South Carolina Journal of International Law and Business

Volume 16 | Issue 3 Article 6

2020

GAME THEORY FOR INTERNATIONAL ACCORDS

Uri Weiss
Van Leer Jerusalem Institute and The Hebrew University of Jerusalem

Joseph Agassi
Tel Aviv University and York University

GAME THEORY FOR INTERNATIONAL ACCORDS

Uri Weiss[*] *and Joseph Agassi*[+]

Realpolitik is the claim that agreements in international relations are worthless since there is no institution to enforce them. Game theoretician Robert J. Aumann suggests in his 2006 Nobel lecture that "the fundamental insight is that repetition is like an enforcement mechanism."[1] *The application of this insight to international relations allows for the improvement of their applicability and it, thus, refutes Realpolitik.*

Early game theory appeared as an alternative to the social sciences. However, it is better anchored within social science—as a useful tool. This renders game-theoretical recommendations irenic. Aumann argues that there is no a priori reason to expect that the agreement to

---

[*] Dr. Uri Weiss, Polonsky Fellow, Polonsky Academy at Van Leer Jerusalem Institute and The Hebrew University of Jerusalem; e-mail: uriw10@gmail.com.

[+] Prof. Emeritus Joseph Agassi, Tel Aviv University and York University, Toronto; e-mail: agass@tauex.tau.ac.il.

[1] Robert J. Aumann, "War and Peace," Prix Nobel 2005, the Nobel Foundation, Stockholm (2006).

cooperate should have practical results.[2] *His claim rests on an additional assumption: at times no improvement is observed. Yet, at times significant improvement is observed. This should encourage the search for the conditions that lead to improvement; it goes well with the proposal to consider game theory part-and-parcel of social science: how does playing in a given game depend on the culture within which it takes place.*

I. THE POLITICAL ASPECT OF GAME THEORY

A common topic of discussion within game theory is *the prisoner's dilemma* and its relevance to cooperation because its rules lead to conduct that reinforce conflict. When political scientists, jurists, or biologists apply game theory to the analysis of cooperation, they usually refer to this specific game. Of course, many other games pertain to cooperation. It is not easy to find out what game describes a situation sufficiently well in the field. It is easier to find out what game is advantageous to play under what circumstances. Such matters are better open to critical discussion and empirical tests.

Before presenting *the prisoner's dilemma*, let us present another, simpler game, *the movies dilemma*, a variant of *the prisoner's dilemma* often present in film. Here it is:

	Coop	Def
Coop	1, 1	3, 2
Def	2, 3	2, 2

	Cooperate	Defect

[2] Robert J. Aumann, Nash Equilibria Are Not Self-Enforcing, in ECONOMIC DECISION-MAKING: GAMES, ECONOMETRICS, AND OPTIMIZATION: CONTRIBUTIONS IN HONOUR OF JACQUES H. DRÈZEE. 201– 206 (J.J. Gabszewicz,, J.-F. Richard, and L. Wolsey, eds. 1990).

Cooperate	freedom, freedom	penalty, reduced penalty
Defect	reduced penalty, penalty	reduced penalty, reduced penalty

Figure 1. *The Movies Dilemma*

In this game, mutual cooperation is best for both players. To achieve mutual cooperation they have to trust each other sufficiently; if they mistrust one another, then they will come to mutual defection;[3] one defection leads to the worst outcome. Thus, if both players expect the other to either cooperate or defect, then their very expectations will make it true.[4]

In *the movies dilemma*, the information that one player has about the decision of the opponent plays a crucial role. Therefore, in variants of this game that allow the police to manipulate players through misinformation, it may lead one player to expect the other to defect. In that case, the expectation is self-fulfilling. Hence, manipulation is unnecessary. It suffices for the police to convince the players that the police will manage to convince one player that the opponent will expect the other player to expect the opponent to lose the trust of the one player.

In contrast, in *the prisoner's dilemma* game, it is worthwhile for each player to defect regardless of what the opponent does. This is the whole of the specification of that game. In the literature, it usually comes with a standard illustration that depicts a situation with four options: no penalty and penalties of three levels: lenient, severe, and medium—lenient penalty for the illegal possession of arms, severe penalty for having used them illegally, and the

---

[3] We use terms such as defection because they are common in the game's theoretical literature, but not because of their moral content. For example, a firm that does not join a cartel defects in the game's theoretical language, although it should not be denounced.

