal mutual insurance in developing countries, international policy coordination, and measures against global warming. Lastly, there is no legal system to enforce cartels or collusion, because the existing legal system refuses to enforce any contract that violates anti-trust laws. Hence a long-tem relationship is the only way to enforce a cartel or collusive agreement.

Is the folk theorem a negative result?

The theory of repeated games based on the folk theorem is often criticized because it does not, as the criticism goes, have any predictive power. The folk theorem basically says that anything can be an equilibrium in a repeated game. One could argue, however, that this criticism is misplaced if we regard the theory of repeated games as a theory of informal contracts. Just as anything can be enforced when the party agrees to sign a binding contract, in repeated games any (feasible and individually rational) outcome is sustained if the players agree on an equilibrium. Enforceability of a wide range of outcomes is the essential property of effective contracts, formal or informal. The folk theorem correctly captures this essential feature.

This criticism is valid, however, in the sense that the theory of repeated games does not provide a widely accepted criterion for equilibrium selection. When we regard a repeated game as an informal contract, where the players explicitly try to agree on which equilibrium to play, the problem of equilibrium selection boils down to the problem of bargaining. In such a context, it is natural to assume that an efficient point (in the set of equilibria) is played. In the vast majority of applied works of repeated games with symmetric stage games (such as the gas stations example), it is common to look at the best symmetric equilibrium. In contrast, when players try to find an equilibrium through trial and error, the theory of repeated games is rather silent about which equilibrium is likely to be selected. A large body of computer simulation literature on the evolution of cooperation, pioneered by Axelrod (1984), may be regarded as an attempt to address this issue.

Imperfect Monitoring

So far we assumed that players can perfectly observe each other's actions. In reality, however, long term relationships are often plagued by *imperfect monitoring*. For example, a country may not verify exactly how much CO_2 is emitted by neighboring countries. Workers in a joint project may not directly observe each other's effort. Electronic appliance shops often offer secret discounts for their customers, and each shop may not know exactly how much is charged by its rivals. In such situations, however, there are usually some pieces of information, or *signals*, which imperfectly reveal what actions have been taken. Published meteorological data indicates the amount of CO_2 emission, the success of the project is more likely with higher effort, and a shop's sales level is related (although not perfectly) to its rivals' prices.

According to the nature of the signals, repeated games with imperfect monitoring are classified into two categories: the case of *public monitoring*, where players commonly observe a public signal, and the case of *private monitoring*, where each player observe a signal that is not observable to others. Hence, the CO_2 emission game and the joint project game are examples with imperfect public monitoring (published meteorological data and the success of the project are publicly observed), while the secret price cutting game by electronic shops is a good example with imperfect private monitoring (one's sales level is private information).

This difference may appear to be a minor one, but, somewhat surprisingly, it is not. The imperfect *public* monitoring case shares many features with the *perfect* monitoring case, and we now have a good understanding of how it works. In contrast, the imperfect private monitoring case is not fully understood, and we only have some partial characterizations of equilibria. In what follows, I will sketch the main results in the imperfect public and private monitoring cases.

Imperfect Public Monitoring

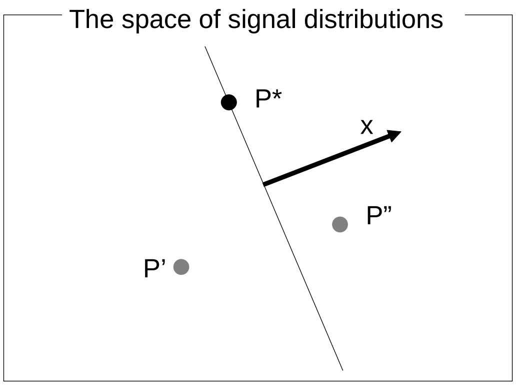
At first sight, this case might look much more complicated than the perfect monitoring case, but those two cases are similar in the sense that they share a *recursive structure*. Consider the set W* of all average payoff profiles associated with the subgame perfect equilibria of a perfect monitoring repeated game. Any point $w \in W^*$ is a weighted average of the current payoff g and the continuation payoff w': $(1 - \delta)g + \delta w'$. The continuation payoff typically changes when a player deviates from g, in such a way that the short-term gain from deviation is wiped out. Subgame perfection requires that all continuation payoffs are chosen from the equilibrium set W*. In this sense, W* is generated by itself, and this stationary or recursive structure turns out to be quite useful in characterizing the set of equilibria.

