Winning Versus Doing Well

Russell Hardin

In both popular and sophisticated circles, the theory of games is associated most prominently with the postwar nuclear standoff between the United States and the Soviet Union. Unfortunately, entry to game theory is not as quick and easy as it too often seems; much of the understanding of the nuclear dilemma has been clouded by crude accounts of verbal game theorists. That is a perverse result, because the value of game theory is that it can clear up conceptual confusions; yet in this realm it seems to have fathered as many confusions as it has cleared up. I wish to discuss two that are particularly pervasive and destructive of understanding.

The first is to think of strategic interactions as games like chess and poker, games that, when played well, are to be won. This confusion we owe directly to John von Neumann, the inventor of game theory, for so naming his analysis of the logical structures of strategic interactions. It is too late in the day to change the name, but it is perhaps not too late to stop thinking of poker and chess when we are involved in strategic interactions of the more interesting sorts that fill our lives.

The second confusion may be related to the first: It is the tendency to see games as played once and in isolation from each other. It may take victories in a series of games to win certain titles or contests, such as the world championship in chess; but each game can be played independently of the others. Play in one game influences play in another in such a series only in that one may learn from one game how an adversary plays and then use that knowledge to play more successfully. Otherwise, the games in a sequence of chess or poker games are strategically decoupled. In political relations, games are seldom decoupled—plays today affect plays tomorrow and, more importantly, anticipation of tomorrow affects actions today.

These two confusions are heavily intertwined in our misunderstandings of various social interactions, especially in our frequent misunderstanding of the nuclear dilemma. I wish to address them in turn and then to bring them together in trying to assess more clearly the actual structure of the nuclear dilemma.
GAMES AS COMPETITIONS

The word *game* provokes thoughts of competition. Game theory is far more general than this association suggests. It is intended to represent strategic interactions of all kinds: conflict, cooperation, and coordination. Typical social interactions often involve elements of all three of these. Indeed, cooperation may best be defined as the resolution of partial conflicts when there are possibilities of coordination. That sounds like a recipe for daily life—and for international relations.

It has commonly been assumed, without much critical ado, that the superpower conflict is a *Prisoner's Dilemma* or *Chicken* interaction. Certain other obvious game theoretic possibilities are almost never canvassed, presumably because casual game theorists take the unfortunate term "game" too seriously and focus on winning rather than on doing well, which generally is the point of life. This is a blunder of potentially grievous proportions. One might be confident of winning a fight with a young bobcat—but one would do well to avoid it.

The most important check to the casual game theorists' prescriptions on nuclear weapons may be to take more seriously the task of analyzing the plausible game structures of the nuclear dilemma. Let us begin such an analysis by supposing that we can specify the values of the outcomes of various strategies by each side. For many purposes it will be sufficient to specify those values only ordinarily, that is, to say that one outcome is preferred to another, without saying by how much it is preferred.

The first insight of game theory is that, of course, neither side can choose the outcome it wants. Rather, each can only choose a strategy and the outcome that is reached is the joint result of the strategy choices. Much of the discussion of nuclear weapons policy focuses on the choice of strategies almost as though it were strategies and not outcomes that are of interest and that should be ranked in our preferences. On the contrary, the only reason for concern with strategies is that they will eventually determine outcomes in their interaction with the strategy choice of the other side. To begin at the beginning, therefore, we should rank outcomes. When we have done so, we will automatically have determined the form of the game or strategic interaction in which we are involved.

As a crude first step, we can suppose that the outcomes of interest are four: that in which both sides are armed, that in which both sides are disarmed, and those in which one side is armed and the other disarmed. What are plausible constraints on our ranking of these? Three that should be unobjectionable are as follows.

