

Chapter 36

Convention (An Excerpt on Coordination and Higher-Order Expectations)

David Lewis

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.

(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).

This chapter is an excerpt of David Lewis’ PhD thesis *Convention*. It contains the primary content of Chapter 1: *Coordination and Convention*, where Lewis presents the idea of conventions as a mutually expected regularity in coordination in a recurrent situation.

D. Lewis (deceased)
Princeton University, Princeton, NJ, USA

(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.

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.

To exclude trivial cases, a coordination problem must have more than one coordination equilibrium. But that requirement is not quite strong enough. Figure 36.1 shows two matrices in which, sure enough, 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 coordination—since if they try for different equilibria, some equilibrium 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

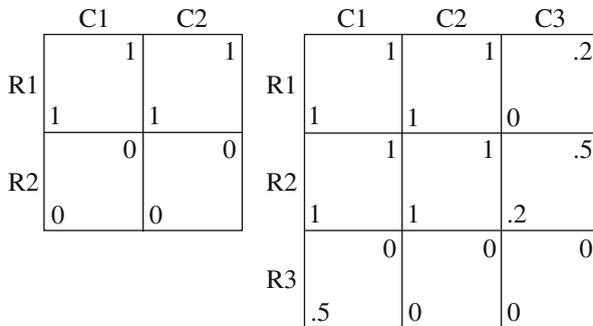


Fig. 36.1

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

Fig. 36.2

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 Fig. 36.1, 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 Fig. 36.2 ought to be counted as essentially similar to our clear examples of coordination problems, despite the impropriety of its equilibrium $\langle R3, C3 \rangle$. The two proper coordination equilibria— $\langle R1, C1 \rangle$ and $\langle R2, C2 \rangle$ —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 Fig. 36.3. But a satisfactory restriction would be complicated and would entail too many qualifications later. And situations like those of Fig. 36.3 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 Fig. 36.4, which does have two proper coordination equilibria. The right-hand matrix can be consolidated in a similar way. But matrices like the one in Fig. 36.5, which are ruled out by the strong restriction, and ought to be ruled out, cannot be rescued by any such consolidation.

To sum up: Coordination problems—situations that resemble my 11 examples in the important respects in which they resemble one another¹—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.

¹See 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.

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

Fig. 36.3

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

Fig. 36.4

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

Fig. 36.5

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).

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 Fig. 36.6, Row-chooser should do his part of the coordination equilibrium $\langle R1, C1 \rangle$ 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 Fig. 36.7, Row-chooser should choose $R1$ only if he has more 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 $\langle R1, C1 \rangle$ in order to hedge against the possibility that his expectation was wrong. And in the coordination problem shown in Fig. 36.8, Row-chooser might be sure that if Column-chooser fails to do

	C1	C2
R1	1	0
R2	0	1

Fig. 36.6

	C1	C2
R1	1	-8
R2	0	1

	C1	C2
R1	1	0
R2	0	9

	C1	C2
R1	3	-26
R2	0	1

Fig. 36.7

	C1	C2	C3
R1	1	0	-8
R2	0	1	9

Fig. 36.8

his part of $\langle R1, C1 \rangle$, 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 n th-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 Fig. 36.9. (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 Fig. 36.10. 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 Fig. 36.11. And so on. The whole of my reasoning (simplified by disregarding the rationality premises) may be represented as in Fig. 36.12 for whatever finite number of stages it may take for

