IV | Convention and Communication

1. Sample Signals

So far we have been considering conventions in general. Now we turn to one especially important class of conventions: those whereby we give to suitable actions the status of signals.

Communication by conventional signals is a commonplace phenomenon, so much so that we must make an effort not to take it for granted. We could exercise our tacit understanding all we want without ever making it more explicit. That is what would happen if we started by saying that actions are signals when we endow them with meanings. This truism will bring us no nearer to describing the phenomenon of signaling without depending on our prior tactit understanding thereof. So let us describe the phenomenon in other terms and leave meaning to look after itself.

Consider a communicator and his audience—for instance, the sexton of the Old North Church and Paul Revere. The sexton acts according to some contingency plan, such as:

R1: If the redcoats are observed staying home, hang no lantern in the belfry.
If the redcoats are observed setting out by land, hang one lantern in the belfry.
If the redcoats are observed setting out by sea, hang two lanterns in the belfry.

or

R2: If the redcoats are observed staying home, hang one lantern in the belfry.

If the redcoats are observed setting out by land, hang two lanterns in the belfry.
If the redcoats are observed setting out by sea, hang no lantern in the belfry.

or

R3: If the redcoats are observed staying home, hang one lantern in the belfry.
If the redcoats are observed setting out by land, hang no lantern in the belfry.
If the redcoats are observed setting out by sea, hang two lanterns in the belfry.

(There are three more contingency plans with no lantern, one lantern, and two lanterns, plus any number of further plans involving other actions—hanging three lanterns, hanging colored lanterns, waving lanterns, hanging a flag, etc.) Paul Revere acts according to a contingency plan, such as:

C1: If no lantern is observed hanging in the belfry, go home.
If one lantern is observed hanging in the belfry, warn the countryside that the redcoats are coming by land.
If two lanterns are observed hanging in the belfry, warn the countryside that the redcoats are coming by sea.

or

C2: If no lantern is observed hanging in the belfry, warn the countryside that the redcoats are coming by sea.
If one lantern is observed hanging in the belfry, go home.
If two lanterns are observed hanging in the belfry, warn the countryside that the redcoats are coming by land.

or

C3: If no lantern is observed hanging in the belfry, warn the countryside that the redcoats are coming by land.
If one lantern is observed hanging in the belfry, go home.
If two lanterns are observed hanging in the belfry, warn the countryside that the redcoats are coming by sea.

It matters little to either which contingency plan he follows, provided their two plans combine to ensure that Paul Revere warns the countryside that the redcoats are coming by land if and only if the sexton observes them setting out by land, and that Paul Revere warns the countryside that the redcoats are coming by sea if and only if the sexton observes them setting out by sea (so far as interference and error permit). Each must act according to a contingency plan chosen with regard to his expectation about the other's choice—an expectation formed, in this case, by an explicit agreement. Their situation is a coordination problem represented in Figure 38. Successful com-

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

Figure 38

munication occurs if they achieve one of the coordination equilibria; for instance, the historical $\langle R1, C1 \rangle$, or $\langle R2, C2 \rangle$, $\langle R3, C3 \rangle$ . . . These combinations ensure that Paul Revere will give the right warning. Conditionally successful communication occurs if they achieve one of the combinations like $\langle R1, C3 \rangle$: Paul Revere may or may not give the right warning, depending on what the redcoats do. The outright failures are combinations like $\langle R1, C2 \rangle$, which ensure that Paul Revere will give a wrong warning no matter what.

I have now described the character of a case of signaling without mentioning the meaning of the signals: that two lanterns meant that

the redcoats were coming by sea, or whatever. But nothing important seems to have been left unsaid, so what has been said must somehow imply that the signals have their meanings.

It should not disturb us that this is a problem of coordinating contingency plans, not—as in the previous coordination problems—of coordinating actions themselves. We are treating the agent's *choice* of a contingency plan as an action—one that is part of his action of choosing a contingency plan and then acting on it. An agent's action according to a contingency plan is regarded as consisting of two phases: choice of a contingency plan and fulfillment of the chosen plan. His choice is concentrated in the first phase; his use of his ability to tell which contingency holds is concentrated in the second. The agents' coordination problem pertains only to the first phase. For instance, if the sexton hangs two lanterns after observing the redcoats setting out by sea, he might be doing it according to *R1* or *R3* or any number of other contingency plans. We count the difference in contingency plan as making a difference in his action, even though it is a difference that turns out not to matter.

In this case, the sexton and Paul Revere agreed upon signals for a single occasion. In other cases, the same signaling system—preeminently analogous coordination equilibrium combinations of contingency plans for a communicator and an audience—may occur repeatedly, without need for fresh agreement every time. The regularity whereby communicators and audiences use such a pair of contingency plans is a convention.

Let us consider five examples of signaling conventions.

(1) The International Code of Signals lists a correspondence of flag hoists and certain predicaments. That is, it gives a contingency plan for ships (strictly, for ships' officers) of the form: if in such-and-such predicament, hoist such-and-such flags. There is a complementary contingency plan for ships that observe flags on nearby ships: if a ship hoists such-and-such flags, act as would be appropriate on the assumption that it is in such-and-such predicament. Ships do regularly act according to these two complementary plans. A ship on a speed

trial hoists a flag white on the left and blue on the right, and other ships seeing that flag keep out of her way. A ship loading or unloading explosives hoists a flag white on the left and red on the right, and other ships seeing that flag take appropriate precautions. A ship whose crew has mutinied hoists a flag with a yellow cross on red above a flag with diagonal yellow and red stripes, and other ships seeing those flags give whatever help they can.

The regularity of acting according to this pair of contingency plans is a convention. Any ship does conform when she is in one of the designated predicaments or when she observes flags on nearby ships. And she conforms because she expects other ships to conform too—specifically, those other ships related to her as communicator to audience or as audience to communicator—and she prefers to conform to this, or any similar, regularity if they do. They share this preference. And it is common knowledge among ships that these conditions hold.

The convention governing flag hoists is a relatively formal affair. It originated by explicit agreement among representatives of some of those now party to it; it is explicitly codified; violations are punishable independently of conformity by others. For instance, a ship might be held responsible for an accident if she had failed to hoist the prescribed flag, even if that part of the code had fallen into disuse. But if it is important that ships respond appropriately to others' predicaments, then once again this independent incentive would be redundant if it agreed with convention, would be outweighed if it did not, would not be decisive either way, and therefore would not detract from the conventionality of the code.

(2) Now we turn to a very informal signaling convention: that whereby a helper standing behind a truck gestures to the driver to help him steer the truck into a narrow space. Here there is no agreement, no authoritative reference book, no institutionalized enforcement, no teaching. But the elements of a system of conventional signals are here: the helper's gestures depend on the position of the truck and the driver's steering depends on the helper's observed

gestures. So the steering depends mediately on the position of the truck, and this dependence is such as to serve the common interest of the helper and the driver in getting the truck into the space.

The helper could have made his gestures depend differently on the position of the truck. The driver could have made his steering depend differently on the helper's gestures. Some of the other combinations of dependences would have given the same dependence of the steering on the position of the truck, and would have done about as well for getting the truck into the space. The helper gestures as he does because he expects the driver to respond as he does, and the driver responds as he does because he expects the helper to gesture as he does. They have derived these expectations from their experience with helpers and drivers in the past who set a precedent by gesturing and responding according to the same dependences.

Their behavior conforms to a conventional regularity. It would be a hard regularity to describe, if only because it is hard to describe a gesture or a way of steering well enough to identify it. Yet experienced helpers and drivers do conform, expect conformity, and recognize conformity, presumably without the aid of descriptions of their regular behavior.

We can say that helpers and drivers among us act according to their respective contingency plans, provided we understand that these contingency plans are no more than descriptions that *could be given* of the way their actions depend upon their observations. In example (1) the communicator's and audience's contingency plans were more than this: they were descriptions that *had been given* of a dependence between predicaments and flags, and the dependence was maintained partly because the agents could consult that description to find out what to do.

