

NOTES ON THREATS, COMMITMENTS AND "MOVES"

by Glenn H. Snyder

In discussing threats and commitments, Schelling says:

"The threat differs from the ordinary commitment, however, in that it makes one's course of action conditional on what the other player does. While the commitment fixes one's course of action, the threat fixes a course of reaction, of response to the other player. The commitment is a means of gaining first move in a game in which first move carries an advantage; the threat is a commitment to a strategy for second move."¹

This is not very helpful. It implies that a commitment always binds one to do something "no matter what", i.e., regardless of the other party's behavior. There are some commitments of this sort, as when a country becomes "committed" to go to war regardless of the other party's pre-war behavior. But most commitments of interest in the bargaining process are commitments to "second move", i.e., they are commitments to do something in a certain contingency; whether the action is taken depends on whether the adversary does or does not do something.

But this is virtually equivalent to Schelling's definition of "threat" as a commitment to second move. This definition is defective, first, because a threat does not necessarily have to carry a commitment with it, and secondly, because it confuses the essential character of threats and commitments. A threat is an act of communication; a commitment essentially is an arrangement of incentives such that it is rational for one to carry out the act one is committed to. This is the essential distinction, not whether one is committing oneself to first move or second move.

1. Thomas C. Schelling, The Strategy of Conflict (Cambridge: Harvard University Press, 1960), p. 124.

Commitment

The term "commitment" can refer either to a state of being committed or an act of committal. It is useful to keep these two meanings separate. One is in a state of commitment (i.e., one is committed) when one's incentive structure is such one would carry out a particular act--it is "rational" for one to do so. An "act of commitment" is an act which modifies a party's incentive structure so that he becomes committed in this sense when he was not before.

There are two kinds of acts of committal, broadly speaking. One is a change in one's values (incentive structure) which makes it less costly to perform the act in question than to fail to do so. The other is an act which forecloses all alternatives except the act in question, without changing the underlying value or payoff structure. In effect, the latter amounts to "crossing off the matrix" all options but the one committed to.

Schelling uses the term "irrevocable commitment" in both of these senses, although most often he seems to have the latter meaning in mind (as in "burning bridges" or maneuvering one's army up against a deep ravine). Sometimes, however, he uses the other meaning as when a party engages enough values behind a threatened course of action (prestige bargaining reputation, etc.) that the cost of non-fulfillment becomes prohibitive. Both meanings are legitimate, of course, but they should be kept separate.

One puzzling question is whether it is possible or correct to speak of different degrees of commitment. Schelling's employment of the adjective "irrevocable" suggests that there are degrees, that there are other kinds of commitment which are not irrevocable. And in

common usage, there are frequent references to different degrees of "firmness" of commitments. But in the sense in which the term is defined here, commitments are always "irrevocable". A rational player does not violate his incentive structure, or alternatively, he cannot choose other options if these options are foreclosed. What can be said to reconcile this definition with common usage?

It is possible that the value structure of an actor is such that the act committed to is favored by only a small margin. Then an actor is committed, but not "firmly" in the sense that a small change in his payoff structure may de-commit him, and such small changes can easily occur in the course of bargaining or decision-making. In this sense the degree of firmness is a function of the size of the "gap" between the costs of fulfillment and the costs of non-fulfillment.

Another source of "firmness variability" may lie in the actor's uncertainty or vagueness about his own incentive structure. At the time of committal he may be uncertain as to whether his values have actually been re-arranged so as to bind himself. Furthermore, he may be uncertain (even if he is certain at the time of committal) whether his incentive structure will stay put during the time period between committal and the occasion for fulfilling the commitment, or whether, when the occasion for fulfillment confronts him, his values may not appear entirely different than they did prior to that occasion arising. He knows that many things can happen during the process of bargaining or interaction which may change the values he attaches to particular acts or outcomes.

In colloquial usage, the notion of "degrees of firmness" may refer to the degree of ambiguity in threats to an enemy or pledges to an ally. Maximum ambiguity leaves maximum opportunity to renege when the

occasion for fulfillment arises. Ambiguity may be decreased (and "firmness" increased) by a sequence of actions or declarations over time. Thus, Britain was not very firmly committed to the Triple Entente before military conversations with France were begun, but became more firmly committed (although perhaps not "irrevocably") after the start of these talks and the division of tasks between the British and French fleets.

Finally, degrees of firmness may refer to the perceptions of others as to the probability that the actor is committed or not. Thus, regardless of whether the actor is actually committed or not, he is thought to be "firmly" committed if the probability of his carrying out the act in question is thought to be high, but "not so firmly" if the probability is low. However, this interpretation is somewhat unsatisfying theoretically because we have defined "commitment" in terms of the actor's own options and incentive structure, not in terms of their perceptions by others.