[4] This game appears in movies in diverse variants. For example, one prisoner may seemingly betray the other, but without losing the other's trust. This variant of the game may end with the trust rewarded, and it may result with the trusting party alone receiving full penalty, thus, leading to a new game of revenge. In all variants of the movies dilemma, the information that one player has about the decision of the other player plays a crucial role.

reduction of the severe penalty that leaves it still harsher than the lenient one.[5] Consider two persons detained for possession of illegal weapons near a bank in which an armed robbery just took place. The police have strong enough evidence to charge them with the lenient penalty, but not enough evidence to charge them with the severe penalty, so the police try to encourage them to testify against each other. To achieve this, the police isolate them and propose to each of them a plea-bargain. The options that the game offers are these: if they both defect, they will both receive medium penalty; if they cooperate with each other and keep silent then they will both receive the lenient penalty. There are four levels of possible results, from 1 to 4:

	Coop	Def
Coop	2, 2	4, 1
Def	1, 4	3, 3

Figure 2. *The Prisoner's Dilemma*

Thus, the wish to maximize individual payoff imposes on each player in *the prisoner's dilemma* game the betrayal of the other regardless of the strategy of the other.[6] A strategy like the one

[5] Stuart Oskamp & Daniel Perlman, Factors Affecting Co- operation in the Prisoner's Dilemma, 9 J. OF CONFLICT RESOL. 29–31 (1965).

[6] This idea of strategy is as old as game theory. According to the definition of Von Neumann and Morgenstern set forth in 1944, a strategy is a player's plan, which specifies what choices to make in every possible situation, for all possible information available at the moment decision is called for. The strategy conforms to the pattern of information that the rules of the game prescribe. Thus, a strategy is a comprehensive policy, a plan for action in every possible situation that the rules of the game allow. Obviously, then, the project of Von Neumann and Morgenstern is utopian. As Kenneth Arrow has noted, such a strategy is impossible even for chess—a problem-situation much simpler than some real-life ones. Von Neumann and Morgenstern postulated that comprehensive strategies are always parts of games. This limits the applicability of game theory to the very simplest

described here is not the only one available. It is dominant in the sense that in all permissible situations a player will gain from it more than from any alternative strategy; therefore, in this game a player cannot gain anything from the information about the opponent.[7] Hence, in *the prisoner's dilemma* game, rationality precludes the socially optimal result: it leads to the socially worst result. This is why it is intriguing; the unpleasant aspect of the situation in *the prisoner's dilemma* is that the distrust inherent in it is irreparable, since it imposes a result not improvable by soliciting trust.[8]

In some similar games, raising the level of trust might improve matters. The most common illustration for this is the variant of *the prisoner's dilemma* known as *the stag-hunt game*[9] (what makes game theory interesting is that it offers many variants of this game with different results; a little change in the game may, at times, lead to a completely different result). In it, cooperation brings the best payoff for each of them; the unilateral betrayal of one meets the defector the second-best payoff and the other the worst payoff, and mutual betrayal gives both the third-best payoff. For this, again, four possible outcomes are required. This is illustrated by two hunters who choose simultaneously whether to hunt a stag or rabbits. They succeed only if they both go for a stag, and each player achieves the best result—the stag. A player who goes for a stag alone is met with absolute failure. A player who goes for a rabbit alone wins all the rabbits, which is the second-best result, while both going for the

---

games, thus, limiting severely the intellectual challenge of game theory. Sometimes it is surpassingly possible to write some strategies. The standard examples are the always defect in the repeated prisoner's dilemma, as discussed above, and the tit-for-tat in the same game. *See* JOHN VON NEUMANN & OSKAR MORGENSTERN, THEORY OF GAMES AND ECONOMIC BEHAVIOR (1944) (ebook), https://ebookcentral.proquest.com/lib/lawsc/reader.action?docID=1092486&ppg=2.

[7] This is the equivalent to Savage's "sure thing principle." In the early stages of game theory, it was called "the sure thing strategy."

[8] This may explain the futility in some situations of the good will of peace activists who do not try to act politically, specifically in a way that changes the game.

[9] Richard H. McAdams, Beyond the Prisoner's Dilemma: Coordination, Game Theory and the Law, 82 S. CALIF. L. REV. 13–15 (2008).

rabbits mutually gives every player half of the rabbits—the third-best result. Therefore, it is best for both to go for the stag. For the one who goes for the rabbits, it is better if the other does not, namely, that the opponent goes for the stag (and loses), thus, enabling the one to hunt rabbits unimpeded. Consider then four levels of success, from 1 to 4:

	Coop	Def
Coop	1, 1	4, 2
Def	2, 4	3, 3

Figure 3. *The Stag-Hunt*

Obviously, the absence of trust prevents the achievement of the optimal solution in this game, while if the players trust each other enough, they will achieve it. The important difference between the two games is not in the stories, but in the matrices for it is possible to translate the story of the stag-hunt game to the terms of the prisoner's dilemma game, and it will remain the stag-hunt game. For example, if the two suspects from the prisoner's dilemma game keep silent, they will both walk; if they both sing, they will both receive the usual penalty; and if only one sings, then only that one will receive a lenient penalty, and the other will receive a heavy penalty. Hence, the matrix determines the game, not its illustration.

The most important difference between the unrepeatable prisoner's dilemma and the unrepeatable stag-hunt is that in the former game defection is the dominant strategy—each rational player will defect in any case—whereas in the latter the defection (or its avoidance) depends on the assessment of the interdependent strategies of players. Whereas the one game offers no hope for cooperation, the other offers recognition of the option of raising the incentive for cooperation by raising trust. Hence, it is more important to avoid situations that impose the prisoner's dilemma game rather than the stag-hunt game.[10] Although both games describe conflict situations, the lesson for social science is that in

[10] *See* Joseph Agassi & Abraham Meidan, Philosophy from a Skeptical Perspective 96 (Cambridge Univ. Press, 2008).

some conflict situations action can improve actors' situations all around even without eliminating the conflict.

