3.2

Equilibrium, Common Knowledge, and Optimal Sequential Decisions

JOSEPH B. KADANE AND TEDDY SEIDENFELD

I. INTRODUCTION

In a paper described as an initial exploration of what two Bayesians need to know in order to play a sequential game against each other, DeGroot and Kadane (1983) argue that optimal sequential decisions need not conform to much of what traditional game theory requires of rational play. Specifically:

1. The players' optimal strategies need not form a Nash equilibrium.
2. Nor do the players need to know (or even believe that they know) the optimal choices of their opponent; there is no requirement of "common knowledge," in that sense.

Nonetheless, these authors propose that

3. Reasoning by backward induction succeeds in locating optimal play.

Each of these three claims is a point of active dispute. For example, regarding (1), in an extended defense of a refined equilibrium concept, Harsanyi and Selten (1988) argue that because of common knowledge

We dedicate this paper to the memory of our dear friend and colleague, Morris H. DeGroot.

Research for this work was supported, in part, by NSF Grants DMS-8705646, DMS-8701770, SES-8900025, by ONR Contract N00014-89-J-1851, and by the Buhl Foundation.

From *Knowledge, Belief and Strategic Interaction*, edited by Cristina Bicchieri and Maria Luisa Dalla Chiara (Cambridge: Cambridge University Press, 1992), pp. 27–45.

(of mutual rationality) the agents *ought to* settle on an equilibrium solution – but a refined equilibrium (see also Harsanyi 1989). Aumann (1987), like Harsanyi and Selten, seeks to reconcile Bayesian and game-theoretic rationality but is led to a theory of correlated equilibrium based on an assumption of a "common prior" for different players. Relating to (2), Binmore and Brandeburger (1988), wary of common-knowledge assumptions, seek other grounds for justifying equilibrium solutions. Concerning (3), Bicchieri (1989) questions the validity of backward induction in cases where the agents have too much or too little common knowledge.

The position we take in this chapter is based on the initial exploration of DeGroot and Kadane (1983). We extend the central example of that work to include the Harsanyi–Selten "trembling hand" model of choices. We argue that, in the extended example as in the original one, equilibrium is not a norm for rational play. (This is a position already announced by Kadane and Larkey 1982.) Based on our views about what may serve as states of uncertainty in a common prior across players, we argue that rational play need not result in a correlated equilibrium. And we argue that there is no problem of too much common knowledge. That is, with respect to the standards of expected utility, backward induction is a valid method for arriving at optimal play. In short, we respectfully disagree with each of the authors mentioned above!

The analysis we offer in this essay supports the condition of "rationalizability" for strategies, a view intelligently defended in papers by Bernheim (1984) and Pearce (1984). In addition, a version of the DeGroot–Kadane game that introduces common priors leads us to a conclusion similar to that reported by Rubinstein (1989); namely, approximating common knowledge does not yield strategies that approximate optimal play under common knowledge. Thus, we endorse the attitude expressed by many who emphasize a careful assessment of *what* players in a game know of each other's beliefs and, especially for sequential games, *when* they come to know it.

II. THE DEGROOT–KADANE GAME

The DeGroot–Kadane game is played between two agents by successive moves of a visible pointer located on the real line. Following their (1983) presentation, suppose the game has three moves: First player 1 moves the pointer, then player 2, and last player 1 moves it again. The payoff (a loss given in utiles) to each player is a function of two components: how far the player has moved the pointer and, after the final move, how far the pointer is from that player's designated target value (x for player 1 and y for player 2). Put formally, let s_0 be the initial location of the pointer and let s_i be its location after the three moves (i = 1, 2, 3). Thus, $u = s_1 - s_0$ is player 1's first move, $v = s_2 - s_1$ is player 2's move, and $w = s_3 - s_2$ is player 1's second (and final) move. The payoff (loss) to player 1 is

$$L_1 = q(s_3 - x)^2 + u^2 + w^2 \tag{2.1}$$

and the payoff to player 2 is

$$L_2 = r(s_3 - y)^2 + v^2. \tag{2.2}$$

DeGroot and Kadane examine two version of this game. In both versions, s_0, q, and r are quantities known to both players. However, in the first version each player knows both targets (so that x and y are common knowledge), while in the second version of the game each player knows only his or her respective target and is uncertain about the opponent's target. In both versions of the game the players are (expected) utility maximizers, and each player models the opponent in that way too. Thus, in the language of Pearce (1984) and Bernheim (1984), the players construct rationalizable strategies.

In the simple version, where the targets are common knowledge, DeGroot and Kadane show that optimal play, as identified by backward induction, yields the following strategies:

$$w = \frac{q(x - s_2)}{q+1} \tag{2.3}$$

$$v = (1-k)(m - s_1), \tag{2.4}$$

and

$$u = \frac{(q+1)(x - s_0) + r(x - y)}{(q+1)[1 + (q+1)/qk^2]}, \tag{2.5}$$

where

$$m = (q+1)y - qx \quad \text{and} \quad k = \frac{(q+1)^2}{r + (q+1)^2}.$$

Evidently, these strategies are in equilibrium; that is, each constitutes the best reply if the opponent's play is as specified above. Because the

solutions are unique, these strategies make for a "strong equilibrium" in Harsanyi's (1977, p. 104) sense. Moreover, the game has no maximin value for either player. That is, given a proposed value $V < 0$, there are "silly" moves the opponent is permitted to make that force the player's loss to exceed V.