In bargaining theory, "commitment" usually refers to something that one does vis-a-vis an adversary. However, in ordinary parlance, it most often refers to a pledge or promise that is made to an ally or non-enemy party--typically a pledge to defend that party against the common enemy. When people refer to "the commitments of the United States" that is usually what they have in mind. To "be committed" in this sense means that other parties hold expectations about one's future behavior which have arisen as the result of a past promise or pledge. The commitment has a certain moral quality. To renege on the pledge would incur high moral costs, so high that one does not have the alternative of reneging, unless the other party "releases" one from the pledge.

However, in international relations, commitments to allies and friends are at the same time commitments vis-a-vis an enemy, so there is no contradiction between the theoretical usage and the ordinary usage. They are simply two aspects of the notion of being committed. I.e., they represent two sets of costs which are suffered if the commitment is not fulfilled--one vis-a-vis the enemy and one vis-a-vis the ally. And for both aspects it is still possible to distinguish the act of committal from the state of being committed. The act of committal faces two ways--toward the ally and toward the enemy--but the "state of being committed" is, in a sense, a "commitment to oneself" since it results from a particular arrangement of one's own incentive structure. In this sense, it produces an intention. Thus there is close relation between the concept of intention and the concept of commitment, although, as we shall point out presently, neither intentions nor commitments necessarily result from specific "acts of committal."

Threat

A threat, strictly speaking, is simply an act of communication. It specifies a demand, plus a sanction that is to be inflicted if the demand is not met. Thus a threat is quite different than a commitment, which is not essentially a communicative act by a particular arrangement of values. However, the two are linked because a threat may create a commitment by engaging certain values of the threatener which would be sacrificed if the threat is not carried out. We might then distinguish between the communication quality of a threat and its committal quality. In its communicative dimension, a threat is simply the conveyance of information about purported intentions. Its committal quality refers to its effect on the incentives and preferences of the threatener.

Schelling distinguishes quite sharply between a "threat" and a "warning." For him, "the distinctive character of a threat is that one asserts that he will do, in a contingency, what he would manifestly prefer not to do if the contingency occurred. . . . Furthermore, a threat seems always to incur a commitment. "Like the ordinary commitment, the threat is a surrender of choice, a renunciation of alternatives, that makes one worse off than he need be in the event the tactic fails. . . ."² In short, before the threat is made, the threatener has no incentive to do what is threatened, but the making of the threat alters the actor's incentive structure so that he then becomes "committed."

A "warning", on the other hand, according to Schelling, is a statement that one will do something painful to another party if he fails to comply, but in situations where the incentives to carry out the punishment already exist before the statement is made. One simply "points out" to the other party, or "reminds" him, that these incentives exist. In our terminology, a warning has a communicative dimension but no committal dimension. Schelling argues that is useful to keep warnings separate from threats because they do not constitute "moves", since they do not change the warner's incentive structure.³

It seems questionable whether this distinction is useful. In ordinary usage, the term "threat" is often applied to declarations which Schelling calls "warnings." It is not advisable to deviate from common usage unless there are good theoretical reasons for doing so, and it seems doubtful that such reasons exist in this case. Warnings and threats both have the same communication structure--a demand and a

2. Schelling, op. cit., p. 123.

3. Ibid., pp. 123-124.

sanction. In both cases, as Schelling admits, there remains the problem of convincing the party that the incentives for actually carrying out the threatened (warned) action do in fact exist. Empirically, it will often be difficult to distinguish a warning from a threat by Schelling's definitions.

If we drop Schelling's condition that a threat must always create a commitment, (and I think we should) then it becomes unnecessary to distinguish so sharply between warnings and threats. A warning then appears only as a particular kind of threat, not a class of declaration different from threats. By dropping this condition of Schelling's it becomes possible to distinguish three classes of threats according to associated effects on incentive structure, as follows:

Type I: Threats in which the incentive to carry out the threat already exists prior to the communication (Schelling's "warning").

Type II: Threats which modify the incentive structure of the threatener but not enough to commit.

Type III: Threats which modify the incentive structure enough to create a commitment which did not exist before (Schelling's "threat").

(Thus we have added a third class--Type II--which does not appear at all in Schelling's definitions, a type which, one strongly suspects, comprises the great majority of "threats" in the real world)

We can now reserve the term "warning" for another class of declarations which are consistent with the ordinary meaning of the word and which need a label. These are statements which convey information about something bad which may happen to the other party (or to both parties) which is assertedly beyond the control of the first party. E. g., "the war could not be limited", or "events may get out of hand", etc. While such declarations may be intended to influence the behavior

of another, they are not threats, but rather statements which draw attention to a shared risk.

