"Game Theory in a Liberal Arts Education"

in *International Handbook on Teaching and Learning Economics*

(Gail Hoyt and Kim Marie McGoldrick, editors)

Joseph E. Harrington, Jr.
Department of Economics
Johns Hopkins University
Baltimore, MD 21218
410-516-7615, -7600 (Fax)
joe.harrington@jhu.edu

Introduction

A redesign of the liberal arts curriculum would first and foremost include a course on statistics and, almost as compelling, a course on game theory. It is left to a statistician to argue the former and, as regards the latter, there are several reasons why game theory should be part of a general college education. First, it intersects with many disciplines including economics, sociology, political science, anthropology, history and biology. Second, it provides a way to understand social situations and to systematically analyze and engage in strategic reasoning, which goes beyond being ubiquitous in society to being a defining feature of society. Third, though it is applied mathematics, game theory – in all its glory and not a dumbed down version – can be conveyed with a level of mathematics that students will have mastered by their junior year in high school. This is not to say it is easy but the difficulties are conceptual, not mathematical. Finally, game theory is fun. People routinely engage in strategic reasoning for enjoyment; playing board games, solving puzzles and riddles, discussing strategy in sports and politics, reading mystery novels and the like. While game theory can be presented in a sufficiently technical manner so that such pleasure is abstracted away, it is unnecessary to do so. It was in that spirit that I wrote my undergraduate textbook, Games, Strategies and Decision Making. My approach is to convey the principles of game theory in a lighthearted though rigorous manner, with a minimum of mathematical fuss. In short, game theory is fundamental and fun and deserves to be part of a general college education.

Course Objectives

The first and primary course objective is to teach students how to teach themselves when it comes to analyzing a strategic situation. Game theory provides a framework for thinking about strategic issues and it is that framework, more than a catalog of results, that is likely to have a long half-life in students' memories and to have relevance for their future self, whether as a manager, doctor, lawyer, engineer, you name it. The course's emphasis is not on answers per se but rather on the problem solving process, that is, how one gets an answer. What is to be seared into a student's mind is the logic of a game-theoretic argument, the process by which an answer is derived to a puzzle about human behavior. This is fully consistent with a microeconomics education which promotes a style of reasoning involving a proper appreciation of incentives and equilibrating forces. Game theory, with its drilled down focus on strategic reasoning, is in that spirit.

A second objective is for students to develop their logical reasoning muscles. The game-theoretic framework is a set of tools, and proper use of those tools requires practice in conducting careful, rigorous, systematic analysis. It is all too easy for people to evaluate and decide based on their gut or an emotional response. While there may well be intelligence embedded in visceral reactions, it is one thing to listen to one's gut and quite another to feel that it obviates the need to engage in logical reasoning. Students should be encouraged to regularly and strenuously exercise their reasoning muscles and hone them for purposes of understanding and deciding.

While students should learn how to generate insight on their own, both from having learned a framework and how to use it, the provision of some broadly relevant insight is a third course objective. Science is all about finding commonality, understanding how phenomena are alike because they are driven by the same set of underlying forces, and determining when apparent differences are irrelevant. There is then a fair amount of insight to be conveyed during the course which is applicable to a broad class of situations. Let me provide a few examples.

One of the most significant contributions of game theory is understanding the many cooperative norms and institutions in society, why they occur, where they occur and what allows them to persist. It is valuable is to understand the incentive to cooperate (that is, most equilibria to games are Pareto-inefficient because of externalities) and when cooperation can be a stable outcome even though there is a short run incentive to deviate. Students can learn how sustaining cooperation requires the prospect of future interactions, effective monitoring of choices or outcomes, and the presence of mechanisms to punish.