There must, in general, be some mediating mechanism to make an agent's actions depend on his observations in a definite way. The mechanism must be sufficiently under the agent's control that he can set it to produce the dependence he wants. *One* such mechanism is a description of the desired dependence for the agent to consult. But

in the present example, we have no reason to think that the helper's or the driver's dependence-producing mechanism contains a description of his desired dependence, or anything like one.

Our first two examples illustrated the standard sort of signaling, in formal and informal versions. The third is a slight variation; the fourth, fifth, and sixth are more serious ones.

(3) A man blazes a trail by making marks that depend on the route he thinks it best for others to take. His actions are done once and for all, but their traces last a long time. When others follow his trail afterwards, the route they take depends on the observed traces of his actions, so that they take the route he thought best for them.

In the previous examples we had sequences of independent coordination problems between communicator-audience pairs. In this example, each trail gives us many two-person coordination problems with the trail blazer's side in common. The several trail followers must all choose their contingency plans to achieve coordination with the trail blazer's original choice of a contingency plan. The trail blazer must choose his contingency plan to achieve coordination with each—or as many as possible—of the trail followers' choices of contingency plans on various occasions. There are as many different coordination problems as there are occasions on which the trail is followed. These coordination problems are extended through time, so that the trail followers must have contingency plans for responding to the observed traces of the trail blazer's actions, not to observations of those actions themselves.

(4) A railroad installs automatic signals: semaphores and the machinery to make their position depend on the occupancy of the track ahead. Instead of a communicator who does observable actions according to a contingency plan, there is the original agent who acts to install the machinery and there is the machinery which subsequently operates according to a contingency plan. As in example (3), there are any number of two-person coordination problems with one side—that of the agent who chose a contingency plan to be built into the machinery—in common. But in this example the agent who

chooses a contingency plan for the signaling machinery does not himself act according to that plan.

Or the trains which stop and go on in response to the semaphores could be automated. On a railroad with automated trains and manual semaphores, every agent who operates a semaphore is involved in a two-person coordination problem with the agent who chose, once and for all, a contingency plan to be built into the trains. (On a railroad with both automated trains and automated semaphores, there is only the single coordination problem between the agent who chooses a contingency plan to be built into the trains and the agent who chooses a contingency plan to be built into the semaphores.)

(5) An automatic stoplight is installed to regulate traffic at a crossroad. The light always shows red to one road, green to the other. It changes periodically, not according to any contingency plan. The drivers at the crossroad (at any time) *do* act according to a conventional contingency plan: stop on red, go on green. Thereby they get through the crossing safely and easily. The drivers—the stoplight's audience—achieve a coordination equilibrium. (Once more the independent incentives provided by the highway patrol are superflous if they agree with convention, outweighed if not, and hence do not make the regularity any less conventional.)

In this case the coordination is entirely among members of the audience. No one plays the role of communicator, neither the stoplight nor whoever had it installed. Nobody chose a contingency plan for the stoplight; it has none. It is just a feature of the scene which the drivers can use to mediate their own coordination. It is inessential that someone is responsible for its presence: if we built everlasting stoplights and our descendants forgot they were artifacts, the lights could still regulate traffic and they would still be signals. Even the rising of the moon can be a signal—to begin an uprising, say—though it would be a prearranged one-shot signal, not a conventional one.

(6) Many men take turns using the same horses. When they want to turn a team right they yell "gee"; when they want to turn a team left they yell "haw." When a horse fails to turn right on "gee" or

left on "haw," they hit him. The men are the communicators and the horses are the audience, and both act according to complementary contingency plans. But the coordination problem of choosing contingency plans, and the convention that solves it, exist only among the men. The horses are only beasts, and they react as they have been trained to react. The men must coordinate in order to keep the horses trained and in order to take advantage of their training.

2. Analysis of Signaling

Now that we have made the acquaintance of signaling problems, signaling systems, and signaling conventions, how shall we define them?

In this section I shall limit my discussion to *two-sided* signaling, in which the coordination needed is coordination between communicator and audience. The example of Paul Revere and the sexton, and the first three examples of signaling conventions, were two-sided. The last three examples illustrated *one-sided* signaling, in which the coordination needed was either between communicators or between members of the audience. Having defined two-sided signaling problems, systems, and conventions, it would be straightforward but tedious to give the analogous definitions for both kinds of one-sided signaling and for signaling in which a mixture of two-sided and one-sided coordination is needed. I shall leave it to the examples to show how that would be done. But what is said about signaling henceforth is meant to apply *mutatis mutandis* to signaling in general.

A *two-sided signaling problem* is a situation S involving an agent called the *communicator* and one or more other agents called the *audience*, such that it is true that, and it is common knowledge for the communicator and the audience that:

> Exactly one of several alternative states of affairs $s_1 \ldots s_m$ holds. The communicator, but not the audience, is in a good position to tell which one it is.

Each member of the audience can do any one of several alternative actions $r_1 \ldots r_m$ called *responses*. Everyone involved wants the audience's responses to depend in a certain way upon the state of affairs that holds. There is a certain one-to-one function F from $\{s_i\}$ onto $\{r_j\}$ such that everyone prefers that each member of the audience do $F(s_i)$ on condition that s_i holds, for each s_i.

The communicator can do any one of several alternative actions $\sigma_1 \ldots \sigma_n$ $(n \geq m)$ called *signals*. The audience is in a good position to tell which one he does. No one involved has any preference regarding these actions which is strong enough to outweigh his preference for the dependence F of audience's responses upon states of affairs.

Note that the preferred response is specified to be the same for all members of the audience. This may seem restrictive, but it is not. If the preferred action of a member of the audience depends on his circumstances, his role in the situation, or anything else besides the state of affairs in $\{s_i\}$, that dependence should be built into the specifications of the responses $\{r_j\}$. If a warship is in distress, another warship's preferred response may be this: to take her place in battle if it is possible to do so effectively; otherwise to come to her aid. We count this as one response, specified by a pair of conditionals.

A *communicator's contingency plan* is any possible way in which the communicator's signal may depend upon the state of affairs that he observes to hold. It is a function Fc from $\{s_i\}$ into $\{\sigma_k\}$. A communicator in S acts according to Fc if, for each s_i, he does $Fc(s_i)$ if he observes that s_i holds. Since he is in a position to tell which s_i does hold, and since $Fc(s_i)$ is his own action, he should be able to act according to any contingency plan Fc. If Fc is a one-to-one function, we call it *admissible*.

Similarly, an *audience's contingency plan* is any possible way in which the response of a member of the audience may depend upon the signal he observes the communicator to give. It is a one-to-one

function *Fa* from part of $\{\sigma_k\}$ into $\{r_j\}$. A member of the audience in *S* acts according to *Fa* if, for each σ_k in the domain of *Fa*, he does $Fa(\sigma_k)$ if he observes that the communicator gives σ_k. Since he is in a position to tell which σ_k is given, and since $Fa(\sigma_k)$ is his own action, he should be able to act according to any contingency plan *Fa*. If the range of *Fa* coincides with the range of *F*, we call *Fa admissible*.

Suppose the communicator acts according to a contingency plan *Fc* and all members of the audience act according to a contingency plan *Fa*. Then the audience's response will depend on the state of affairs that holds according to the function $Fa|Fc$ obtained by composition of *Fa* and *Fc*. (For any two functions f and g, and any argument x such that $g(x)$ is in the domain of f, $f|g(x)$ is defined as $f(g(x))$; $f|g$ is undefined on other arguments.)

If the actual state of affairs happens to be *s*, and if $Fa|Fc(s) = F(s)$, the audience's response will be one that is preferred in the actual state of affairs. So each agent will be acting according to one of the contingency plans which is best given the others' contingency plans and the actual state of affairs.

Better, suppose $Fa|Fc = F$ uniformly over $\{s_i\}$. (It follows that the range of *Fc* must coincide with the domain of *Fa*.) Then the audience's response will be one that is preferred in the actual state of affairs, no matter which state of affairs in $\{s_i\}$ that may happen to be. Each agent will be acting according to the contingency plan that is best given the others' contingency plans and *any* state of affairs. Whenever *Fc* and *Fa* combine in this way to give the preferred dependence of the audience's response upon the state of affairs, we call $\langle Fc, Fa\rangle$ a *signaling system*.