Because the players cannot cooperate in this game (there are no binding agreements), it is not surprising that optimal play does not yield Pareto efficiency. For example, as noted by DeGroot and Kadane (1983, Thm. 2-ii, p. 202), if $s_0 < x < y$ then player 1's first move u is negative – *player 1 moves the pointer away from both targets* – if and only if $x - s_0/(y - x) < r/(1 + q)$. Such a move is clearly Pareto inefficient, as it leads to increased losses for each player compared to what they can achieve subject to binding agreements.

III. AN OBJECTION BY BICCHIERI

In response to a challenge raised by Bicchieri (1989) concerning the legitimacy of backward induction, we note that these strategies are optimal under the assumptions of the model for the game. As we understand Bicchieri's worry, it is that hypothetical reasoning used with backward induction does not accurately reflect how players would react were the hypothetical conditions realized. In terms of the simple DeGroot–Kadane game, we believe her objection takes the following form.

The optimality of player 2's move, given by (2.4), depends upon the assumption that player 1 makes the final move according to (2.3). (That choice of w is determined, according to backward induction, by the fact that – regardless of what has happened on the first two moves – player 1 minimizes L_1 by adhering to (2.3).) Then, regardless of player 1's initial move, player 2 does best by conforming to (2.4). However, according to (our interpretation of) Bicchieri's objection, were player 2 to observe that the initial move by player 1 fails to satisfy (2.5), then the assumption that player 1 will satisfy (2.3) also becomes questionable; hence the backward induction reasoning leading to (2.4) is undone by player 2's observation of the initial move by player 1. To conclude (our account of) Bicchieri's analysis: The hypothetical reasoning that (2.4) is best for player 2, *regardless* of what player 1 chooses as an initial move, is not correct; hence, backward induction is not valid.

Our response to this objection is that the backward induction reasoning, as illustrated by the argument that player 2's move should agree with (2.4), does not indicate what player 2 should do if the model fails, as would be indicated by player 1 failing to satisfy (2.5). Rather, backward induction is used by DeGroot and Kadane merely as an algorithm for determining what is optimal *under the model* for their game. Backward induction employs hypothetical reasoning of the form: "If player 1 has moved to s_1 and this accords with the model (i.e., if that is best for player 1), then what is player 2's best move?" Of course, we discover that player 1's final move accords with the model if and only if it satisfies (2.3), player 2's move accords with the model if and only if it satisfies (2.4), and player 1's first move accords with the model if and only if it satisfies (2.5). But we can use backward induction (and hypothetical reasoning) to discover this without requiring the model be consistent with each (physically) possible move that the players are capable of making.

What is "the model" for the game? In the DeGroot–Kadane game it is the combined assumptions that the players know the initial position s_0, targets (x, y), and the permissible moves, they know the loss functions (2.1) and (2.2), and they accept that each player is concerned to minimize his or her loss and knows this of the other. Under the DeGroot–Kadane model, the solution (2.3)–(2.5) is optimal; that is, it maximizes each player's utility and backward induction correctly identifies the solution. Reasoning by backward induction to arrive at the DeGroot–Kadane solution is not to be confused with reasoning how (for example) player 2 would play when the model fails.

Let us illustrate the difference. Suppose, contrary to the DeGroot–Kadane model, that by some action (internal or external to the game) player 1 can cause player 2 to believe that the final move might not conform to (2.3). Under this alternative model, optimal play for player 1 can differ from that prescribed by (2.5).

Consider the game with $q = r = 1$, $s_0 = 0$, $x = 1$, and $y = 2$. Optimal play (under the DeGroot–Kadane model) yields choices $u = 4/33$, $v = 19/33$, and $w = 5/33$, with a payoff (loss) to player 1 of $L_1 = 2/33$ (≈ 0.06) and to player 2 of $L_2 \approx 1.66$. Suppose, in fact, that player 2 adopts the DeGroot–Kadane model for the game and that player 1 knows it. What would result if the first player were to depart from the strategy (2.5), $u = 4/33$, and instead make the surprising move $s_1 = s_0$? What would happen if player 1 chose not to move the pointer on the first round ($u = 0$)?

With (2.5) failed, it would establish conclusively for player 2 that the

hypothesized DeGroot–Kadane model is false. How would player 2 react? Might not player 2 come to doubt that player 1 will maximize in choosing w? Might not player 2 come to think player 1 will again refuse to move the pointer and choose $s_3 = s_2$ (corresponding to $w = 0$)? If that is player 2's reaction to the surprise of $u = 0$ (i.e., if then player 2 predicts the choice $w = 0$), the very best player 2 can do is to make the move $v = 1$ and earn the loss $L_2 = 2$. But that yields player 1 the best possible score, $L_1 = 0$, by choosing $w = 0$ – ironically, just as player 2 predicts. Can player 1 anticipate player 2's reaction to the initial move ($u = 0$) and fool player 2 into this false model for the choice of w (which coincidentally happens to yield a correct prediction for the third move if player 2 chooses $v = 1$)? Such a reaction by player 2 improves player 1's payoff over the strategy (2.3)–(2.5). Does this hypothetical reasoning refute the backward induction argument leading to the strategies (2.3)–(2.5)?

