

WHEN STRATEGIC MISTAKES SAVE THE PLAY APPLICATION OF GAME THEORY IN SARTRE'S *DIRTY HANDS*

Stéphane Pillet

(translated by Balthazar De Ley)

As we know, Sartre's theatre is a theatre of situations, a theatre of liberty where each character "qui fait un acte est persuadé qu'il a raison de le faire" ["who takes action is convinced he is right to do so"]. Each situation represents an appeal to the characters' freedom of action. From another point of view, each of their decisions must take into consideration the presence of the other, thus creating a tight relational structure. It is in this way, precisely, that game theory players must reflect on their choices and their consequences, as well as on the choices of the other players. From this point of view, each choice is individual but at the same time depends on the situation given. Indeed, Nicolas Eber defines game theory as a prediction and decision-making tool to analyze interactions among individuals with divergent interests (120). Roger Myerson, one of the three 2007 Nobel Prize winners for their works on mechanism design, a branch of game theory, notes that his interest in the field lies in the fact that it promises a deeper understanding of the logic of conflict and cooperation that could possibly help build better social structures. In part, as Gaël Giraud explains, game theory studies the reasons for which different players are involved in or resist agreements or decision, and predict their actions and strategies (19). Hebert De Ley comments that the winning player seems to be the master of his destiny; he knows before acting what the good strategy is and what is the value of his game. "La théorie des jeux," he adds, "insiste sur le choix rationnel du héros" (138) ["Game theory insists on the rational choices of the hero"]. This is precisely the approach that Sartre develops in his theatre. Therefore, my aim in this study is to use game theory to determine whether, in *Dirty Hands*, the coalition was as logical and as inevitable as traditional criticism of the play suggests, and whether

Hugo's action was truly unexpected.¹

One of the essential passages of *Dirty Hands* is of course the negotiation between three hitherto hostile Illyrian political factions. In an initial stage the Regent's group, the Pentagon group, and the Proletarian party have waged war mercilessly. The Proletarian party leader, Louis, speaks of "Trois groupes d'hommes d'intérêts inconciliables, trois groupes d'hommes qui se haïssent" ["Three groups of men with irreconcilable interests, three groups of men who hate each other"] (49). The Prince reminds Karsky that "vous avez organisé l'an dernier deux attentats contre mon père" ["last year you organized two assassination attempts against my father"] (141). And Hoederer observes that "Quand un type du Pentagone rencontrait un gars de chez nous, il y en avait toujours un des deux qui restait sur le carreau" ["When a guy from the Pentagon met one of our guys, there was always one who didn't get up"] (146).

The Pentagon hates the Proletarians and the Regent; the Regent hates the Pentagon and the Proletarians, and the Proletarians in turn hate the other two. The game presented in normal form in Figure 1 shows the possible decisions for any pair of these players and also their preferences before their negotiations and possible alliance. The dominant strategy for each group is the same, that is, to defeat the other. For this reason, player A may just as well represent the Prince as Karsky or Hoederer, and the same is true for player B. Apparently we have situations of total conflict, since A's preference is symmetrical to B's. Indeed either one's best choice (4) is also the worst choice (1) for the other. These are, therefore, zero sum games.

Figure 1. The Game of the Parties

		B	
		Pouvoir	Non pouvoir
A	Pouvoir	(3, 3)	(4, 1)
	Non pouvoir	(1, 4)	(2, 2)

The situation just described, which seems perfectly clear, could have lasted indefinitely, as long as each belligerent party sought the

¹*Les Mains sales* (Paris: Folio, 1971), p. 30. Henceforth in text. De Ley, *Le Jeu classique. Jeu et théorie des jeux au Grand Siècle* (Paris, Seattle, Tübingen, 1988), p. 138.

maximum payoff. However, the matrix also shows us that the second choice (3) is the same for each player, that is, that each would prefer to retain power, even if the other group(s) retain power also. This is interesting, since it opens up the possibility of compromise and, specifically, negotiation. It means also that if some outside event intervened to change the situation, the two groups might unite rather than disappear from the scene altogether. And if two or more parties unite, the nature of the Game of the Parties is changed.