Communications, Incentive Structures and Perceptions

The discussion above suggests that it is useful to distinguish between three separate phenomena in the bargaining process: the incentive structures of the parties (their "payoffs" for alternative outcomes), the perceptions of each party about the other's incentive structure, and communications. The primary role of communications is to modify incentive structures or the other party's perceptions about them.

Incentive structures and perceptions exist independently of communications and are more basic than the latter. Ultimately, behavior depends on the incentives and perceptions of the parties; although the bargaining process usually is thought of as a process of communication, communication tactics are subordinate in that they only modify, and perhaps only marginally, the other two basic elements. Incentive structures and perceptions are implicit in any situation even if nothing is explicitly communicated. They reside in "interests" which are to some extent known to or perceived by the other party. Known interests imply incentive structures which in turn imply the alternatives for action of each party and the probability that certain alternatives will be chosen. E.g., the United States will defend India and the Russians and Chinese undoubtedly perceive this even though we have made no explicit "threats" or "commitments" concerning India.

In short, incentive structures and beliefs about them are the basic stuff of international politics, they are relatively permanent and continuous through non-crisis and crisis periods, and they are the basic determinants of behavior in a crisis. Communications may be

thought of as a kind of "veneer" of explicit activity which serves to modify marginally these basic determinants, although such modifications may be extremely important in a crisis.

Moves

Schelling defines a "strategic move" as follows:

"A strategic move is one that influences the other person's choice, in a manner favorable to oneself, by affecting the other person's expectations on how one's self will behave."⁴

This is too limited. It includes modifying one's own incentive structure and the other's perception of it, and influencing the probabilities the other attaches to one's strategy choices. But it leaves out modifying the other party's incentive structure. It also seems to omit those moves which change the alternatives confronting the other party, or perhaps both parties. That is, the Schelling definition seems to have behind it an image of a single matrix, and does not explicitly include the possibility of moves which create a new matrix, in effect a "new game" by changing the situation so that the parties confront a new set of alternatives.

Furthermore, the Schelling definition applies only to "strategic moves", i.e., moves intended to coerce or influence. It does not include what I would call "bidding moves". These are demands, offers and proposals for settlement which the parties make in the process of moving toward a compromise. The act of "capitulation" or "concession" is a bidding move. Often, bidding moves are small or marginal concessions prompted by the other's strategic pressures. In short, "strategic moves" are the application of power behind given bids. Bidding moves involve

4. Ibid., p. 160.

a changing of the bids (although they may also change perceptions of the strategic power supporting the bids.)

I would redefine strategic moves as follows: "A strategic move is one that attempts to influence the other person's choice, in a manner favorable to oneself, by affecting his perceptions of one's own incentive structure and the probabilities of one's own alternative choices, or by modifying the other party's incentive structure, or by changing the basic alternatives available to the other or both parties."

This is rather long and clumsy, but it captures much more of the bargaining process, in a more explicit way, than the Schelling definition. Having made this definition, it is useful to break down "strategic moves" further into "basic moves" and "communication moves."

Basic moves are moves which irreversibly change the structure of the game--in terms of alternatives, payoffs and probabilities (and perceptions of them) but particularly in terms of alternatives. In effect, they set up a new game with new alternatives for each party. Examples of basic moves would be the establishment of the Berlin blockade in 1948-49, the counter-move of the airlift, the convoying of Chiang's ships to Quemoy island in 1958, the Soviet emplacement of missiles in Cuba, the establishment of the U. S. blockade in the Cuban crisis, the construction of the Berlin Wall, etc. There are likely to be only a few of these basic moves in any particular crisis, perhaps no more than two or three.

Basic moves can be conceived in terms of "decision points" on a "game tree." Theoretically, at the beginning of a crisis, each party could visualize a sequence of possible basic moves and responses. Also theoretically, this imagined sequence could be "collapsed" into a "game in normal form" with the payoffs representing the values of ultimate

outcomes, each of these payoffs being appropriately discounted by the probabilities of certain moves being taken during the course of the sequence. A decision point is an occasion for making a basic move. At each decision point, each party could theoretically have in mind a matrix in normal form, with the usual set of alternatives and payoffs, the payoffs representing expectations about ultimate outcomes, appropriately discounted by probabilities attached to future moves. The basic move taken at each decision point changes the alternatives available to the parties and in general rather sharply modifies the ultimate matrix visualized by each of the parties.

