Learning How to Cooperate: Optimal Play in Repeated Coordination Games

Vincent P. Crawford, Hans Haller

Econometrica, Volume 58, Issue 3 (May, 1990), 571-595.

Econometrica, Vol. 58, No. 3 (May, 1990), 571–595

LEARNING HOW TO COOPERATE: OPTIMAL PLAY IN REPEATED COORDINATION GAMES

BY VINCENT P. CRAWFORD AND HANS HALLER[1]

This paper proposes a characterization of optimal strategies for playing certain repeated coordination games whose players have identical preferences. Players' optimal coordination strategies reflect their uncertainty about how their partners will respond to multiple-equilibrium problems; this uncertainty constrains the statistical relationships between their strategy choices players can bring about. We show that optimality is nevertheless consistent with subgame-perfect equilibrium. Examples are analyzed in which players use precedents as focal points to achieve and maintain coordination, and in which they play dominated strategies with positive probability in early stages in the hope of generating a useful precedent.

KEYWORDS: Coordination, equilibrium refinements, focal points, multiple equilibria, repeated games.

1. INTRODUCTION

MUCH HAS BEEN LEARNED in recent years about the use of strategies with memory to enforce cooperation in long-term relationships whose short-term incentives impede cooperation. Much less is known about strategies for learning *how* to cooperate, despite evidence (see, for example, Macaulay (1963)) that this problem is equally important in practice. This paper proposes a characterization of optimal strategies for playing certain two-person repeated coordination games, and uses it to solve examples that help to explain some commonly observed features of the learning process.

The term *coordination game* refers here to a game in which players have identical preferences over strategy combinations, with two or more (Nash equilibrium) combinations at which each player's strategy choice is a unique best reply if the other player correctly anticipates it, but not in general otherwise. Coordination games present the problem of learning how to cooperate in its purest form: They involve no incentive problems, as these are normally characterized, because their efficient outcomes are supportable as equilibria. Nevertheless, playing them often involves real difficulties.[2] Similar difficulties lie at the heart of many questions normally analyzed under the implicit assumption that players can coordinate their strategy choices on any desired equilibrium. These include the design of incentive schemes; the characterization of which outcomes can be supported by implicit contract in a long-term relationship; and the determination of whether, and how, bargainers share the surplus from making an agreement.

[1] We owe thanks to Joseph Farrell, Yong-Gwan Kim, David Kreps, John McMillan, Michael Rothschild, Dennis Smallwood, Joel Sobel, and two referees for helpful comments. Crawford's work was supported in part by NSF Grants SES 8408059 and SES 8703337.

[2] This is strikingly confirmed, for games whose efficient outcomes are supportable as equilibria, in Van Huyck, Battalio, and Beil's (1988, 1990) coordination experiments; see also Crawford (1989).

Convincing answers to these questions must go beyond the observation that if rational players have commonly known, identical beliefs about each other's strategies, those beliefs must be consistent with some equilibrium in the game. But the usual game-theoretic techniques for analyzing games with multiple equilibria maintain these assumptions on beliefs, and proceed by "refining" Nash's notion of equilibrium to ensure that if it is already common knowledge that a particular equilibrium is expected to govern play, it is actually in each player's interest to play his equilibrium strategy.[3] Coordination games, by definition, have more than one equilibrium that passes this kind of test; as a result, the usual refinement techniques are not much help in playing them. Players in a coordination game may thus bear significant uncertainty about how other players will respond to its multiplicity of equilibria, even with complete information. This *strategic uncertainty* undermines the arguments that players should play according to any given equilibrium, and even calls into question the rationale for playing an equilibrium strategy.

One suggested resolution of this problem involves relaxing the assumption that behavior is fully rational, focusing instead on the adaptive process by which players learn from the history of play in analogous games. Predictions are then derived from the dynamics of players' strategy choices, based on the sets of initial conditions that lead to each possible final outcome. Examples of this approach include Fudenberg and Kreps (1988), who considered the possibility of constructing adaptive justifications for orthodox equilibrium refinements; and Sugden (1986) and Crawford (1989), who used adaptive models to study coordination problems. The models of Sugden and Crawford are closely related to those of evolutionary game theory (see, for example, Maynard Smith (1982)), the standard approach to coordination problems in biology.

Another possible resolution is suggested by the work of Harsanyi and Selten (1988). They proposed a theory of rational play that is *orthodox*, in that it assumes that players are rational, and always come to have identical, commonly known beliefs about each other's strategies before play begins (and therefore choose strategies that are in Nash equilibrium). But the mental tâtonnements Harsanyi and Selten use to model the convergence of players' beliefs respond in plausible ways to strategic uncertainty. Also, their theory always yields unique prescriptions, so that the existence of multiple equilibria does not immediately undermine their arguments for playing as it recommends.

Even though Harsanyi and Selten's theory captures many common intuitions about rationality, it has been less widely accepted and less often applied than other equilibrium refinements. This seems due mainly to its complexity and to another difficulty inherent in defining rationality in games with multiple equilibria: Players' rational responses to strategic uncertainty must be sensitive to their prior beliefs about how other players will respond to it; but it is difficult to

[3] Kohlberg and Mertens (1986) explain why the optimality of playing equilibrium strategies does not follow immediately from the definition of an equilibrium as a strategy combination made up of best replies to commonly held beliefs, and provide an introduction to the literature on equilibrium refinements.

specify these priors precisely with any confidence, and correspondingly difficult to offer a convincing characterization of rationality.

In this paper, we take a middle course between these two approaches, focusing on the dynamics of optimal learning in the presence of strategic uncertainty, but studying the learning process for repeated games in an orthodox framework like Harsanyi and Selten's. We sidestep the difficulties of specifying players' prior beliefs by focusing on games with symmetries, in which the principle of insufficient reason determines priors in a precise, convincing way; and we avoid most of the other difficulties Harsanyi and Selten encountered by studying coordination games.

The perfect coherence of players' preferences in coordination games suggests the working hypothesis that rational players should choose strategies that jointly maximize their expected payoffs, given their priors. We shall call such strategy combinations "optimal." Although the ability to use the principle of insufficient reason to specify priors is highly sensitive to deviations from the symmetries of the games we study, the priors themselves, and the qualitative features of the optimal strategies they imply, seem to us unlikely to be. Our approach therefore seems likely to shed some light on the implications of strategic uncertainty in more general games.

The rest of the paper is organized as follows. Section 2 introduces our approach to coordination by using it to solve for the optimal strategy combinations in a leading example. Section 3 then shows how to extend the arguments that underlie Section 2's solution to other repeated coordination games. Section 4 studies the relationship between our solutions of these games and subgame-perfect equilibrium, and Section 5 solves additional examples that illustrate some general features of learning in the presence of strategic uncertainty. Section 6 concludes the paper by discussing related work.

2. AN EXAMPLE

Our approach is best introduced by example. Consider the situation of two players who must play a repeated coordination game. In each stage, each player chooses (possibly randomly) between two actions, and players' actions determine their stage-game payoffs according to the payoff matrix

1, 1	0, 0
0, 0	1, 1

Each player observes both players' actions and the resulting payoffs in each stage, and has perfect recall of them. Each chooses a strategy that determines the choice probabilities of his actions in each stage as a function of the history of the game up to that stage. Players discount future payoffs at the same constant rate, each seeking to maximize the expected discounted sum of his stage-game payoffs. The structure of the game is common knowledge.

The game just described is a complete model of players' interaction. They cannot communicate except by playing it,[4] and there are no commonly observed variables on which to condition their strategy choices and thereby correlate them. In particular, even though the structure of the game is common knowledge, players have no common-knowledge description of it when they begin play. (In what follows, such a description will be called a "common language.") Instead, the structure is described to each of them privately and independently, so that their descriptions cannot focus their expectations on any outcome. It makes no sense for a player to try to coordinate on the efficient "top-left" equilibrium, for instance, because he has no reason to believe that "top" and "left" have the same meanings for his partner.

The essence of the coordination problem in our example is that, because players must choose their strategies independently, strategic uncertainty constrains the statistical relationship between their strategy choices, even though their strategy choices are unconstrained. The symmetries of the game in our example make it possible to model these constraints by assuming that players have no common language—our only departure from the assumptions of a standard repeated-game analysis. It is shown in Section 3 that this is equivalent to a joint restriction on players' strategy choices, in effect limiting them to strategy combinations that can be described without reference to their privately observed descriptions of the game. Players' private descriptions are the only way to discriminate between the symmetric actions and positions in this section's example. The constraints therefore effectively limit players to strategy combinations that respect these symmetries.