The difference between variants of a game may, thus, be significant. The decision about which variant describes a given political situation already determines attitudes towards it. Thus, bellicose game theoreticians set the game one way, and the irenic ones set it the other way. This is Mario Bunge's criticism of game theory: it encourages arbitrariness.[11] The description of a real-life situation as a game will, thus, be less arbitrary if it includes options—whenever these are possible—for players to choose what game to play, with whom, and with what payoffs. This decision as to what game to play—this super game—describes some situations better than the games prescribed in standard game-theoretical texts.[12] This requires the recognition that at times some players are able to choose what game to play next.

This is also the choice available to scholars who wish to use game theory in order to analyze given situations: they may (and possibly should) ask what games are available to players and what game is better for a player to play. This will prescribe for scholars the decision as to the choice of game to analyze—the most important in the field. They may then help players or social planners improve their lots by offering good advice. For example, in the sphere of litigation, it is more important for students of jurisprudence to analyze the *asymmetric* litigation game than the *symmetric* one, even if the symmetric games are more frequent.[13] Only the asymmetric

---

[11] *See* MARIO BUNGE, SOCIAL SCIENCE UNDER DEBATE: A PHILOSOPHICAL PERSPECTIVE 180 (Univ. of Toronto Press, 1998) (ebook), https://ebookcentral.proquest.com/lib/southcarolina/detail.action?docID=4671968)

[12] *See id*. at 176–80.

[13] For example, Weiss analyzed the appeal game as an asymmetric one, while Shavell analyzed the appeal game as a symmetric one. *See* Uri Weiss, *The Regressive Effect of Appealability*, SSRN ELECTRONIC JOURNAL 1 (2011), http://dx.doi.org/10.2139/ssrn.1688877. Shavell noted that his model is not valid in a case of heterogeneous litigants, and, nevertheless, derives general policy recommendations from this model. The difference is

game may hide unacceptable consequences for the weaker party. The legal system may allow for situations in which weak litigants cannot realize their rights or at least it is not worthy for them. Legislators, judges, and attorneys for the weak litigants should try to prevent these situations as the initial (super) game. This is a worthy moral for the "law and economics" movement that aims to assess which legal rules are economically efficient. The analysis of the symmetric game—where options are the same for each side—is elegant, easy, natural, and relatively easy to apply, but it is not the most important game in town. Legal theory will benefit more from research that will reduce the number of unavoidable injustices of the system, and these are the asymmetric cases where financially comfortable litigants have many more options, including those who are less risk-averse due to their richness than ones who happen to be financially constrained.[14] This may lead the weak parties to forego the use of all the legal advantages that they have and settle for much less than what the law entitles them. This is also the case when one party is a one-time player, and the other party is a repeat player[15] (ironically, the literature considers this case not a part of "law and economics" but a part of "law and society"; obviously, it is both). Any move intended to compensate the less well-off litigant is a revision that will lead jurists to prevent games that end up in patent injustice. This is not limited to any specific society; the Bible mentions asymmetric litigation: "seek judgment, relieve the oppressed, judge the fatherless, plead for the widow."[16]

Admittedly, asymmetric games are usually mathematically less elegant, but they are socially more important, at least from the humanist perspective. Unlike the prisoner's dilemma, many situations of war and peace comprise asymmetric games. In many

---

not only that one analyzes this game and the other analyzes another game: the important question is what game should be analyzed in the theory of litigation. What game should we see when we recommend rules of litigation? *See* Steven Shavell, *The Appeals Process and Adjudicator Incentives*, 35 THE J. OF LEGAL STUDIES. 1, (2006).

[14] *See* Uri Weiss, *The Regressive Effect of Legal Uncertainty*, TEL AVIV UNIV. L. FAC. PAPERS, 2005, at 1.

[15] *See* Marc Galanter, *Why the "Haves" Come out Ahead: Speculations on the Limits of Legal Change*, 9 L. & SOC'Y REV. 95 (1974).

[16] *Isaiah* 1:17.

cases of violence, the strong party sees the game as asymmetric but presents it as symmetric in efforts to fend off the police, the courts, or public opinion. Even kids who are bullies in school do that when facing school authorities. Under attack, then, it is often useful to change the game by making a credible threat to involve the police, the courts, and public opinion. Making a conflict visible may even render an asymmetric game symmetric and thereby reduce violence dramatically.