The first is that the Soviet and American preference orderings over these outcomes are symmetric. A game representation of the interaction will be anonymous in the sense that, if we were to switch the labels of the two sides in the game matrix, the new matrix would still represent the interaction. This is a powerful constraint that most people seem to find compelling in the abstract but that many commentators violate in their actual discussions of the nuclear dilemma. For example, many Americans seem to think that the Soviet Union wants victory whereas the United States merely wants peace; many Russians reputedly hold the contrary view. The implication is that one side supposedly prefers mutual armament to mutual disarmament while the other side has the opposite preference. It is easily conceivable that particular leaders on the two sides could have such asymmetric views, yet it seems implausible that they could correctly hold such views if they were sufficiently concerned with the loss of life that would result from an actual nuclear conflict, and if they had similar expectations about the likely implications of war for the two regimes. In what follows, I will assume that both parties have similar views on the objective facts of various regimes. It may be of interest, of course, to assess to what extent this assumption is correct.

The second constraint is that each side prefers mutual disarmament to mutual armament. Under present conditions, it is implausible to suppose that either side would disarm fully in a regime of mutual disarmament. Rather, each would simply disarm down to some much lower level that would still secure deterrence. Hence, mutual disarmament would involve much lower expenditures for nuclear weapons and, more importantly, a much smaller likelihood of a devastating nuclear war that might result from unwonted reactions in a tense situation. What makes it difficult to be sure that both sides prefer mutual disarmament to mutual armament is that both steadily pursue policies of arms expansion. However, they presumably do so primarily because each has control only over its own policies and cannot rely on the other not to expand. Given the next constraint, they face a risky choice that may recommend increased arms even when they would prefer mutual disarmament that could be safely guaranteed.

The third constraint is that each side would most prefer to be armed while the other is not armed. I think that the applicability of this constraint is beyond question. There may be a philosopher or saint on one side or the other who has the contrary preference, but it is implausible that any leader, East or West, in a position to influence policy over the past several decades has had the contrary preference. Nor can there be very many citizens in the Soviet Union or the United States who have the contrary preference.

What do these three constraints do for our understanding of the nuclear dilemma? They tell us what strategic structures that dilemma can have. The dilemma can be represented as a simple two-by-two $(2 \times 2)$ game, that is, a game in which two players have two strategy choices each. If each player ranks the outcomes ordinarily without ties, there are 576 possible structures for such $2 \times 2$ games. (There are four possible outcomes in such a game.) The three constraints above reduce the 576 possible structures to only three that we need
consider (see Appendix). They are presented in Table 9.1. In each game in the table, D represents the strategy of disarmament and A represents that of armament. Each outcome presents the ordinal payoff to each party, with 1 as the most-preferred and 4 as the least-preferred outcome.

If our constraints are reasonable, it should come as no surprise that the three games in Table 9.1 include Prisoner's Dilemma and Chicken, the two games most often associated with the superpower dilemma. However, they also include one unequal coordination game (Game 3) that we might call the Nuclear Pax. The two best outcomes in it are armament by one superpower and disarmament by the other. It is a coordination game because both sides prefer either of these outcomes to the other outcomes, but the coordination is unequal because one side would prefer a Nuclear Pax Americana and the other side would prefer a Nuclear Pax Sovietica.

Briefly consider the three games in Table 9.1. In each game, stable or equilibrium outcomes are circled. An outcome is stable if neither party has an incentive to switch strategies unilaterally to try to improve its payoff. An outcome may be stable even though both parties could be made better off by switching their strategies simultaneously. For example, in the Prisoner's Dilemma (Game 1) both parties would be better off if they switched from the stable outcome of mutual armament to the unstable outcome of mutual disarmament. The latter is unstable, however, because each party would prefer unilaterally to switch its strategy from disarmament to armament if they have reached the outcome of mutual disarmament.

The superpower dilemma is a Prisoner's Dilemma if, in addition to the three constraints above, each side considers the outcome in which it is unilaterally disarmed to be the worst of all the outcomes. This would be true if each side suspected the other would attack it in order to preempt the possibility of its rearming with sufficient probability and sufficient devastation as to outweigh the risk of a regime of mutual armament. Until very recently, this was prob-

ably the way virtually all leaders and most citizens on both sides viewed the nuclear dilemma. The regime of mutual armament has seemed relatively stable in the sense that it has seemed unlikely to lead to nuclear war. But neither side would have trusted the other as a sole nuclear power.