All and only admissible contingency plans belong to signaling systems. Proof:

> Let *Fc* be a communicator's contingency plan. If it is admissible, it is one-to-one and it has an inverse $Fc\check{}$; if so, $\langle Fc, F|Fc\check{}\rangle$ is a signaling system. If it is inadmissible it is not one-to-one; so for any *Fa*, either $Fa|Fc$ is not one-to-one or $Fa|Fc$ is not

defined for some state of affairs; in neither case can $Fa|Fc$ coincide with F; so $\langle Fc, Fa\rangle$ is not a signaling system.

Let Fa be an audience's contingency plan. If it is admissible, let $F\check{a}$ be its inverse; $\langle F\check{a}|F, Fa\rangle$ is a signaling system. If it is inadmissible its range does not coincide with the range of F; so for any Fc, the range of $Fa|Fc$ does not coincide with the range of F; so $\langle Fc, Fa\rangle$ is not a signaling system.

In a signaling problem with m states of affairs and n signals, there are $n!/(n - m)!$ signaling systems. Proof:

Construct Fc thus: for any s in $\{s_i\}$, if $F(s)$ is the kth member of the range of F, let $Fc(s)$ be σ_k. Fc is an admissible communicator's contingency plan.

Take any function G which maps the range of Fc one-to-one into $\{\sigma_k\}$. Let $\check{G}$ be the inverse of G; let $F\check{c}$ be the inverse of Fc. Then $\langle G|Fc, (F|F\check{c})|\check{G}\rangle$ is a signaling system. Every signaling system can be obtained in this way. There are $n!/(n - m)!$ different one-to-one functions from the range of Fc into $\{\sigma_k\}$, so there are that many different signaling systems.

We may think of the communicator and the audience as acting by first choosing contingency plans and then acting according to their chosen plans. If so, their choices of plans make a coordination problem. We can represent it by a matrix in which the rows represent the communicator's contingency plans and the columns, levels, and so on, represent contingency plans for all members of the audience. Some of these plans are admissible, others are not. An agent's payoff for any combination depends on his payoffs for that combination in each of the states of affairs in $\{s_i\}$, weighted by the probabilities he assigns to the different states of affairs. (For some combinations, we will have to consider what he expects members of the audience to do in response to a signal not in the domains of their chosen plans.) It is easy to see that every signaling system is a proper coordination equilibrium in this coordination problem.

But signaling systems may not be the only coordination equilibria. Certain pairs of inadmissible contingency plans may be coordination equilibria. In our example of Paul Revere and the sexton, suppose a defense against a land attack would also work fairly well against a sea attack, but not vice versa, and suppose the two attacks seem equally likely. Then there may be a coordination equilibrium between this plan for the sexton:

> If the redcoats are observed staying home, hang no lantern in the belfry.
> If the redcoats are observed setting out by land, hang one lantern in the belfry.
> If the redcoats are observed setting out by sea, hang one lantern in the belfry.

and this plan for Paul Revere:

> If no lantern is observed hanging in the belfry, go home.
> If one lantern is observed hanging in the belfry, warn the countryside that the redcoats are coming by land.

The plans combine to give a dependence of Paul Revere's response upon the observed state of affairs which, although not the preferred dependence, is not too bad. But neither one can reach the preferred dependence just by changing his plan; each is acting according to a plan that is best given the other's plan. So the combination is a coordination equilibrium. (It is improper; given the sexton's plan, Paul Revere might just as well have chosen a different plan:

> If no lantern is observed hanging in the belfry, go home.
> If one lantern is observed hanging in the belfry, warn the countryside that the redcoats are coming by land.
> If two lanterns are observed hanging in the belfry, warn the countryside that the redcoats are coming by sea.

His response would have been just the same, no matter what the sexton observed.) It is not a signaling system. At least, it is not a

signaling system in our *original* signaling problem. If we reclassify states of affairs and responses, forgetting the difference between the two kinds of attacks and the two kinds of warnings, we get a new signaling problem with a less detailed preferred dependence function; and in this new problem, the combination of plans we are considering is a signaling system and is a proper coordination equilibrium.

So we have two kinds of coordination equilibria in signaling problems. There are signaling systems; and there are improper coordination equilibria that are not signaling systems but would become signaling systems under a less detailed classification of states of affairs and responses. I do not know whether there can be coordination equilibria of any other kind. One relevant result is this: if the audience has just one member, then no combination of one admissible plan and one inadmissible plan can be an equilibrium; for another combination on the same row or column would be a signaling system and would be preferred to the given combination.

Given the definition of signaling systems, we can define a *signaling convention* as any convention whereby members of a population P who are involved as communicators or audience in a certain signaling problem S do their parts of a certain signaling system $\langle Fc, Fa \rangle$ by acting according to their respective contingency plans. If such a convention exists, we also call $\langle Fc, Fa \rangle$ a *conventional signaling system.*

If some signaling problems do turn out to have coordination equilibria that are not signaling systems—not even under a suitable less detailed classification of states of affairs and responses—then we would have to say that there could be conventions which are not signaling conventions even though they govern the choice of contingency plans in a signaling problem. I see no great harm, however, in leaving that possibility open.

When an agent acts in conformity to a signaling convention—and in any other case in which he solves a signaling problem by agreement, salience, or precedent—his action is justified by the reasoning shown in Figure 39 (for a communicator) or in Figure 40 (for a

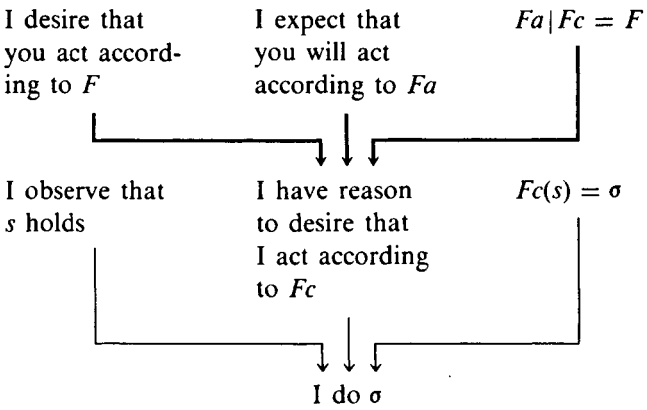

Figure 39

member of the audience). That is, we are treating the agent's choice of an action as a rational choice of a whole contingency plan, followed by action to carry out his chosen plan in the light of particular information about the situation. This is the best treatment when we want to include signaling within the general theory of coordination and convention.

But we could treat the agents' justifications of their choices more economically. The particular information the agent needs—which

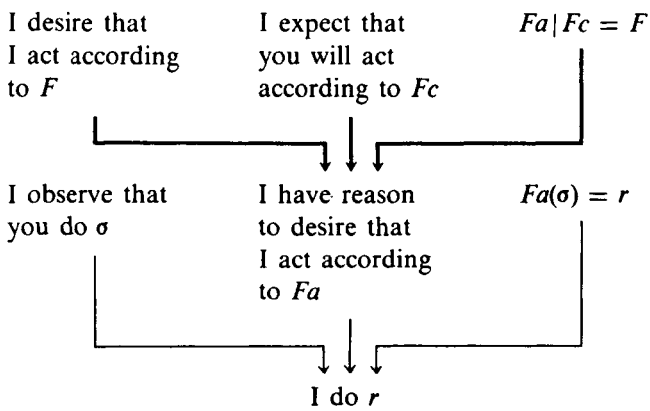

Figure 40

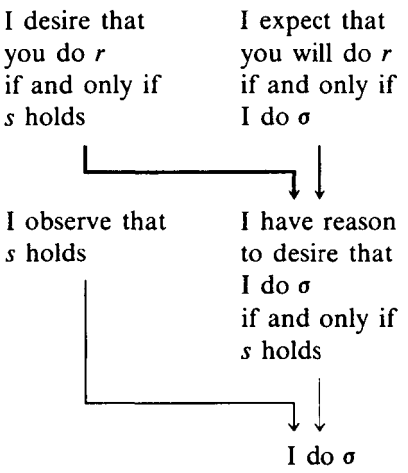

Figure 41

state of affairs holds or which signal was given—is available to him all along; so his choice is justified if he justifies only the relevant part of his contingency plan. Throwing out the irrelevant contingencies, we get the reasoning shown in Figure 41 (for a communicator) or in Figure 42 (for a member of the audience). These justifi-