II. A COMPARISON BETWEEN OUR APPROACH AND THE RELATED LITERATURE

Aumann (2009) claims:

> incentive . . . has to be there, and that is what is represented by the prisoner's dilemma in very stark, obvious language . . . . Absolutely, you must create incentives for stopping CO_2. There is one very simple way to do it. Just tax the emissions. You could impose a much higher tax on gasoline. And there are other ways to tax emissions. Do not overtax them, but tax them at the true cost of these emissions. Absolutely, you have to give incentives. Not by fear: that is not going to work. What is going to work is giving people incentives. Precisely game engineering.[17]

We assume our readers are familiar with this, especially since incentives can appear in different places and grow at different paces depending on extant social and political conditions. Incentives can be chosen as part of the game, such as in the case that a player chooses a conditional strategy in the prisoner's dilemma, and they can be chosen in order to prevent a particular kind of games. Let us sharpen that in the example Aumann described, the social planner actually supplies an incentive in order to prevent an undesirable game, so it is actually a super-game.

Similarly, Aumann and Shapley show the need for social science in order to explain the stability of the repeated prisoner's

---

[17] Robert J. Aumann, *Game Engineering*, *in* DISCUSSION PAPER: CENTER FOR THE STUDY OF RATIONALITY no. 518 (2009).

dilemma—as due to the cooperation between players imposed by the rule that requires penalty for those who do not punish:

> it . . . should be noted . . . that not only are defections from the cooperative sequence punished, but also defections from any punishing sequence are punished. A player who 'should' punish and does not do so will himself be punished. This is what provides the motivation for the punisher actually to carry out the penalty, and so keeps [the equilibrium].[18]

To this we add its converse: the same rules can destabilize the prisoner's dilemma itself and even eliminate it almost totally.

As to the context of any game, Aumann and Drèzee (2008) observe this:

> Formally, a game is defined by its strategy sets and payoff functions. But in real life, many other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game. For example, in a parliamentary democracy with three parties, the winning coalitions are the same whether the parties each hold a third of the seats in parliament, or, say, 49 percent, 39 percent, and 12 percent, respectively. But the political situations are quite different. The difference lies in the attitudes of the players; in their expectations about each other; in custom; and in history, though the rules of the game do not distinguish between the two situations.[19]

Let us comment on this: in Aumann's example (or even in a more extremist case of seats divided to 49%, 49%, 2%), traditional game theory may deem the three political parties in possession of

---

[18] Robert J. Aumann & Lloyd S. Shapley, *Long-Term Competition—A Game-Theoretic Analysis*, *in* ESSAYS IN GAME THEORY, 1– 15 (Megiddo N. ed., 1994).

[19] Robert J. Aumann & Jacques H. Drèzee, *Rational Expectations in Games*, 98 AM. ECON REV. 1, 72–86 (2008).

equal power, since no party can establish a winning coalition by itself, and each party can establish a winning coalition with every other party.[20] Clearly, this is a mistake. Nevertheless, game theory is right in considering the small party in this case as having much more power than the number of its seats suggest and in its explanation of this fact; but, game theory ignores the constraint on the power of the small party that social norms of fairness impose. Game theory also ignores the incentives that this situation provides to change the situation radically. Thus, members of the big parties may defect and establish small parties or the majority may change the voting system. This invites interesting questions. How does the prevalent view of fairness influence the situation? What is the right view of fairness? How should it influence the situation? These questions and their likes pull us out of the mathematical world of game theory and lead us to apply social science. This illustrates the fruitfulness of traditional game theory as well as its limitation. Hence, to be fruitful, game theory should become part and parcel of social science. Otherwise, game theory may generate more mistakes than it can prevent.

In the conclusion of their paper, Aumann and Drèzee add this: "The fundamental object of study in game theory should be the game situation G rather than its underlying game G," while in the paper itself they define game situation as "a game played in a specific context."[21]

As young as game theory is, it already has a tradition. That tradition rests on its initial aim that was tacit. It was, we say, to replace the explanatory model of the social sciences (indeed, one of the early names of game theory was "social physics"). Von Neumann and Morgenstern said of its applications that they are of two kinds: "On the one hand to games in the proper sense, on the other hand to economic and sociological problems as well . . . . We hope to establish satisfactorily . . . that the typical problems of economic behavior become strictly identical with the mathematical notions of suitable games of strategy . . . ." For economic and social problems, the games fulfill—or should fulfill—the same function,

---

20 *See id*.
21 *Id.* at 72, 82.

which various geometrical-mathematical models have successfully performed in the physical sciences.[22]

The 2008 paper of Aumann and Drèzee just cited is possibly a challenge to the tradition of Von Neumann and Morgenstern, a step-in effort to revise it. In line with this we try to anchor the theory—as is or in a revised version—within traditional social science. To that end, we draw attention to the difference between Aumann[23] (1990) and Aumann and Drèzee (2008). Aumann (1990) claimed—quite rightly—that agreement to play the stag-hunt game in mutual cooperation is not self-enforcing.[24] He added that the agreement to cooperate while playing the stag-hunt game does not bring about any improvement.[25] This we deem somewhat incorrect since it is an oversight of the agreement that may change the mutual expectations of players that the result of the game depends on. Aumann's argument is this: both players will gladly agree to cooperate, whether or not they later keep their word while playing; hence, their explicit agreement conveys no information: "To say that a game is non-cooperative means that there is no external mechanism available for the enforcement of agreements . . . . Incentives can be changed by changing either the payoffs or the information of the players."[26]