A second piece of insight concerns signalling. Teaching games of incomplete information is notoriously difficult but it offers some of the most exciting applications. A fundamental insight that signalling games provide is that there can be information in individuals' actions but in order to extract that information you need to understand their incentives, not only in terms of the intrinsic value they attach to an action but also the temptation to deceive. These incentives can imply that a "good" type of person, in order to distinguish herself from a "bad" type, must take such an extreme action that a bad type wouldn't choose it even if she was inferred to be a good type. This mechanism and its behavioral implications are quite common in society. An example known all to well to college students is the considerable effort expended to build up a list of high school extra-curricular activities designed to impress a college admissions officer.

A third piece of insight comes from analyzing basic games, and then determining whether a particular real-world strategic situation has the rudiments of that game. For example, in learning about and how to apply Nash equilibrium, it is standard to consider some classic games: prisoners' dilemmas, coordination games, battles of the sexes, and out-guessing games (a term to capture situations of pure conflict). Students learn that players have a dominant strategy in a prisoners' dilemma and it leads to a Pareto-inferior outcome; that what players do is secondary to doing the same thing as other players in a coordination game; that players want to coordinate in the battle of the sexes but they differ on what it is they want to coordinate on; and, for an out-guessing game, the desire to do something different from what the other player thinks you're going to do prevents the existence of a stable collection of (pure) strategies. Rarely do people face situations as simple as those classic games suggest, but we often encounter situations for which a simple game captures part of what is going on. For example, in spite of its industry clout, IBM's operating system OS/2 partly failed because it fell prey to a coordination failure; a software developer didn't create an application because it didn’t expect other applications to be written and, without many applications, few people would buy OS/2 and, therefore, few applications would be sold. That was one equilibrium but there was another equilibrium in which software developers did write applications because each expected others to do so.

Along the way to providing a framework for producing insight, developing students' abilities to reason logically, and delivering some kernels of insight, a fourth course objective is to convince students that all this matters. The case should be made that game theory is relevant as both a descriptive and prescriptive tool, and it can shed light on the playful (soccer penalty kicks), the practical (FCC spectrum auctions), the pernicious (preventing nuclear war), and the profound (existence of God). If students are going to be put through the pain of head-aching trains of logical thought, they need to be convinced that it is worth doing.

The fifth and final course objective, which is the primary reason I teach game theory, is to have fun! I want us to revel in clever reasoning that we can't wait to tell our friends (in the card game Concentration, you may want to flip a card that you know will not result in a matched pair), examine an everyday situation with a fresh look to show that something deeper is going on (how price-matching guarantees can stifle rather than promote price competition), or to state a conundrum and then solve it (how people can act racist while not being racists).

Teaching Methods

A three-pronged approach is deployed to teach game theory: motivate (generate enthusiasm), build (gradually develop a concept or piece of insight), and deliver (provide something rewarding at the end of this process).

Motivating Students

If students are to learn then their minds must be focused, and the best way to do that is to get them interested. To spark enthusiasm, I begin with a puzzle. It could be an inexplicable phenomenon (such as why, on occasion, soldiers achieved a truce in the trenches of World War I) or a mystifying claim (price-matching guarantees can be anti-competitive) or describing a strategic situation and asking what one should do (how should you bid at a second-price auction). As there is a Gatesian-size wealth of game-theoretic applications from which to draw, an example should be appealing to students, either because it is relevant to them (dating dilemmas, team projects), of some importance or significance (brinkmanship, Galileo and the Inquisition), or just cool or intriguing (penalty kicks, Lord of the Rings). Of course, an example can be mesmerizing either in the sense of being enthralling or in the sense of putting you into a deep coma, and which applies may vary across students. There is no satisfying everyone with any particular game, which argues to presenting a diverse collection of applications during the course.