The set of equilibria in an imperfect public monitoring game also shares the same structure. Consider the equilibria where the public signal determines which continuation equilibrium to play. When a player deviates from the current equilibrium action, it affects both her current payoff and (through the public signal) her continuation payoff. The equilibrium action should be enforceable in the sense that any gain in the former should be wiped out in the latter, and this is easier when the continuation payoff admits large variations. Formally, given the range of continuation payoffs W, we can determine the set B(W) of enforceable average payoffs. The larger the set W is, the more actions can be enforced in the current period (and therefore the larger the set B(W) is). As in the perfect monitoring case, the equilibrium payoff set W=W* generates itself: it satisfies the *self-generation condition* of Abreu, Pearce and Stacchetti (1990) $W \subseteq B(W)$. W* is the largest (bounded) set satisfying this condition, and the condition is in fact satisfied with equality. Conversely, it is easy to show that any (bounded) set satisfying the self-generation condition is contained in the equilibrium payoff set W*.

This provides a simple and powerful characterization of equilibria, which is an essential tool to prove the folk theorem in the imperfect public monitoring case. The folk theorem shows that, despite the imperfection of monitoring, we can achieve any feasible and individually rational payoff profile under a certain set of conditions.

Before presenting a formal statement, let me sketch the basic ideas behind the folk theorem. When monitoring is imperfect, players have to be punished when a "bad" signal outcome ω is observed, and this may happen with a positive probability even if no one defects. For example, in the joint project game, the project may fail even though everyone works hard. A crucial difference between the perfect and imperfect monitoring cases is that, in the latter, punishment occurs *on the equilibrium path*. The resulting welfare loss, however, can be negligible under certain conditions.

Consider a two-player game, where the probability distribution of the signal $\omega \in \Omega = \{\omega^1, \ldots, \omega^K\}$, when no one defects, is given by $P^* = (p^*(\omega^1), \ldots, p^*(\omega^K))$ in the Figure. Suppose that each player's defection changes the probability distribution to exactly the same point P'. Then, there is absolutely no way to tell which player deviates, so that the only way to deter a defection is to *punish all players simultaneously*, when a "bad" outcome arises. This means that surplus is thrown away, and we are bound to have substantial welfare loss. Now consider a case where different players' actions affect the signal asymmetrically: player 1' defection leads to point P', while the defection by player 2 leads to P". In this asymmetric case, one can *transfer* future payoff from player 1 to 2 when player 1's defection is suspected. Under such a transfer, surplus is never thrown away, and this enables us to achieve efficiency.



Figure

More precisely, consider the normal vector x of the hyper plane separating P' and P" in the figure, and let $w_1 = x$ and $w_2 = -x$ be the continuation payoffs of player 1 and 2 respectively. The Figure indicates that player 1's expected continuation payoff $P \cdot w_1 = P \cdot x$ is reduced by her own defection ($P' \cdot x < P^* \cdot x$). Similarly, player 2's defection reduces her expected continuation payoffs ($P^* \cdot (-x) > P'' \cdot (-x)$). Note that this asymmetric punishment scheme does not reduce the joint payoff, because by construction $w_1 + w_2$ is identically equal to 0. This is an essential idea behind the folk theorem under imperfect public monitoring: *When different players' deviations are statistically discriminated,* *asymmetric punishment deters defections without welfare loss.*

When can we say that different players' deviations are statistically discriminated? Note well that the above construction is impossible when P" is exactly in between P* and P' (i.e., when P" is a convex combination of P* and P'). Such a case can be avoided if P*, P' and P" are linearly independent. The linear independence of the equilibrium signal distribution (P*) and the distributions associated with the players' unilateral deviations (P' and P"), is a precise formulation of what it means that the signal "statistically discriminates different players' deviations".

