2 Dynamic Deliberation: Equilibria

> A static theory deals with equilibria. The essential characteristic of an equilibrium is that it has no tendency to change, i.e. that it is not conducive to dynamic developments. An analysis of this feature is, of course, inconceivable without the use of certain rudimentary dynamic concepts.
>
> —John von Neumann and Oskar Morgenstern (1947, p. 45)

Deliberational Dynamics and Game Theory

Let us suppose that one deliberates by calculating expected utility. In the simplest cases, deliberation is trivial; one calculates expected utility and maximizes. But in more interesting cases, the very process of deliberation may generate information that is relevant to the evaluation of the expected utilities. Then, processing costs permitting, a Bayesian deliberator will feed back that information, modify his probabilities of states of the world, and recalculate expected utilities in light of the new knowledge.[1]

In the presence of informational feedback Bayesian deliberation becomes a dynamical process. The decisionmaker starts in a state of indecision; calculates expected utility; moves in the direction of maximum expected utility; feeds back the information generated by his move and recalculates; and so forth. In this process, his probabilities of doing the various acts evolve until, at the time of decision, his probability of doing the selected act becomes virtually one.

The theory of dynamic deliberation carries with it an *equilibrium* principle for individual decision. The decisionmaker cannot decide to do an act that is not an equilibrium of the deliberational process.[2] If he is about to choose a nonequilibrium act, deliberation carries him away from that decision. This sort of equilibrium requirement for individual decision can be seen as *a consequence of the expected utility principle*. It is usually neglected only because the process of informational feedback in deliberation is usually neglected. In cases in which there is no informational feedback, simple choice of the act with initial maximum expected utility automatically fulfills the equilibrium requirement.

One sort of situation where such informational feedback is relevant is that envisioned by von Neumann and Morgenstern in their indirect argument for the Nash equilibrium concept. Each player calculates the *prima facie* optimum act; but then there is sufficient common knowledge for each player to work out the others' calculations, feed back this information, and recalculate; but then each knows that this will have happened, works out these calculations, feeds back this information, and so forth.

We know from the discussion in Chapter 1 that sufficient common knowledge to make sense of this story must include considerably more than the common knowledge of rationality assumed by von Neumann and Morgenstern. We will see in this chapter how, under suitably strengthened assumptions of common knowledge, *a joint deliberational equilibrium on the part of all the players corresponds to a Nash equilibrium point of the game.* This is the sort of justification that von Neumann and Morgenstern desired, and it is based on the expected utility principle. Furthermore, strengthening the assumptions slightly to make the players qualitatively "more Bayesian" leads in a natural way to refinements of the Nash equilibrium. There is also an important connection with *correlated equilibria.*

Deliberational Equilibria

Let us model the deliberational situation in an abstract and fairly general way. A Bayesian has to choose between a finite number of acts: $A_1 \ldots A_n$. Calculation takes time for her, although its cost is negligible. We assume that she is certain that deliberation will end and she will choose some act (perhaps a mixed one) at that time. Her *state of indecision* will be a probability vector assigning probabilities to each of the n acts, which sum to one. These are to be interpreted as her probabilities now that she will do the act in question at the end of deliberation. A state of indecision, P, carries with it an expected utility, the expectation according to the probability vector $\mathbf{P} = \langle p_1 \ldots p_n \rangle$ of the expected utilities of the acts $A_1 \ldots A_n$. The expected utility of a state of indecision is thus computed just as that of the corresponding mixed act. Indeed, the adoption of a mixed strategy can be thought of as a way to turn the state of indecision for its constituent pure acts to stone. We will call the mixed act corresponding to a state of indecision the state's *default mixed act.*[3]

A person's state of indecision will evolve during deliberation. In the first place, on completing the calculation of expected utility, she will believe more strongly that she will ultimately do that act (or one of those acts) that are ranked more highly than her current state of indecision. If her calculation yields one act with maximum expected utility, will she not simply become sure that she will do that act? She will not *on pain of incoherence* if she believes that she is in an informational feedback situation and if she assigns any positive probability at all to the possibility that informational feedback may lead her ultimately to a different decision. (I will return to this topic, in more detail, in Chapter 4.) So, she will typically in one step of the process move in the direction of the currently perceived good, but not all the way to decision.

We assume that she moves according to some simple dynamical rule for "making up one's mind," as opposed to performing an elaborate calculation at each step. This rule should, however, be "qualitatively Bayesian" in various ways. It should reflect her knowledge that she is an expected utility maximizer and the status of her present expected utility values as her expectation of her final utility values.

For the moment, we will assume that we have a dynamical rule that *seeks the good,*[4] in the following modest sense:

1. the rule raises the probability of an act only if that act has utility greater than that of the status quo;
2. the rule raises the sum of the probabilities of all acts with utility greater than that of the status quo (if any).

All dynamical rules that seek the good have the same fixed points: those states in which the expected utility of the status quo is maximal.

As a concrete example of such a rule we can take the function that Nash (1951) used to prove the existence of equilibria for finite non-zero-sum games. Define the *covetability* of an act, A, to a person in a state of indecision, P, as the difference in expected utility between the act and the state of indecision if the act is preferable to the state of indecision, and as zero if the state of indecision is preferable to the act, or cov (A) $= \max[U(A) - U(P), 0]$. Then the Nash map takes the decisionmaker from state of indecision P to state of indecision P', where each component p_i of **P** is changed to:

$$p'_i = \frac{p_i + \text{cov}\,(A_i)}{1 + \Sigma_i \text{cov}(A_i)}$$

Here a bold revision is hedged by averaging with the status quo.

We can get a whole family of Nash maps by allowing different weights for the average:

$$p'_i = \frac{kp_i + \mathrm{cov}(A_i)}{k + \Sigma_i \mathrm{cov}(A_i)}$$

The constant k $(k > 0)$ is an index of caution. The higher k is, the more slowly the decisionmaker moves in the direction of acts that look more attractive than the status quo. In continuous time, one has the corresponding Nash flows:

$$\frac{dp(A)}{dt} = \frac{\mathrm{cov}(A) - p(A)\Sigma_j \mathrm{cov}(A_j)}{k + \Sigma_j \mathrm{cov}(A_j)}$$

The Nash rules are not the only rules which seek the good,[5] and we shall see later that a more refined Bayesian analysis may lead elsewhere. But we will use the rules as a source of easily realized concrete examples. The Appendix shows how to implement the Nash dynamics on a personal computer, and the interested reader is urged to try it out on various games.

The decisionmaker's calculation of expected utility and subsequent application of the dynamical rule constitutes new information. The new information may affect the expected utilities of the pure acts by affecting the probabilities of the states of nature, which together with the act determine the payoff. In the typical game-theoretical contexts, states of nature consist of the possible actions of the opposing players. For simplicity, we will assume here a finite number of states of nature.

The decisionmaker's *personal state* is then, for our purposes, determined by two things: her state of indecision and the probabilities that she assigns to states of nature. Her personal state space is the product space of her space of indecision and her space of states of nature. Deliberation defines a dynamics on this space. We could model the dynamics as either discrete or continuous, but for the moment we will focus on discrete dynamics. We assume a dynamical function, ϕ, which maps a personal state $\langle x,y \rangle$ into a new personal state $\langle x',y' \rangle$ in one unit of time. The dynamical function, ϕ, has two associated rules: (1) the adaptive dynamical rule,[6] D, which maps $\langle x,y \rangle$ onto x' and (2) the informational feedback process, I, which maps $\langle x,y \rangle$ onto y' [where $\langle x',y' \rangle = \phi\langle x,y \rangle$].

A personal state $\langle x,y \rangle$ is a *deliberational equilibrium* of the dynamics, ϕ, if and only if $\phi\langle x,y \rangle = \langle x,y \rangle$. If D and I are continuous, then ϕ is continuous and it follows from the Brower fixed point theorem that a deliberational equilibrium exists. Let N be the Nash dynamics for some $k > 0$. Then if the informational feedback process, I, is continuous, the dynamical function $\langle N,I \rangle$ is continuous and has a deliberational equilibrium. Then, since N *seeks the good*, for any continuous informational feedback process, I, $\langle N,I \rangle$ has a deliberational equilibrium $\langle x,y \rangle$ whose corresponding mixed act maximizes expected utility in state $\langle x,y \rangle$. This is a point from which process I does not move y and process N does not move x.

But if process N does not move x, then no other process which seeks the good will either (whether or not it is continuous). So, we have—*à la Nash*—a general existence result for deliberational equilibria:

> If D seeks the good and I is continuous, then there is a deliberational equilibrium, $\langle x,y \rangle$, for $\langle D,I \rangle$. If D' also seeks the good, then $\langle x,y \rangle$ is also a deliberational equilibrium for $\langle D',I \rangle$. The default mixed act corresponding to x maximizes expected utility at $\langle x,y \rangle$.

Games Played by Bayesian Deliberators

Suppose that two (or more) Bayesian deliberators are deliberating about what action to take in a noncooperative non-zero-sum matrix game. We assume that each player has only one choice to make, and that the choices are causally independent in that there is no way for one player's decision to *influence* the decisions of the other players. Then, from the point of view of decision theory, for each player the decisions of the other players constitute the relevant *state of the world* which, together with her decision, determines the *consequence* in accordance with the payoff matrix.

