I | Coordination and Convention

1. Sample Coordination Problems

Use of language belongs to a class of situations with a conspicuous common character: situations I shall call *coordination problems.* I postpone a definition until we have seen a few examples. We begin with situations that might arise between two people—call them "you" and "I."

(1) Suppose you and I both want to meet each other. We will meet if and only if we go to the same place. It matters little to either of us where (within limits) he goes if he meets the other there; and it matters little to either of us where he goes if he fails to meet the other there. We must each choose where to go. The best place for me to go is the place where you will go, so I try to figure out where you will go and to go there myself. You do the same. Each chooses according to his expectation of the other's choice. If either succeeds, so does the other; the outcome is one we both desired.

(2) Suppose you and I are talking on the telephone and we are unexpectedly cut off after three minutes. We both want the connection restored immediately, which it will be if and only if one of us calls back while the other waits. It matters little to either of us whether he is the one to call back or the one to wait. We must each choose whether to call back, each according to his expectation of the other's choice, in order to call back if and only if the other waits.

(3) An example from Hume's *Treatise of Human Nature:* Suppose you and I are rowing a boat together. If we row in rhythm, the boat goes smoothly forward; otherwise the boat goes slowly and erratically,

we waste effort, and we risk hitting things. We are always choosing whether to row faster or slower; it matters little to either of us at what rate we row, provided we row in rhythm. So each is constantly adjusting his rate to match the rate he expects the other to maintain.

Now we turn to situations among more than two people.

(4) Suppose several of us are driving on the same winding two-lane roads. It matters little to anyone whether he drives in the left or the right lane, provided the others do likewise. But if some drive in the left lane and some in the right, everyone is in danger of collision. So each must choose whether to drive in the left lane or in the right, according to his expectations about the others: to drive in the left lane if most or all of the others do, to drive in the right lane if most or all of the others do (and to drive where he pleases if the others are more or less equally divided).

(5) Suppose we are campers who have gone looking for firewood. It matters little to anyone in which direction he goes, but if any two go in the same direction they are likely to cover the same ground so that the one who gets there later finds no wood. Each must choose a direction to go according to his expectations about the others: one different from anyone else's.

(6) Suppose several of us have been invited to a party. It matters little to anyone how he dresses. But he would be embarrassed if the others dressed alike and he dressed differently, since he knows that some discreditable explanation for that difference can be produced by whoever is so inclined. So each must dress according to his expectations about the way the others will dress: in a tuxedo if the others will wear tuxedos, in a clown suit if the others will wear clown suits (and in what he pleases if the others will dress in diverse ways).

(7) Suppose we are contented oligopolists. As the price of our raw material varies, we must each set new prices. It is to no one's advantage to set his prices higher than the others set theirs, since if he does he tends to lose his share of the market. Nor is it to anyone's advantage to set his prices lower than the others set theirs, since if he does he menaces his competitors and incurs their retaliation. So each must

set his prices within the range of prices he expects the others to set.

(8) An example from Rousseau's *Discours sur l'inégalité:* Suppose we are in a wilderness without food. Separately we can catch rabbits and eat badly. Together we can catch stags and eat well. But if even one of us deserts the stag hunt to catch a rabbit, the stag will get away; so the other stag hunters will not eat unless they desert too. Each must choose whether to stay with the stag hunt or desert according to his expectations about the others, staying if and only if no one else will desert.

(9) Suppose we take it to be in our common interest that some scarce good, say grazing land, should be divided up somehow so that each of us can count on having the exclusive use of one portion. (Suppose nobody ever thinks it would be in his interest to help himself to someone else's portion. The struggle, the harm to his neighbor, the bad example, the general loss of confidence, invariably seem to outweigh any gain.) It matters little to anyone who uses which portion, so long as people never try to use the same portion and no portion ever goes to waste. Each must choose which portion to use according to his expectations about the portions others will use and the portion they will leave for him.

(10) Suppose we are tradesmen. It matters little to any of us what commodities he takes in exchange for goods (other than commodities he himself can use). But if he takes what others refuse he is stuck with something useless, and if he refuses what others take he needlessly inconveniences his customers and himself. Each must choose what he will take according to his expectations about what he can spend—that is, about what the others will take: gold and silver if he can spend gold and silver, U.S. notes if he can spend U.S. notes, Canadian pennies if he can spend Canadian pennies, wampum if he can spend wampum, goats if he can spend goats, whatever may come along if he can spend whatever may come along, nothing if he can spend nothing.

(11) Suppose that with practice we could adopt any language in some wide range. It matters comparatively little to anyone (in the

long run) what language he adopts, so long as he and those around him adopt the same language and can communicate easily. Each must choose what language to adopt according to his expectations about his neighbors' language: English among English speakers, Welsh among Welsh speakers, Esperanto among Esperanto speakers, and so on.

2. Analysis of Coordination Problems

With these examples, let us see how to describe the common character of coordination problems.

Two or more agents must each choose one of several alternative actions. Often all the agents have the same set of alternative actions, but that is not necessary. The outcomes the agents want to produce or prevent are determined jointly by the actions of all the agents. So the outcome of any action an agent might choose depends on the actions of the other agents. That is why—as we have seen in every example—each must choose what to do according to his expectations about what the others will do.

Some combinations of the agents' chosen actions are *equilibria:* combinations in which each agent has done as well as he can given the actions of the other agents. In an equilibrium combination, no one agent could have produced an outcome more to his liking by acting differently, unless some of the others' actions also had been different. No one regrets his choice after he learns how the others chose. No one has lost through lack of foreknowledge.

This is not to say that an equilibrium combination must produce an outcome that is best for even one of the agents (though if there is a combination that is best for everyone, that combination must be an equilibrium). In an equilibrium, it is entirely possible that some or all of the agents would have been better off if some or all had acted differently. What is not possible is that any one of the agents would have been better off if he alone had acted differently and all the rest had acted just as they did.

We can illustrate equilibria by drawing *payoff matrices* for coordination problems between two agents. Call the agents *Row-chooser* and *Column-chooser.* We represent Row-chooser's alternative actions by labeled rows of the matrix, and Column-chooser's by labeled columns. The squares then represent combinations of the agents' actions and the expected outcomes thereof. Squares are labeled with two *payoffs,* numbers somehow measuring the desirability of the expected outcome for Row-chooser and Column-chooser.[1] Row-chooser's payoff is at the lower left, Column-chooser's at the upper right.

Thus the matrix of Figure 1 might represent a simple version of example (1), where *R1*, *R2*, and *R3* are Row-chooser's actions of

	C1	C2	C3
R1	1 meet 1	0 0	0 0
R2	0 0	1 meet 1	0 0
R3	0 0	0 0	1 meet 1

Figure 1

going to places *P1, P2,* and *P3* respectively, and *C1, C2,* and *C3* are Column-chooser's actions of going to places *P1*, *P2*, and *P3* respectively. The equilibria are the three combinations in which Row-

[1] My account will demand no great sophistication about these numerical measures of desirability. If a foundation is required, it could be provided by decision theory as developed, for instance, by Richard Jeffrey in *The Logic of Decision* (New York: McGraw-Hill, 1965). I take it that decision theory applies in some approximate way to ordinary rational agents with imperfectly coherent preferences; our payoffs need never be more than rough indications of strength of preference.

chooser and Column-chooser go to the same place and meet there: ⟨*R1*, *C1*⟩, ⟨*R2*, *C2*⟩, and ⟨*R3*, *C3*⟩. For instance, ⟨*R2*, *C2*⟩ is an equilibrium by definition because Row-chooser prefers it to ⟨*R1*, *C2*⟩ or ⟨*R3*, *C2*⟩, and Column-chooser prefers it to ⟨*R2*, *C1*⟩ or ⟨*R2*, *C3*⟩. Both are indifferent between the three equilibria.

But suppose we change the example so that Row-chooser and Column-chooser care where they go, though not nearly so much as they care whether they meet. The new payoff matrix might be as shown in Figure 2. The equilibria remain the same: ⟨*R1*, *C1*⟩,

	C1	C2	C3
R1	1.5 meet 1.5	.2 .5	0 .5
R2	.5 .2	1.2 meet 1.2	0 .2
R3	.5 0	.2 0	1 meet 1

Figure 2

⟨*R2*, *C2*⟩, and ⟨*R3*, *C3*⟩. But Row-chooser and Column-chooser are no longer indifferent between the equilibria. ⟨*R1*, *C1*⟩ is the best possible outcome for both; ⟨*R3*, *C3*⟩ is the worst equilibrium outcome for both, though both prefer it to the nonequilibrium outcomes. Or if the payoff matrix were as shown in Figure 3, then ⟨*R1*, *C1*⟩ would be Row-chooser's best outcome and Column-chooser's worst equilibrium outcome; ⟨*R3*, *C3*⟩ would be Column-chooser's best outcome and Row-chooser's worst equilibrium outcome. No outcome would be best for both.