Of course just such an outside event occurs with the arrival of the Russians, requiring a partial change in strategy offering different payoffs. In *Dirty Hands*, the Russians are presented as a force of nature, all powerful, enigmatic, whose intentions can only be guessed at, but whose intentions can nevertheless be guessed with some certainty. Confronting the overwhelming power of the Russians, each group wants to retain power, but can only do this now if they accept the other(s) as allies. The objective for Karsky and the Prince is to retain power, and this can happen only if they find agreement with the Proletarian Party, to avoid being swept aside, purely and simply, by the Russians. Each non-proletarian group must therefore adopt a mixed strategy, safeguarding its own interests while cooperating with the others. This is why the Prince declares that “Aujourd’hui les positions se sont brusquement rapprochées...” [“Today, the positions have suddenly drawn closer together”] and he assures Hoederer that “Karsky estime que les divisions intestines ne peuvent que desservir la cause de notre pays” [“Today the positions are suddenly much closer... Karsky believes that internal dissension can only weaken the cause of our country”] (142-143).

This diplomatic discourse inviting everyone to come together in peace does not totally substitute for Karsky’s much less “fraternal” offer to Hoederer. Karsky is more inclined to play with Hoederer the classic game theory Game of the Wholesaler and the Retailer. In this game, two players have a certain interest in reaching agreement, but also seek to find the most advantageous possible price to charge or to pay. Naturally it often happens that the wholesaler, thanks to his greater resources, seeks to impose a price on the weak retailer. This is precisely what happens when Karsky, on the strength of his party’s support by fifty-seven per cent of the population, proposes “deux voix pour votre Parti dans le Comité National Clandestin que nous allons constituer” [“two votes for your party in the National Clandestine Committee we’re going to form”] (143-144). In Karsky’s plan, this will be two votes out of twelve. Karsky also wants “un remaniement et une fusion par la base de nos deux organisations clandestines. Vos hommes entrèrent dans notre dispositif pentagonal” [“a restructuring

and a fusion at the grass roots level of our two clandestine organizations. Your people will join the Pentagon forces”] (144).

The game of wholesaler and retailer does not stop, of course, with the first offer, and if the wholesaler’s price is too high the retailer can make a counteroffer or even refuse to play. Hoederer does indeed reject Karsky’s proposal. However, he continues to play the game as he changes the stakes:

Un comité directeur réduit à six membres. Le Parti Prolétarien disposera de trois voix; vous vous repartirez les trois autres comme vous voudrez. Les organisations clandestines resteront rigoureusement séparées et n’entreprendront d’action commune que sur un vote du Comité Central. (145)

[A steering committee reduced to six members. The Proletarian Party will control three votes; you can divide up the three others any way you want. The clandestine organizations will remain strictly separate. They can act together only on a vote of the Central Committee.]

Hoederer reverses the roles to his advantage and reinvents the coalition on his own terms: he proposes the conditions, and the others can only accept joining him or refusing to play. This position is perfectly rational if we take into account another notion of game theory, which is that of the Value of Shapley. Shapley’s value (v) represents the gain which a given player brings to a coalition. It is of capital importance in *Dirty Hands*, where indeed it determines the nature of the coalition. As explained by Andrzej Novak and Tadeusz Radzik, two mathematicians using game theory, the classic form of a game of several players, in which various forms of cooperation are possible, is a function v which assigns to each coalition S a single real number which is the value—or the power—of S (150). As they study the formation of a coalition in an n -person game, they see a sequential process in which each player brings some value to the coalition, but this value depends not only on the players’ contribution(s) but also on the order in which the coalition is formed.

Taking this situation into account, the nature of the negotiations becomes obvious. Hoederer knows that he is, so to speak, the first member of the coalition. His adversaries are actually in a situation of weakness. The Prince has been a German sympathizer and Karsky represents primarily the bourgeoisie. Both are therefore in a poor position vis-à-vis the Russians—unlike the Proletarian Party, which has always supported the communist line. Clearly, a coalition would serve no purpose without the Proletarians. Karsky has tried to avoid a disadvantageous arrangement, but risks the possibility that Hoederer may abandon him in favor of an agreement with the Prince—who has the most to atone for and who therefore might make a more favorable agreement.

