Backgammon: The Optimal Strategy for the Pure Running Game*

Edward O. Thorp

[Edward O. Thorp & Associates]

Abstract. The pure running game (end game, no-contact game) occurs in backgammon when no further contact between opposing men is possible. It compares roughly in importance to the end game in chess. This paper presents a complete exact solution to the pure running game with the doubling cube under the simplifying assumption (Model $1'$) that dice totals can be arbitrarily subdivided in moving men and that men can be borne off even though not all are in the inner table.

The actual pure running game (i.e., without these simplifying assumptions) can be solved to high approximation by the methods of this paper. The strategies obtained are superior to those employed by the world's best players.

Our strategies and methods for the end game lead to substantially improved middle-game (where contact is still possible) play. It is feasible to find exact solutions to some middle-game situations. By also employing positional evaluation functions, it seems feasible to program an existing computer to become the world backgammon champion.

1 Introduction

The game of backgammon, along with chess and checkers, is among the oldest games known. A primitive backgammon-like game dated 2600 B.C. is on display in the British Museum (Jacoby and Crawford 1970, pp. 8–9).

In place of the traditional board we use the equivalent representation of 26 linearly ordered cells numbered $0, 1, \ldots, 25$. The traditional board corresponds to cells $1, \ldots, 24$. White men travel from higher to lower numbered cells on White turns and Black men move in the opposite direction. To preserve the symmetry, we also use the coordinate systems $i = Wi = B(25 - i)$. See Figure 1.

The object of White is to move all men to $W0$ and the object of Black is to move all men to $B0$. The first player to get all men to his 0 cell wins. The

* Presented at the Second Annual Conference on Gambling, South Lake Tahoe, Nevada, June 15–18, 1975.

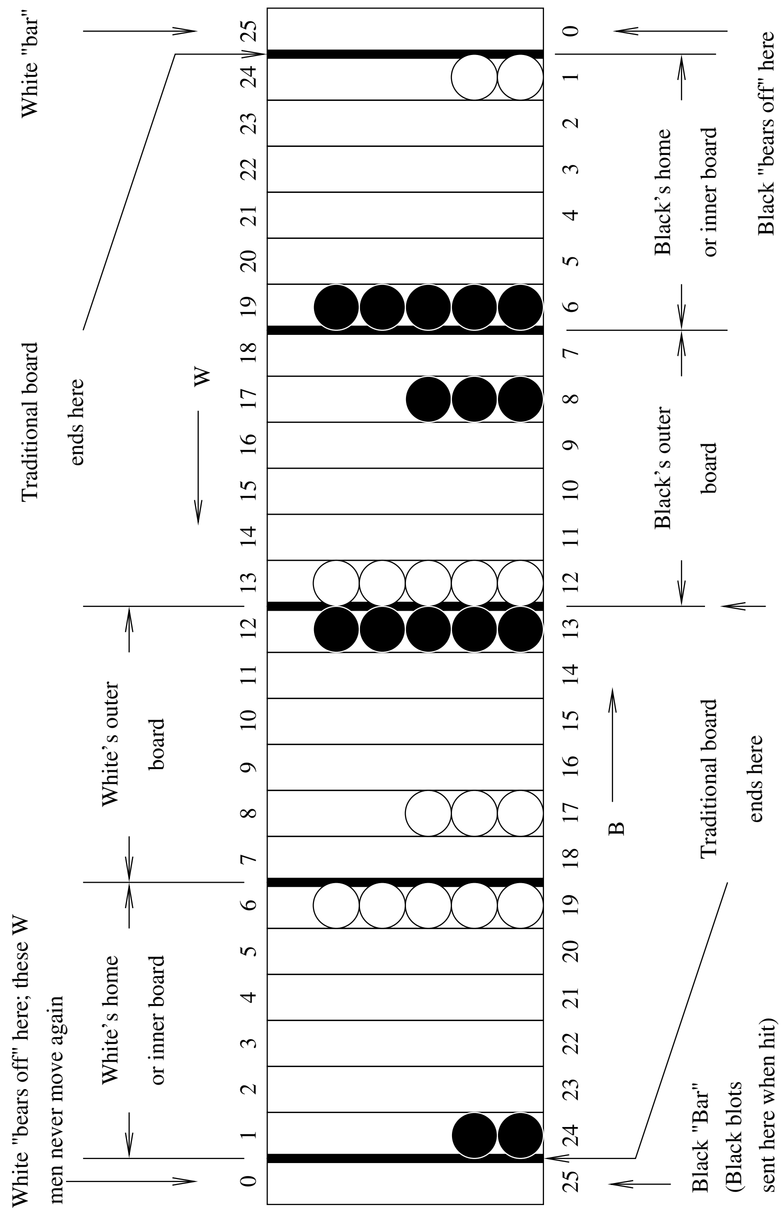

Fig. 1. Backgammon board. White travels left on his turn. Black travels right on his turn.

men are initially as in Figure 1. The additional rules (see, e.g., Holland 1973, Hopper 1972, Jacoby and Crawford 1970, Lawrence 1973, Obolensky and James 1969, or Stern 1973) are assumed known to the reader and will be used, without a complete formal statement, as we proceed.

[In Section 2 we give bounds for the duration of simplified models of the game. In Section 3 we give bounds for the number of board configurations. Section 4 gives a recursive solution to the Model 1 end game, i.e., the one-checker pure race model, and Section 5 extends this to Model 1′, which adds the doubling cube. Section 6 discusses possible extensions of the research.]

2 Duration of the game

It is clear from Figure 1 that each player requires a minimum of 167 steps to bear off all his men. The expected number of steps on the first player's first turn is 7; for each subsequent roll of the dice pair it is 8 1/6. Dividing 167 by 8 1/6 gives 20.45 so it is plausible to suppose that the lower bound for the expected duration of the game might have order of magnitude 20.

More precisely, let $X_1, X_2, \ldots, X_n, \ldots$ be the totals rolled by the second player. The X_i are independent identically distributed random variables. Let $M = 167$ and let n^* be the least n such that $X_1 + \cdots + X_n \geq M$. Then $E(n^*)$ is the expected time for player 2 to finish assuming no "interference" and no "waste" (i.e., Model 1, below). Thus $E(n^*) \geq 20.45$ in this case and à fortiori in the standard game. Similarly for the first player $(E(m^*) - 1)E(X_i) + 7 \geq M$ or $E(m^*) \geq 20.59$. [Define the *duration* of the game to be $\min(m^*, n^*)$.] Note, however, that the expected duration of the standard game, or even of Model 1 below, could still be less than 20.45 because the game ends when either player finishes and in general $E(\min(m^*, n^*)) < \min(E(m^*), E(n^*))$ for random variables m^* and n^*.

Definition 1 (Model 1). *The random walk idealization of the game assumes that two players (F and S for "first" and "second") alternately roll the pair of dice. They start with totals of zero and accumulate points according to the outcomes of their rolls, valued as is done subsequent to the first roll of the actual game. The goal of F is a total $N(F)$; that of S is $N(S)$. The first player to accumulate a total which is greater than or equal to his goal wins.*

Model 1 omits several features of the game: (1) there is no blocking and no hitting of blots; in Model 1 opposing men can both occupy the same cell; the "flows" of opposing men do not interact; (2) there are no restrictions on bearing off in Model 1 whereas in the game all men not already borne off must be in the player's home board before he can continue to bear off; (3) Model 1 allows the total to be arbitrarily subdivided for use in moving men whereas in the game each move of a man must be in a jump that uses the entire number on a die. For instance, rolling $(5, 3)$ the admissible jumps are 5 and 3. Rolling $(4, 4)$ ("doubles") they are 4, 4, 4, and 4.

Model 1 may be played as a race between F and S, each on his own individual track. The tracks have $N(F)+1$ and $N(S)+1$ cells, respectively. The initial configuration is a single man in cell 0 of each track. Players roll alternately, moving the total number of steps indicated by the dice. The first player to reach the last cell of his track wins.

The expected duration $E(D_1)$ of the Model 1 game is a lower bound for the expected duration $E(D_G(s))$ of the standard game G where s is any pair of strategies for Black and White. To see this, note that in Model 1 each roll of the dice is fully utilized in reducing the corresponding player's pip count (pips needed to win) whereas in the standard game part or all of the roll of the dice may not be utilized. Reasons include blocked points and the possible waste of some of the roll in bearing off. Also the player on roll can hit the other player's blots thereby increasing the other player's pip count. Thus each Model 1 player's pip count will always be less than or equal to the corresponding player's standard game pip count so the Model 1 game will end no later than the standard game. So $D_1 \leq D_G(s)$ for each realization of the standard game hence $E(D_1) \leq E(D_G(s))$ for all pairs of player strategies s.

Model 1 corresponds to a random path on the lattice points of the non-negative quadrant with absorption at $x = 167$ (F wins) and at $y = 167$ (S wins). The odd numbered steps are horizontal and to the right distances $X_1, X_3, \ldots$, and the even numbered steps are vertically up distances $Y_2, Y_4, \ldots$, where $X_1, X_3, \ldots$ and $Y_2, Y_4, \ldots$ are independent random variables distributed like the dice outcomes. A computer simulation will now readily yield a fairly accurate value of $E(D_1)$.

What is the least number n of moves possible in the standard game? We know $n \geq \lceil 167/24 \rceil = 7$ in the Model 1 game because 24 points (double sixes) is the maximum per roll. (The symbol $\lceil x \rceil$, the "ceiling" of x, is the smallest integer k such that $k \geq x$.) Hence 7 is a lower bound.

To see that the game cannot end in 7 moves, consider a somewhat more realistic model.

