THE THEORY OF GAMES

By

Richard Bellman

P-1062

April 15, 1957

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

SUMMARY

The purpose of this paper is to present an expository account of the fundamental ideas of the theory of games, together with a discussion of some of the unresolved aspects of the theory.

CONTENTS

THE THEORY OF GAMES

By

Richard Bellman

§1. Introduction.

In recent years, a new branch of mathematical analysis has been developed and intensively studied. It possesses a great deal of intrinsic interest and a number of close ties with other parts of mathematics and various cognate fields such as mathematical economics and mathematical statistics. In this article we propose to give a brief sketch of some of the basic ideas of what is now called the Theory of Games.

Although a systematic foundation of the theory was begun by Borel, [8], independently von Neumann [22] presented his own formulation, and derived the basic result which is the cornerstone of the theory. This result, whose proof had eluded Borel, is called the Min-Max theorem, and will be discussed extensively below.

One of the attractive features of the theory of games, shared by some other fields of mathematics such as number theory and topology, is that it is not possible to gauge the depth of problems which can be formulated in terms of quite simple ideas, and in very few words. The result is that some fairly simply stated questions lead in some cases to quite recondite analysis, and in other cases to the very boundary of the unknown.

In the course of the chapter we shall refer to a number of books and research papers which explore in detail a number of the subjects we mention quite briefly, and which contain a great deal more besides.

The most interesting and entertaining account of the fundamental concepts, valuable for amateur and professional alike, is that contained in the book by Williams, [26].

§2. Matching Coins.

In order to illustrate the type of problems encountered in the theory of games and the concepts used to analyze these problems, let us begin our discussion with the perennial diversion of matching coins.

As we know, the game proceeds in the following fashion. At each stage of the game, each player chooses to show a head or a tail, with his choice unknown to the other player. It is agreed upon in advance that one player wins if the coins match, and that the other player wins if they do not.

Suppose that one player is suddenly inspired to analyze the game mathematically in the hope of gaining an advantage in this way.

How does one analyze a process of this type mathematically? This question is asked deliberately, and not rhetorically, to emphasize first of all the fact that before a satisfactory method of analysis is discovered, it is often not at all clear how to proceed, and further to emphasize the related fact that even after one good method has been found, it may still be true that another still better method is required to treat questions of a higher level of difficulty. No theory then should be regarded as either inevitable in its formulation, or final in its conclusions.

Returning to coin-matching, it is clear that if we reject extrasensory perception and telekinesis as being at best unproved and undependable, there is no way in which either player can gain an advantage playing the game once against an unknown opponent. The way out of this cul-de-sac, which would seem to block any mathematical study, is to focus our attention upon games which are played a large number of times, and concentrate upon gaining an advantage in the long run. In other words, each player is to play in such a way as to maximize an <u>average</u> return.

This, of course, is one of the guiding principles of gaming, either at the card table, on the football field, or on a baseball diamond. It sustains a bridge master when he makes an unfortunate finesse; it sustains the poker expert who calls a straight with aces over fours; it sustains the baseball manager who calls for a hit-and-run which ends in a double play. The players who disregard the averaging process, and rely instead upon hunches, make a number of spectacular gains, which are duly advertised by the players themselves, and sometimes the newspapers. In the long run, they lose consistently, and consistently blame their misfortune upon bad luck, inferior partners, poor teamwork, etc., etc., etc.

Let us agree then that the analytic player decides to play in such a way as to maximize the average amount won at any particular play of the game, which is equivalent to the assertion that he intends to play the game a large number of times and maximize the average amount won in the course of this large number of plays, or acts as if he intended to follow this pattern. We shall discuss below some of the complications connected with this point of view.

Consider the situation of this player, taken to be the one who wins if the coins match. He can argue as follows:

"The principle of insufficient reason assures me that the other player will be displaying a head or a tail with equal probability, since there seems to be no reason why he should show one or the other. Consequently, no matter what I do, I have equal probability of winning or losing. I might just as well show heads all the time."