Of course, one may see the custom of keeping promises as irrational in any one-time game. This is a mistake. Expectations regarding cooperation that rest on agreement are too common to dismiss. Also, it will be beneficial for any specific society as well as for the international community to reform the culture in a manner that generates expectations to cooperate. That reform would render the reliance on promises eminently rational. As such, agreements tend to raise expectations; they improve the likelihood of achieving cooperation even in the repeated prisoner's dilemma game. This has a significant effect also for international relations, where institution to enforce contracts are still rather ineffective. This is in agreement with Aumann: "In the international relations literature, the game has

---

22 VON NEUMANN & MORGENSTERN, *supra* note 6, at 2.
23 Aumann, *supra* note 2.
24 *Id.*
25 *Id.*
26 *Id.*

been called the 'security dilemma.'"[27] Contrary to Aumann 1990, however, we argue that international agreements in stag-hunt situations improve the disposition to cooperate and that, therefore, game theory rejects *Realpolitik* in international relations (*Realpolitik*, to repeat, is the recommendation to consider all agreements altogether worthless[28]).

Aumann is quite right in asserting that there is no *a priori* reason to expect agreement to cooperate to lead to cooperation.[29] The very need to come to agreement may already signal potential mistrust and, thus, mistrust and doubt as to the expectation that promises lead to cooperation. Thus, Aumann's assertion that there is no *a priori* reason to expect agreement to lead to cooperation requires completion; at times, but only at times, there is *a posteriori* reason for that.[30] This then is an argument for the proposal to consider game theory, part and parcel, of social science. How a given player will behave in a given game, thus, depends on the culture within which the game takes place. Hence, the conclusion from the rules of the game to the conduct of its players depends on tacit suppositions that represent the social conditions under which they play the game. These are better specified explicitly. The rules of the game called game theory should be altered to include this demand. This will lead to the proliferation of variants of many games that have, thus far, already been considered exhaustively.

For example, in the traditional wording of the stag-hunt game, the description of the set of alternatives is too sketchy: the option of agreement is missing without notice. Therefore, when one mentions it, one implicitly indicates that the game is not a closed system; it is, then, not mathematics; at best, it is social science. Considered pure mathematics, it does not have a unique solution: the conclusion that agreement will lead to improvement is questionable and depends on the expectation the agreement creates. In this regard, we agree with Aumann.

---

27 *Id.*
28 Aumann, *supra* note 2.
29 Aumann, *supra* note 2, at 619–20.
30 *See id.*

We can conclude from the above discussion that it is better to play the variant of stag-hunt with the option of preliminary communication than the stag-hunt without the option, and that these are indeed two different games. In the stag-hunt game with an option of preliminary communication, words are not merely cheap talk, but, they are in the one-time prisoner's dilemma game with an option of preliminary communication. To be precise, we should not ignore the variant of prisoner's dilemma played publicly with unenforceable agreements to cooperate: in this variant of the game, players will respect their agreement to cooperate in cultures in which the refusal to honor one's commitment will damage one's reputation considerably.

In Aumann's Nobel lecture we read, "the fundamental insight is that repetition is like an enforcement mechanism."[31] This insight of Aumann is a clear refutation of *Realpolitik* that assumes that since there is no institution to enforce agreements in international relations, those agreements are worthless. Aumann's insight explains why covenants without sword waving can serve as much more than mere words: they add significant strength to much needed security. What we said contradicts Watkins assertion,[32] which states that game theory endorses the claim of Hobbes: "covenants, without the sword, are but words, and of no strength to secure a man at all." We argue that game theory leads to the contrary conclusion: that covenants may prevent war even without sword waving, more in line with the observation of Hobbes.[33]

---

[31] Aumann, *supra* note 1, at 354.

[32] John Watkins, *Imperfect Rationality*, *in* EXPLANATION IN THE BEHAVIORAL SCIENCES 202–03 (Robert Borger and Frank Cioffi ed., Cambridge Univ. Press 1970).

[33] THOMAS HOBBES, THE ELEMENTS OF LAW NATURAL AND POLITIC 78 (Ferdinand Tonnies ed., Frank Cass and Company Limited 1969): "In contracts that consist of such mutual trust, as that nothing be by either party performed for the present, when the contract is between such as are not compellable, he that performeth first, considering the disposition of men to take advantage of everything for their benefit, doth but betray himself thereby to the covetousness, or other passion of him with whom he

Game theory conflicts with the *Realpolitik* idea that international agreements are not worth the paper on which they are written.[34] Game theory similarly conflicts with the *Realpolitik* idea that the rule of law does not matter since it can do no more than reflect and legitimize extant balances of forces active between the nations with no ability to change them.[35] This is the social philosophy of Hegel that is popular today among the legal realist movement.[36] Fortunately, this view meets with a very simple refutation: a new enforceable law can prevent, or at least reduce, situations of prisoner's dilemma, which is agreeable to all parties involved, so such a law has a great likelihood of changing an undesirable Nash equilibrium in many games.[37]

III. ADDITIONAL MORALS TO LEARN FROM AUMANN'S NOBEL LECTURE (2005)

We offer two morals from Aumann's Nobel lecture. The first corresponds with his conclusion of his analysis of a particular repeated game:

---

contracteth. And therefore such covenants are of none effect. For there is no reason why the one should perform first, if the other be likely not to perform afterward. And whether he be likely or not, he that doubteth, shall be judge himself . . . . But when there shall be such power coercive over both the parties, as shall deprive them of their private judgments in this point; then may such covenants be effectual; seeing he that performeth first shall have no reasonable cause to doubt of the performance of the other that may be compelled thereunto."