Definition 2 (Model 2). *The* no-contact *model assumes the usual valuation of, and use of the outcomes of the rolls. Hence the men must move in jumps corresponding to the dice faces. This is the only difference from Model 1. A board configuration and who has the move is specified. There are no stakes and no doubling cube. Men can be borne off even though not all have reached the inner table. Unlike Model 1, there may be "waste" when men are borne off: If a six is used to bear off a man from cell 5, the unused step cannot be applied to another man.*

Visualize Model 2 as a game where the players' pieces are on two separate boards. Men then move as in the standard game except that not all men need to reach the inner table before men can be borne off. This can be an advantage. For instance suppose W has only two men, on $W1$ and $W13$. Now W rolls $(6,1)$. He has a choice of $W1 \to W0$ and $W13 \to W7$ or of $W13 \to W6$ whereas in the standard game only the latter is allowed. But the configuration "one man on $W7$" is better than "men on $W1, W6$" in that $n^*(W7) \leq n^*(W1, W6)$ and strict inequality holds with positive probability. This will be proven (easily) later. For now, note that a roll of $(5, 2)$ bears off one man on $W7$ but with men on $W1$ and $W6$ it leaves a man on $W1$.

In the preceding example, the Model 2 player upon rolling $(6, 1)$ had a choice as to whether to change $\{W1, W13\}$ into $\{W0, W7\}$ or into $\{W1, W6\}$ and we saw that $\{W0, W7\}$ was better. The game situations in which a player has a decision to make are "states." The rule the player uses to make his choices in each state is his strategy.

Definition 3. *A* configuration *is an arrangement of the 15 White men and the 15 Black men on the extended board of 26 ordered cells. A* position *is a configuration and the specification of who has the doubling cube and what its value is. We always assume the initial stake is 1 unit. A* state *is a position and the specification of who is to move.*

[A configuration may be thought of as a photograph of the board without the doubling cube. Then a position is a photograph that also shows (the upper face of) the doubling cube.]

The game can be unambiguously continued when the state is specified.

Definition 4. *A* strategy *for a player is a rule that specifies how the player is to move for each state of the game in which he is to move. For a given state the strategy may be pure or mixed. In a* pure *strategy the player always makes the same moves for a given roll. In a* mixed *strategy he chooses his moves for a given roll from among several alternatives, each having a specified probability of being chosen. Strategies also may be stationary or nonstationary. A* stationary strategy *is one which does not depend on the time (i.e., turn number). Each time a state repeats, a stationary strategy has the same rule for selecting moves. A* stationary pure strategy *selects exactly the same moves. A* nonstationary strategy *does depend on the time. A* strategy pair $s = (s_1, s_2)$ *is a specification of strategies for players 1 and 2 respectively.*

We shall generally limit ourselves to stationary pure strategies in subsequent sections. This is justified by the facts that (1) there is a strategy pair $s = (s_1^*, s_2^*)$ such that s_1^* and s_2^* are optimal for the respective players (see von Neumann and Morgenstern 1964 for explanation of "optimal" and for proof) and (2) there is an optimal pair of pure strategies among the class of all pairs of optimal strategies.

Proceeding as in the Model 1 case shows that (a) the expected duration $E(D_2)$ of Model 2 is less than or equal to $E(D_G)$; (b) for each sequence of rolls the Model 2 game ends at the same time or sooner than the corresponding standard game. Thus if L_2 is the least number of turns in any Model 2 game and L_G is the least for G, then $L_2 \leq L_G$. Now count the number of jumps needed in Model 2 to bear off. Assuming the best case, all rolls $(6, 6)$, the number of jumps required for a (say) W man at coordinate x is $\lceil x/6 \rceil$, which gives a total of $5 + 3 \cdot 2 + 5 \cdot 3 + 2 \cdot 4 = 34$. Thus $L_G \geq L_2 = \lceil 34/4 \rceil = 9$ rolls. It is easy to prescribe sequences of rolls and strategies such that either B or W can win G in 9 moves. Therefore $L_G = 9$.

We have shown that $E(D_1) \leq E(D_2(s)) \leq E(D_G(s))$. Note that strategies aren't relevant for D_1.

Theorem 1. *For any strategy pair s, $E(D_1) < E(D_2(s)) < E(D_G(s))$.*

Proof. To show strict inequality, it suffices to give a finite sequence of rolls such that $D_1 < D_2(s)$ for each s and to give a sequence such that $D_2(s) < D_G(s)$ for each s. If $(6, 5)$ is the opening roll and double sixes only are rolled thereafter, then $D_1 = 7$ and $D_2(s) \geq 9$. If $(6, 5)$ is the opening roll and then double fives only are rolled, $D_2(s) = 11$ but in G the men on 1, 6, 19, and 24 can never move so we can make $D_G(s) \geq N$ for arbitrary N. □

In both Model 1 and Model 2 the length of the game is bounded above and the game trees are finite. The proof of the theorem shows this is not the case for G.

Example 1. A prime (of player X) is 6 or more consecutive cells each "blocked," or occupied, by two or more X men. Suppose F and S have all their men in adjacent primes, each separating the other from his 0 cell. If $(6, 6)$ is rolled forever by both players, no one moves and the game does not end.

Table 1 gives an alternative "more active" example.

Table 1. Initial configuration: all men borne off except 1 W at 22, 1 B at 3. All cell numbers are W coordinates. The cycle $W2 \cdots B4$ is repeated forever.

move no.	roll	choice
$W1$	$(3, 3)$	$W22 \to 10$
$B1$	$(4, 3)$	$B3 \to 10$, $W10 \to 25$
$W2$	$(2, 1)$	$W25 \to 22$
$B2$	$(2, 3)$	$B10 \to 15$
$W3$	$(4, 3)$	$W22 \to 15$, $B15 \to 0$
$B3$	$(2, 1)$	$B0 \to 3$
$W4$	$(2, 3)$	$W15 \to 10$
$B4$	$(4, 3)$	$B3 \to 10$, $W10 \to 25$
$W5$	$(2, 1)$	$W25 \to 22$

This leads to the question of whether the expected duration of the game is finite. Conceivably this could depend on which strategies the players choose.

Here, *but for a gap*, is a proof that the expected duration of the game is finite for arbitrary stationary strategies.[2]

Assume first that the values of the doubling cube are limited to the *finite* set $\{2^k,\ k = 0, 1, \ldots, n\}$. Then the number of possible states is finite. The absorbing states are those in which (a) all the men of the player who has just moved are in his 0 cell and not all the other player's men are in his zero cell. The transition probability P_{ij} from state i to state j depends only on the states i and j. The game is therefore a finite Markov chain with stationary transition probabilities. We conjecture (this is the gap in the proof) that all the other states are transient. It follows at once from this that the expected time to reach an absorbing state, and hence the expected duration of the game, is finite (e.g., Kemeny and Snell 1960, Theorem 3.1.1, p. 43.).

If instead the possible doubling cube values increase without limit ($2^k,\ k = 0, 1, 2, \ldots$), the same proof works under the additional hypothesis that for sufficiently large N, the strategy pair is the same for all $k \geq N$. Then we group together all states that have $k \geq N$ and are otherwise the same. Now apply the preceding "proof."

3 The number of backgammon configurations

We shall see that the number of possible configurations is enormous. Thus any attempt to solve the game by direct calculation would seem to be limited to parts of the end game. This does not, however, rule out solutions to larger portions of the game via theorem proving, or by efficient algorithms. Such solutions might be exact, approximate, or involve inequalities.

Some configurations cannot arise in the course of the game. The rules specify that no one of cells $1, 2, \ldots, 24$ can contain both Black and White men. The rules also specify that the game ends when one player's men are in his 0-cell, therefore the configuration where each player has all his men in his 0-cell is impossible.

The impossibility of other configurations may be deduced. For instance, suppose all men are on the bar. Whoever just moved had to, by the rules, attempt (successfully in this case) to move a man or men off the bar. This is a contradiction. As a more complex illustration, suppose both players have men on the bar and have primes on their own home boards. What happened on the last move? The player whose move it was had men on the bar and

[2] *Editor's note*: The gap can be filled. A standard result in Markov chains is that if state i is recurrent and leads to state j in a finite number of steps with positive probability, then state j is recurrent *and leads back to state i in a finite number of steps with positive probability.* In particular, if every nonabsorbing state leads to an absorbing state in a finite number of steps with positive probability, then the conjecture that every nonabsorbing state is transient is correct. So all that needs to be verified is that from any state there is a finite sequence of dice rolls that leads to one player or the other winning the game, assuming the specified strategy. This requires a bit of thought but can be verified to be correct.

attempted to move them. He did not succeed since his opponent had a prime after the move, hence before. Therefore the configuration did not change. By induction, the board configuration was always thus, contradicting the initial configuration specified in the rules. Hence these configurations are impossible.

If we know who has just moved, additional configurations are impossible. For instance, a player cannot complete his turn with all his men on the bar when the opponent has no blocked points on his home board.

Some questions about possible configurations: Suppose W has i men on the bar, B has j men on the bar, and X ($= W$ or B) is to move. Describe this by $(i, j; X)$. Which $(i, j; X)$ can arise in the course of the game? In particular what are the maximal such triples $(i, j; W)$ and $(i, j; B)$? Presumably all submaximal triples can also arise.

[Following a suggestion of the referee we find a simple upper bound for the number of configurations. Using the fact that the number of ways of placing r balls in n ordered cells is $\binom{r+n-1}{r}$ (Feller 1968), we can place the 15 B men in the 26 cells $\{0, \ldots, 25\}$ in $\binom{40}{15}$ ways. The same is true for the 15 W men. This includes all the possible configurations plus some impossible ones, such as those with B and W men in the same cell. The resulting number $M = \binom{40}{15}^2 \approx 1.62 \times 10^{23}$ is an upper bound for the number of configurations. Using the rule that opposing men may not occupy the same cell reduces this somewhat. The bound given in the original paper, via a more complex combinatorial calculation, was about 1.85×10^{19}.][3]

An upper bound for the number of states, assuming a limit of n values for the doubling cube and initial stakes of one unit, is then $2Mn$.

We turn now to the important special case of the pure running game.