Then the following disturbing thought occurs:

"Suppose that the other player is irrational, or suppose that I meet a string of coin-matchers who on the basis of philosophical principles, or as a result of election bets, have pledged themselves to show tails all the time. I will then lose on every play. How can I guard against this contingency?"

A small amount of reflection shows that the player can guard against situations of this type by showing heads or tails with equal probability in a random fashion. In this way, no matter what group of opponents he encounters, on the average his gain or loss will be zero.

This is not a very encouraging result as far as indicating the utility of mathematical analysis, but let us persevere. Something may come of it.

The analysis above was highly plausible and completely intuitive. How can we obtain these results in a systematic fashion which puts less strain upon our intelligence and more upon our mathematics? The advantage in developing a systematic approach lies in the possibility that this same approach may possibly be useful in connection with other processes of less trivial nature.

Let us now suppose that we have two players guided by the same principle, that of maximizing average return. Let the first player play heads with probability a and tails with probability 1-a, and let the second player show heads with probability b, and tails with probability 1-b.

The probability that the first player wins is then

$$(1) \qquad ab + (1-a)(1-b),$$

the total probability that the two coins match, while the probability that the other player wins is given by

$$(2) \qquad a(1-b) + b(1-a).$$

Let us agree to credit the first player with +1 for a win and -1 for a loss, so that the average gain per play to the first player will be

$$(3) \qquad E(a,b) = ab + (1-a)(1-b) - [a(1-b) + b(1-a)]$$

and the average gain per play to the second player will be the negative of this.

The first player will then choose a so as to maximize E(a,b), and the second player chooses b so as to minimize. The least he can receive is then

$$(4) \qquad E = \min_{0 \le b \le 1} \max_{0 \le a \le 1} E(a,b).$$

Let us now perform some elementary calculations. We have

(5) $$E(a,b) = (2a-1)(2b-1)$$

so that

(6) $$\max_{0 \le a \le 1} E(a,b) = \max(2b-1, 1-2b).$$

Hence the minimum over b of this expression occurs where $b = 1/2$. Thus

(7) $$\min_{0 \le b \le 1} \max_{0 \le a \le 1} E(a,b) = 0.$$

Now consider the situation of the second player. He proceeds to choose b so as to minimize $E(a,b)$, yielding the first player an average return of

(8) $$\min_{0 \le b \le 1} E(a,b) = \min(2a-1, 1-2a).$$

It follows that the first player to maximize his return must choose $a = 1/2$.

From these calculations, we see that each player, if he does not know what the other person is doing, is forced to protect himself against loss by using the equal-probability policy. This guarantees that on the average he will win as often as he loses.

Observe something fascinating about this situation. Suppose that the first player announces openly that he is using this equal-probability policy. Then the other player cannot improve his average return even with this additional information at his disposal, and vice versa.

If one player can induce the other player to depart from a safe policy by using over a number of plays a policy which appears to be of one type, but which is actually of another, then he can gain. This is, however, a risky maneuver.* The mathematical theory of these more complicated processes is part of the study of "learning processes", of which one aspect is the theory of sequential analysis developed by Wald, [25]. For other aspects of the theory of learning processes, see Johnson and Karlin, [12], Bellman, Harris and Shapiro, [7], Bellman, [1], and the book by Bush and Mosteller, [10].

§3. Unsymmetric Coin-Matching.

Suppose we now consider a coin matching game of the following type. If two heads occur, the first player receives 3 units; if two tails, he receives 1 unit; if head-tails, or tail-heads, he gives the other player 2 units.