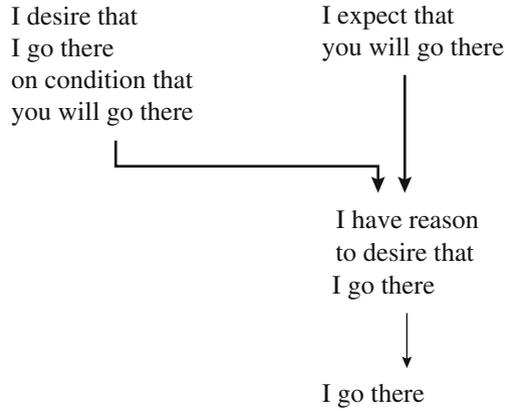


Fig. 36.9

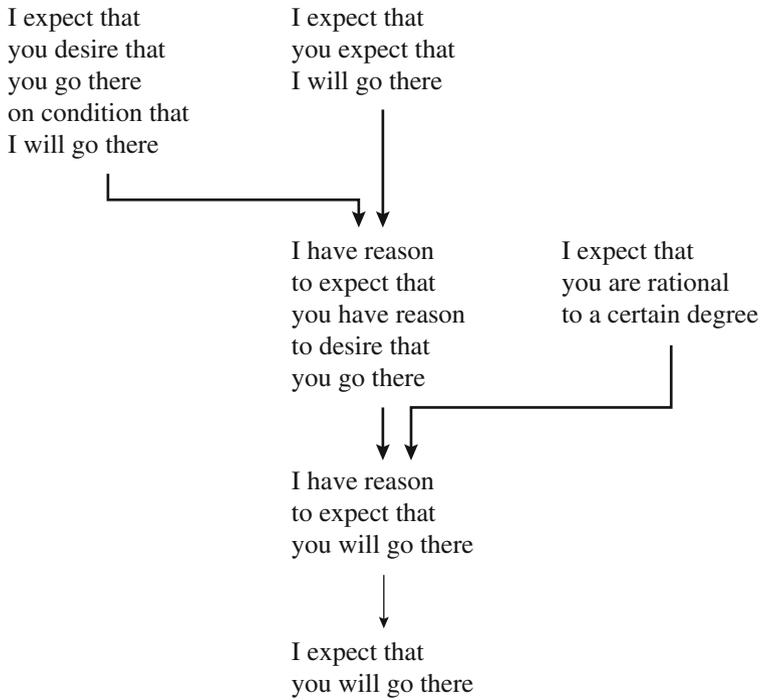


Fig. 36.10

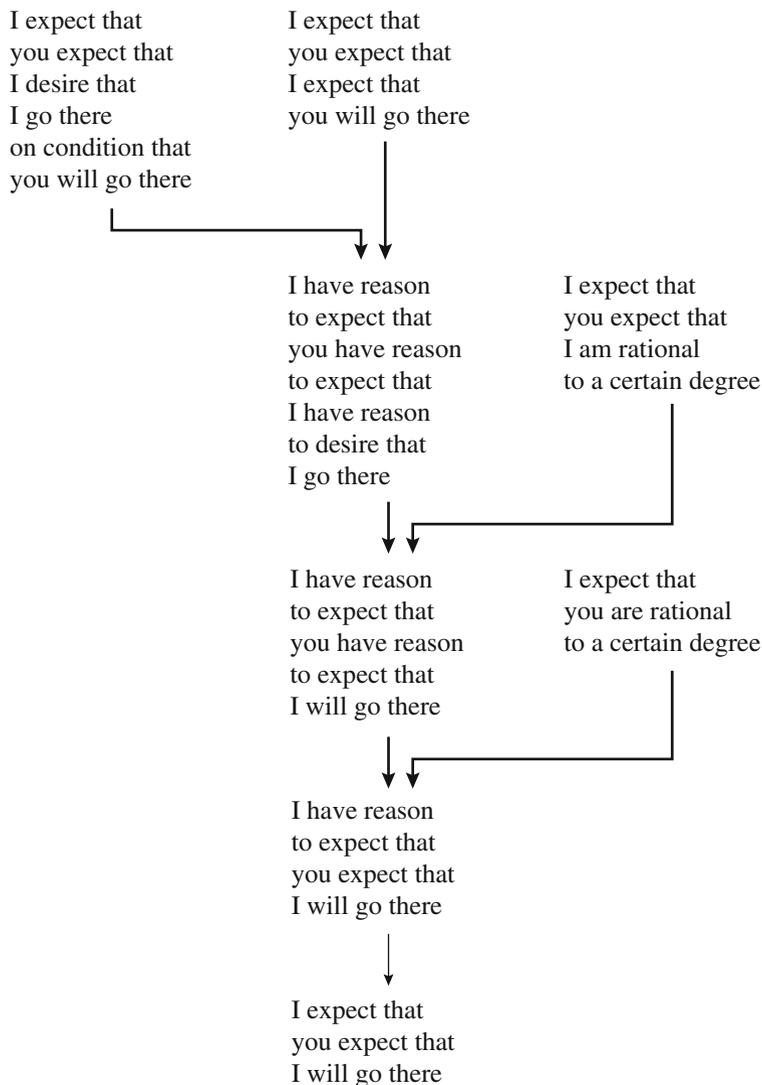


Fig. 36.11

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) contains 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

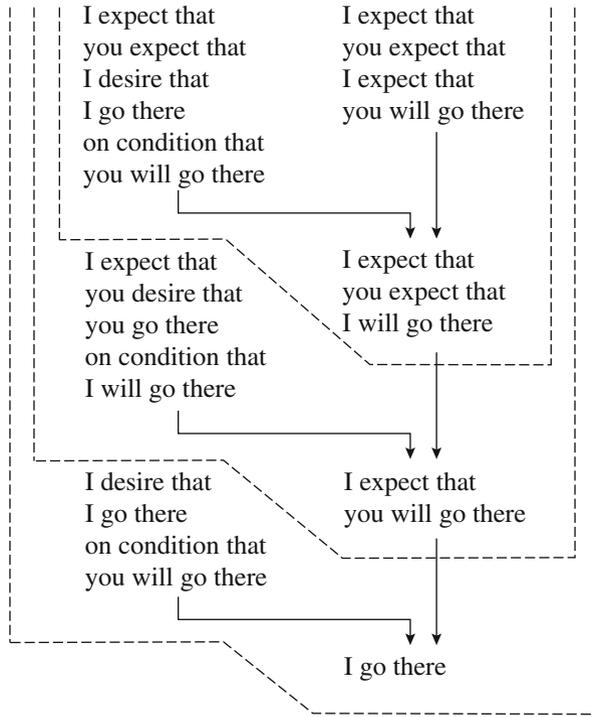


Fig. 36.12

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 $x_1, x_2 \dots$ 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 n th-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 n th-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.² 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.

²Michael 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 n th-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.

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 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.

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 .

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.