Definition 5. *The* pure running game *(or* end game *or* no-contact game*) is that portion of the game where no further contact between men is possible, i.e., when there exists a real number x such that all W men have coordinates less than x and all B men have coordinates greater than x. (Here all coordinates are W coordinates.)*

The end game is finite and cannot take more than 63 moves for W, 62 for B. This case occurs if $x = 12.5$, all 15 B men are on 13, all 15 W men are on 12, the only subsequent rolls are $(1, 2)$ and B and W assemble all 15 men on $B1$ and $W1$ respectively, before bearing off. It follows that the game tree arising from each initial end-game position is *finite*.

To get an upper bound K for the number of end-game configurations, relabel the cells $1, \ldots, 26$.

---

[3] *Editor's note*: A better bound is

$$M = \sum_{i=1}^{15} \binom{26}{i} \binom{15-1}{i-1} \binom{15+26-i-1}{15} \approx 1.14 \times 10^{19}.$$

The three factors correspond to the number of ways to choose the i cells occupied by W, the number of ways to place the 15 W men in these i cells with at least one in each, and the number of ways to distribute the 15 B men into the remaining $26 - i$ cells.

(a) Pick $1 \le i < j \le 26$.
(b) Put 1 W man in cell i, put 14 W men in cells $1, \ldots, i$.
(c) Put 1 B man in cell j, put 14 B men in cells $j, \ldots, 26$. Then

$$K = \sum_{1 \le i < j \le 26} \binom{13+i}{14} \binom{40-j}{14} \approx 1.403 \times 10^{15}.$$

There is one (and possibly other) impossible end-game configuration: all W men in 0 and all B in 25. As before we believe this upper bound is a good approximation to the number of possible configurations.

According to the rules, the player can only bear off men (i.e., move them to his 0-cell) if all men not borne off are in the home board. Thus all 15 W men, e.g., are in cells 0 to 6 which can happen in $\binom{21}{15}$ ways. There are $\binom{21}{15}^2 \approx 2.945 \times 10^9$ configurations and $2n\binom{21}{15}^2 \approx 5.889 \times 10^9 n$ states, not all of which are possible, given n values for the doubling cube.

An interesting special problem is to find the optimal strategy and expectation for those states where each player has exactly one man left to bear off. [When B and W are out of contact there are $12 \cdot 23 = 276$ configurations, $276n$ positions, and $552n$ states. In these cases Model $1'$ and the actual game coincide and Section 5 below gives the complete solution.]

4 Solution to the Model 1 end game

Label Model 1 states $(i, j; W)$ with $i \ge 1$, $j \ge 0$ or $(i, j; B)$ with $i \ge 0$, $j \ge 1$, where i is the White goal, j is the Black goal, and W and B designate who is to move. States $(i, 0; W)$ and $(0, j; B)$ are terminal states in which B and W, respectively, have won. If $i, j \ge 0$ then a White move goes from state $(i, j; W)$ to $(\max(i-k, 0), j; B)$ and a Black move goes from state $(i, j; B)$ to $(i, \max(j-k, 0); W)$, where k is distributed according to the dice probabilities $A(k)$ and $3 \le k \le 24$. For reference $36A(k)$ is respectively $2, 3, 4, 4, 6, 5, 4, 2, 2, 1$ for $k = 3, \ldots, 12$; 1 for $k = 16, 20, 24$; and 0 otherwise.

[For the player who is to move next, the expected payoff for the game (assuming a unit stake) will be denoted for White by $E(i, j; W)$ and $E(i, j; B)$. These will be calculated recursively on a pair of lattices as indicated schematically in Figures 2 and 3.] Note that the probability $P(i, j; X)$ of White winning satisfies $P(i, j; X) = (E(i, j; X) + 1)/2$, where $X = W$ or B. The two lattices describe all Model 1 states. Then $E(i, j; W) = \sum_k E(\max(i-k, 0), j; B)A(k)$ and $E(i, j; B) = \sum_k E(i, \max(j-k, 0); W)A(k)$ subject to the boundary conditions $E(i, 0; W) = -1$ and $E(0, j; B) = 1$. Equivalently,

$$E(i, j; W) = \sum_{k=3}^{i-1} E(i-k, j; B)A(k) + \sum_{k=i}^{24} A(k) \tag{1}$$

and

$$E(i,j;B) = \sum_{k=3}^{j-1} E(i,j-k;W)A(k) - \sum_{k=j}^{24} A(k), \tag{2}$$

where an empty sum is taken to be zero.

To start the recursion, refer to Figures 2 and 3 and note that for zone (a1) $E(i,j;W) = 1$ for $i = 1,2,3$. Then to find values for zone (b1) we use (2) to find $E(i,j;B)$ for $i = 1,2,3$. Next to find values for zone (a2) calculate $E(i,j;W)$ for $i = 4,5,6$, then for zone (b2) find $E(i,j;B)$ for $i = 4,5,6$, etc. The matrices of values are readily determined with the aid of a computer.

The matrices will satisfy (error checks)

- $|E(i,j;X)| \le 1$ (the maximum gain or loss is 1)
- $E(i,j;W) = -E(j,i;B)$ (antisymmetry)
- $E(i,j;W) \ge E(i,j;B)$ (it is better for White if White is to move)
- $E(i,j;X) \ge E(i+m,j;X)$, $m \ge 0$ (it is better for White to be closer to his goal)
- $E(i,j;X) \le E(i,j+n;X)$, $n \ge 0$ (it is better for White if Black has farther to go).

The antisymmetry property allows us to replace (1) and (2) by

$$E(i,j;W) = -\sum_{k=3}^{i-1} E(j,i-k;W)A(k) + \sum_{k=i}^{24} A(k). \tag{3}$$

Then $E(i,j;W)$ is calculated recursively by finding (1) the first three rows, (2) the first three columns, (3) the next three rows, (4) the next three columns, etc.

We calculated $E(i,j;W)$ recursively to four places (some roundoff error in fourth place) by equations (1) and (2) for all i and j such that $1 \le i,j \le 168$. The calculation is extremely fast. Tables 2 and 3 show results rounded to two places. Table 4 shows a portion of the full table from which Tables 2 and 3 were extracted. [The rows and columns of Tables 2–4 vary monotonically so values not given can be obtained by interpolation or extrapolation.]

An alternative to the recursion calculation is to simulate a large number of Model 1 games on a computer, to determine the matrices of Figures 2 and 3. Equivalently we can determine the matrix $(P(i,j;F))$ of probabilities that the first player F will win given that his goal is i and that the goal of the second player S is j. Such a matrix is computed for selected (i,j) and given as Table 5. Note that an exact $(P(i,j;F))$ matrix follows from Table 2 and the equation $P(i,j;F) = (E(i,j;F)+1)/2$.

The entries in Table 5 are each based on $n = 10{,}000$ games. The standard deviation of the sample mean from the true mean $p = P(i,j;F)$ is $\sqrt{p(1-p)/n} \le 0.005$ so the third digit is in doubt. For large (i,j), simulation is a useful alternative to the recursion method. It provides values for a specified (i,j) pair without having to calculate values for any, not to say all, (i',j') pairs with $i' \le i$ and $j' \le j$. Further, $P(i',j;F) \ge P(i,j;F)$ if $i' \le i$ and

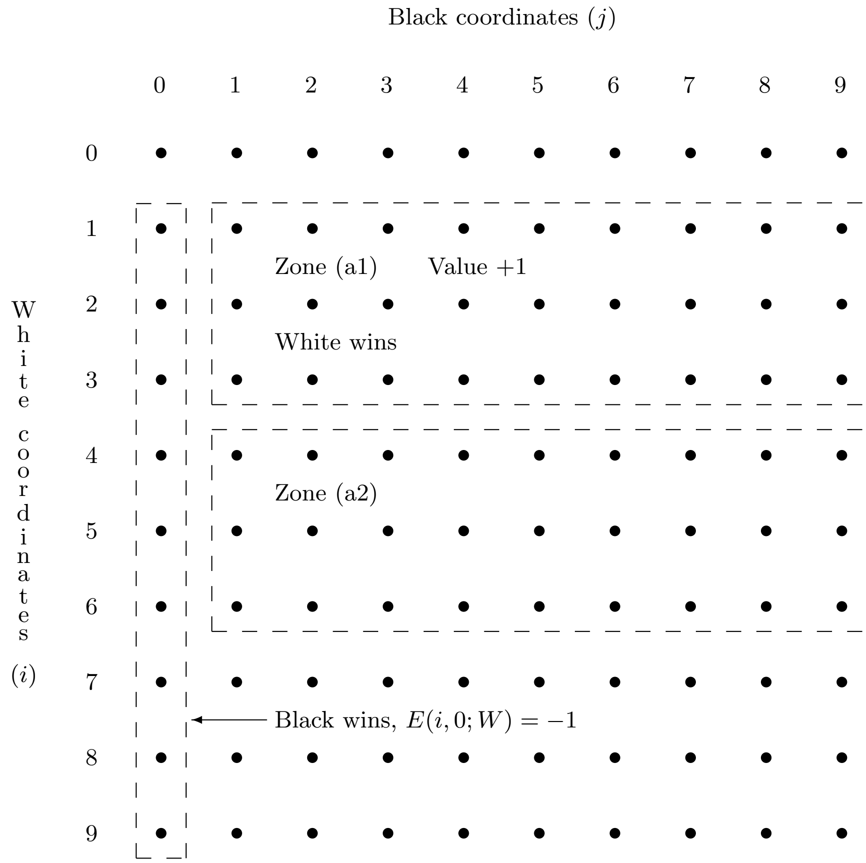

Fig. 2. Lattice of states $(i, j; W)$, $i \geq 1$, $j \geq 0$, White to move. Expectation for White is $E(i, j; W)$.

$P(i, j'; F) \leq P(i, j; F)$ if $j' \leq j$ so given $P(i', j'; F)$ and $P(i, j; F)$ we have $P(i', j'; F) \leq P(r, s; F) \leq P(i, j; F)$ for $i \leq r \leq i'$ and $j' \leq s \leq j$. Therefore a table of selected values provides useful information about the remaining values.

