The Prisoner's Dilemma . . . or Another Game

Adam Brandenburger*

Version 12/21/15

1 Introduction

Two-by-two games are the first encounter many people have with the language of game theory. One of the most famous two-by-two games is the Prisoner's Dilemma. In these notes, we will look at a way to sort all possible two-by-two games into a compact **classification**. We will then use our classification scheme in an application.

2 Symmetric Two-By-Two Games

We begin with a particular subworld of the world of all two-by-two games — namely, the subworld consisting of games of the form depicted in Figure 1:

Ann \ Bob	Y	X
Y	a, a	b, c
X	c, b	d, d

Figure 1

Following game-theory convention, we call the two players Ann and Bob. Each player has two possible strategies, labeled X and Y. The game is **symmetric** in the following sense. If both players choose the same strategy, then they get the same payoff: If they both choose X, they both get a payoff of d; if they both choose Y, they both get a payoff of a. If Ann chooses Y and Bob chooses X, then she gets a payoff of b and he gets a payoff of c. Symmetrically, if Ann chooses X and Bob chooses Y, then she gets a payoff of c and he gets a payoff of b. We will assume that $a > d$. (If $d > a$, we can just flip the matrix around.)

*Stern School of Business, Tandon School of Engineering, NYU Shanghai, New York University, New York, NY 10012, U.S.A., adam.brandenburger@stern.nyu.edu, adambrandenburger.com. This note is based on joint work with Ken Corts. Lawrence White provided important input.

Every different choice of the four numbers a, b, c, and d defines a different matrix. But, to get a useful classification scheme, we need to group together different matrices which we deem to be **strategically equivalent.** The approach we will adopt will be to group together matrices which may differ in the exact values that a, b, c, or d take, but where certain rankings among these four numbers are the same. One ranking we consider is between a and c. If $a > c$, then if Bob plays Y, Ann will prefer to play Y over X. This preference is indicated by the upward-pointing arrow in Figure 2:

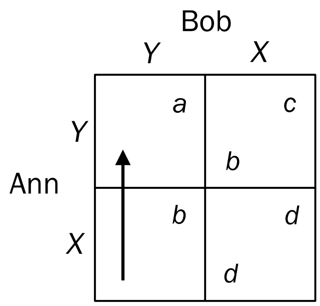

Figure 2

If $a < c$, we would draw the arrow pointing downwards. (We could add the case $a = c$ by extending our classification scheme, but, to keep things simple, we will ignore this case and other cases of equality for now.) The other ranking we consider is between b and d. If $b > d$, then if Bob plays X, Ann will prefer to play Y over X. This preference is indicated by the upward-pointing arrow in Figure 3 (the arrow would be downward pointing if $b < d$):

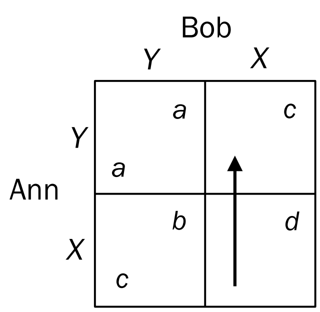

Figure 3

We can now describe our **classification scheme**: We will view two matrices as strategically equivalent if they yield the same configuration of arrows for both players, even if the two matrices differ in the values that a, b, c, or d take. That is, we will put any two such matrices into a single group in our classification scheme.

Exercises:

- Into how many different groups does this scheme sort the world of symmetric two-by-two matrices?
- Where do some of the familiar two-by-two games — such as the Prisoner's Dilemma and the Battle-of-the-Sexes — appear in this scheme?

It is important to bear in mind that there is no one 'correct' classification scheme. Different schemes are possible depending on what properties of a game matrix we wish to distinguish. The test of a particular scheme is its value in application. (Just to make this point clear, consider two extreme classification schemes. One distinguishes any two matrices that differ in any of the four numbers a, b, c, or d. Another extreme classification scheme groups together all matrices regardless of what values these numbers take. The first scheme contains infinitely many groups, while the second comprises just one group. Useable schemes most likely contain a small finite number of groups.)

3 Asymmetric Two-By-Two Games

So far, we have looked only at two-by-two games that are symmetric in the way we defined. The general — asymmetric — two-by-two game is defined by 8 rather than 4 numbers, as in Figure 4:

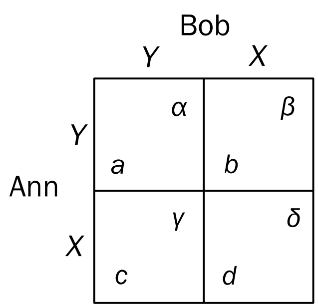

Figure 4

Exercises:

- Using our arrow-based classification scheme, how many additional groups do we need to create to include all asymmetric two-by-two games in the scheme?
- Where do some of the familiar asymmetric two-by-two games — such as Matching Pennies — appear in this enlarged scheme?

