

## OPTIMAL DOUBLING STRATEGY AGAINST A SUBOPTIMAL OPPONENT

KONSTANTINOS V. KATSIKOPOULOS,\* *University of Massachusetts, Amherst and  
Max Planck Institute for Human Development*  
ÖZGÜR ŞİMŞEK,\*\* *University of Massachusetts, Amherst*

### Abstract

For two-person zero-sum games, where the probability of each player winning is a continuous function of time and is known to both players, the mutually optimal strategy for proposing and accepting a doubling of the game value is known. We present an algorithm for deriving the optimal doubling strategy of a player who is aware of the suboptimal strategy followed by the opponent. We also present numerical results about the magnitude of the benefits; the results support the claim that repeated application of the algorithm by both players leads to the mutually optimal strategy.

*Keywords:* Gambling; doubling; optimal strategy

2000 Mathematics Subject Classification: Primary 60G40  
Secondary 91A05

### 1. Introduction

We consider a zero-sum game between two players  $A$  and  $B$ , in which, at any time  $t \geq 0$ , the probability  $p(t)$  that  $A$  will win is known to both players. We assume that  $p(t)$  is a continuous function of  $t$ . The objective of each player is to maximize their expected value from the game.

Initially, the value of the game is one unit. Either player may propose a doubling of the game value (or a ‘double’) each time it is his or her turn to play. When one player proposes a double, the other player may refuse it, in which case the game ends and the refusing player loses the current value of the game, or may accept the double, in which case the game continues with twice the value. When a double is proposed and accepted, the player who accepted gains the exclusive right to propose the next double. In other words, a player cannot propose two consecutive doubles.

This model may be used to approximately describe a number of gambling games in discrete time, in which the value of a game may increase exponentially through repeated use of the doubling device. Backgammon is such a game, with many dedicated players. Intuitive arguments on doubling strategies have been discussed by analytically oriented backgammon experts (see [1] and [4]), but there does not appear to be a consensus on which strategy is the best one [5]. Optimal doubling strategies in backgammon have also been studied formally by applied probabilists (see [2], [6], [7]).

---

Received 16 October 2001; revision received 14 March 2005.

\* Postal address: Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Lentzeallee 94, Berlin, 14195, Germany. Email address: katsikop@mpib-berlin.mpg.de

\*\* Postal address: Department of Computer Science, University of Massachusetts, 140 Governors Drive, Amherst, MA 01003, USA.

Keeler and Spencer [2] showed that, if both play optimally, then  $A$  should double when  $p(t) \geq 0.8$ , and should decline  $B$ 's double when  $p(t) \leq 0.2$ ;  $B$ 's optimal strategy mirrors this. The main tool in their argument was the following lemma.

**Lemma 1.** *Suppose that  $p(t) = p$  and the game continues indefinitely. Then, for  $s > t$ , the probability that  $p(s) = p + x$  before  $p(s) = p - y$  is  $y/(y + x)$ , assuming that  $0 \leq p - y \leq p \leq p + x \leq 1$ .*

Keeler and Spencer [2] also hinted that a player aware of the opponent's suboptimal strategy can adopt a doubling strategy better than the mutually optimal one. Here, we solve this problem and present an algorithm that, given  $p(t)$  and the suboptimal strategy of the opponent, outputs the adaptive optimal strategy.

In the next section, we state the problem and outline the solution. Then we present the details of the algorithm and provide some numerical results.

### 2. The problem and its solution

We assume that  $p(t)$ , the probability of player  $A$  winning, is currently  $p$ , that the maximum value of  $p(t)$  at which player  $B$  will accept  $A$ 's double is  $a'$ , and that the maximum value of  $p(t)$  at which player  $B$  will propose a double is  $d'$ . We seek to determine, for player  $A$ , the optimal minimum values of  $p(t)$  for accepting ( $a$ ) and proposing ( $d$ ) doubles. This is shown schematically in Figure 1.

We consider the situation where  $d' < p < a'$  and  $a < p < d$ . These assumptions are reasonable in a number of situations, for example at the beginning of a game between two equally skilled players (i.e.  $p = 0.5$ ). We also assume that no doubles have been proposed yet and, thus, either player may propose the next double. The algorithm will output the optimal values of  $a$  and  $d$  for this situation. It will be obvious how the algorithm may be adapted for other relative values of  $d', a', p, d$ , and  $a$ , and in the case where only one player is eligible to propose the next double.

Player  $A$  collects the current value of the game,  $v$ , if  $p(t) = 1$ , or if they propose a double that is refused. For the latter to happen, it is necessary that  $p(t) = d$  and  $d > a'$ . Conversely,  $A$  collects  $-v$  if  $p(t) = 0$ , or if they refuse a double that is proposed by  $B$ . For the latter to happen, it is necessary that  $p(t) = d'$  and  $a > d'$ .