Even though our characterization of the effects of strategic uncertainty is motivated by assuming that players begin play without a common language, it can also serve as a model of the example's strategic uncertainty when players do have a common language, but lack a common-knowledge understanding about how to use it. In this more general interpretation, the function of the characterization's constraints is to filter out "focal-point" solutions based on the description of the game (see Schelling (1960)), isolating for study those that are description-invariant.

Because an orthodox analysis assumes common knowledge of players' beliefs about each other's strategies, it does not fully reflect the constraints implied by strategic uncertainty. In the search for a "plausible" equilibrium, the coherence of players' preferences in our example then creates a strong temptation to identify plausibility with efficiency, although standard refinements do not require this. By contrast, respecting the symmetries in our example prevents

[4] Explicitly modeled pre-play communication would involve players in a higher-order coordination problem; the rationale for ruling it out is discussed further in Section 6.

players from realizing any of its efficient outcomes, even though they are subgame-perfect equilibrium combinations of feasible strategies.

We now consider how strategic uncertainty limits what players can accomplish in our example. The analysis of the one-stage version is trivial. There, because each player must consider his partner's actions equally likely when he chooses his own action, his prior probability of coordination is one-half no matter what. Any strategy combination therefore yields the same probability distribution of payoff outcomes as the combination in which each player chooses each of his actions with equal probability, the only one that respects the symmetries of the game.

This conclusion extends to the first stage of the multi-stage version of our example, where it of course remains possible to make the probability of coordination one-half in each stage. But players can improve on this outcome in the multi-stage version by using their perfect recall and knowledge of the structure of the game to maintain coordination forever once they locate a pair of coordinated actions. They can do this either by repeating those actions, or by alternating deterministically between them and the other coordinated pair.

If players maintain coordination once they learn how to coordinate, their repeated-game payoffs are determined by how long it takes them to learn. The symmetries of the game effectively restrict players to identical repeated-game strategies that treat their actions symmetrically ex ante. Because players' first-stage choices do not affect their payoffs, given their priors, there is no loss of generality in characterizing their strategies by the probabilities with which they switch actions in each stage if they have not yet coordinated. Numbering the first stage 0, and letting s_t denote players' common switching probability in the tth stage and a their common discount factor, the sequence $\{s_t\}_{t=1}^{\infty}$ yields them expected repeated-game payoffs of

$$(1) \qquad \frac{1}{2}\frac{1}{1-a}+s_1(1-s_1)\frac{a}{1-a}+[1-2s_1(1-s_1)]s_2(1-s_2)\frac{a^2}{1-a} \\ +[1-2s_1(1-s_1)][1-2s_2(1-s_2)]s_3(1-s_3)\frac{a^3}{1-a}+\cdots.$$

Because raising $s_t(1-s_t)$, for any t, moves probability mass from more heavily discounted to less heavily discounted stage-game payoffs, players' payoffs are maximized, for any discount factor, when the s_t maximize each of these terms independently. This in turn requires that $s_t = 1/2$ for all t. Players' common repeated-game payoffs can be computed by substituting $s_t = 1/2$ into (1). Alternatively, denoting their payoffs for the remaining horizon by V and noting that these strategies make V stationary with an infinite horizon,

$$(2) \qquad V=\frac{1}{2}\frac{1}{1-a}+\frac{1}{2}aV=\frac{1}{(1-a)(2-a)}.$$

The role of randomization in this solution is to ensure that players' switching decisions are statistically independent. (From the point of view of dynamic programming, this randomization is an optimal response to the constraints on transition probabilities implied by our characterization of the example's strategic uncertainty.) Identical deterministic strategies would always make players switch simultaneously, and would therefore never bring about coordination, unless by chance at the start. The expected number of stages without coordination would therefore be half the total number of stages, infinite with an infinite horizon. By contrast, the best randomized strategy combinations yield coordination after only one stage on average.

It is instructive to compare this solution to what players could accomplish if they began play with a common language in which to identify their actions and/or their positions, assuming that this makes it possible for them to implement any strategy combination it allows them to distinguish. It is not clear that players can implement such combinations, but the results provide at least an upper bound on their payoffs, and are useful later on.

Note first that the efficient strategy combinations in which players coordinate in every stage are all symmetric across players. Identifying actions therefore allows instant coordination, even without identifying positions. Maintaining coordination by repeating actions then yields infinite-horizon repeated-game payoffs of $1/(1-a)$.

Identifying positions, but not actions, allows strategy combinations in which one player, specified in advance, stands pat and the other switches actions if coordination does not occur in the first stage. This ensures coordination by the second stage. Coordination can then be maintained by repeating actions. Because the probability of coordination in the first stage is one-half no matter what, this is the best possible outcome in this case. It yields infinite-horizon repeated-game payoffs of

$$\frac{1}{2}\frac{1}{1-a}+\frac{1}{2}\frac{a}{1-a}=\frac{1}{2}\frac{1+a}{1-a}. \tag{3}$$

There is half a stage without coordination on average. Players' payoffs in this case are less than their payoffs of $1/(1-a)$ when they can discriminate between actions and greater than their payoffs of $1/[(1-a)(2-a)]$ when they cannot discriminate between actions or positions. Identifying positions is thus an imperfect substitute for identifying actions.

These solutions have several interesting features. In each case, there is an optimal strategy combination whose strategies are in subgame-perfect equilibrium. (Verifying this requires extending the strategies in our solutions to unreached subgames in a way that is consistent with equilibrium in them; this is easily done.) This conclusion is not an immediate consequence of common knowledge of rationality, because we have not assumed common knowledge of players' beliefs about each other's strategies (see, for example, Aumann (1987) and Brandenburger and Dekel (1987)). Nor does it follow from the standard argument that a jointly optimal strategy combination must be an equilibrium when players have identical preferences, because the strategic uncertainty in our example effectively constrains the relationship between players' strategies in ways not allowed by that argument. However, it is shown in Section 4 that the set of optimal strategy combinations in the games we study always includes a subgame-perfect equilibrium.

In each case, players begin by searching for a pair of actions to serve as a coordination precedent and then use this precedent to maintain coordination. In effect, they use asymmetric history to "label" actions that cannot be distinguished at the start. The precedents on which these labels are based are focal points, in one of the senses in which Schelling used the term. Although "position labeling" does not occur in our solutions, the analysis shows that if players could identify their positions, they might benefit by using strategy combinations that treat them asymmetrically. This suggests that there might be games in which position labeling allows them to take advantage of more effective continuation strategies. Section 5 solves examples that confirm this, and illustrate some of the other uses of focal points.

Finally, our approach yields prescriptions that are less ambiguous than those that follow from an orthodox analysis. The game in our example has a large set of efficient equilibrium strategy combinations, which are essentially nonunique, in that maximizing expected repeated-game payoffs requires coordination of players' choices between them. Our characterization of strategic uncertainty yields optimal strategy combinations that require much less coordination. There are two kinds of nonuniqueness in our solutions when players cannot discriminate between positions or actions. The nonuniqueness implied by the irrelevance of players' first-stage choices is clearly inessential. The nonuniqueness implied by the existence of more than one way to maintain coordination is essential, but is easily resolved by adopting the common-sense rule of thumb of not experimenting with different actions once a combination with the highest possible stage-game payoffs has been located. Thus, if optimality is assumed, allowing for strategic uncertainty eliminates most of the need for coordination of strategy choices in our example when players cannot discriminate between positions or actions.

Players' optimal first-stage strategy choices remain essentially unique when they have a common language in which to identify positions, but not actions. In this case, however, there is a more troublesome nonuniqueness in later stages, with two essentially different but isomorphic sets of solutions, depending on which player stands pat if players do not coordinate in the first stage. The difficulty of coordinating on one of these sets raises doubts about the rationality of playing as our solutions suggest. Our analysis of this case is thus vulnerable to a higher-level version of our criticism of the orthodox analysis. We do not resolve the issues raised by this criticism, but Section 5's analysis of examples suggests that this more troublesome kind of nonuniqueness is rare when players have no common language.

The importance of strategic uncertainty, the usefulness of equilibrium analysis, and the difficulty of formalizing the idea of a focal point make it worth some effort to evaluate the generality of these results. To this end, Section 3 extends the techniques used to solve our example to a more general class of repeated coordination games, in preparation for the analysis of optimal learning in Sections 4 and 5.