There seems to be a difference between equilibrium combinations in which every agent does the same action and equilibrium combinations in which agents do different actions. This difference is spurious, however. We say that the agents do the same action if they do actions

	C1	C2	C3
R1	1 meet 1.5	.2 .5	.5 .5
R2	0 .2	1.2 meet 1.2	.5 .2
R3	0 0	.2 0	1.5 meet 1

Figure 3

of the same kind, particular actions falling under a common description. But actions can be described in any number of ways, of which none has any compelling claim to primacy. For *any* combination of actions, and *a fortiori* for any equilibrium combination of actions, there is *some* way of describing the agents' alternative actions so that exactly those alternative actions in the given combination fall under a common description. Any combination, equilibrium or not, is a combination of actions of *a* same kind (a kind that excludes all the agents' alternative actions). Whether it can be called a combination in which every agent does the same action depends merely on the naturalness of that classification.

Consider example (2). If we have in mind these action-descriptions,

R1 or *C1*: calling back
R2 or *C2*: not calling back

we draw the payoff matrix shown in Figure 4 and think of the case as one in which the equilibria ⟨*R1*, *C2*⟩ and ⟨*R2*, *C1*⟩ are combinations in which the agents do different actions. But if we have in mind these action-descriptions,

R1′ or *C1′*: calling back if and only if one is the original caller
R2′ or *C2′*: calling back if and only if one is not the original caller

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

Figure 4

we draw the payoff matrix shown in Figure 5 and think of the case as one in which the equilibria ⟨*R1′*, *C1′*⟩ and ⟨*R2′*, *C2′*⟩ are combi-

	C1′	C2′
R1′	1 1	0 0
R2′	0 0	1 1

Figure 5

nations in which the agents do the same action. But what makes the first pair of action-descriptions more natural than the second? And so what if it is?

We might say that coordination problems are situations in which several agents try to achieve uniformity of action by each doing whatever the others will do. But this is a dangerous thing to say, since it is true of a coordination problem only under suitable descriptions of actions, and sometimes the descriptions that make it true would strike us as contrived—so, for instance, in examples (2), (5), (9), and perhaps (4). What is important about the uniform combinations we are interested in is not that they are—under some description—uniform, but that they are equilibria.

Of course this is not to say that coordination problems are distinguished by the presence of equilibria. Indeed the bulk of the mathematical theory of games is precisely the theory of equilibrium combinations (known also as *saddle points* or *solutions*) in situations of the opposite kind: pure conflict of interest between two agents, as in Figure 6.

	C1	C2	C3
R1	0 0	−.5 .5	−.5 .5
R2	.5 −.5	1 −1	−1 1
R3	.5 −.5	−1 1	1 −1

Figure 6

In general, pure conflict can be represented by a payoff matrix in which the agents' payoffs (perhaps after suitable linear rescaling) sum to zero in every square.[2] This is to say that one agent's losses are the others' gains, and vice versa. Yet there are equilibria in pure conflict. In the example shown, ⟨*R1*, *C1*⟩ is an equilibrium: Row-chooser prefers it to ⟨*R2*, *C1*⟩ or ⟨*R3*, *C1*⟩, and Column-chooser prefers it to ⟨*R1*, *C2*⟩ or ⟨*R1*, *C3*⟩.

Schelling argues for a "reorientation of game theory" in which games—problems of interdependent decision—are taken to range over a spectrum with games of pure conflict and games of pure coordination as opposite limits.[3] *Games of pure conflict,* in which the

[2]There is no point in changing the definition to let the sum be a constant other than zero. By allowing rescaling, we already have full generality. Without rescaling, we would not reach full generality just by allowing nonzero constant sums. And by allowing linear rescaling, we make clear why—despite appearance—our definitions do not depend on any problematic interpersonal comparison of desirabilities.

[3]*Strategy of Conflict,* pp. 83-118, 291-303.

agents' interests are perfectly opposed, can be defined as we have just seen. *Games of pure coordination,* in which the agents' interests coincide perfectly, are games in which the agents' payoffs (perhaps after suitable linear rescaling) are equal in every square. Other games are mixtures in varying proportions of conflict and coordination, of opposition and coincidence of interests.

My coordination problems such as (1)–(11) are among the situations at or near the pure coordination end of Schelling's spectrum. I do not want to require perfect coincidence of interests. For instance, I allowed imperfect coincidence of interests in those versions of example (1) in which Row-chooser and Column-chooser care somewhat where they go, though much less than they care whether they meet. We recall the payoff matrices of Figures 2 and 3 (pp. 10–11). In several squares, the payoffs are not quite equal. No linear rescaling of either matrix could make them equal in every square at once.

I want, however, to confine my attention to situations in which coincidence of interest predominates: that is, in which the differences between different agents' payoffs in any one square (perhaps after suitable linear rescaling) are small compared to some of the differences between payoffs in different squares. So they are in the matrices of Figures 2 and 3; the largest difference within one square is .5, whereas the largest difference between payoffs in different squares is 1.5.

An equilibrium, we recall, is a combination in which no one would have been better off had he alone acted otherwise. Let me define a *coordination equilibrium* as a combination in which no one would have been better off had *any one* agent alone acted otherwise, either himself or someone else. Coordination equilibria are equilibria, by the definitions. Equilibria in games of pure coordination are always coordination equilibria, since the agents' interests coincide perfectly. Any game of pure coordination has at least one coordination equilibrium, since it has at least one outcome that is best for all. But coordination equilibria are by no means confined to games of pure coordination. They are common in situations with mixed opposition and coincidence of interests. They can occur even in games of pure conflict: ⟨*R1*, *C1*⟩ in Figure 7 is a coordination equilibrium.

	C1	C2
R1	0 / 0	0 / 0
R2	0 / 0	−1 / 1

Figure 7

Most versions of our sample coordination problems are not games of pure coordination; but they all have coordination equilibria. We have noticed that the versions of the meeting-place problem shown in Figures 2 and 3 are not games of pure coordination; but their equilibria—⟨*R1*, *C1*⟩, ⟨*R2*, *C2*⟩, and ⟨*R3*, *C3*⟩ in both versions—are coordination equilibria.

This is not to say that *all* the equilibria in a coordination problem must be coordination equilibria. Take still another version of example (1). Suppose there is a fourth place, *P4*. Row-chooser and Column-chooser both like to go to *P4* alone, but a meeting at *P4* would detract from their enjoyment of going to *P4* and *P4* would be of little use as a meeting place. So we have the matrix shown in Figure 8, with the usual coordination equilibria ⟨*R1*, *C1*⟩, ⟨*R2*, *C2*⟩, ⟨*R3*, *C3*⟩ and a new *non*coordination equilibrium ⟨*R4*, *C4*⟩. It is an equilibrium because Row-chooser prefers it to ⟨*R1*, *C4*⟩, ⟨*R2*, *C4*⟩, or ⟨*R3*, *C4*⟩, and Column-chooser prefers it to ⟨*R4*, *C1*⟩, ⟨*R4*, *C2*⟩, or ⟨*R4*, *C3*⟩. It is not a coordination equilibrium because not all—in fact, none—of these preferences are shared by Row-chooser and Column-chooser. Yet this version of (1) does not seem significantly different from the others. The situation still has that distinctive character which I introduced by means of my eleven examples. So let us tolerate noncoordination equilibria in coordination problems.

	C1	C2	C3	C4
R1	1 meet 1	0 0	0 0	.5 0
R2	0 0	1 meet 1	0 0	.5 0
R3	0 0	0 0	1 meet 1	.5 0
R4	0 .5	0 .5	0 .5	.2 meet .2

Figure 8

All my sample coordination problems have two or more different coordination equilibria. This multiplicity is important to the distinctive character of coordination problems and ought to be included in their definition. If there is no considerable conflict of interest, the task of reaching a unique coordination equilibrium is more or less trivial. It will be reached if the nature of the situation is clear enough so that everybody makes the best choice given his expectations, everybody expects everybody else to make the best choice given his expectations, and so on. These conditions do not ensure coordination if there are multiple coordination equilibria, as we shall see.

Many of the situations with unique coordination equilibria are still

	C1	C2
R1	−8 −8	−10 −1
R2	−1 −10	−2 −2

Figure 9

more trivial (and more deserving of exclusion). For instance, any situation in which all the agents have [*strictly*] *dominant* choices—actions they prefer no matter what the others do—can have only one equilibrium (and *a fortiori* only one coordination equilibrium), namely, the combination of dominant choices. A combination of dominant choices must be an equilibrium; but it might not be a coordination equilibrium, as in the well-known Prisoner's Dilemma, shown in Figure 9, in which *R1* and *C1* (treacherous confession, in the usual story) are dominant and their combination ⟨*R1*, *C1*⟩ is a noncoordination equilibrium.

