U.S. AIR FORCE
PROJECT RAND
RESEARCH MEMORANDUM

TWO EXAMPLES CONCERNING BEHAVIOR STRATEGIES
F. B. Thompson and R. Wagner

RM-747

18 December 1951

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 It may be expanded, modified, or withdrawn at any time.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

Copyright 1952
The RAND Corporation
Summary: Some questions concerning properties of behavior strategies are answered by means of an example.

TWO EXAMPLES CONCERNING BEHAVIOR STRATEGIES

F. E. Thompson and R. Wagner

So far little seems to be known about behavior strategies in finite games. The authors feel that in such a young domain the exhibition of examples is not out of order. There are a number of questions which thus can be quickly answered and the way cleared for deeper inquiry.

The notion of behavior strategy was first formalized by Kuhn in his paper "Extensive Games" [1]. However it is implicitly used by earlier investigators [2]. It will be assumed that the reader is familiar with the concepts involved.

1. Notation

The two examples to be presented differ only in payoff function. Thus we shall exhibit here the game-tree with the player and information partitions shown thereupon:
The $i^{th}$ "player" or team is indicated by $p_i$. The alternatives at each move are to be thought of as numbered in the clock-wise order. A behavior strategy for $p_1$ consists of a pair $<< a^i_1, a^i_2 >>$, $<< a^i_1, a^i_2 >>$, where $a^i_j$ is the probability that $p_1$ plays the $j^{th}$ alternative at his $i^{th}$ move. Evidently $a^i_1 + a^i_2 = a^2_1 + a^2_2 = 1$. Similarly a behavior strategy for $p_2$ is $<< b^1_1, b^1_2 >>$, $<< b^2_1, b^2_2 >>$.

If the players were both to use optimal mixed strategies then the expected payoff to $p_1$ would be a fixed number $v$ which is usually called the value of the game. It is the maximum $p_1$ can insure himself and the minimum to which $p_2$ can restrict $p_1$. Let "$v_1$" denote the maximum which $p_1$ can insure himself when using behavior strategies, and let "$v_2$" denote the minimum to which $p_2$ can restrict $p_1$ when using behavior strategies. It is intuitively evident (and follows formally from a result of Dalkey's) that $v_1 \leq v \leq v_2$. Dresher, Helmer and Wagner [3] have exhibited a game for which $v_1 \neq v_2$. By an optimal behavior strategy for $p_1$ we shall mean any behavior strategy which insures him $v_1$. By an optimal behavior strategy for $p_2$ we shall mean any behavior strategy for $p_2$ which restricts $p_1$ to at most $v_2$.

Pure strategies for player $p_1$ will be denoted by "$< i, j >$" where $i, j = 1, 2$; $< i, j >$ indicating alternative $i$ at $p_1$'s first move and alternative $j$ at $p_1$'s second move. Similarly for $p_2$. 
2: First example.

<table>
<thead>
<tr>
<th>Payoff matrix:</th>
<th>&lt;1,1&gt;</th>
<th>&lt;1,2&gt;</th>
<th>&lt;2,1&gt;</th>
<th>&lt;2,2&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1,1&gt;</td>
<td>1</td>
<td>1</td>
<td>-2</td>
<td>0</td>
</tr>
<tr>
<td>&lt;1,2&gt;</td>
<td>3</td>
<td>2</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>&lt;2,1&gt;</td>
<td>0</td>
<td>-1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>&lt;2,2&gt;</td>
<td>0</td>
<td>-2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

For this game the following values can be verified:

(i) \( v_1 = 0, \ v = \frac{1}{2}, \ v_2 = 1; \) thus \( v_1 < v < v_2; \)

(ii) a behavior strategy for \( p_1 \) is optimal if and only if it is of the form: \( \langle \frac{1}{2}, \frac{1}{2} \rangle, \langle h, 1-h \rangle \) where \( 0 \leq h \leq 1; \)

(iii) a behavior strategy for \( p_2 \) is optimal if and only if it is of the form: \( \langle \frac{1}{2}, \frac{1}{2} \rangle, \langle k, 1-k \rangle \) where \( 0 \leq k \leq 1; \)

(iv) if \( p_1 \) uses \( \langle \frac{1}{2}, \frac{1}{2} \rangle, \langle h, 1-h \rangle \), and \( p_2 \) uses \( \langle \frac{1}{2}, \frac{1}{2} \rangle, \langle k, 1-k \rangle \), then the expected payoff is: \( \frac{1}{2} - (h - \frac{1}{2})(k - \frac{1}{2}) \).