3. REPEATED COORDINATION GAMES

This section describes the class of games to be studied, introduces the notions of symmetry used to specify players' prior beliefs in them, and shows how to characterize the resulting restrictions on beliefs. Where indicated below, definitions and arguments are stated more formally in the Appendix.

The games to be studied are finitely or infinitely repeated coordination games. These games have two players, whom we shall call R and C, for Row and Column, their positions in the stage game as *we* describe it. In each stage, R and C simultaneously choose actions, possibly randomly, from finite sets, A_R and A_C. R's payoff in the stage game when he chooses action i and C chooses action j is denoted r_{ij}; the analogous payoff for C is denoted c_{ij}. Thus, the matrices $r \equiv [r_{ij}]$ and $c \equiv [c_{ij}]$ are R's and C's stage-game payoff matrices. Unless otherwise noted, we assume throughout that R and C have identical preferences, so that c can be taken equal to r without loss of generality.

In the repeated game, each player observes both players' actions and payoffs in each stage, and players have perfect recall. Pre-play communication is not allowed, and there are no commonly observed variables on which to base correlated strategy choices. Players' (mixed) strategies are defined in the usual way, as probability distributions over the pure strategies that are feasible in the repeated game. Players' repeated-game payoffs are expected discounted sums of their stage-game payoffs; the discount factor is $a \in [0, 1)$, assumed constant over time and across players. R's and C's repeated-game strategies are denoted p and q, respectively, and their repeated-game payoffs are denoted $R(p, q)$ and $C(p, q)$.

All of the structure just described is common knowledge. This is usually implicitly assumed to include common knowledge of a description of the game, normally the analyst's. However, in justifying the restrictions on players' beliefs we use to model strategic uncertainty we assume that, even though players have common knowledge of the structure of the game, their descriptions of it are wholly or partly privately observed. When players begin play with a common-knowledge description of part of the structure, we say that they have a *common language* in which that part's actions or positions can be *identified*.

Some additional definitions are needed to formalize our restrictions on beliefs. Recall that players' preferences are identical, so that their stage-game payoff matrices can be taken to be equal. Two of a player's actions are *identical* if they yield the same stage-game payoffs for each of the other player's actions. In what follows, any of a player's actions that are identical are treated as a single action; thus players' common payoff matrix has distinct rows and columns. Two of a player's actions are *symmetric* if the associated rows or columns of the payoff matrix are equal up to a permutation of the columns or rows of his partner's *symmetric* actions. Thus, players' sets of symmetric actions are simultaneously determined; it is shown in the Appendix that they are always well-defined. Actions are *payoff-distinguished* if they are not symmetric. Players' positions in the game are *symmetric* if interchanging them yields (possibly after permuting rows and columns) an identical payoff matrix. Positions are *payoff-distinguished* if they are not symmetric. Finally, actions or positions are *distinguished* if they are payoff-distinguished or if players begin play with a common language in which they can be identified. Note that the sets of distinguished actions and positions are completely determined by players' common knowledge, and are therefore the same for both players.

These concepts are illustrated in the following examples, in which the numbers in the payoff matrices represent both the r_{ij} and the c_{ij}:

EXAMPLE 1:

1	0	0	0
0	1	0	0
0	0	1	0
0	0	1	1

Players' positions, and each player's first two actions, are symmetric. Each player's first two actions are payoff-distinguished from his last two actions, which are payoff-distinguished from each other as well. (The second and third actions, for example, are payoff-distinguished from each other because their payoffs are related to those of the fourth action in different ways.)

EXAMPLE 2:

3	0	2	1
0	3	1	2

Players' positions are payoff-distinguished. R's actions are symmetric, as are C's first two and last two actions. Each of C's first two actions is payoff-distinguished from each of his last two actions.

Thus, a game may have action-symmetry (of varying degrees) without position-symmetry, as in Example 2; the reverse, as in Example 1; both, as in Section 2's example; or (as is generically true) neither.

An *attainable* strategy combination is one in which any two of a player's undistinguished actions enter his strategy symmetrically, and in which players whose positions are undistinguished play identical strategies. Thus, an attainable strategy combination is one that can be described entirely in terms of players' common knowledge when they begin play, whether or not this includes a common language in which some or all of their positions or actions can be identified. An *optimal* attainable strategy combination is an attainable combination that maximizes both players' repeated-game payoffs over the set of attainable combinations. A *subgame-optimal* attainable combination is one that is optimal in the entire game and in every subgame. Players' optimal attainable strategies are said to be *essentially unique* if failures to coordinate their choices among optimal attainable strategy combinations do not reduce their expected payoffs in the entire repeated game. Attainability, optimality, and subgame-optimality (including the underlying notion of attainability in subgames) are formally defined in the Appendix.

In the analysis of Sections 4 and 5, we maintain the working hypothesis that players play an optimal (sometimes, a subgame-optimal) attainable strategy combination. Optimality was motivated in the Introduction. The restriction to attainable strategy combinations can be justified as follows. Fix a game, and assume that everything about it is common knowledge except how players, and the analyst, describe it to themselves. Players' descriptions of the game are privately observed permutations of the analyst's labelings of players' positions and actions. These permutations are independently drawn from commonly known distributions, which assign equal prior probability to alternative labelings of undistinguished positions or actions, but always assign the same labels for each player to distinguished positions or actions (whether payoff-distinguished or identified in a common language), thereby preserving the distinctions between them as common knowledge. The resulting game is one of incomplete information: Although its structure is common knowledge, the fact that players do not observe the analyst's description makes it impossible for them to compute their payoffs in terms of it.

Suppose that it is common knowledge that players subscribe to a given theory of rational play. Such a theory must map the data of a player's situation into beliefs about his partner's strategy choice, and thence into a set of expected payoff-maximizing strategies. This mapping must respect the limits imposed by players' information. In particular, a player's beliefs in the game of incomplete information just described must be independent of both his own private information and his partner's, in the first case because his partner does not observe his private information, and in the second because he does not observe his partner's. Thus, players' beliefs and their sets of expected payoff-maximizing strategies depend only on their common knowledge. Finally, there is no loss of generality in assuming that each player plays a mixed strategy identical to his partner's beliefs, because a player's uncertainty about his choice among expected payoff-maximizing strategies affects neither his own nor his partner's expected payoff. Players are therefore effectively restricted to attainable strategy combinations—those that can be described entirely in terms of their common knowledge of the structure and description of the game.

Because players have identical preferences, our assumption that they play an optimal attainable strategy combination is implied, given Theorem 1 below, by

Harsanyi and Selten's (1988) assumption that they play an equilibrium that is efficient in the set of equilibria in the game of incomplete information we use to model the original game's strategic uncertainty.

4. OPTIMAL AND EQUILIBRIUM PLAY

This section studies the relationship between optimal and equilibrium play. The main result can be stated as follows:

THEOREM 1: *Any subgame-optimal attainable strategy combination in a repeated coordination game is in subgame-perfect equilibrium in that game.*

REMARKS: Subgame-perfect equilibrium is defined for the original game of complete information, with no restrictions on players' strategy choices. Aumann (1987) and Brandenburger and Dekel (1987) have shown that, if players have common knowledge of rationality and of each other's beliefs, their behavior can be described by an appropriately chosen equilibrium. Theorem 1 shows that, at least in repeated coordination games with essentially unique optimal attainable strategies, these common-knowledge assumptions can be replaced by common knowledge that players play optimal attainable strategies.

PROOF: We first show that optimal attainable strategy combinations are in equilibrium, and then extend the argument to show that subgame-optimality implies subgame-perfectness. Formal definitions of the sets used to characterize players' attainable strategy combinations and the required extension of attainability to subgames are given in the Appendix.

Let P denote the set of R's strategies in which all of the actions in each of his groups of undistinguished actions enter symmetrically, and let Q denote the analogous set for C. Thus, the diagonal of $P \times P$, where $\times$ denotes Cartesian product, is the set of attainable strategy combinations when players' positions are not distinguished, and $P \times Q$ is the set of attainable strategy combinations when they are distinguished. When players' positions are not distinguished, an optimal attainable strategy combination solves the problem

$$\max_{p \in P} R(p,p). \tag{4}$$

R's repeated-game payoff function can represent both players' preferences, because they are identical. Because the restriction to attainable strategy combinations already reflects players' uncertainty about each other's strategy choices, and $R(p,p)$ already reflects any randomization that occurs in p, there is no need to take expectations in (4).

