U. S. AIR FORCE

PROJECT RAND

RESEARCH MEMORANDUM

ON NON-ZERO-SUM GAMES AND STOCHASTIC PROCESSES

Richard Bellman and J. LaSalle

RM-212

19 August 1949

Assigned to ______________________

This is a working paper. It may be expanded, modified, or withdrawn at any time. The views, conclusions, and recommendations expressed herein do not necessarily reflect the official views or policies of the United States Air Force.

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

ON NON-ZERO-SUM GAMES AND STOCHASTIC PROCESSES

Richard Bellman and Joseph LaSalle

Summary

The major objective of the report is to introduce the reader to a class of stochastic games which we believe possesses a factor (of realism) that has been neglected. These stochastic games arose in an attempt to find an approach to non-zero-sum games. An explanation of this will be found in the introduction. The remainder of the report presents some examples which illustrate avenues of investigation, raise new problems, indicate general results, and demonstrates even for the case of a zero sum how unrealistic it might be to neglect the resources of the players.

§1. Introduction.

The theory of non-zero-sum games is at present little developed. This is not for want of interest, nor for lack of sufficient importance; rather it seems difficult to pose the problem mathematically.

After wondering why it is that a comprehensive theory of non-zero-sum games is difficult, if not impossible, to obtain, we decided to examine a few simple non-zero-sum games. Our hope was that the examples would indicate the nature of the difficulties, that they might suggest problems of interest, and possibly yield general procedures.

Intuitively, we felt the importance of considering the external factors which cause the players to participate in a non-zero-sum game and which determine their objectives.

Attempt to conceive of a situation which is not zero-sum. The valuations in the game may be stated as so many dollars, planes, etc. If it is a "price war" or a "battle of attrition" - the Battle of Britain - one readily sees that a proper evaluation of their preferences may not be those given by the data which reports the pay-offs and that these evaluations will depend upon the resources of the participants. Even if the data makes the game zero-sum, one may doubt that a dollar has the same value to all players.

In the usual treatment of games and strategy, an a priori assumption is that proper evaluations have been placed on the preferences of the players. We attempt to go further. Our mathematical model includes in the description of the game the objectives of the players. From this information one determines the proper evaluations and the best strategies. In particular, our model is adapted to the situation where the resources of the players is finite and is an important factor in determining the best strategies.

In a conflict between two opponents, there will be many engagements which when isolated are certainly not zero-sum. They can be understood only as part of the larger conflict; it is the larger conflict which supplies the motivation and objectives of the players. As an example of this, consider a non-zero-sum game to be immersed in a larger game. The larger game is made up of a series of plays of the non-zero-sum game. This can be represented as a two-dimensional, random walk with absorbing barriers.* Each play of the non-zero-sum

* Were each step zero-sum, the representation would be one-dimensional.

game is a step in the stochastic process. The game is concluded when the absorbing barrier is reached and the pay-off is a function of where the game ends.

§2. A simple non-zero-sum game.

We consider the following game based upon matching coins:

When the coins match, player II loses one and player I receives nothing. When I has heads and II tails, they both lose one. When I has tails and II heads, I loses one and II receives nothing.

The pay-offs to players I and II versus their (pure) strategies are:

I \ II	H	T
H	0	-1
T	-1	0

PAY-OFF TO I

I \ II	H	T
H	-1	-1
T	0	-1

PAY-OFF TO II

TABLE 1

In a zero-sum-game, players I and II would be "completely opposed" to each other; if I confines his attention to maximizing his pay-off, he has automatically minimized that of his opponent. Here this is not the case; II has no preference, as indicated by the pay-off, between (H, H) and (H, T), while I prefers (H, H).

If players I and II are in "complete opposition", the objection which can be raised immediately is that the above pay-offs as prescribed by the game cannot represent their preferences. Faced only with this information, it is doubtful that one can proceed without taking into account some external factors or without some assumptions concerning the objectives of the players.

Let us illustrate this further by introducing more information concerning the players. Suppose that I and II are the same person; the game is now a one-person game and the solution is trivial: play anything except (H, T).

A less trivial illustration is to assume that I and II judge a pay-off to their opponent of amount x is equivalent to receiving -x. The pay-offs would then be:

	H	T
H	1	0
T	-1	1

PAY-OFF TO I

	H	T
H	-1	0
T	1	-1

PAY-OFF TO II

TABLE 2

The game is now a zero-sum-game.