Faced with an invitation to play this game, the first player must decide whether he wants to play or not. Let us see the type of analysis he might employ. As above, we make the assumption that average return is to be the criterion for both players. Let a be the probability with which the first player displays heads, and b the probability with which the second player displays heads. Then the expected return to the first player is

$$(1) \qquad E(a,b) = 3ab + (1-a)(1-b) - a(1-b)-b(1-a).$$

Again, as above, the first player wishes to maximize this function over all values of a in $[0,1]$ and the second player wishes to minimize it over b in the same interval. Depending upon which

*But one which has consistently been used in warfare.

point of view one takes, we see that we have to determine the two quantities

$$(2) \qquad \min_{0\le b\le 1} \max_{0\le a\le 1} E(a,b),$$

and

$$(3) \qquad \max_{0\le a\le 1} \min_{0\le b\le 1} E(a,b).$$

Carrying out the computation, as above, we see that the two quantities are rather surprisingly equal, and furthermore that the common value is attained for the following probabilities:

$$(4) \qquad a = 5/8, \quad b = 3/8.$$

It follows that each player can announce behavior without giving any advantage to the other.

§4. Saddlepoints.

The substance of the above result is that the function $E(a,b)$ possesses a saddlepoint over the square defined by $0 \le a, b \le 1$. This is to say that there is a point $[a',b']$ with the property that

$$(1) \qquad \begin{aligned} E(a',b') &\le E(a',b), \quad 0 \le b \le 1, \\ E(a',b') &\ge E(a,b'), \quad 0 \le a \le 1. \end{aligned}$$

In more complicated games, optimal strategies need not be unique, as they have been in these two cases considered above.

§5. A General Two-Choice Game.

Let us now consider a general two-choice game, where the two choices may be thought of as heads or tails. Let us define

(1) a_{ij} = the return, or pay-off, to the first player if he makes the i-th choice and the second player makes the j-th choice, i,j, = 1,2.

The matrix

$$(2) \qquad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

is called the payoff matrix. The pay-off matrix for the second player will be the negative of this. This type of game is called zero-sum, and is the only type for which there is a satisfactory theory at the present time.

At each play of the game, let the first player make the first choice with probability x_1 and the second choice with probability x_2. A set of values (x_1, x_2) with $x_1 + x_2 = 1$, $x_1, x_2 \geq 0$, is called a strategy. Similarly, let the second player make his first choice with probability y_1 and his second choice with probability y_2. Then the expected return to the first player is

$$(3) \qquad E(x,y) = a_{11}x_1y_1 + a_{12}x_1y_2 + a_{21}x_2y_1 + a_{22}x_2y_2.$$

We leave it as a simple exercise in algebra, or analytic geometry, for the reader to prove that

$$(4) \qquad \max_x \min_y E(x,y) = \min_y \max_x E(x,y),$$

where the maximum is taken over the region $x_1 + x_2 = 1$, $x_1, x_2 \geq 0$, $y_1 + y_2 = 1$, $y_1, y_2 \geq 0$. The common value of these two expressions is called the <u>value</u> of the game, and occasionally denoted by $v(A)$.

As above, it follows that the first player possesses a strategy which guarantees him an expected return of at least v, while the second player possesses a strategy which guarantees that he will not sustain a loss of more than v. If v is negative, we can, if we wish, interchange the terms "return" and "loss".

These strategies need not be unique.

§6. <u>The General Finite Game</u>.

Having introduced the above notation, it is now easy to continue to a discussion of a more general situation which each player possesses a finite number of choices.

Assume that the first player may make one of M choices and that the second player may make any of N choices. Let the pay-off matrix be, as above

$$(1) \qquad A = (a_{ij}),$$

where a_{ij} is the return to the first player if he makes the i-th choice and the second player makes the j-th choice. The negative of this will then be the return to the second player.