Let us now generalize this observation. Given an action profile (for simplicity of exposition, assume it is pure) to be sustained, there is an associated signal distribution P*. Consider any pair of players i and j, and let $|A_k|$ be the number of player k's actions (k=i, j) in the stage game. Since each player k=i, j has $|A_k| - 1$ ways to deviate, we have $|A_i|+|A_j| - 2$ signal distributions associated with their unilateral deviations. If those distributions and the equilibrium distribution P*, altogether $|A_i|+|A_j| - 1$ vectors, are linearly independent, we say that the signal can discriminate between deviations by i and deviations by j. This is called the *pairwise full rank condition*. This holds only when the dimension of the signal space ($|\Omega|$, the number of signal outcomes) is larger than the number of those vectors (i.e., $|\Omega| \geq |A_i|+|A_j| - 1$). Conversely, if this inequality is satisfied, the pairwise full rank condition holds "generically" (i.e., it holds unless the signal distributions have a very special structure, such as exact symmetry). This leads us to the folk theorem under imperfect public monitoring (this is a restatement of Fudenberg, Levine and Maskin (1994) in terms of genericity):

Folk theorem under imperfect public monitoring: Suppose that the signal space is large enough in the sense that $|\Omega| \geq |A_i|+|A_j| - 1$ holds for each pair of players i and j. Then, for a generic choice of the signal distributions and the stage game, any feasible and individually rational payoff profile $v \in V^*$ can be asymptotically achieved by a sequential equilibrium as the discount factor δ tends to 1.

In contrast to the perfect monitoring case, the proof is non-constructive. Rather than explicitly constructing equilibrium strategies, the theorem is proved by showing that any smooth subset of V* is self-generating. In fact, the exact structure of the equilibrium strategy profile to sustain, for example, an efficient point is not so well understood. Sannikov (2005) shows that detailed structure of equilibrium strategies can be obtained if the model is formulated in continuous time.

Imperfect Private Monitoring

Now consider the case where each player receives a private signal about the opponents' actions. Although this has a number of important applications (a leading example is the secret price cutting model), this part of research is still in its infancy. Hence, rather than just summarizing definitive results as in the previous subsections, I explain in somewhat more technical detail the source of difficulties and the nature of existing approaches.

The difficulties come from a subtle but crucial difference from the perfect or public monitoring case. I will explain below the difference from a couple of viewpoints, in the increasing order of technicality.

(i) In the perfect or public monitoring case, players share a mutual understanding about when and whom to punish. They can coordinate to implement a specific punishment, and, more importantly, they can mutually provide the incentives to carry out the punishment. This convenient feature is lost when players have diverse private information about each other's action.

(ii) In the perfect or public monitoring case, public information directly tells the opponents' future action plans. In the private monitoring case, however, each player has to draw statistical inferences about the history of the opponents' private signals to estimate what they are going to do. The inferences quickly become complicated over time, even if players adopt relatively simple strategies.

(iii) In the perfect or public monitoring case, the set of equilibria has a recursive structure, in the sense that a Nash equilibrium of the repeated game is always played after any history. Now consider a Nash equilibrium of, for example, the repeated prisoners' dilemma with imperfect private monitoring. After the equilibrium actions in the first period, say (C,C), players condition their action plans on their private signals ω_1 and ω_2. Hence the continuation play is a *correlated equilibrium*, where it is common knowledge that the probability distribution of the correlation device (ω_1,ω_2) is given by $p(\omega_1,\omega_2 | C,C)$. When player 1 deviates to D in the first period, however, the distribution of correlation device is *not* common knowledge: player 1 knows that it is $p(\omega_1,\omega_2 | D,C)$, while player 2 keeps the equilibrium expectation $p(\omega_1,\omega_2 | C,C)$. Hence, after a deviation, the continuation play is no longer a correlated equilibrium in the usual sense. In addition, the space of the correlation device (the history of private signals) becomes increasingly rich over time. Therefore, the equilibria in the private monitoring case do not have a compact recursive structure; a continuation play is chosen from a different set, depending on the history.