At this point in the negotiations, it is important to compare the traditional critiques of this passage with the lessons of game theory. For Françoise Bagot and Michel Kail, for example, Hoederer represents the purest realism. Their opinion is shared by Mark Buffat, for whom the negotiations “sanctionne la justesse des vues politiques de Hoederer” [“confirm the wisdom of Hoederer’s political project”] (72). He adds that “Hoederer “leur présente une image adéquate du réel. Et s’ils finissent par accepter ses propositions, c’est qu’ils en reconnaissent la vérité” [“He presents them with an adequate image of the real”] (74). Thus for the critics, Hoederer’s analysis as presented up to this point is lucid and adequate to the real. He is objective. He recognizes the divergences of the two other parties and does not seek to play them off against each other, but rather tries to integrate them in order to go beyond them. It appears that everyone, in the end, recognizes the correctness of his approach—the Regent, Karsky, and later Jessica and Louis and Olga. At the same time, for reasons of his own and to which I will return in a moment, Hugo remains the single exception.

It seems to me, however, that the success of the coalition depends less on the Prince and Karsky’s political realism than on Hoederer’s somewhat different political analysis. Actually, neither Karsky nor the Prince ever asks himself why Hoederer accepts negotiation. This is particularly curious when Hoederer asks the two others to remember that

quand les armées soviétiques seront sur notre territoire, nous prendrons le pouvoir ensemble, vous et nous, si nous avons travaillé ensemble; mais si nous n’arrivons pas à nous entendre, à la fin de la guerre mon parti gouvernera seul. (148)

[when the Soviet armies are in our territory, we will take power together, you and us, if we have worked together; but if we don’t reach agreement, then at the end of the war my party will govern alone.]

From one point of view Hoederer’s best strategy would be to adopt a passive policy and wait for the arrival of the Russians—and, of course, to make the coalition fail. Yet Hoederer does exactly the opposite, apparently without raising the slightest doubt among his adversaries.

The reason is that Hoederer possesses, in game theory terms, perfect information while the others possess imperfect information. In slightly different terms, one might say that Hoederer plays a different game than his opponents. Rather than laying out a position of pure realism, Hoederer has other views and has taken care not to say everything. As he explains later to Hugo, the Party’s taking power alone would place it in an untenable situation. It would be the party of a foreign power, a party imposed by the Russians. Moreover, the country is in ruins, and the measures necessary to rebuild it can

only be unpopular. If the Party takes power, it will be swept away by insurrection. Hoederer's solution will be to avoid responsibility for any drastic measures and remain in opposition even as it tacitly takes part in government: "Une minorité, voilà ce que nous devons être" ["A minority, that's what we should be"] (193) will be the temporary slogan of the Party. Of course, such a strategy should not be communicated to the Regent nor to Karsky, since it would be their turn to refuse any negotiation, which refusal would push Hoederer and his supporters into power, much to Hoederer's great regret.

The foregoing demonstrates in addition how the possession of perfect information leads to inertia. This remark is supported by the research of Barton Lipman, economic theorist and specialist in game theory, who has demonstrated that common understanding of rationality is different from common knowledge of the facts—for what is rational is a function of what is known (116). For this reason, he concludes, one may expect convergences to fail when mutual knowledge approaches infinity.

In this connection the reader may also want to think about the limits of game theory in a situation of perfect information opposing two players pursuing the same objective. Louis-Jean Calvet, in an only partly mathematical analysis, finds game theory too preoccupied with guessing the choice of the other, which may lead to *immobilisme*. He posits that the rational player may think that "s'il pense que je vais faire le choix 1 alors il va faire le choix 2 donc je vais faire le choix 3" ["If he thinks I'm going to choose one, then he will make choice two, so I will make choice three"] (193). And so on. Calvet quotes Michel Plon, who suggests that such reflections draw the players into a "mouvement perpétuel fondé sur le calcul des intentions de l'autre" ["perpetual movement based on guessing the intentions of the other"] (193). Plon and Lipman's ideas are described by some others as the problem of infinitely rapid feedback, which may lead not to perfection but to erratic reactions.