We do not think the question of how player 2 would respond to a failure of his model for the game is to be answered by the logic of decisions. How player 2 would react to the move $u = 0$ is instead an empirical matter. Our point here is simple: A correct interpretation of the backward induction argument is to see it as reasoning used to identify optimal play *under the conditions of a model for the game*. Backward induction does not include (or require) the counterfactual reasoning that is needed when a player's model of the game is falsified, so the strategy (2.3)–(2.5) is optimal under the conditions of the DeGroot–Kadane model. In Section IV we illustrate how a small change in the DeGroot–Kadane game leads to a model of rational play consistent with all possible observations.

IV. TREMBLING HANDS IN THE SIMPLE DEGROOT–KADANE GAME

John Harsanyi and Reinhard Selten question the adequacy of Nash's equilibrium concept when applied to the normal-form version of an extensive-form game. They deny the equivalence of normal and extensive game forms. Instead, they advocate a refined equilibrium concept for extensive-form games, based on a "trembling-hand" model of choice. (But even such refined equilibria are subject to criticism, as illustrated by Pearce 1984, p. 1044.) An equilibrium for extensive forms is acceptable, according to their account, provided it is robust over small perturbations in choice. Specifically, in order to avoid "imperfect

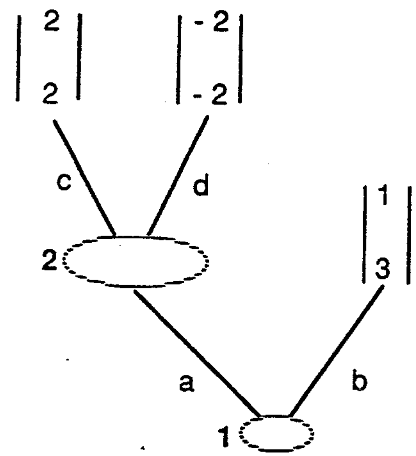

Figure 1

	C	D
A	2, 2	-2, -2
B	1, 3	1, 3

Figure 2

equilibria," they alter the basic moves in a game so that an agent selects one from a set of distributions (on pure options); a player chooses a mixed strategy rather than a pure option.

One of their examples beautifully illustrates the difference between the two kinds of equilibria. Figures 1 and 2 report (respectively) the extensive and normal forms of their game. Figure 3 gives the normal form for the perturbed game, where players may choose one of two mixed strategies in a perturbed extensive-form game (not pictured).

	C *	D *
A *	$2-5\varepsilon+4\varepsilon^2$, $2+3\varepsilon+4\varepsilon^2$	$-2+7\varepsilon-4\varepsilon^2$, $-2+9\varepsilon-4\varepsilon^2$
B *	$1+\varepsilon-4\varepsilon^2$, $3-\varepsilon-4\varepsilon^2$	$1-3\varepsilon+4\varepsilon^2$, $3-5\varepsilon+4\varepsilon^2$

Figure 3

In their game, each player has two pure strategies. In the extensive form, strategy $\{a, b\}$ for player 1 and – provided 1's information set is reached (provided player 1 chooses a) – strategy $\{c, d\}$ for player 2; in the corresponding normal form, $\{A, B\}$ for player 1 and $\{C, D\}$ for player 2. (Note that the normal form fails to distinguish between the extensive form of Figure 1 and a different game where both play simultaneously, i.e., where player 2's information set does not reflect whether player 1 chooses a or b.) In the perturbed game, the normal-form options given in Figure 3 arise by using a two-point distribution, with probabilities $(1 - \varepsilon)$ and ε assigned to each pure option in the corresponding perturbed extensive form.

Observe that, corresponding to the normal-form Figure 2, there are two equilibria: the pairs $\{A, C\}$ and $\{B, D\}$. However, the latter is "imperfect" in the extensive form of Figure 1, as that requires player 2 to (threaten to) play option d in case choice node **2** is reached. Of course, at node **2**, player 2 maximizes by playing option c instead, and player 1 knows this fact.

In the perturbed versions of the game, this difference between the two solution pairs (which are in equilibrium in the game form of Figure 2) is made evident. In the normal form of Figure 3, only the pair $\{A^*, C^*\}$ is in equilibrium. The $\{B^*, D^*\}$ pair is not in equilibrium because, when player 1 chooses B^*, player 2 improves 2's (expected) payoff by shifting from D^* to C^*; that is, D^* is not player 2's best response to B^*.

The Harsanyi–Selten idea is that imperfect equilibria are deficient because, in extensive game forms, they require a player to choose an outcome that fails to maximize utility. Nonetheless, the suspect choice is justified by Nash's criterion of equilibrium in the corresponding normal form – and it can be viewed as "threat" in the extensive form.

We agree with the Harsanyi–Selten objection to such imperfect equilibria. In the extensive form of their game, player 2 does not maximize utility by choosing option d (if node **2** arises) – d is an idle threat. That move is inconsistent with the assumption that the players are utility maximizers. In contrast, the trembling-hand model of choice eliminates the imperfect equilibrium from the normal form of the game. When the choices are mixed options, the imperfect equilibria fail to be Nash equilibria.