[34] *See* Erik Ringmar, *The Relevance of International Law: a Hegelian Interpretation of a Peculiar Seventeenth-Century Preoccupation*, 21 REV. OF INT'L STUDIES 87 n.1 (1995).

[35] *Id.* at 91.

[36] *See id.* at 101–02.

[37] A game is in a Nash equilibrium only if no player has incentive to change strategy unilaterally. *See* John Nash, *Non-Cooperative Games*., 54 ANNALS OF MATHEMATICS S2 n.2, 286 (1951).

> "What is maintaining the equilibrium in these games is the threat of punishment. If you like, call it 'MAD'—mutually assured destruction, the motto of the cold war."[38]

In the game, he analyzed it is indeed feasible to implement the advice to punish the party that plays the repeated prisoner's dilemma with a hostile strategy. It may nevertheless be infeasible in international relations, for example, in cases where punishment leads to a response from a third player such as an umpire (it will not be a repeated prisoner's dilemma, but a mere variant of it). In an effort to achieve a result of cooperation in the game, a player may be ready to punish the opponent severely. Other parties may then block the whole game, even in cases in which mutual cooperation is achievable with relative ease. Even the option of lenient penalty may be politically and scarcely feasible then. Therefore, an umpire may prevent the game and sometimes lead one player to always cooperate and the other to always defect. Let us propose these two games that may be enforced by the umpire: a repeated unilateral stag-hunt and a repeated unilateral prisoner's dilemma. Therefore, it is a super game; the umpire may force the states to play one of these games; in other situations, the teacher may force the pupils to play one of these games. In these games, one player can choose between cooperation and defection, and the second player has only an option of cooperation. The payoffs of the possible results of these games are such that the payoffs of these results in the prisoners' dilemma or stag-hunt. Actually, the umpire deletes one of the lines in the matrix of the game and by this makes it a new matrix—a new game.

This will be the matrix of the unilateral prisoner's dilemma:

	Coop	Def
Coop	1, 2	2, 1

Figure 4. *The Unilateral Prisoner's Dilemma*

This will be the matrix of the unilateral stag-hunt:

	Coop	Def

[38] Aumann, *supra* note 1, at 354.

Coop	1, 1	2, 2

Figure 5. *The Unilateral Stag-Hunt*

One of the advantages of the variant of a repeated stag-hunt with an umpire whose task is to force one player to avoid punishing the other player, to enforce the repeated unilateral stag-hunt, on such a variant of a repeated prisoner's dilemma, is this: in the repeated unilateral stag-hunt, the players will reach mutual cooperation, while in the repeated unilateral prisoner's dilemma they will reach the result in which one player will always defect and the other will always cooperate. While mutual cooperation is a possible result (as well as mutual defection) in the repeated prisoner's dilemma, the only possible result in the repeated unilateral prisoner's dilemma is that one player will always defect (this is their dominant strategy) and the other player will always cooperate. Hence, international intervention will be more desirable if it prevents the repeated prisoners' dilemma than if it prevents one side unilaterally from defecting in the repeated prisoner's dilemma. Furthermore, if a state believes that the umpire prevents them from defecting in a repeated prisoner's dilemma, the state should prevent this game when possible.

Notice that although in a prisoner's dilemma game the response to always-defect by always-defecting is reasonable and is possibly the best winning strategy, it still poses a possible penalty. Similarly, raising the reward for mutual cooperation or for being betrayed unilaterally may make tit-for-tat the reasonable strategy even in the prisoner's dilemma. This is so since the risk of the tit-for-tat strategy that incurs is reasonable: a player who adopts it takes a risk of losing in the first round, but he gains the opportunity to achieve the payoff of mutual cooperation, an opportunity that is not achievable by the always-defect strategy. The rational choice between these two options then depends, not only on the expectation that the opponent will play tit-for-tat, but also on the time discount and on the distance between the different payoffs (this fully accords with the complaint of Bunge 1998 that game theoreticians do not consider sufficiently critically the numbers that they write as examples for payoffs).[39]

---

[39] BUNGE, *supra* note 11, at 178.

The second moral from Aumann's theory is sober. The mutual cooperation in the repeated prisoner's dilemma depends on mutual threat, sometimes a threat to use force or to punish. Therefore, a change of the rules of the game that stabilizes mutual cooperation is beneficial even when its players choose mutual cooperation as the status quo. This is a challenge to the observation of Von Neumann and Morgenstern that any "game is simply the totality of the rules which describes it": they obviously overlooked the possibility of changing the rules of a game.[40] Constitutions often include some formal rules for change, and every constitution is open to a revolution.[41] This is so since even if the players achieve a Nash equilibrium of cooperation; the equilibrium may not be stable for some changes. Furthermore, there are equilibria of cooperation that rest on mutual threats, and there are those that rest on mutual trust; from a social science point of view, the latter is more stable and, thus, more desirable.