By way of example, when Russia confronted the "decision point" of "mobilize or not" in 1914, she had to consider, in calculating the payoffs, the probabilities and ultimate costs associated with the following possible moves by other parties: German mobilization alone, German attack, French participation (or non-participation) on the side of Russia, British participation (or non-participation), etc. Once the Russian move was made, each of the parties confronted a "new game" i.e. a new set of alternatives, a new set of probability calculations, and a new set of visualized payoffs for the ultimate outcome. Similar statements could be made, say, for the U. S. decision point of "blockade, air strike, or invasion" in the Cuban missile crisis. After the blockade was established, the U. S. was at the decision point "sink Russian ships or not", and conceivably after that one, she would have been faced with "air strike or not", etc. The Russians, of course, would have been faced with their own sequence of decision points: run the blockade or not, sink U. S. ships or not, retaliate after U. S. air strike or not (and how), etc.

Communication moves differ from basic moves in that they do not irreversibly change the fundamental alternatives available to the parties

and they modify payoffs and probabilities (and their perceptions) only marginally, not drastically. Putting it another way, communications moves modify payoffs and perceptions in a given game; they do not transform the situation into a different game (different matrix). Communications moves may be designed to influence the other party's behavior at his next decision point. I.e., the basic game is that which exists at the next decision point, and communication moves may be designed to influence the other party's payoffs, and his estimate of one's own payoffs and probabilities of choices in that game. However, ~~it~~ would not be entirely accurate to limit communications moves to the immediate decision point, since they may also have reference to subsequent decision points (E.g., Kennedy, in his TV speech, not only warned the Russians of the consequences of attempting to run the blockade but also the consequences of not taking the missiles out of Cuba.)

Having distinguished between basic moves and communication moves, the next question is what ~~kind~~ of moves to put into each category. Are basic moves always physical? Is an ultimatum a basic move? A solemn declaration of firm support for an ally? It appears that basic moves are always physical since only physical moves can initiate a line of action which is irreversible and which changes the basic structure of the game. Verbal moves, no matter how solemn or dramatic, can always be retracted.

In this light, moves of "automation" or "burning bridges" are always physical acts of some sort and are basic moves. If the U. S. did not ~~have~~ troops in Europe, but in a crisis introduced them, this would be a basic move because it would irreversibly change the alternatives open to the U. S. and would drastically alter the Russian's perceptions of payoffs and probabilities.

However, the definition of "physical" needs to be quite broad. If the U. S. delegated to lower commanders the authority to fire nuclear weapons this would be a basic move. A U. S. nuclear response would have become not quite automatic but the probability of it would increase sharply in the Soviet minds.

Are communication moves always verbal? Obviously not, because shows of force, visits by dignitaries, etc. are communication moves. One might say that all moves which do not change the fundamental alternatives confronting the parties are communication moves. This would make the transformation of alternatives the essential criterion for distinguishing between the two types of moves.

Still, there is a grey area. Threats which engage enough values to commit the threatener irrevocably are pretty basic and they do sharply modify the alternatives available to the threatener. And it is also conceivable that some physical acts of commitment or automation could be retracted. (In case of a Russian attack in Europe, the U. S. could withdraw its troops. The U. S. Navy in the Cuban blockade could have been ordered not to fire on a Soviet missile-carrying ship which refused to stop. Actually, it is hard to think of acts in international politics which are absolutely irreversible short of the actual use of violence). Thus, in terms of ability to commit, certain verbal acts may be stronger than some physical acts. Or, the committal quality of certain physical acts may reside not so much in their physical irrevocability but, as in the case of threats, in the costs which would be incurred if they were retracted. Thus, the distinction is not absolutely clear-cut, but nevertheless it seems valid enough as a point of departure.

The answer to the question Which moves to select for in-depth

analysis in our case studies? now becomes somewhat clearer. Obviously, basic moves must be analyzed exhaustively. As for communication moves, all which are intended to modify the adversary's valuation of outcomes (payoffs), or change his perception of one's own payoffs and probable strategy choices, must at least be recorded. Those which have major effects in these respects must be analyzed in detail. We might call these primary communication moves. To analyze in detail means at least to ascertain (1) the decision-making process and factors which led to the choice of move, (2) the intended effects of each component of the move in terms of "bargaining function", (3) the actual effects on the communicator's intentions, (4) the effects on the perceptions and payoffs of the recipient, and (5) effects on the behavior of the recipient--i.e., effects on his choice of his next move.

A typology of communications moves and associated "bargaining functions" will be presented in a subsequent paper.