We might guess that there is dominance in *any* game of pure coordination with a unique equilibrium: that all, or at least some, agents have dominant, or at least dominated, choices. (A [*strictly*] *dominated* choice is one such that, no matter how the others choose, you could have made some other choice that would have been better. If one choice is dominant, another must be dominated; but not vice versa, since *which* other choice would have been better for you may depend on how the others chose.) There is this much truth in the guess: in any finite two-person game of pure coordination with a unique equilibrium, at least one action of one of the agents is dominated. Proof:

> Let $P(\langle Rj, Ck\rangle)$ represent the payoff at the combination $\langle Rj, Ck\rangle$, equal for Row-chooser and Column-chooser.
>
> Take a suitable game with *m* rows and *n* columns. Assume without loss of generality that its rows and columns are so arranged that for any combination $\langle Ri, Ci\rangle$ on the diagonal and any combination $\langle Rj, Ck\rangle$ such that $j \geq i$ and $k \geq i$, $P(\langle Rj, Ck\rangle) \leq P(\langle Ri, Ci\rangle)$. In particular, $\langle R1, C1\rangle$ must be the unique equilibrium, and $P(\langle R1, C1\rangle)$ must exceed every other payoff in the game.
>
> If $\langle R1, C1\rangle$ is the only diagonal combination that is either a row-maximum or a column-maximum, then *Rm* (if $m \geq n$) or *Cn* (if $n \geq m$) must be dominated.

Otherwise let $\langle Ra, Ca\rangle$, $a \neq 1$, be the rightmost diagonal combination which is either a row-maximum or a column-maximum. It is not both, since it is not an equilibrium. Suppose without loss of generality that it is a row-maximum.

Unless Ra is strictly dominated, there is a column-maximum on Ra; let $\langle Ra, Cb\rangle$ be the rightmost one. $\langle Ra, Cb\rangle$ is not a row-maximum since it is not an equilibrium, so $P(\langle Ra, Ca\rangle) > P(\langle Ra, Cb\rangle)$.

Unless Cb is strictly dominated, there is a row-maximum on Cb; let $\langle Ra', Cb\rangle$ be the lowest one. Since $\langle Ra, Cb\rangle$ is a column-maximum, $P(\langle Ra, Cb\rangle) \geq P(\langle Ra', Cb\rangle)$, so $P(\langle Ra, Ca\rangle) > P(\langle Ra', Cb\rangle)$.

Unless Ra' is strictly dominated there is a column-maximum on Ra'; let $\langle Ra', Cb'\rangle$ be the rightmost one; $P(\langle Ra, Ca\rangle) > P(\langle Ra', Cb'\rangle)$.

Unless Cb' is strictly dominated, there is a row-maximum on Cb'; let $\langle Ra'', Cb'\rangle$ be the lowest one; $P(\langle Ra, Ca\rangle) > P(\langle Ra'', Cb'\rangle)$.

Unless Ra'' is strictly dominated, there is a column-maximum on Ra''; let $\langle Ra'', Cb''\rangle$ be the rightmost one; $P(\langle Ra, Ca\rangle) > P(\langle Ra'', Cb''\rangle)$. And so on.

If $\langle Rj, Ci\rangle$ is a column-maximum and $P(\langle Ra, Ca\rangle) > P(\langle Rj, Ci\rangle)$, then $\langle Rj, Ci\rangle$ is above the diagonal. For otherwise $j \geq i$, so $P(\langle Rj, Ci\rangle) \leq P(\langle Ri, Ci\rangle)$. And since $\langle Rj, Ci\rangle$ is a column-maximum, $P(\langle Rj, Ci\rangle) = P(\langle Ri, Ci\rangle)$. Then $\langle Ri, Ci\rangle$ is also a column-maximum, and it is to the right of $\langle Ra, Ca\rangle$ since $P(\langle Ra, Ca\rangle) > P(\langle Ri, Ci\rangle)$. But that is contrary to our choice of $\langle Ra, Ca\rangle$.

In particular: $\langle Ra, Cb\rangle$, $\langle Ra', Cb'\rangle$, $\langle Ra'', Cb''\rangle$, etc. are above the diagonal.

By a parallel argument, if $\langle Rj, Ci\rangle$ is a row-maximum and $P(\langle Ra, Ca\rangle) > P(\langle Rj, Ci\rangle)$, then $\langle Rj, Ci\rangle$ is below the diagonal. In particular: $\langle Ra', Cb\rangle$, $\langle Ra'', Cb'\rangle$, etc. are below the diagonal.

Therefore the sequence of combinations we were constructing moves back and forth across the diagonal, as shown in Figure 10, so that $a < a' < a''$. . . and $b < b' < b''$. . . . Since the game is finite, these sequences terminate, which can happen only if one of *Ra*, *Cb*, *Ra'*, *Cb'*, *Ra''*, *Cb''* etc. is strictly dominated.

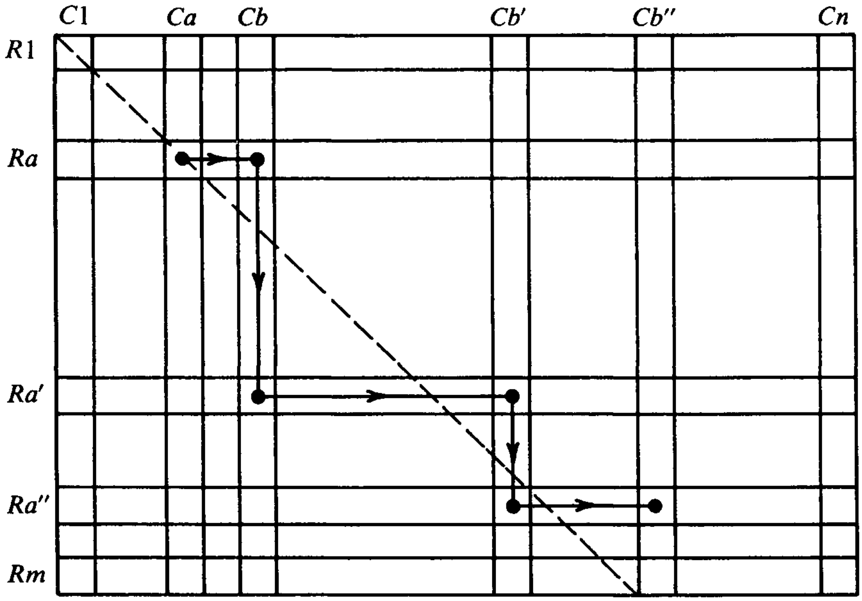

Figure 10

The deletion of a dominated action in a finite two-person game of pure coordination with a unique equilibrium leaves a new game, which is itself a finite two-person game of pure coordination with a unique equilibrium. So the deletion can be repeated. By successive deletions of dominated actions, the game is transformed into a situation that is patently trivial because Row-chooser and Column-chooser each have only one available action. The outcome is determined by the fact that everybody ignores dominated actions, everybody expects everybody else to ignore dominated actions, and so on.

The result just proved cannot, unfortunately, be strengthened in any of the ways one might hope. It does not carry over to infinite two-person games; Figure 11 is a counterexample. It does not carry over to finite three-person games; Figure 12 is a counterexample.

	C1	C2	C3	C4	C5 ...
R1	64, 64	0, 0	0, 0	0, 0	0, 0
R2	32, 32	16, 16	0, 0	0, 0	0, 0
R3	0, 0	8, 8	4, 4	0, 0	0, 0
R4	0, 0	0, 0	2, 2	1, 1	0, 0
R5	0, 0	0, 0	0, 0	.5, .5	.25, .25
⋮					

Figure 11

L1	C1	C2
R1	5, 5, 5	0, 0, 0
R2	0, 0, 0	1, 1, 1

L2	C1	C2
R1	4, 4, 4	3, 3, 3
R2	1, 1, 1	2, 2, 2

Figure 12

(Call the third agent's choices *levels L1* and *L2*; write his payoffs in the centers of the squares.) It cannot be strengthened for the finite two-person case; Figure 13 is an example with no dominant action and only a single dominated action (and that one is dominated only

	C1	C2	C3
R1	16 16	0 0	0 0
R2	8 8	4 4	0 0
R3	0 0	2 2	1 1

Figure 13

by all the alternatives together). Therefore we cannot say that dominance is responsible for all cases of unique equilibria in games of pure coordination.

To exclude trivial cases, a coordination problem must have more than one coordination equilibrium. But that requirement is not quite strong enough. Figure 14 shows two matrices in which, sure enough,

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

	C1	C2	C3
R1	1 1	1 1	.2 0
R2	1 1	1 1	.5 .2
R3	0 .5	0 0	0 0

Figure 14

there are multiple coordination equilibria (two on the left, four on the right). Yet there is still no need for either agent to base his choice on his expectation about the other's choice. There is no need for them to try for the same equilibrium—no need for coördination—since if they try for different equilibria, some equilbrium will nevertheless be reached. These cases exhibit another kind of triviality, akin to the triviality of a case with a unique coordination equilibrium.

A combination is an equilibrium if each agent likes it *at least as well as* any other combination he could have reached, given the others' choices. Let us call it a *proper* equilibrium if each agent likes it *better than* any other combination he could have reached, given the others' choices. In a two-person matrix, for instance, a proper equilibrium is preferred by Row-chooser to all other combinations in its column, and by Column-chooser to all other combinations in its row. In the matrices in Figure 14, there are multiple coordination equilibria, but all of them are improper.