Now,  $v = 2^k$ , where  $k \geq 0$ , is the number of doubles accepted in the game. For  $A$  to accept a double, it is necessary that  $p(t) = d'$  and  $a \leq d'$ . For  $B$  to accept a double, it is necessary that  $p(t) = d$  and  $d \leq a'$ .

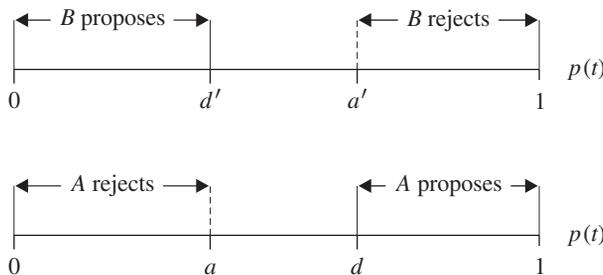


FIGURE 1: The doubling strategy of each player.

Thus, given  $(a', d')$ ,  $A$ 's choice of  $(a, d)$  leads to exactly one of the following four possibilities:

- I.  $(a > d', d > a')$ ,
- II.  $(a \leq d', d > a')$ ,
- III.  $(a > d', d \leq a')$ , or
- IV.  $(a \leq d', d \leq a')$ .

For each case, we use Lemma 1 when  $d \leq p \leq d'$  to find the probability that, given  $p(t) = p$ , we have  $p(s) = d$  or  $p(s) = d'$  before  $p(s) = 1$  or  $p(s) = 0$ . This enables us to compute the corresponding values of  $E[v]$ , the expected value of the game for  $A$ .

### 3. The algorithm

As stated above, the dynamics of the game depend on the relative values of  $a$  and  $d'$ , and of  $a'$  and  $d$ . We now consider each of the four possibilities.

In case I, both players refuse when their opponent proposes a double. Thus,  $A$  collects  $-1$  if  $p(t) = d'$  before  $p(t) = d$ , and  $+1$  otherwise. Thus, from Lemma 1,

$$E_I[v] = \frac{d - p}{d - d'}(-1) + \frac{p - d'}{d - d'}(+1). \tag{1}$$

In case II,  $A$  accepts  $B$ 's doubles, but  $B$  refuses  $A$ 's doubles. Thus, if  $p(t) = d$  before  $p(t) = d'$ , then  $A$  proposes a double that is refused and  $A$  collects  $+1$ . If  $p(t) = d'$  before  $p(t) = d$ , then  $B$  proposes a double that is accepted, and the game terminates when  $p(t) = d$  with  $A$  collecting  $+2$ , or when  $p(t) = 0$  with  $A$  collecting  $-2$ . Then, application of Lemma 1 and some simple algebra yield

$$E_{II}[v] = \frac{d - p}{d - d'} \left( \frac{4d'}{d} - 2 \right) + \frac{p - d'}{d - d'}(+1). \tag{2}$$

In case III,  $A$  refuses  $B$ 's doubles, but  $B$  accepts  $A$ 's doubles. We obtain the following equation:

$$E_{III}[v] = \frac{d - p}{d - d'}(-1) + \frac{p - d'}{d - d'} \left( \frac{4(d - d')}{1 - d'} - 2 \right). \tag{3}$$

Finally, in case IV, all doubles are accepted, and the game continues until  $p(t) = 0$  or  $p(t) = 1$ . We will express the expected value of the game for player  $A$ , conditioning on which player doubles first or, equivalently, on the value of  $p(t)$  when the first double is proposed.

The value of  $p(t)$  when the first double is proposed is  $d'$  with probability  $(d - p)/(d - d')$ , and  $d$  with probability  $(p - d')/(d - d')$ . Thus, if  $E_{d'}[v]$  is the expected value of what  $A$  collects in the former case, and  $E_d[v]$  is the expected value of what  $A$  collects in the latter case, then

$$E_{IV}[v] = \frac{d - p}{d - d'} E_{d'}[v] + \frac{p - d'}{d - d'} E_d[v]. \tag{4}$$

