Prisoner's Dilemma Variations and Their Applications

Andrew Elias

13 May 2011

The Prisoner's Dilemma (P.D.) is a game studied as a prime example of Game Theory and its applications to real-life situations. It has been studied by sociologists, psychologists, economists, mathematicians, biologists, computer scientists, as well as those from many other fields. It does a remarkable job of helping people study what happens when individual incentives are in direct conflict with group incentives. The Prisoner's Dilemma is also a perfect illustration of how simple rules and motives at the micro level can lead to chaos and seemingly irrational behavior at the macro level.

The basic Prisoner's Dilemma involves a simple situation: two players, Al and Bob, are paired up, and each is given a choice between two options. Each person can either choose "Cooperate," indicating that they will cooperate with the other player, or "Defect," indicating that they have chosen to betray the other player. Because they can cooperate, the players are not called "opponents," but instead "peers." The name "Prisoner's Dilemma" comes from a specific formulation in which prisoners are tempted to betray each other. The game only has four possible outcomes: "Both players cooperate," "Both players defect," "Al cooperates and Bob defects," or "Al defects and Bob cooperates." It will soon be shown that the narrow range of outcomes in no way implies that the game is easy to play or understand.

The available moves are tied to economic incentives. For example, the moderator of the game will have some means of either giving out money (money is a simple example, but in real life, the incentives are more abstract), taking it away, or both. If both players defect, they each get nothing. If both players cooperate, they stand to each gain three dollars. At this point, it should be clear that "Both cooperate" is the best option so far. But here's the kicker: if Al cooperates and Bob defects, Bob gains five dollars while Al loses three dollars. And if Bob cooperates and Al defects, Al gains five dollars while Bob loses three dollars. These final two situations complicate things; now, Al and Bob have substantial motivation to defect against each other; Al, for example, sees that no matter what Bob does, Al's payoff is better when he defects. Bob sees the same thing. The rational choice for each player is to defect, and "Both defect" leaves each player with nothing. This is a disheartening situation, because, if only the two players could work together and both choose Cooperate, their total economic sum, six dollars, would far exceed the zero dollars they are getting now. But, even knowing this, Al could never choose Cooperate, because he knows that Bob, in his rationality, will always play Defect. After all, when Al chooses Cooperate, Bob can make two extra dollars ($5 - 3 = 2$) by playing Defect. And Al doesn't want that, because it means he loses three dollars!

It is clear to see why the Prisoner's Dilemma is called a dilemma; while each player aspires to receive the benefits which arise when Cooperate is played, neither of them can rationally play Cooperate. For this conundrum to arise, certain conditions must be met. First, "Both cooperate" must be the best outcome for the group's cumulative wealth (+\$6 in our example), and "Both defect" must be the worst outcome for the group's cumulative wealth (+\$0). The Temptation, the individual's benefit from being the sole defector (+\$5), must be the best outcome for each individual. The Sucker's Payoff, the individual's cost from being the sole cooperator (−\$3), must be the worst outcome for each individual. Finally, although this is not initially obvious, the average of the Temptation and the Sucker's Payoff (+\$1) must not exceed the amount each individual gains when both cooperate (+\$3) [3, p.204]. This ensures that "Both cooperate" is the one and only best outcome for the group as a whole. All of these constraints can be notated in the following way, where $(C, D)_X$ denotes X's individual gain when he played Cooperate while the other played Defect, and $(C, D)_G$ denotes the group's cumulative gain when X played Cooperate while the other played Defect:

For each player X,

$$(C, D)_X < (D, D)_X < (C, C)_X < (D, C)_X$$

and

$$(D, D)_G < (C, D)_G < (C, C)_G$$

While these may seem a bit restrictive, there are actually plenty of examples of the Prisoner's Dilemma in everyday life. For example, when two corporations that rely on the environment cooperate to preserve it, they can both benefit greatly. But if one of them defects by polluting the environment, it can reap substantial benefits, while the cooperating corporation suffers losses due to its somewhat futile environmental efforts, which were ineffective due to the low cooperation rate. And when both take the rational route and pollute their environment, they both suffer greatly. While both corporations want to help the environment, they are in a situation where they unintentionally force each other to harm it.