However, recent speculations about just how devastating a substantial nuclear war would be seem to have shaken many in this sanguine view. It is plausible that the worst of all outcomes is that in which both sides are armed, if a substantial nuclear exchange would be as devastating as recent scenarios on nuclear winter suggest and if the regime of mutual armament is as unstable (in the sense of its leading to war) as some have begun to believe in the light of recent American and Soviet arms build-ups. Hence, the superpower dilemma would not be Prisoner's Dilemma but, rather, either Chicken or Nuclear Pax. In both of these games, as in Prisoner's Dilemma, the regime of mutual disarmament is unstable. The game would seem to be Chicken if the instability were to lead to dominance by one side or the other. It would seem to be Nuclear Pax if the instability were most likely to lead to a renewed and explosive arms race as each tried to be first as the sole armed power. Again, the degree of certainty of knowledge of the other side's plans and actions might be the chief factor in determining which game we are in. If we are sufficiently uncertain, then each may fear a regime of mutual disarmament almost as much as that of mutual armament. In this case we are in the Nuclear Pax game and both sides may have great incentive to strive to beat the other to become armed enough to be the sole hegemonic power.

In sum, if our uncertainties are grievous, we must be in either the Prisoner's Dilemma or the Nuclear Pax. If, additionally, the likelihood that a regime of mutual armament will lead to nuclear winter is high enough, we are in the Nuclear Pax.

The final prospect is so different from anything that has been taken seriously into consideration over the past several decades that its significance has yet to be faced in the literature. Oddly, however, it is a view that makes consistent two apparently different positions taken by Bertrand Russell. In Common Sense and Nuclear Warfare, Russell assumed that all would concede the following three propositions:

1. A large-scale nuclear war would be an utter disaster, not only to the belligerents, but to mankind, and would achieve no result that any sane man could desire.
2. When a small war occurs, there is a considerable risk that it may turn into a great war; and in the course of many small wars the risk would ultimately become almost a certainty.
3. If all existing nuclear weapons had been destroyed and there were an agreement that no new ones should be manufactured, any serious war would, nevertheless, become a nuclear war as soon as the belligerents had time to manufacture the forbidden weapons.²
Russell concluded that, "if we are to escape unimaginable catastrophes, we must find a way of avoiding all wars, whether great or small and whether intentionally nuclear or not" and that "an International Authority ... must be the ultimate goal of all those who wish to preserve the world from the disasters of nuclear war."

Earlier, Russell had supposed that the future of mankind depended on having the United States seize world hegemony while it had a near monopoly on nuclear weapons. The reasoning was very much the same as in his later argument: The alternative to hegemony by a single power is catastrophe because mere disarmament is unstable. Once the Soviet Union developed its own nuclear weapons and delivery systems, however, the form that hegemony would have to take was through mutual creation of a supranational agency. In the late 1940s, the United States could possibly have acted alone to become the dominant power in the world and to have prevented any other from becoming a nuclear threat.

In both his accounts, Russell seems consistently to have viewed the situation as either the Chicken or the Nuclear Pax game. He himself introduced the notion of the Chicken game in this context. However, the relative certainty with which he foresees catastrophe, even if we achieve a voluntary regime of mutual disarmament, suggests that his view better fits the Nuclear Pax game. He supposes that the results of a nuclear war would be as grim as the nuclear winter and that such a war under a regime of either mutual armament or mutual disarmament is nearly certain. Hence, sufferance under a hegemonic power would be preferable to either mutual armament or mutual disarmament.