There is no need to stipulate that all equilibria in a coordination problem must be proper; it seems that the matrix in Figure 15 ought to be counted as essentially similar to our clear examples of coordina-

	C1	C2	C3
R1	2, 2	0, 0	0, 0
R2	0, 0	2, 2	0, 0
R3	0, 0	1, 1	1, 1

Figure 15

tion problems, despite the impropriety of its equilibrium ⟨*R3, C3*⟩. The two proper coordination equilibria—⟨*R1, C1*⟩ and ⟨*R2, C2*⟩—are sufficient to keep the problem nontrivial. I stipulate instead that a coordination problem must contain at least two proper coordination equilibria.

This is only one—the strongest—of several defensible restrictions. We might prefer a weaker restriction that would not rule out matrices like those in Figure 16. But a satisfactory restriction would be com-

	C1	C2	C3
R1	1, 1	0, 0	0, 0
R2	0, 0	1, 1	1, 1
R3	0, 0	1, 1	1, 1

	C1	C2	C3	C4
R1	1, 1	1, 1	0, 0	0, 0
R2	1, 1	1, 1	0, 0	0, 0
R3	0, 0	0, 0	1, 1	1, 1
R4	0, 0	0, 0	1, 1	1, 1

Figure 16

plicated and would entail too many qualifications later. And situations like those of Figure 16 can be rescued even under the strong restriction we have adopted. Let *R2′* be the disjunction of *R2* and *R3*, and *C2′* the disjunction of *C2* and *C3* in the left-hand matrix. Then the same situation can be represented by the new matrix in Figure 17, which does have two proper coordination equilibria. The

	C1	C2′
R1	1, 1	0, 0
R2′	0, 0	1, 1

Figure 17

right-hand matrix can be consolidated in a similar way. But matrices like the one in Figure 18, which are ruled out by the strong restriction, and ought to be ruled out, cannot be rescued by any such consolidation.

	C1	C2	C3
R1	1, 1	1, 1	0, 0
R2	0, 0	1, 1	1, 1
R3	0, 0	0, 0	1, 1

Figure 18

To sum up: Coordination problems—situations that resemble my eleven examples in the important respects in which they resemble one another[4]—are situations of interdependent decision by two or more agents in which coincidence of interest predominates and in which there are two or more proper coordination equilibria. We could also say—though less informatively than one might think—that they are situations in which, relative to *some* classification of actions, the agents have a common interest in all doing the same one of several alternative actions.

3. Solving Coordination Problems

Agents confronted by a coordination problem may or may not succeed in each acting so that they reach one of the possible coordination equilibria. They might succeed just by luck, although some of them choose without regard to the others' expected actions (doing so perhaps because they cannot guess what the others will do, perhaps because the chance of coordination seems so small as to be negligible).

[4]See Michael Slote, "The Theory of Important Criteria," *Journal of Philosophy,* 63 (1966), pp. 211–224. Slote shows that we commonly introduce a class by means of examples and take the defining features of the class to be those distinctive features of our examples which seem important for an understanding of their character. That is what I take myself to be doing here and elsewhere.

But they are more likely to succeed—if they do—through the agency of a system of suitably concordant mutual expectations. Thus in example (1) I may go to a certain place because I expect you to go there, while you go there because you expect me to; in example (2) I may call back because I expect you not to, while you do not because you expect me to; in example (4) each of us may drive on the right because he expects the rest to do so; and so on. In general, each may do his part of one of the possible coordination equilibria because he expects the others to do theirs, thereby reaching that equilibrium.

If an agent were completely confident in his expectation that the others would do their parts of a certain proper coordination equilibrium, he would have a decisive reason to do his own part. But if—as in any real case—his confidence is less than complete, he must balance his preference for doing his part if the others do theirs against his preferences for acting otherwise if they do not. He has a decisive reason to do his own part if he is *sufficiently* confident in his expectation that the others will do theirs. The degree of confidence which is sufficient depends on all his payoffs and sometimes on the comparative probabilities he assigns to the different *ways* the others might not all do their parts, in case not all of them do. For instance, in the coordination problem shown in Figure 19, Row-chooser should

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

Figure 19

do his part of the coordination equilibrium ⟨*R1*, *C1*⟩ by choosing *R1* if he has more than .5 confidence that Column-chooser will do his part by choosing *C1*. But in the coordination problems shown in Figure 20, Row-chooser should choose R1 only if he has more

	C1	C2
R1	1, 1	−8, −8
R2	0, 0	1, 1

	C1	C2
R1	1, 1	0, 0
R2	0, 0	9, 9

	C1	C2
R1	3, 3	−26, −26
R2	0, 0	1, 1

Figure 20

than .9 confidence that Column-chooser will choose *C1*. If he has, say, .8 confidence that Column-chooser will choose *C1*, he would do better to choose *R2*, sacrificing his chance to achieve coordination at ⟨*R1*, *C1*⟩ in order to hedge against the possibility that his expectation was wrong. And in the coordination problem shown in Figure 21, Row-chooser might be sure that if Column-chooser fails to do

	C1	C2	C3
R1	1, 1	0, 0	−8, −8
R2	0, 0	1, 1	9, 9

Figure 21

his part of ⟨*R1*, *C1*⟩, at least he will choose *C2*, not *C3*; if so, Row-chooser should choose *R1* if he has more than .5 confidence that Column-chooser will choose *C1*. Or Row-chooser might think that if Column-chooser fails to choose *R1*, he is just as likely to choose *C3* as to choose *C2*; if so, Row-chooser should choose *R1* only if he has more than .9 confidence that Column-chooser will choose *C1*. Or Row-chooser might be sure that if Column-chooser does not choose *C1*, he will choose *C3* instead; if so, Row-chooser's minimum sufficient degree of confidence is about .95. The strength of concordant expectation needed to produce coordination at a certain equilibrium is a measure of the difficulty of achieving coordination there, since however the concordant expectations are produced, weaker expectations will be produced more easily than stronger ones. (We can imagine cases in which so much mutual confidence is required to achieve coordination at an equilibrium that success is impossible. Imagine that a millionaire offers to distribute his fortune equally among a thousand men if each sends him $10; if even one does not, the millionaire will keep whatever he is sent. I take it that no matter what the thousand do to increase their mutual confidence, it is a practical certainty that the millionaire will not have to pay up. So if I am one of the thousand, I will keep my $10.)

We may achieve coordination by acting on our concordant expectations about each other's actions. And we may acquire those expectations, or correct or corroborate whatever expectations we already have, by putting ourselves in the other fellow's shoes, to the best of our ability. If I know what you believe about the matters of fact that determine the likely effects of your alternative actions, and if I know your preferences among possible outcomes and I know that you possess a modicum of practical rationality, then I can replicate your practical reasoning to figure out what you will probably do, so that I can act appropriately.

In the case of a coordination problem, or any other problem of interdependent decision, one of the matters of fact that goes into determining the likely effects of your alternative actions is my own action. In order to figure out what you will do by replicating your practical reasoning, I need to figure out what *you* expect *me* to do.

I know that, just as I am trying to figure out what you will do by replicating your reasoning, so you may be trying to figure out what I will do by replicating my reasoning. This, like anything else you might do to figure out what I will do, is itself part of your reasoning. So to replicate your reasoning, I may have to replicate your attempt to replicate my reasoning.

This is not the end. I may reasonably expect *you* to realize that, unless I already know what you expect me to do, I may have to try to replicate your attempt to replicate my reasoning. So I may expect you to try to replicate my attempt to replicate your attempt to replicate my reasoning. So my own reasoning may have to include an attempt to replicate your attempt to replicate my attempt to replicate your attempt to replicate my reasoning. And so on.

Before things get out of hand, it will prove useful to introduce the concept of *higher-order expectations*, defined by recursion thus:

> A first-order expectation about something is an ordinary expectation about it.
>
> An $(n + 1)$th-order expectation about something $(n \geq 1)$ is an ordinary expectation about someone else's nth-order expectation about it.

For instance, if I expect you to expect that it will thunder, then I have a second-order expectation that it will thunder.

Whenever I replicate a piece of your practical reasoning, my second-order expectations about matters of fact, together with my first-order expectations about your preferences and your rationality, justify me in forming a first-order expectation about your action. In the case of problems of interdependent decision—for instance, coordination problems—some of the requisite second-order expectations must be about my own action.