Suppose, in addition, that each player has an adaptive rule, D, which seeks the good (each one need not have the same rule) and that what kind of Bayesian deliberator each player is is common knowledge. Suppose also that each player's initial state of indecision is common knowledge, and that other players take a given player's state of indecision as their own best estimate of what that player will ultimately do. Then initially there is a probability assignment to all the acts for all the players that is shared by all the players and is common knowledge.[7]

Under these strong assumptions of common knowledge, an interesting informational feedback process becomes available. Starting from the initial position, player 1 calculates expected utility and moves by her adaptive rule to a new state of indecision. She knows that the other players are Bayesian deliberators who have just carried out a similar process, and she knows their initial states of indecision and their updating rules. So she can simply go through their calculations to see their new states of indecision and update her probabilities of their acts accordingly. We will call this sort of informational feedback process *updating by emulation*. Suppose that all the players update by emulation. Then, in this ideal case, the new state is common knowledge as well and the process can be repeated.

Since the joint state of all players is common knowledge at all times, the von Neumann–Morgenstern reasoning applies:

> In a game played by Bayesian deliberators with a common prior, an adaptive rule that seeks the good, and a feedback process that updates by emulation,[8] with common knowledge of all the foregoing, each player is at a deliberational equilibrium at a state of the system if and only if the assignment of the default mixed acts to each player constitutes a Nash equilibrium of the game.

There is an alternative interpretation of the mathematics of this model, which fits nicely with an alternative interpretation of mixed equilibria. In this interpretation, the application of the "adaptive rule" represents not a given player's rule for changing beliefs in her probabilities of her own actions, but rather the other players' shared inductive rule for modifying predictions of her action.[9] Her updating by emulation then tells her what other players have as shared probabilities for her actions. The requirement that dynamical rules "seek the good" is then a somewhat more modest and perhaps more credible version of "best-response" reasoning.

The notion of a "default mixed act" falls away if we adopt a reinterpretation of mixed equilibria suggested by Aumann.[10] He advocates a point of view in which the probabilities in a player's mixed strategy are thought of as shared probabilities of the other players' strategies.[11] Mixed equilibria are then thought of as equilibria in beliefs. If we adopt this point of view together with the reinterpretation of deliberational dynamics, we can conclude that in the situation envisioned above the players are at a joint deliberational equilibrium just in case their beliefs constitute an equilibrium in this sense.

Equilibrium Selection

It is worth taking a closer look at the way in which dynamic deliberation, in the sort of setting under consideration, deals with the problem of multiple equilibria in non-zero-sum games. Consider the following game:

Battle of the Sexes

	C1	C2
R1	2,1	0,0
R2	0,0	1,2

A woman, Row, and a man, Column, are each deciding where to go for an evening's entertainment. The woman would rather go to event 1 and the man to event 2, but each would prefer to go to the event where he or she will meet the other. (Back in the 1950s, the conventional way of telling this story made event 1 an opera and event 2 a prizefight!) This game is a nice mixture of competitive and cooperative motivations. There is an equilibrium at [R1, C1], which the woman prefers, one at [R2, C2], which is favored by the man, and a mixed equilibrium with each tossing a fair coin.

Suppose that they start deliberating with a commonly known probability of 0.8 that the woman will choose R1 and 0.6 that the man will choose C2. Notice that if each were simply to maximize expected utilities on these probabilities, without informational feedback, the woman would choose R2 and the man C1, with the result that they will both end up with a payoff of zero. Now suppose that it is common knowledge that they are Nash deliberators (with a fairly high index of caution, which is also common knowledge). Then deliberation will carry them along the orbit indicated in Figure 2.1 to [R1, C1]. Equilibrium selection is effected by the dynamics in virtue of the strong assumptions of common knowledge in force. That strong assumptions are required to overcome the difficulties of classical non-zero-sum game theory should not come as a surprise. We will be interested in subsequent chapters in investigating the results of weakening these assumptions in various ways, but in this chapter we will be interested in principled Bayesian reasons for slight modifications and strengthenings of the condition that the dynamics "seeks the good."

The Bayes Dynamics

A Bayesian updates by Bayes' rule. Thus, if the new information that a player gets by emulating other players' calculations, updating his probabilities on their actions, and recalculating his expected utilities is e, then his new probabilities that he will in the end do an act A, $p_2(A)$, in terms of his old probabilities, $p_1(A)$, should be:

$$B: \quad p_2(A) = p_1(A)\frac{p(e|A)}{\Sigma_i p(A_i)p(e|A_i)}$$

where $\{A_i\}$ is a partition of alternative acts.[12]

Our story has been that the deliberator does not have the appropriate proposition, e, in a large probability space that defines the likelihood, $p(e|A)$. Instead, our "bounded Bayesian" deliberator uses a simple dynamical rule at this stage of deliberation. But such a simple dynamical

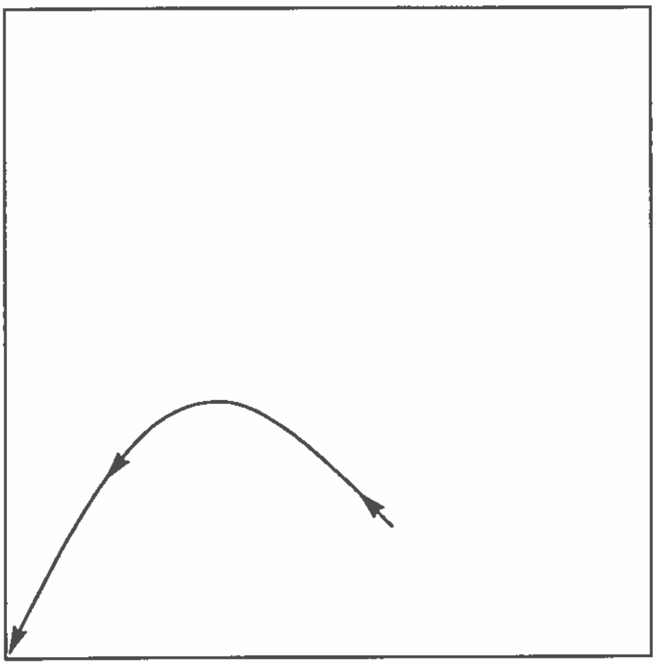

Figure 2.1. Battle of the Sexes: orbit of [0.2,0.6] goes to [0,0]

rule can still, in a sense, be more or less Bayesian according to which qualitative features of a full Bayesian analysis it retains.

For instance, one can argue that updating by Bayes' rule (on evidence with positive prior probability) cannot raise zero probabilities. In this respect, the Nash dynamics is un-Bayesian because it can and does raise zero probabilities. Thus, no hypothesis one could make about the likelihoods, $p(e_i|A)$, could embed the Nash dynamics in a larger setting in which it coincided with Bayes' rule. In a way, the Nash dynamics tries too hard. If a deliberator starts out with probability one that he will do an act that has utility less than that of the status quo, Nash dynamics will pull that probability down and raise the zero probabilities of competing acts (as will every rule which "seeks the good"). From the standpoint of adaptive behavior this effort seems laudable, but with respect to coherent updating it may leave something to be desired.[13]

Indeed, one can argue that if a deliberator is absolutely sure which act he is going to do he needn't deliberate, and if he is absolutely sure he won't do one of a set of alternative acts his deliberations should concern only the others. Putting it the other way around, if a decisionmaker thinks that there is any chance that deliberation might change his probabilities of an act, he should have given the act a probability different from zero or one. According to this account, deliberation should start at some point in the *interior* of the space of indecision, and the desideratum that a dynamic rule should seek the good should be restricted to moving probabilities that are properly movable. Notice that when deliberation originates in the interior of the space of indecision, the Nash dynamics stays in the interior so it will never get a chance to display unruly behavior with respect to probabilities of zero and one.

The Nash dynamics can lead to un-Bayesian updating in another, more subtle way when there are more than two possible acts under consideration. If two or more acts have utility less than that of the status quo, (SQ), they all get the same covetability—namely, zero—even if their expected utilities are quite different. (Remember that the $\mathrm{cov}(A) = \max[0, U(A) - U(SQ)]$.) This does not square well with the Bayesian story we give to motivate dynamic rules seeking the good on the interior of the space of indecision.

Our deliberator supposes that her calculations yielding new expected utilities give new information about what she will finally do because she believes that she will move in the direction of the apparent good until the moment of truth. It is possible—perhaps likely—that deliberation will have reached an equilibrium by the moment of truth, in which case her decision will be a best response. On the other hand, in the absence of special knowledge, it is no more likely if the moment of truth arrives before equilibrium that she will make a worse response rather than a better one. The present expected utilities just calculated may not be the ones which will obtain at the moment of truth, but they are in a sense the decisionmaker's best estimate of them.[14]

On the basis of such general considerations, the decisionmaker acts as if the likelihood, $p(e|A)$, is an increasing function of the newly calculated expected utility of A. We could call this the assumption of a *tendency toward better response.* The reasoning deserves to be looked at in more detail, and we will do so later. For the moment, I just want to point out that to the extent the assumption supports pumping up the probability of what looks best, it also supports pumping up the ratios of the probabilities of second to third place, and so on. By Bayes' theorem:

$$\frac{p_2(A)}{p_2(B)} = \frac{p_1(A)}{p_1(B)} \cdot \frac{p_1(e|A)}{p_1(e|B)}$$

If the likelihood is an increasing function of the expected utility of an act, then the Bayes dynamics will modify the probabilities of those acts with expected utility less than that of the status quo in a way that the Nash dynamics does not.

It would be nice to have a concrete example of this sort of Bayesian dynamics. The simplest way to make the likelihood an increasing function of the expected utility is to set the likelihood equal to the expected utility (with respect to some appropriate utility scale). This gives:

$$D: \qquad p_2(A) = p_1(A)\frac{U(A)}{U(SQ)}$$

where utility should be measured on a scale which is nonnegative, and positive on the interior of the space of indecision.[15]

Readers familiar with the evolutionary game theory of Maynard Smith (1982) will recognize this as a dynamics that Nature implements through the process of evolution. (The payoffs are in terms of reproductive fitness.) For this reason we will call it the Darwin map. It's nice to know that Mother Nature is a rough-and-ready Bayesian.

