*Int. Journal of Game Theory, Vol. 5, Issue 2/3, page 107–115. Physica-Verlag, Vienna.*

# On the Subject of Non Optimal Play in Zero Sum Extensive Games: "The Trap Phenomenon"[1])

By *J.-P. Ponssard*, Paris[2])

*Abstract*: This paper discusses the best reply to be played after a mistake is committed in a zero sum extensive game. First it is shown that the minimax strategies do not, in general, provide specific answers. Two possible criteria are proposed to narrow the set of optimal strategies. Then, a surprising feature of the selected optimal strategy may be observed: it may set the stage in advance so that, if the mistake is committed, the penalty will be maximal. This is called: the "trap phenomenon".

## § 1. Introduction

This paper discusses the topic of non optimal play in zero sum extensive games. It deals mainly with examples and suggestions for a general approach though this is not attempted here.

Our interest for the subject comes from the study of sequential games with incomplete information [*Ponssard*, 1975; *Ponssard* and *Zamir*, 1973] and our examples are drawn from this class. In such games the two players initially receive some private information about the final payoffs of the game tree on which they are playing. Then the personnal moves are played in an alternate fashion. Note that the players learn the moves as they are played, so that they may detect a "mistake": the play of a dominated move (this possibility exists in most extensive games). Using a behavioral strategy how can they take advantage of mistakes? This question appears to have no simple answer, except for the trivial case of games with perfect information. Therefore it seems interesting to investigate what may be the mathematical meaning of the optimal moves selected by the minimax principle. It will be shown that the optimal strategies developed for the "static" normal form, generate ambiguous "dynamic" moves. Two suggestions to systematically exploit the mistakes will be proposed: the first one is inspired from the idea of perturbed games due to J. Harsanyi [*Ponssard*, 1974; *Selten*, 1975], the second one is based on a lexicographic application of the minimax principle originally developped for matrix games by *Dresher* [1961]. These formalisations lead to a surprising result: some particular optimal strategy is ordinarily prefered and it will be shown by means of an example that the corresponding behaviour may be quite

[1]) This paper benefited from the comments of the participants of the International Workshop on Basic Problems in Game Theory (Bielefeld, September 1974).

[2]) Dr. *J.-P. Ponssard*, Centre de Recherche en Gestion, Ecole Polytechnique, 17, Rue Descartes, F–75005 Paris.

sophisticated. It operates like a trap by setting the stage in advance so that, if the mistake is committed, the penalty is maximal. Now, contrarily to a psychological approach [*Luce* and *Raiffa*, 1957], such a strategy will lead to no loss of optimality in case the mistake were avoided by the opponent.

## § 2. The Ambiguity of an Optimal Reply after a Dominated Move and two Suggestions

### 2.1 An Example

The zero sum game described by the following game tree may be interpreted as a one stage simplified poker. Player 1 receives one card which may be low ($L$) with probability 2/3 or high ($H$) with probability 1/3. Then he may drop ($D$), raise 1 unit ($R1$) or raise two units ($R2$). If Player 1 raised then Player 2 may drop ($d$) or call ($c$). There is one unit in the pot at the beginning of the game and the payoffs have been computed so that Player 1 is the maximizer.

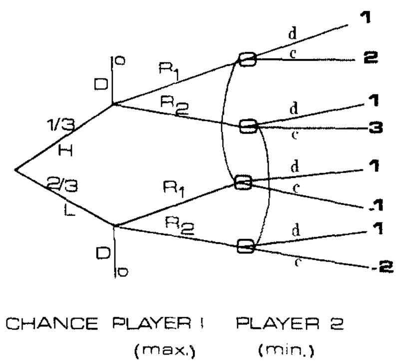


Fig. 1

This game will be solved in the traditional way, using the normal form:

| Player 2 / Player 1 | *dd*[2]) | *dc* | *cd* | *cc* |
|---|---|---|---|---|
| *DD*[1]) | 0 | 0 | 0 | 0 |
| *DR*1 | 2/3 | 2/3 | −2/3 | −2/3 |
| *DR*2 | 2/3 | −4/3 | 2/3 | −4/3 |
| *R*1*D* | 1/3 | 1/3 | 2/3 | 2/3 |
| *R*1*R*1 | 1 | 1 | 0 | 0 |
| *R*1*R*2 | 1 | −1 | 4/3 | −2/3 |
| *R*2*D* | 1/3 | 1 | 1/3 | 1 |
| *R*2*R*1 | 1 | 5/3 | −1/3 | 1/3 |
| *R*2*R*2 | 1 | −1/3 | 1 | −1/3 |