4 An Application

The Prisoner's Dilemma game was created by two mathematicians, Merrill Flood and Melvin Dresher, in 1950. They were working at the RAND Corporation, a prominent think tank set up immediately after World War II to assemble intellectual talent with which to formulate U.S. strategy in the new age of atomic weapons. The game became highly influential in making a nuclear arms race between the U.S. and the U.S.S.R. seem more or less inevitable. The influence of the game spread even further, and it became a canonical model of interaction in economics, political science, evolutionary biology, and elsewhere.[1]

The reaction of the mathematicians themselves to the Prisoner's Dilemma was to treat it as something of a challenge to the direction in which game theory was moving at the time, which was away from von Neumann's original theory that allowed players to act in a way that maximizes their joint interest, to Nash's new theory that allowed only individually motivated choices by players.[2] The mathematicians saw the Prisoner's Dilemma as problematic for the new theory because in experiments they ran at RAND, they found a level of cooperation by players which was hard to reconcile with Nash's theory.[3]

While the mathematicians debated what the Prisoner's Dilemma meant for game theory, physicists had been debating, even before the game itself was invented, the right way to think about relations between countries in the atomic age. One of these physicists was the famous Niels Bohr, whose 1913 work on the structure of the atom had been central to the creation of quantum mechanics.[4] Later, Bohr proposed what he called the principle of complementarity — that whether a subatomic object such as an electron should be viewed as a particle or a wave depends on how the experimenter interacts with it — as the fundamental feature of the subatomic world.[5] By 1944, as the U.S. project to make an atomic bomb (called the Manhattan Project) was making big strides, Bohr was thinking about the implications of this new weapon:

> By this time he [Bohr] had thought very seriously about the postwar situation and had had what he regarded as a revelation about the "complementarity" of atomic bombs — a revelation as important, he believed, as his earlier epiphany regarding the complementarity of subatomic particles. Just as electrons are at one and the same time waves and particles, so, Bohr now believed, atomic bombs were at one and the same time the greatest danger to mankind and the greatest boon. Atomic bombs could put an end to civilization and human life itself, or, precisely because of that, they could bring an end to war. What was needed, Bohr felt, was a spirit of cooperation and, above all, *openness*. If the power of atomic bombs was made clear to everybody, Bohr reasoned, there would be at least the possibility of cooperation and therefore the possibility that this terrible weapon could turn out, because of its very terribleness, to be the best thing mankind had ever invented.[6]

In September 1949, the U.S. learned that the U.S.S.R. had exploded its first atomic bomb. A big debate immediately ensued in U.S. circles as to the best response, with one group pushing for the U.S. to regain its superior position by developing the much more powerful hydrogen bomb.[7] Another group argued against this course. One member of the second group was the leading political strategist George Kennan,[8] who argued that:

> [T]he Soviet Union, given its economy and industry were still in ruins after the devastation of the war, might not want to embark on an expensive arms race and might be willing to negotiate an agreement that ensured that neither side developed the hydrogen bomb.[9]

Also on this side of the debate was the physicist J. Robert Oppenheimer, who had headed the Manhattan Project that had produced the world's first atomic bomb back in 1945:

> [Oppenheimer] said that he did not think the United States should build the hydrogen bomb, and the main reason he gave for this ... was that if we built a hydrogen bomb, the

Russians would build a hydrogen bomb, whereas if we did not build a hydrogen bomb, then the Russians would not build a hydrogen bomb.[10]

Exercises:

- Find an example where someone has described a real-world situation in terms of a two-by-two matrix, and you agree with the choice of two-by-two matrix used in the description.
- Find an example where someone has described a real-world situation in terms of a two-by-two matrix, and you disagree with the choice of two-by-two matrix used in the description.

Notes

[1]Erickson, P., J. Klein, L. Daston, R. Lemov, T. Sturm, and M. Gordin, *How Reason Almost Lost Its Mind: The Strange Career of Cold War Rationality*, The Chicago of University Press, 2013, pp.133-136.

[2]Op.cit. p.141. Nash, J., “Non-cooperative Games,” *Annals of Mathematics*, 54, 951, 286-295.

[3]Flood, M., “Some Experimental Games,” *Management Science*, 5, 1958, 5-26.

[4]Bohr, N, “On the Constitution of Atoms and Molecules, Parts I-III,” *Philosophical Magazine*, 26 1913, 1-24, 476-502, 857-875.

[5]Bohr, N.,“Discussions with Einstein on Epistemological Problems in Atomic Physics,” In Schilpp, P. *Albert Einstein: Philosopher-Scientist*, Open Court, 3rd edition, 1998.

[6]Monk, R., *Robert Oppenheimer: A Life Inside the Center*, Anchor Books, 2012, p.414.

[7]Op.cit. p.566.

[8]George Kennan was most famous for his 1946 “Long Telegram” which laid the foundation for the U.S. Cold War strategy of containment towards the U.S.S.R.. See, e.g., Gaddis, J.L., *Strategies of Containment: A Critical Appraisal of American National Security Policy during the Cold War*, Oxford University Press, 2005.

[9]Monk, p.569.

[10]Monk, p.570.