Suppose that p^* solves (4), but that (p^*, p^*) is not an equilibrium. Then there exists a p such that $R(p, p^*) > R(p^*, p^*)$. It is not difficult to show that, given Section 3's incomplete-information characterization of strategic uncertainty, there also exists a $\hat{p} \in P$ such that $(\hat{p}, p^*)$ induces the same probability distribution over outcomes as (p, p^*). Thus, because players have symmetric positions and identical preferences,

$$(5) \qquad R(\hat{p}, p^*) = R(p^*, \hat{p}) = R(p, p^*) > R(p^*, p^*).$$

Let $\bar{p} \equiv (1-\delta)p^* + \delta\hat{p}$, where $0 \leqslant \delta \leqslant 1$. Then, using (5),

$$(6) \qquad R(\bar{p}, \bar{p}) = (1-\delta)^2 R(p^*, p^*) + 2\delta(1-\delta) R(p^*, \hat{p}) + \delta^2 R(\hat{p}, \hat{p}),$$

where the equality follows from the bilinearity of $R(p, q)$ in p and q and the linearity of the expectations operator. It follows that, for δ sufficiently small, $R(\bar{p}, \bar{p}) > R(p^*, p^*)$. Because $(\bar{p}, \bar{p})$ is attainable by construction, this contradicts the hypothesized optimality of (p^*, p^*), proving the "equilibrium" part of the theorem for players with undistinguished positions.

When players' positions are distinguished, each of their optimal attainable strategy combinations solves the problem

$$(7) \qquad \max_{\substack{p \in P \\ q \in Q}} R(p, q),$$

with R's repeated-game payoff function again representing both players' preferences. There is again no need to take expectations, given the constraints in (7) and the definition of $R(\cdot)$. Let $(\bar{p}, \bar{q})$ solve (7) and suppose that this strategy combination is not an equilibrium, so that for one player, R without loss of generality, there exists a strategy p such that $R(p, \bar{q}) > R(\bar{p}, \bar{q})$. It can again be shown that there also exists a $\hat{p} \in P$ such that $(\hat{p}, \bar{q})$ induces the same probability distribution over outcomes as $(p, \bar{q})$, so that $R(\hat{p}, \bar{q}) = R(p, \bar{q}) > R(\bar{p}, \bar{q})$. This contradicts the hypothesized optimality of $(\bar{p}, \bar{q})$ and completes the proof of the "equilibrium" part of the theorem.

To prove subgame-perfectness, note that attainability allows the use of repeated-game strategies that use asymmetric history to (implicitly) label undistinguished actions or positions. It is shown in the Appendix that these labels function in subgames just like a more extensive common language (in which labeled actions or positions can be identified) in the entire repeated game. Because our proof of the "equilibrium" part of the theorem already allows for this possibility, it also shows that a subgame-optimal attainable strategy combination is in equilibrium in every subgame. *Q.E.D.*

Considering the symmetric one-stage Prisoner's Dilemma reveals that Theorem 1's conclusion fails, in general, when players have identical preferences only over attainable strategy combinations. The proof breaks down in this case because the strategy combinations $(\hat{p}, p^*)$ and $(p^*, \hat{p})$ are not attainable when player's positions are not distinguished. Thus, unless the game is a coordination game in the strong sense in which the term is used here, R's payoff function can no longer be used to represent C's preferences.

The proof makes clear that, as long as players' preferences are identical and their strategy spaces are convex, all that is really required is that each player's preferences are quasiconcave in his own strategy and jointly continuous in both players' strategies. Continuity is needed to ensure that, for small δ, the effect on players' preferences of playing $(\hat{p}, \hat{p})$ with probability δ^2 is dominated by the effect of playing $(\hat{p}, p^*)$ or $(p^*, \hat{p})$, each with probability $\delta(1-\delta)$. Quasiconcavity is needed to ensure that players are made better off by moving probability mass from the less preferred combination (p^*, p^*) to the more preferred combinations $(\hat{p}, p^*)$ and $(p^*, \hat{p})$. In our framework, continuity and quasiconcavity are implied by the linearity of expected payoffs in players' mixed-strategy probabilities.

These assumptions cannot be dispensed with, in general. To see what goes wrong without quasiconcavity, consider a one-stage game in which R and C simultaneously choose $x \in [0,1]$ and $y \in [0,1]$, respectively, to maximize the common payoff function $f(x,y) \equiv x(1-y)^2 + (1-x)^2 y$. Players' choice variables x and y are "pure" actions, and randomization is not allowed, to prevent the von Neumann-Morgenstern independence axiom that underlies the hypothesis that players maximize expected payoffs from restoring quasiconcavity. It is easily verified that this game has two pure-strategy equilibria, $(x,y) = (0,1)$ and $(x,y) = (1,0)$. Each player's actions are payoff-distinguished, but players' positions are not payoff-distinguished, because $f(x,y) \equiv f(y,x)$. Thus, in the absence of a common language in which positions can be identified, an optimal attainable strategy combination solves the problem

$$(8) \qquad \max_{x \in [0,1]} f(x,x) \equiv 2x(1-x)^2,$$

which has the unique solution $x = 1/3$. The optimal attainable strategy combination $(x,y) = (1/3, 1/3)$ is not an equilibrium.[5]

We close this section with a simple result about the convergence of optimal learning:

THEOREM 2: *In an infinite-horizon repeated coordination game, an optimal attainable strategy combination, with probability one, converges in finite time to a single action combination or to a set of action combinations that all yield players the same stage-game payoffs.*

PROOF:[6] Suppose, to the contrary, that there exists an optimal attainable strategy combination that, in some contingency with positive prior probability, returns infinitely often to two or more action combinations with different

[5] It may appear that we could construct an example with the same implication in which randomization is allowed, using Machina's (1982) non-expected utility characterization of decision-making under uncertainty, which relaxes the independence axiom in a way that allows quasiconvexity. The adjustments needed to do equilibrium analysis in this framework (see Crawford (1990)), however, yield Theorem 1's conclusion even with quasiconvex preferences. The inconsistency of optimal and equilibrium strategies in the example is thus due ultimately to players' inability to randomize, not to quasiconvexity.

[6] We are grateful to a referee for pointing out a lacuna in our earlier proof and helping us to fill it.

stage-game payoffs. Then there also exists an attainable strategy combination that remains at a (with positive probability) recurring action combination with the highest of these payoffs once it is located. Further, the expected discounted gains of locating action combinations with still higher payoffs are bounded. It follows that the conditional probability of locating a combination with higher payoff must remain bounded away from zero forever in a contingency with more than one recurring payoff, to justify foregoing the highest currently attainable payoff. But in the supposed contingency, the recurrence of a higher payoff has zero probability. Therefore, by continuity, the conditional probability of locating an action combination with higher payoff must converge to zero. That probability becomes sufficiently small, in finite time, so that optimality requires remaining from then on at the action combination with the highest payoff yet located. This contradicts the hypothesized optimality of the initial strategy combination, completing the proof. *Q.E.D.*

5. ADDITIONAL EXAMPLES

This section analyzes additional examples of repeated coordination games, with the goal of learning more about optimal learning. The examples show that it is sometimes optimal to forsake an efficient strategy combination forever in favor of an inefficient one that is less costly to locate; that it is sometimes optimal to use one coordination problem as a "trial run" for another whose solution has a higher payoff but is more costly to locate; that it is sometimes optimal to use precedents to label positions as well as actions; and that it is sometimes optimal to play dominated stage-game actions in early stages in the hope of generating a useful precedent.

The examples have infinite horizons; they could easily be modified to make the same points with finite horizons. There is no common language in which to identify positions and actions, which are therefore distinguished if and only if they are payoff-distinguished. We compute only those parts of players' optimal attainable strategy combinations that are needed to characterize the optimal path; our solutions are easily extended to ensure optimality in unreached subgames.

We begin with larger versions of Section 2's example; even this apparently straightforward generalization contains some surprises. We refer to the version in which the stage game has k actions for each player as the $k \times k$ case; the stage-game payoff matrix in the $k \times k$ case (again using a single number to represent both the r_{ij} and the c_{ij}) is a $k \times k$ identity matrix. Neither players' positions nor their actions are distinguished in these games.