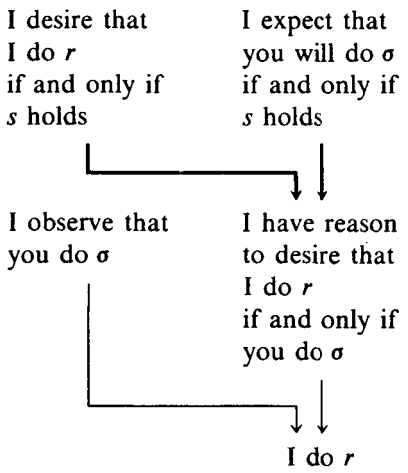

Figure 42

cations may be rearranged slightly, to eliminate the choice of even part of a contingency plan. The premises available can be applied in a different order. For the communicator, we then get the reasoning shown in Figure 43. For a member of the audience, we get the

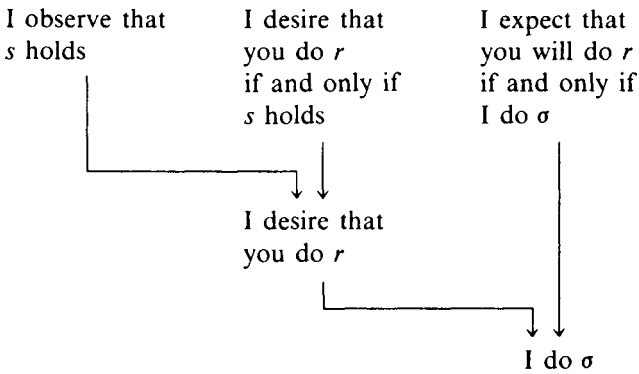

Figure 43

reasoning shown in Figure 44. In the communicator's case, the state of affairs *s* gives him a reason for trying to produce the audience's response *r*; and his expectation that the audience will respond to σ according to *Fa* gives him reason to believe that σ would be an

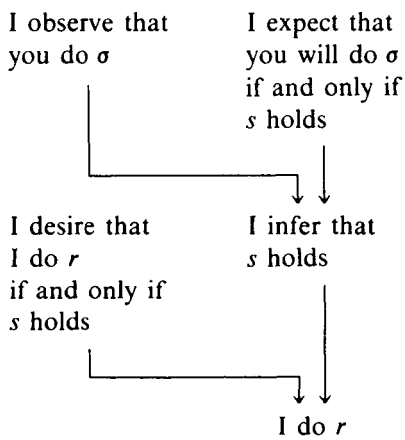

Figure 44

effective means of producing *r*. So he does σ. In the case of a member of the audience, his expectation that the communicator will act according to *Fc* gives him reason to take σ as good evidence that *s* holds; and having inferred that *s* holds, he has reason to do *r*.

These new schemata of justification, like the old ones, can be replicated; one agent's replication of the other's reasoning provides him with the central premise for his own reasoning. But these replications must be conditional: one agent figures out how the other *would* reason if he were given a premise. The communicator figures out how the audience would reason in response to a signal; the audience figures out how the communicator would reason in response to a state of affairs. These conditional replications can be replicated in turn—and so on, for any number of stages. Consider the communicator's reasoning in justification of his decision to give σ; we can carry it two stages back, as shown in Figure 45.

We have now represented the agents' justifications in two quite different ways. The first way is based on the fact that the agent's choices of contingency plan constitute a coordination problem, and the conventional signaling system ⟨*Fc*, *Fa*⟩ is a coordination equilibrium therein. The second way is based on the fact that, since it is common knowledge that the audience acts according to *Fa*, it is common knowledge that a signal σ is a good means of producing the response *Fa*(σ); and that, since it is common knowledge that the communicator acts according to *Fc*, it is common knowledge that a signal σ is good evidence that the state holds which is mapped onto σ by *Fc*.

Both representations are somewhat artificial, however, as representations of an agent's *actual* reasoning. Fortunately, they are artificial in opposite and complementary ways. So if one is realistic, the other is not. Take an agent who does not merely follow his chosen contingency plan blindly. After he knows which is the actual contingency, he deliberates on the likely outcomes of his alternative actions and does the one that seems best. Must he really be said to choose—or renew his choice of—his whole contingency plan? Or only the part

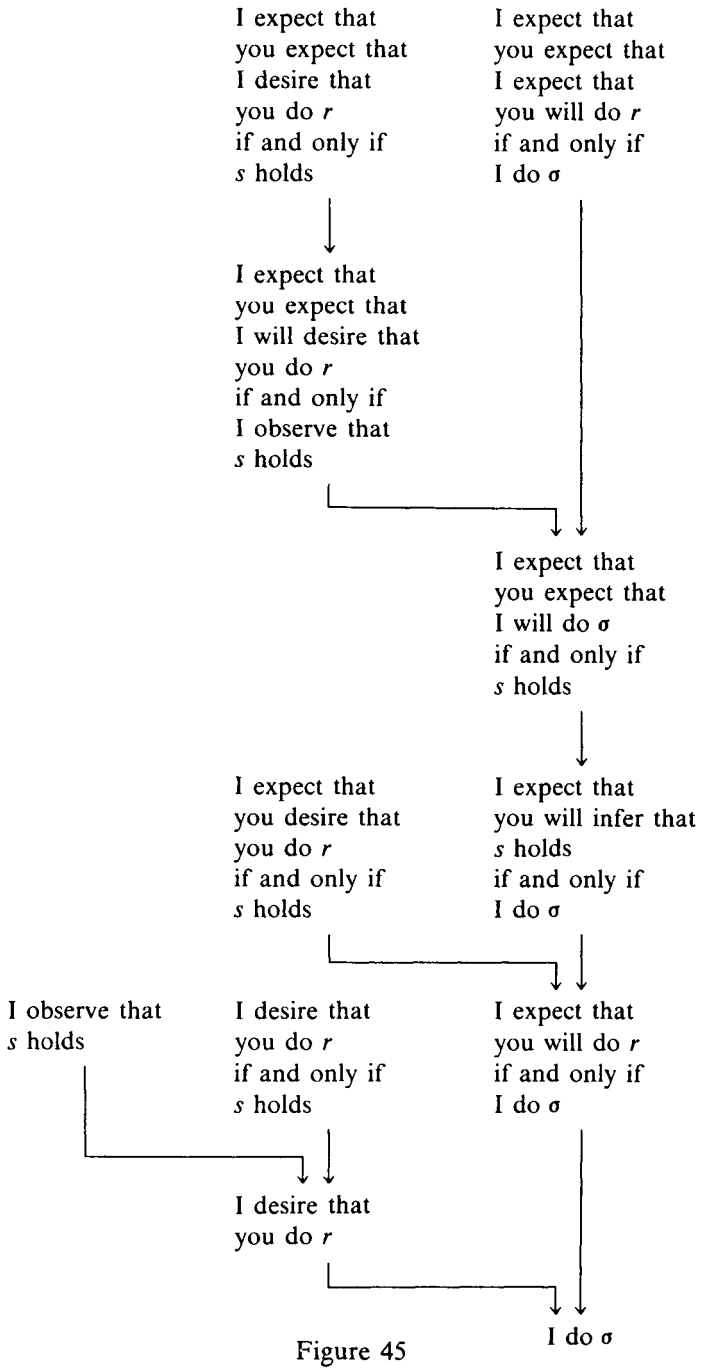

Figure 45

that covers the actual contingency? Take, on the other hand, an agent whose signaling is dull routine work, done with his mind elsewhere. Does he stop to think: "This action will get the audience to do so-and-so, which is the best thing under the circumstances, so I had better do this action" (if he is a communicator), or "The communicator must be doing that because he has observed such-and-such, in which case this would be the best action for me to do, so I had better do this action" (if he is in the audience)? No: he relaxes and follows his contingency plan by habit. He is aware in a general way that the plan he is following is the best one, given the plan his partner is following. If he received evidence to the contrary, he would stop and think, and he would begin to act according to a different contingency plan. But as long as matters go well, he does not think about how what he is doing will be for the best on any particular occasion.