This enthusiasm can be amplified by augmenting an intriguing phenomenon or situation with rich context. This can take the form of facts, quotations, anecdotes, empirical and experimental evidence, and multimedia. For historical episodes, one can read actual transcripts of people, whether it is from meetings of the Joint Chiefs of Staff during the Cuban Missile Crisis or the journal entries of World War I soldiers. It could be an anecdote from a film, book, or newspaper - a prisoners' dilemma in the cartoon TV show Dilbert, chicken in the timing of the release of Dreamworks' Shark Tale and Disney's The Incredibles, Rock-Paper-Scissors in The Simpsons, or a coordination game among East German civil protestors in 1989. These cases can be described or, better yet, presented using video footage. (For the use of such technologies in the classroom, see the chapter by Mateer and Calhoun.) It could be statistics, such as the percentage of times in English Premier League play that players shoot to the left, center, or right on a penalty kick and the frequency with which the goalkeeper dives to his left, right, or remains in the center; and such facts can be complemented with YouTube clips from World Cup play. All this serves to enrich the question being posed: What should the kicker do? What should the keeper do? Even if it is purely for entertainment, it engages students and that is reason enough. Of course, there is not always time in class to cover background material, which is why a textbook with detailed applications and supplemental reading material are instrumental.

Explaining Concepts

The concepts of game theory can be effectively presented at the lowest common denominator among college students which means, at most, high school algebra. At the same time, it is important to present central concepts regardless of their subtlety or complexity (games of incomplete information and repeated games with non stage game Nash equilibrium punishments immediately come to mind). To deal with this pedagogical challenge, an incremental approach is used which I refer to as “building an explanation.” This may mean (when it is possible) breaking a concept down into its constituent parts, presenting each of those parts, and then putting them together to construct the concept. Or it may involve working with a series of models, going from the simple to the complex, to eventually reach the concept or insight. Or it could mean offering a stripped down version of the concept, which captures some but not all of its essence, and then gradually enriching it.

Consider two examples in which a concept is broken down into its constituent parts. Prior to ever mentioning Nash equilibrium, we first examine what can be learned from more basic primitives; specifically, the game is common knowledge and the rationality of players is common knowledge. As "rationality is common knowledge" encompasses an infinite hierarchy of assumptions (players are rational, players know players are rational, players know players know players are rational, etc.), it is straightforward to incrementally work towards the ultimate goal of understanding what is implied by rationality being common knowledge. Armed only with the assumption that players are rational, the class is asked what we can infer about behavior. The prisoners' dilemma, in the form of Puccini's opera Tosca, is presented and it is quite transparent that assuming only rationality is enough to conclude what Tosca and Scarpia will do. But then another game is introduced, Steven Bram's inventive scenario between man and God on the matter of whether man should believe in God. Here we find that assuming man and God are rational is insufficient to say what will happen. However, if the assumption that man believes God is rational and God believes man is rational is added, then the game can be solved. In this manner we gradually move towards more subtle reasoning. The next setting is a three-player game among athletes who are deciding whether or not to take steroids. Solving it necessitates three levels of knowledge: athletes are rational, athletes believe other athletes are rational, and athletes believe other athletes believe other athletes are rational. The final step is to define rationality is common knowledge and introduce the iterative deletion of strictly dominated strategies (IDSDS). It is then pointed out how the previous three analyses are all special cases of the IDSDS. Starting simple and special and moving toward complex and general gives students a better chance of reaching that final destination by being able to mentally keep pace with the increasing sophistication of the argument.

Another place where this "breaking down the concept into smaller parts" approach is taken is with one of the most challenging but exciting areas of game theory: games with incomplete information. I originally did not cover this material at the undergraduate level because of its difficulty but ultimately came around to doing so because private information is integral to so many strategic situations. Again the discussion begins with a particular scenario; here it is the negotiations leading up to the 1938 Munich Agreement between Neville Chamberlain and Adolf Hitler. In deciding whether to propose concessions to Hitler in exchange for a peace agreement, Chamberlain is described as being uncertain of Hitler's ultimate objectives or, in other words, his payoffs. (In the spirit of gradualism, the private information is initially one-sided.) The initial situation presented before students is Figure 1. We then ask what possibilities Chamberlain may be considering when it comes to assessing Hitler's intentions, of which we consider two. First, Hitler is *amicable* and the payoffs for Hitler are 3,4,2,1 (reading from left to right); second, he is *belligerent* and the payoffs are 4,2,3,1. At this point, we introduce the Harsanyi trick of "turning back the clock" to when Nature determines Hitler's type. Pictorially, this means attaching the *amicable* and *belligerent* extensive forms together with an initial move by Nature that determines Hitler's type (the interested reader is referred to Figures 10.2 and 10.3 in Harrington, 2009). By gradually constructing a game of incomplete information, students better understand it by observing the components that underlie it.