The continuous counterpart of the Darwin map is the Darwin flow:

$$\frac{dp(A)}{dt} = k\frac{U(A) - U(SQ)}{U(SQ)}$$

The Appendix shows how to implement the Darwin dynamics on a personal computer. One can, of course, construct all sorts of other Bayesian dynamics with more or less plausibility by setting the likelihood equal to other continuous monotonic functions of the utility.

The connection between deliberational dynamics and the Nash equilibrium concept becomes slightly more complicated if we accept these Bayesian modifications of our viewpoint. Since Bayesian dynamics seeks the good on the interior of the space of indecision, the connection between deliberational equilibria and Nash equilibria remains the same for points in the interior. But with respect to points not in the interior, we must focus on limiting behavior. If Darwin dynamics starting in the interior converges to a point, then that point corresponds to a Nash equilibrium of the game.

Refinements of the Nash Equilibrium for the Normal Form

Models of dynamic deliberation provide a setting which may make more sense of the project of refining the Nash equilibrium concept than does the metaphor of the trembling hand. If we think of perfection as being motivated by considerations of a slight probability of irrationality on the part of other players, we will have trouble making sense of the concept. If the probability of irrational play *really is zero*, why not stick with the Nash equilibrium? If it really isn't zero, we open up a Pandora's box whose contents cannot be adequately dealt with by the concept of perfection. The situation is even more paradoxical with respect to *proper* equilibria. Here the "trembles" must be considered more likely in the direction of least loss, an assumption that requires a kind of rational control of irrationality. In models of dynamic deliberation, however, there is no irrationality—only uncertainty—as the players deliberate. From this point of view, we see the examples used to motivate refinements of the Nash equilibrium in a new light.

Let us begin by reconsidering a matrix game used in Chapter 1 to motivate Selten's notion of a perfect equilibrium:

	C1	C2
R2	0,0	0,0
R1	1,1	0,0

We can get an idea of the deliberational dynamic structure under Nash deliberation by examining the orbits plotted in Figure 2.2. Every point in the interior of the space of indecision is carried by dynamic deliberation to the perfect equilibrium [R1, C1]. This is not a peculiarity of the Nash dynamics, but is true for any dynamics which seeks the good. Of course, if the players are both absolutely sure that [R2, C2] will be played, then act 2 has maximal expected utility for each player. But in this case, deliberation does not make sense. If, as I argued in the previous section, Bayesian deliberation must start in the interior of the space of indecision, *dynamic deliberation cannot lead to* [R2, C2].

Thus there is a natural motivation for a refinement of the Nash equilibrium concept in the theory of deliberational dynamics—that is, *an equilibrium which one can converge to by deliberation starting at a completely mixed state of indecision.* Let us call such an equilibrium *accessible.* In the foregoing example the odd equilibrium [R2, C2] is not accessible under Nash or Darwin deliberation. Does Selten's concept of perfection coincide with some variety of accessibility?

Now let us consider the following example of the kind Myerson (1978) used:

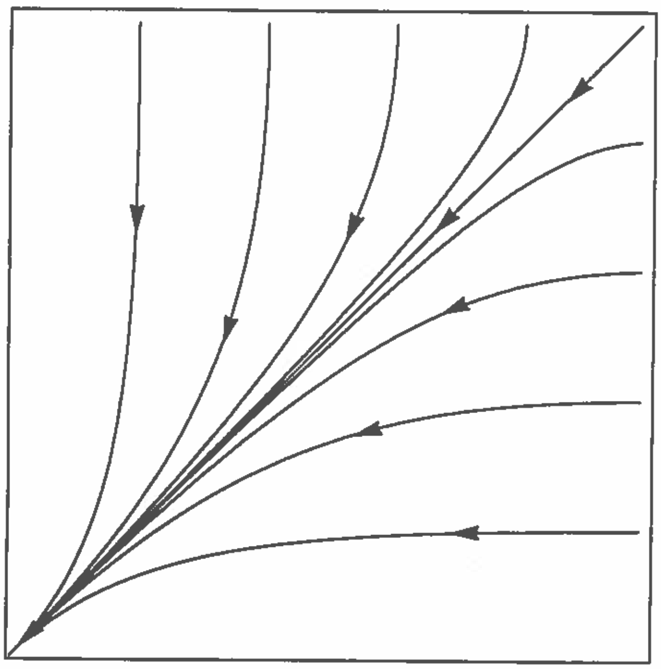

Figure 2.2. Perfect vs. Nash equilibrium

	C1	C2	C3
R3	−9, −9	−4, −4	−4, −4
R2	0,0	0,0	−4, −4
R1	1,1	0,0	−9, −9

Nash deliberation can lead to both the proper equilibrium at [R1, C1] and the improper equilibrium at [R2, C2]. Figure 2.3 shows the orbit starting at $p(\text{R1}) = 0.01$, $p(\text{R2}) = 0.5$, $p(\text{R3}) = 0.49$, $p(\text{C1}) = 0.01$, $p(\text{C2}) = 0.5$, $p(\text{C3}) = 0.49$ converging to the improper equilibrium at [R2, C2]. (Because of the symmetry of the game and the starting points, Row's orbit on the subspace with vertices R1, R2, and R3 is identical to

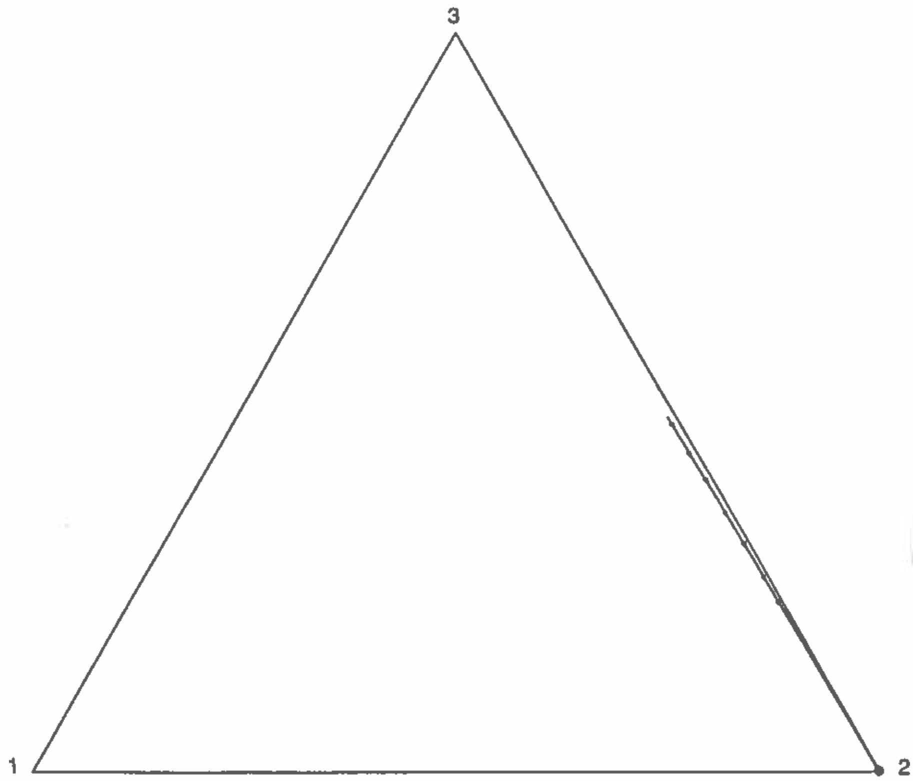

Figure 2.3. Myerson's game with Nash dynamics: orbit of [(0.01,0.5,0.49),(0.01,0.5,0.49)] converges to a perfect but improper equilibrium

Column's orbit on the subspace with vertices C1, C2, and C3.) Convergence to [R2, C2] is possible because the Nash dynamics is not affected by the relative attractiveness of acts with expected utility less than that of the status quo. Compare the orbit of the same point under Darwin deliberation (with the utilities appropriately rescaled), shown in Figure 2.4. As the probability of act 2 increases for a given player, the relative attractiveness of act 1 over act 3 increases for the other player. When the relative probability of act 1 over act 3 for a player gets large enough, act 1 becomes more attractive to the other player than act 2. The orbit then "turns the corner" and heads for the proper equilibrium.[16] Thus we have a natural motivation in deliberational dynamics for another refinement: equilibrium that can be reached by the Bayes dynamics starting