But we do not have to represent the agents' actual reasoning. We have to consider only the rational justifications of their choices by practical reasoning they *could* go through, given their beliefs and desires. Yet this is not to renounce an interest in explaining their choices. Justifications do explain choices, whether or not the agent actually goes through a process of reasoning following the justification. For it is a fact of human nature that we tend to act in ways justified by our beliefs and desires, even when we do not think through the justification. I may put it negatively: whatever may be the habitual processes that actually do control our choices, if they started tending to go against our beliefs and desires they soon would be overridden, corrected, and retrained by explicit practical reasoning.

3. Verbal Signaling

An action is suited to be a signal if it or its traces are easy for an audience to observe; and if it is intrinsically unimportant, so that no party has strong extraneous preferences for or against its being done.

Two sorts of actions are ideally suited to be signals, in many

situations: production of a string of vocal sounds within earshot of the audience; and production of a string of lasting marks on a smooth surface that will be visible to the audience. In short, speaking and writing.

Let us define a *verbal expression* as any finite sequence of types of vocal sounds or types of marks.[1] To utter or to inscribe a verbal expression is to produce a string of sounds reverberating in the air or a string of marks on a surface. A signal that is an action of uttering or inscribing a verbal expression may be called a *verbal signal.*

Officially, the signal is the action. But we usually need not be careful to distinguish the action, the string of sounds or marks produced by the action, and the verbal expression uttered or inscribed; all three can be called the signal. We can also allow "uttering" to cover inscribing as well as vocal uttering.

Had circumstances been otherwise, the sexton and Paul Revere might have found it best to use a prearranged or conventional verbal signaling system. As a signal to be given if the redcoats were observed setting out by sea, the sexton might have whispered sounds in Paul Revere's ear or sent him a piece of paper on which marks were inscribed. These sounds or marks might be some special prearranged code, or they might be "The redcoats are coming by sea." In the latter case, the sexton's verbal signaling would be indistinguishable from ordinary use of language.

If we endow a hypothetical community with a great many verbal signaling conventions for use in various activities, with verbal expressions suitably chosen *ad hoc*, we shall be able to simulate a community of language users—say, ourselves—rather well. An observer who stayed in the background watching these people use conventional verbal signals as they went about their business might take a long time to realize that they were not ordinary language users. But an observer who tried to converse with them would notice some deficiencies. He would find that every verbal expression they used was conventionally associated with some readily observable state of

[1] On this definition, a verbal expression can exist which is never uttered. See Quine, *Word and Object,* pp. 194–195.

affairs, or with some definite responsive action, or with both. And he would find that they could use only finitely many verbal expressions, so that the conventions governing their verbal signaling could be described by mentioning each expression they used.

We shall consider later how verbal signaling falls short of the full use of language. Yet it remains true that our hypothetical verbal signalers do not do anything we do not do. We just do more. Their use of language duplicates a fragment of ours.

Their use of language is covered by signaling conventions. What about ours, on those occasions when ours is just like theirs? I see no reason to deny that we too have verbal signaling conventions. What *would* be conventional signaling if it were our only use of language does not have a different character because we use language in other ways as well. Examining the definition of a verbal signaling convention, we can verify that it is satisfied by part of our use of language.

But even if some conventions governing our use of language are verbal signaling conventions, it does not follow that these would appear in a complete specification of our conventions of language. Probably they would be subsumed under more general conventions. We recall from Chapter II.5 that specializations of conventions may be conventions in their own right. A general convention of language, covering infinitely many different verbal expressions and a wide range of situations, would have consequences restricted to finite classes of expressions and narrowly specific situations. Our verbal signaling conventions could be consequences of this kind.

4. Conventional Meaning of Signals

I have been trying to demonstrate that an adequate account of signaling need not mention the meanings of signals—at least, not by name. But of course signals *do* have meanings.

Suppose $\langle Fc, Fa \rangle$ is a conventional signaling system; σ is a signal therein; s is the state of affairs mapped onto σ by Fc; r is the response onto which σ is mapped by Fa.

Then we might call σ—in general and on any particular occasion

on which it is given in conformity to the convention—a *conventional signal that s* holds; and we might say that σ *conventionally means that s* holds. For σ is evidence that *s* holds, by virtue of the fact that the communicator acts according to *Fc*, hence by virtue of the convention. Or we might call σ a *conventional signal to do r*; and we might say that σ *conventionally means to do r.* For σ is a good means of getting the audience to do *r*, by virtue of the fact that the audience acts according to *Fa*, hence by virtue of the convention.

Which way should we give the meaning of σ: as a signal-that or as a signal-to? Sometimes we can properly give it both ways—for instance, in all our examples so far. But in some cases it is only proper to describe σ as a signal *that* a state of affairs holds; in such cases, let us call σ an *indicative* signal (in the signaling system in question). In other cases it is only proper to describe σ as a signal *to* do something; in such cases, let us call σ an *imperative* signal. (Let us call σ a *neutral* signal if it is equally properly called a signal-that or a signal-to.) The difference seems to lie in the character of the plans *Fc* and *Fa.* How much discretion do they allow the communicator or the audience?

A contingency plan may or may not be *discretionary*; that is, it may or may not require an agent to deliberate about which course of action would be best for himself and his partners. I suggest that if *Fa* is discretionary and *Fc* is not, then σ is indicative; if *Fc* is discretionary and *Fa* is not, then σ is imperative; if neither or both are discretionary, then σ is neutral.

Take the sexton's signal of hanging two lanterns in the belfry. (Pretend now that the signaling system was a conventional one, not confined to a single occasion.) If the signaling system is ⟨*Fc*, *Fa*⟩, with

> *Fc*: If the redcoats are observed staying home, hang no lantern in the belfry.
>
> If the redcoats are observed setting out by land, hang one lantern in the belfry.

If the redcoats are observed setting out by sea, hang two lanterns in the belfry.

and

Fa: If no lantern is observed hanging in the belfry, go home.
If one lantern is observed hanging in the belfry, warn the countryside that the redcoats are coming by land.
If two lanterns are observed hanging in the belfry, warn the countryside that the redcoats are coming by sea.

then hanging two lanterns is a signal that the redcoats were observed setting out by sea, and it is a signal to warn the countryside that the redcoats are coming by sea. The two descriptions seem equally proper; the signal is neutral.

But if the signaling system is $\langle Fc, Fa' \rangle$, Fa' being a discretionary variant of *Fa,*

Fa': If no lantern is observed hanging in the belfry, do whatever seems best on the assumption that the redcoats were observed staying home.
If one lantern is observed hanging in the belfry, do whatever seems best on the assumption that the redcoats were observed setting out by land.
If two lanterns are observed hanging in the belfry, do whatever seems best on the assumption that the redcoats were observed setting out by sea.

then hanging two lanterns is properly called a signal that the redcoats were observed setting out by sea. It would be strange to call it a signal to do whatever seems best on the assumption that the redcoats were observed setting out by sea. In $\langle Fc, Fa' \rangle$, hanging two lanterns is an indicative signal.

On the other hand, if the signaling system is $\langle Fc', Fa \rangle$, Fc' being a discretionary variant of *Fc*,

Fc': If it seems best that Paul Revere should go home, hang no lantern in the belfry.
If it seems best that Paul Revere should warn the countryside that the redcoats are coming by land, hang one lantern in the belfry.
If it seems best that Paul Revere should warn the countryside that the redcoats are coming by sea, hang two lanterns in the belfry.

then hanging two lanterns is properly called a signal to warn the countryside that the redcoats are coming by sea. It would be strange to call it a signal that it seems best that Paul Revere should warn the countryside that the redcoats are coming by sea. In $\langle Fc', Fa \rangle$, hanging two lanterns is an imperative signal.