The dilemma also occurs in personal relationships. When two people get divorced, the best interaction would be to cooperate by dividing possessions evenly, refraining from "slinging mud," and certainly avoiding dragging expensive lawyers into the situation. But each individual sees that, no matter what the other does, hiring a lawyer will result in the better individual payout. So they each spend loads of money hiring their own lawyer, they throw each other's proverbial dirty laundry all over the public record, and, in the end, they end up dividing possessions evenly anyway. If only they could both cooperate, the experience would be better for both of them, and furthermore, the experience would be better for each individual.

Because the only equilibrium in the Prisoner's Dilemma is "Both defect," there have been countless efforts to adjust the situation such that both players can confidently play Cooperate. The solution to the environmental dilemma established earlier is government regulation. On the surface, pollution-preventing laws prevent corporations from defecting, which is good. Going one level deeper, the law actually serves a secondary purpose: it gives each corporation the confidence that its peers are cooperating, therefore reducing the fear that it will be stuck with the Sucker's Payoff. Even if the players don't personally trust their respective peers, the players can trust their peers' fear of the law. The end result is that, when the government steps in, players, the environment, and relationships all benefit. Because the Prisoner's Dilemma so readily shows that government regulation can be used to improve upon laissez-faire capitalism in this way, it is often used as support for socialist ideas. Unfortunately, not all solutions can be solved in such a manner; no law will ever be passed preventing divorcers from hiring lawyers, even though in most cases, such a law would benefit the divorcers.

It should now be clear that the Simple Prisoner's Dilemma is a crucial relationship dynamic that can be used for studying sociological interactions, but it only scratches the surface. It only covers single-interaction, two-participant relationships in which both players act simultaneously. There are many variations on the theme, which can be extended to cover almost any relationship involving a dilemma.

The first important variation is the Infinitely Iterated Prisoner's Dilemma. This is when two players, Al and Bob, repetitively engage in the Simple Prisoner's Dilemma ad infinitum. The crucial difference here is that players have a means of ensuring trust. Although Al still has the short-term incentive to defect, he now has a more strategic option; he can sacrifice his first move, playing Cooperate, thus ridding Bob of the idea that Al must play Defect. The same logic works from Bob's point of view, and both players cooperate. The players each realize that they are in a great situation: if the payoffs were the same as the example at the beginning of this paper, they are now each gaining three dollars per move instead of zero dollars. But what happened to the Temptation to defect? In the Simple P.D., Defect was a lucrative move from an individual's perspective. But now that both players have a good thing going, they both see that Defect, while leading to a short-term individual gain of two extra dollars, would ultimately lead to long-term losses (or rather, lack of gains) when the trust is broken. It is important to notice that the Infinitely Iterated P.D. only exists when both players recognize that they are locked into an infinite sequence of games with the same peer. If they have no way of recognizing each other, they will assume that, at each iteration, they are facing a fresh, new peer. This is no different from playing the Simple Prisoner's Dilemma a bunch of times, and the outcome is equally disheartening; both players will surely defect.

An even more interesting variation is the Finitely Iterated Prisoner's Dilemma, which only lasts a certain number of moves and then ends. It is natural to think that this should be almost no different from the Infinitely Iterated P.D., but that is not the case. Because the final iteration is exactly the same as the Simple P.D., its outcome is no different: both players must defect. And because there is therefore no true choice in the last iteration, the second-to-last iteration is also the same as the Simple Prisoner's Dilemma, again resulting in both players defecting. By backwards induction, all moves should theoretically be locked-in from the start; both players should always defect. To the chagrin of logicians, people actually play the Finitely Iterated Prisoner's Dilemma the same way they would play the Infinitely Iterated Prisoner's Dilemma, up until the last few moves. It has been suggested that this implies that human gut-feelings do not discern a practical difference between long periods of time and infinitely long periods of time. Some also say that this shows that humans do not naturally tend to plan for the distant future, although it does not prove the statement.