A physical system is in an equilibrium when the net force on each body in it is zero. It is stable if a small temporary deviation from it does not destroy it. It is unstable if it does (the equilibrium is indifferent if this deviation leads to another equilibrium). Moreover, equilibrium is relative to the forces in question: a system can be stable regarding only one set of extant forces. A game is in a Nash equilibrium if, and only if, no player has incentive to change strategy unilaterally. However, not all Nash equilibria are stable. Consider not only strategy change but also changes in the rules. Some equilibria remain stable even after such a change, but not after a change in the mutual expectations. Thus, stability is a relative matter.

One great advantage of the repeated stag-hunt game over the repeated prisoner's dilemma game is that only in the repeated stag-hunt game does each player always mutually cooperate, resulting in a Nash equilibrium. Thus, pacifist players will gain most from preferring to play stag-hunt over playing prisoner's dilemma: a

---

[40] VON NEUMANN & MORGENSTERN, *supra* note 6, at 49.

[41] *See generally* Herbert John Spiro, *Constitution*, ENCYCLOPEDIA BRITANNICA (Feb. 5, 2020), https://www.britannica.com/topic/constitution-politics-and-law.

player committed to play "always cooperate" will achieve the best result in all interactions, even where the received norm is mutual defection. And then, remarkably, all parties to the game are better off when they move from an equilibrium of mutual cooperation in the repeated prisoner's dilemma game to an equilibrium of mutual cooperation in a repeated stag-hunt game.

One may question this observation by noticing that those two equilibria allot the same payoffs to both players: this suggests there is no advantage in the shift from the one game to the other. The preference of more stable equilibria over less stable ones will lead to the rejection of this suggestion. This generally holds true as long as the more stable equilibrium does not impose stagnation; otherwise, the objection to stagnation may change the preference. Game theory is understandably an idealization, and, thus, it is not sufficiently sensitive to account for the difference in degrees of stability of the repeated game; this is no reason to overlook this difference, however. It is generally a political mistake to overlook degrees of stability, and it seems game theory can hardly help here without first inviting some development or change. As it happens, this oversight is common. Politicians systematically propose to end a war by reinstating the *status quo* in hopes of avoiding the repetition of past failed efforts at stability. At times, this hope for better stability rests on better considerations of the balance of powers between warring states. Game theory in its current state is unable to critically examine such considerations, as it is not sensitive enough to compare degrees of stability. It even overlooks the price for the achievement and maintenance of mutual cooperation in games of the prisoner's dilemma. Parties to this sort of game may make aggressive threats, which are costly even when there is no intention to follow them up. And then, players have to weigh the cost of war against the cost of the equilibrium within which peace depends on the fragile tool of threats to fight back (this resembles the equilibrium of peace in the repeated prisoner's dilemma). Since the consideration of waging war is expensive, it is wiser, whenever possible, to change the situation to enable players to rely on trust, which is the transition from the prisoner's dilemma game to the stag-hunt game. This happened in Europe after World War II, it seems. For now, peace is recognized as the best option for every European Union country, even where an attack on a neighbor would lead to an immediate surrender. This situation is obviously the best goal for all international relations, as it achieves the most stable situation. In this

situation, the peace will be stable, even if the two sides assess the outcome of a potential war as advantageous for themselves (even to the degree they are both convinced that they will definitively win the war). This is quite intriguing because in most other sorts of games this optimistic assessment of the results of wars usually leads to wars. Game theory, to repeat, does not succeed in accounting for the stability of the kind of game that leads to a Nash equilibrium of peace since no degree of stability is intentionally built as yet into the system of game theory.

We therefore recommend rendering game theory more sensitive to degrees of stability of its equilibria. This includes stability of the equilibrium when the rules are unstable or when players are misinformed, commit common mistakes, or change their preferences midgame.[42] It also includes stability of the equilibrium when new players enter the game or when the available set of alternatives for current players change. The development should be more fruitful as a toolbox to achieve stable world peace.[43] The ideal

---

[42] Howard *et al*. discuss a meta-game for which prospective players may choose their emotions, preferences, and even rationality. Players' self-interest will influence these and make some of their threats and promises credible; they will then rationally promote their chosen preferences. Howard *et al*. say, "often (as a player) one would be better placed strategically if one's preferences (P) were replaced by other preferences (P'). With preferences P', one would be in a stronger position to pursue one's original preferences P. Fundamentally, this happens because players can make use of each other's preferences as a means to obtain their ends." Now, if in the former game the players can adopt such moves, they do not improve their situation in the central game but prevent the central game; they make it another game. This is so since objective rules and options do not suffice to determine the game, as payoff for players signify too. Oddly, Howard *et al*. dismiss this rather obvious consideration. Nigel Howard ET AL., *Manifesto for a Theory of Drama and Irrational Choice*, 44 J. OPERATIONAL RES. SOC'Y, Jan. 1993, at 99, 100. *See also*, Nigel Howard, *The Present and Future of Metagame Analysis*, EUR. J. OPERATIONAL RES. 32.1, 1987, at 1.