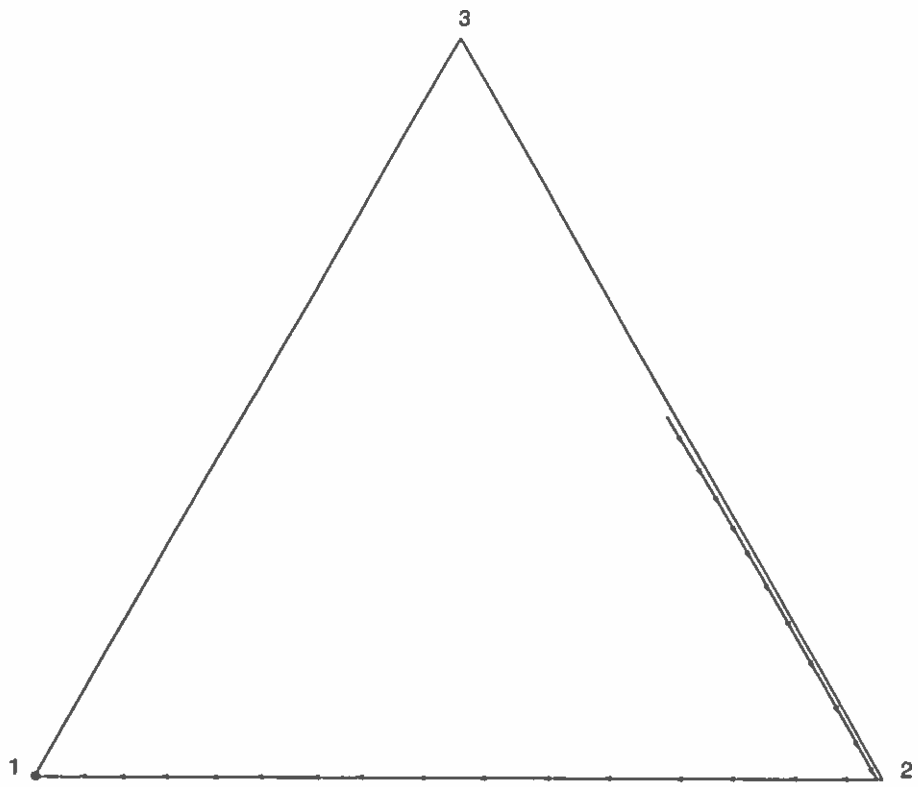

Figure 2.4. Myerson's game with Darwin dynamics: orbit of [(0.01,0.5,0.49),(0.01,0.5,0.49)] converges to a proper equilibrium

from a completely mixed point. Does Myerson's concept of proper equilibrium coincide with accessibility under a Bayes dynamics?

Analysis by Larry Samuelson (1988)[17] shows that our Bayesian accessibility concepts do not coincide with the refinements of the Nash equilibrium that have been proposed by Selten and Myerson. Darwin deliberation can lead to an equilibrium that is not only improper but also imperfect. Consider the following game:

	C1	C2
R2	0,0	0,0
R1	1,1	−1,−1

As shown in Figure 2.5, the orbit starting at $p(R2) = 0.5$, $p(C2) = 0.99$ is carried to an imperfect equilibrium by Darwin deliberation. The reason is that although C1 looks better than the status quo to Column at any completely mixed strategy, the velocity $dp(C1)/dt$ depends on *how much* better it looks. Along the orbit leading to the imperfect equilibri-

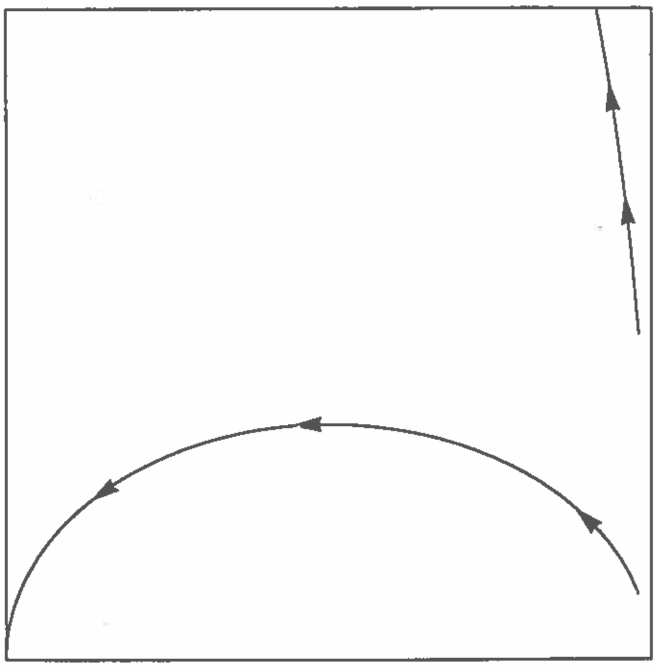

Figure 2.5. Darwin can lead to an imperfect equilibrium

um, R2 looks better than R1 to Row, and it looks enough better so that the orbit converges to the imperfect mixed equilibrium shown.

Compare this outcome with the results of Nash deliberation for the same game shown in Figure 2.6. The orbit of $p(\text{R2}) = 0.05$, $p(\text{C2}) = 0.99$ goes nicely to the perfect equilibrium, as do the other cases shown. Notice, however, that this is due to the un-Bayesian nature of Nash deliberation. The covetability of C2 is identically zero throughout the interior of the space, even though C2 looks as nearly as good as C1 when $p(\text{R2})$ is near one.

This example suggests that there may be a class of adaptive rules such that for rules in that class the deliberationally accessible equilibria are just the perfect ones. In fact, for two-person games Samuelson (1988) has isolated such a class of rules. A key feature of this class is *ordinality*: the velocity of probability change of a strategy depends only on the ordinal ranking among strategies according to their expected utilities.

It is, however, not clear to me that the orbit of the Darwin deliberator in this example is in any way unreasonable. In this respect, this example

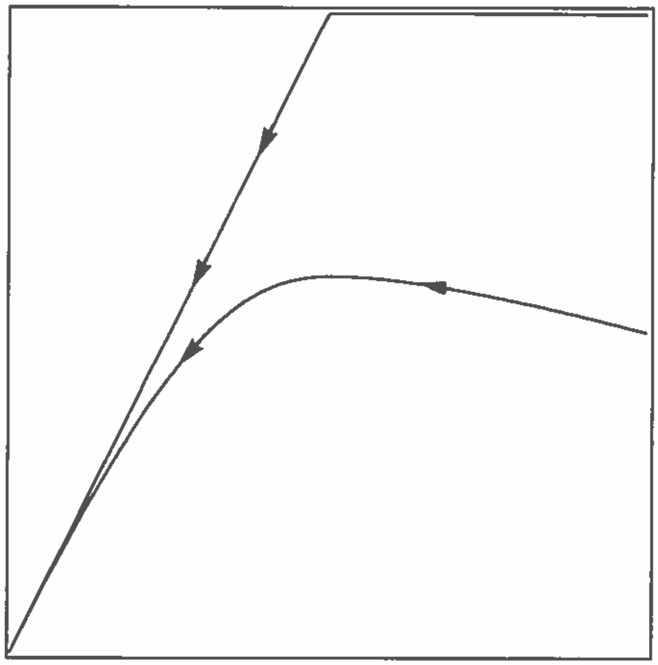

Figure 2.6. Nash converging to a perfect equilibrium

is quite different from the one given earlier. I will be bold enough to suggest that perhaps the correct response to this sort of example may be to reconsider the motivation for the definitions of perfect and proper equilibrium. In each of these definitions, it is required that the relative probabilities of a better and a worse prospect change in a way that is not sensitive to the magnitude of the difference. From the standpoint of dynamic Bayesian deliberation, the ordinal nature of these definitions is hard to justify.

The model of the dynamic deliberator may make more sense out of the program of refining the Nash equilibrium concept than does the metaphor of the trembling hand, but it also suggests that important boundaries are to be drawn in somewhat different places.

Refinements of the Nash Equilibrium for the Extensive Form

Dynamic deliberation for games in extensive form is a straightforward generalization of deliberational dynamics for games in normal form. The essential difference is that at an information set, the player's expected utilities are calculated using probabilities *conditional on being at that information set.*[18] This is simply a matter of respecting the informational structure that is specified in the game tree, but it clearly leads to deliberational inequivalences between an extensive form game and its strategic normal form.

To illustrate the process, let us reconsider the simple extensive-form game from Chapter 1 (Figure 2.7). Let us compare deliberation in the extensive form with deliberation in the strategic normal form:

	B1 if A2	B2 if A2
A1	0,0	0,0
A2	1,1	−1, −1

This is a familiar matrix. There are two Nash equilibria, a sensible one [A2, B1 if A2] and a questionable one [A1, B2 if A2]. For deliberation under Nash dynamics on the strategic normal form, both coincide with deliberational equilibria with the sensible one being strongly stable and the questionable one being unstable. But under Nash deliberation on the extensive form game, the questionable Nash equilibrium [A1, B2 if A2] is not a deliberational equilibrium at all. The reason is that the probability that A does A2 *conditional on B's information set where he chooses between* B1 *and* B2 must be one, no matter what the unconditional probabilities are. Using these conditional probabilities, B is faced with a choice of payoffs between 1 and −1. Deliberation on the strategic normal form uses the unconditional probabilities and ignores the informational structure of the tree.

This difference between deliberation on the strategic normal form and that on the extensive form is evident in a more subtle way when we consider accessibility from a completely mixed point under Darwin deliberation. This is the normal-form matrix that I used to make Samuelson's point. There are imperfect mixed equilibria, such as [A1, $p(B2|A2) = 0.96$], that are accessible under the Darwin dynamics. These equilibria are *not* accessible under Darwin dynamics for deliberational dynamics based on the true extensive form. Figure 2.8 shows the phase portraits for this game under Darwin deliberation for (A) the strategic normal form and (B) the extensive form. The difference between them is this: in simultaneous deliberation about strategies (Figure 2.8A), B's expected utility for the strategy B2 if A2 is a weighted average of B's payoff from this strategy if A does A1 and B's payoff from the strategy if A does A2; in extensive-form deliberation (Figure 2.8B), B uses his expected utilities *at his information set* so that the utility of B1 if A2 is 1 and the utility of B2 if A2 is −1 throughout deliberation. It is true that B1 if A2 looks better than B2 if A2 to both deliberators, but the *magnitude* of the difference shrinks to zero for deliberation in the normal form but remains constant for deliberation in the extensive form. Deliberation with a dynamical rule like Darwin, for which these relative magnitudes are crucial, puts the difference between extensive form and strategic normal form in the spotlight.