It will first be noted that each player has a continuum number of optimal behavior strategies. If both players play optimal behavior strategies, the payoff will be some value between \( \frac{1}{4} \) and \( \frac{3}{4} \) and any value in the closed interval \( \left[ \frac{1}{4}, \frac{3}{4} \right] \) is obtainable. Thus if both play optimal behavior strategies the payoff will actually be greater than \( v_1 \) and less than \( v_2. \)

One is immediately led to the question of "preferred" optimal behavior strategy. This suggests the induced game over the unit square whose payoff is \( H(h, k) = \frac{1}{2} - (h - \frac{1}{2})(k - \frac{1}{2}) \). This game has a saddle-point: \( h = k = \frac{1}{2} \).

Thus one might conclude that the "preferred" optimal behavior strategies are respectively \( \langle \frac{1}{2}, \frac{1}{2} \rangle, \langle \frac{1}{2}, \frac{1}{2} \rangle, \langle \frac{1}{2}, \frac{1}{2} \rangle, \langle \frac{1}{2}, \frac{1}{2} \rangle \).
Now an examination of the above game raises the following comments and questions:

(i) A convex combination of optimal mixed strategies is always an optimal mixed strategy. We note that in our first example the convex combination of optimal behavior strategies is an optimal behavior strategy. Is this always the case?

(ii) We note that the induced game we considered to determine the "preferred" optimal behavior strategies had a saddle-point solution. The corresponding consideration for optimal mixed strategies trivially yields the same result since optimal mixed strategies when played against one-another always give v. Will there always be such a saddle-point?

3: Second example.

Payoff matrix:

<table>
<thead>
<tr>
<th></th>
<th>&lt; 1,1 &gt;</th>
<th>&lt; 1,2 &gt;</th>
<th>&lt; 2,1 &gt;</th>
<th>&lt; 2,2 &gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 1,1 &gt;</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>&lt; 1,2 &gt;</td>
<td>2</td>
<td>2</td>
<td>-2</td>
<td>0</td>
</tr>
<tr>
<td>&lt; 2,1 &gt;</td>
<td>-1</td>
<td>-2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>&lt; 2,2 &gt;</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

For this game the following values can be verified:

(i) \( v_1 = \frac{2}{3}, \ v = 1, \ v_2 = \frac{4}{3} \);

(ii) \( p_1 \) has exactly two optimal behavior strategies:
    \( \alpha_1 = << 0,1 >, < \frac{1}{3}, \frac{2}{3} >> \) and \( \alpha_2 = << 1,0 >, < \frac{2}{3}, \frac{1}{3} >> \);

(iii) \( p_2 \) has exactly two optimal behavior strategies:
    \( \beta_1 = << \frac{2}{3}, \frac{1}{3} >, < 0,1 >> \) and \( \beta_2 = << \frac{1}{3}, \frac{2}{3} >, < 1,0 >> \);

(iv) the payoff to \( p_1 \) if \( p_1 \) plays \( \alpha_1 \) and \( p_2 \) plays \( \beta_j \) is given by the following table:
We thus are in a position to answer the questions in Section 2. In fact:

(1) The convex combination of optimal behavior strategies (in whatever meaningful way this may be defined) is not necessarily an optimal behavior strategy.

(ii) The induced game, whose normal form is

\[
\begin{array}{cccc}
\alpha_1 & \beta_1 & \beta_2 \\
\frac{2}{3} & \frac{11}{9} \\
\frac{4}{3} & \frac{2}{3} \\
\end{array}
\]

does not have a saddle-point solution. Thus there appears to be no way for either player to pick out a "preferred" optimal behavior strategy. The existence of a game with this property seems to us to indicate a deficiency either in the usefulness of behavior strategies or in their theory. It seems rational to expect \( p_2 \) to play an optimal behavior strategy (unless \( p_2 \) is playing under conditions where mixed strategies can be used, a situation which is excluded here). In our first example \( p_1 \) could choose a behavior strategy which would not only insure \( v \) against all behavior strategies of \( p_2 \) but insure \( v \) against all optimal behavior strategies of \( p_2 \); while \( p_2 \) can prevent \( p_1 \) from gaining more than \( v \) if \( p_1 \) plays an optimal behavior strategy. This is not the case in example two where \( p_1 \) can insure himself no more than \( \frac{2}{3} \) whereas \( p_2 \) can only be sure \( p_1 \) receives no more than \( \frac{11}{9} \).
Thus the whole question of how a team should best play a game in which team members are separated from one another over several plays seems again to be thrown open.

There is an observation concerning the second example which seems interesting. As is well known there is a mixture of pure strategies of $p_1$ such that if played it insures $p_1$ at least $v$. Now pure strategies can be considered as a subclass of the class of behavior strategies. Thus there is a mixture of behavior strategies which if played insures $p_1$ an expected payoff of at least $v$. However if one computes the optimal mixture of $p_1$'s optimal behavior strategies one obtains the mixture $\frac{6}{11} \alpha_1$ with $\frac{5}{11} \alpha_2$, which if played insures $p_1$ no more than $\frac{30}{33}$, thus less than $v$.