A table equivalent to Tables 2, 3, and 5 appears as Table 1 of Keeler and Spencer (1975). Their table was computed by simulation and the method was different than used for our Table 5. However, the agreement with our Table 5 and our exact Tables 2 and 3 is excellent. Their figures are off by at most 1 in the second place.

The first draft of this paper and the paper by Keeler and Spencer were each written without knowledge of the other. This paper was then revised to refer to their results.

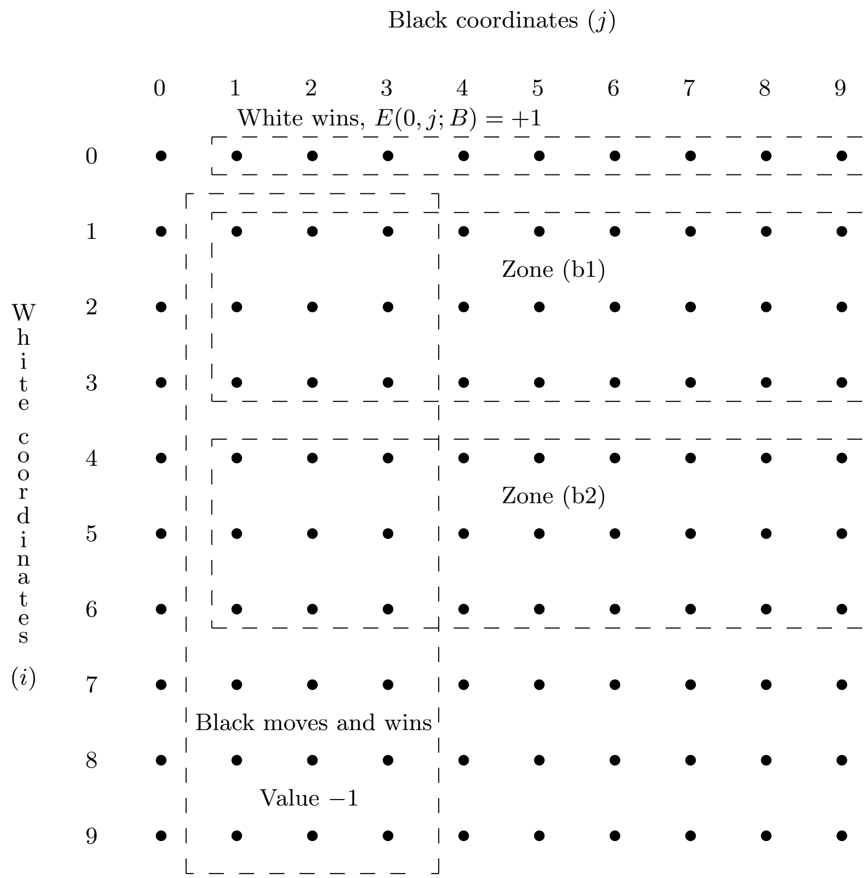

Fig. 3. Lattice of states $(i, j; B)$, $i \geq 0$, $j \geq 1$, Black to move. Expectation for White is $E(i, j; B)$.

The values in parentheses in Table 5 are those given by Jacoby and Crawford (1970, p. 118). There is a systematic discrepancy between the two sets of numbers, with the latter set of probabilities noticeably closer to 0.5 than ours. To "explain" this, note first that Jacoby and Crawford (1970, p. 119) say that they do not claim their values are "completely accurate" and that they are based on "experience." Thus their figures may have substantial errors. On the other hand, our values are for Model 1 and Model 1 neglects "blots." Model 1 also neglects wastage in bearing off. In the real game, with such large i, j, there will be on average several blots hit on each side. This will have the effect of increasing the effective i and j, very possibly enough to give values consistent with those of Jacoby and Crawford.

Table 2. $E(i, j; W)$ for Model 1 rounded to two decimal places. Decimal points is two places from right and has been omitted.

White	Black total																
total	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85
5	76	94	98	98	100	100	100	100	100	100	100	100	100	100	100	100	100
10	−36	56	81	90	99	99	100	100	100	100	100	100	100	100	100	100	100
15	−71	−6	40	73	88	94	98	99	100	100	100	100	100	100	100	100	100
20	−85	−59	−11	37	66	82	92	97	99	100	100	100	100	100	100	100	100
25	−97	−79	−51	−10	32	60	78	89	95	98	99	100	100	100	100	100	100
30	−99	−90	−74	−45	−7	29	56	74	86	93	97	99	99	100	100	100	100
35	−100	−97	−87	−68	−41	−7	26	52	71	84	91	96	98	99	100	100	100
40	−100	−99	−94	−83	−64	−37	−6	24	50	68	81	90	94	97	99	99	100
45	−100	−100	−98	−92	−80	−61	−35	−6	23	47	66	79	88	93	96	98	99
50	−100	−100	−99	−97	−90	−77	−58	−33	−5	22	45	63	77	86	92	96	98
55	−100	−100	−100	−98	−95	−87	−74	−55	−31	−5	21	43	61	75	84	91	95
60	−100	−100	−100	−99	−98	−94	−85	−72	−53	−30	−5	20	41	59	73	83	90
65	−100	−100	−100	−100	−99	−97	−92	−83	−69	−51	−28	−4	19	40	57	71	81
70	−100	−100	−100	−100	−100	−99	−96	−90	−81	−67	−49	−27	−4	18	39	56	70
75	−100	−100	−100	−100	−100	−99	−98	−95	−89	−79	−65	−47	−26	−4	18	37	54
80	−100	−100	−100	−100	−100	−100	−99	−97	−94	−87	−77	−63	−46	−25	−4	17	36
85	−100	−100	−100	−100	−100	−100	−100	−99	−97	−93	−86	−76	−62	−44	−25	−4	17
90	−100	−100	−100	−100	−100	−100	−100	−99	−98	−96	−92	−84	−74	−60	−43	−24	−4
95	−100	−100	−100	−100	−100	−100	−100	−100	−99	−98	−95	−90	−83	−72	−59	−42	−23
100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−99	−97	−94	−89	−82	−71	−57	−41
105	−100	−100	−100	−100	−100	−100	−100	−100	−100	−99	−99	−97	−94	−88	−80	−70	−56
110	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−99	−98	−96	−93	−87	−79	−68
115	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−99	−98	−96	−92	−86	−78
120	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−99	−97	−95	−91	−85
125	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−99	−99	−97	−94	−90
130	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−99	−98	−97	−94
135	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−99	−98	−96
140	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−99	−99	−98
145	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−99	−99
150	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−99
155	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100
160	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100
165	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100	−100

Table 3. Continuation of Table 2.

White	Black total															
total	90	95	100	105	110	115	120	125	130	135	140	145	150	155	160	165
5	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
15	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
20	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
25	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
30	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
35	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
40	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
45	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
50	99	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100
55	97	98	99	100	100	100	100	100	100	100	100	100	100	100	100	100
60	94	97	98	99	100	100	100	100	100	100	100	100	100	100	100	100
65	88	93	96	98	99	99	100	100	100	100	100	100	100	100	100	100
70	80	87	92	95	97	98	99	100	100	100	100	100	100	100	100	100
75	68	78	86	91	95	97	98	99	99	100	100	100	100	100	100	100
80	53	67	77	85	90	94	96	98	99	99	100	100	100	100	100	100
85	35	52	65	76	84	89	93	96	98	99	99	100	100	100	100	100
90	16	34	50	64	75	83	88	93	95	97	98	99	99	100	100	100
95	−4	16	34	49	63	73	82	88	92	95	97	98	99	99	100	100
100	−23	−4	15	33	48	62	72	80	87	91	94	96	98	99	99	100
105	−40	−22	−3	15	32	47	60	71	80	86	91	94	96	98	99	99
110	−55	−39	−21	−3	13	31	46	59	70	79	85	90	93	96	97	98
115	−67	−54	−38	−21	−3	14	31	46	58	69	78	84	89	93	95	97
120	−77	−66	−53	−37	−21	−3	14	30	45	57	68	77	84	89	92	95
125	−84	−76	−65	−52	−37	−20	−3	14	30	44	57	67	76	83	88	92
130	−89	−83	−75	−64	−51	−38	−20	−3	13	29	43	56	66	75	82	87
135	−93	−89	−82	−74	−63	−50	−35	−19	−3	13	28	43	55	65	74	81
140	−96	−93	−88	−81	−73	−62	−49	−35	−19	−3	13	28	42	54	65	73
145	−97	−95	−92	−87	−80	−72	−61	−48	−34	−19	−3	13	28	41	53	64
150	−98	−97	−95	−91	−86	−80	−71	−60	−47	−33	−18	−3	12	27	41	53
155	−99	−98	−97	−94	−91	−80	−79	−70	−59	−47	−33	−18	−3	12	27	40
160	−99	−99	−98	−96	−94	−90	−85	−78	−69	−58	−46	−32	−18	−3	12	26
165	−100	−99	−99	−98	−96	−93	−89	−84	−77	−68	−58	−45	−32	−17	−3	12

Table 4. Part of the four place table of $E(i, j; W)$ for Model 1. The decimal place is four places from the right and has been omitted. The complete table (49 pages like this one) covers all i, j from 1 to 168.