We may incorporate trembling hands in the DeGroot–Kadane game by limiting players to distributions for the location of the pointer, rather than supposing that player moves fix the pointer location exactly. Suppose a player moves by determining the mean of the distribution for the pointer, and suppose that distribution has a fixed and finite variance. Thus, a player may aim as follows:

the player may fix $E_k(s) = k$
by aiming the pointer at location k; (4.1)

and

the distribution has known, finite variance,
$E_k\left[(s-k)^2\right] = c \quad (0 \le c < \infty)$,
where c does not depend upon k but may reflect the stage of the game and past locations of the pointer. (4.2)

The DeGroot–Kadane game (of Section II) is a special case, where $c = 0$. The loss functions for the modified game are again given by (2.1) and (2.2), as formulated in terms of the successive, observed pointer locations. We do this in order to preserve complete information in the game. (We understand the sequential game to have *complete information* if the payoffs to each player are a known function of the public outcomes, outcomes that both players observe. This condition does not require that the players' *choices* be known to both, as the following example illustrates.) The feature of complete information would be lost if, instead, losses were defined through the unobserved aiming points.

Aside. The class of distributions specified by (4.1) and (4.2) is more general than Harsanyi–Selten's version of trembling hands. In our version, by contrast, (i) we do not require either that there be a point mass concentrated on a single pure strategy – there does not have to be probability mass assigned to point k; (ii) nor do we suppose that errors are symmetrically distributed across the alternative pure strategies.

Under this modification of the simple version of the DeGroot–Kadane game, backward induction leads to the same choice of aiming points as in the degenerate case, where $c = 0$.

Theorem. *For the (modified) simple version of the DeGroot–Kadane game, with moves specified by* (4.1) *and* (4.2), *the optimal aiming points are given by* (2.3)–(2.5), *which is the same solution as in the original game where* $c = 0$.

Proof. The reason for the theorem is the well-known fact (about point estimates) that the mean of a distribution minimizes expected squared error.

We illustrate the calculation for player 1's final move, on the third round of the game, given the observations of s_0, s_1, and s_2 (respectively, the initial location of the pointer, its location after player 1's first move, and after player 2's move). The prospective loss to player 1 after the final move, given by (2.1), is $L_1 = q(s_3 - x)^2 + u^2 + w^2$. Player 1's selection of moves is the choice of where to aim s_3 $(=s_2 + w)$, where $E_k[s_3] = s_2 + E_k[w] = k$. Thus, the expected loss in choosing $E_k(s_3) = k$ is

$$E_k(L_1) = E_k\left[q(s_2 + w - x)^2 + u^2 + w^2\right]$$
$$= q(s_2 - x)^2 + 2q(s_2 - x)E_k(w) + (q+1)E_k[w^2] + u^2$$
$$= q(s_2 - x)^2 + 2q(s_2 - x)(k - s_2) + (q+1)\left(c + [k - s_2]^2\right) + u^2.$$

Solving the equation $0 = dE_k[L_1]/dk$ (to minimize expected loss) yields

$$w = (k - s_2) = \frac{q(x - s_2)}{q+1},$$

as required by (2.3). Equations (2.4) and (2.5) follow in a similar fashion. □

Thus, the DeGroot–Kadane solution (2.3)–(2.5) is also the refined, perfect equilibrium solution advocated by the Harsanyi–Selten theory. (This result obtains because (2.3)–(2.5) are, trivially, limit points of the solutions generated by letting $c \to 0$.) There is, nonetheless, an interesting difference between the two forms of the (perfect information) DeGroot–Kadane game. In the original version, with $c = 0$, the model for the game is consistent with exactly one line of play, as dictated by (2.3)–(2.5). But, in the modified form of the simple game, if the error distribution has full support on the real line (e.g., using a normally distributed shot aimed at its mean), then each (logically) possible combination of locations for s_1, s_2, and s_3 is consistent with the model that both players are utility maximizers. No matter what player 2 sees for the location s_1 – that is, no matter where the pointer stops after player 1's first move – that outcome is consistent with the hypothesis that player 1 chose optimally in accord with (2.5).

Hence, in the trembling-hands version of this simple game ($c > 0$), with suitable error distributions, backward induction reasoning may be used to answer hypothetical questions of the form, "What would player 2 do if u is observed equal to u_1?" In other words, with such trembling-hand moves, Bicchieri's concern is satisfied: Within the model that players are utility maximizers and know this of each other, backward induction accommodates all possible moves. No possible outcome leads to a counterfactual situation.

Of course, some outcomes will be surprising though consistent under the model. Even so, we do not require of the players that they retain their belief in the model, regardless of the observed outcomes. Again, we advocate the strategies (2.3)–(2.5) as optimal under the model. Our view on this matter is no different from our view regarding statistical models generally. Sometimes observations force reevaluation of the statistical model; other times, as with "outliers," it is reasonable to do so even though the data are (formally) consistent with the model.

V. VERSION 2 OF THE DEGROOT–KADANE GAME: TARGETS ARE NOT COMMON KNOWLEDGE

In the first versions of the DeGroot–Kadane game (with and without "trembles"), optimal play according to expected utility theory leads to strategies that are in perfect equilibrium. That consequence does not obtain when the game is modified and a target point is known *only* to the player for whom it is the target – where player 1 knows x but not y and player 2 knows y but not x, and this information difference itself is common knowledge. Let us rehearse the DeGroot–Kadane solutions

to the second game (where $c = 0$) to see why optimal play will not form a Nash equilibrium.