While we have this zero-sum-game before us, let us make good on the assertion that the "standard" solution may not at all be the one the players should use. This is a warning against writing down the matrix and presenting someone with the solution without investigating the true nature of the game.

For a single play of the game in Table 2, the best strategy would be:

$$\text{I plays } \frac{2}{3}\,H + \frac{1}{3}\,T$$

$$\text{II plays } \frac{1}{3}\,H + \frac{2}{3}\,T\ ;$$

the value of the game is $\frac{2}{3}$.

Suppose, however, that examining the game further one finds that I starts with an amount 1 and II with 2, and that they play until one player is ruined. Then it can be shown very easily that if (a) when I has 1, he plays H with probability $1 - \varepsilon$, (b) when he has 2, he plays H with probability $\frac{1}{2}$, then he is assured that his expectation will be greater than $\frac{2 - 4\varepsilon}{1 + \varepsilon}$; i.e. by making $\varepsilon > 0$ sufficiently small, he can assure that his expectation is as close to 2 as he pleases.

§ 3. A stochastic game.

Rather than making direct assumptions on the evaluations of their preferences, it is quite realistic and of greater interest to introduce external conditions which determine what these evaluations should be. Let us suppose that the non-zero-sum game originally proposed is but one play of a larger game. A simple example of this is the following:

I <u>begins the game with amount</u> x <u>and</u> II <u>with</u> y. <u>They play the non-zero-sum game</u> (<u>of</u> § 2) <u>until one of them is ruined</u>. <u>If</u> I <u>is ruined</u>

first, I loses b and II wins b; if II is ruined first, II loses a and I wins a; in case of a tie, neither wins nor loses.

This stochastic game is a zero-sum-game. The steps are not zero-sum. Let the point (x, y) represent I and II's resources; each play of the non-zero-sum game determines one of the steps (-1, 0), (0, -1), (-1, -1). The x and y-axes are the absorbing barriers which determine the end of the game.

§4. No information during the progress of the game.

As one can readily visualize, the players may know only the starting point (x, y) and not know the results of the individual steps of the game.

Suppose that it is only feasible for the players to select the probability of playing heads or tails (the same probabilities at each step). Let

ξ = probability that I plays H

η = probability that II plays H

$\phi(\xi, \eta)$ = pay-off to I

This is a continuous game with polynomial pay-off function; the best strategies for I will be to play a finite number of probabilities $\xi_1, \dots, \xi_n$ with probabilities $\alpha_1, \dots, \alpha_n$. II's best strategies have a similar form.*

* See R-115, Theorem 5.2. For a solution of a particular case of this game see RM-215.

§5. Information on the outcome of each step.

Here we assume that the opponents know the result of each step (the result of each play of the non-zero-sum game). In many respects, the game is now simpler to analyze. Let

$$v(x, y) = \text{value of the game to I if the game started at } (x, y).$$

To decide on his first play (a play of the non-zero-sum game), I would examine the pay-off matrix

	H	T
H	$v(x, y-1)$	$v(x-1, y-1)$
T	$v(x-1, y)$	$v(x, y-1)$

It can be shown by an induction argument that $v(x, y)$ is defined (a best strategy exists) and that

$$v(x-1, y) \leq v(x-1, y-1) \leq v(x, y-1) \text{ for all } (x, y).$$

Letting

$$\xi_1(x, y) = \text{probability that I plays heads when at } (x, y),$$

we then see that the best strategy for I is given by

$$\xi_1(x, y) = \frac{1}{1 + \frac{v(x, y-1) - v(x-1, y-1)}{v(x, y-1) - v(x-1, y)}} \geq \frac{1}{2}$$

and

$$v(x, y) = \xi_1(x, y)v(x, y-1) + \big(1-\xi_1(x, y)\big)v(x-1, y).$$

This gives rise to non-linear difference equations. A general solution looks to be difficult to obtain, though particular cases could be computed.

When should I or II participate in the game if $x \leq y$? If $x > y$, then I is always sure of winning. But how much larger than x does y have to be, in order that II have a positive expectation by participating in these "battles of attrition"? Also, comparing the differences between the values of the games in §5 and §4 would provide a model for the value of reconnaissance. It is also possible to demonstrate by this model the harm done in overestimating or underestimating your own or your opponent's losses and resources.