The value of this game is 5/9. Player 1's optimal mixed strategy is unique: (2/3 $R2D$, 1/3 $R2R2$)[3]).

[1]) ($D$, $D$) stands for (drop with a high card, drop with a low card) and so on.
[2]) ($d$, $d$) stands for (drop if Player 1 raises 1, drop if Player 1 raises 2) and so on.
[3]) (2/3 $R2D$, 1/3 $R2R2$) stands for (use pure strategy $R2D$ with probability 2/3 and pure strategy $R2R2$ with probability 1/3) and so on.

Player 2's optimal mixed strategy set has four extremal points

$MS1 : (1/3dd, 0dc, 1/3cd, 1/3cc)$

$MS2 : (0dd, 1/3dc, 2/3cd, 0cc)$

$MS3 : (1/6dd, 1/3dc, 1/2cd, 0cc)$

$MS4 : (1/2dd, 0dc, 1/6cd, 1/3cc)$.

However its optimal behavioral strategy set has only two extremal points

$BS1 : (1/2d, 1/2c \mid R1)$ and $(2/3d, 1/3c \mid R2)$

$BS2 : (1/3d, 2/3c \mid R1)$ and $(2/3d, 1/3c \mid R2)$

(It is easily seen that $MS1$ and $MS2$ collapse into $BS1$ and, $MS3$ and $MS4$ into $BS2$).

Note that move $R1$ is non optimal irrespective of Player 1's card. Thus if $R1$ is played Player 2 knows for sure that it is a mistake. His optimal reply as determined by the minimax principle is flexible: $(1/2d, 1/2c)$ or $(1/3d, 2/3c)$ or any convex combination. What mathematical properties determine these points? To understand, first note that whatever Player 1's strategy, Player 2 is guaranteed to pay at most 5/3 if the card is high and 0 if it is low (this may be read in the game tree by taking the expectation over Player 2's moves and then look at the worst case). Now the exploitation of the mistake should somewhat improve these security levels. In fact, looking at figure 2[3]), it is easily seen that the two extremal replys correspond to whether the mistake is exploited in case of a high card ($BS1$) or a low card ($BS2$) under the contraint that $R1$ remains a mistake (the conditional expectation given $R1$ should be lower than 5/3 in $H$ and 0 in $L$). A convex combination between $BS1$ and $BS2$ will generate a positive penalty for Player 1 whether he has a high or a low card. Then it is easy to understand that if Player 2 commits himself to such a minimax strategy, move $R1$ becomes dominated for Player 1. Such a strategy might appear good enough in a static situation, here in the extensive form the situation is somewhat different: Player 2 knows that the mistake is made, can he exploit it further?

A (may be naive) bayesian approach to this problem could start with some assumption on the likelihood ratio that the mistake is made with a high or a low card. Let $q_{HL}$ be this ratio, $q_{HL}$ = Prob $(R1|H)$ / Prob $(R1|L)$. From this assumption one derives the a posteriori probabilities Prob $(H|R1)$ and Prob $(L|R1)$.

---

[3]) The graphical interpretation of games with incomplete information is detailed in *Ponssard* [1975]. Player 2's strategies are represented by linear functions of the probability distribution on the state of nature.