When $k = 3$, any strategy combination yields the same probability of coordination in the first stage, $1/3$ in this case, as when each player plays each of his actions with equal probability. However, in the 3×3 case, even when positions cannot be identified, there is an attainable combination that ensures coordination from the second stage onward, and is therefore optimal. In it, each player uses the strategy: "Play each of your actions with equal probability in the first stage. If coordination results, maintain it by repeating your first-stage action. If not, rule out your first-stage action and the action that would have yielded coordination, given your partner's first-stage action; then play the action not ruled out from the second stage onward." Note that implementing this strategy requires only that each player knows the structure of the game and has perfect recall, in terms of his own description of the game. Letting V_3 denote each player's repeated-game payoff,

$$(9) \qquad V_3 = \frac{1}{3}\frac{1}{1-a} + \frac{2}{3}\frac{a}{1-a} = \frac{1+2a}{3(1-a)};$$

on average, there are two-thirds of a stage without coordination. Curiously, V_3 in (9) is greater than V_2 in (2) whenever $a > 1/2$: With high discount factors, players actually do better in the 3×3 case than in the 2×2 case!

The optimal attainable strategies just described are essentially unique: The prescribed method for finding a pair of actions to coordinate on in the second stage is plainly the only one that is guaranteed to work, and repeating actions is the only sure way to maintain coordination. The 3×3 case is unusual, however, in that knowing its structure allows players to identify a way to coordinate in the second stage from how they failed to coordinate in the first stage. This unrealistic feature is no longer present in sufficiently larger versions of this example. When $k \geqslant 6$, the attainable combination in which each player adopts this strategy is optimal: "Play each of your actions with equal probability in the first stage. If coordination results, maintain it by repeating your first-stage action. If not, rule out all of your actions but your first-stage action and the action that would have yielded coordination, given your partner's first-stage action. Then play the resulting 2×2 game using your part of its optimal attainable combination."

It is again easy to verify that each player has enough information to implement this strategy. The proof of optimality uses the facts that, when $k \geqslant 6$, failing to coordinate neither teaches players how to coordinate nor creates any asymmetric history they can use to label their positions, and it is always better to play a 2×2 version of the game than a $(k-2) \times (k-2)$ version.[7] There are, on average, $2(k-1)/k$ stages without coordination in this case. Letting V_k denote each player's repeated-game payoff in the $k \times k$ case,

$$(10) \qquad V_k = \frac{1}{k}\frac{1}{1-a} + \frac{k-1}{k}\frac{a}{(1-a)(2-a)} = \frac{2+(k-2)a}{k(1-a)(2-a)}.$$

[7] We do not discuss the 4×4 case, because its solution involves ambiguities that do not seem to yield general insights. The analysis of the 5×5 case is complicated by the fact that the solution of the 3×3 case plays a role in its solution whenever $a > 1/2$, again in a way that does not seem informative.

V_k is decreasing in k, with $V_k \to aV_2$ as $k \to \infty$. Thus, our intuition about the greater difficulty of solving larger coordination problems is confirmed for sufficiently large versions of this example, but the difficulty is bounded, equivalent in the limit to solving the 2×2 problem with an added one-stage delay. This is in part an artifact of the special structure of the example, in which all ways to coordinate are equally rewarding. It can be shown that the value for V_k given in (10) is also valid for $k = 4$, but lower than the true value for $k = 5$ when $a > 1/2$, for reasons like those noted above for $k = 3$.

It is interesting to note that, in this example, Section 2's rule of thumb of not experimenting when the relationship is working well actually follows from optimality in the subgames that follow coordination in the first stage. When players do not coordinate in the first stage, however, the subgame that results is equivalent to the 2×2 case, and shares its essential nonuniqueness; this of course is still resolved by Section 2's rule of thumb.

Our next example is a multi-stage generalization of a one-stage game discussed by Schelling (1960, Appendix C). Its stage-game payoff matrix is

1	0	0
0	1	0
0	0	b

.

Players' positions are not distinguished, and, unless $b = 1$, each player's first two actions are distinguished from his third, but not from each other. (When $b = 1$ and there is no common language, this example is equivalent to the 3×3 case of the example just analyzed.)

Players' optimal attainable strategy combination can be computed as follows. When $b > 1$, optimality plainly requires that each player play his "b" action in every stage. This yields repeated-game payoffs of $b/(1-a)$. When $b < 1$, it may be worth searching for a higher stage-game payoff. To see when this is so, note first that if each player plays his "b" action, or if players play uncoordinated "non-b" actions, neither learns anything that is useful in subsequent play: Players can distinguish their "b" actions at the start, and failures to coordinate with "non-b" actions neither teach them how to coordinate nor allow them to label their positions or reduce the scope of their search. If, however, one player happens to play his "b" action in a given stage while the other plays a "non-b" action, this history creates a positional asymmetry that allows players to ensure coordination from the next stage onward. They can do this by adopting the continuation strategy, "If the last stage was the first in which you played 'b' and your partner did not, or vice versa, then from now on play the action that would have yielded coordination with your partner's last-stage action if you played 'b', and repeat your last-stage action if you did not play 'b'."

Letting W denote players' repeated-game payoffs, these arguments reduce the game to a one-stage game with payoff matrix

	$\bar{s}$	$\bar{s}$	$1-2\bar{s}$
$\bar{s}$	$\frac{1}{1-a}$	aW	$\frac{a}{1-a}$
$\bar{s}$	aW	$\frac{1}{1-a}$	$\frac{a}{1-a}$
$1-2\bar{s}$	$\frac{a}{1-a}$	$\frac{a}{1-a}$	$b+aW$

Attainability restricts players to identical strategies, and requires each to play his undistinguished actions with equal probability, denoted $\bar{s}$, until history labels them. Thus, letting $s \equiv 2\bar{s}$, the game can be further reduced to

	s	$1-s$
s	$\frac{1}{2}\frac{1}{1-a}+\frac{1}{2}aW$	$\frac{a}{1-a}$
$1-s$	$\frac{a}{1-a}$	$b+aW$

The optimal combination is determined by choosing $s \in [0,1]$ to maximize

$$W = s^2\left[\frac{1}{2}\frac{1}{1-a}+\frac{1}{2}aW\right] + 2s(1-s)\frac{a}{1-a} + (1-s)^2(b+aW) \tag{11}$$

$$= \frac{s^2 + 2b(1-a)(1-s)^2 + 4as(1-s)}{(1-a)\left[2-as^2-2a(1-s)^2\right]}.$$

Straightforward calculations reveal that $\partial W/\partial s$ has the same sign as

$$a(1-b)s^2 + \left[(2+a)b+1-4a\right]s + 2(a-b). \tag{12}$$

When $b > a$, $(2+a)b+1-4a > (2+a)a+1-4a = (1-a)^2 > 0$, so the expression in (12) rises monotonically from $2(a-b) < 0$ when $s=0$ to $1-a > 0$ when $s=1$. Thus, W is maximized either at $s=0$ or $s=1$ when $b > a$. Because $s=0$ makes $W = b/(1-a)$ and $s=1$ makes $W = 1/[(1-a)(2-a)]$, $s=0$ is optimal whenever $b > 1/(2-a)$ and $s = 1$ is optimal whenever $b < 1/(2-a)$; both are optimal when $b = 1/(2-a)$. (Because $1/(2-a) > a$, these cases are all consistent with $b > a$.) In view of Theorem 1, it is interesting to note that both $s = 0$ and $s = 1$ are equilibria in the reduced game (given the associated values of W) when $b > a$.

It is easiest to analyze the case where $b \leqslant a$ by noting that $\partial^2 W/\partial s\,\partial b \leqslant 0$ for all $s \in [0,1]$, with strict inequality unless $s = 1$. Because $s = 1$ is optimal whenever $a < b < 1/(2-a)$, this implies that $s = 1$ remains optimal for all $b \leqslant a$. It is easily verified that $s = 1$ is always in equilibrium in the reduced game in this case as well.

To sum up, $s = 1$ is uniquely optimal for all $b < 1/(2-a)$, both $s = 0$ and $s = 1$ are optimal when $b = 1/(2-a)$, and $s = 0$ is uniquely optimal for all $b > 1/(2-a)$ except $b = 1$ (in which case the optimal attainable strategy combination takes a different form). W is independent of b until it reaches $1/(2-a)$ and increasing in b from then on, except for a single downward discontinuity at $b = 1$. This discontinuity arises because at that point players' "b" actions are not payoff-distinguished; it would not arise if they began play with a common language in which to identify their "b" actions.