This is also only a relatively simple example of a stochastic game. The absorbing barriers can be made more complicated, the steps can be made more complicated functions of the players' strategies, the pay-off function can be generalized, and so on.

§6. <u>A slightly different model</u>.

Consider the following game based upon matching coins:

<u>When the coins match head-head, both I and II lose one. When the coins match tail-tail, I gains one, II loses one. When the coins</u>

match head-tail, or tail-head, I loses one, II gains one.

Let us assume that I starts with an amount A, II with an amount B. Both players are allowed an initial strategy, which must then remain fixed for the duration of the game, consisting of the choice of probabilities x, y, respectively, with which heads will be displayed.

For each choice of x, y we have a "gambler's ruin", or "random walk" process with three directions permitted from any point. The probabilities associated with the directions are:

(a) xy that I, II both lose one,

(b) $(1-x)(1-y)$ that I gain one, II lose one,

(c) $x(1-y) + (1-x)y$ that I lose one, II gain one.

The axes are the absorbing barriers here.

Although we do not recall having seen explicit formulas for the probabilities

$P_1(x, y)$ = probability that I is ruined before II,

$P_2(x, y)$ = probability that II is ruined before I,

there does not seem to be any great difficulty to deriving these probabilities.

If the players are restricted to selecting a single x and y, the problem that now confronts both players is that of choosing a strategy which will make the probability that the other is ruined first a max-min, and the alternate probabilities a min-max. In general, this will be impossible — and it is precisely difficulties of this type which have been the stumbling-blocks to a general theory of non-zero-sum games.

Let us examine the game a little more closely and note some elementary facts.

It is manifest that the game is unfair to I. If II's initial resources surpass those of I, II has a trivial method of ruining I, namely that of playing heads constantly.

Nevertheless, there is still a problem here, since it is possible that more economical strategies exist which make $P(x, y) = 1$.

Here we have essentially a battle of pure attrition, the Grant vs. Lee at Richmond type.

Let us now consider the case where I's initial amount is greater than that of II, and that I now concentrates on making $P(x, y)$ as large as possible. It is no longer true that a max-min relation holds,

$$\underset{x}{\text{Max}}\ \underset{y}{\text{Min}}\ P_2(x, y) \neq \underset{y}{\text{Min}}\ \underset{x}{\text{Max}}\ P_2(x, y).$$

What replaces this is a principle of the following sort: "As $A/B \to \infty$, $P_2(x, y) \longrightarrow 1$."

These and similar topics will be discussed in more detail in forthcoming papers.

§ 7. A simple three-person game.

The three, and more general n-, person game has previously been little treated, due to a lack of mathematical formulation of the term "good strategy".

Let us indicate briefly that the above ideas are applicable to some types of three-person games.

Consider the three-person game where three players match coins, odd man winning one token from the other two.

Assuming again that each player is allowed an initial strategy, which must then be adhered to for the remainder of the game, which consists in choosing the probability with which heads will be displayed, it is easy to see that I and II can form a coalition against III which will make III's expectation zero.

However, if, for example, I will always display heads, and II tails, III will always lose. Nevertheless, if he plays heads consistently, I will also always lose. Consequently unless I and II agree to share the loot, at periodic intervals a new game will be formed with a new coalition. This type of behavior bears a remarkable resemblance to European power politics of the last four hundred years, and in particular spotlights Britain's historic rôle on the continent as the preserver of the balance of power.

Even without allowing subsequent division of the spoils, I and II may still form a profitable coalition where each has equal expectation of 1/2, by choosing a random sequence of heads and tails, I playing heads, II playing tails.

Consequently two players can always combine to form a profitable coalition against the third.

We would like to point out that the above ideas are the result of several conversations with David Blackwell on this topic (although it <u>is</u> a bit unfair to hold him responsible).

Let us now show that we can, by imposing the condition that each player have a limited amount at the beginning of the game, rescue the situation from triviality.

It is intuitively clear that if III has a large enough initial amount, compared to the quantities of the other two players, the probability will be one that either I or II will be ruined before III is, after which the game will reduce to a two-person game, with III again winning if the process of matching continues. This assertion, of course, requires proof which we shall furnish subsequently.

The above criterion is only one of many possible. We state it explicitly to indicate some of the possible directions of future research.

In any case, it is clear that considerations of the above type fundamentally alter the nature of the game. The strategies to be followed depend crucially upon the initial resources.

RB-JL:rb