Applied Probability Trust (15 March 2005)

OPTIMAL DOUBLING STRATEGY AGAINST A SUB-OPTIMAL 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 sub-optimal 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 from both players leads to the mutually optimal strategy.

Keywords: gambling; doubling; optimal strategy

AMS 2000 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 his or her expected value from the game.

* Postal address: Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany.

** Postal address: Department of Computer Science, University of Massachusetts, 140 Governor's Drive, Amherst, MA 01003.

Initially, the value of the game is one unit. Either player may propose a doubling of the game value each time it is his or her turn to play. When one player proposes a double, the other player may refuse the double, in which case the game ends and the refusing player loses the current value of the game, or 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 can not propose two consecutive doubles.

This model may be used to approximately describe a number of gambling games in discrete time in which the value 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 [1, 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 [2, 7, 6].

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

Lemma. Suppose $p(t) = p$ 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 $0 \leq p - y \leq p \leq p + x \leq 1$.

Keeler and Spencer [2] also hinted that a player aware of the opponent's sub-optimal 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 sub-optimal 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 and the maximum value of $p(t)$ at which player B will accept A's double is a' and 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 (then, $p = 0.5$). We also assume that no doubles have been proposed yet and thus, both players 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 positions of d', a', p, d, a, and for only one player being eligible to propose the next double.

Player A collects the current value of the game, v, if $p(t) = 1$, or if he/she proposes 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 he/she refuses 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 $d \leq a'$.

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

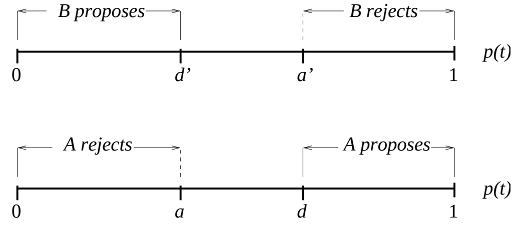

FIGURE 1: The doubling strategy of each player.

$(a > d', d > a')$, $(a \leq d', d > a')$, $(a > d', d \leq a')$, and $(a \leq d', d \leq a')$, which we label I-IV respectively. For each case, we use the Lemma above when $d \leq p \leq d'$ to find the probability that, given $p(t) = p$, then $p(s) = d$ 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 to 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. Below, we consider each of the four possibilities.

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

$$E_{\mathrm{I}}[v] = \frac{d-p}{d-d'} \cdot (-1) + \frac{p-d'}{d-d'} \cdot (+1) \tag{1}$$

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 which is refused and A collects $+1$. If $p(t) = d'$ before $p(t) = d$, then B proposes a double which is accepted, and the game terminates when $p(t) = d$ with A collecting $+2$, or $p(t) = 0$ with A collecting -2. Then, application of the Lemma and some simple algebra yields the following:

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

In case III, A refuses B's doubles, but B accepts A's doubles and we obtain the following equation:

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

Finally, in case IV, all doubles are accepted, and the game continues until $p(t)$ equals 0 or 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 $\frac{d-p}{d-d'}$, and d with probability $\frac{p-d'}{d-d'}$. 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, the following holds.

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

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 consider separately the cases where both players double the same number of times, and where player B doubles one more time than A. In equation (5), $E_{d'}[v]$ 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. $E_d[v]$ is expressed in a similar manner in equation (6).

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

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

We make the assumption of a maximum finite number of doubles to prevent possible non-convergence of the (otherwise infinite) series in (5) or (6).

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

Input: d', p, 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: a, d.

4. Numerical Results

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

Simulations suggest that player A should adapt to B's sub-optimal strategy, and that benefits vary considerably depending on B's strategy. We present results for $p = 0.5$ and $n = 5$. First 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 $d' < 0.2$, B is conservative and proposes only at too high a probability of winning. On the other hand, if $d' > 0.2$, B is eager and proposes at too low a probability of winning.

Figure 2a shows, for different values of d', the expected value of the game for 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 non-zero at 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 ($d' = 0$), 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 sub-

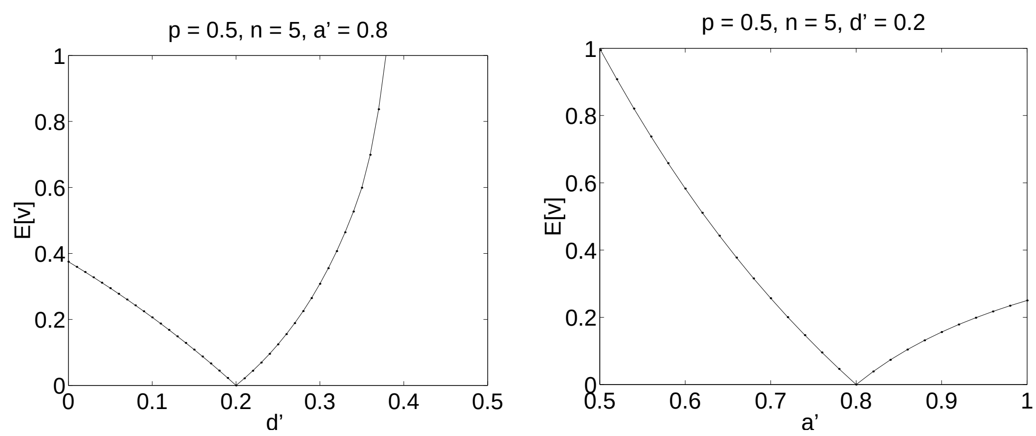

FIGURE 2: Expected game value for player A for some strategies of B.

optimal in accepting doubles. Figure 2b 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 at rejecting 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 ($a' = 1$), 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 probability of winning rises just above a' by an infinitesimal ϵ.

Finally, Figure 3 provides a contour plot of the expected game value for player A

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

TABLE 1: Optimal responses to some sub-optimal strategies; ϵ denotes an infinitesimal positive number.

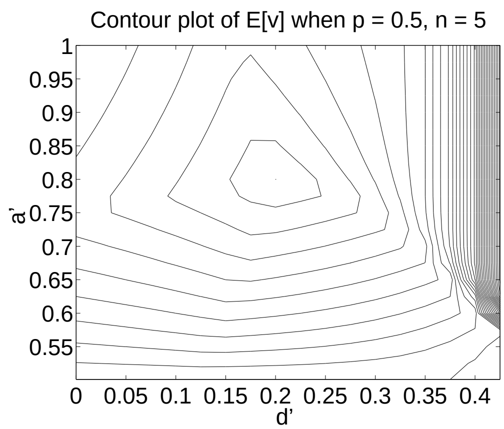

FIGURE 3: Contour plot for expected game value for player A.

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 we investigated 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 an ϵ value of 0.01).

5. Summary

We 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 sub-optimal strategy. We formally derived an algorithm that the other player, A, may use in order to adapt to B's strategy. We also provided numerical results for the benefits from 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 from both players leads to the mutually optimal strategy.

We note here that caution should be exercised in translating this work to practical advice for games that unfold in discrete time. Especially for backgammon, the features of gammons and backgammons further complicate the analysis, as does the fact that A's win probability 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. *Oper. Res.* **23,** 1063-1071.

[3] Luce, R.D. and Raiffa, H. (1957) Games and Decisions, 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. *Mgmt. Sci.* **23,** 986-993.

[7] Zadeh, N. and Kobliska, G. (1977). On optimal doubling in backgammon. *Mgmt. Sci.* **23,** 853-858.