According to backward induction, at the final move with u and v given, in order to minimize loss L_1, player 1 takes no interest in player 2's target y. Hence player 1's last move w is determined once again by (2.3):

$$w = \frac{q(x - s_2)}{q+1}. \tag{5.1}$$

Player 2 does not know x, but player 2 knows that player 1 will choose w to minimize L_1. Thus the argument leading to (2.4) does not apply directly. However, as an expected utility maximizer, player 2 has a personal probability for x (given the datum u) with mean $E_2(x|u)$. In light of the squared-error form of the loss function (2.2) and knowing that player 1 will choose w according to (5.1), player 2 minimizes expected loss by choosing

$$v = (1-k)[M(u) - s_1], \tag{5.2}$$

where $M(u) = (q + 1)y - qE_2(x|u)$.

How shall player 1 determine the initial move u? Player 1 knows neither y nor the quantity $E_2(x|u)$. But, knowing that player 2 solves (5.2) to find v, player 1 establishes an optimal choice for u in terms of the personal joint probability for these two quantities: y and $E_2(x|u)$. The resulting optimization is given by solving the equation

$$0 = \frac{1}{2}\frac{d}{du}E_1\{K^2(u)\} + \frac{u(q+1)}{q}, \tag{5.3}$$

where $K(u) = (s_0 + u)k - x + (1 - k)M(u)$.

Although these moves are the best responses to what a player believes about the opponent's moves – that is, they maximize the subjective expected utilities of each player (under the common model that they are subjective utility maximizers) – these strategies do not form a Nash equilibrium. The strategies are not in equilibrium for the simple reason that the model does not include the targets as common knowledge. The model does not result in players knowing what the other will do, even under optimal play.

For example, if at stage **2** of the game player 1's rule for choosing w (formula (5.3)) were made known to player 2 by revealing the target x, this would alter player 2's belief set – unless player 2 thinks u reveals where x is or that x is already known, in which case $\text{Var}_2[x|u] = 0$ – with the result that player 2's move would change to agree with (2.4). Likewise, if player 1 learns both 2's target y and that player 2 learns x prior to 2's best move at stage **2** of the game, then w is selected according to (2.5), not according to (5.3).

In short, exposing details of the opponent's strategy, as the Nash condition requires for ascertaining whether a replay is also a "best response," radically changes the epistemic conditions of the game. The common-knowledge assumptions leading to strategies (5.1)–(5.3) are not consistent with players verifying that theirs is a best response. The epistemic change required for satisfying the Nash condition is inconsistent with the model for the second version of the game.

This feature of our analysis is not affected by the use of trembling-hand moves. That is, if (as in Section IV) a player moves by choosing an aiming point rather than by fixing the quantity u, v, or w for certain, then optimal play does not form a Nash equilibrium, just as optimal play according to (5.1)–(5.3) does not result in Nash equilibrium. This is shown by the following.

Theorem. *For the (modified) second version of the DeGroot–Kadane game – where target points are not common knowledge – with trembling-hand moves specified by* (4.1) *and* (4.2), *the optimal aiming points are again given by* (5.1)–(5.3).

Proof. As before, the solution arises because the mean squared error of an estimate is minimized at the mean. In particular, player 1's final move (the choice of where to aim s_3, given u and v) is optimized by aiming w so that $0 = dE_k[L_1]/dk$; hence w satisfies (2.3). Likewise, at stage **2**, player 2 knows how the first player will aim the last shot, though player 2 may remain uncertain of player 1's target x. Nonetheless, player 2 has a personal probability for x, given u, whose mean $E_2(x|u)$ enters the optimization just as in the previous version of the game (the version without trembling hands). Because the mean minimizes the expected loss, player 2 chooses the aiming point for v according to (5.2). Equation (5.3), governing the first aiming point, is obtained in the same fashion. □

To repeat the point of this exercise, under a model for the DeGroot–Kadane game where players have common knowledge that they are expected utility maximizers but where they lack common knowledge of their targets, and where moves are subject to trembles, optimal play does not result in a Nash equilibrium.

VI. AUMANN'S CORRELATED EQUILIBRIUM AND THE DEGROOT–KADANE GAME

Aumann (1987) proposes an original unification of the game-theoretic and Bayesian decision-theoretic viewpoints. He identifies the game-theoretic perspective with a generalized account of (Nash) equilibrium, leading to what Aumann terms *correlated equilibrium*. These are best-response strategies that may rely on correlated (rather than independent) joint distributions to form mixed options. That is, the distribution used by player 1 to create a mixed strategy can be correlated with the distribution used by player 2. Aumann's account of Bayesian rationality in games leads to the result that Bayes-rational players will adopt strategies that are in correlated equilibrium. Moreover, each correlated equilibrium can be a model (with specific informational constraints on the individual players) for Bayes-rational play. Hence, there is a reconciliation of the two viewpoints.

We agree with Aumann that Bayesian decision theory should apply to games; the logic of choice is the same whether our uncertainty is about "Nature" or an opponent's moves (see Kadane and Larkey 1982). However, we take issue with (what we understand to be) Aumann's formulation of Bayes rationality in games. He requires a very rich space Ω of states of the world:

> The term "state of the world" implies a definite specification of all parameters that may be the object of uncertainty on the part of any player of [the game] G. In particular, each ω includes a specification of which action is chosen by each player of G at the state ω. Conditional on a given ω, everybody knows everything; but in general, nobody knows which is really the true ω. (1987, p. 6).

Though agents are permitted private information about Ω, Aumann requires that (each) player i's personal probability (over Ω), here denoted by $p_i(\Omega|D_i)$, is a conditional probability that arises from a common prior: $p_i(\Omega) = p(\Omega)$ given i's (perhaps) private data $d_i \in D_i$; however, the prior is the same for each player i. That is, apart from private evidence, the players are required to have the same opinions about the set of states Ω. Because (by the severe assumption that) each stage ω specifies "all parameters that may be the object of uncertainty on the part of any player," Aumann argues that the information sets D_i (though not the private information d_i) also are common knowledge to all players.