These optimal strategies are essentially unique, with two exceptions. (Recall that essential uniqueness was defined in terms of the payoffs in the entire repeated game, and thus does not restrict strategy combinations in unreached subgames.) When $b < 1/(2-a)$, the nonuniqueness of the 2×2 case appears again; and in the knife-edge case in which $b = 1/(2-a)$, there is a more serious essential nonuniqueness, involving two nonisomorphic solutions.

This example illustrates a useful technique and makes some important points. Whenever $b > 1/(2-a)$, it is optimal for players to play their "b" actions from the start, even though when $b < 1$, this means that they will never find the highest possible stage-game payoff; in this case, a focal-point solution is better than searching for a strategy combination that is efficient in the stage game. Thus, there is no reason, in general, to expect optimal learning to lead to an efficient outcome in the stage game.

The optimal attainable strategy combinations in our earlier examples implicitly use history to label undistinguished actions. The example just analyzed contains the interesting but imperfectly realized possibility that optimal strategies may use history to label undistinguished positions as well. In particular, if just one player played "b" in a given stage, subgame-optimal strategies would use this asymmetric history to single out a player to switch actions to coordinate in the next stage; hence the $a/(1-a)$ entries in our reduced payoff matrix. If it were ever optimal to set s strictly between zero and one, this position-labeling would occur with positive prior probability. Unfortunately, whenever raising s above zero yields an improvement in this example, it is even better to raise it all the way to one, so that position-labeling "occurs" only in unreached subgames.

The principles that underlie our solution of this example extend to more general games in which players must choose between coordination problems and solving one yields no information that is useful in solving the other.

Consider, for instance, the stage-game payoff matrix

1	0	ζ	0
0	1	0	ζ
ζ	0	$d+e$	d
0	ζ	d	$d+e$

with $0 < e \ll d < d+e < 1$. (ζ is a parameter, specified below.) Players' positions are not distinguished, and each player's first two actions are distinguished from his last two but not from each other, and vice versa.[8]

When $\zeta = 0$, solving the coordination problem in the lower right 2×2 block does not help to solve the coordination problem in the upper left 2×2 block, and vice versa. An analysis like that of the previous example shows that it is then optimal to specialize in the upper left block if $1/[1-a][2-a] > [(2-a)d+e]/[1-a][2-a]$, and otherwise to specialize in the lower right block. As before, the associated optimal strategies are essentially unique, except for the nonuniqueness they inherit from Section 2's example. When, on the other hand, $\zeta > 0$, finding a "$d+e$" allows players to use their knowledge of the structure of the game to locate a "1", and vice versa, so that solving one coordination problem reveals how to solve the other. (Players can use a "$d+e$" to find a "1" by adopting the continuation strategy, "If the last stage was the first in which your payoff was $d+e$, then from now on play the action that would have yielded ζ, given your partner's last-stage action.") Because $d+e<1$, optimality requires transferring to a "1" immediately after a "$d+e$" is located. It can be shown, using techniques like those used to analyze the previous example, that when e and ζ are sufficiently small, the optimal attainable strategy combination when $d > 1/2$ first solves the "safer" coordination problem in the lower right block and then transfers to a "1", and the optimal combination when $d+e < 1/2$ specializes in the upper left block and follows the solution of Section 2's example. Intuitively, when solving one coordination problem reveals how to solve the other, and vice versa, optimality requires using the one with higher expected payoffs when not yet solved as a trial run for the one with higher expected payoffs once solved.

There is a new kind of essential nonuniqueness in this solution, which is not resolved by Section 2's rule of thumb as stated: The continuation strategy in which "yielded 0" replaces "yielded ζ" is also consistent with optimality.

Our final example shows that position-labeling can occur with positive probability in optimal play, and that it is sometimes optimal to play actions that are dominated in the stage game with positive probability, in the hope of generating

[8] We are grateful to a referee for suggesting this example.

a useful precedent. The stage-game payoff matrix is a $k \times k$ identity matrix bordered by a row and a column of zeros. (These zeros could be perturbed to make players' weakly dominated "0" actions strictly dominated, without altering the results.) Each player's "0" action is distinguished from his non-"0" actions, but his non-"0" actions are not distinguished from each other. Players' positions are not distinguished.

We assume, for simplicity, that $k \geqslant 6$. (The same point could be made with $k = 4$ or 5, but not with $k = 2$ or 3; see footnote 7.) Attainability restricts players to identical strategies and requires each player to play his k non-"0" actions with equal probability in the first stage. We can therefore describe an attainable strategy combination by specifying this probability; let z denote k times the probability with which each player plays each of his non-"0" actions or, equivalently, the total probability with which he plays these actions. Letting U_k stand for players' repeated-game payoffs, the game reduces to a one-stage game with payoff matrix

$$\begin{array}{c|c|c|}
 & z & 1-z \\
\hline
z & \dfrac{1}{k}\dfrac{1}{1-a} + \dfrac{k-1}{k}\dfrac{a}{(1-a)(2-a)} & \dfrac{a}{1-a} \\
\hline
1-z & \dfrac{a}{1-a} & aU_k \\
\hline
\end{array}\ .$$

The reduced payoff matrix for this example reflects the facts that, if exactly one player plays his "0" action in a given stage, there is a continuation strategy (analogous to the one described above for the "b" example) that ensures coordination from the next stage onward; that players learn nothing of value if both play "0"; and that if they play uncoordinated non-"0" actions, our analysis of the "b" example with $b = 0$ shows that their best continuation strategy is to play the 2×2 game labeled by those actions using its optimal attainable strategy combination.

The optimal combination is determined by choosing $z \in [0,1]$ to maximize

$$(13)\qquad \begin{aligned} U_k &= z^2\left[\frac{1}{k}\frac{1}{1-a} + \frac{k-1}{k}\frac{a}{(1-a)(2-a)}\right] \\ &\quad + 2z(1-z)\frac{a}{1-a} + (1-z)^2 aU_k \\ &= \frac{z^2 + 2akz(1-z) + a(k-1)z^2/(2-a)}{k(1-a)\left[1-a(1-z)^2\right]}. \end{aligned}$$

Straightforward calculations reveal that $\partial U_k/\partial z$ has the same sign as

$$(14)\qquad a(2-ak)z^2 + (2-2a-3ak+2a^2k)z + a(2-a)k;$$

this implies that U_k qualitatively resembles a cubic, with $\partial U_k/\partial z > 0$ when $z = 0$, and $\partial U_k/\partial z < 0$ when $z = 1$ if and only if $k > 2/a$. Intuitively, it is never optimal to play the "0" action with probability one, but it is advantageous to play it with positive probability whenever $a/(1-a) > V_k$ as given by (10), because small probabilities make it much more likely that exactly one player will end up playing his "0" action than that both will.

When $k > 2/a$, the expression in (14) is positive at $z = 0$ and negative at $z = 1$, which implies, because it is quadratic, that it has a unique zero strictly between $z = 0$ and $z = 1$. As this value of z is also a zero of $\partial U_k/\partial z$, with $\partial U_k/\partial z > 0$ for lower values of z and $\partial U_k/\partial z < 0$ for higher values, it determines an optimal value of z, which can be solved for by equating expression (14) to zero and choosing the relevant root. In the strategy combination associated with this value of z, players play their dominated "0" actions with positive probability in the hope that it will generate history that can be used to label their positions. In this example, such history also labels their actions, allowing them to coordinate from the next stage onward.

When $k \leqslant 2/a$, each player's first strategy in the reduced payoff matrix displayed above weakly dominates his second strategy, because $1/(1-a) > U_k$. It then follows from Theorem 1, recalling the restriction to identical strategies for undistinguished players, that $z = 1$ is uniquely optimal and players' weakly dominated "0" actions are irrelevant.

The optimal attainable strategies in this example are essentially unique, except that its 2×2 subgames inherit the essential nonuniqueness of Section 2's example; as before, this can be eliminated by adopting the rule of thumb proposed in Section 2. Note in particular that, even though the solution involves using history to label positions in the subgames that follow the first use of a "0" action, this entails no nonuniqueness, essential or otherwise: In this case (unlike in Section 2's example with identified positions), optimality uniquely determines which player must switch actions.

This example makes several important points: Players may find it useful to play strategies designed to generate history that can be used to label positions as well as actions. Even when players have identical preferences, optimality can require them to play actions that are dominated in the stage game with positive probability in early stages, so deleting such actions before analyzing the game may eliminate their best coordination strategies. Finally, optimal learning in the face of strategic uncertainty may use history in far more complex ways than an orthodox analysis suggests.