[43] There are different sorts of equilibria in game theory, and they may be perceived to present different degrees of stability, but not in a way that will be fruitful for handling the problems we present here. We invite our readers to challenge us and correct our mistakes; we will be grateful for this.

of international relations would naturally be a solution of this kind, whenever possible. The hope, that we suggest will make it more viable, is that this situation is achievable by building widespread expectations for the application in international courts of strict laws against aggressive national leaders. At the very least, we should welcome efforts to minimize all incentives for political leaders to break international law or to ignore its summons or rulings. All this is easier said than done, of course. Our point, however, is that it is common sense and obvious from the viewpoint of game theory, as it should be. The generally received observation is that no one wants game theory to make recommendations that conflict with the public interest. And it is almost a consensus that the public interest is to make peace a top priority in all cases except for intolerable situations like enslavement or destabilization that worsens the situation (as symbolized by the compromise that Britain accepted in Munich in 1938).[44] If there is a situation in which war is better than peace, this should be subject to critical discussion, together with all possible answers to the question, what compromise is tolerable. Can game theory in its current version help the search for a reasonable answer to such a discussion? The answer, it seems on its face, is present-day game theory is useless for that purpose. We have argued that this is not true: present-day game theory may help rethink how to mitigate situations that threaten peace, admittedly, when degrees of stability signify greatly this is the case. And then, we say, it need not be so since game theory can nevertheless help one rethink the extent of the desirability of raising the degree of stability of peace and, thus, the cost that it is worthwhile to meet that end. And, observe,

---

[44] *See*, *Munich Agreement*, Fr.-Ger.-Gr. Brit.-It., Sept. 29, 1938, The Avalon Project, Lillian Goldman Law Library, Yale Law School.

Abraham Wald has shown the way.[45] Considering chains of games and sub-games will be a useful extension of current game theory.[46]

We suggest then that the most significant achievement of game theory is not in the design or in the applications of games, but in the suggestions of what games are unwise to play. Here, we follow Popper (Popper 1945), who said, politically, preventing pain or suffering has priority over creating pleasure. Obviously, in game theory, prevention is also much easier than application because every game requires some conditions for its very applicability, and these are never too clear and are seldom part of game theory proper. The games we consider unwise to play are obviously dangerous, as they may lead to war. The paradigm case here is chicken/brinkmanship. To our regret, game theoreticians are often more concerned with the best way to play them. Even if they are right, we prefer not to join them, but to recommend the proposal to avoid playing them when possible. At times, the game theoreticians in question stress that peaceful games fit some utopian situations so that in the meantime war is inevitable. We say, even if some war is inevitable, we should do our best to try to prevent every specific case of impending war, giving the good Lord the benefit of the doubt.

IV. CONCLUSION

Here then is our major corollary to game theory: the tools for achieving cooperation are incentives that generate strong and significant expectations: in brief, hope. The incentives may be

---

[45] *See* Abraham Wald, Statistical Decision Functions Which Minimize the Maximum Risk, Annals of Mathematics, 46 ANNALS MATHEMATICS, Second Series, Apr. 1945, at 265. *See also* Abraham Wald, Statistical Decision Functions (1950). *See also* Jacob Wolfowitz, J 1952. Abraham Wald, 1902–1950, The ANNALS OF MATHEMATICAL STATISTICS 23, 1–13 (1952).

[46] *See* Eilon Solan & Nicolas Vieille, *Quitting Games*, 26 MATHEMATICS OPERATIONS RES., May 2001, at 265. Solan and Vieille discuss the system in which players have the choice between quitting and continuing to play. They impose on games limitations that increase their mathematical elegance. Alas, these limitations lose the empirical character of games that gamblers play and of games that are important in social studies.

supplied by the legal system and by the norms and customs of civil society. The expectations may be products of institutions created to raise trust and join the educational or the diplomatic system. Surprisingly, a little success in trust-building may have a huge, dramatic, and positive impact on situations like the repeated prisoner's dilemma. Is this moral from game theory true? This is an empirical question not discussed here. That it deserves such discussions is obvious from the huge success of every educator who tried to reach neglected youths. Still, it is important to notice that the theory suggests that trust is superior to defection as the default option, thus, opening a venue to its empirical tests.

We recommend adding hope to the incentives and expectations of standard economic theory. Of course, appropriate incentives may generate hope, but they may also generate despair—intentionally or not. People can expect the best (that sounds hopeful), and they can expect the worst. Yet the logic of the ascription of expectations to rational agents differs from that of hope, since, unlike expectation theory, the theory of hope requires the will to live as more basic than any expectation, rational or not. As it happens, game theory evolved during the Cold War under the strong influence of economic theory and expectations theory. The theory of hope awaits proper development. We suggest that this step will also promote peace.