White	Black total											
total	25	26	27	28	29	30	31	32	33	34	35	36
25	3235	3914	4511	5080	5586	6034	6435	6828	7188	7503	7789	8061
26	2461	3174	3813	4423	4972	5465	5912	6345	6744	7096	7418	7721
27	1676	2414	3086	3732	4323	4860	5353	5828	6267	6657	7015	7350
28	871	1623	2323	3002	3633	4214	4753	5272	5752	6182	6576	6945
29	72	830	1548	2254	2920	3542	4124	4684	5205	5674	6104	6506
30	−706	48	776	1500	2194	2851	3471	4069	4627	5133	5599	6033
31	−1456	−713	15	751	1464	2147	2798	3429	4020	4560	5060	5527
32	−2183	−1456	−733	6	731	1433	2107	2765	3386	3957	4489	4988
33	−2865	−2161	−1450	−718	10	721	1411	2088	2731	3330	3890	4419
34	−3493	−2816	−2126	−1408	−688	23	719	1406	2066	2686	3271	3828
35	−4074	−3427	−2763	−2066	−1363	−661	32	723	1391	2027	2634	3215
36	−4616	−4001	−3365	−2696	−2014	−1329	−644	42	714	1360	1983	2585
37	−5119	−4537	−3933	−3294	−2639	−1974	−1305	−628	41	693	1327	1946
38	−5576	−5029	−4458	−3854	−3229	−2590	−1941	−1279	−617	33	674	1304
39	−6002	−5490	−4953	−4382	−3789	−3179	−2552	−1909	−1260	−615	27	663
40	−6401	−5922	−5418	−4882	−4321	−3739	−3137	−2515	−1882	−1247	−609	27
41	−6775	−6329	−5857	−5352	−4822	−4269	−3693	−3094	−2479	−1857	−1229	−597
42	−7115	−6701	−6260	−5787	−5288	−4764	−4214	−3640	−3047	−2442	−1827	−1205
43	−7429	−7045	−6635	−6192	−5722	−5226	−4704	−4155	−3586	−3001	−2403	−1795
44	−7722	−7367	−6985	−6571	−6129	−5661	−5165	−4642	−4098	−3536	−2957	−2366
45	−7992	−7665	−7310	−6923	−6509	−6067	−5597	−5101	−4581	−4043	−3487	−2915
46	−8233	−7933	−7604	−7245	−6857	−6442	−5999	−5530	−5036	−4523	−3990	−3440
47	−8450	−8176	−7873	−7540	−7179	−6790	−6374	−5931	−5465	−4977	−4469	−3941
48	−8648	−8398	−8121	−7813	−7477	−7115	−6725	−6309	−5869	−5407	−4924	−4419
	37	38	39	40	41	42	43	44	45	46	47	48
25	8316	8530	8726	8906	9075	9206	9318	9417	9509	9579	9637	9693
26	8005	8248	8471	8675	8866	9020	9152	9272	9381	9467	9540	9610
27	7663	7935	8184	8414	8628	8805	8960	9102	9230	9336	9426	9511
28	7288	7588	7865	8121	8359	8562	8741	8906	9056	9182	9291	9393
29	6879	7209	7514	7797	8061	8290	8495	8684	8857	9005	9135	9255
30	6437	6798	7132	7442	7733	7989	8221	8436	8633	8805	8957	9096
31	5962	6353	6718	7057	7376	7661	7920	8161	8384	8579	8754	8915
32	5454	5877	6272	6642	6991	7304	7592	7860	8108	8328	8526	8709
33	4917	5372	5799	6199	6577	6920	7236	7531	7805	8050	8273	8478
34	4354	4840	5298	5730	6137	6509	6854	7176	7476	7746	7993	8221
35	3768	4284	4773	5234	5671	6072	6445	6794	7119	7415	7687	7939
36	3163	3706	4223	4714	5179	5608	6009	6385	6736	7058	7354	7630
37	2544	3111	3654	4171	4663	5120	5547	5950	6326	6673	6995	7296
38	1917	2505	3069	3610	4126	4608	5062	5490	5891	6264	6611	6937
39	1286	1889	2472	3033	3570	4075	4553	5006	5432	5830	6202	6553
40	656	1268	1865	2442	2997	3524	4024	4500	4950	5373	5770	6145
41	32	649	1254	1843	2413	2957	3478	3975	4448	4895	5316	5716
42	−581	35	644	1240	1821	2380	2917	3434	3928	4397	4842	5265
43	−1181	−570	38	637	1225	1795	2347	2880	3393	3882	4349	4794
44	−1766	−1164	−561	38	629	1206	1769	2316	2845	3353	3839	4305
45	−2331	−1742	−1148	−554	37	618	1188	1746	2289	2812	3316	3801
46	−2875	−2301	−1719	−1133	−547	34	608	1174	1726	2263	2781	3282
47	−3397	−2841	−2273	−1698	−1119	−541	33	602	1161	1707	2238	2753
48	−3897	−3359	−2808	−2246	−1678	−1106	−535	34	597	1149	1689	2215

The exact calculation for $i = j = 167$ shows a Model 1 win probability for F of .5592 (compare Table 5 value .564). This suggests that the value

Table 5. Probabilities of success for first player in Model 1. Each figure based on 10,000 simulated games, assuming player who moves has goal $N(F)$ and other player has goal $N(S)$. Standard deviation of error is less than or equal to .005. Parenthetical numbers are corresponding probabilities given by Jacoby and Crawford (1970).

$N(F)$	$N(S)$													
	50	60	70	80	90	100	110	120	130	140	150	160	167	170	180
50	.610	.817	.930	.977	.995	.999	.999								
60	.346	.598	.801	.913	.969	.992	.998								
		(.545)		(.889)											
70	.169	.355	.592	.779	.902	.957	.986								
			(.524)		(.857)										
80	.068	.190	.365	.582	.764	.883	.953								
				(.512)		(.833)									
90	.019	.081	.201	.382	.578	.754	.875								
					(.508)		(.818)								
100	.007	.027	.094	.222	.383	.574	.737								
						(.506)		(.800)							
110	.001	.008	.037	.104	.225	.399	.568								
							(.505)		(.792)						
120								.571	.716	.841	.915	.957		.983	.993
130								.404	.574	.710	.822	.906		.956	.982
140								.259	.412	.564	.713	.822		.901	.955
167													.564		

of the actual game for F is positive. Note, however, that in the actual game $E(X) = 7$ for the initial roll rather than the 8 1/6 of subsequent rolls, whereas Tables 2, 3, and 5 assume 8 1/6 for all rolls. (This last is reasonable because Model 1 is meant to represent the end game.) Examination of the full Tables 2 and 3 suggests that changing the initial roll to conform with the actual game would reduce the F win probability to about .542, giving a value of .084 to E.

However, we cannot make useful inferences from this about the value (assuming both players choose optimal strategies) of the actual game to F. The additional features of the actual game make it substantially different from Model 1. Hitting blots, loss of some or all of a roll when men are sufficiently blocked, and "wastage" all tend to make $N(F)$ and $N(S)$ substantially greater than in Model 1. However, symmetry shows that $p = P(i, i; W) > .50$ for $i \geq 1$. Thus we might conjecture that the probability p that F wins satisfies $.55 > p > .50$ whence the value $V(F)$ of the game to F (without doubling or gammons) would satisfy $.10 > V(F) > 0$.

Doubling and gammons both should alter this considerably. If X doubles and Y folds it is an ordinary win. In scoring the actual game, there are three cases if no one folds. If X wins and some Y men have reached $Y0$ it is an *ordinary win* and X receives the value on the cube. If no Y men have reached $Y0$ and all Y men are in cells $Y1$ through $Y18$ it is a *gammon* and X gets twice the value on the cube. If no Y men have reached $Y0$ and some Y men are in $Y19$ through $Y25$ it is a *backgammon* and X receives three times the value on the cube.

Even without doubling or gammons, $V(F)$ is not easy to judge. We can't even tell if $V(F) > 0$. For instance, Cooke and Bradshaw (1974) classify initial rolls according to how good they judge them to be. There is fairly general agreement on the approximate classification and as to whether initial rolls are good or bad. The good rolls for F are 3-1, 4-2, and 6-1 which have joint probability 6/30, since the initial F roll in the actual game cannot be doubles. For the second player all first rolls of doubles, except double fives, are also good to very good, for a probability of 11/36 that S begins with a good roll. Thus the advantage of the initial roll is offset. But how much? Could the value of the game even be positive for the second player?

5 Solution to the Model 1′ end game

We extend the method to solving Model 1′, which we define as Model 1 extended to incorporate the doubling cube. The doubling cube is the most interesting, subtle, and least understood aspect of the game. Here for the first time we will learn things contrary to the beliefs of most players.

Either player may make the initial double. If one player doubles, the other player may accept or fold. If he accepts, the game continues for doubled stakes and the player who accepted now possesses the doubling cube, i.e., is the only player who can make the next (re)double. This must be done just prior to his roll by the player who possesses the cube. If the doubled or redoubled player folds, he loses the current stake prior to the double.

The states for Model 1′ are described by six lattices, $(i,j;X)$, $(i,j;X,D)$, and $(i,j;X,O)$, where $X = W$ or B. If the last coordinate is X, no one has yet doubled and therefore no one has the cube. If the last coordinate is D the player who is to move has the doubling cube; if it is O the other player has it. The moves of the game, i.e., the allowable transformations (between lattices), are:

a. given W
   White doubles: $(i,j;W)$ to $(\max(i-k,0),j;B,D)$
   White does not double: $(i,j;W)$ to $(\max(i-k,0),j;B)$
b. given B
   Black doubles: $(i,j;B)$ to $(i,\max(j-k,0);W,D)$
   Black does not double: $(i,j;B)$ to $(i,\max(j-k,0);W)$
c. given W,D
   White doubles: $(i,j;W,D)$ to $(\max(i-k,0),j;B,D)$ or
   White does not double: $(i,j;W,D)$ to $(\max(i-k,0),j;B,O)$
d. given W,O
   $(i,j;W,O)$ to $(\max(i-k,0),j;B,D)$

e. given B, D
   Black doubles: $(i, j; B, D)$ to $(i, \max(j-k, 0); W, D)$ or
   Black does not double: $(i, j; B, D)$ to $(i, \max(j-k, 0); W, O)$
f. given B, O
   $(i, j; B, O)$ to $(i, \max(j-k, 0); W, D)$