Suppose for a moment that Russell's view is generally right. There is presumably a real difference between being a citizen of the hegemonic power and being a citizen of a former superpower now under the sway of the sole nuclear power. If the difference is great enough, it may seem worthwhile to run some real risk of nuclear war while jockeying to become the hegemonic power rather than acquiescing in the hegemony of the other superpower. In playing the game in this risky way to try to do better, however, one may too readily lose sight of the issue of merely doing better and tend to focus on the competition for victory. That is, if the only acceptable outcomes are the lower-left and the upper-right cells of Game 3 (or Game 2), we may tend to look at the games as a choice merely between these two outcomes and to forget about the other possibilities.

If our choice were genuinely between only these two outcomes, we would be in a game of pure conflict. Chess and poker are games of pure conflict and, as we all know, the point of playing these games is to win. There are no possibilities for point gains; there is only the possibility that one of us gains at the expense of the other. Much of the contemporary rhetoric about nuclear arms policies in the United States seems to be based on the assumption that these are indeed the only outcomes at stake: Either the United States will prevail or the Soviet Union will prevail. But, while striving to win, we may neither one do well because we may increase the prospect of our winding up in the lower-right cell of mutual disaster.

Herman Kahn says that, according to Russell, such striving to win in Chicken is the game of "degenerates in America and nations everywhere." Kahn defends the practice on the part of statesmen as necessary: "One must be willing to play the 'game' in some form or another, or surrender." Winning is evidently all there is.

In the literature of international relations, especially of the so-called realist school, it is often argued that international relations are an anarchic Hobbesian war of all against all. Robert Keohane supposes that this implies that "cooperation on the basis of shared purposes should not exist except as part of a larger struggle for power," so that allies would cooperate with each other but there would be no other cooperation. If there is any lesson to be gained from game theory in this context, it is that this claim simply does not follow unless international relations are a matter of pure conflict without any hope of mutual benefit. There is no representation of the superpower dilemma over nuclear weapons that can sensibly be considered pure conflict. Only a blinkered focus on the two theoretically stable outcomes of Game 3, the Nuclear Pax, would lead one to see the superpower dilemma as a game of pure conflict.

**DOING WELL IN THE LONG RUN**

Perhaps the most striking difference between games of pure conflict like chess and poker and strategic interactions like international politics is that, in the latter, one's choice of strategy may depend profoundly on whether there are to be further plays. If I have a winning hand in this round of poker, I should play it for all it is worth, for I will gain nothing in the next round by playing down my present strength. However, in a game in which there are potential gains from coordination, or, in the language of traditional economics, gains from trade, playing several rounds of the game can affect one's choice of strategy in the first and subsequent rounds. Given the second constraint on the superpower dilemma—that both would prefer mutual disarmament to mutual armament—it follows that the dilemma is not an interaction of pure conflict. Indeed, that is what makes it a dilemma. Hence, it is a mistake to suppose we could understand what strategy is best in the interaction simply from the strategic analysis of a one-shot play of the game.

Reexamine briefly Russell's second proposition: When a small war occurs, there is a considerable risk that it may turn into a great war; and in the course of many small wars the risk would ultimately become almost a certainty. This is a conclusion from the law of large numbers: What is only slightly likely to occur on one occasion is almost certain to occur in one of a large number
of similar occasions. Russell supposes that repetition of the conditions of the nuclear stand-off over many years will eventually produce nuclear war. In doing so, he supposes that what happens in each of the repetitions must be essentially independent of what happens in the others, that these events are strategically decoupled. Although commonplace, this is a fallacy of often staggering proportions.

The implication of the fallacy in Russell's discussion is that the regime of mutual disarmament is seen as no better than that of mutual armament. In either regime, Russell supposes, we will eventually suffer catastrophe. In at least one of our plausible representations of the nuclear dilemma, the Prisoner's Dilemma of Game 1, this conclusion does not follow. Indeed, we know very well that, in many interactions that have the strategic structure of the Prisoner's Dilemma, iteration over many plays of the game produces cooperation, as a careful strategic analysis suggests it should. If the superpower dilemma is a one-shot Prisoner's Dilemma, then we choose now and for all time what we are to do and the resulting outcome inescapably determines our payoff. The usual analysis of the one-shot Prisoner's Dilemma strongly recommends cooperation if an enforceable agreement can be reached, and noncooperation if no such agreement can be reached.