If the first player employs a strategy $x = (x_1, x_2, \ldots, x_M)$ and the second player employs the strategy $y = (y_1, y_2, \ldots, y_N)$, the expected return to the first player will be

$$(2) \qquad E_1(x,y) = \sum_{i,j=1}^{M,N} a_{ij} x_i y_j .$$

The fundamental result in the theory of games is

<u>Theorem (Min-Max Theorem of von Neumann). The function $E_1(x,y)$ possesses a saddle-point over the region defined by</u>

(4) a. $x_i, y_i \geq 0$

b. $\sum_{i=1}^{M} x_i = 1, \quad \sum_{i=1}^{N} y_i = 1.$

<u>Hence</u>

(5) $$\min_y \max_x E_1(x,y) = \max_x \min_y E_1(x,y) = v.$$

<u>Consequently, the first player possesses at least one strategy which guarantees him an expected return of at least</u> v, <u>regardless of what the second player does and the second player possesses likewise at least one strategy which guarantees that his expected loss is not more than</u> v, <u>regardless of what the first player does</u>.

There are no simple proofs of this result although there are elementary ones. The shortest proofs require fixed-point theorems borrowed from topology.

This result was established by Von Neumann in 1928, in the general case, while the particular cases N = 2,3,4 were considered by Borel. Unfortunately, Borel at first believed that the general case was not true. For an interesting discussion of questions of priority see the article by Frechet, [11] and the rebuttal by Von Neumann, [23].

§7. Computational Algorithms.

The determination of the value of a game associated with a matrix of even moderate size, say ten-by-ten, is not an easy task.

No explicit analytic representation of the value exists, nor does there exist any analytic representation of the set of optimal strategies. Furthermore, even if these analytic representations did exist there is no guarantee that they would actually be useful for computational purposes. Consider, as an illustration of this, the simpler problem of solving a system of ten-by-ten linear equations. The explicit solution by means of Cramer's rule can be used effectively in only very rare circumstances to determine the numerical solution.

Consequently, the problem of computing the value of a game resolves itself into a hunt for effective numerical algorithms. One of the most important algorithms exploits the connection between a multi-stage game and the original game, viewed as the steady state version of the dynamic game. This procedure was inaugurated by Brown and von Neumann, [9], and its validity was established by J. Robinson [17]. A continuous version of this process was also considered by von Neumann, [9], cf. also Bellman, [2] for a generalization. There are a number of other techniques based upon the connection between the theory of games and linear programming, cf. [20].

In any particular case, a great deal can usually be done by the use of dominance arguments, which greatly simplify the search for a solution by eliminating certain feasible, but obviously inefficient, strategies at the outset.

§8. Continuous Games.

In the preceding sections, we have assumed that each player had a finite number of choices. Let us now consider a more general situation where each player has a continuum of choices.

Suppose that each player is to choose a number in the interval [0,1]. If the first player chooses x and the second player chooses y, the payoff to the first player is determined by the function K(x,y), with the negative of this the return to the second player.

In order to mix choices, each player chooses a distribution function, dG(x) for the first player and dH(y) for the second player. The expected return for the first player is then given by

$$(1) \qquad \int_0^1 \int_0^1 K(x,y)dG(x)dH(y)$$

The analogue of the fundamental result of von Neumann is the result, due to Ville, [21], that

$$(2) \qquad \underset{G}{\text{Max}}\ \underset{H}{\text{Min}} \int_0^1 \int_0^1 K(x,y)dG(x)dH(y) = \underset{H}{\text{Min}}\ \underset{G}{\text{Max}} \int_0^1 \int_0^1 K(x,y)dG(x)dH(y$$

provided that K(x,y) is continuous over the square $0 \leq x,y \leq 1$.

One of the reasons for considering continuous games lies in the fact that the solution may be considerably simpler to obtain than in the discrete case. Here, the continuous case is to be considered as an approximation to the discrete case.

An interesting study of games of timing, arising from the study of duels, is contained in Shiffman, [19]. A discussion of continuous poker games is contained in the references cited in §11.

§9. Non-Zero Sum Games.

The analysis of the foregoing sections applied to games in which the players were in direct opposition, in the sense that a gain for one player meant a loss for the other. In a large number of applications of game theory, say to economic situations or to military situations, this is not the case.