One way to get around these problems is to allow communication (Compte (1998) and Kandori and Matsushima (1998)). In their equilibrium, players truthfully communicate their private signal outcomes in each period. The equilibrium is constructed in such a way that each player's report of her signal is utilized to discipline *other* players and does *not affect one's own continuation payoff.* This implies that each player is indifferent about what to report, and therefore truth telling is *a* best reply. Such an equilibrium, which depends on the history of publicly observable messages, works in much the same way as the equilibria in the public monitoring case. Hence, with communication, the folk theorem is obtained in the private monitoring case.

The remaining issue is to characterize the equilibria in the private monitoring case without communication. From the viewpoint of potential applications, this is important, because collusion or cartel enforcement is a major applied area of repeated games, where communication is explicitly prohibited by the anti-trust law.

One may expect that, when players' private information admits sufficient positive correlation, an equilibrium can be constructed in a similar way as in the public monitoring case. Sekiguchi (1997) is the first to construct a non-trivial (and nearly efficient) equilibrium in the private monitoring game without communication, and his construction is basically built on such an idea. Strong correlation of private information is, however, not assumed in his model but is derived endogenously. He assumed that private signals provide nearly perfect observability and considered *mixed* strategies. In such a situation, the privately observed random variables, the action-signal pairs, are strongly correlated (because a player's random action is strongly correlated with another player's signal under nearly perfect observability). Mailath and Morris (2002) showed that, in general, there is "continuity" between the public and private but sufficiently correlated monitoring cases, in the sense that any strategy with a *finite memory* works in either case.

Those papers are examples of the *belief-based approach*, which directly addresses the statistical inference problem (see point (ii) above). There are some other papers to follow this approach, and they provide judiciously constructed strategies in rather specific examples, where the inference problem becomes tractable. Aside from the case with near perfect correlation, however, we are yet to have generally applicable results or techniques along this line of approach.

More successful has been the *belief-free approach*, where an equilibrium is constructed in such a way that the inference problem becomes *irrelevant*. As a leading example, I illustrate Ely and Valimaki's work (2002) on the repeated prisoners' dilemma with imperfect private monitoring. Each player's strategy is a Markov chain with two states, R (reward) or P (punishment). A specific action is played in each state (C in R, and D in P), and the transition probabilities between the states depend on the realization of the player's private signal. Choose those transition probabilities in such a way that the *opponent* is always indifferent between C and D *no matter which state the player is in*. This requirement can be expressed as a simple system of dynamic programming equations, which has a solution when the discount factor is close to 1 and the private signal is not too uninformative. By construction, any action choice is optimal against this strategy after any history, and in particular this strategy is a best reply to itself (so that it constitutes an equilibrium). Note that one's incentives do not depend on the opponent's state, and therefore one does not have to draw the statistical inferences about the history of the opponent's private signals.

There are certain difficulties, however, to obtain the folk theorem with such a class of equilibria. First, players may be punished simultaneously in this construction, and our discussion about the public monitoring case shows that some welfare loss is inevitable (unless monitoring is nearly perfect). Second, even if we restrict our attention to the nearly perfect monitoring case, there is a certain set of restrictions imposed on the action profiles that can be sustained by such a belief-free equilibrium.