This limitation of game theory throws some light on the situation of the three political factions. Rather than enlighten his adversaries, Hoederer chooses to feed their own ideas back to them and make himself the interpreter of their own views. Hoederer's supposedly lucid vision is perhaps less to be admired than Karsky and the Prince are to be faulted for their own limited understanding. Their strategic weakness is perhaps all the more remarkable since it reveals the lack of good sense in leaders who do not even ask themselves the initial question of any coalition formation: why would my adversary wish to accept? A game theoretician might even see in this state of affairs a weakness in the strategic problem presented by Sartre: the game

depends on Karsky and the Prince's failure to assess correctly their own position.

This weakness of strategic thought is perceived, in a sense, by Bagot and Kail. These critics see it, precisely, as a form of inertia. They emphasize that Karsky and the Prince are "voués à la défense d'un ordre social ... qui contient l'essentiel et n'a plus qu'à perdurer à la faveur du principe d'inertie" ["committed to defense of a social order ... which contains everything essential and needs only to continue by inertia"] (87).

From this point of view they see time as the immutable order of the past; what was, is and will be so indefinitely. They are the representatives of an order which, according to themselves, cannot be modified. In such a situation the future has, therefore, no importance. The paradox is that the inertia of their thought allows the movement of the play and advances the strategy of Hoederer.

Whereas Hoederer plays masterfully his game with Karsky and the Prince, taking advantage of their imperfect information, their limited analysis, and a mental inertia not totally unknown to game theory, he plays a different and more complex game with Hugo, with somewhat different results.

For Hoederer the end justifies the means. "Tous les moyens sont bons," he declares, "quand ils sont efficaces" ["Any means is good, as long as it works"] (197). Hoederer is pleased, therefore, when he learns that his people have taken over an arms stockpile from Karsky's group. Karsky whines that Hoederer's men outnumbered his ten to one. Hoederer responds that "Quand on veut gagner, il vaut mieux se mettre à dix contre un, c'est plus sûr" ["When you want to win, it's better to be ten to one: better to be sure"] (140). Hoederer has a corresponding dislike for ideas and idealism. When Hugo expresses his political ideals, Hoederer concludes: "C'est vrai: tu as des idées. Ça te passera" ["It's true, you have ideas. But you'll get over it"] (195).

Examined more closely Hugo's attitude toward ideas is complex. He believes in ideas; he believes in absolute values, applicable in all situations. Naturally this makes him an idealist, someone who acts as if things happened in reality just as they might in an ideal world. As a result, Hugo lives in abstractions--an absolutely simple-minded view of reality. Hugo has nothing concrete to propose, since for him everything is abstract, even the other and perhaps especially the future. In a certain sense, his idea of reality is a fiction.

At the same time, however, Hugo is an intellectual who refuses his own way of thinking. Although formed as a thinker, he would prefer

to be like Kyo in Malraux's *La Condition humaine*. In *La Condition humaine* Kyo is the opposite of an intellectual. He sees the world through action, never asking questions, sustained by a true communist faith. As Francis Jeanson notes, Hugo would like to "tuer en lui-même ce pouvoir de réflexion... Seule l'action le délivrerait de son mal mais ce mal est précisément ce qui le rend incapable d'agir. Du reste, ce mal est celui de tous les intellectuels... Et naturellement l'action dont ils rêvent c'est d'emblée la plus héroïque" ["kill in himself this power of reflection... Only action can deliver him from his sickness, but it is precisely this sickness that makes him incapable of action. In fact this problem is the same for all intellectuals... And naturally the action they dream about has to be the most heroic"] (46).

Hoederer's objective in the Game of Hugo and Hoederer is to get Hugo to move from the illusion of idealism to the reality of politics. He wants to persuade Hugo and disdains the use of force: "Je lui parlerai," he says, "Je veux le convaincre" ["I'll talk to him... I want to convince him"] (213). What Hugo needs to be convinced of is, of course, that he would be wrong to kill Hoederer. And Hoederer wants him to discover this through the use of his own reason. In order to do this, acting just as he did with the Prince and Karsky, Hoederer wants to play back to Hugo a version of Hugo's own ideas—what Hugo would think, say, or do if he acted rationally, that is, according to the rationality of Hoederer.

However, in reality, each player plays, so to speak, in a different dimension. Hoederer is rational in his own terms and believes Hugo is irrational. Hugo thinks the same of him. Each one lives in a logical system which makes each one believe the other is wrong. The situation can easily be modeled in a game matrix as in Figure 2.