Let us consider two simple ways in which we can be forced to study non-zero sum games. Returning to a discussion of the two-choice game, assume that the payoff matrix for the first player is as before

$$(1) \qquad A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

but that the payoff matrix for the second player is not the negative of this, but rather the matrix

$$(2) \qquad B = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$$

This means that the players measure the outcomes of decisions in different ways. This is the usual situation.

The first player as before wishes to play so as to maximize the quantity $ab + (1-a)(1-b) - (a(1-b)) - b(1-a)$, but the second player wishes to play so as to maximize the quantity $-2ab - 2(1-a)(1-b) + a(1-b) + b(1-a)$. In the case where $B = -A$, these two aims were in direct opposition, so that we could combine these two types of play into one min-max situation. This we can no longer do.

What then do we do? The answer is that nobody knows for sure. There are a number of tentative proposals, cf. von Neumann-Morgenstern, [24], Nash, [15], which yield interesting and informative results in some cases, but there is no uniform satisfying theory corresponding to the zero-sum case. An interesting discussion is contained in McKinsey, [3].

§10 Different Criteria.

We encountered difficulties in the previous section because we allowed different pay-off matrices for the players. Suppose we insist that $B = -A$, but now assume that the two players have different theories as to how one should proceed in processes of this type.

For example, one player may be perfectly willing to maximize his expected return, while the other player may wish to maximize the probability that he wins a certain amount, or being a conservative type, may wish to minimize the probability that he loses more than a certain quantity.

Using these different criteria, we are once again faced with a situation in which the two players are not in direct opposition, and there is the same lack of a definitive theory noted in the preceding section.

§11. N-Person Games — $N \geq 3$.

In applications, particularly of an economic nature, we encounter processes in which there are more than two players. Examples of this are furnished by bidding on industrial contracts, and by a number of games of social nature such as bridge and poker.

Analyzing the problem in a purely rational manner, it turns out that the obvious thing to do in some cases is for two of the players to form a coalition against the third, or in other cases for all three of the players to form a coalition against the consumer. However, the rules of the game may forbid this simple solution. How then to play?

Again, there are a number of tentative theories, cf. von Neumann-Morgenstern, [24], Nash, [15], and Shapley, [18], but all have a number of drawbacks. At the present time it appears as if there will never be any unitary theory of N-person games, but only a number of theories, each satisfactory within its domain, but incapable of being stretched to cover the entire region of interest.

§12. Poker.

As soon as one hears the term "theory of games", one is intrigued by the possibility of applying this theory to the treatment of such pastimes as poker. From what we have said above, it is not to be expected that much can be done in connection with actual poker, where there are six, seven or eight players, each with quite different utility functions. However, it is interesting to analyze some simple two-person games in the hopes of being able to understand some aspects of such characteristic features as bluffing, and, generally, as a matter of intellectual curiosity.

A considerable amount of material on poker may be found in the book by von Neumann-Morgenstern, [24]. It turns out, however, that it is considerably simpler to consider some continuous versions of poker, cf. Bellman-Blackwell, [6], and Bellman, [3].

An analysis of a three-person poker game using the equilibrium point theory of Nash may be found in Nash-Shapley, [16], where other references may be found.

As far as applying these results to actual play is concerned, let us state the general rule that the only way to play poker is to play according to the opponents and not according to some rigid preconceived theory.

§13. Games of Survival.

Another very interesting class of games are those which have been given the name "games of survival". These correspond to the classic "gambler's ruin" in which two players sit down and play until one or the other of the players has all the money in the game. What distinguishes this type of game from those considered above, is not so much the multi-stage aspect, but the fact that there is a correlation between stages due to the fact that the choices available to each player at each stage depends upon the amount of money he has at this stage.

For a discussion of multi-stage games and games of survival, we refer the reader to Bellman, [4], Milnor and Shapley, [14]. The functional equation technique of dynamic programming, [5], is useful in the discussion of these processes.