As a rule, it is reasonable to assume that no enforceable agreement can be reached between the United States and the Soviet Union. We can, of course, sign treaties, but there is no higher power available to enforce the treaties. Indeed, arms control treaties between the United States and the Soviet Union generally include a clause permitting easy unilateral abrogation. Beginning with the Limited Test Ban Treaty of 1963, arms control treaties have had the appearance of a "marriage contract with a protocol for divorce." They are clearly so far from being enforceable in a meaningful sense that they leave future actions and choices of strategy open. Hence, available treaty mechanisms do not permit us to consider the game as played once and for all. These mechanisms are useful only as part of an ongoing relationship of iterated play. It follows that, if the superpower dilemma is a one-shot Prisoner's Dilemma, each side can serve its own interest best by refusing to cooperate and armimg to whatever level it thinks best deters the other side.

To suppose that the interaction is one-shot, however, is prima facie silly. The superpowers did not decide on their strategies for all time in, say, 1954. Rather, they have repeatedly adjusted their armaments, not only to adapt to changing technology, but more importantly to adapt to each other's strategy choices over time. We have played the game as though our strategy choices were free to be changed at virtually any moment. Indeed, a large part of the explanation of the rapidity of technological change in weapons and delivery systems over the past three decades is that each side has wanted to change its position relative to the other. Our interaction has been iterated and dynamic, not one-shot.

However, game theoretic analyses that suppose that the interaction is Prisoner's Dilemma or Chicken argue more often than not exclusively from the one-shot analysis of those games. Even to mention the iterated game, as Anatol Rapoport does, is rare. Verbal game theory in this realm has typically abused, rather than aided, our understanding. It may not have radically influenced policy over the past decades, but it may have defaulted on any possibility that it could have provided a positive influence. In particular, the careful game theoretic understanding of how best to play in an iterated Prisoner's Dilemma has largely been ignored in political and strategic circles. Quick and dirty game theory has been dismal game theory.

Suppose the structure of the superpower dilemma is that of the iterated Prisoner's Dilemma. Each side has strong incentive to play cooperatively in this round of the game in order to encourage the other to be cooperative in future rounds. Here it is necessary to distinguish two aspects of the interaction that, in a sense, split the interaction into two simultaneous games with different payoffs: those from using the weapons and those from developing and deploying them. These are not equivalent. The latter are a function of the former and of the likelihood that the weapons will be used.

First, consider the game of using the weapons. If we do not have a first-strike system capable of winning a war, then the payoffs from each side's using the weapons are essentially sure to be the destruction of both sides. In the first round of play, one side attacks; in the second round, the other side retaliates. Absent a genuine first-strike force, and the game of using the weapons is inherently iterated, not one-shot. If both sides are playing this game rationally, they cannot plausibly choose other than to cooperate so long as the other side cooperates—that is, they must choose to hold the weapons only for retaliation. Because it is inherently iterated, the usual analysis of the one-shot Prisoner's Dilemma—that one should defect—fails to apply to the game of using nuclear weapons.

Another way to put this issue is to say that each side's actions in the game of using nuclear weapons are in certain respects contingent on the other side's actions: If one uses the weapons, the other will too. Hence, we can probably suppose that the asymmetric outcomes in which one superpower dominates the other are ruled out as possibilities. Therefore, the interaction reduces to a simple coordination problem in which both sides fare badly or both sides fare well. This seems to be the way the interaction is viewed by many people, such as Robert McNamara and George Kennan from past administrations in the United States and Richard DeLauer from the Reagan administration, although one should be cautious in attributing precise, clear understandings to those whose statements are, at best, put in ordinary language.