Consider our first sample coordination problem: a situation in which you and I want to meet by going to the same place. Suppose that after deliberation I decide to come to a certain place. The fundamental practical reasoning which leads me to that choice is shown in Figure 22. (In all diagrams of this kind, heavy arrows represent implications; light arrows represent causal connections between the mental states or actions of a rational agent.) And if my premise for this reasoning—my expectation that you will go there—was obtained by replicating your reasoning, my replication is shown in Figure 23. And if my premise for this replication—my expectation that you will expect me to go there—was obtained by replicating your replication of my reasoning, my replication of your replication is shown in Figure 24. And so on. The whole of my reasoning (simplified by disregarding the rationality premises) may be represented as in

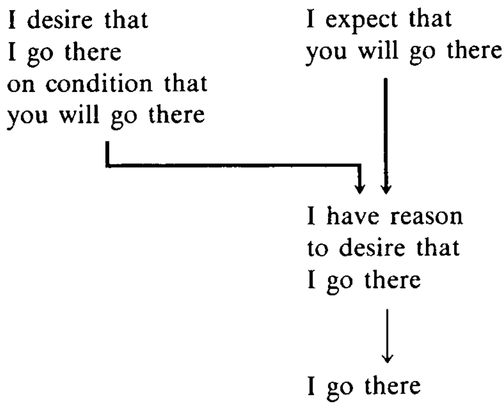

Figure 22

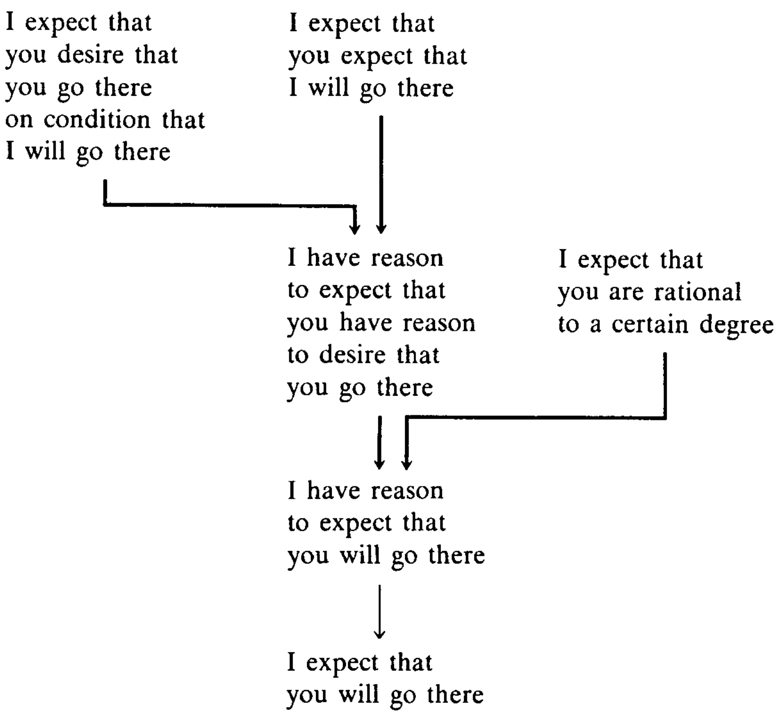

Figure 23

I expect that
you expect that
I desire that
I go there
on condition that
you will go there

I expect that
you expect that
I expect that
you will go there

I have reason
to expect that
you have reason
to expect that
I have reason
to desire that
I go there

I expect that
you expect that
I am rational
to a certain degree

I have reason
to expect that
you have reason
to expect that
I will go there

I expect that
you are rational
to a certain degree

I have reason
to expect that
you expect that
I will go there

I expect that
you expect that
I will go there

Figure 24

Figure 25 for whatever finite number of stages it may take for me to use whatever higher-order expectations may be available to me regarding our actions and our conditional preferences. Replications are nested to some finite depth: my reasoning (outer boundary) con-

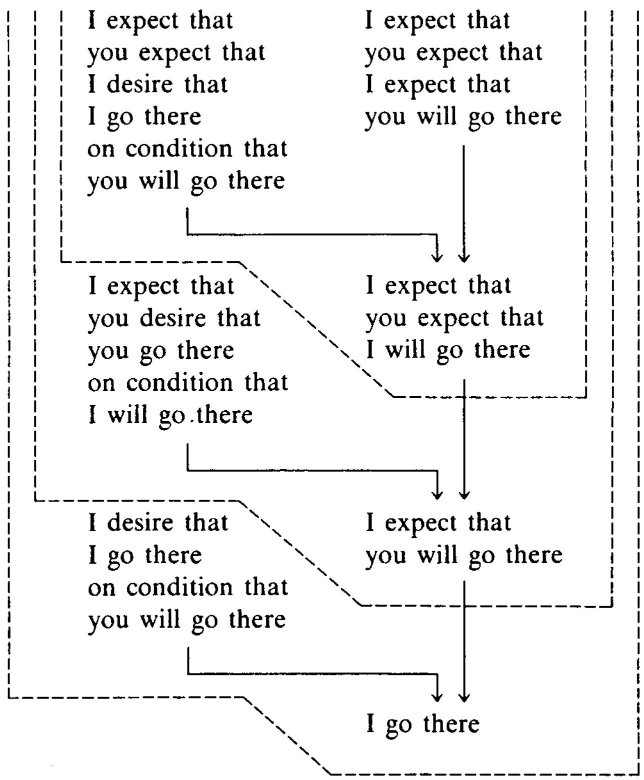

Figure 25

tains a replication of yours (next boundary), which contains a replication of your replication of mine (next boundary), and so on.

So if I somehow happen to have an *n*th-order expectation about action in this two-person coordination problem, I may work outward through the nested replications to lower- and lower-order expectations about action. Provided I go on long enough, and provided all the needed higher-order expectations about preferences and rationality are available, I eventually come out with a first-order expectation about your action—which is what I need in order to know how I should act.

Clearly a similar process of replication is possible in coordination problems among more than two agents. In general, my higher-order expectations about something are my expectations about x_1's expectations about x_2's expectations . . . about it. (The sequence x_1, x_2 . . . may repeat, but x_1 cannot be myself and no one can occur twice in immediate succession.) So when m agents are involved, I can have as many as $(m - 1)^n$ different nth-order expectations about anything, corresponding to the $(m - 1)^n$ different admissible sequences of length n. Replication in general is ramified: it is built from stages in which $m - 1$ of my various $(n + 1)$th-order expectations about action, plus ancillary premises, yield one of my nth-order expectations about action. I suppressed the ramification by setting $m = 2$, but the general case is the same in principle.

Note that replication is *not* an interaction back and forth between people. It is a process in which *one* person works out the consequences of his beliefs about the world—a world he believes to include other people who are working out the consequences of their beliefs, including their belief in other people who . . . By our interaction in the world we acquire various high-order expectations that can serve us as premises. In our subsequent reasoning we are windowless monads doing our best to mirror each other, mirror each other mirroring each other, and so on.

Of course I do not imagine that anyone will solve a coordination problem by first acquiring a seventeenth-order expectation from somewhere and then sitting down to do his replications. For one thing, we rarely do have expectations of higher order than, say, fourth. For another thing, any ordinary situation that could justify a high-order expectation would also justify low-order expectations directly, without recourse to nested replications.

All the same, given the needed ancillary premises, an expectation of arbitrarily high order about action does give an agent *one* good reason for a choice of action. The one may, and normally will, be one reason among the many which jointly suffice to justify his choice. Suppose the agent is originally justified somehow in having expectations of several orders about his own and his partners' actions. And suppose the ancillary premises are available. Then each of his original expectations independently gives him a reason to act one way or another. If he is lucky, all these independent reasons will be reasons for the same action.[5] Then that action is strongly, because redundantly, justified; he has more reason to do it than could have been provided by any one of his original expectations by itself.

I said earlier that coordination might be rationally achieved with the aid of concordant mutual expectations about action. We have seen that these may be derived from first- and higher-order expectations about action, preferences, and rationality. So we generalize: coordination may be rationally achieved with the aid of a system of concordant mutual expectations, of first or higher orders, about the agents' actions, preferences, and rationality.

The more orders of expectation about action contribute to an agent's decision, the more independent justifications the agent will have; and insofar as he is aware of those justifications, the more firmly his choice will be determined. Circumstances that will help to solve a coordination problem, therefore, are circumstances in which the agents become justified in forming mutual expectations belonging to a concordant system. And the more orders, the better.

In considering how to solve coordination problems, I have postponed the answer that first comes to mind: by agreement. If the agents can communicate (without excessive cost), they can ensure a common understanding of their problem by discussing it. They can choose a coordination equilibrium—an arbitrary one, or one especially good for some or all of them, or one they can reach without too much

[5] Michael Scriven, in "An Essential Unpredictability in Human Behavior," *Scientific Psychology: Principles and Approaches*, ed. B. B. Wolman (New York: Basic Books, 1965), has discussed mutual replication of practical reasoning between agents in a game of conflict who want *not* to conform to each other's expectations. There is a cyclic alternation: from my $(n + 4)$th-order expectation that I will go to Minsk to my $(n + 3)$th-order expectation that you will go to Pinsk to my $(n + 2)$th-order expectation that I will go to Pinsk to my $(n + 1)$th-order expectation that you will go to Minsk to my nth-order expectation that I will go to Minsk . . . Scriven notices that we cannot both act on complete and accurate replications of each other's reasoning. He takes this to prove human unpredictability. But perhaps it simply proves that the agents cannot both have enough time to finish their replications, since the time either needs increases with the time the other uses. See David Lewis and Jane Richardson, "Scriven on Human Unpredictability," *Philosophical Studies*, 17 (1966), pp. 69-74.

mutual confidence. And each can assure the rest that he will do his part of the chosen equilibrium. Coordination by means of an agreement is not, of course, an alternative to coordination by means of concordant mutual expectations. Rather, agreement is one means of producing those expectations. It is an especially effective means, since it produces strong concordant expectations of several orders.