We calculate the expectation for White *per unit* currently (i.e., at point i, j) bet. A decision to double is made by comparing the expectation from doubling, per current unit, with that from not doubling. The expectation, if positive, is increased by doubling and decreased by the loss of the doubling cube. Thus whether to double will depend in general on the results of each calculation. In case (a), White's expectation if he doubles is $E^D \equiv E^D(i, j; W) = \min(1, \overline{E}^D)$, where $\overline{E}^D(i, j; W) = 2\sum_k E(\max(i-k, 0), j; B, D)A(k)$, and his expectation if he does not double is

$$E^0(i, j; W) = \sum_k E(\max(i-k, 0), j; B)A(k).$$

In case (c) White's expectation if he doubles is

$$E^D \equiv E^D(i, j; W, D) = \min\left(1, \overline{E}^D(i, j; W, D)\right),$$

where

$$\overline{E}^D(i, j; W, D) = 2\sum_k E(\max(i-k, 0), j; B, D)A(k),$$

and if he does not double, his expectation is

$$E^0(i, j; W, D) = \sum_k E(\max(i-k, 0), j; B, O)A(k).$$

We assume each player's goal is to maximize his own expectation. Then in cases (a) and (c) White should double if $E^D > E^0$, in which case we assign the value E^D to $E(i, j; W)$ or $E(i, j; W, D)$. He is indifferent if $E^D = E^0$ and he should not double if $E^D < E^0$. In these cases we set $E(i, j; W) = E^0$. Black should accept a double if $\overline{E}^D = E^D < 1$, he is indifferent if $\overline{E}^D = E^D = 1$, and should fold if $\overline{E}^D > 1$. Let $A(i, j) = \{m, n\}$ stand for these alternatives, where $m = 0, 1, 2$ according to whether White should not double, is indifferent, or should double, and $n = 0, 1, 2$ according to whether Black if doubled should fold, be indifferent, or accept a double. Then for cases (a) and (c) we record the $A(i, j)$ strategy matrices as the E matrices are calculated.

In case (d) White cannot double so

$$E(i, j; W, O) = \sum_k E(\max(i-k, 0), j; B, D)A(k).$$

The E matrices for Black to move are useful auxiliaries in the recursion but they can be deduced from those for White to move. This follows, just as in Model 1, from the analogous antisymmetries for Model 1$'$:

$$E(i,j;W) = -E(j,i;B), \tag{4}$$
$$E(i,j;W,D) = -E(j,i;B,D), \tag{5}$$
$$E(i,j;W,O) = -E(j,i;B,O). \tag{6}$$

For reference, the recursion equations for Black to move are:
Case (b):

$$E(i,j;B) = \min(E^0, E^D),$$

where

$$E^0 \equiv E^0(i,j;B) = \sum_k E(i, \max(j-k,0); W)A(k)$$

and

$$E^D = \max(-1, \overline{E}^D)$$

with

$$\overline{E}^D(i,j;B) = 2\sum_k E(i, \max(j-k,0); W, D)A(k).$$

Case (e):

$$E(i,j;B,D) = \min(E^0, E^D),$$

where

$$E^0 \equiv E^0(i,j;B,D) = \sum_k E(i, \max(j-k,0); W, O)A(k)$$

and

$$E^D = \max(-1, \overline{E}^D)$$

with

$$\overline{E}^D(i,j;B,D) = 2\sum_k E(i, \max(j-k,0); W, D)A(k).$$

Case (f):

$$E(i,j;B,O) = 2\sum_k E(i, \max(j-k,0); W, D)A(k).$$

It is persuasive that it is at least as good, and generally better, to possess the doubling cube than not to. This ought to be true in the standard game as well as in the various simpler models we discuss. This is generally taken for granted by players but a formal proof remains to be given.

We give a proof for Model $1'$, in which case it is equivalent to the simultaneous inequalities

$$E(i,j;W,D) \geq E(i,j;W) \geq E(i,j;W,O) \tag{7}$$
$$E(i,j;B,D) \geq E(i,j;B) \geq E(i,j;B,O) \tag{8}$$

Theorem 2. *Inequalities (7) and (8) hold for Model $1'$, i.e., it is always at least as good to have the doubling cube as not to.*

Proof. For $i = 1, 2, 3$ all parts of (7) are 1. This corresponds to zone (a1) of Figure 2, in each of the lattices $\{(i, j; W, D)\}$, $\{(i, j; W)\}$, and $\{(i, j; W, O)\}$. Now let $i = 1, 2, 3$ and suppose Black follows the (not necessarily optimal) strategy of not doubling when in states $(i, j; B, D)$ or $(i, j; B)$. Then his expectation is, respectively, $-\sum_k E(i, \max(j-k, 0); W, O)A(k)$ or $-\sum_k E(i, \max(j-k, 0); W)A(k)$, whereas the expectation of a Black player in state $(i, j; B, O)$ is $-\sum_k E(i, \max(j-k, 0); W, D)A(k)$. Comparing term by term and using what was proved about "zone (a1)" establishes (8) for $i = 1, 2, 3$, i.e., for the regions of $(i, j; B, y)$, $y = D$ or blank or O, corresponding to zone (b1) of Figure 3. We continue with zone (a2), zone (b2), etc. just as in the original recursion. This concludes the proof. □

The player who is doubled should accept the double if as a result his expectation $E > -1$, he is indifferent if $E = -1$, and he should reject the double if $E < -1$. In this last instance the game ends and the other player receives $+1$ so this value, not $-E$, is assigned to the appropriate state for use in subsequent recursions. Thus for each state in each of the matrices $\{(i, j; W, D)\}$ and $\{(i, j; B, D)\}$ we compute E^D and E^0. Then we will list whether the player doubles (-1 = NO, 0 = INDIFF, $+1$ = YES), whether the doubled player accepts (-1 = NO, 0 = INDIFF, $+1$ = YES), and the expectation E.

Tables 6 and 7 are selected portions of four-place tables for $E(i, j; W, D)$ and $E(i, j; W, O)$, respectively. [A comparison of corresponding entries in the two tables shows the enormous value of the doubling cube.]

As a check, the various Model $1'$ tables were shown to satisfy (4), (5), (6), (7), (8) as well as monotonicity:

$$E(i, j; X, Y) \geq E(i + k, j; X, Y) \tag{9}$$

$$E(i, j; X, Y) \leq E(i, j + k; X, Y). \tag{10}$$

and that a player with three or less steps wins on the next move:

$$E(i, j; W, Y) = 1 \text{ if } i = 1, 2, 3 \tag{11}$$

$$E(i, j; B, Y) = -1 \text{ if } j = 1, 2, 3. \tag{12}$$

As a further test we have:

Lemma 1. *If* $0 < E(i, j; W, O) < 1/2$*, then* $E(i, j; W, D) \geq 2E(i, j; W, O)$ *and* W *should double.*

Proof. If W cannot double, the next state will be of the form $(m, n; B, D)$. If W can and does double, the next state will also be of this form with the same probabilities of m, n. Thus W gets $0 < E(i, j; W, O) < 2E(i, j; W, O) < 1$ by doubling.

However (W, D) can choose not to double in which case the next state is $(m, n; B, O)$. It can and does happen that this sometimes gains more (see Tables 6 and 7 for many (i, j) examples) hence $E(i, j; W, D) > 2E(i, j; W, O)$ may occur. □

The optimal strategies were computed directly as previously described. However an optimal strategy can be derived from the E tables. To illustrate, we do this using Tables 6 and 7 for the case when White has the doubling cube and is to move. The procedure is to analyze cases (let $E(i,j;W,D) = E_D$, $E(i,j;W,O) = E_O$), as in Table 8.

Applying these cases to Table 6 produces the shaded area. It also is another check on the tables.

Keeler and Spencer (1975) approach the Model 1′ problem by replacing the actual dice with a random variable having values 5, 10, and 20 with probabilities .55, .36, and .09. (The text suggests that their Model 1 simulation also used this simplification. However the difference in probabilities between Table 6 and 7 suggests otherwise.) These probabilities are "... set to make the mean = 8 1/6 and the variance = 19, which are the real backgammon values." Actually the correct variance is $18.47\overline{2}$. This comes from the formula $(\sum_i x_i^2 - n\overline{x}^2)/n$. The incorrect value 19 arises for the actual dice if one uses the formula $(\sum_i x_i^2 - n\overline{x}^2)/(n-1)$, which is the appropriate formula for the different situation of independent sampling from a population whose mean is *unknown*.

Replacing the actual dice by the "5, 10, 20" dice greatly simplifies the problem: only (i,j) multiples of 5 need to be considered. Thus 170×170 matrices become 34×34 matrices. The matrix storage sizes are reduced by a factor of 25. Also the recursion to calculate each entry involves sums of only three terms, rather than 14 so we have another reduction by a factor of almost 5. Thus the total calculation is reduced by a factor of more than 100. Yet the Keeler–Spencer results are an excellent approximation.

For instance the probabilities for equal counts in their Table 2 are off by at most one in the second place. When the opponent has 5 more points they are within 2 in the second place (all the errors here are overestimates). When the opponent has 10 more points their figures overestimate by 1 to 4 in the second place.

The correct Model 1′ strategy for doubling, redoubling, and folding, has a somewhat irregular pattern. Gillogly and Keeler (1975) give a simplified rule based, presumably, on the Keeler–Spencer calculations:

Approximate rule (Gillogly–Keeler): With pip count C let $D = C$ if $C \leq 25$ and $D = 1.1C$ if $C > 25$. Then make the first double if the opponent's count $j \geq D - 2$. Redouble if $j \geq D - 1$. The opponent should fold if (doubled and) $j \geq D + 2$.

Prior to this paper, Keeler and Spencer (1975), and Gillogly and Keeler (1975), the rule recommended by experts is the one given by Jacoby and Crawford (1970):

Approximate rule (Jacoby–Crawford): With a pip count of C double if $j \geq 1.075C$; redouble if $j \geq 1.1C$; the opponent should fold if $j \geq 1.15C$.