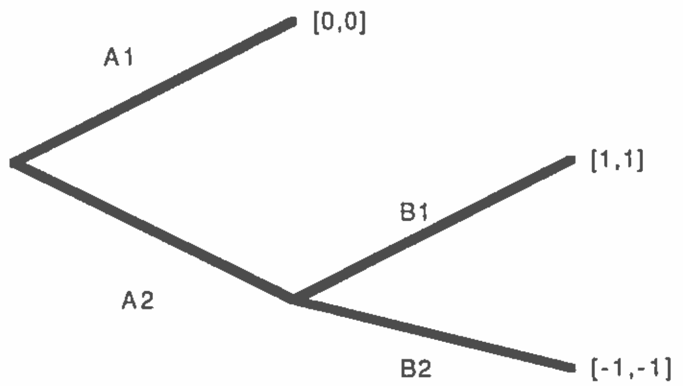

Figure 2.7. A challenge to strategic normal form

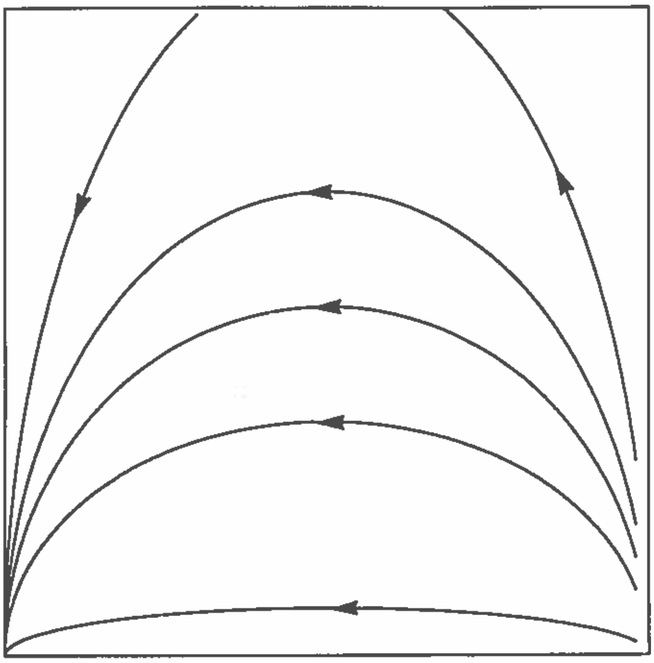

Figure 2.8A. Darwin dynamics on strategic normal form

As another illustration of the same sort of phenomenon, consider the sinister toy example of nuclear deterrence presented in Figure 2.9. At the onset, A can either attack or not. Then B can retaliate or not. If A attacks and B retaliates, A is devastated by B (A's payoff is -10) and B is devastated by A and suffers additional loss from the fallout from his own strike (B's payoff is -11). If A attacks and B doesn't retaliate, A gains an advantage ($+1$) and B is devastated (-10). If A doesn't attack and B "retaliates" anyway, A is devastated (-10) and B gains a slight advantage but perhaps also invites retaliation from A—let us say B's payoff is -5. If A doesn't attack and B doesn't retaliate, both get a payoff of 0. In strategic normal form there are three pure Nash equilibria: MAD—A doesn't attack and B retaliates if and only if attacked; First strike 1—A attacks and B does not retaliate whether attacked or not; and First strike 2—A attacks and B retaliates if and only if not attacked. But strategic normal form conceals the fact that MAD rests on a noncredible threat. Deliberation on the extensive-form game will lead from a state of initial uncertainty to First strike. Once we see retaliation as a non-

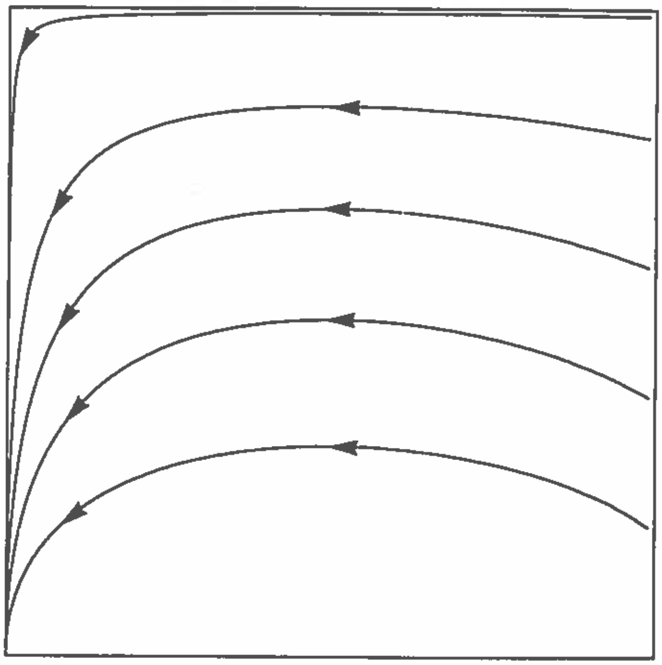

Figure 2.8B. Darwin dynamics on extensive form

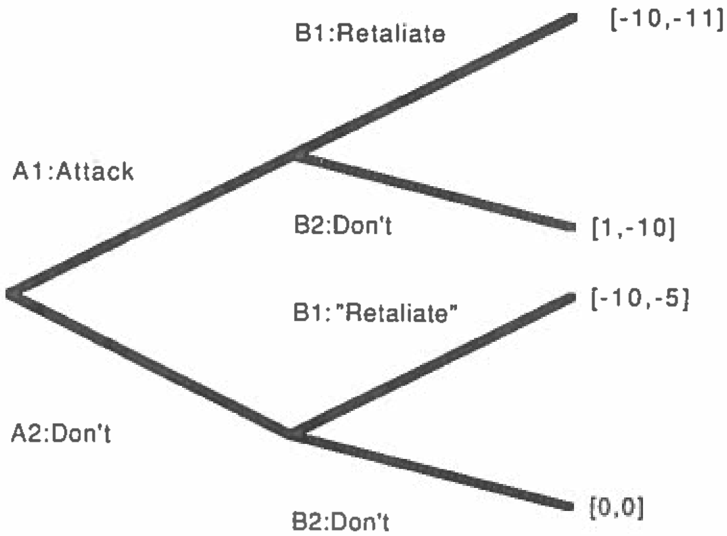

Figure 2.9. MAD vs. First strike

credible threat we might reconsider the payoffs when A doesn't attack and B "retaliates" anyway. B's payoff here should be just +1 for the advantage gained. I invite you to explore the deliberational dynamics of this modified game and to think of further modifications or alternatives to it.

Things become more interesting in games where some information sets are not unit sets. In some such games the same point can be made without worrying about the probabilities on the information set. Consider the game due to Kreps and Wilson (1982b, p. 871) given in Figure 2.10. One Nash equilibrium has A playing A1 and B having the strategy of play B2 if he finds himself at the information set where A has played either A2 or A3. B's strategy is not credible. No matter how B's probabilities might tilt between A2 and A3 at this information set, B will be better off choosing B1 than B2, since B1 gives him a better payoff than B2 in any case. Thus B would play B1 at this information set, and A, realizing this, will play A2 and assure herself a payoff of 12. This more credible equilibrium [A2, B1 if A2 or A3] is *sequential* in the sense of Kreps and Wilson, whereas the noncredible one is not.[19] The non-

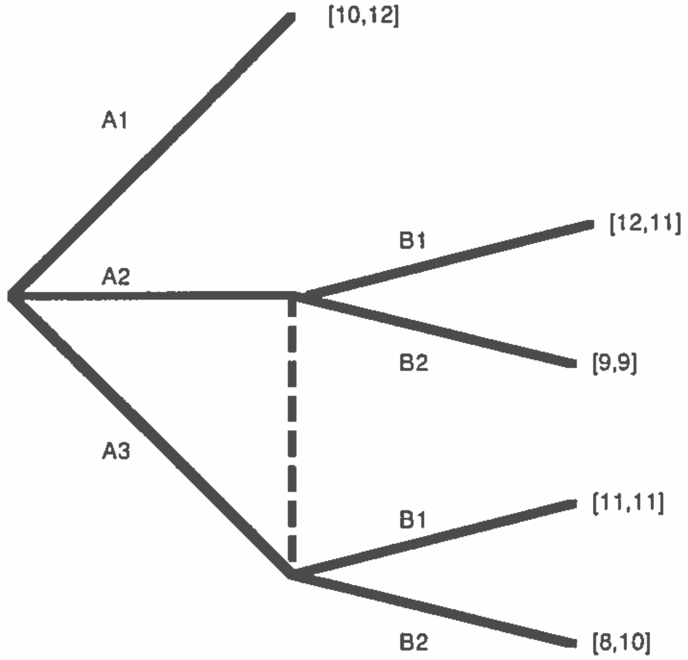

Figure 2.10. Sequential equilibrium vs. [subgame perfect] Nash equilibrium

credible equilibrium has B's agent at the information set [A2 or A3] make a choice that cannot maximize expected utility for any probability over [A1, A2]. As is to be expected, deliberational dynamics will never lead to the "bad" equilibrium and will always lead to the "good" one.

The power of dynamic deliberation is underutilized in the foregoing examples, because although A has to worry about what B will do, B doesn't really have to worry about what A has done. Things are different in the example due to Kohlberg and Mertens (1986) in Figure 2.11. Here A must worry about what B will do at information set [A2 or A3], and B must worry about what A has done to get him in that information set. There is an equilibrium at [A1, B2 if A2 or A3], and one at [A2, B1 if A2 or A3]. Both equilibria are sequential, but there is nevertheless something wrong about the first one. If A plays A2, A will get a better payoff than if A plays A3, no matter what B does. A can figure this out and B can figure out that A can figure it out. So B should make his probability of A2 conditional on A2 or A3 high, which will lead him to play B1 if he finds himself at the information set [A2 or A3]. A should be able to figure *this* out, and so will play A2 to secure a payoff of 16.

