Fall 1999

Joe P. Buhler (jpb@msri.org)
Elwyn Berlekamp (berlek@math.berkeley.edu)

Note: These solutions are a work in progress; comments, references, etc. are appreciated. Most references and figures can be found with the problem statements.

Problem 1. In contemplating the increase in average rental rates from 1998 to 1999 in 5 Bay Area cities, Jackie Blue finds that in each city the average rent in 1999 was at least 10% larger than the average rent in 1998. Prove or disprove: The average rent in those five cities taken together in 1999 was at least 10% larger than the average rent in 1998.

Discussion: The assertion is false. The easiest way to see this is to recast the statement with 0% instead of 10% everywhere (so instead of "at least 10% larger" the premise says simply "larger".) Even if all rents increase everywhere, if the population in the low-rent areas increases faster, that might tilt the average enough to yield a decrease in the overall average.

Here is a counterexample. Let the cities be A, B, C, D, and E, with the statistics as tabulated below; here # denotes the number of renters in thousands, and $ denotes the monthly rent.

City	1998 #	1999 #	1998 $	1999 $
A	20	20	1,500	1,520
B	20	20	1,500	1,520
C	20	20	1,500	1,520
D	20	20	1,500	1,520
E	100	120	960	985

In 1998 the total rental income is $4 \cdot 20 \cdot 1500 + 100 \cdot 960 = 216000$ and the average rent is therefore $216000/180 = \$1200$. The 1999 average rent is $(4 \cdot 20 \cdot 1520 + 120 \cdot 985)/200 = \$1999$, and we see that the average rent has actually gone down.

This initially counterintuitive result is known as Simpson's paradox, and the reader can get a sense of the vast literature on it, and its innumerable variants, by searching for it on the web. The curious aspects of this "paradox" have been noticed in a number of real-world contexts.

Problem 2. Find the 98-th digit to the right of the decimal point in the decimal expansion of $(\sqrt{2}+1)^{500}$.

Discussion: The number $x := (1-\sqrt{2})^{500} + (1+\sqrt{2})^{500}$ is an integer. The first term is positive and very small; easy estimates show that it is less than 10^{-100}. It follows that $10^{98}(1+\sqrt{2})^{500}$ is just below an integer, and in fact within .001 of an integer, so the asked-for digit is a 9.

Problem 3. Your wealthy father-in-law gives you $10,000 dollars to place a "double-or-nothing" bet on the Boston Red Sox in the world series. In other words, your father-in-law expects to receive $20,000 if the Sox win, and $0 if they lose.

To your horror, you then discover that the only kind of bets that are allowed at your neighborhood casino are double-or-nothing wagers on individual games. (For simplicity, we assume that the casino accepts bets up until the start of each game, and that the house cut is negligible.) How much should you bet on the first game? I.e., what strategy should you follow so that your final outcome is certain to be equivalent to an overall double-or-nothing bet on the series. (We remind readers that the world series is a "best-of-seven" event in which two teams play a series of games and the first team to win four games is the winner.)

Discussion: For nonnegative integers j, k, not both 0, let $M(j,k)$ denote the amount of money that we currently need to satisfy our possible obligations when we are in the state that our team (the Red Sox) will win the series if it wins k games, and the opponent will win they win j games. For notational simplicity, let $x = \$20{,}000$ be our initial state, so that $M(j,0) = x$, i.e., we need \$20,000 if our team wins. Similarly, we need nothing when our team loses, i.e., $M(0,k) = 0$. If it comes down to a one game series, its clear that we can need $x/2$ dollars (conveniently, the stake that our father-in-law has given us), and we obviously must bet it all, i.e., $M(1,1) = x$.

With a little more thought about a best-of-three series, we find that $M(1,2) = x/4$, $M(2,1) = 3x/4$, and $M(2,2) = x/2$.

In general, we need to fill in the following table of values $M(j,k)/x$:

$j \backslash k$	4	3	2	1	0
4					1
3					1
2			1/2	3/4	1
1			1/4	1/2	1
0	0	0	0	0	

If we are at state (j,k) and wager w then our team will either win, in which case we move to state $(j, k-1)$ are our capital increases to $M(j,k) + w$, or our team will lose, in which case we move to state $(j-1, k)$ are our capital decreases to $M(j,k) - w$. Adding the equations

$$M(j,k) + w = M(j,k-1), \qquad M(j,k) - w = M(j-1,k) - w$$

shows that $M(j,k)$ is the average of the values below and to the right of it in the table, and it shows that the wager amount is half the difference of those values.

We use this recursion to fill in the table, working along SW/NE diagonals.