Suppose you and I want to meet tomorrow; today we happen to meet, and we make an appointment. Each thereby gives evidence of his interest in going where the other goes and of his intention to go to a certain place. By observing this evidence, we form concordant first-order expectations about each other's preferences and action. By observing each other observing it, we may also form concordant second-order expectations. By observing each other observing each other observing it, we may even form concordant third-order expectations. And so on; not forever, of course, but limited by the amount of reasoning we do and the amount we ascribe to each other—perhaps one or two steps more. The result is a system of concordant mutual expectations of several orders, conducive to coordination by means of replication.

The agents' agreement might be an exchange of formal or tacit promises. But it need not be. Even a man whose word is his bond can remove the promissory force by explicit disavowal, if not otherwise. An exchange of declarations of present intention will be good enough, even if each explicitly retains his right to change his plans later. No one need bind himself to act against his own interest. Rather, it will be in the interest of each to do just what he has led the others to expect him to do, since that action will be best for him if the others act on their expectations.

If one does consider himself bound by a promise, he has a second, independent incentive. His payoffs are modified, since he has attached the onus of promise breaking to all but one choice. Indeed, he may modify his payoffs so much by promising that the situation is no longer a coordination problem at all. For instance, the agent's promised action might become his dominant choice: he might wish to keep his promise no matter what, coordination or no coordination.

If such a strong promise is made publicly, the others will know that they must go along with the one who has promised, for they know what he will do. Such forceful promising is a way of getting rid of coordination problems, not a way of solving them.

Explicit agreement is an especially good and common means to coordination—so much so that we are tempted to speak of coordination otherwise produced as *tacit* agreement. But agreement (literally understood) is not the only source of concordant expectations to help us solve our coordination problems. We do without agreement by choice if we find ourselves already satisfied with the content and strength of our mutual expectations. We do without it by necessity if we have no way to communicate, or if we can communicate only at a cost that outweighs our improved chance of coordination (say, if we are conspirators being shadowed).

Schelling has experimented with coordination problems in which the agents cannot communicate. His subjects know only that they share a common understanding of their problem—for instance, they may get instructions describing their problem and stating that everyone gets the same instructions. It turns out that sophisticated subjects in an experimental setting can often do very well—much better than chance—at solving novel coordination problems without communicating. They try for a coordination equilibrium that is somehow *salient*: one that stands out from the rest by its uniqueness in some conspicuous respect. It does not have to be uniquely *good*; indeed, it could be uniquely bad. It merely has to be unique in some way the subjects will notice, expect each other to notice, and so on. If different coordination equilibria are unique in different conspicuous ways, the subjects will need to be alike in the relative importance they attach to different respects of comparison; but often they are enough alike to solve the problem.

How can we explain coordination by salience? The subjects might all tend to pick the salient as a last resort, when they have no stronger ground for choice. Or they might expect each other to have that tendency, and act accordingly; or they might expect each other to expect each other to have that tendency and act accordingly, and act accordingly; and so on. Or—more likely—there might be a mixture of these. Their first- and higher-order expectations of a tendency to pick the salient as a last resort would be a system of concordant expectations capable of producing coordination at the salient equilibrium.

If their expectations did produce coordination, it would not matter whether anyone really would have picked the salient as a last resort. For each would have had a good reason for his choice, so his choice would not have been a last resort.

Thus even in a novel coordination problem—which is an extreme case—the agents can sometimes obtain the concordant expectations they need without communicating. An easier, and more common, case is that of a *familiar* coordination problem without communication. Here the agents' source of mutual expectations is precedent: acquaintance with past solved instances of their present coordination problem.

4. Convention

Let us start with the simplest case of coordination by precedent and generalize in various ways. In this way we shall meet the phenomenon I call *convention*, the subject of this book.

Suppose we have been given a coordination problem, and we have reached some fairly good coordination equilibrium. Given exactly the same problem again, perhaps each of us will repeat what he did before. If so, we will reach the same solution. If you and I met yesterday—by luck, by agreement, by salience, or however—and today we find we must meet again, we might both go back to yesterday's meeting place, each hoping to find the other there. If we were cut off on the telephone and you happened to call back as I waited, then if we are cut off again in the same call, I will wait again.

We can explain the force of precedent just as we explained the force of salience. Indeed, precedent is merely the source of one important kind of salience: conspicuous uniqueness of an equilibrium because we reached it last time. We may tend to repeat the action that succeeded before if we have no strong reason to do otherwise. Whether or not any of us really has this tendency, we may somewhat expect each other to have it, or expect each other to expect each other to have it, and so on—that is, we may each have first- and higher-order expectations that the others will do their parts of the old coordination equilibrium, unless they have reason to act otherwise. Each one's expectation that the others will do their parts, strengthened perhaps by replication using his higher-order expectations, gives him some reason to do his own part. And if his original expectations of some order or other were strong enough, he will have a decisive reason to do his part. So he will do it.

I have been supposing that we are given a coordination problem, and then given the same problem again. But, of course, we could never be given exactly the same problem twice. There must be this difference at least: the second time, we can draw on our experience with the first. More generally, the two problems will differ in several independent respects. We cannot do exactly what we did before. Nothing we could do this time is exactly like what we did before—like it in every respect—because the situations are not exactly alike.

So suppose not that we are given the original problem again, but rather that we are given a new coordination problem analogous somehow to the original one. Guided by whatever analogy we notice, we tend to follow precedent by trying for a coordination equilibrium in the new problem which uniquely corresponds to the one we reached before.

There might be alternative analogies. If so, there is room for ambiguity about what would be following precedent and doing what we did before. Suppose that yesterday I called you on the telephone and I called back when we were cut off. Today you call me and we are cut off. We have a precedent in which I called back and a precedent—the same one—in which the original caller called back. But this time you are the original caller. No matter what I do this time, I do something analogous to what we did before. Our ambiguous precedent does not help us.

In fact, there are always innumerable alternative analogies. Were it not that we happen uniformly to notice some analogies and ignore others—those we call "natural" or "artificial," respectively—precedents would always be completely ambiguous and worthless. *Every* coordination equilibrium in our new problem (every other combination, too) corresponds uniquely to what we did before under *some* analogy, shares *some* distinctive description with it alone. Fortunately, most of the analogies are artificial. We ignore them; we do not tend to let them guide our choice, nor do we expect each other to have any such tendency, nor do we expect each other to expect each other to, and so on. And fortunately we have learned that all of us will mostly notice the same analogies. That is why precedents can be unambiguous in practice, and often are. If we notice only one of the analogies between our problem and the precedent, or if one of those we notice seems far more conspicuous than the others, or even if several are conspicuous but they all happen to agree in indicating the same choice, then the other analogies do not matter. We are not in trouble unless conflicting analogies force themselves on our attention.

The more respects of similarity between the new problem and the precedent, the more likely it is that different analogies will turn out to agree, the less room there will be for ambiguity, and the easier it will be to follow precedent. A precedent in which I, the original caller, called back is ambiguous given a new problem in which you are the original caller—but not given a new problem in which I am again the original caller. That is why I began by pretending that the new problem was like the precedent in all respects.

Salience in general is uniqueness of a coordination equilibrium in a preeminently conspicuous respect. The salience due to precedent is no exception: it is uniqueness of a coordination equilibrium in virtue of its preeminently conspicuous analogy to what was done successfully before.

So far I have been supposing that the agents who set the precedent are the ones who follow it. This made sure that the agents given the second problem were acquainted with the circumstances and outcome of the first, and expected each other to be, expected each other to expect each other to be, and so on. But it is not an infallible way and not the only way. For instance, if yesterday I told you a story about people who got separated in the subway and happened to meet again at Charles Street, and today we get separated in the same way, we might independently decide to go and wait at Charles Street. It makes no difference whether the story I told you was true, or whether you thought it was, or whether I thought it was, or even whether I claimed it was. A fictive precedent would be as effective as an actual one in suggesting a course of action for us, and therefore as good a source of concordant mutual expectations enabling us to meet. So let us just stipulate that somehow the agents in the new problem are acquainted with the precedent, expect each other to be acquainted with it, and so on.

So far I have been supposing that we have a single precedent to follow. But we might have several. We might all be acquainted with a class of previous coordination problems, naturally analogous to our present problem and to each other, in which analogous coordination equilibria were reached. This is to say that the agents' actions conformed to some noticeable regularity. Since our present problem is suitably analogous to the precedents, we can reach a coordination equilibrium by all conforming to this same regularity. Each of us wants to conform to it if the others do; he has a *conditional preference* for conformity. If we do conform, the explanation has the familiar pattern: we tend to follow precedent, given no particular reason to do anything else; we expect that tendency in each other; we expect each other to expect it; and so on. We have our concordant first- and higher-order expectations, and they enable us to reach a coordination equilibrium.

It does not matter *why* coordination was achieved at analogous equilibria in the previous cases. Even if it had happened by luck, we could still follow the precedent set. One likely course of events would be this: the first case, or the first few, acted as precedent for the next, those for the next, and so on. Similarly, no matter how our precedents came about, by following them this time we add this case to the stock of precedents available henceforth.