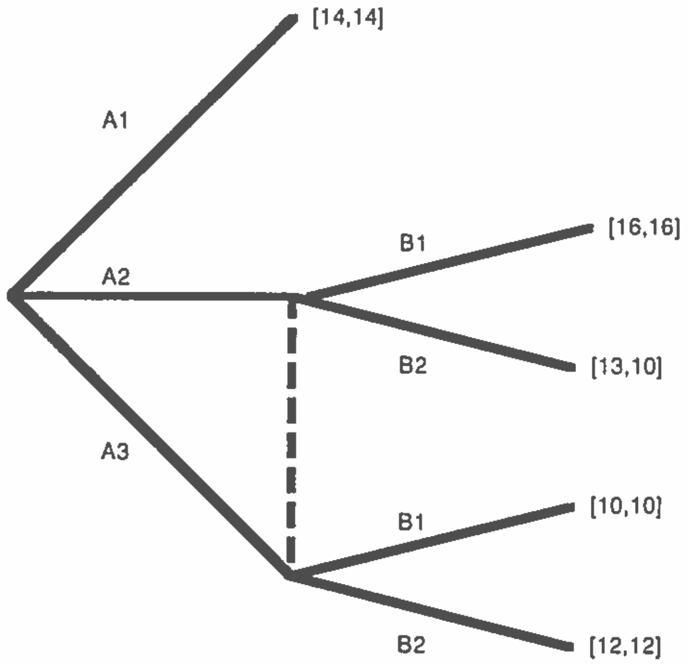

Figure 2.11. Proper vs. sequential equilibrium

This sort of reasoning can be implemented by deliberational dynamics provided we implement some version of the *Bayes dynamics* with the assumption of a *tendency toward better response*.[20] Figure 2.12 shows the action of the Darwin deliberation when this game is started near the "bad" sequential equilibrium. This is most striking when viewed in real time. The players appear to sit on the bad equilibrium for a long time, mulling it over, and then they suddenly start moving to the good *proper sequential equilibrium*. What is really happening is that very near the "bad" equilibrium, the ratios of the very tiny probabilities of A2 and A3 are being adjusted until [B1 if A2 or A3] begins to look better to B than the alternative. Probabilities are adjusted until A2 looks best and then the system moves rapidly toward the proper equilibrium.

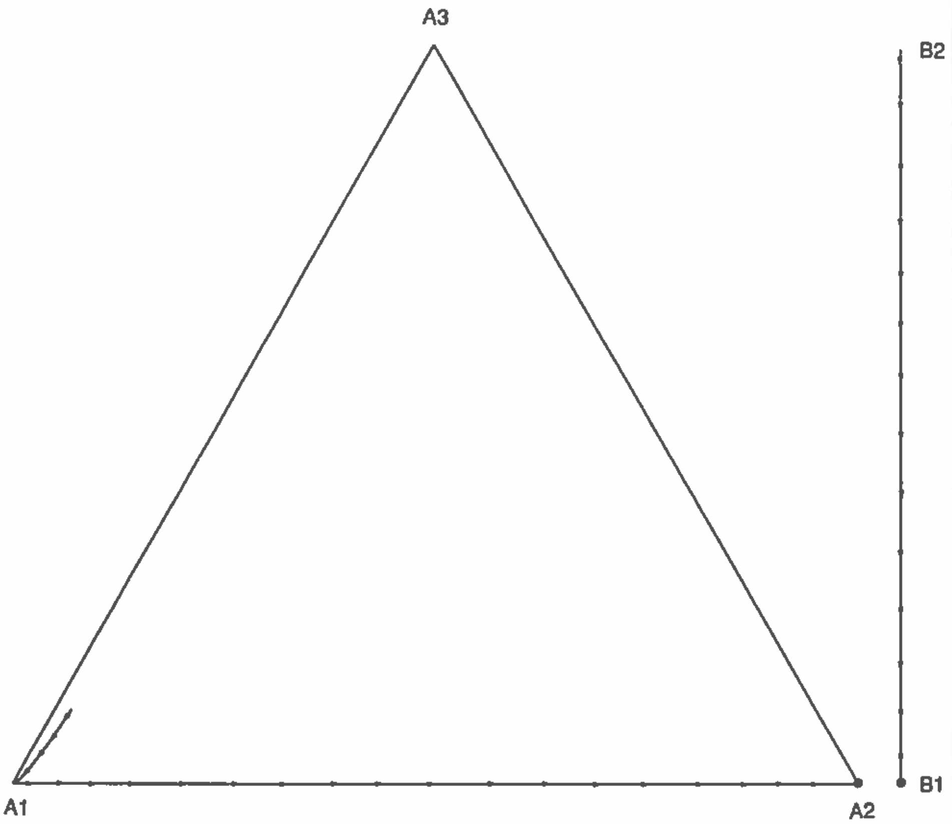

Figure 2.12. Darwin deliberators opt for a proper sequential equilibrium

So far so good, but Kohlberg and Mertens have an example in which the supposedly bad equilibrium is sequential and proper. In this example, Bayesian deliberational dynamics can lead to the "bad" equilibrium. I believe, however, that the arguments given here against the "bad" equilibrium are not conclusive, and that whether it is bad or not depends on the prior probabilities at the start of deliberation or on the mode of deliberation. The example is given in Figure 2.13. There is a perfect sequential equilibrium at [A1, B2 if A2 or A3] and another at [A2, B1 if A2 or A3]. Kohlberg and Mertens argue as follows: A1 strictly dominates A3 (in other words, no matter what B does, A1 gives B a greater payoff than A3). Therefore, B should know that A will play A2 or A3 only if she plays A2. Accordingly, B will play B1 rather than B2, and knowing this A will play A2 rather than A1.

Notice the difference between the reasoning in this example and that given in the preceding one. In the example given in Figure 2.12 it was argued that upon reflection we would have to conclude that the probability of A2 conditional on A2 or A3 should be high because A2 strictly dominates A3. In this example it is being argued that the same condi-

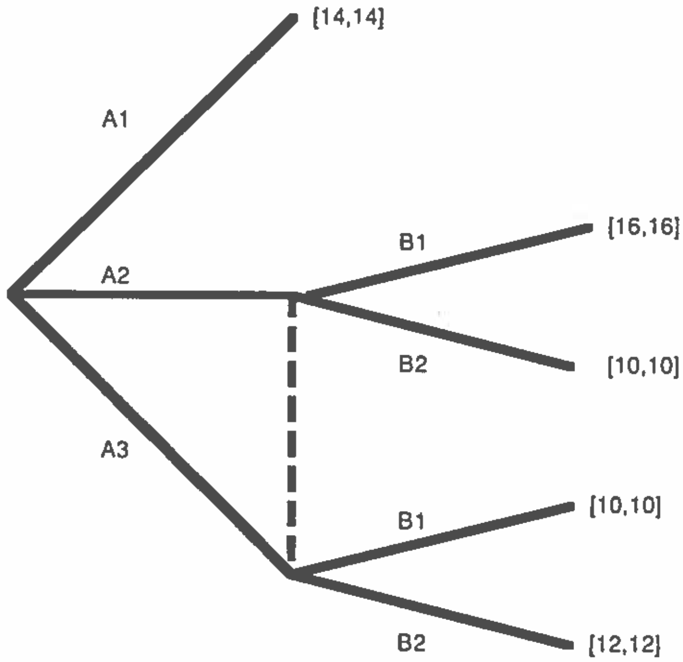

Figure 2.13. Dominance and initial degrees of belief

tional probability should be high because *A1* dominates A3. This argument is a *non sequitur* unless it can be supplemented with some extra assumptions.

Darwin deliberation can take players to either equilibrium. Questions of timing here are crucial. *If* before deliberation begins player A has been looking for strictly dominated strategies, and giving them probabilities that are virtually zero and very small relative to the probabilities of all other strategies, then deliberation will begin with p(A2|A2 or A3) large. In this case, deliberational dynamics will lead to the "good" equilibrium at [A2, B1 if A2 or A3]. On the other hand, the players may have started deliberation without noticing dominance or they may for some reason have started deliberation with A2 having a very small probability (or both these conditions may hold), so that initially p(A3|A2 or A3) is large. Suppose also that for whatever reasons, the players' initial p(B2|A2 or A3) is large. Then during deliberation dominance will be reflected by the ratio p(A1)/p(A3) getting large, but at the same time the magnitudes of the expected utilities will also lead to p(A1)/p(A2) getting larger. In this case deliberation will lead to the "bad" equilibrium at [A1, B2 if A2 or A3]. Figure 2.14 shows the two kinds of orbit.

On the other hand, rather than pushing these considerations into the predeliberational beliefs, one might focus on deliberators who build temporal precedence of considerations of dominance into the deliberational rules. For example, the players might begin deliberation with a routine for iterated elimination of strictly dominated strategies and then proceed to apply Nash or Darwin deliberation to the remaining problem. Such deliberators would go through essentially the same reasoning as Kohlberg and Mertens and end up at the preferred equilibrium. This sort of two-stage procedure constitutes one reasonable way for deliberation to proceed; but is it the *only* reasonable way for deliberation to proceed? If not, the Kohlberg-Mertens reasoning applies only to a proper subclass of rational deliberators.