Figure 2. The Game of Hoederer and Hugo (I)

		Hugo	
		Realpolitick	Idealism
Hoederer	Realpol.	(4, 1)	(3, 3)
	Idealism	(2, 2)	(1, 4)

The realist Hoederer wants to persuade Hugo. His first choice therefore, is that he choose Realpolitick and Hugo do the same (4). His worst choice would be to adopt Hugo's idealism (1). In the same

way, Hugo would like Hoederer to be idealistic, and indeed Hugo respects Hoederer enough to go to some trouble to try to persuade the older man (4). Hugo's worst choice would be to share Hoederer's political cynicism (1). In Figure 2, the game matrix exposes one possible model of this situation: both players have dominant strategies, and these strategies nowhere coincide. One might say with Pascal that the two players' strategies belong to different orders of magnitude. Hoederer pursues the greater good through unattractive means. Hugo refuses the "dirty hands" of the political schemer and prefers the clean solution of the revolution.

Translated into action, however, the two players' strategies are somewhat different. In one possible modeling, the pragmatist Hoederer may care little for Hugo's philosophy, but only whether Hugo will shoot him or not. In trying to avoid getting shot, Hoederer may nevertheless prefer to recruit Hugo, or simply replace the father that Hugo detests. Under these conditions, Hoederer may want to persuade Hugo (4), but in any case to keep him from shooting (3). His worst alternative is perhaps to persuade Hugo but get shot anyway, which is one possible description of what actually happens in the play. Hugo, for his part, might simply choose whether to shoot or not. This second version of the game of Hoederer and Hugo is modeled in Figure 3.

Figure 3. The Game of Hoederer and Hugo (II)

		Hugo	
		Shoot	Don't shoot
Hoederer	Persuade	(1, 4? 3?)	(4, 2)
	Neutralise	(3, 3? 4?)	(2, 1)

From another point of view Hugo, the intellectual, may have some insight on Hoederer's position. He is attracted to Hoederer and would perhaps like to accept Hoederer's views and abandon his plans to shoot him. For Hoederer, Jessica lays out a different version of the Game of Hoederer and Hugo. As she sees it, Hugo would like to be persuaded. One possible form of Jessica's game is modeled in Figure 4. It gives Hugo a dominant strategy: don't shoot.

Figure 4. The Game of H and H (Jessica version)

		Hugo	
		Shoot	Don't shoot
Hoederer	Persuade	(2, 2)	(4, 4)
	Don't	(1, 1)	(3, 3)

In yet another view, however, Hoederer abandons realism to become idealistic. In this latter version of the Game of Hoederer and Hugo, it is Hoederer's turn to play with imperfect information. He disdains the "c'est plus sûr" which conditioned his thinking about Karsky, and he is at first unaware and then fails to take into account the circumstances which led Hugo to become his secretary. The latter, in accepting and even volunteering to shoot Hoederer, placed himself in a zero-sum situation: to kill or be put out of the way by the Party, as Olga reminds him. Her ultimatum is without appeal: "Si dans vingt-quatre heures tu n'as pas terminé ta besogne, on enverra quelqu'un pour la finir à ta place" ["If you haven't finished the job in twenty-four hours, they'll send somebody to finish it for you"]. If Hugo doesn't act, he has a good chance to "se faire buter... comme un petit imbecile dont on se débarrasse par crainte de ses maladresses" ["take a hit like the little imbecile you get rid of so he won't cause trouble"] (173). For Hugo, killing Hoederer is not only an idealistic act, but also a way of saving his own life.

Finally Hoederer is unaware of or forgets Hugo's original motive: by means of murder he wants to make himself a heavyweight, to give reality to his being. Thus Hoederer, when it comes to Hugo, fails to calculate his adversary's intentions. If he had, he would have known that Hugo might not be persuaded, and he would also have known how to avoid becoming his victim. Hugo has, of course, an internal conflict triggered by his own thoughts. As an intellectual, he asks himself "Si je le tue... Qu'est ce qu'il faut que je fasse? Que ferais-tu?" ["If I kill him... What must I do? What would you do?"] (179). This way of thinking makes Hugo someone who is not cut out to be an assassin. As an intellectual, he is incapable of acting in cold blood, without thinking. Indeed, he thinks so much that he does nothing, since to act he would have to answer all the questions he has asked himself, all these and more: What am I doing here? Am I right to do what I'm doing? Am I lying to myself?