Several precedents are better than one, not only because we learn by repetition but also because differences between the precedents help to resolve ambiguity. Even if our present situation bears conflicting natural analogies to any one precedent, maybe only one of these analogies will hold between the precedents; so we will pay attention only to that one. Suppose we know of many cases in which a cut-off telephone call was restored, and in every case it was the original caller who called back. In some cases I was the original caller, in some you were, in some neither of us was. Now we are cut off and I was the original caller. For you to call back would be to do something analogous—under one analogy—to what succeeded in some of the previous cases. But we can ignore that analogy, for under it the precedents disagree.

Once there are many precedents available, without substantial disagreement or ambiguity, it is no longer necessary for all of us to be acquainted with precisely the same ones. It is enough if each of us is acquainted with some agreeing precedents, each expects everyone else to be acquainted with some that agree with his, each expects everyone else to expect everyone else to be acquainted with some precedents that agree with his, etc. It is easy to see how that might happen: if one has often encountered cases in which coordination was achieved in a certain problem by conforming to a certain regularity, and rarely or never encountered cases in which it was not, he is entitled to expect his neighbors to have had much the same experience. If I have driven all around the United States and seen many people driving on the right and never one on the left, I may reasonably infer that almost everyone in the United States drives on the right, and hence that this man driving toward me also has mostly seen people driving on the right—even if he and I have not seen any of the *same* people driving on the right.

Our acquaintance with a precedent need not be very detailed. It is enough to know that one has learned of many cases in which coordination was achieved in a certain problem by conforming to a certain regularity. There is no need to be able to specify the time and place, the agents involved, or any other particulars; no need to be able to recall the cases one by one. I cannot cite precedents one by one in which people drove on the right in the United States; I am not sure I can cite even one case; nonetheless, I know very well that I have often seen cars driven in the United States, and almost always they were on the right. And since I have no reason to think I encountered an abnormal sample, I infer that drivers in the United States do almost always drive on the right; so anyone I meet driving in the United States will believe this just as I do, will expect me to believe it, and so on.

Coordination by precedent, at its simplest, is this: achievement of coordination by means of shared acquaintance with the achievement of coordination in a single past case exactly like our present coordination problem. By removing inessential restrictions, we have come to this: achievement of coordination by means of shared acquaintance with a *regularity* governing the achievement of coordination in a class of past cases which bear some conspicuous analogy to one another and to our present coordination problem. Our acquaintance with this regularity comes from our experience with some of its instances, not necessarily the same ones for everybody.

Given a regularity in past cases, we may reasonably extrapolate it into the (near) future. For we are entitled to expect that when agents acquainted with the past regularity are confronted by an analogous new coordination problem, they will succeed in achieving coordination by following precedent and continuing to conform to the same regularity. We come to expect conforming actions not only in past cases but in future ones as well. We acquire a general belief, unrestricted as to time, that members of a certain population conform to a certain regularity in a certain kind of recurring coordination problem for the sake of coordination.

Each new action in conformity to the regularity adds to our experience of general conformity. Our experience of general conformity in the past leads us, by force of precedent, to expect a like conformity in the future. And our expectation of future conformity is a reason to go on conforming, since to conform if others do is to achieve a coordination equilibrium and to satisfy one's own preferences. And so it goes—we're here because we're here because we're here because we're here. Once the process gets started, we have a metastable self-perpetuating system of preferences, expectations, and actions capable of persisting indefinitely. As long as uniform conformity is a coordination equilibrium, so that each wants to conform conditionally upon conformity by the others, conforming action produces expectation of conforming action and expectation of conforming action produces conforming action.

This is the phenomenon I call convention. Our first, rough, definition is:

> A regularity R in the behavior of members of a population P when they are agents in a recurrent situation S is a *convention* if and only if, in any instance of S among members of P,
>
> (1) everyone conforms to R;
> (2) everyone expects everyone else to conform to R;
> (3) everyone prefers to conform to R on condition that the others do, since S is a coordination problem and uniform conformity to R is a proper coordination equilibrium in S.

5. Sample Conventions

Chapter II will be devoted to improving the definition. But before we hide the concept beneath its refinements, let us see how it applies to examples. Consider some conventions to solve our sample coordination problems.

(1) If you and I must meet every week, perhaps at first we will make a new appointment every time. But after we have met at the same time and place for a few weeks running, one of us will say, "See you here next week," at the end of every meeting. Later still we will not say anything (unless our usual arrangement is going to be unsatisfactory next week). We will just both go regularly to a certain place at a certain time every week, each going there to meet the other and confident that he will show up. This regularity that has gradually developed in our behavior is a convention.

In this case the convention that sets our meeting place holds in the smallest possible population: just two people. In other cases, larger populations—perhaps with changing membership—have conventional meeting places. What makes a soda fountain, coffeehouse, or bar "in" is the existence of a convention in some social circle that it is the place to go when one wants to socialize. The man in the song—"Standing on a corner with a dollar in my hand / Looking for a woman who's looking for a man"— is standing on *that* corner in conformity to a convention among all the local prostitutes and their customers.

(2) In my hometown of Oberlin, Ohio, until recently all local telephone calls were cut off without warning after three minutes. Soon after the practice had begun, a convention grew up among Oberlin residents that when a call was cut off the original caller would call back while the called party waited. Residents usually conformed to this regularity in the expectation of conformity by the other party to the call. In this way calls were easily restored, to the advantage of all concerned. New residents were told about the convention or learned it through experience. It persisted for a decade or so until the cutoff was abolished.

Other regularities might have done almost as well. It could have been the called party who always called back, or the alphabetically first, or even the older. Any of these regularities could have become the convention if enough of us had started conforming to it. It would have been a bit less convenient than our actual convention; if the original caller calls back, he may still remember the number and he must at least know where to find it. But the inconveniences of another convention would not have outweighed the advantage of achieving a coordination equilibrium by calling back if and only if one's partner does not.

This example illustrates the possibility that (describing actions in any natural way) a conventional regularity may specify different actions under different conditions. In this case it specifies what we would naturally call different actions for agents involved in situation S in different *roles*. Except for *ad hoc* descriptions like "action in conformity to such-and-such regularity," the actions conforming to a conventional regularity do not have to share any common natural description. Therefore, when we speak of a convention to do an action A in a situation S, it must be understood that A may stand for an unnaturally complex action-description.

(3) If the two rowers in Hume's boat manage somehow to fall into a smooth rhythm and maintain it for a while, they "do it by an agreement or convention, though they have never given promises to each other." A regularity in their behavior—their rowing in that particular rhythm—persists because they expect it to be continued and they want to match their rhythms of rowing. "This common sense of interest . . . known to both . . . produces a suitable resolution and behavior" in which "the actions of each . . . have a reference to those of the other, and are performed upon the supposition that something is to be performed upon the other part."

This convention is peculiar. It holds in a very small population for a very short time—between two people for a few minutes—and the regularity is one we would find it very hard to describe, though we can easily catch on to it. But these oddities do not detract from its conventionality.

(4) We drive in the right lane on roads in the United States (or in the left lane on roads in Britain, Australia, Sweden before 1967, parts of Austria before a certain date, and elsewhere) because we do not want to drive in the same lane as the drivers coming toward us, and we expect them to drive on the right.

There is a complication: if we do not drive on the right, the highway patrol will catch us and we will be punished. So we have an independent incentive to drive on the right, and this second incentive is independent of how the others drive. But it makes no important difference. If I expected the others to be on the left, I would be there too, highway patrol or no highway patrol. My preference for driving on the same side as the others outweighs any incentive the highway patrol may give me to drive on the right. And so it is for almost everyone else, I am sure. The highway patrol modifies the payoffs in favor of driving on the right; but there are still two different coordination equilibria. The punishments are superfluous if they agree with our convention, are outweighed if they go against it, are not decisive either way, and hence do not make it any less conventional to drive on the right. The same goes for other considerations favoring one coordination equilibrium over the other: the fact that our cars have left-hand drive, the fact that we are mostly right-handed, and so on.

(5) If four men who camp together find that often they waste effort by covering the same ground in search of firewood, they may get fed up and agree once and for all: let Morgan look to the north, Jones to the east, Owen to the south, Griffith to the west. From that day on, each goes his proper way without further discussion. A regularity has begun by explicit agreement. At first, perhaps, it persists because each man feels bound by his promise and takes no account of the advantages of keeping it or breaking it. But years pass. They forget that they agreed. Morgan is replaced by Thomas, who never heard of the agreement and never promised anything. Yet whenever they need firewood each still goes off in his proper direction, because he knows that is how to have the ground to himself. As the force of their original promises fades away, the regularity in their behavior becomes a convention.

(6) Wanting to attend parties dressed as the others will be dressed, we wear whatever is conventional dress for the occasion; in picking our clothes we act in conformity to a convention of our social circle. By means of a conditional conventional regularity specifiying the style of clothes worn in various circumstances, we satisfy our common interest in being dressed alike.