We object to Aumann's condition that there be a common prior (across players) in games. He recognizes this challenge in Section 5 of his paper. Concerning ordinary decisions, we believe there is no basis within (say) Savage's decision theory for that assumption, regardless of the detail with which states (of Nature) are defined. Savage's opposition to what he called "necessary" Bayesian theory (Savage 1954, p. 61; 1962, p. 102; 1967) leads us to think he rejected a common-prior requirement even in the structured setting of parametric statistical inference, where likelihoods are specified, a fortiori in less structured game settings where likelihoods are not so determined.

Our second concern is with consequences of demanding that Ω be as detailed as Aumann proposes. In particular, we are uncomfortable with the prospect that agents are required to hold (nontrivial) probabilities over their own current choices. (Again, we observe that Savage's theory is carefully formulated to distinguish between acts and states; states but not acts are assigned personal probability.) We do not see a problem when an agent assigns personal probabilities (more accurately, personal conditional probability) now to future choices, because the agent cannot now make those states true or false. Nor do we find a conceptual problem in assigning a personal probability to past choices, since the agent may have forgotten those past choices. The difficulty with personal probability over one's current choices is that such probabilities do not support the familiar betting-odds interpretation. (See Spohn 1977, Kadane 1985, and Levi 1989 for related discussions.)

The second version of the DeGroot–Kadane game serves to illustrate our position on this issue. Recall that in the second version, players know their respective targets but are uncertain of the other's target, and this informational structure is itself common knowledge. Recall also that it is the uncertainty about the opponent's target that alone differentiates the two versions of the game. In the first version of the game, when the informational structure of the game includes common knowledge of the targets, the optimal strategies (2.3)–(2.5) are common knowledge, too; there is no uncertainty for either player about what he or she will do.

We propose, therefore, to analyze the second version of the game (without trembles) using pairs of targets for the states $\Omega' = \{(x, y)$: x is

player 1's target, y is player 2's target$\}$. We introduce a common prior $p(\bullet)$ over these states to allow a comparison with Aumann's theory. As we make clear shortly, the set Ω' is not Aumann's set of states Ω for this game. (Also, to agree with Aumann's presentation, we are prepared to use the game's normal form. That is, we see the selection of "states" as the relevant issue here, not the collapse of extensive to normal form.)

Suppose the two players begin their analysis with a common prior over Ω'; that is, they do not yet know their targets, yet they share the following background information: It is given that both players are utility maximizers, that their respective loss functions are L_1 and L_2, that the initial pointer location is s_0, and that all this is common knowledge. For simplicity, before learning their targets, assume the players have a (common) joint distribution $p(x, y)$ which is bivariate normal (μ, Σ), with known means $\mu = (x_0, y_0)$, with known and equal variances $(\sigma_x = \sigma_y = \sigma)$, and with (x, y) independent $(\rho_{xy} = 0)$.

Then, after player 1 learns x, 1's probability for player 2's target, $p_1(y|x) = p_1(y)$, is normal $N(y_0, \sigma^2)$, since x and y are independent. Likewise, after learning y, player 2 has uncertainty about x, denoted by $p_2(x|y) = p_2(x)$, which is normal $N(x_0, \sigma^2)$. These distributions are common knowledge. In particular, prior to any moves, player 1 knows that player 2's expected value for x is x_0; that is, player 1 knows $E_2(x|y) = E_2(x) = x_0$ and player 2 knows $E_1(y|x) = E_1(y) = y_0$.

Despite the common prior, this common knowledge does not induce a correlated equilibrium with respect to Ω'. That the addition of a common prior for Ω', even one that makes (x, y) independent variables, fails to yield a correlated equilibrium is explained by tracing the impact of the common prior on the solutions (5.1)–(5.3). With respect to player 1's choice of a final move w, the prior $p(x, y)$ is irrelevant because x is known and y plays no role in minimizing L_1 through the choice of s_3. At the second move, when player 2 is contemplating the choice of v, what is relevant is the quantity $E_2(x|u, y)$. But the common prior $p(x, y)$ does not determine this expectation! It fails to do so since it leaves open what might be player 2's beliefs about player 1's choice of u. That is, all of $p_2(u)$, $p_2(u|y)$, and $p_2(u|x, y)$ are underdetermined by the common prior on Ω'.

For instance, both players know that

$$p_2(x|u, y) = p_2(u|x, y) \cdot p_2(x|y)/p_2(u|y).$$

Also, it is common knowledge that $p_2(x|y) = p_2(x)$, where x is a normal $N(x_0, \sigma^2)$ distribution. But the common prior in (x, y) does not fix the ratio $p_2(u|x, y)/p_2(u|y)$, which is known to player 2 only. The terms $p_2(u|x, y)$ and $p_2(u|y)$ cannot be derived using Bayes's theorem merely by giving player 2 privileged information about (x, y). Specifically, the probability $p_2(u|x, y)$ should not be confused with the (point mass) solution for u, given by (2.5) from the first version of the DeGroot–Kadane game (where both targets are common knowledge and player 2 knows u for certain). It is important to correctly interpret the compound conditioning event in $p_2(u|x, y)$. That conditioning event specifies both targets, but it leaves x known to player 1 only. It is important to distinguish two cases:

1. Conditioning on the event (x, y) when these are common knowledge, as in the first version of the DeGroot–Kadane game, leading to (2.3)–(2.5).
2. Conditioning on the event (x, y), when target x is known to player 1 alone and y is known to player 2 alone.