The most plausible explanation for the discord between the logically deduced outcome and the empirically shown outcome of the Finitely Iterated P.D. involves the type of information to which the players have access. When two people each know a fact, they possess "individual knowledge" of that fact, whereas if, in addition, they each know that the other knows the fact, and they each know that the other knows that they know the fact (ad infinitum), then they possess "mutual knowledge" of that fact [9]. Common Knowledge Rationality (C.K.R.) exists when both players possess mutual knowledge of the rules and the rational strategy for a certain game [10]. Most theoreticians, including the ones attempting the aforementioned backwards-induction proof, assume C.K.R. when studying the Prisoner's Dilemma. Varoufakis explains that, while both players possess rationality in real-life scenarios, they do not possess Common Knowledge Rationality; each player doubts that the other player has the knowledge that brings him to the decision to play Defect. David Kreps, in an article entitled, "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma," examines the possibility that people subconsciously calculate a Bayesian probability that the other player is rational and well-informed. Because the probabilities are usually less than one, the players can play the Finitely Iterated P.D. with a strategy roughly similar to the one they would use to play the Infinitely Iterated P.D. But there is still a time-horizon at which things change. When there are, say, three moves left in the game, Al sees that the game will end soon, but more importantly, Al sees that Bob sees this. In suspicion that Bob will start to defect for the remaining iterations, Al defects, because he has nothing to gain by cooperating. Bob comes to the same conclusion at roughly the same time, and, like a switch, the Finitely Iterated P.D. stops resembling the Infinitely Iterated P.D. and starts resembling the Simple P.D. The amount of time between this time-horizon and the end of the game is arbitrary and, in reality, impossible to predict. The underlying cause of the collapse is the destruction of the doubt of Common Knowledge Rationality.

An important example of this phenomenon is the collapse of economic bubbles. Bubbles form when speculators invest in a stock or commodity (think dot-com startups or gold) that is already priced too highly. They all make this seemingly irrational choice in hopes that others will make the same seemingly irrational choice, thus driving up the price. Notice that this is not a completely irrational choice; when others invest, the price goes up, increasing the value of the shares held. In this situation, all investors are playing what they think of as an Infinitely Iterated P.D., and their mutual cooperation allows them to all benefit. The problem occurs when the arrival of new knowledge about the stock/commodity destroys the investors' doubts of Common Knowledge Rationality. All at once, investors realize that other investors have realized that selling (defecting) is now the logical move, as the price will soon plummet. This causes the price to plummet, which causes further destruction to the 'trust' established between the investors, which in turn causes more selling, causing further price drops. The feedback loop can create a massive crash in the stock/commodity price, which only stops when its inherent value has been reached.

The Finitely Iterated P.D. can also be applied to sociological relationships. Racism, for example, can be considered the result of a Prisoner's Dilemma between two races. I once witnessed an instance when football players from a predominantly white high school were given the opportunity to skip school to attend the pep rally for a predominantly black high school across town. The white football players refused to support their black peers, citing their suspicion that the black football players would act in a racist manner towards them. It's a classical example of a self-fulfilling prophecy: the fear of racism actually caused racism. The white players were defecting for the sole reason that they thought their peers would defect. If only the two races could build up enough trust in each other, they could both benefit indefinitely from mutual cooperation. But it seems that in this example, the Finitely Iterated P.D. between the races has long-ago crossed its time-horizon, and the only way to break the streak of mutual defection would be to establish doubt of Common Knowledge Rationality, or in other words, to fool people into thinking that racism doesn't already exist. This supports Morgan Freeman's famous quote, "Stop talking about race, and racism will end." Unfortunately, if and when racism does end, the world will be in a sort of 'bubble' of trust which, similarly to economic bubbles, could burst at the slightest hint of C.K.R., leading to disastrous consequences.