But we must distinguish two cases. If each of us wants to dress like the majority and wants everyone else to dress like the majority too, then we achieve a coordination equilibrium when we all dress alike: our regularity is a genuine convention. Suppose, however, that many of us are nasty people who want to dress like the majority but also want to have a differently dressed minority to sneer at. We still achieve an equilibrium when we all dress alike, but it is not a coordination equilibrium: nobody wishes he himself had dressed otherwise, but the nasty ones wish that a few other people—say, their worst enemies—had dressed otherwise. The regularity whereby we achieve this equilibrium is not a genuine convention by my definition, because the element of conflict of interest prevents it from being a means of reaching a *coordination* equilibrium.

It may not be obvious that our regularities of dress should not be called conventions if there are many people who want to see them violated. But when our analysis has shown us how the presence of substantial conflict makes a disanalogy between this case and other clear cases of convention, and makes an important analogy between this case and clear cases of nonconvention like the one to be examined in Chapter III.5, I think we ought to end up agreeing with the analysis even against our first impressions. If the reader disagrees, I can only remind him that I did not undertake to analyze anyone's concept of convention but mine.

(7) If we are contented oligopolists who want to maintain a uniform but fluctuating price for our commodity, we dare not make any explicit agreement on prices; that would be a conspiracy in restraint of trade. But we can come to a tacit understanding—that is, a convention—by our ways of responding to each others' prices. We might, for instance, start to follow a price leader: one firm that takes the initiative in changing prices, with due care to set a price in the range that is satisfactory to all of us.

In this example, it becomes seriously artificial to divide our continuous activity into a sequence of separate analogous coordination problems, related only by force of precedent. (The difficulty will reappear in examples [9], [10], and [11]; it was present somewhat in [3] and [4].) We can actually set or reconsider prices at any time. How long is a coordination problem? Pretend, already idealizing, that we set our prices every morning and cannot change them later in the day. Then each business day is a coordination problem. But a day is too short. Our customers take more than a day to shop around;

they compare my price for today with yours for yesterday and someone else's for tomorrow. We are leaving out most of the coordination: coordination of one's action on one day with another's action on another nearby day. If, on the other hand, we take longer stretches as the coordination problems, then—contrary to the definition—everyone has time for several different choices within a single coordination problem. We might pretend that everyone starts each week by choosing a contingency plan specifying what to do in every possible circumstance during the week (a *strategy* in the sense of the theory of games), and then follows his plan all week without making any further choice. Then a business week is a coordination problem in which everyone makes only his one initial choice of a contingency plan. But this treatment badly misdescribes what we do; and it still leaves out the coordination between, say, my prices for Friday and yours for next Monday. A better remedy, scheduled for Chapter II.3, goes deep. We can forget about individual coordination problems; instead of saying that uniform conformity to a regularity *R* constitutes a coordination equilibrium in every instance of a situation *S*, we can say approximately the same thing in terms of conditional preferences for conformity to *R*.

(8) If Rousseau's stag hunters stay with the hunt every time, they do so by a convention. Each stays because he trusts the others to stay as they did before, and he will eat better by staying and taking his share of the stag when it is caught.

But, less obviously, if they always split up and catch rabbits separately, that is a convention too. If the stag hunt fails unless all take part, there is no point in joining unless all the others do. Each prefers to catch rabbits if even one of the others does, and *a fortiori* if all the others do. For each to catch his rabbit is not a good coordination equilibrium. But it is a coordination equilibrium nonetheless, so long as catching a rabbit is better than going off on a one-man stag hunt that is bound to fail. So rabbit catching is, by definition, a convention.

(9) On the hypothesis that each of us wants the exclusive use of

some land and that nobody ever thinks it worth the trouble to try taking over the use of some land from another, any *de facto* division of the land is a convention. Each goes on using a certain portion and keeping off the rest in the knowledge that, since others will go on using everything else, that is the only way he can meet his needs and stay out of trouble. A better convention might provide a regular way to deal with changes in the population of land users. It might be part of the convention, for instance, that when any man dies, the oldest boy not yet using land begins to use the vacant portion.

I have not called these portions of land *property.* At its simplest—say, among anarchists—the institution of property might be nothing more than a convention specifying who shall have the exclusive use of which goods. This seems to be Hume's theory of property. For us, the institution of property is more complicated; we have built it into an elaborate system of laws and institutions. We do not say that a squatter owns the land he farms, though he enjoys the exclusive use of it by a convention, since another claimant is entitled by law to call on the police to kick the squatter off. I therefore shall not *define* property as goods reserved by convention for someone's exclusive use.

(10) A medium of exchange—say, coin of the realm—has its special status by a convention among tradesmen to take it without question in return for goods and services. Some conventional media are better than others: bulky or perishable ones are bad; ones that would retain some use if the convention collapsed are good—but the inconvenience of accepting a bad medium of exchange is less than the inconvenience of refusing it when others take it, or of taking what one can neither use nor spend. Again, as in (4), there is the complication of legal sanctions. Refusal to accept legal tender makes a debt legally unenforceable. But again, such sanctions are superflous if they agree with convention, are outweighed if they go against it, are not decisive either way, and therefore do not make our regularity any the less conventional.

I suppose we may safely define a *medium of exchange* as any good that is conventionally accepted in some population in return for goods and services. This definition raises an annoying question: is it right to say that we have a convention to accept our media of exchange in return for goods and services? It is false to say that our convention is that we accept our media of exchange in return for goods and services. For what follows "that" does not state any convention because it is true, by definition, of any population. On the other hand, it is true to say *of* our media of exchange that our convention is that we accept *them* in return for goods and services. My question was ambiguous. It can be read opaquely or transparently.[6] It is like the question whether Hegel knew that the number of planets is greater than seven. He did not know that the number of planets is greater than seven. But he did know, *of* the number of planets—namely nine—that *it* is greater than seven.

(11) A population's common use of some one language—Welsh, say—is a convention. The Welshmen in parts of Wales use Welsh; each uses Welsh because he expects his neighbors to, and for the sake of communication he wishes to use whatever language his neighbors use.

Does he not rather wish to use whatever language his neighbors will *understand?* Yes; but as a fact of human nature, he and his neighbors will best understand the language they use. So the right thing to say is that he wishes to use the language they use *because* that is the language they will understand. It follows that this is another case of coordination over time: he wishes to use the language they have been using most over a period in the past, a period long enough for them to have become skilled in its use.

To say that he wishes to use whichever language his neighbors use is not to say that if they switch suddenly, somehow, he would wish to switch immediately. He would not wish to, because he could not; he would have to practice their new language. Besides, he could count on them to understand Welsh for a time after they had ceased to

[6] See W. V. Quine, "Quantifiers and Propositional Attitudes," *Journal of Philosophy,* 53 (1956), pp. 177-187.

use it. But probably he would wish to switch as soon as he easily could. And if it suddenly came to pass that his neighbors had been using their new language for twenty years—while he, let us say, had been sleeping like Rip Van Winkle—he would try to conform with the utmost urgency.

I do not deny, of course, that a man may prefer one language to another—say, the language of his fathers to the language of their conquerors. But that does not matter. Different coordination equilibria do not have to be equally good—only good enough so that everyone is ready to do his part if the others do. There are few who would give up communication out of piety to the mother tongue, if it came to that.

Certainly not every feature of a language is conventional. No humanly possible language relies on ultrasonic whistles, so it is not by convention that the Welshmen do not. We do not yet know exactly which features of languages are conventional and which are common to all humanly possible languages; Noam Chomsky and his school have argued that there is less conventionality than one might have thought.[7] But so long as even two languages are humanly possible, it must be by convention that a population chooses to use one or the other.

In saying that Welshmen use Welsh by convention, I do not say it is a convention that Welshmen use Welsh. This, or something similar but more complicated, might perhaps be true by definition of "Welsh." Rather, I say *of* Welsh that it is a convention among Welshmen that *they* use *it*. The difference is the same ambiguity between opaque and transparent readings that arose in (10).

If using Welsh is to be a convention, it must be a regularity in behavior. It is not, of course, a regularity that fully determines a Welshman's behavior. He can say a variety of things, or remain silent, and he can respond to utterances in a variety of ways, and still be conforming to the conventional regularity. But that is nothing special.

[7] Noam Chomsky, "Recent Contributions to the Theory of Innate Ideas," *Synthese*, 17 (1967), pp. 2-11.

No convention determines every detail of behavior. (The meeting-place convention, for instance, does not specify whether to walk or ride to the meeting place.) This convention, like any other, restricts behavior without removing all choice. There is more choice, and more important choice, in this case than in some others; but there is no difference in kind.

A convention is a regularity in behavior. I do not want to say that the users of Welsh are conforming to their convention when and only when they are rightly said to be "using Welsh." A man lying in Welsh is using Welsh, but he is violating its convention; a man who remains silent during a conversation may be conforming to the convention although he is not using Welsh. In due course we shall see how the convention of a language may be described; here I will say only that it is a regularity restricting one's production of, and response to, verbal utterances and inscriptions. Linguistic competence consists in part of a disposition to conform to that restriction with ease; and in part of an expectation that one's neighbors will be likewise disposed, with a recognition of their conformity as the reason for one's own. No doubt a child or an idiot may conform without reason; if so, he is not party to the convention and his linguistic competence is incomplete.