Hoederer can succeed with Hugo either by persuading him or by neutralizing him. He can neutralize Hugo either by calling his bodyguards or simply by keeping Hugo in some kind of state of inertia. The latter strategy has to work only for a limited period of time since, as Hoederer tells Jessica, “S’il ne fait pas son coup ce matin, il ne le fera jamais” [“If he doesn’t do it this morning, he’ll never do it”] (213). The indicated strategy, perhaps his best strategy short of the use of force (2), is to keep Hugo thinking, or in other words to prevent Hugo from not thinking (3).

The intellectual Hugo’s best choice meanwhile would be to kill Hoederer in full awareness of his act, for then he would think and conclude in favor of action (4). His second choice would be to kill Hoederer without thinking, if only to save himself from Louis’ men. His last choice would be the one in which he actually finds himself, that is to want to kill without being able to (1). This last version of Hugo and Hoederer’s game is modeled in Figure 5. Figure 5 is a two-by-four matrix, since Hugo possesses four alternatives.

Figure 5. The Game of Hugo and Hoederer (IV)

		Hugo			
		Shoot		Don't shoot	
		unconc.	conc.	unconc.	conc.
Hoederer	Persuade	(1, 3)	(1, 4)	(4, 1)	(4, 2)
	Neutralise	(3, 1? 2?)	(3, 2)	(2, 1)	(3, 2)

It is of course only Jessica’s intervention which stops Hugo thinking and permits him to act. Instead of choosing his worst solution (1), he moves to his second best (3): kill without thinking.

If the play were perfectly in agreement with the principles of game theory and with the rational decisions that Sartre wanted for his theatre, the action would have been blocked on two occasions. First the negotiation might never have concluded, since Karsky and the Prince would have understood Hoederer’s true utilities and his “he who loses wins” strategy. One might wonder, moreover, what became of the counselors each must necessarily have had to keep them from strategic blunders. On the second occasion Hoederer could have remained in a certain immobilisme and avoided senselessly provoking Hugo. Indeed Hoederer makes a manful effort to do so, before giving in to temptation. By giving up thinking himself, Hoederer gives

up his winning strategy(ies). And by giving up thinking, Hugo acts totally illogically with respect to his own previous behavior, perhaps a happy outcome for a Sartrean hero. He ends up with his second best choice (3), but no doubt for the greater good of the play, which continues and finds its ending.

That is, until Hugo gets out of prison. It is at that time, in the last tableau, that Hugo is able to become momentarily a Party heavyweight and give a political rationale to his action. In so doing, in his famous "Non-récupérable," he moves from his second best choice to the choice that was first for him all along.

Stéphane Pillet

University of Puerto Rico at Mayagüez
Puerto Rico

Works Cited

- Bagot, Françoise and Michel Kail. *Jean-Paul Sartre. Dirty Hands*. Paris: P.U.F., 1985.
- Buffat, Marc. *Dirty Hands de Jean-Paul Sartre*. Paris: Folio, 1991.
- Calvet, Louis-Jean. *Les Jeux de la société*. Paris: Payot, 1978.
- De Ley, Herbert. *Le Jeu classique. Jeu et théorie des jeux au Grand Siècle*. Paris: Biblio 17, 1988.
- Eber, Nicolas. *Théorie des jeux*. Paris: Dunod, 2004.
- Giraud, Gaël. *La Théorie des jeux*. 2 ed. Paris: Flammarion, 2004.
- Jeanson, Francis. *Sartre par lui-même*. Paris: Seuil, 1955.
- Lipman, Barton. "A Note on the Implications of Common Knowledge of Rationality," *Games and Economic Behavior* 6 (1994): 114-129.
- Novak, Andrzej and Radzik Tadeusz. "The Shapley Value for N-Person Games in Generalized Function Form," *Games and Economic Behavior* 6 (1994): 150-161.
- Sartre, Jean-Paul. *Dirty Hands*. Paris: Folio, 1971.
- _____. *Un théâtre de situations*. Paris: Gallimard, 1973.