Fa and *Fa'* are stated in terms of different, crosscutting classifications of Paul Revere's action. *Fc* and *Fc'* are stated in terms of different, crosscutting classifications of the states of affairs. So the three signaling systems above belong to different signaling problems. Each is a signaling system in its own signaling problem, since it makes Paul Revere's action depend on the state of affairs in the best way available within the given classifications.

I have ignored $\langle Fc', Fa' \rangle$ because I am not sure it gives the preferred dependence in its signaling problem—it is too likely that the sexton and Paul Revere will cross each other up. If it is a signaling system, its signals are neutral.

An agent who acts according to a discretionary contingency plan must use his judgment to respond flexibly to the whole situation as he sees it, so that he can take special action to deal with unforeseen circumstances as they arise. Paul Revere, acting according to *Fa'* and observing two lanterns, might decide it was more urgent to hide the leaders than to warn the countryside. The sexton, acting according to *Fc'* and observing the redcoats setting out by sea, might discover that they were headed for the wrong place; so it would be best not to hinder them and to wait for a more important attack later. The

cost of flexibility, of course, is the risk of misinformation or misjudgment.

An agent might use his judgment even if his conventional signaling system gives him a nondiscretionary contingency plan. That is, he might decide it would be for the best to violate the signaling convention in view of unforeseen circumstances; and he might be right. But if he did, his exercise of judgment would be unexpected; whereas if he were acting according to a discretionary contingency plan in conformity to a signaling convention, it would be common knowledge that he would use his judgment.

It is not at all necessary to confine ourselves to conventional signaling systems in defining meaning for signals. Consider any signaling system $\langle Fc, Fa \rangle$ for a signaling problem S, whether or not $\langle Fc, Fa \rangle$ happens to be conventionally adopted in any population. If Fa is discretionary and Fc is not, the signals of $\langle Fc, Fa \rangle$ are indicative. Consider a signal σ and the state of affairs s mapped onto σ by Fc. We can call σ a *signal in* $\langle Fc, Fa \rangle$ *that* s holds; and we can say that σ *means in* $\langle Fc, Fa \rangle$ *that* s holds. If Fc is discretionary and Fa is not, the signals of $\langle Fc, Fa \rangle$ are imperative. Consider a signal σ and the response r onto which σ is mapped by Fa. We can call σ a *signal in* $\langle Fc, Fa \rangle$ *to do* r; and we can say that σ *means in* $\langle Fc, Fa \rangle$ *to do* r. Finally, if neither or both of Fc and Fa are discretionary, the signals of $\langle Fc, Fa \rangle$ are neutral; they can equally well be called signals-that or signals-to in $\langle Fc, Fa \rangle$; and they can equally well be said to mean-that or to mean-to in $\langle Fc, Fa \rangle$. For the most part, I shall ignore neutral signals henceforth.

We would expect that by giving the meanings of indicative signals, we give their truth conditions. And so we do. Let σ be an indicative signal that s holds, in a signaling system $\langle Fc, Fa \rangle$ for signaling problem S. Then we can call σ *true* in any instance of S in which s does hold, and *false* in any instance of S in which s does not hold.

Officially, we recall, the signal σ is an action. So we are ascribing truth and falsity to actions. We did not decide whether signals were types of actions or particular actions; nor must we decide now. The

action-type σ is true or false relative to any instance of *S*, according as the state of affairs *s* does or does not hold therein. There is no reason to confine ourselves to those instances of *S* in which σ happens to be performed. A particular action σ of that type is true or false according as the state of affairs *s* does or does not hold in the instance of *S* in which the action σ is performed. If σ is a verbal signal, we can indulge in another harmless ambiguity: we can say that the string of sounds or marks produced by the action σ in a particular instance of *S* is true or false, and we can say that the verbal expression uttered in the action σ is true or false relative to any instance of *S*. Philosophers have argued at length over the question of which entities are the proper bearers of truth and other semantic properties; I adopt an eclectic policy, letting several different kinds of entities bear semantic properties. I see no danger, provided we are prepared to convert semantic properties into semantic relations (for instance, of verbal expressions to instances of *S*) whenever it is appropriate to do so.

In any instance of *S*, by definition of a signaling problem and a signaling system, one of the signals of ⟨*Fc*, *Fa*⟩ is true and the rest are false. Suppose ⟨*Fc*, *Fa*⟩ is a conventional signaling system in some population. Then we can describe the conventional regularity in their behavior thus: in any instance of *S* among them, the communicator tries to give whichever signal is true in that instance, and every member of the audience responds by doing whatever seems best on the assumption that the communicator's signal is true.

But this does not seem right. How can truthfulness be a convention? What is the alternative? Systematic *un*truthfulness? Suppose we have arranged the signals of ⟨*Fc*, *Fa*⟩ in some sort of cyclic order. Then the members of *P* could get on very happily if, whenever any signal σ was true, they gave not σ but rather the signal just next to σ in the order (say, next on the right)—a false signal. This sort of systematic untruthfulness, if uniform and uniformly expected, would do just as well as systematic truthfulness. But that is because it would *be* systematic truthfulness—not in the original signaling system ⟨*Fc*, *Fa*⟩, but

in an alternative signaling system obtained from ⟨*Fc*, *Fa*⟩ by permuting the signals.

There is the solution: I have been talking throughout not about truth in general, but about truth in the signaling system ⟨*Fc*, *Fa*⟩. I should have stated the convention thus: in any instance of *S* among members of *P*, the communicator tries to give whichever signal is true in ⟨*Fc*, *Fa*⟩ in that instance, and the audience responds by doing whatever seems best on the assumption that he has succeeded in so doing. What was called a convention of truthfulness is more accurately called a convention of truthfulness in ⟨*Fc*, *Fa*⟩.

Given this more careful statement of the convention, we no longer have any trouble finding its alternatives. The alternative to a convention of truthfulness in ⟨*Fc*, *Fa*⟩ would be a regularity of truthfulness in some other signaling system ⟨*Fc'*, *Fa'*⟩. This may be any signaling system in the problem *S* which is sufficiently unlike ⟨*Fc*, *Fa*⟩—different enough so that it is almost always impossible to conform both to the convention and to its alternative. It is enough to require that no signal of ⟨*Fc*, *Fa*⟩ which is a signal of ⟨*Fc'*, *Fa'*⟩ as well can be true in both. It is more than enough to require that the two have no signals in common; indeed, that may be impossible because there may not be enough different signals available in *S*. If the two signaling systems have some or all of their signals in common, then, as we foresaw, systematic truthfulness in one will sometimes or always coincide with a certain kind of systematic untruthfulness in the other.

What does *not* have an alternative, and is not a convention, is this regularity: in any instance of *S* among members of *P*, the communicator tries to give whichever signal is true *under the prevailing convention* in that instance, and the audience responds by doing whatever seems best on the assumption that he has succeeded in so doing. This regularity holds whenever the members of *P* have any conventional signaling system for *S* with indicative signals, no matter what their convention may be. We can say that a signal σ is *true* in *P* in an instance of *S* if and only if there is some suitable signaling system

that is conventionally adopted in P and σ is true in that signaling system in that instance of S.

An analogous treatment is available for imperative signals. Let σ be an imperative signal to do r in a signaling system $\langle Fc, Fa \rangle$ for a signaling problem S. We can call σ *true* in $\langle Fc, Fa \rangle$ in an instance of S if almost every member of the audience does do r in that instance, and *false* in $\langle Fc, Fa \rangle$ in that instance of S if not.

Why "true"? Why not "obeyed"? My use of "true" is admittedly not standard; but "obeyed" would not be quite what I mean. Is a conditional imperative, "Do so-and-so in case such-and-such," obeyed if the antecedent turns out false? Perhaps not, but I want to call it true. Is an imperative obeyed if the imperative itself is not part of the obeyer's reason for doing what he was told to do? Perhaps not, but I want to call it true. Nicholas Rescher introduces a new expression, "terminated," to play the role I assign to "true";[2] anyone unwilling to tolerate my stretching of "true" may substitute "true or terminated." But my terminology has the advantage of pointing to a certain symmetry between imperatives and indicatives: the signal either asserts or commands a certain state of affairs to hold, and is true if that state of affairs does hold.