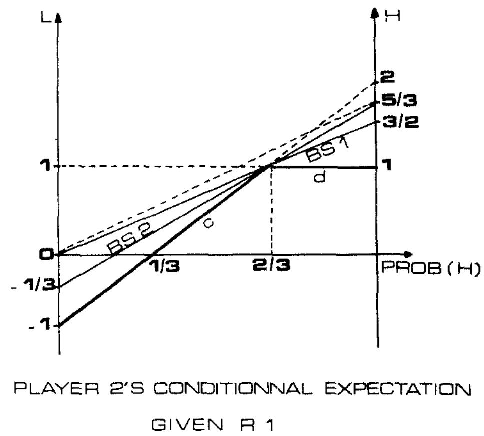


Fig. 2

$$\text{Prob}\ (H|R1) = \text{Prob}\ (R1|H)\ \text{Prob}\ (H)\ /\ (\text{Prob}\ (R1|H)\ \text{Prob}\ (H) + \text{Prob}(R1|L)\text{Prob}(L))$$

$$= q_{HL} \cdot 1/3\ /\ [q_{HL} \cdot 1/3 + 2/3]$$

$$= q_{HL}\ /\ (q_{HL} + 2)$$

Then:

if Prob $(H|R1) > 2/3$ one would select the move $d$

if Prob $(H|R1) < 2/3$ one would select the move $c$

if Prob $(H|R1) = 2/3$ one would be indifferent between $d$ and $c$.

Comparing this bayesian approach to the minimax principle gives the result that the minimax strategies ($BS1$ and $BS2$), which requires randomization, would only be consistent with the assumption Prob $(H|R1) = 2/3$ that is $q_{HL} = 4$. In other words the minimax strategy implies that the mistake is made in a very specific way: four times more often with a high card than with a low card. The game theoretic solution seems to severely restrict the potential for irrationality!

Two suggestions to put this problem in proper perspective will now be explored.

## 2.2 A Restricted Equilibrium Principle for Extensive Games

This restriction is based on the idea of small mistakes [*Selten*,1975]. Assume that Player 1 is constrained to play move $R1$ with small probabilities, say $(\epsilon_H, \epsilon_L)$, in case of a high or a low card respectively, both players being aware of this constraint. The suggestion is to solve the constrained zero sum game letting $(\epsilon_H, \epsilon_L)$ go to zero.

As $(\epsilon_H, \epsilon_L)$ go to zero, Player 1's optimal constrained strategy will converge to his optimal unconstrained strategy.

But, as may be expected, the ratio Prob $(R1|H)$ / Prob $(R1|L)$ will converge to $4$[4]. As for Player 2 the situation is summarized in figure 3. Depending on the position of $(\epsilon_H, \epsilon_L)$ with respect to the line $\epsilon_H = 4\epsilon_L$, $BS2$ or $BS1$ will be optimal (on the line any convex combination is optimal).

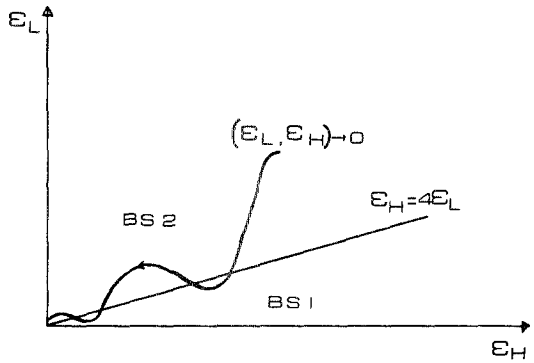


Fig. 3

This extension suggests a constrained bayesian approach for the exploitation of the mistakes:

(i) restrict Player 2's strategy set to the equilibrium strategies,
(ii) for a given likelihood ratio $q_{HL}$ select the bayesian best

reply. In this example,

if $q_{HL} > 4$ play $BS1$
if $q_{HL} < 4$ play $BS2$
if $q_{HL} = 4$ play any convex combination of $BS1$ and $BS2$.

We may say that this procedure is a mixture of game theoretic and behavioral arguments. One argument alone would not solve the problem of the exploitation of the mistake.