Note that when we restricted consideration, above, to the two desirable outcomes in Game 3, Nuclear Pax, we produced a game of pure conflict. Now let us restrict consideration to the two outcomes in Prisoner's Dilemma in
which the players receive at least as good payoffs as they can guarantee themselves by their own independent, unilateral action. These are the outcomes at lower right (3,3) and upper left (2,2). These two outcomes in isolation comprise a game that is pure coordination. It would be ill-considered to the point of extravagance for either side to fail to coordinate on the better of these two outcomes.

Then why is the superpower dilemma commonly seen as a perverse rather than cooperative version of the Prisoner’s Dilemma? Because the difficult choices we face are not whether to use the weapons but whether to develop and deploy them, especially whether to develop and deploy new weapons to add to or replace extant weapons. This, too, is an inherently iterated Prisoner’s Dilemma, but it is one in which we are conspicuously not playing cooperative strategies. This is the perversity, because such a game should generally be played cooperatively. Let us briefly consider the perversity of the arms-race Prisoner’s Dilemma to see why it is played so differently from the nuclear-war Prisoner’s Dilemma.

The central problem in the arms race is that it is largely driven by technological change and by the hope and fear that such change might lead to genuine capacity to win a nuclear war by destroying enough of the other side’s weapons to prevent serious retaliation. Neither side can stop research and development while the other side continues them. But it may be that neither side can really be sure that the other is not continuing research and development even if it is not. If we are radically unsure of what the other side is to do tomorrow, we may as well play today as though we are in an unenforceable one-shot game.

Oddly, the arms race may even be exacerbated by the fact that both sides are generally convinced that the game of using the weapons is, for the time being, one of pure coordination. Actual deployments of new weapons are of little serious interest other than to the extent that actual production and deployment contribute to research and development, because deployment does not imply eventual use. Hence, one can decide on deployment without regard to anything more than negligible effects on actual use. More often than not, systems may even reduce the riskiness of accidental launch. Because new systems have technologically more sophisticated controls, typically carry smaller, cleaner warheads, and are more precise than older systems, they will thus be less subject to accidental launch in normal times and cause less threatening damage if they are accidentally launched.

These advantages may, of course, be offset by the increasing numbers of weapons or of weapons systems actually deployed over time. They may also be offset by the increased scale of damage that is possible in the event of an intentional, as opposed to accidental, attack, or of an inadvertent attack during a crisis, although the scale of possible damage has long since passed into the region of overkill and possibly even nuclear winter.

Note an implication of the changes that have taken place. If the increases in potential devastation from new systems have not been accompanied by decreases in the likelihood of all-out nuclear war, then the order of expected outcomes in the game of using the weapons may have changed to make either superpower rationally prefer unilateral disarmament to a continuation of mutual armaments. It may long have been true that the best joint payoff in the game of deciding on a level of armaments would be under a regime with a single superpower, if joint payoffs are counted in total number of expected lives lost or some such crude objective indicator. This is, of course, the claim of recent moral philosophers such as Douglas Lackey, who argue that, on a utilitarian account, unilateral disarmament is the rational, self-interested—not only the moral—policy. As Herbert Simon supposes, “A plausible inference (from the nuclear winter scenarios) is that such a weapon is worse than none, either for attack or deterrence, and that the way is now open for nuclear disarmament, even unilateral partial disarmament.” The only caveat on this claim is the major caveat that the policy most in one side’s interest is to convince the other side that they should be the ones to disarm, perhaps by convincing them that one’s own side is too crazy to follow its interests. If there really were only one nuclear power, it could use the weapon with some impunity, without fear of causing nuclear winter, since one or two bombs could humble any nation unable to retaliate.