We make the economically reasonable assumption that each player may double for, at most, the same finite number of times,  $n$ . To compute  $E_{d'}[v]$ , we separately consider the cases where both players double the same number of times, and where player  $B$  doubles one more time than  $A$ . In (5), below,  $E_{d'}[v]$  is expressed as a sum of three terms. The first two terms refer to

the realizations of the process where both players double the same number of times: exactly  $n$  in the first term and less than  $n$  in the second term. The third term refers to the case where player  $B$  doubles one more time than player  $A$ . The expectation  $E_d[v]$  is expressed in a similar manner in (6), below. We obtain

$$\begin{aligned}
 E_{d'}[v] &= \left(\frac{d'}{d}\right)^n \left(\frac{1-d}{1-d'}\right)^n (2d'-1)2^{2n} \\
 &+ \sum_{k=0}^{n-1} \left(\frac{d'}{d}\right)^k \left(\frac{1-d}{1-d'}\right)^k \left(\frac{d'd-d'}{d(1-d')}\right) 2^{2k+2} \\
 &+ \sum_{k=0}^{n-1} \left(\frac{d'}{d}\right)^k \left(\frac{1-d}{1-d'}\right)^k \left(\frac{d-d'}{d}\right) (-2^{2k+1})
 \end{aligned} \tag{5}$$

and

$$\begin{aligned}
 E_d[v] &= \left(\frac{d'}{d}\right)^n \left(\frac{1-d}{1-d'}\right)^n (2d-1)2^{2n} \\
 &+ \sum_{k=0}^{n-1} \left(\frac{d'}{d}\right)^k \left(\frac{1-d}{1-d'}\right)^k \left(\frac{1-d}{1-d'} \frac{d-d'}{d}\right) (-2^{2k+2}) \\
 &+ \sum_{k=0}^{n-1} \left(\frac{d'}{d}\right)^k \left(\frac{1-d}{1-d'}\right)^k \left(\frac{d-d'}{1-d'}\right) 2^{2k+1}.
 \end{aligned} \tag{6}$$

We make the assumption of a maximum finite number of doubles to prevent possible nonconvergence of the (otherwise infinite) series in (5) and (6).

An algorithm for computing  $a$  and  $d$  for  $d' < p < a'$  and  $a < p < d$  is as follows.

- Input: Choose  $d'$ ,  $p$ , and  $a'$  such that  $d' < p < a'$ .
- Step 1: Use (5)–(6) to compute  $E_{d'}[v]$  and  $E_d[v]$ .
- Step 2: Use (1)–(4) to compute  $(a^*, d^*) = \arg \max_{(a,d)} E[v]$ .
- Step 3: Set accepting point  $a = \min a^*$  and doubling point  $d = \min d^*$ .
- Output: Return  $a$  and  $d$ .

### 4. Numerical results

From the above algorithm, three questions naturally arise. Firstly, is there a departure from the mutually optimal strategy (0.2, 0.8) for player  $A$  when  $A$  adapts to  $B$ 's suboptimal strategy? Secondly, what are the benefits to  $A$  for such departures? Thirdly, since  $B$  may also run the algorithm using  $(a, d)$  as input to compute a new  $(a', d')$ , and so on, does repeated use of the algorithm by both players lead to convergence to the mutually optimal strategy?

Simulations suggest that player  $A$  should adapt to  $B$ 's suboptimal strategy, and that benefits vary considerably depending on  $B$ 's strategy. We present results for  $p = 0.5$  and  $n = 5$ . Firstly, consider the case where  $a' = 0.8$ , but  $d' \neq 0.2$ . That is, player  $B$  accepts  $A$ 's doubles in accordance with the mutually optimal strategy, but departs from this strategy in proposing doubles. If  $d' < 0.2$  then  $B$  is conservative and proposes doubles only at too high a probability of winning. On the other hand, if  $d' > 0.2$  then  $B$  is eager and proposes doubles at too low a probability of winning.

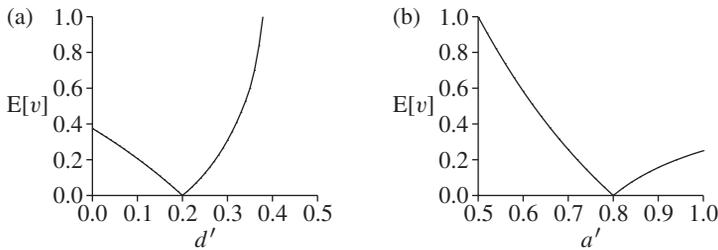


FIGURE 2: Expected game value for player A for some strategies of B, where (a)  $p = 0.5, n = 5, a' = 0.8$  and (b)  $p = 0.5, n = 5, d' = 0.2$ .

TABLE 1: Optimal responses to some suboptimal strategies.

Player B's strategy	Player A's optimal adaptive strategy
$d' = 0.2, a' < 0.8$	any $a \in [0, d']$ , $d = a' + \varepsilon$
$d' = 0.2, a' > 0.8$	any $a \in [0, d']$ , $d = a'$
$d' < 0.2, a' = 0.8$	any $a \in (d', 1]$ , $d = a' - \varepsilon$
$d' > 0.2, a' = 0.8$	any $a \in [0, d']$ , $d$ varies with $d'$

Figure 2(a) shows, for different values of  $d'$ , the expected value of the game for player A when A optimally adapts to B's strategy. As anticipated, when  $d' = 0.2$ , the expected value of the game is zero. However, this value is nonzero for all other  $d'$ , and rapidly increases as  $d'$  moves away from 0.2. The increase is much more dramatic when B is eager (rather than conservative), as many more doubles are expected in that case. In fact, the slope of the curve becomes close to infinity when  $d'$  approaches 0.4, which is why the plot is truncated. At the other extreme, if B never doubles ( $d' = 0$ ), then the expected value of the game is 0.375.