Suppose $\langle Fc, Fa \rangle$ is a conventional signaling system in some population. Then we can describe the conventional regularity in their behavior thus: in any instance of S among them, every member of the audience tries to respond in such a way that the communicator's signal is true in $\langle Fc, Fa \rangle$ in that instance, and the communicator gives whichever signal seems best on the assumption that the audience will succeed in so doing. Once again, we can describe their convention as a convention of truthfulness in $\langle Fc, Fa \rangle$. But in the indicative case it was up to the communicator to see to it that his signal was true, by choosing the correct signal to give; whereas in the imperative case it is up to the audience to make the communicator's signal true by responding to it correctly. Once again, the alternative to a convention of truthfulness in $\langle Fc, Fa \rangle$ would be a regularity of truthfulness

[2] *The Logic of Commands* (New York: Dover, 1966), pp. 52–61.

in some other signaling system $\langle Fc', Fa' \rangle$ for the same signaling problem, $\langle Fc', Fa' \rangle$ being sufficiently unlike $\langle Fc, Fa \rangle$.

If $\langle Fc, Fa \rangle$ is any signaling system with indicative or imperative signals, we can identify $\langle Fc, Fa \rangle$ by specifying three things. (1) We must specify the set of signals of the system; that is, the set which is to be the range of *Fc* and the domain of *Fa*. (2) We must specify, for each signal of the system, whether it is indicative or imperative. (3) We must specify, for each signal of the system, the state of affairs in which it is true. We can take this state of affairs to be a certain set of possible instances of *S*—those in which, as we say, the state of affairs holds. In other words, we specify truth conditions for the signals of the system.

Given these three pieces of information, *Fc* and *Fa* are determined. By looking at the truth conditions of all the signals of the system, we find all the possible instances of the signaling problem to which the system applies. If the signals are specified as indicative, we reconstruct *Fc* by looking at the truth conditions of the signals; if the signals are identified as imperative, we reconstruct *Fa* by looking at the truth conditions of the signals. Given either *Fc* or *Fa*, we find the other just by looking for a discretionary contingency plan that combines with the given plan to give the preferred dependence of responses on states of affairs.

It is possible to specify $\langle Fc, Fa \rangle$ by means of a single function $\mathfrak{L}$ constructed as follows. The domain of $\mathfrak{L}$ is to be the set of signals of the system. Given any signal σ of the system, $\mathfrak{L}$ is to assign it an *interpretation* $\langle \mu, \tau \rangle$. The component μ, called a *mood*, indicates whether σ is indicative or imperative. It does not matter just what thing μ is. Let us take it to be a code number: 0 for indicative, 1 for imperative. The component τ of an interpretation, called a *truth condition*, indicates the state of affairs in which σ is true. We can take τ to be a set of possible instances of some signaling problem: namely, those instances in which σ is true in the sense appropriate to its mood.

If $\langle Fc, Fa \rangle$ is a verbal signaling system, its signals are actions of

uttering verbal expressions. Then we may make a slight change. Let the domain of £ be the set of verbal expressions uttered in signals of the system, rather than the signals themselves. We can call any verbal expression in the domain of £ a *sentence* of £. When we have made this change, it is natural to call the function £ a *language*: the language associated with ⟨*Fc*, *Fa*⟩. Not that every language is thus associated with a verbal signaling system. But some rudimentary languages are, and these we have now examined in some detail. It remains to be seen how these verbal signaling languages fall short of more interesting languages.

A sentence σ of a verbal signaling language £ is *true* in £ in a particular instance of the signaling problem to which £ applies if and only if that instance belongs to the truth condition assigned to σ by £; otherwise σ is false in £ in that instance of the problem. Suppose the associated signaling system is conventionally adopted in some population; then we can say they have a convention to use the language £. Their convention can be restated thus: in any instance of the signaling problem to which £ applies, one party—communicator for indicatives, audience for imperatives—tries to make sure that the communicator utters a sentence that is true in £ in that instance; the other party acts as seems best on the assumption that the first party has succeeded in so doing. This is the familiar signaling convention, now redescribed as a convention of truthfulness in £; its alternatives would be regularities of truthfulness in other verbal signaling languages which apply to the same problem as £ but which are sufficiently unlike £.

5. Meaning$_{nn}$ of Signals

H. P. Grice, in his paper "Meaning," draws our attention to the manifest difference illustrated in the following contrasting pairs.[3]

(1A) Herod presents Salome with the head of St. John the Baptist on a charger.

[3] *Philosophical Review*, 66 (1957), pp. 377–388.

(1B) Herod says to Salome, "He's dead."
(2A) Feeling faint, a child lets its mother see how pale it is (hoping that she may draw her own conclusions and help).
(2B) A child says to its mother, "I feel faint."
(3A) I leave the china my daughter has broken lying around for my wife to see.
(3B) I say to my wife, "Our daughter has broken the china."
(4A) The bus conductor taps the pane behind the driver so that he will turn around and notice that the bus is full.
(4B) The bus conductor rings a bell three times to inform the driver that the bus is full.
(5A) A policeman stops a car by standing in its way.
(5B) A policeman stops a car by waving.

The difference we want must be a subtle one. Our A and B cases are alike in many respects. We can contrast them with these cases, for instance:

(1C) Herod leaves the head somewhere; Salome happens to see it.
(1D) Herod leaves the head where he knows Salome will see it, correctly supposing she will not realize he left it for her to see.
(1E) Herod leaves the head where he knows Salome will see it, mistakenly supposing she will not realize he left it for her to see.

In all our A and B cases, and also in (1C), (1D), and (1E), this first condition is met:

> Someone does some action that produces a belief or some other response in an audience.

In all our A and B cases, and also in (1D) and (1E) but not (1C), this second condition is met:

> He intends—expects and wants—to produce that response by his action.

In all our A and B cases, and also in (1E) but not (1C) or (1D), this third condition is met:

> The audience recognizes his intention to produce that response by his action.

In all our A and B cases, but not in (1C), (1D), or (1E), this fourth condition is met:

> He intends that the audience should recognize his intention to produce that response by his action.

The difference Grice has discovered between the B cases and the A cases is given by a fifth condition:

> He intends the audience's recognition of his intention to produce that response to be effective in producing that response. He does not regard it as a foregone conclusion that his action will produce the intended response, whether or not his intention is recognized.

This fifth condition holds in every B case, but not in the A cases and not in cases like (1C), (1D), and (1E).

Whenever the fifth condition holds, Grice says that the agent *means*$_{nn}$ something by his action. More concisely, someone means$_{nn}$ something by an action if and only if he intends the action "to produce some effect in an audience by means of the recognition of this intention [namely, to produce the effect]."[4] "Means$_{nn}$" is short for "means nonnaturally"—not to suggest anything supernatural, but by contrast with the so-called natural meaning of natural signs: spots that mean measles, smoke that means fire, and the like.

I would have pointed to another difference between the A and B cases. In the B cases, but not the A cases, the audience's response is produced by means of a conventional signal, given in conformity to a signaling convention. But I am in no disagreement with Grice; for meaning$_{nn}$ is a consequence of conventional signaling.

[4] "Meaning," p. 385.

Let σ be a signal of a conventional signaling system $\langle Fc, Fa \rangle$. Let there be a state of affairs s and a response r such that σ is $Fc(s)$ and r is $Fa(\sigma)$. Suppose I am the communicator and you are the audience in a signaling problem of the proper kind; and, having observed that s holds, I give σ in conformity to our convention. Suppose that the case is completely normal: none of the exception clauses provided in Chapter II.4 apply to it. Then I mean$_{nn}$ something by σ. I intend σ to produce your response r by means of your recognition of this intention. Proof:

> The intention with which I do σ can be established by examining the practical reasoning that justifies me in doing it. I need not actually go through that reasoning to have an intention; actions done without deliberation are often done with definite intentions. So examine my justification, in one of the versions considered in section 2. It was shown in Figure 45 (repeated below).
>
> My decision to do σ, having observed s, is premised on my expectation that I can thereby produce r and on my desire to produce r. So I do σ with the intention to produce r.
>
> I expect you to infer s upon observing that I do σ. I expect you to recognize my desire to produce r, conditionally upon s. I expect you to recognize my expectation that I can produce r by doing σ. So I expect you to recognize my intention to produce r, when you observe that I do σ.
>
> My expectation that I can produce r by doing σ is also premised on my expectation that you will recognize my desire to produce r, conditionally upon s, and on my expectation that you will recognize my expectation that I can produce r by doing σ. These two premises constitute both my expectation that you will recognize my intention and my reason for expecting my intention to be fulfilled. So I expect your recognition of my intention to be effective in producing your response. I do not regard it as a foregone conclusion that my action will produce r, whether or not you recognize my intention to produce r.