Two conclusions follow for the arms race, depending on whether it is a Prisoner’s Dilemma or a Nuclear Pax. If it is the former, it should be cooperatively resolved if both sides can be relatively certain of the cooperative behavior of the other and it should be uncooperatively resolved if they cannot. By far the most important consideration in pushing us one way or the other, then, is the quality of surveillance of what the other side is doing. On many accounts, each side now can be quite confident of knowing whether the other side has made any dramatic advances in weapons systems that would permit it to risk a first-strike attack without fear of retaliation. If this is true, we should strive to play the arms race with a tit-for-tat cooperative strategy in which we unilaterally reduce arms in the expectation that the other side will follow suit. The only reason for one side to refuse to reduce arms would be its hope that it could alone achieve serious first-strike capacity for victory, which is surely a dim hope at best, even if the scenarios of the nuclear winter have been exaggerated.

If the arms race has now turned into a Nuclear Pax game—because, for example, the likelihood of nuclear winter is great—then our prospects may be far worse. We face a pure conflict game in the determination of which superpower is to be the hegemonic power. Neither is likely to step aside easily to let the other become dominant. But until one side does so, we face the risk of explosive war in a bad moment, a war that might obliterate us both and no doubt others along with us—all because we are in a game in which “win-
ning” has meaning, in which doing well is not the only possibility. We still cannot hope to win an actual nuclear war, but we might hope to win the contest of nerves for which power gains hegemony. Russell and many others have supposed we might instead resort to a world government. Theirs cannot be an optimistic hope.

The most important elements in determining which state of affairs we are in are the actual risks we run of failing to achieve deterrence with present weapons and the risks we run of facing a credible first-strike threat from future weapons. On the second issue, I doubt that victory is plausible in the foreseeable future, a doubt that is only reinforced if the prospect of nuclear winter is at all serious. The chief problem in judging the first issue, the likelihood of the failure of deterrence, is assessing risks, which, alas, is a heavily subjective enterprise. On this problem, again, game theory can contribute clarity.

RISKS OF TRYING TO DO WELL

Presumably, we may take as given that almost no military leader anywhere would agree with Simon that “such a weapon is worse than none.” Many might wish it did not exist but, given that it does, all would rather have it than not. On the admittedly distorted evidence of American national electoral politics, we can probably also take as given that the vast majority of citizens in the United States similarly would rather the United States have it than not. Hence, we “must do the best (we) can in making security sense of a weapon that cannot be disinvented.”

Total disarmament is bound to be quite unstable while there is no supranational enforcement agency because a nation that has had the capacity to build the weapons would, once disarmed, have the residual capacity to create several of them on very short notice, possibly before the other side could respond in kind even with fairly good surveillance. To anyone who claims that the weapons are worse than none, one may simply counter that to a lone nuclear power they might be extremely valuable. Hence, the incentive for secret rearmament in a regime of total disarmament would be so great as to make anyone doubt the intelligence of total disarmament that is secured merely by a voluntary agreement between nations.

It is sometimes supposed that under present conditions “the only purpose that a nuclear armory serves is to deter an opponent from using his.” This is surely a false supposition—the weapons must have some value in extended deterrence simply by virtue of raising the stakes for certain transgressions. But if the only purpose of the weapons were to deter the other side from using its weapons, then far fewer weapons could do what our present arsenals do, especially if both sides had fewer weapons. Indeed, given the diversity of current systems and the relative invulnerability of many of them, one side alone could safely disarm down to much lower levels, without fear of losing its capacity to retaliate with sufficient devastation as to deter the other side from attacking it.

If one side continued to develop and deploy new technologies while the other was satisfied with a modest deterrent force, the latter side might eventually find itself facing a serious prospect of first-strike destruction. Such a development, however, would take years and the passive power would know that the other was aggressively developing new systems. Hence, there is virtually no risk in the short term that a policy of unilateral arms reductions would harm one of the superpowers. But because the chief motive either side has to continue to develop arms is the fear that the other is doing so—therefore, we rightly speak of an arms race—a less aggressive policy on one side should encourage a less aggressive policy on the other.