Next consider the case where  $d' = 0.2$  and  $a' \neq 0.8$ . That is, player B is suboptimal in accepting doubles. Figure 2(b) shows the expected value of the game for A for different values of  $a'$ , when A optimally adapts to B's strategy. Benefits increase as  $a'$  moves away from 0.8 and are considerably larger when B is eager to reject doubles ( $a' < 0.8$ ). Benefits, however, are not as large as when  $d' > 0.2$  and  $a' = 0.8$  because, overall, a smaller number of doubles is proposed. If B always accepts ( $a' = 1$ ) then the value of the game is 0.25.

Table 1 shows the optimal strategy for player A in the above two cases. Note that the optimal response is often simple, leaving open the possibility that humans might be able to behave optimally or near-optimally. For example, the optimal response to an opponent eager to reject doubles ( $a' < 0.8$ ) is to propose doubles when the probability of winning rises just above  $a'$  by a positive infinitesimal  $\varepsilon$ .

Finally, Figure 3 provides a contour plot of the expected game value for player A, when we allow both  $d'$  and  $a'$  to vary simultaneously. The plot shows contour lines at increments of 0.1 in game value.

The last question to investigate is whether the repeated use of the algorithm by both players converges to the mutually optimal strategy. It is known that such an iterative process in a zero-sum two-player game need not converge [3]. In our case, a formal analysis of convergence is not straightforward because the dependence of  $a$  and  $d$  on  $a'$  and  $d'$  is not captured by a single formula. We have, however, observed convergence in simulations for all settings of  $a'$  and  $d'$  in increments of 0.01 (using  $\varepsilon = 0.01$ ).

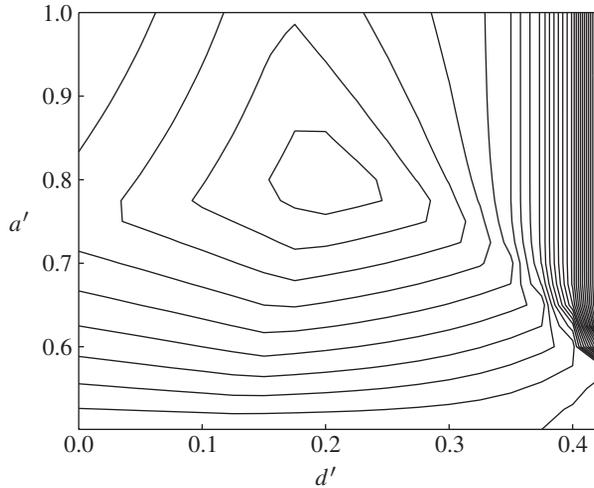


FIGURE 3: Contour plot of the expected game value  $E[v]$  for player A, where  $p = 0.5$  and  $n = 5$ .

### 5. Summary

We have extended previous work on optimal strategies for proposing and accepting doubles in two-person zero-sum, completely observable, continuous games, by considering situations where one of the players,  $B$ , uses a suboptimal strategy. We formally derived an algorithm that the other player,  $A$ , may use in order to adapt to  $B$ 's strategy. We have also provided numerical results for the benefits of the use of the algorithm, and showed how they depend on  $B$ 's strategy. The results also provide support for the claim that repeated application of the algorithm by both players leads to the mutually optimal strategy.

We note here that caution should be exercised in translating this work into practical advice for games that unfold in discrete time. The features of gammon and, especially, backgammon further complicate the analysis, as does the fact that the probability of  $A$  winning varies discretely, not continuously.

### References

- [1] JACOBY, O. AND CRAWFORD, J. R. (1970). *The Backgammon Book*. Viking Press, New York.
- [2] KEELER, E. AND SPENCER, J. (1975). Optimal doubling in backgammon. *Operat. Res.* **23**, 1063–1071.
- [3] LUCE, R. D. AND RAIFFA, H. (1957). *Games and Decisions*. John Wiley, New York.
- [4] MAGRIEL, P. (1985). *Backgammon*. X-22 Publishing, New York.
- [5] ROBERTIE, B. (1997). *Backgammon for Serious Players*, 2nd edn. Cardoza Publications, New York.
- [6] ZADEH, N. (1977). On doubling in tournament backgammon. *Manag. Sci.* **23**, 986–993.
- [7] ZADEH, N. AND KOBILSKA, G. (1977). On optimal doubling in backgammon. *Manag. Sci.* **23**, 853–858.