REFERENCES

1. Bellman, R., A Problem in the Sequential Design of Experiments, Sankhya, Vol. 16, 1956, pp. 221-229.

2. ___________, On a New Iterative Algorithm for Finding the Solutions of Games and Linear Programming Problems, RAND Paper P-473, 1953.

3. ___________, On Games Involving Bluffing, Rendiconti del Circolo Matematico di Palermo, Series 2, Vol. 1, 1952.

4. ___________, The Theory of Dynamic Programming, Bull. Amer. Math. Soc., Vol. 60, 1954, pp. 503-516.

5. ___________, Functional Equations in the Theory of Dynamic Programming—III, Multi-Stage Games, Rendiconti del Circolo Matematico di Palermo, 1957.

6. Bellman, R., and D. Blackwell, Some Two-Person Games Involving Bluffing, Proc. Nat. Acad. Sci., Vol. 35, 1949, pp. 600-605.

7. Bellman, R., T. E. Harris and H. N. Shapiro, Some Functional Equations Occurring in the Theory of Decision Processes, RAND Paper P-382.

8. Borel, E., Applications auz jeux de hasard, Traite du Calcul des Probabilities et de ses Applications, Gauthier-Villars, 4f. II 1938, Paris.

9. Brown, G. W. and J. von Neumann, Solutions of Games by Differential Equations, Contributions to the Theory of Games, Annals of Mathematics Studies 24, Princeton University Press, 1950. see also,

Brown, G. W., Iterative Solutions of Games by fictitious Play, Cowles Commission Monograph, Activity Analysis of Production and Allocation, John Wiley and Sons, 1953.

10. Bush, R. R., and F. Mosteller, Stochastic Models for Learning, J. Wiley and Sons, 1955, New York.

11. Frechet, M., Econometrica, Vol. 21, 1953, pp. 97-127.

12. Johnson, S. and S. Karlin, A Mathematical Treatment of Learning Models, RAND Research Memorandum RM-921, 1954.

13. McKinsey, J. C. C., Some Notions and Problems in Game Theory. Bull. Amer. Math. Soc., Vol. 58, 1952, pp. 519-532.

14. Milnor, J. and L. Shapley, Games of Survival, Contributions to the Theory of Games, 1957 (to appear).

15. Nash, J., Equilibrium Points in N-Person Games, Proc. Nat. Acad. Sci., Vol. 36, 1950, pp. 48-9.

16. Nash, J. and L. Shapley, A Simple Three-Person Poker Game, Contributions to the Theory of Games, Annals of Mathematics Studies 24, Princeton University Press, 1950, pp. 105-116.

17. Robinson, J., An Iterative Method of Solving a Game, Annals of Mathematics, Vol. 54, 1951, pp. 269-301.

18. Shapley, L. A Value for N-Person Games, Annals of Mathematics Studies 28, Princeton University Press, 1953.

19. Shiffman, M., Games of Timing, Contributions to the Theory of Games, Annals of Mathematics Studies 24, Princeton University Press, 1950.

20. Tucker, A. W. and D. Gale and H. Kuhn, Linear Programming and the Theory of Games, Activity Analysis of Production and Allocation, John Wiley and Sons, 1953.

21. Ville, J., Sur la theorie generale des jeux ou intervient l'habilite des joueurs, Traite du Calcul des Probabilitis et de ses Applications, E. Borel and collaborators, Paris, 1938, Vol. 2, No. 5, pp. 105-113.

22. von Neumann, J., Zur Theorie der Gesellshaftsspiele, Math. Annalen, Vol. 100, 1928, pp. 295-320.

23. ________________, Econometrica, Vol. 21, 1953, pp. 97-127.

24. von Neumann, J. and O. Morgenstern, The Theory of Games and Economic Behavior, Princeton University Press, 1944.

25. Wald, A., Statistical Decision Functions, John Wiley and Sons, New York, 1950.

26. Williams, J. D., The Compleat Strategyst, McGraw-Hill, New York, 1954.