A second way in which gradualism can be deployed is to work with a series of models before getting to the model that delivers the insight or encompasses the concept. This approach is useful to convey how cooperation can be sustained through repeated interaction. The discussion revolves around the following puzzle: How was it that there were sustained episodes of informal truces in the trenches of World War I? A video clip from the classic film All Quiet on the Western Front can be used to provide context by depicting the horror of trench warfare. That carnage is then juxtaposed with cases in which peace "broke out" in some trenches. To provide emphasis, passages are read from soldiers' diaries describing the presence of this truce and how it operated. This then leads us to ask: How do we explain this observed peaceful behavior?

The approach is to construct a plausible model for which an equilibrium has soldiers not trying to kill each other. It starts with a one-shot trench warfare game in which each side decides whether or not to shoot to kill. The game is a prisoners' dilemma and the dominant strategy is to shoot to kill. As we have failed to explain the observed behavior, this leads to an in-class discussion to identify what is missing from our model. Some students will say that payoffs are not as assumed, that these soldiers are pacifists, but then that assumption would fail to explain why, at other times, these same soldiers did try to kill the enemy. If no one has proposed the intended direction to the discussion, the following question is then posed: What was unique about trench warfare as opposed to other theaters of war? Someone will eventually point out that it involved the same people fighting each other day in, day out. This observation suggests a model of soldiers interacting repeatedly over time. We then consider repeating the original trench warfare game twice which, of course, has a unique subgame perfect equilibrium of both sides shooting to kill. After exploring why the same outcome has emerged, the argument is then extended to any finitely repeated game. Finally, we consider a game with an indefinite (or infinite) horizon which ultimately produces the cooperative outcome displayed by troops.

What is appealing about this approach is that students gain a deeper understanding as to why the prospect of future interactions is the foundation for cooperation among people. Taking such an approach can be time consuming but if students are to truly learn a concept, rather than simply results (such as the Folk Theorem), it is important to patiently describe how to go from a question to a model to a solution.

Delivering Value

The third and last facet to my teaching method is to deliver something to students at the end of all of the analysis. If you're going to put them through a long train of logic, which you should, be sure to deliver something worthwhile at the end. The entire exercise should pass a student's internal cost-benefit analysis if we want them to willingly go along on the next thought adventure. The reward at the end could be an insightful answer to a puzzle, as with understanding cooperative behavior in the trenches of World War I. At the same time, you shouldn't be afraid to point out where game theory doesn't work, to tell students there is a phenomenon we don't understand or evidence that runs counter to theory. Thus, presenting experimental evidence of cooperation in a finitely repeated setting or how behavior in the ultimatum game runs counter to the theory can initiate a spirited in-class discussion – why doesn't the theory work? is it preferences? rationality? – and that can be the reward at the end of this process.

The deliverable may be clever reasoning. It could be a cheeky argument for why man should not believe in God or how 1960s "white flight" in response to a few black families moving into a neighborhood could occur even if both whites and blacks actually preferred racial integration. Or what is delivered at the end may be pure entertainment. After solving the Tosca game, a video clip of the scene in which Tosca stabs Scarpia can be shown. Appealing to students' senses can reinforce concepts wrestled within the associated game theoretic model.