When (x, y) are not common knowledge, as in the second version of the game, it is the second of these two cases that the players face when evaluating the term $p_2(u|x, y)$. For some discussion on the range of values $p_2(u|x, y)$ can take (all of which are unknown to player 1), see Corollaries 1 and 2 in DeGroot and Kadane (1983, p. 206).

Thus, we see the impact of Aumann's selection of fine-grained states Ω on his result equating Bayes rationality (subject to a common prior) with correlated equilibria. For Aumann's theorem to apply, the agents must include player 1's choice of u, as well as the targets x and y, in the states of uncertainty. Then, with a common prior over the refined states $\Omega = \{(u, x, y)\}$, the problematic term $E_2(x|u, y)$ becomes common knowledge. However, to demand a common prior over Ω mandates two conditions that we find unwarranted for rational play in this game. Aumann's analysis mandates: (i) that player 2's beliefs about player 1's choice u are transparent to player 1; and (ii) that player 1 holds nontrivial probabilities about 1's own actions. What is the basis for demanding condition (i)? What is the interpretation, from player 1's perspective, of assigning (nontrivial) probabilities to the choice u?

VII. ON RUBINSTEIN'S THESIS: ALMOST COMMON KNOWLEDGE IS NOT GOOD ENOUGH

The preceding section explored a consequence of imposing a common prior distribution $p(x, y)$ on the set of target states Ω' for the second version of the DeGroot–Kadane game. The upshot of that analysis is that a common prior on Ω' is insufficient for defining player 2's choice of move v, since it leaves open player 2's conditional distribution for player 1's move u, given targets x and y. Thus, the common prior on Ω' also leaves open player 1's first move u, since that depends upon player 1's expectation of player 2's expectation of u, and so forth.

This argument is valid for each value $\sigma^2 > 0$. (Recall that σ^2 is the common variance for the targets.) However, the first version of the game, with targets (x_0, y_0) common knowledge, corresponds to the limiting distribution $\sigma^2 = 0$. Thus, the first version of the DeGroot–Kadane game is not necessarily the limit of the second-version games with common priors, where $\sigma^2 \to 0$. That is, the optimal strategy (2.3)–(2.5) for the first game (where targets are common knowledge) may not equal the limit (as $\sigma^2 \to 0$) of optimal strategies for the second version of the game, constrained by a common prior.

Let us illustrate the point. To simplify the formulas, take $q = r = 1$ and $s_0 = 0$. With the variance $\sigma^2 > 0$ given, denote with subscripts u_σ and v_σ the choices for the first two moves. And, with some slight abuse of notation, use the subscripted u_0 and v_0 to denote the limit of these moves as $\sigma \to 0$. Suppose player 2 reasons as follows.

In the first version of the DeGroot–Kadane game, with targets common knowledge, according to (2.5) my opponent's first move u is linear in the targets x and y. That is, were our targets known, player 1 would choose

$$u = (3x - y)/8.25. \tag{7.1}$$

So, I'll take my expectation for x to be linear in u_σ and y:

$$E_2(x|u_\sigma, y) = a_\sigma + b_\sigma u_\sigma + c_\sigma y. \tag{7.2}$$

Then my move v_σ satisfies

$$v_\sigma = 0.2[2y - a_\sigma - (1 + b_\sigma)u_\sigma - c_\sigma y]. \tag{7.3}$$

The move v_σ, (7.3), contrasts with player 2's choice (from 2.4) of

$$v = 0.2[2y - u - x] \tag{7.4}$$

for the case where targets are common knowledge. Recall that x and y are uncorrelated. Therefore, in order to make (7.3) equal (7.4) as $\sigma \to 0$ (i.e., for $v_0 = v$), it is necessary and also sufficient that $a_\sigma \to 0$, $b_\sigma \to 1$, $c_\sigma \to 0$, and $u_\sigma \to x$.

Now, in case player 1 knows that player 2 has the linear expectation (7.2) (without in general knowing the coefficients a_σ, b_σ, and c_σ), DeGroot and Kadane (1983, p. 206) have shown that player 1's optimal choice of u_σ, given x, satisfies

$$u_\sigma = \frac{E_1\{(0.2b_\sigma - 0.8)\cdot(0.2[(2-c_\sigma)y - a_\sigma] - x)\}}{2 + E_1\{(0.2b_\sigma - 0.8)^2\}}. \tag{7.5}$$

For the limiting values of the coefficients necessary to make (7.3) and (7.4) agree, this yields

$$u_0 = E_1\{(0.6x - 0.24y)/2.36\}. \tag{7.6}$$

However, $u_0 \neq u$, (7.6) does not agree with (7.1) (which is player 1's choice for u when targets are common knowledge), and neither does $u_0 = x$, as is necessary for (7.3) to agree with (7.4).

In short, the limit of optimal play (with $\sigma \to 0$) here does not correspond to the optimal play at $\sigma = 0$. The singularity (at $\sigma^2 = 0$) occurs because merely shrinking the variance ($\sigma^2 \to 0$) of the prior distributions for the targets does not suffice also for shrinking player 2's conditional probability $p_2(u|x, y)$ to the point mass for u concentrated at the solution (2.5). It fails to do so because, in part, the correct interpretation of this conditional probability in the second version of our game is not to be confused with the first-version interpretation, which corresponds to common knowledge of the targets.