$j \backslash k$	4	3	2	1	0
4	64/128	42/64	26/32	15/16	1
3	22/64	16/32	11/16	7/8	1
2	6/32	5/16	4/8	3/4	1
1	1/16	1/8	1/4	1/2	1
0	0	0	0	0	

We write the terms with constant powers of two in the denominators, on diagonals, so that the reader might guess that the $M(j,k)/x$ is 2^{-j-k} times a cumulative sum of binomial coefficients (i.e., removing denominators, and taking differences gives Pascal's triangle; more precisely,

$$M(j,k) = 2^{-j-k} \sum_{i=0}^{j} \binom{j+k}{i}$$

which is easily proved by induction from the recursion.

So, if we need \$20K when our team wins the series, we'll need 64/128 of this, or 10K when the series begins, and the initial wager is the difference between \$10K and $\$13,125 = (42/64) \cdot \$20K$ that we'll need if our team wins the first game, i.e., \$3,125.

Problem 4. The aforementioned father-in-law gives you \$10,000 to play in the following card game. The cards of a standard deck are turned over one by one. Each card then remains face up until all are turned over. Before each card you can bet any fraction of your current capital on a "double-or-nothing" bet on the color of the next card. How should you bet?

Discussion: (**First solution.**) Let $C(b, r)$ be the amount of current capital needed to **guarantee** that, with b remaining black cards and r remaining red cards, and using an optimal strategy, we will have \$1 when the deck is exhausted, no matter what might be the order of the remaining cards. Then the initial condition of our recurrence is the terminal condition of the game, namely $C(0, 0) = 1$. Next, $C(0, 1) = C(1, 0) = 1/2$, because on the last card we can bet all of our current capital and double our money.

In general, $C(b, r) = (C(b-1, r) + C(b, r-1))/2$, because by betting $|C(b-1, r] - C(b, r-1)|/2$ on the more likely color, we can ensure meeting our new capital requirements no matter whether we win or lose the current bet. This is the same recursion that occurred in the previous problem, and we find that

$$C(b, r) = 2^{-b-r}\binom{b+r}{b}.$$

So to be certain of ending with \$1, our starting capital will need to be

$$2^{-52}\binom{52}{26}.$$

By Stirling's formula, this is very close to $1/\sqrt{26\pi} \simeq 1/9.04$. So, starting with \$1, we can ensure finishing with a few pennies over \$9. Or, if we need to end with \$1, we'll need a small fraction of a penny more then eleven cents at the start.

What fraction of our current capital, C, should we bet that the next card is red? Plugging into our formula and simplifying yields the elegant answer, $(R-B)/(R+B)$. When this is zero, we don't bet, and when it is negative, we bet on black. In general, the better the odds, the more we will bet.

(**Second solution.**) There are

$$N := \binom{52}{26}$$

permutations of the colors in the deck of cards. They are all equally likely. So suppose we divide our initial capital, C, into N strategies, each beginning with C/N dollars. Then one of our large collection of strategies will get every color correct, and double its value 52 times, yielding a final payoff of $2^{52}C/N$.

Remark: This problem illustrates an important investment concept: It's very important to strive not only for returns with high expectation, but also for low variance. In this peculiar problem, good strategy attains a variance which is zero, while also maximizing the expected value.

Problem 5. Prove that any automorphism of a finite group of order n has order strictly less than n.

Discussion: This appears in a paper Horosevskii in *Math. Sb.* **93** (1974), along with other results on orders of automorphisms of finite groups. Marty Isaacs has a somewhat easier proof that will appear here at some point.

Problem 6. At the end of the column, the following diagram was given, together with the caption "Move and Win."

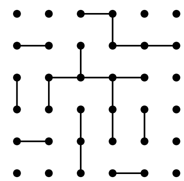

Figure 1: The problem

Discussion: Figure 1 shows a problem in the familiar children's game called Dots-and-Boxes. This game began with an array of 6×6 dots, and 18 moves have already been made. Each move consists of a unit vertical or horizontal line between an adjacent pair of dots. Altogether, on this board there are 60 moves, and the game will end when they have all been made.

Whenever a player's move completes the fourth side of a unit square, called a box, then he must place his initial in that box and make another move. If this move completes another box, he again places his initial in that box and makes another move. If two boxes are completed on the same move, the player gets only one extra move.

As explained in the first chapter of the reference (see last page), the most important elementary insight into this game is that long chains of three of more boxes behave quite differently than short chains of only one or two boxes. In a game between two good players, the decisive part of the game is typically a battle over the parity of the number of long chains. On boards with an odd number of boxes, including the 25-box board which appears in this problem, the first player strives for an even number of long chains, and the second player strives for an odd number of long chains.

Since no boxes have yet been completed in Figure 1, it is evident that all moves have been on separate turns. So 18 turns have been taken, and it is now the first player's turn to move. He should seek an even number of long chains.