## 2.3 A Restricted Minimax Principle for Extensive Zero Sum Games

In this restriction we shall take the point of view that if the mistake is committed then the move is no longer in the hands of a rational opponent but in the hands of chance. The objective will be to maximize the minimum penalty over the possible states of nature, here the two possible cards. The matrix of this game is easily com-

[4]) Let $\epsilon = \text{Min}(\epsilon_H, \epsilon_L)$, Player 1's optimal constrained strategy may be computed to be $(0D, \epsilon R, (1-\epsilon)\, R2|H)$ and $(2(1-11\epsilon/2)\,/\,3D, 4\epsilon R1, (1-\epsilon)\,/\,3\,R2|L)$. Note that if $\epsilon_H \neq \epsilon_L$, one of the two additional constraints will be unbinding, generating a "volontary deviation" from the minimum mistake [*Selten*, 1975].

puted using Player 2's conditional security levels for $H$ or $L$, that is 5/3 and 0 respectively,

| Player 2 \ state of nature | $H$ | $L$ |
|---|---|---|
| $d$ | $\frac{5}{3} - 1$ | $0 - 1$ |
| $c$ | $\frac{5}{3} - 2$ | $0 + 1$ |

That is

| Player 2 \ state of nature | $H$ | $L$ |
|---|---|---|
| $d$ | 2/3 | −1 |
| $c$ | −1/3 | 1 |

The value is 1/9 and Player 2's optimal strategy is $(4/9d, 5/9c)$ or $(2/3BS1, 1/3BS2)$. Note that chance's optimal strategy is $(2/3H, 1/3L)$. In this example, Player 2's conditional security levels are improved of 1/9, irrespective of Player 1's card.

This procedure may be interpreted as a lexicographic restriction of the minimax principle, indeed it may be viewed as the extensive form counterpart of the normal form restriction suggested by *Dresher* [1961] (see *Ponssard* [1974] for a formal presentation in the special case of zero sum sequential games with incomplete information).

## § 3. Some Implications for Multistage Games ("The Trap Phenomenon")

The previous example shows that the problem of the exploitation of a mistake in an extensive game is not simple. A strict minimax approach concentrates on the point that the dominated move should remain dominated: it merely delineates the set of replies such that this dominance property is satisfied. Moreover, it was shown that such a constraint is ambiguous since, in return, it implies that the mistake should be made in some specific way. This ambiguity is somethat levelled by embedding the problem in a larger framework: two suggestions were presented to refine the game theoretic solution concept. Then the problem of the exploitation of the mistake is made precise and ordinarily a unique optimal reply is selected among those delineated by the minimax principle.

In this section we shall show that if at the beginning of the game a player specifies his criterium for exploiting the potential mistakes, in the sense of this paper, then not only will the resulting optimal strategy specify a particular way to take advantage a posteriori of a mistake but it may also specify some particular *a priori* moves which will make the a posteriori advantage maximal. We may say that such a strategy operates like a trap. The presentation of such a sophisticated behavior as resulting from formal considerations is the first objective of this paper.

## A Two Stage Game with a Trap

There are two cards, say one white and one black, which will be presented in sequence to Player 1. The sequence is chosen at random. Player 1 will announce the color of the card and he may say the truth or lie. After each card is announced Player 2 will say correct or not correct. Each stage payoff is given by the following table:

| Player 2 / Player 1 | $c$ correct | $\bar{c}$ not correct |
|---|---|---|
| truth | 0 | 1 |
| lie | 1 | −1 |

Player 1 (Player 2) wants to maximize (minimize) the expected payoff over the two stages. It is assumed that intermediary payoffs are not revealed.

A graphical interpretation of this game is given in figure 4.

The first card is white

The first card is black

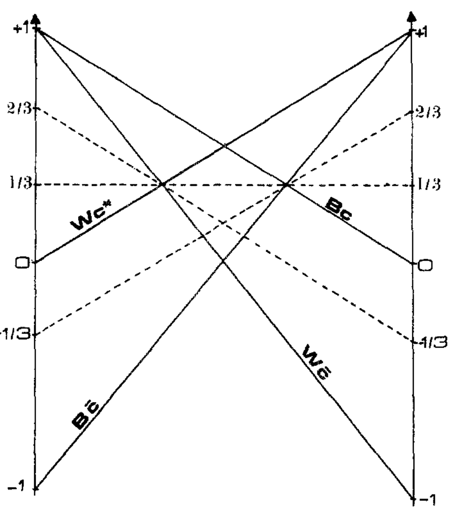


Fig. 4

* *Wc* stands for Player 1 announces white and Player 2 says correct. The corresponding line gives the expected payoff as a function of the probability that the card is black. And so on.