The next theoretical question to arise is this: given that streaks of trust, or 'bubbles,' can be mutually beneficial but are inherently risky, what is the best strategy for playing the Finitely Iterated P.D.? Concise mathematical proofs are of no use here, because a player's optimal strategy is highly dependent upon that of his peer. In 1980, political scientist Robert Axelrod set out to run a computer-simulated, round-robin tournament, pitting many human-designed strategies against one another. There were two round-robin tournaments, each preceded by a call for entries [1] [2]. Tournament entries came from sociologists, psychologists, economists, mathematicians, computer scientists, and biologists alike [1, p.21-24]. The ultimate result was a list of properties generally describing the highest-earning strategies in the Finitely Iterated P.D. Surprisingly, there were very few properties having much to do with defecting often; it was shown that nice, forgiving, provocable (a more fair/forgiving version of 'vengeful'), and optimistic strategies were ultimately the most successful. The overall winner was a very simple strategy called Tit-for-Tat, which, on each move, copies its peer's previous move. It is optimistic, meaning it plays Cooperate on the first move by default, so it is inclined to be nice. But, to its merit, it has the capacity to 'punish' defectors by defecting against them on the next move. Axelrod noted that if Tit-for-Two-Tats, and even nicer and more forgiving strategy, had been entered into the tournament, it would have won. And when Marcus Frean, in an unrelated study, examined the Asynchronous Iterated P.D., which is usually more realistic, he also came to a similar set of positive traits for a strategy to have. He concluded that strategies should be apologetic, nonexploitative, retaliative, and exhibit guarded generosity.

Because round-robin tournaments are not very good at representing the dynamical environment of players existing and evolving in a real life environment, Axelrod set up a third simulation, this time providing an evolutionary environment in which populations of certain strategies could wax and wane [2]. The idea was that certain strategies would emerge as Evolutionarily Stable Strategies [8], or in other words, strategies which, when dominating the population, would continue to dominate the population even amongst other strategies' efforts to overtake them. The population graph that resulted was much more chaotic than expected; no single strategy stayed on top. Some strategies rose and maintained a strong hold, but then crashed when a 'predator' strategy arose to take advantage of them. Sometimes two strategies worked together in a symbiotic relationship to dominate the population. When one of these would start to die off, the other would as well, and a new 'best strategy' would emerge. Naturally some strategies were better than others, and it turned out that the good traits from the previous two experiments were the same in the third experiment. But there was no longer a definite measure for answering "What is the best strategy?" Instead, Axelrod showed that the ecological environment plays a crucial role in determining what the best strategy is. Dawkins adds that, while Axelrod only considered the global environment, local environments also play a crucial role; a population of Always-Defect players is an Evolutionarily Stable Strategy, but a local population of Tit-for-Tat players could benefit off of each other to the point where they can overtake the large population of Always-Defect. Clearly Iterated Prisoner's Dilemmas are capable of producing complex dynamics which cannot be reliably predicted.

There is yet another notable variation on the Prisoner's Dilemma that makes it able to model more real-life situations. The N-Person Prisoner's Dilemma occurs when more than two people engage in a game in which there exists a discord between the group's interests and the individual's interests. Often, the N-Person P.D. is studied as a two-person P.D in which one player is 'Al' and the other player is 'the rest of the world.' This is generally effective when examining the situation from Al's point of view, because the discord, and therefore the dilemma, mostly still exists. But a key difference occurs when Al has the capability to band together with some other people who he trusts, to form a subgroup [4]. When taking n to be the total number of people cooperating and N to be the total number of people in the population, the existence of the Prisoner's Dilemma is sometimes dependent on n. In the Simple P.D., it has already been shown that there are necessary conditions implying that $C(1) < C(2) < D(1)$, where $C(x)$ indicates the payout to each Cooperator when x (qty.) people cooperate, and $D(x)$ indicates the payout to each Defector when x (qty.) people cooperate [4]. In the N-Player P.D., these constraints can be generalized to $C(n) < C(n+1) < D(n)$. The dilemma exists for all $n \leq N$ which satisfy these constraints. But it turns out that, when one person can influence r (qty.) people, including himself, to decide on a particular choice, the constraints change to $C(n) < C(n+r) < D(n)$. This is critically different from the previous case, because for most n, there exists an r making these constraints not hold. In other words, if Al can band together enough friends to cooperate together, they can avoid the Prisoner's Dilemma altogether; when they convert their decision from Defect to Cooperate, their individual payoffs increase, but because $D(n) \leq C(n+r)$, they don't have the motivation to Defect against each other or the rest of the population. This has some subtle kinks which Hamburger addresses, but overall, the logic is sound. And depending on the graphs of $D(x)$ and $C(x)$, the N-Player P.D. can have dynamics with complexities rivaling that of the Finitely Iterated P.D. For example, there may be a range of possible values of n for which the standard conditions, $C(n) < C(n+1) < D(n)$, do not hold, but the r-modified ones do. Then, once $n+r$ exceeds that range, the standard conditions do hold for all value of n from that point to N. Essentially, multi-person banding can push the economy into a situation in which the Prisoner's Dilemma no longer exists, even for individuals!