Those difficulties can be resolved when we consider *block strategies*. Block strategies treat the stage games in T consecutive periods as if they were a single stage game, or a block stage game, and applies the belief-free approach with respect to those block stage games. It is now known that, by using the block strategies, the folk theorem under private monitoring holds in the nearly perfect monitoring case (Horner and Olszewski (2004)) and for some two-player games where monitoring is far from perfect (Matsushima (2004)). In the former, the block structure of the stage game helps to satisfy the restrictions imposed on the actions in belief-free equilibria. In the latter, an equilibrium is constructed where players choose constant actions in each block. This means that players have T samples of private signals for the constant actions, so that the observability practically becomes nearly perfect when T is large. With this increased observability and some restrictions on payoff functions, the folk theorem is obtained. For this construction to be feasible, the signals have to satisfy certain strong conditions, such as independence (across players).

The general folk theorem, or a general characterization of equilibria, for the private monitoring case is yet to be obtained, and it remains to be an important open question in economic theory. A comprehensive technical exposition of the perfect monitoring, imperfect public monitoring, and private monitoring cases can be found in

Mailath and Samuelson (2006).

Kandori, Michihiro

Bibliography

Abreu, D., Dutta, P., and Smith, L. 1994. The Folk Theorem for Repeated Games: A NEU condition, Econometrica 62, 939-948.

Abreu, D., Pearce, D. and Stacchetti, E. 1990. Towards a Theory of Discounted Repeated Games with Imperfect Monitoring, Econometrica 58, 1041-1064.

Aumann, R. 1959. Acceptable Points in General Cooperative N-Person Games, in Contributions to the Theory of Games IV, ed. by R. D. Luce and A. W. Tucker, Princeton: Princeton University Press.

Axcelrod, R. 1984. Evolution of Cooperation, New York: Basic Books

Benoit, J.P. and Krishna, V. 1985. Finitely Repeated Games, Econometrica 53, 905-922.

Compte, O., 1998. Communication in Repeated Games with Imperfect Private Monitoring, Econometrica 66, 597-626.

Ely, J. and Valimaki, J. 2002. A Robust Folk Theorem for the Prisoner's Dilemma, Journal of Economic Theory 102, 84-105.

Friedman, J. 1971. A Non-Cooperative Equilibrium for Supergames, Review of Economic Studies 38, 1-12.

Fudenberg, D., Levine, D. and Maskin, E. 1994. The Folk Theorem with Imperfect Public Information, Econometrica 62, 997-1040.

Fudenberg D. and Maskin, E. 1986. The Folk Theorem in Repeated Games with Discounting or with Incomplete Information, Econometrica 54, 533-554.

Horner, J. and Olszewski, W. 2004. The Folk Theorem for Games with Private

Almost-Perfect Monitoring, mimeo., Northwestern University.

Kandori, M. and Matsushima, H., 1998. Private Observation, Communication and Collusion, Econometrica 66, 627-652.

Macauley, S. 1963. Non-contractual Relations in Business: A Preliminary Study, American Sociological Review 28, 55-67.

Mailath, G. and Morris, S. 2002. Repeated Games with Imperfect Private Monitoring: Notes on a Coordination Perspective, Journal of Economic Theory 102, 189-228.

Mailath, G. and L. Samuelson, 2006. Repeated Games and Reputations: Long-Run Relationships, Oxford: Oxford University Press.

Matsushima, H. 2004. Repeated Games with Private Monitoring: Two Players, Econometrica 72, 823-852.

Pearce, D. 1990. Repeated Games: Cooperation and Rationality, in Advances in Economic Theory, ed. by J. Laffont, Cambridge: Cambridge University Press.

Rubinstein, A. 1979. Equilibrium in Supergames with Overtaking Criterion, Journal of Economic Theory 21, 1-9.

Sannikov, Y. 2005 Games with Imperfectly Observable Actions in Continuous Time, mimeo., University of California, Berkeley.

Sekiguchi, T 1997. Efficiency in Repeated Prisoner's Dilemma with Private Monitoring, Journal of Economic Theory 76, 345-361.