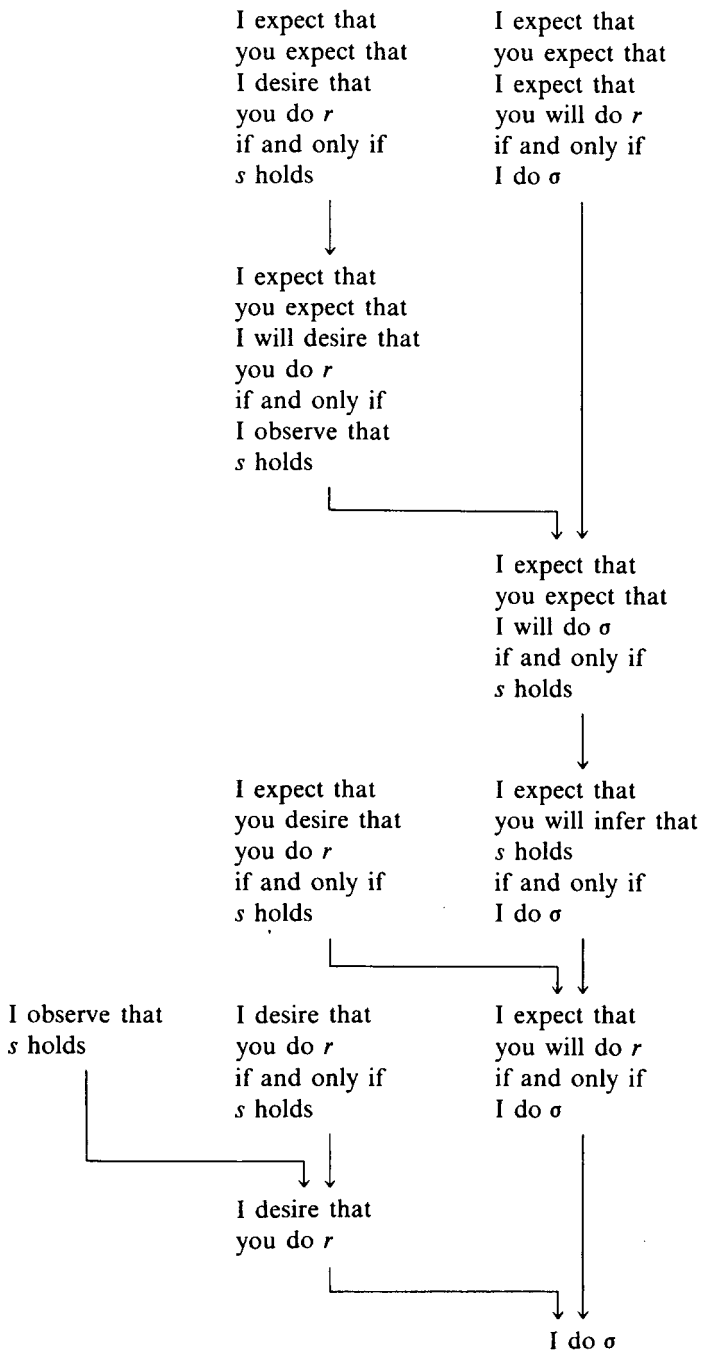

Figure 45

A conventional signal, given in normal conformity to the appropriate convention, must mean$_{nn}$ something. Is the converse true? No, for several reasons.

I might be deceiving you. My signal would mean$_{nn}$ something as usual, but I would be violating the convention.

Or I might be mistaken about the prevailing conventions and mean$_{nn}$ something by an action that I wrongly took to be a conventional signal. My reasoning, expectations, preferences, and hence intentions will look the same whether I am right or wrong. The only difference will be that I will fail to produce the intended response.

Or I might (rightly or wrongly) expect my audience to be mistaken about the prevailing conventions. That is what happens in this example from John Searle's "What Is a Speech Act?"

> Suppose that I am an American soldier in the Second World War and that I am captured by Italian troops. And suppose also that I wish to get these troops to believe that I am a German officer in order to get them to release me. What I would like to do is to tell them in German or Italian that I am a German officer. But let us suppose I don't know enough German or Italian to do that. So I, as it were, attempt to put on a show of telling them that I am a German officer by reciting those few bits of German that I know, trusting that they don't know enough German to see through my plan. Let us suppose I know only one line of German, which I remember from a poem I had to memorize in a high school German course. Therefore I, a captured American, address my Italian captors with the following sentence: "Kennst du das Land, wo die Zitronen blühen?" I intend to produce . . . the effect of believing that I am a German officer; and I intend to produce this effect by means of their recognition of my intention.[5]

It is inessential that the soldier is trying to deceive his captors. He could just as well be a real German officer who has forgotten the rest of his German from shell shock.

[5] *Philosophy in America*, ed. Max Black (Ithaca, New York: Cornell University Press, 1965), pp. 229–230.

Or I might mean$_{nn}$ something without thinking that I was conforming to a convention, and without expecting my audience to think so either. P. F. Strawson gives an example of this sort in his "Intention and Convention in Speech Acts."

> *S* . . . arranges convincing-looking "evidence" that *p*, in a place where *A* is bound to see it. He does this, knowing that *A* is watching him at work, but knowing also that *A* does not know that *S* knows that *A* is watching him at work. He realizes that *A* will not take the arranged "evidence" as genuine or natural evidence that *p*, but realizes, and indeed intends, that *A* will take his arranging of it as grounds for thinking that he, *S*, intends to induce in *A* the belief that *p* . . . He knows that *A* has general grounds for thinking that *S* would not wish to make him, *A*, think that *p* unless it were known to *S* to be the case that *p*; and hence that *A*'s recognition of his (*S*'s) intention to induce in *A* the belief that *p* will in fact seem to *A* a sufficient reason for believing that *p*. And he intends that *A*'s recognition of his intention . . . should function in just this way.[6]

In this case an abnormality would show up if we carried *S*'s justification of his action back one more replication. It is *not* the case that *S* expects that *A* expects that *S* expects that *A* expects that *S* will arrange "evidence" if he knows that *p*.

Or I might mean$_{nn}$ something without thinking that I was conforming to a convention, without expecting my audience to think so either, and with exactly the same sort of justification by replications as if I were conforming to a convention. Suppose I have come upon a patch of quicksand and I know of no conventional warning signal. I put a scarecrow up to its chest in the quicksand, hoping that whoever sees it will catch on. There is no convention to mark quicksand with half-submerged scarecrows; nor do I think there is or expect my audience to think so. But I do intend my action to produce awareness that this is quicksand by means of the recognition of my intention

[6] *Philosophical Review*, 73 (1964), pp. 446–447.

to produce that awareness. I have done my part of a signaling system in a signaling problem; and I hope my future audience will do its part. But if so, coordination will be achieved not by force of precedent but by force of salience.

Finally, I might mean$_{nn}$ something by an action in conformity to a convention of language, but not in conformity to any signaling convention. We have found it plausible that some use of language, but not all, is covered by verbal signaling conventions as well as by more general conventions of language. Even after we have considered what the general conventions of language might look like, I shall not try to show that any signal given in conformity to them must mean$_{nn}$ something; but I conjecture that this is true.

Searle draws this moral from his example of the soldier: "we must capture both the intentional and the conventional aspects [of communication] and especially the relationship between them."[7] I have been arguing that once we capture the conventional aspect, we are done. We have captured the intentional aspect as well.

[7] "What Is a Speech Act?" p. 230.