The direct implication of the N-Person P.D. is that small enclaves of communism can sustainably benefit society, including themselves, even in a prominently capitalist economy. This can be seen in charities, nonprofit organizations, and petitions. It also provides yet another pragmatic reason for governments to preserve the freedom of assembly.

So far the variations on the Prisoner's Dilemma have come to some remarkable conclusions. They have shown that laissez-faire capitalism can be improved upon by tossing in elements of socialism and small enclaves of communism. They have given economists a way of studying economic bubbles, and they have connected that research to hostility in interpersonal relationships. All this research doesn't even include Asymmetrical Prisoner's Dilemmas, which exist when the payoffs are different for each player. Adding this element can make the models much more realistic. Even when the literal monetary payoffs are equal between players, in iterated games this can still lead to asymmetrical utility functions, which are often determined by the geometric payoffs instead of the arithmetic payoffs [6].

As a whole, the Prisoner's Dilemma provides many tools for examining the conflict between group and individual interests. Simplifying the problem into a game with tangible rewards allows researchers to study the fundamental economic forces at play. And when more realistic aspects are tossed into the mix, unsurprisingly, the conclusions are more applicable to real life interactions. And even devoid of applications, the game's chaotic nature has an intellectual appeal.

About his experiment, Axelrod said, “The sheer number of surprises was itself a surprise.” This trait typifies game theory as a whole, and can lead to further research, ultimately producing a bank of knowledge about social interactions which can help humans cooperate to reach their maximum potential.

References

[1] Axelrod, Robert. "Effective Choice in the Prisoner's Dilemma". The Journal of Conflict Resolution. 24.1 (1980): 3-25.

[2] Axelrod, Robert. "More Effective Choice in the Prisoner's Dilemma". The Journal of Conflict Resolution. 24.3 (1980): 379-403.

[3] Dawkins, Richard. *The Selfish Gene*. 1976. New York: Oxford UP, 2006.

[4] Hamburger, Henry. "N-Person Prisoner's Dilemma". Journal of Mathematical Sociology. 3 (1973): 27-48.

[5] Frean, Marcus R. "The Prisoner's Dilemma without Synchrony". Proceedings: Biological Sciences. 257.1348 (1994):75-79.

[6] Kelly, J. L. "A New Interpretation of Information Rate". Information Theory, IRE Transactions on. 2.3 (1956): 185-189.

[7] Kreps, David M., et al. "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma". Journal of Economic Theory. 27 (1982): 245-252.

[8] Maynard Smith, John. *Evolution and the Theory of Games*. 1982. Cambridge University Press, 2004.

[9] Pinker, Steven. "The Stuff of Thought: Language as a Window into Human Nature". 2010 The Royal Society for Encouragement of the Arts, Manufactures, and Commerce. Youtube. 8 April 2011.

[10] Varoufakis, Yanis. "Modern and Postmodern Challenges to Game Theory". Erkenntnis. 38.3 (1993): 371-404.