6. RELATED WORK

This section concludes the paper with a discussion of related work.

Perhaps the most interesting approach to coordination not already mentioned is the literature on pre-play communication; see for example Kalai and Samet (1985), Farrell (1987, 1988), Farrell and Gibbons (1989), Matthews (1989), and Matthews and Postlewaite (1989). These models typically allow one or more stages of costless communication ("cheap talk") before any directly payoff-relevant decisions are made. Players begin play with a common language rich enough to identify any desired agreement. Agreements are non-binding, but the communication stage is tied to the directly payoff-relevant part of the game by studying only equilibria in which any agreement that players have incentives to follow is implemented. The central theme of these papers is that, as in Crawford and Sobel's (1982) analysis of costless strategic communication, communication facilitates coordination, but its usefulness is limited by differences between players' preferences.

In related work, Dixit and Shapiro (1986) study learning by actual, costly play in a dynamic model of firms' entry decisions (compare Farrell (1987)). Van Damme (1987, Chapter 10) shows that strategic stability in the sense of Kohlberg and Mertens (1986) has some cutting power in repeated coordination games, in that a player can signal by his play in early stages, via "forward induction," how he expects the rest of the game to be played. Van Damme (1989) and Ben-Porath and Dekel (1988) also study costly communication, showing that its possibility can facilitate coordination (even when no costs are actually incurred in equilibrium!) in strategically stable repeated-game equilibria. Farrell and Saloner (1988) allow both costless communication and directly payoff-relevant preemptive moves, showing that using these coordination techniques together can yield better outcomes than either technique by itself.

The possibility of pre-play communication raises an important question about our analysis. Because our players have identical preferences, sensible use of costless communication would presumably solve the coordination problems we study. But characterizing the effects of costless communication requires overcoming formidable multiple-equilibrium problems, as in Crawford and Sobel (1982). Ruling out pre-play communication allows a characterization of players' rational responses to strategic uncertainty that seems to us more convincing than those now available for more general models. In real relationships, differences in players' preferences impede truthful communication. Further, for reasons like those that make complete contingent contracting an unrealistic solution to the problems addressed by contract theory, it is normally impossible or uneconomical to communicate completely how the relationship is to be organized. As a result, significant coordination problems remain after players have taken full advantage of the communication possibilities. Something like the understanding of coordination provided by our approach seems essential in understanding these problems.

Dept. of Economics, University of California, San Diego, La Jolla, CA 92093, U.S.A.

and

Dept. of Economics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, U.S.A.

Manuscript received September 1987; final revision received July, 1989.

APPENDIX

In this Appendix, we formalize some of the notions used in the text.

Two of a player's actions are symmetric if the associated rows or columns of the payoff matrix are equal up to a permutation of the columns or rows of the other player's symmetric actions. To formalize this idea, using the notation introduced in Section 3, write $A_R \equiv \{1,\ldots,n\}$ and $A_C \equiv \{1,\ldots,m\}$. Then $r=[r_{ij}]$ is an $n\times m$ matrix, with ith row $r^i \equiv (r_{i1},\ldots,r_{im})$ and jth column $r_j \equiv (r_{1j},\ldots,r_{nj})'$, where the prime denotes transposition. Consider a partition $\boldsymbol{R}\equiv(R_1,\ldots,R_N)$ of A_R and a partition $\boldsymbol{C}\equiv(C_1,\ldots,C_M)$ of A_C. The pair $(\boldsymbol{R},\boldsymbol{C})$ is *compatible* if $\boldsymbol{R}$ and $\boldsymbol{C}$ share the following two properties:

1. If $1\leqslant k\leqslant N$, $R_k\in\boldsymbol{R}$, and $i,j\in R_k$, then there exist (possibly trivial) permutations γ_l of C_l, $l=1,\ldots,M$, and a permutation γ: $A_C\to A_C$ of A_C such that $\gamma(t)=\gamma_l(t)$ if $t\in C_l$, $l=1,\ldots,M$, and $r_{i,t}=r_{j,\gamma(t)}$ for $t=1,\ldots,m$. In other words, for any $i,j\in R_k\in\boldsymbol{R}$, the rows r^i and r^j are equal up to a permutation of columns that respects the partition $\boldsymbol{C}$.

2. If $1\leqslant l\leqslant M$, $C_l\in\boldsymbol{C}$, and $i,j\in C_l$, then there exist (possibly trivial) permutations δ_k of R_k, $k=1,\ldots,N$, and a permutation δ: $A_R\to A_R$ of A_R such that $\delta(t)=\delta_k(t)$ if $t\in R_k$, $k=1,\ldots,N$, and $r_{t,i}=r_{\delta(t),j}$ for $t=1,\ldots,n$. In other words, for any $i,j\in C_l\in\boldsymbol{C}$, the columns r_i and r_j are equal up to a permutation of rows that respects the partition $\boldsymbol{R}$.

Now the set P_R of partitions of A_R is partially ordered by defining $\boldsymbol{R}'\geqslant\boldsymbol{R}$ if and only if $\boldsymbol{R}'$ is coarser than or equal to $\boldsymbol{R}$; an analogous partial order is defined on the set P_C of partitions of A_C. The set $P_R\times P_C$ is partially ordered by defining $(\boldsymbol{R}',\boldsymbol{C}')\geqslant(\boldsymbol{R},\boldsymbol{C})$ if and only if $\boldsymbol{R}'\geqslant\boldsymbol{R}$ and $\boldsymbol{C}'\geqslant\boldsymbol{C}$. These partially ordered sets are lattices with $(\boldsymbol{R},\boldsymbol{C})\wedge(\boldsymbol{R}',\boldsymbol{C}')=(\boldsymbol{R}\wedge\boldsymbol{R}',\boldsymbol{C}\wedge\boldsymbol{C}')$ and $(\boldsymbol{R},\boldsymbol{C})\vee(\boldsymbol{R}',\boldsymbol{C}')=(\boldsymbol{R}\vee\boldsymbol{R}',\boldsymbol{C}\vee\boldsymbol{C}')$ for all $\boldsymbol{R},\boldsymbol{R}'\in P_R$ and $\boldsymbol{C},\boldsymbol{C}'\in P_C$. An obvious but important observation is that if $(\boldsymbol{R},\boldsymbol{C})$ and $(\boldsymbol{R}',\boldsymbol{C}')$ are compatible pairs, then their supremum, $(\boldsymbol{R},\boldsymbol{C})\vee(\boldsymbol{R}',\boldsymbol{C}')$, is also a compatible pair. A compatible pair always exists: for example, the discrete partitions of A_R and A_C.

It follows from the above observation and the finiteness of P_R and P_C that there exists a $\geqslant$-maximal compatible pair, which we denote $(\boldsymbol{R}^*,\boldsymbol{C}^*)$ in what follows. $(\boldsymbol{R}^*,\boldsymbol{C}^*)$ reflects the maximal degree of action-symmetry in the game. We call two actions $i,j\in A_R$ *symmetric* if and only if there is an $R_k\in\boldsymbol{R}^*$ such that both i and j belong to R_k. Similarly, two actions $i,j\in A_C$ are *symmetric* if and only if there is a $C_l\in\boldsymbol{C}^*$ such that both i and j belong to C_l.

Players' positions are *symmetric* if the matrix r is square and, possibly after permuting rows and columns, coincides with its transpose, r'. For the present, we identify position-symmetry with symmetry of r, in the usual sense that $r=r'$. (We discuss the modifications required by our more general notion of position-symmetry below.) Position-symmetry then implies that $\boldsymbol{R}^*=\boldsymbol{C}^*$, because if $(\boldsymbol{R},\boldsymbol{C})$ is compatible, then $(\boldsymbol{C},\boldsymbol{R})$ and $(\boldsymbol{R}\vee\boldsymbol{C},\boldsymbol{R}\vee\boldsymbol{C})$ are also compatible.