Obolensky and James (1969) gives no rule; Lawrence (1973) and Genud (1975) repeat the rule without acknowledgment (a common practice in gaming circles). Stern (1973) gives a complex and unclear set of maxims.

Table 6. Model 1′ selected portion of table of $E(i, j; W, D)$ values (White expectation, White to move, White has doubling cube). White should double if and only if (i, j) is within or above the shaded area. Black should accept unless (i, j) is above the shaded area, in which case he should fold.

White	Black total															
total	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
30	5068	5897	7001	8234	9354	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000
31	4130	5018	5860	6850	8060	9168	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000
32	3180	4096	4991	5819	6680	7872	8971	9985	10000	10000	10000	10000	10000	10000	10000	10000
33	2253	3170	4083	4955	5761	6499	7679	8774	9778	10000	10000	10000	10000	10000	10000	10000
34	1367	2266	3171	4057	4901	5693	6423	7494	8576	9594	10000	10000	10000	10000	10000	10000
35	526	1396	2279	3156	4011	4837	5631	6352	7309	8401	9427	10000	10000	10000	10000	10000
36	−273	563	1418	2274	3123	3959	4775	5556	6282	7147	8245	9277	10000	10000	10000	10000
37	−1028	−227	595	1427	2257	3088	3913	4723	5495	6223	7004	8102	9121	10000	10000	10000
38	−1736	−972	−183	618	1427	2243	3063	3879	4675	5445	6171	6871	7953	8969	9927	10000
39	−2403	−1676	−920	−148	635	1431	2237	3046	3846	4635	5399	6120	6780	7805	8820	9777
40	−3030	−2339	−1617	−873	−115	657	1444	2239	3031	3821	4599	5352	6062	6720	7660	8673
41	−3620	−2961	−2272	−1558	−826	−76	688	1464	2242	3021	3797	4561	5300	6003	6660	7518
42	−4168	−3543	−2885	−2202	−1498	−774	−32	722	1481	2245	3011	3772	4519	5248	5946	6601
43	−4680	−4087	−3462	−2808	−2134	−1437	−721	10	749	1495	2246	2997	3742	4477	5198	5892
44	−5164	−4600	−4004	−3382	−2736	−2068	−1378	−672	45	772	1506	2243	2978	3711	4439	5151
45	−5621	−5083	−4516	−3923	−3306	−2665	−2003	−1323	−629	78	794	1515	2238	2962	3686	4403
46	−6047	−5539	−4999	−4433	−3845	−3233	−2598	−1943	−1274	−588	109	814	1523	2236	2951	3664
47	−6443	−5965	−5455	−4918	−4357	−3773	−3165	−2537	−1890	−1227	−550	138	833	1534	2237	2942
48	−6812	−6365	−5886	−5379	−4845	−4287	−3706	−3104	−2482	−1840	−1183	−512	167	854	1546	2239

Table 7. Model 1′ selected portions of four place table of $E(i, j; W, O)$ values (White expectation when White is to move and Black has the doubling cube). The last digit in this table and in Table 6 may be off by one because of machine round-off errors.

White	Black total															
total	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
30	2088	2819	3500	4117	4677	5190	5659	6091	6486	6858	7200	7519	7802	8063	8303	8520
31	1246	2028	2760	3425	4030	4584	5093	5561	5990	6393	6768	7118	7433	7725	7996	8242
32	374	1201	1980	2691	3340	3936	4485	4993	5457	5894	6302	6685	7032	7355	7656	7933
33	−511	351	1169	1922	2612	3249	3840	4387	4889	5361	5804	6219	6599	6953	7285	7592
34	−1396	−513	336	1124	1853	2529	3160	3747	4288	4797	5274	5723	6136	6521	6883	7220
35	−2292	−1385	−518	299	1063	1777	2447	3074	3654	4201	4714	5196	5641	6057	6449	6815
36	−3212	−2268	−1383	−546	248	997	1705	2370	2990	3573	4122	4638	5116	5563	5984	6378
37	−4157	−3174	−2253	−1400	−584	196	938	1641	2298	2918	3502	4051	4561	5038	5488	5911
38	−5094	−4097	−3138	−2249	−1420	−618	155	891	1583	2238	2855	3436	3977	4484	4963	5414
39	−5994	−5029	−4043	−3112	−2254	−1437	−641	123	846	1533	2182	2793	3365	3902	4410	4889
40	−6830	−5923	−4965	−3993	−3096	−2253	−1443	−658	92	807	1485	2125	2727	3294	3830	4336
41	−7570	−6760	−5848	−4889	−3951	−3072	−2243	−1445	−674	66	769	1436	2065	2661	3225	3759
42	−8169	−7499	−6679	−5751	−4822	−3903	−3042	−2230	−1447	−688	38	729	1384	2006	2598	3159
43	−8605	−8117	−7420	−6574	−5672	−4751	−3854	−3013	−2221	−1451	−706	6	685	1333	1951	2537
44	−8911	−8571	−8060	−7318	−6486	−5594	−4686	−3811	−2991	−2215	−1459	−728	−28	643	1285	1897
45	−9119	−8892	−8531	−7968	−7231	−6400	−5521	−4626	−3773	−2974	−2212	−1469	−751	−59	605	1240
46	−9261	−9106	−8866	−8462	−7886	−7147	−6322	−5455	−4571	−3741	−2960	−2211	−1480	−771	−86	570
47	−9352	−9255	−9089	−8819	−8403	−7809	−7072	−6250	−5390	−4524	−3714	−2948	−2210	−1487	−786	−111
48	−9435	−9352	−9250	−9065	−8780	−8349	−7742	−7004	−6178	−5328	−4482	−3687	−2935	−2206	−1493	−802

Table 8. Analysis of cases.

case	optimal strategy	value of E_D
$E_O < 0$	do not double	$\geq E_O$
$E_O < 0$, $E_D \neq 0$	do not double	> 0
$E_O = 0$, $E_D = 0$	any	0
$\frac{1}{2} > E_O > 0$, $E_D = 2E_O$	double; B should accept	$1 > E_D = 2E_O > 0$
$\frac{1}{2} > E_O > 0$, $E_D > 2E_O$	do not double	$1 \geq E_D > 2E_O > 0$
$E_O = \frac{1}{2}$	double; not doubling* may or may not also be optimal; folding* is optimal for Black; accepting may or may not also be.	1
$E_O > \frac{1}{2}$	double; not doubling* may or may not also be optimal; B should fold if doubled.	1

*Doubling gives the opponent the chance to err; not doubling does not, so if both are optimal doubling is recommended.

Both the J–C and G–K strategies are quite close to the exact results for Model $1'$. Note that in the actual game the effective i and j will tend to be greater than the pip counts, due to "wastage" in bearing off. This tends to slightly increase the double, redouble and fold numbers in Table 5.3. We expect subsequent more realistic calculations to give an increase of 1 just before some of the change-over points. For instance, when W has 43 White might redouble at 5 instead of 4 and Black might wait until 8 to fold instead of 7. Of course the use of the corrected count (Section 6, below) is a more precise way to incorporate this wastage effect.

Table 9 compares the approximate rules with the exact results if $1 \leq C \leq 65$, for the case when White redoubles and Black considers accepting or folding. This region of small C is the most irregular, and even here the G–K approximate rule is generally good.

For fixed White count i, let D_i be the least Black count j for which White should double. Let R_i be the least Black count for which White should redouble, and let F_i be the least Black count for which White should fold. It is plausible to expect the $D_i \leq R_i \leq F_i$ for all i. The last inequality corresponds to the obvious fact that there is nothing (i.e., no expectation) to be gained by folding without being redoubled. Similarly we prove $D_i \leq F_i$. Note however that if a player's expectation is -1, whence he has no probability of winning he may as well fold voluntarily to save time unless his opponent is a far from optimal player.

This establishes the first part of the following theorem.

Table 9. Comparison of Gillogly–Keeler approximate rule and exact rule, for redoubling (R) and folding (F), if $1 \leq C \leq 65$.

W has	W: R if B has			B: F if B has		
	Exact	G–K	J–C	Exact	G–K	J–C
1–6	Any	−1	1	Any	2	1
7	−6	−1	1	1	2	2
8	−1	−1	1	1	2	2
9	−1	−1	1	0	2	2
10	−2	−1	1	0	2	2
11–12	−2	−1	2	0	2	2
13	−2	−1	2	1	2	2
14–18	−1	−1	2	2	2	3
19	−1	−1	2	3	2	3
20	0	−1	2	3	2	3
21–22	0	−1	3	3	2	4
23–25	1	−1	3	4	2	4
26	1	2	3	4	5	4
27	1	2	3	5	5	5
28–30	2	2	3	5	5	5
31	2	3	4	5	6	5
32–33	2	3	4	6	6	5
34–37	3	3	4	6	6	6
38	3	3	4	7	6	6
39–40	4	3	4	7	6	6
41–43	4	4	5	7	7	7
44–46	5	4	5	8	7	7
47–50	5	4	5	8	7	8
51	5	5	6	8	8	8
52–53	6	5	6	9	8	8
54–59	6	5	6	9	8	9
60	6	5	6	10	8	9
61–65	7	6	7	10	9	10

Theorem 3. *It is an optimal strategy to fold only when the expectation would be less than* -1 *by not folding. With this folding strategy and any optimal redoubling strategy,*

$$R_i \leq F_i \textit{ for all } i.$$

The inequality $D_i \leq R_i$ *holds for all* i *for either of the following optimal strategy cases:*

(a) White redoubles only when he gains by doing so (i.e., he does not redouble if he is indifferent). He doubles whenever he gains and he doubles or not, as he pleases, when he is indifferent.

(b) White redoubles whenever he gains and does as he pleases when he is indifferent. He doubles whenever he gains or is indifferent.