A similar case occurs after reviewing network effects in the context of Windows vs Mac operating systems. For the simple game presented, there are two equilibria (all consumers buy Windows and all consumers buy Mac) and the emphasis is on consumers' expectations being self-fulfilled. If all consumers believe Windows will dominate then Windows will indeed dominate, and if instead all consumers believe Mac will dominate then Mac will indeed dominate. It follows that companies selling products with network effects will try to figure out ways to influence consumers' beliefs. What is crucial, however, is not simply convincing a consumer that one's product is great but also that other consumers believe one's product is great. Thus, a company wants to advertise when a consumer watching the ad knows that many other consumers are also watching it. At that point, the 1984 Apple commercial is shown which was aired during the Super Bowl – the ultimate event during which people know other people are watching. Delivering such a visual exclamation mark to the lecture may help students retain the insight that you've just delivered.

It really goes without saying (but, of course, I'm now going to say it) that in-class experiments are a valuable multidimensional tool for teaching game theory. First, experiments can be used to document a phenomenon which will then be examined. For example, have students play the ultimatum game and then summarize and present the experimental output. After reviewing the standard theory (which is sure to run counter to the more egalitarian offers made and the rejection of proposals observed in the experiments) the stage is set for an in-class discussion as to why the theory performs so poorly. Second, experiments can produce introspection among students which is conducive to them learning about strategic reasoning and how to solve a particular game; having to make decisions themselves, they are likely to discover on their own what incentives are at work, to what extent players' interests coincide or conflict and other issues relevant to behavior. An example is the repeated prisoners' dilemma; students think about signaling a desire to cooperate, recognize the temptation to cheat and witness the role of punishments. When the instructor moves on to discuss these issues, students have already begun to think about them.

Running experiments faces two constraints: money and time. The former is easy to solve as students' performance in experiments can be a (small) part of their final grade, which obviates the necessity of monetary payments. The time constraint is more challenging, especially when the class is large. Fortunately, there are experiments that can be conducted even for a class exceeding a hundred students, and some of them are collected in the Instructor's Guide to my textbook. (The payoffs have also been calibrated so that the payoff from each experiment can be given equal weight in determining a student's grade.)

Another teaching tool is to require a capstone project. Quoting from my syllabus:

> For the project, you are to use game theory to model and make predictive statements about the behavior of people for either a real-world, historical, or fictional situation. A real-world situation is one that routinely occurs in human or non-human society. A fictional situation may be drawn from a story, poem, play, television show, movie, or computer software program but it is not to be a product of your imagination. Your imagination may be used to model a situation but not

> in creating the situation. Most critically, the situation cannot be one that we have gone over in class. The project is meant to be original work and will be graded on: i) how creative, sophisticated, and accurate is your model; and ii) how compelling, insightful, and correct is your analysis.

In contrast to lectures, readings, problem sets, and exams, where students are given a situation to consider and a model to analyze, now they must discover the situation and develop a model for the purpose of analysis. This is a chance for students to use the framework developed in class to explore what interests them.

Opening and Closing Acts

In concluding, let me offer a few tips on the first and last classes of the semester. Game theory is a unique subject, both in terms of its content and its potential for intellectual fun (not that the Gauss-Markov theorem isn't scintillating) , and this ought to be made exceedingly clear by the end of the first class. For this purpose, I recruit two students to play the Centipede game in front of the class. With 20 one dollar bills for all to see, a dollar is placed on the table and one of the volunteers is given the chance to take it. The right to either take or leave the money alternates between them and, each time the pile of money remains, a dollar is added to it. When eventually a student grabs the money, he or she walks away with the cash. (The singularity of the class has already been conveyed since instructors don't usually dole out money.) The lecture then turns to talking about strategic reasoning by mentally walking through the decision-making process faced by these volunteers. Students now know what strategic reasoning is and what the course is about. Game theory is then presented as a mathematical tool for exploring strategic reasoning.

The second phase of the first class is to review some of the strategic situations to be analyzed over the course of the semester. My intent is to get them excited and realize the course is truly interdisciplinary; economic settings are just one of many venues to be explored. A clip may be shown from Raiders of the Lost Ark where Indiana Jones must decide whether to throw the idol to Satipo on the hope that Satipo will throw him the whip; and I mention that, later in the semester, we'll see where Indy went wrong. A variety of other situations are briefly mentioned including the medieval law merchant, reciprocal altruism among vampire bats, and the tragic episode in 1964 in which Kitty Genovese was murdered. In each case, a puzzle is posed to be solved later in the semester.