For a one-stage game, the sets P and Q used to describe attainable strategy combinations in the proof of Theorem 1 can be defined as follows:

$$P\equiv\left\{p\in\mathbb{R}^n_+ \text{ s.t. } \sum_{i=1}^{n}p_i=1 \text{ and } i,j\in R_k\in\boldsymbol{R}^* \text{ imply that } p_i=p_j\right\}\tag{A1}$$

and

$$Q\equiv\left\{q\in\mathbb{R}^m_+ \text{ s.t. } \sum_{i=1}^{m}q_i=1 \text{ and } i,j\in C_l\in\boldsymbol{C}^* \text{ imply that } q_i=q_j\right\}.\tag{A2}$$

An *attainable* strategy combination is then a pair $(p,q)\in P\times Q$ when players' positions are distinguished, and a pair $(p,p)\in P\times Q$ when players' positions are not distinguished.

The concept of attainability is defined analogously for repeated games. In repeated games, actions or positions that are not distinguished at the start of play can be distinguished in subgames by asymmetric history. Any formal argument using "subgame-optimality" requires extending the concept of attainability to subgames. This is done as follows.

For a given history, h, the distinctions between actions made possible by history are described by partitions $\boldsymbol{R}^*(h)$ of A_R and $\boldsymbol{C}^*(h)$ of A_C such that $(\boldsymbol{R}^*,\boldsymbol{C}^*)\geqslant(\boldsymbol{R}^*(h),\boldsymbol{C}^*(h))$ and $(\boldsymbol{R}^*(h),\boldsymbol{C}^*(h))$ is a compatible pair. Two actions $i,j\in A_R$ are *distinguished at history h* if they belong to different elements of $\boldsymbol{R}^*(h)$. Similarly, $i,j\in A_C$ are *distinguished at history h* if they belong to different elements of $\boldsymbol{C}^*(h)$. If history h' is a continuation of history h, then $\boldsymbol{R}^*(h')$ and $\boldsymbol{C}^*(h')$ are respectively finer than or equal to $\boldsymbol{R}^*(h)$ and $\boldsymbol{C}^*(h)$. With no or "0" history at the start of play, the information so far revealed by history is described by the partitions $\boldsymbol{R}^*(0)$ and $\boldsymbol{C}^*(0)$; these equal

$\{A_R\}$ and $\{A_C\}$, respectively, when players begin play with no common language. Finer partitions $\boldsymbol{R}^*(0)$ and $\boldsymbol{C}^*(0)$ can be used to describe a "prehistoric" common language.

Perfect recall allows us to restrict attention to *behavior strategies*, which specify, for each stage, probability distributions over actions that depend in any desired way on the history of all previous stages. The behavior strategies *attainable at history h* are given, suppressing their dependence on history, by

$$\text{(A3)} \qquad P(h) \equiv \left\{ p \in \mathbb{R}_+^n \text{ s.t.} \sum_{i=1}^{n} p_i = 1 \text{ and } i, j \in R_k \in \boldsymbol{R}^*(h) \text{ imply that } p_i = p_j \right\}$$

and

$$\text{(A4)} \qquad Q(h) \equiv \left\{ q \in \mathbb{R}_+^m \text{ s.t} \sum_{i=1}^{m} q_i = 1 \text{ and } i, j \in C_l \in \boldsymbol{C}^*(h) \text{ imply that } q_i = q_j \right\}.$$

An *attainable* strategy combination in the repeated game is a pair of strategies p and q such that, at any history h, p prescribes a behavior strategy $p(h) \in P(h)$ and q prescribes a behavior strategy $q(h) \in Q(h)$.

Several tedious details have so far been omitted from our discussion. We now note these omissions and, when deemed necessary, explain how to rectify them without changing our results and arguments:

A. If players have symmetric positions and there is no common language, then $\boldsymbol{R}^*(0) = \boldsymbol{R}^* = \boldsymbol{C}^* = \boldsymbol{C}^*(0)$. $\boldsymbol{R}^*(h)$ and $\boldsymbol{C}^*(h)$ can differ for histories other than 0. However, for any history h, there is a symmetric history $\bar{h}$ such that $\boldsymbol{R}^*(\bar{h}) = \boldsymbol{C}^*(h)$ and $\boldsymbol{C}^*(\bar{h}) = \boldsymbol{R}^*(h)$. Symmetric play by symmetric players should be understood in this wider sense.

B. In the particular examples analyzed in the text, it is clear which partitions $\boldsymbol{R}^*(h)$ and $\boldsymbol{C}^*(h)$ are associated with a particular history h. For our general results it is not necessary to—and we do not—specify an algorithm to determine these partitions for arbitrary games.

C. Suppose that the matrix r is not symmetric, but that players' positions are symmetric in the more general sense that r is obtained by applying a row permutation, γ, and a column permutation, δ, to its transpose, r'. Then $\boldsymbol{C}^* = \gamma^{-1}(\boldsymbol{R}^*)$ and $\boldsymbol{R}^* = \delta^{-1}(\boldsymbol{C}^*)$, with analogous modifications in the rest of the discussion.

REFERENCES

AUMANN, ROBERT (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55, 1–18.

BEN-PORATH, ELCHANAN, AND EDDIE DEKEL (1988): "Coordination and the Potential for Self Sacrifice," Research Paper 984, Graduate School of Business, Stanford University.

BRANDENBURGER, ADAM, AND EDDIE DEKEL (1987): "Rationalizability and Correlated Equilibria," *Econometrica*, 55, 1391–1402.

CRAWFORD, VINCENT P. (1989): "An 'Evolutionary' Explanation of Van Huyck, Battalio, and Beil's Experimental Results on Coordination," Discussion Paper 89–28, University of California, San Diego.

——— (1990): "Equilibrium without Independence," *Journal of Economic Theory*, 50, 127–154.

CRAWFORD, VINCENT P., AND JOEL SOBEL (1982): "Strategic Information Transmission," *Econometrica*, 50, 1431–1451.

DIXIT, AVINASH, AND CARL SHAPIRO (1986): "Entry Dynamics with Mixed Strategies," in *The Economics of Strategic Planning*, ed. by L. Thomas. Lexington, MA: Lexington Books.

FARRELL, JOSEPH (1987): "Cheap Talk, Coordination, and Entry," *Rand Journal of Economics*, 18, 34–39.

——— (1988): "Communication, Coordination, and Nash Equilibrium," *Economics Letters*, 27, 209–214.

FARRELL, JOSEPH, AND ROBERT GIBBONS (1989): "Cheap Talk Can Matter in Bargaining," *Journal of Economic Theory*, 48, 221–237.

FARRELL, JOSEPH, AND GARTH SALONER (1988): "Coordination through Committees and Markets," *Rand Journal of Economics*, 19, 235–252.

FUDENBERG, DREW, AND DAVID KREPS (1988): *A Theory of Learning, Experimentation, and Equilibrium in Games*, manuscript, MIT and Stanford University.

HARSANYI, JOHN, AND REINHARD SELTEN (1988): *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: The MIT Press.
KALAI, EHUD, AND DOV SAMET (1985): "Unanimity Games and Pareto-Optimality," *International Journal of Game Theory*, 14, 41–50.
KOHLBERG, ELON, AND JEAN-FRANCOIS MERTENS (1986): "On the Strategic Stability of Equilibria," *Econometrica*, 54, 1003–1037.
MACAULAY, STEWART (1963): "Noncontractual Relations in Business: a Preliminary Study," *American Sociological Review*, 28, 55–67.
MACHINA, MARK (1982): "'Expected Utility' Analysis without the Independence Axiom," *Econometrica*, 50, 277–323.
MATTHEWS, STEVEN (1989): "Veto Threats: Rhetoric in a Bargaining Model," *Quarterly Journal of Economics*, 104, 347–369.
MATTHEWS, STEVEN, AND ANDREW POSTLEWAITE (1989): "Pre-Play Communication in Two-Person Sealed-Bid Double Auctions," *Journal of Economic Theory*, 48, 238–263.
MAYNARD SMITH, JOHN (1982): *Evolution and the Theory of Games*. New York: Cambridge University Press.
SCHELLING, THOMAS (1960): *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.
SUGDEN, ROBERT (1986): *The Economics of Rights, Cooperation, and Welfare*. Oxford: Basil Blackwell.
VAN DAMME, ERIC (1987): *Stability and Perfection of Nash Equilibria*. Berlin: Springer Verlag.
——— (1989): "Stable Equilibria and Forward Induction," *Journal of Economic Theory*, 48, 476–496.
VAN HUYCK, JOHN, RAYMOND BATTALIO, AND RICHARD BEIL (1988): "Strategic Uncertainty, Equilibrium Selection Principles, and Coordination Failure in Average Opinion Games," Working Paper 88-20, Texas A&M University.
——— (1990): "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *American Economic Review*, 80, 234–248.