Suppose that the first stage only were played. Then the value of the game would be 1/3 and Player 2's optimal strategy would be to say correct with probability 1/3 and not correct with probability 2/3. Note that the conditional security levels associated with this strategy are (1/3, 1/3) and a simple graphical inspection shows that Player 2 cannot guarantee to pay less than 1/3 if the probability of the first card being black or white is 1/2. (For instance, the conditional security levels associated with the strategy "always say correct" are (+1, +1) generating an unconditional security level of +1 instead of the value 1/3).

It may be proved that the value of the two stage game is 2/3 [*Ponssard* and *Zamir*, 1973]. Taking this result for granted let us concentrate on the possible ways Player 2 can guarantee to pay at most 2/3. A simple immediate way to obtain 2/3 is to play the same behavioral strategy at the two stages. This will guarantee 1/3 per stage. Note that if such a strategy is used then whatever pure strategy is used by Player 1, the expectation is always 2/3. In fact, Player 2 need not guarantee 1/3 per stage but only 2/3 over the two stages. A compensation is possible. The following strategy achieves such a compensation while remaining globally optimal:

1st stage: say "correct" with probability 1/3 and "incorrect" with probability 2/3.

2nd stage: if Player 1 reverses his choices (he says black after he said white or vice versa), say "correct"; if Player 1 does not reverse his choices (he says black twice or white twice), say "correct" with probability 1/3 and "incorrect" with probability 2/3.

So that if for instance the first card is black, the second being necessarily white, Player 2's strategy will generate the following conditional payoffs for Player 1's four possible choices

| Player 1's choices | | Conditional payoffs | | | | |
|---|---|---|---|---|---|---|
| first card (black) | second card (white) | 1st stage | + | 2nd stage | | |
| black | white | 2/3 | + | 0 | = | 2/3 |
| black | black | 2/3 | + | (−1/3) | = | 1/3 |
| white | black | −1/3 | + | 1 | = | 2/3 |
| white | white | −1/3 | + | 2/3 | = | 1/3 |

A similar result obtains for the sequence (white, black). Observe that Player 2's strategy seems to only guarantee a security level of 2/3 at the first stage and 1 at the second stage, that is an overall payoff of 5/3, whereas in fact it guarantees 2/3 over the two stages against a rational opponent and 1/3 against a poor player.

Indeed, there is a significant difference between this second strategy and the first one which generated 2/3 whatever Player 1's choices were. Here the mistakes (black, black) or (white, white) are penalized. But this penalization can *only* be obtained if Player 2 properly correlates his choices over the two stages. If the trap is not opened at the first stage, it *cannot* operate at the second stage. It may be proved that this second strategy is the unique optimal strategy in this game when Player 2 uses the minimax restriction proposed in section 2 [*Ponssard,* 1974][5]).

## References

*Dresher, M.*: Games of Strategy: Theory and Applications, 1961.

*Luce, D.*, and *H. Raiffa*: Games and Decisions, 1957.

*Ponssard, J.P.*: $\epsilon$-Bayesian Solutions of Zero Sum Games in Extensive Form, W.P. 74–37, August 1974, I.I.A.S.A., Laxenburg, Austria.

– : Zero Sum Games with 'Almost' Perfect Information, Management Science, March 1975.

*Ponssard, J.P.,* and *S. Zamir*: Zero Sum Sequential Games with Incomplete Information, Journal of Game Theory **2** (2), 1973.

*Selten, R.*: Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games, Journal of Game Theory **4** (1), 1975.

Received January, 1975
(revised version March, 1976)

[5]) An application of this restriction to Kuhn's Simplified Poker also reduces to one point the set of optimal strategies. Noteworthly this unique optimal strategy involves underbidding which accordingly may be interpreted as a trap.