Correlated Equilibria

I will close this chapter with a discussion of the relation of dynamic deliberation to a rather different equilibrium concept, Aumann's (1974) notion of a *correlated equilibrium*. Two different points of view may be adopted in discussions of "solution concepts" for games. One is the point of view of the players themselves as rational actors. The other is that of a disinterested rational observer or theorist. I have so far discussed deliberational dynamics from the viewpoint of the deliberators themselves, but the subject has interesting consequences for the external point of view as well. I shall illustrate these consequences in two cases where the external point of view is taken by a philosopher or social theorist: the question of the possibility of convention and the question of the nature of the "state of nature." In each case, rational deliberation generates correlation. This phenomenon can be described generally using the notion of a *correlated equilibrium.*

How is convention possible? Quine (1936) challenged conventionalist accounts of language to provide a satisfactory account of how the relevant conventions are set up and maintained that does not presuppose linguistic communication or competency. David Lewis (1969) replied that convention is possible without communication. The mutual expec-

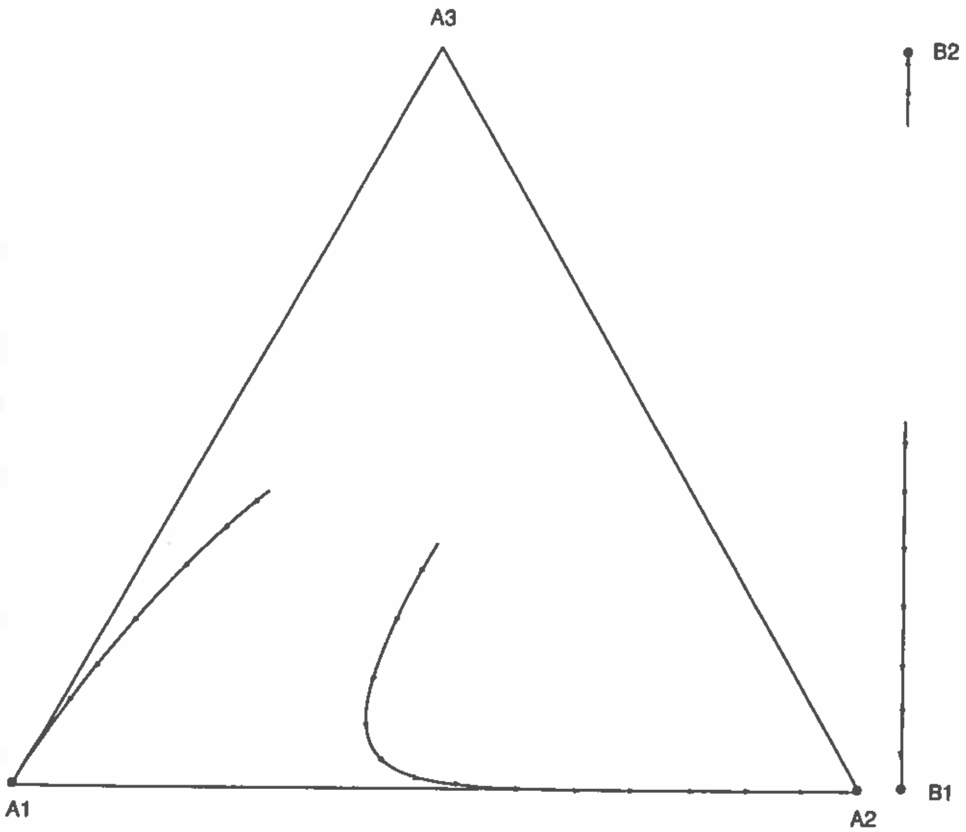

Figure 2.14. Darwin deliberators in Kohlberg and Mertens' game

tations of rational agents can explain the maintenance of a convention at a game-theoretic equilibrium. Consider the pure coordination game:

The Winding Road

	Left	Right
Left	1,1	0,0
Right	0,0	1,1

Two cars approach a blind curve from opposite directions. Each would prefer that they are both driving on the left or both driving on the right. There are two pure equilibria, equally attractive, but if Row goes for one of them and Column goes for the other, they will end up in trouble. If, however, Row believes Column expects him to drive on the left and believes that Column believes him to believe this, and so on, and Column believes likewise about Row, and each believes that the other is rational and that the other believes that he is, then they each have good reason to drive on the left.

The question as to how convention without communication is possible between rational agents has two parts: (1) How can convention without communication be sustained? and (2) How can convention without communication be generated? Lewis gave the answer to the first question in terms of equilibrium (or stable equilibrium) and common knowledge of rationality. His discussion of the second question—following Schelling (1960)—is framed in terms of salience, where a salient coordination equilibrium is "one which stands out from the others in some conspicuous respect." Salience could derive from preplay communication among the players, but it could also arise in other ways. It could arise by precedent. In fact, since salience is a psychological rather than a logical notion, the ways in which salience may arise are as various as the possible psychologies of the players.

The informal discussions of salience by Lewis and Schelling are convincing regarding the plausibility of real-world coordination by salience, but I believe that they give only a partial answer to the second question. Here, deliberational dynamics has something to contribute.

Let us model The Winding Road as a game played by Nash deliberators. (The results would be essentially the same here if we used Darwin deliberation.) Row and Column each have predeliberational probabilities of driving on the left or right. They can be anything at all. At the onset of deliberation each player's initial probabilities of driving left or right are announced and become common knowledge. (This idealization will be weakened later.) You—the philosopher—have some probability distribution over the space of Row's and Column's initial probabilities. You needn't think it likely that they are anywhere near an equilibrium. In fact, we will suppose only that your probability distribution is reasonably smooth (that is, it is absolutely continuous with respect to Lebesgue measure on the unit square), otherwise it can be anything at all. Then you should believe with probability one that the deliberators will converge to one of the pure Nash equilibria, as is evident from Figure 2.15.

It is not surprising here that the players should be led to the state of mutually reinforcing expectations that attend a Nash equilibrium. Coordination is effected by rational deliberation. Precedent and other forms of initial salience may influence the deliberators' initial probabilities, and thus may play a role in determining *which* equilibrium is selected. The answer to the question of how convention can be generated for Bayesian deliberators has both methodological and psychological aspects.

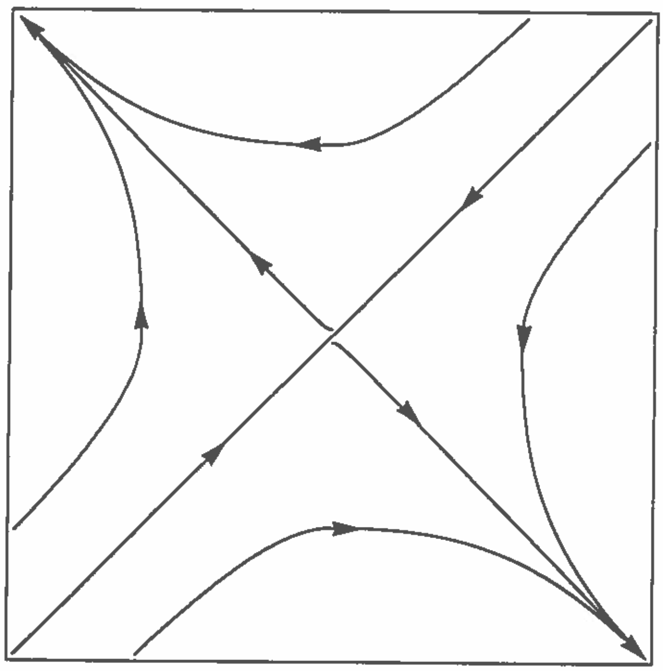

Figure 2.15. The Winding Road

Of course, Bayesian players are not always so lucky as to be involved in pure coordination games. People have conflicting desires and limited altruism. They are roughly equal in their mental and physical powers. Elements of competition intrude. Thomas Hobbes argued that as a consequence rational self-interested decisionmakers in a state of nature, unrestrained by the power of a sovereign, will be engaged in a "war of all against all."

After Darwin, Hobbesian philosophy enjoyed a resurgence. Karl Marx wrote to Engels: "It is remarkable how Darwin has discerned anew among beasts and plants his English society . . . It is Hobbes' *bellum omnium contra omnes*."[21] Marx was being somewhat unfair to Darwin, but "Darwin's Bulldog" lived up to the caricature. In an essay entitled "The Struggle for Existence" (1888), T. H. Huxley popularized Hobbesian Darwinism. Primitive man "fights out the struggle for existence to the bitter end, like any other animal . . . Life was a continual free fight, and beyond the limited and temporary relations of the family, the Hobbesian war of each against all was the normal state of existence" (p. 165). This picture seemed so perverted to Prince Petr Kropotkin that he was moved to write *Mutual Aid* (1902), describing cooperation among animals and also among men in all stages of civilization.

In a fine critical study, Gregory Kavka (1983) found Hobbes' argument inconclusive although, as he points out, many other commentators appear to regard it as obviously correct. Kavka and Gauthier (1969) modeled conflict in the state of nature in terms of Prisoner's Dilemma, but I think the game of Chicken models Hobbes' premises at least as well.

Chicken

	Don't swerve	Swerve
Don't swerve	$-10,-10$	$5,-5$
Swerve	$-5,5$	$0,0$

Each player would like to profit from his opponent's loss. Each would like, at the outset, to appear more aggressive than his opponent. But aggression on the part of both creates an intolerable situation. There are two Nash equilibria in pure strategies: Row swerves and Column doesn't and Column swerves and Row doesn't. There is also a mixed equilibrium where each player has equal chances of swerving and not swerving. If the players are Bayesian deliberators coordination can again be achieved by deliberation, just as in the coordination game. As shown in Figure 2.16, for Nash deliberators every initial point leads to a Nash equilibrium, and almost every initial point leads to a pure Nash equilibrium.