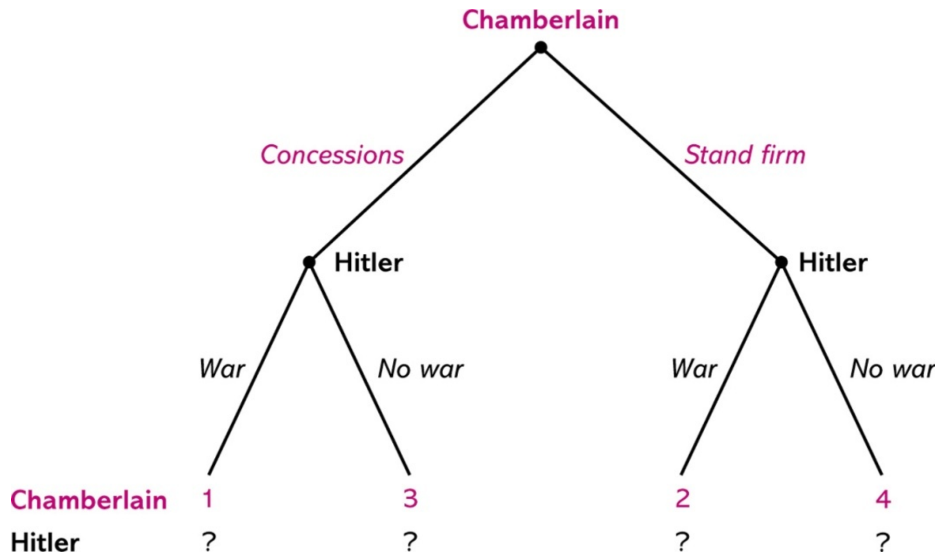
The first class concludes by looking at one strategic situation in some detail: the card game Concentration. This situation is unusual in that game theory can deliver a concrete rule for playing the game. Showing how this rule leads to some unanticipated prescriptions makes the case that game theory can generate fresh insight. But the opportunity is also taken to point out what we can and cannot expect of game theory. The rare simplicity of this setting allows game theory to say exactly what someone should do, but this is presented as an exception, not the rule. Generally, you cannot turn the crank of game theory and out pops a recommended course of action. What game theory *can* deliver is qualitative insight which will help to understand behavior and, from a prescriptive perspective, identify factors and relationships relevant to making an intelligent decision. It is important that students have realistic expectations about game theory before the process of learning is initiated.

My college has a peculiar practice – known as senior option – which allows for a dramatic ending to the semester. If an instructor chooses to offer senior option it means that graduating seniors can skip the final examination and have their final grade based on the remaining class work. Students are told at the first class that whether there is senior option will be determined at the end of the semester. The final class begins by describing a model of the strategic situation faced by the instructor and seniors which, without getting into details, has as its solution the instructor choosing not to offer senior option.The floor is then opened to possible modifications in the model as well as alternative solutions, while maintaining the same set of assumptions on payoffs. Once seniors realize that what happens at that class will influence whether or not there is senior option, they are powerfully energized. The discussion is animated as students, and not only seniors, spew forth clever ideas and argue on their behalf. If you don't have senior option, try to create a situation in which students have something at stake, so as to give them the opportunity and incentive to use all that they've learned in class. As an instructor, it is a most satisfying finality. Students are using game theory in a thoughtful and intelligent manner to solve a real problem. What more could one want?

References

Harrington, Joseph E., Jr., Games, Strategies and Decision Making, Worth Publishers, 2009.

Figure 1



Source: Figure 10.1 from Joseph E. Harrington, Jr., Games, Strategies and Decision Making (Worth Publishers, 2009).