Here we would do well to recall the nature of the payoffs in the superpower games of Table 9.1. The payoffs in a particular outcome are a function of the objective values of the various contingencies that may occur in the outcome. For example (3,3) in Game 1, the Prisoner's Dilemma, takes its value for one of the sides from the probable results that would flow from having both sides heavily armed. To assess what that value might be, we must value each contingency (such as, most obviously, successful deterrence with increasing arms and a failure of deterrence with grievous losses in a nuclear war) and weight it by its likelihood. To suppose that the superpower dilemma is a Prisoner's Dilemma is to suppose that the likelihood of a devastating nuclear war, under conditions of a continued arms race, is not high relative to the likelihood of gruesome suppression of a unilaterally disarmed former superpower.

Of course, the assessment of such likelihoods and their attendant risks is very problematic because it is highly subjective. A major part of the appeal of unilateral efforts to bring about mutual disarmament, down to substantially lower but still deterrent levels, however, is its effect on subjective evaluations. Russell supposes that two disarmed superpowers are as virtually certain to have a devastating nuclear war as are two superpowers locked in an arms race. But two superpowers with far more limited arms, who are not engaged in an arms race, have far less reason to fear they are being outraced by each other. If there is self-evidently no hope of a first-strike elimination of the other side's capacity to retaliate, reason is far more likely to prevail in a hot moment than if there is at least the fleeting hope of victory in a nuclear war, or fear of defeat if one does not strike first.

The way to insure that reason prevails in hot moments is to have it prevail in cooler moments, to reduce the levels of arms and to stem the floodtide of arms developments. Contingent choosing now keeps the arms race going at full tilt; it could as well undergird arms reductions and arms control. Gray asserts that "strategic logic and political judgment meet in a consensus," that,
among other things, "major-scale force reductions are not within the bounds of possibility."21 The United States," he says, "enjoys a consensus on assumptions and risks sufficient to design and conduct a nuclear strategy that has strategic merit and that meets the tests of domestic and allied political acceptability." He does not say just who qualifies for membership in that United States which enjoys a consensus. On any reasonably generous criterion for membership, his claim is false. That it is not worthy of instant dismissal is partly due to the fact that the supposed game theorists of the strategic community have been in default in their game theoretic accounts of the structure of the problems we face.

If the development and deployment of nuclear-weapons systems is an iterated Prisoner's Dilemma, then major-scale force reductions are within the bounds of possibility if we would actually attempt them. If it is such a Prisoner's Dilemma, then it bears little relation to poker, the game theorists' favorite real-life game. In the superpower dilemma we cannot hope to win—but we may hope to do reasonably well.

APPENDIX:
THE SYMMETRIC SUPERPOWER GAMES

There are 576 distinct matrices of $2 \times 2$ games with ordinal payoffs without ties (that is, with ordinally different payoffs in each outcome for either player). Many of these can be obtained by simply switching the labels of the players or of their strategies. Of the 576 games, 48 are symmetric. This set is reduced to the three games of Table 9.1 in the following way.

First, we are concerned with games in which the labels on strategies actually matter. Switching the D and A labels for the Soviet Union in Game 3 produces a game that is, abstractly speaking, strategically symmetric in actual strategies. Eliminating games in which only one side has its strategies relabelled reduces our original set of 48 games to 24. We know that there are only 12 strategically non-equivalent $2 \times 2$ games.22 The 24 games come from doubling this set to include the, abstractly speaking, strategically equivalent games in which both sides' D and A strategies are relabelled.

The set of 24 games is now further cut in half by the condition that both sides prefer mutual disarmament to mutual armament. Finally, these 12 are cut to 3 by the condition that each side would most prefer to be armed while the other is disarmed. Hence, we are left with the 3 games of Table 9.1.

NOTES

1. In many discussions of the nuclear dilemma it appears that the most preferred outcome for either side would be that in which neither is armed. I have discussed this possibility, and the three coordination games that fit it elsewhere. Russell Har-