In this example, it is easy to say which initial points go to which equilibrium. For almost every initial point, one player is initially more likely to swerve and if so that player ends up swerving while the other player does not. In the case in which both players are initially equally likely to swerve, they are carried to a mixed equilibrium where each adopts a random strategy of swerving with chance of 0.5. Here there is a genuine Hobbesian incentive for initial bellicosity. (There is none in Prisoner's Dilemma.) Nevertheless, crashes are almost always avoided as a result of rational deliberation.

Did Hobbes attempt to derive a Prisoner's Dilemma conclusion from Chicken premises? It would be premature to draw this conclusion from such an oversimplified model of the state of nature. A number of complications need to be introduced before we could begin to do Hobbes justice (see Chapter 6). One can find materials for more realistic models

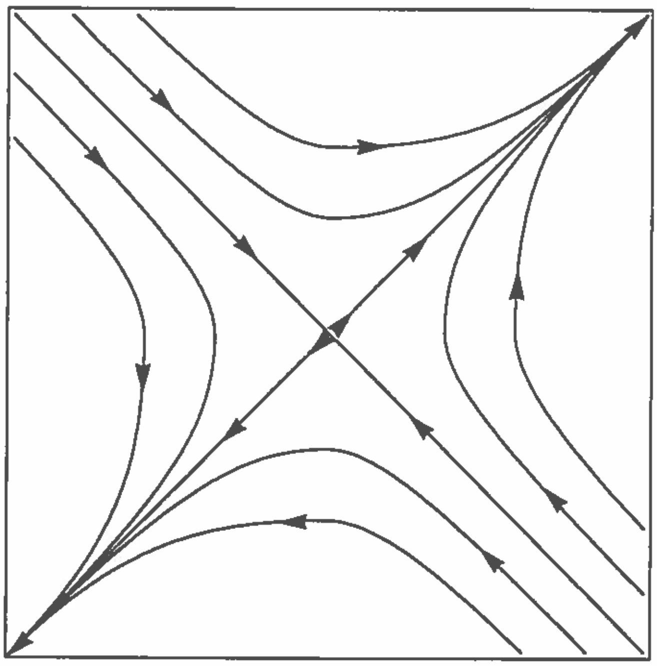

Figure 2.16. Chicken

in ethological descriptions of varieties of animal conflict. Of particular interest is the prevalence of ritualized aggression in which little real damage is done (Lorenz, 1966; Eibl-Eibesfeld, 1970). This *is* part of the state of nature, and it does not agree with Hobbes' description. Kropotkin's remark of 1902 is even more suitable now: "science has made some progress since Hobbes' time, and we have safer ground to stand upon than the speculations of Hobbes and Rousseau."

One might argue that the rationality of humans invalidates the analogy, but the analyses of evolutionary game theory do not support this objection (see Maynard Smith and Price, 1973; Parker, 1974; Maynard Smith, 1982). For these sorts of game-theoretic models, Bayesian deliberators of the kind considered here will decide by deliberational dynamics in a way analogous to the way that Mother Nature decides by evolution. In the games considered by Maynard Smith and Price, self-interested rational deliberators will play in a decidedly un-Hobbesian way. In general, we must agree with the carefully considered conclusion of Kavka (1986, p. 122) that "which strategy is better overall probably cannot be determined a priori for all state of nature situations. Instead, it will depend on the value of a number of important variables and parameters, which will vary according to the version of the state of nature in question."

These considerations, however incomplete they may be with respect to Hobbes, nevertheless make the general point that deliberation may play a role in the genesis of coordination in situations with a considerable amount of competition as well as in pure coordination games. There is a general conception under which all these cases fall. It is a notion usually discussed in *cooperative* game theory—a *correlated equilibrium.* Aumann (1974) suggested that mixed strategies, where the chance devices used by different players are assumed independent, be treated as a special case of *correlated* strategies, where the chance devices may have any joint probability distribution at all.

You might think of a referee observing the outcome of some random process—say, the toss of a many-sided die—and communicating to each player the aspect of the process that is *that player's* random variable. For example, one player might get to know the color of the face, another the number of spots showing. Any correlation of colors and spots in the random device is allowed.

A *random strategy* for player *i* can be thought of as a probability assignment to his space of possible actions, A_i, namely, a random variable mapping some probability space into A_1. A *joint correlated strategy* can be thought of as any probability assignment on the product space of the action spaces of all players, $A_1 \times A_2 \times \ldots \times A_n$, or a mapping of a probability space into sequences of actions. A correlated strategy can obviously be specified by giving the underlying joint probability space together with a sequence of random variables on it: $f_1, f_2, \ldots f_n$, where f_i maps the probability space into the action space of player i. In the special case of ordinary mixed strategies, the probability on the product space is the product measure and the random strategies are independent. An *i-deviation* from a correlated strategy C consists of the same probability space and the same random variables f_i except that each f_i is replaced by some random variable $g(f_i)$ taking values in A_i. An i-deviation represents player i unilaterally deviating from the original joint correlated strategy so that when the original strategy tells him to do one thing he does something else, while all other players stick to the original correlated strategy. A *correlated equilibrium* is a joint correlated strategy such that for each player i expected utility on the joint correlated equilibrium strategy is greater than or equal to his expected utility on any i-deviation from it (the expectation being taken according to the underlying probability space). Thus, a correlated equilibrium is a joint correlated strategy from which no player has anything to gain by unilateral deviation.

In certain situations, it might be to all the players' mutual advantage to agree on a joint correlated equilibrium strategy and then either hire a referee or construct a machine to carry out the random experiment and communicate to each player the action selected for him. On the face of it, it might appear that "for strategies to be correlated there must be some mechanism for communicating and contracting between the players" (Shubik, 1982, p. 247). But, as we have seen in several examples, rational deliberation can play a powerful role in establishing correlation. Let us consider in a general way the sort of situation sketched at the beginning of this section.

An observer, Theo, knows that n players will be induced to play a certain n-person noncooperative game. Theo knows that the players are all Bayesian dynamic deliberators with a common dynamics and that this fact will be common knowledge to the players at the onset of deliberation, as will their prior probabilities. Theo has analyzed the game and knows that in it (as in every example we have seen so far) the dynamics always converges to a Nash equilibrium. Theo may or may not know who the players are. He does not know what their initial probabilities for their possible actions will be, but rather has his own probability measure over the possible initial states of indecision of the system. Although the interpretation of the mathematics is quite different, we nevertheless see that with respect to Theo's probability measure *the players are at a correlated equilibrium.*

When the true initial state of indecision is selected, a recommendation for action is delivered up to each player by deliberational dynamics. Since the dynamics leads from each initial state to a Nash equilibrium, no player has anything to gain by deviating from that recommendation. Thus no i-deviation from the joint correlated strategy defined by Theo's probability is preferable to it for player i, so that joint correlated strategy is, by definition, a correlated equilibrium. This is true no matter what Theo's probability measure over the space of initial states of indecision. *This correlated equilibrium is a general result of the players' common knowledge and Bayesian dynamic deliberation.*

The same result may be obtained without the outside observer if prior to deliberation the players themselves share the role of Theo. For example, Sue and Dora are going to fly to the small country of Freedonia for a vacation and each plans to rent a car. They are to pick up cars at a deserted airport in Freedonia. Sue thinks it likely that Freedonians drive on the left; Dora thinks it likely that Freedonians drive on the right. There may be no one else in Freedonia then because it is a special holiday, and there are no road signs in Freedonia. Sue and Dora prepare to be involved in a game of The Winding Road with one another. They agree that before leaving the airport they will share their then current probabilities of opting for Left or Right and then go their separate ways, deliberate, and do the best they can.

This example brings us close to the point of view of Aumann (1987), who argued that correlated equilibrium is a consequence of a common prior probability together with common knowledge of Bayesian rationality. The latter is taken to be common knowledge that the players will each arrive at a decision that maximizes that player's expected utility. The former includes prior probabilities over what each player will ultimately choose and is, then, itself interpreted as the probability setting up the correlated strategy, with the joint maximization of expected utility assuring that it is an equilibrium.

Aumann's viewpoint is somewhat different from the one presented here in that he does not consider the process of deliberation, but only its result. So there is no analysis of how the players jointly arrive at decisions where each maximizes his expected utility. In contrast, we made additional assumptions to get stronger conclusions. We assumed common knowledge of the dynamical law of deliberation, which is a stronger common-knowledge assumption than that used by Aumann. This is what enables accurate updating by emulation and assures that a state at which each player is at a deliberational equilibrium corresponds to a Nash equilibrium for the game. Consequently, our predeliberational correlated equilibria are mixtures of Nash equilibria. As such they are a proper subclass of Aumann's correlated equilibria, that have especially tight correlation. Considerations of deliberational dynamics add a further dimension to the theory of correlated equilibria and provide an account of one way in which correlated equilibria can be generated.

Equilibria and Rationality

We saw in Chapter 1 that it was hard to justify the Nash equilibrium concept, even for two-person zero-sum games, without making further assumptions. In this chapter we have considered some very simple models of bounded Bayesian deliberators who, under quite strong conditions of prior common knowledge, are at a joint deliberational equilibrium if and only if they are at a Nash equilibrium. Refinements of the deliberational dynamics in a qualitatively Bayesian direction leads naturally to refinements of the Nash equilibrium concept. There is also an important connection between deliberational dynamics and Aumann's concept of a correlated equilibrium. These results about equilibria come from strong assumptions, and one would like to know more about how sensitive they are to small changes in those assumptions.