Proof. Note first that the qualifications (a) or (b) are needed because in general the optimal strategies are not unique. For doubling, redoubling, and folding there are lower and upper numbers, $\underline{D}_i \le \overline{D}_i$, $\underline{R}_i \le \overline{R}_i$, and $\underline{F}_i \le \overline{F}_i$, such that the player is indifferent if and only if $\underline{D}_i \le j < \overline{D}_i$, etc. If these indifference zones overlap, there exist optimal strategies which violate one or both of the inequalities $D_i \le R_i \le F_i$. The proof that $D_i \le R_i$ uses the conditions of cases (a) and (b) to rule out these overlaps.

To establish that $D_i \le R_i$ under conditions (a) or (b) it suffices to prove $\underline{D}_i \le \underline{R}_i$ and $\overline{D}_i \le \overline{R}_i$. To prove $\overline{D}_i \le \overline{R}_i$, note that it is equivalent to showing that

$$\min\left(1, \overline{E}^D(i,j;W,D)\right) > E^0(i,j;W,D)$$

implies

$$\min\left(1, \overline{E}^D(i,j;W)\right) > E^0(i,j;W).$$

This is equivalent to asserting that

$$\min\left(1, 2\sum_k E(\max(i-k,0),j;B,D)A(k)\right) > \sum_k E(\max(i-k,0),j;B,O)A(k)$$

implies

$$\min\left(1, 2\sum_k E(\max(i-k,0),j;B,D)A(k)\right) > \sum_k E(\max(i-k,0),j;B)A(k)$$

or equivalently that

$$2\sum_k E(\max(i-k,0),j;B,D)A(k) > \sum_k E(\max(i-k,0),j;B,O)A(k)$$

implies

$$2\sum_k E(\max(i-k,0),j;B,D)A(k) > \sum_k E(\max(i-k,0),j;B)A(k).$$

But this is equivalent to

$$\sum_k E(\max(i-k,0),j;B,D)A(k) \ge \sum_k E(\max(i-k,0),j;B)A(k).$$

This follows from $E(i,j;B,O) \ge E(i,j;B)$ for all i,j, which in turn follows from (8). Thus $\overline{D}_i \le \overline{R}_i$.

To show $\underline{D}_i \le \underline{R}_i$ we repeat the proof using $\ge$ throughout in place of $>$. This concludes the proof of the theorem. □

An ambiguity that appears frequently in the previous work (e.g., in Jacoby and Crawford 1970, Gillogly and Keeler 1975) is to give the doubling, redoubling, and folding thresholds in terms of the probability of winning by the player who is to move. For example Gillogly and Keeler say that the probability of winning with $i = j = 30$ is .65, and that the cutoffs for D, R, F are respectively probabilities of .65, .69, 75. It seems clear from their text and from comparison with our results that they mean Model 1 probabilities. Note that the Model 1′ probabilities of winning with optimal strategies will in general be different, and may even be multi-valued because they depend on the player's choice of strategy, and even the optimal strategies may be nonunique. In general, the probability of winning may depend strongly on strategy. As an extreme example to make the point, suppose in an actual game both Black and White always double. White always accepts if doubled, but Black always folds. Then the probability is one that White wins.

As an instance from the actual game, suppose White has only one man left, on $W6$ and Black has only one man left, on $B3$. White is to move and has the cube. Then $E(6, 3; W) = 1/2$ (Model 1) because the probability W wins in the next roll is 3/4. Otherwise Black wins.

White should double, giving $\overline{E}^D = 1$. Then if Black plays optimally, he is indifferent as to whether he accepts or folds because in either case his expectation is -1. Suppose Black (by means of a chance device) accepts with probability p and folds with probability $1 - p$. Then he wins the game with probability $p/4$, and White wins with probability $1 - p/4$, $0 \le p \le 1$. Hence, even though both players choose an optimal strategy, the White win probability may assume any value from .75 to 1 depending on the Black choice of optimal strategy. It follows that (a) optimal strategies are nonunique in all the versions of the game which we consider, (b) for any position which can lead to the above position, this is also true, and (c) in all versions of the game that we consider, for any position that can lead to the above position (i.e., in "most" positions), the win probability is not uniquely determined even if the players limit themselves to optimal strategies.

6 Pending research

In a future paper we plan to present exact strategies for bearing off. With these we shall extend the results to Model 2, Model 2′ (= Model 2 plus doubling cube), Model 3 (Model 2 with bearing off allowed only when all men have reached the inner board), and Model 3′ (Model 3 with the doubling cube). Model 3′ is the pure running game with the actual rules. The solutions involve relatively minor corrections or "perturbations" of the Model 1 and Model 1′ solutions. For instance, Gillogly and Keeler (1975) give approximate rules for estimating the "wastage" of an inner board position. This is the amount that should be added to the (Model 1, 1′) pip count i due to the fact that in Models 3 and 3′ some parts of the dice counts may be wasted in bearing off.

Approximate rule for wastage (Gillogly–Keeler): Suppose all n men are on the inner board and cover m points. Let k men be on the one point. Then add w to i where $w = k + (n-1) + (n-m)$.

Thus there are penalties for men on the one point, for extra men, and for extra men clustered on a point.

The rule is not used consistently by its authors. In one example they compare 5 men on the 6 point with 10 men on the one point. The wastage is 8 for the 5 men on the 6 point and 28 for the 10 men on the one point. The adjusted count should be $i = j = 38$. But in the text only the difference of 20 is added to j, giving $i = j = 30$. In our opinion, their "netting out" method is not a correct application of their rule.

The methods of this paper apply mutatis mutandis to numerous variants of the game (Bell 1969). Some earlier forms of the game coincide with Models 1 or $1'$ and are thus completely solved.

I wish to thank Steven Mizusawa for his programming of the computations in this paper.

[Author's afterword: This article was reported on in Thompson (1975) and Thorp (1975). The exact solution to a special case of Model $3'$, the pure running game with the doubling cube, where each player has borne off all but just one or two men in their home boards, appeared in Thorp (1978a,b,c), and was reprinted in Thorp (1984). The exact Model $3'$ solution to pure races when each player is bearing off from his home board has since been computed for hundreds of millions of cases, including up to 9 men for each player, by Hugh Sconyers, a two-time world team backgammon player. See `http://www.back-gammon.info/sconyers/hypergammon.html`. Zadeh and Kobliska (1977) published an important article on optimal doubling in backgammon. Thorp (1977) covers the relation between their paper and this article, as well as unpublished work by Gillogly, Kahn and Smolen.

The original version of this article has been modified slightly for publication here. I have incorporated suggestions from the referee and the editor, as well as some changes of my own. These preserve the historical sense and smooth the reader's path. The longer revisions are set off in brackets.]

References

1. Bell, R. C. (1960) *Board and Table Games from Many Civilizations, Vol. 1.* Oxford University Press, London.
2. Bell, R. C. (1969) *Board and Table Games from Many Civilizations, Vol. 2.* Oxford University Press, London.
3. Cooke, Barclay and Bradshaw, Jon (1974) *Backgammon: The Cruelest Game.* Random House.
4. Feller, William (1968) *An Introduction to Probability and Its Applications, Vol. 1*, third edition. Wiley, New York.

5. Genud, Lee (1974) *Lee Genud's Backgammon Book*. Cliff House Books, Los Angeles.

6. Gillogly, J. J. and Keeler, E. B. (1975) Playing the running game in backgammon. *Popular Bridge* **9** (3) 34–38.

7. Holland, Tim (1973) *Beginning Backgammon*. McKay, New York. Tartan paperback, 1974.

8. Hopper, Millard (1972) *Win at Backgammon*. Dover. Unabridged reproduction of the 1941 original.

9. Jacoby, Oswald and Crawford, John R. (1970) *The Backgammon Book*. The Viking Press, New York.

10. Keeler, Emmett B. and Spencer, Joel (1969) Proper raising points in a generalization of backgammon. The Rand Corporation, 4078 (May).

11. Keeler, Emmett B. and Spencer, Joel (1975) Optimal doubling in backgammon. *Operations Research* **23** (6) (Nov.–Dec.) 1063–1071. [Reprinted in Levy (1988).]

12. Kemeny, John G. and Snell, J. Laurie (1960) *Finite Markov Chains*. Van Nostrand, Princeton, NJ.

13. Lawrence, Michael S. (1973) *Winning Backgammon*. Pinnacle Books, New York.

14. [Levy, David N. L., editor (1988) *Computer Games I*. Springer-Verlag, New York.]

15. Neumann, J. von and Morgenstern, Oskar (1964) *Theory of Games and Economic Behavior*. Wiley (Science Editions).

16. Obolensky, Prince Alexis and James, Ted (1969) *Backgammon, the Action Game*. Macmillan, Collier Books.

17. Stern, Don (1973) *Backgammon, the Quick Course to Winning Play*. Cornerstone Library, Simon and Schuster distributors, Drake Publishers, New York.

18. Thomsen, Dietrick (1975) Beating the game. *Science News* **107** (12) (Mar. 22) 198–199.

19. Thorp, Edward O. (1975) Recent results and open questions for some particular games. Preliminary Report. Paper delivered at American Math. Soc. Annual Meeting, Washington, D.C., Jan. 1975. See *Notices, A.M.S.*, Jan. 1975, Abstract No. 720-90-7, page A-254.

20. [Thorp, Edward O. (1977) Review of Zadeh and Kobliska. *Mathematical Reviews* H449 362.]

21. [Thorp, Edward O. (1978a) End positions in backgammon. *Gambling Times* (Sept.). Reprinted in *Gammon Magazine* preview issue. Reprinted in Levy (1988).]

22. [Thorp, Edward O. (1978b) End positions in backgammon II. *Gambling Times* (Oct.). Reprinted in Levy (1988).]

23. [Thorp, Edward O. (1978c) End positions in backgammon III. *Gambling Times* (Nov.). Reprinted in Levy (1988).]

24. [Thorp, Edward O. (1984) *The Mathematics of Gambling*. Gambling Times, Incorporated.]

25. [Zadeh, Norman and Kobliska, Gary (1977) On optimal doubling in backgammon. *Management Science* **23** (8) (April) 853–858. Reprinted in Levy (1988).]