Though the limit ($\sigma^2 \to 0$) of the common priors is common knowledge of the targets, the limit of the optimal strategies based on these common priors need not be the optimal strategy based on common knowledge of the targets. Rubinstein is correct: Almost common knowledge is not good enough!

Remark. By supplying the two players with additional, common evidence about the targets (x, y), we can implement the dynamics of a common "posterior" with shrinking variance. For example, if both players observe n pairs (x_i', y_i') $(i = 1, \ldots, n)$ of i.i.d. bivariate normal variates, with (unknown) means ($\mu_{x'} = x$, $\mu_{y'} = y$), known (equal) variances σ'^2, and zero correlation ($\rho_{x'y'} = 0$), then their common posterior distribution for the targets will be as independent bivariate normal

variates with a (common) variance that shrinks to 0 as the sample size n grows without bound.

VIII. CONCLUSION

We have used a relatively simple sequential game between two utility maximizers to emphasize that optimal play among Bayesians (who model each other as such) does not put their strategies into equilibrium. Even with a common prior over the uncertain components of the game (which itself is common knowledge), that is, even with a common prior over the target points, optimal play does not require a correlated equilibrium. The optimal extensive-form strategies are rationalizable in the sense of Pearce (1984), as the reasons for (5.1)–(5.3) make clear. That is, those strategies are derived by backward induction using the common knowledge that the opponents are utility maximizers.

It is right to develop Bayesian game theory. A decision against an opponent, rather than against Nature, does not require novel principles. However, especially in sequential games, the challenge of doing Bayesian game theory against a Bayesian opponent is considerable. It has many facets: Not only must players represent their uncertainties about ordinary events (which, in the DeGroot–Kadane game, corresponds to the players' beliefs about each other's target), but each player must be prepared to formalize how his own actions will affect the other's subsequent choices. In order to do that while respecting the model of common knowledge (where each player is an expected utility maximizer), each must think about how the other models himself. That is, I must ponder what the opponent believes about my beliefs, and so on. The complexity of this thought, the depth to which each player must evaluate iterated expectations of beliefs about the other in order to apply backward induction, depends upon the number of turns in the game. Already, in the simple three-move game of this chapter (without common knowledge of the targets), that task is not trivial for player 2.

The subtleties that attend the difference between "common knowledge" and "almost common knowledge" hint at the number of different faces of Bayesian game theory. In this essay we have focused on one game where it is common knowledge that players are rational. Not all games have that form, even when the players are, in fact, all rational. Perhaps this is a direction to look in to gain a better understanding of such tactics as bluffs and feints. We trust the challenges of Bayesian game theory will be met: some through analysis and some through empirical enquiry.

REFERENCES

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

Bernheim, B. D. (1984), "Rationalizable Strategic Behavior." *Econometrica* 52: 1007–28.

Bicchieri, C. (1989), "Backward Induction without Common Knowledge." In A. Fine and J. Leplin (eds.), *PSA 1988*, vol. 2. East Lansing, MI: Philosophy of Science Association. Presented at the workshop on "Knowledge, Belief, and Strategic Interaction" (June 1989). Castiglioncello, Italy.

Binmore, K., and Brandeburger, A. (1988), "Common Knowledge and Game Theory." Technical Report 89-06, Department of Economics, University of Michigan, Ann Arbor.

DeGroot, M. H., and Kadane, J. B. (1983), "Optimal Sequential Decisions in Problems Involving More Than One Decision Maker." In Rizvi, Rustagi, and Siegmund (eds.), *Recent Advances in Statistics*. New York: Academic Press, pp. 197–210.

Harsanyi, J. C. (1977), *Rational Behavior and Bargaining Equilibrium in Games and Social Situations.* Cambridge: Cambridge University Press.

Harsanyi, J. C. (1989), "Game Solutions and the Normal Form." Presented at the workshop on "Knowledge, Belief, and Strategic Interaction" (June 1989). Castiglioncello, Italy.

Harsanyi, J. C., and Selten, R. (1988), *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: MIT Press.

Kadane, J. B. (1985), "Opposition of Interest in Subjective Bayesian Theory." *Management Science* 31: 1586–8.

Kadane, J. B., and Larkey, P. D. (1982), "Subjective Probability and the Theory of Games." *Management Science* 28: 113–20. [Chapter 3.1, this volume]

Levi, I. (1989), "Feasibility." Department of Philosophy, Columbia University, New York.

Pearce, D. G. (1984), "Rationalizable Strategic Behavior and the Problem of Perfection." *Econometrica* 52: 1029–50.

Rubinstein, A. (1989), "The Electronic Mail Game: Strategic Behavior Under 'Almost Common Knowledge'." *American Economic Review* 79: 385–91.

Savage, L. J. (1954), *The Foundations of Statistics.* New York: Wiley.

Savage, L. J. (1962), *The Foundations of Statistical Inference: A Discussion.* London: Methuen.

Savage, L. J. (1967), "Implications of Personal Probability for Induction." *Journal of Philosophy* 64: 593–607.

Spohn, W. (1977), "Where Luce and Krantz Do Really Generalize Savage's Decision Model." *Erkenntnis* 11: 113–34.