The first 18 moves have already separated a small two-box region in the Northeast. These 18 moves have also already created two long chains, which are labeled "A" and "B" in all subsequent figures. As play on either of these long chains is formally "loony" the 23 boxes which are neither in the Northeast nor in any of the 6 boxes already in chains A and B are naturally partitioned into two regions: the West and the Southeast.

As seen in Figure 2a, there is a possibility of a third chain "C" in the West. This chain can be created with the three dotted moves shown in Figure 2a. But since two of these moves complete the box "X" these three moves will entail only two completed turns. As shown in Figure 2b,

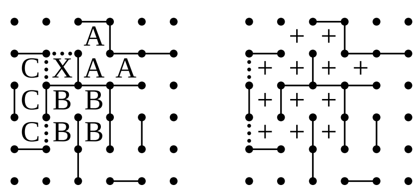

Figure 2: Dangers in the West (+ stands for AB)

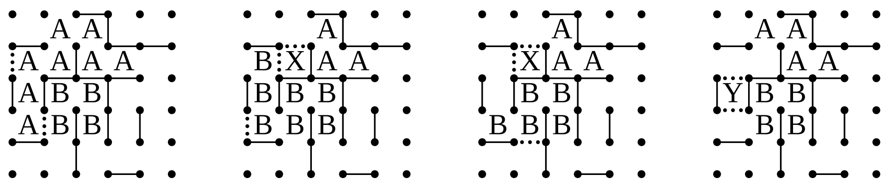

Figure 3: Western dangers averted

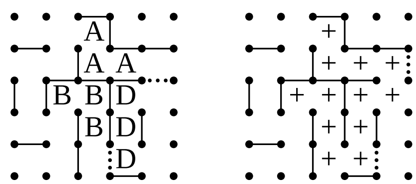

Figure 4: Dangers in the Southeast (+ stands for AB)

there is also the possibility of chains A and B becoming joined in the West. This also entails two complete turns, each consisting of a single move shown as a dotted line in Figure 2b.

The most straightforward way to win the position of Figure 1 is to ensure that the long chains A and B grow separately, and that no more long chains are created. From this perspective, the positions shown in Figure 2 should be regarded as dangers. If, from Figure 1, we were to play any of the dotted moves shown in Figure 2, the opponent could win the game by responding with another dotted move in the same subfigure. So in Figure 1, we should play elsewhere, such as an innocuous move in the Northeast. If the opponent then makes any of the dotted moves in Figure 2, we should regard this as a threat, which we can avert by playing in the West to one of the local configurations shown in Figure 3.

Figure 4 shows that there are similar dangers lurking in the Southeast. We should also refrain from an initial play on any of the four dotted moves shown in this region, because if we did, the opponent could then win by creating the local configuration of Figure 4a or Figure 4b. If instead the opponent begins play in the Southeast with one of the four dotted moves shown in Figure 4, we can thwart his threat with an appropriate local response as in Figure 5.

Of course our opponent might also choose to play innocuous moves, which refrain from doing anything relevant in either the West or in the Southeast. But if he tries this, we can achieve our objective in the West by making the first move there which sacrifices the box Y as shown in Figure 3d. We can also achieve our goal in the Southeast by playing the first move there in between the two boxes labeled “Z” in Figure 5d. This move sacrifices both of those two boxes,

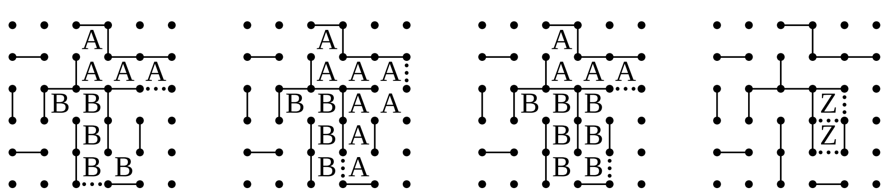

Figure 5: Southeastern dangers averted

but solidifies our two, and only two, long chains at A and B.

So this problem has many solutions.

Those who have read several chapters of the reference may choose to view this as a nimstring problem. The value of Figure 1 is evidently $*$. Any of the many moves to value 0 will win. The only losing moves are those to bigger numbers.

The reason this problem is so tractable is that the position of Figure 1 already decomposes into three pieces: the West, the Southeast, and a trivial third component in the Northeast. Combinatorial game theory provides the tools to solve such problems by doing a relatively detailed analysis of each of the pieces separately, followed by a relatively simple global analysis which combines the analyses of the local pieces.

Reference: Berlekamp: *The Dots-and-Boxes Game: Sophisticated Child's Play*, A K Peters, Ltd. An earlier and much less complete version of this chapter appears as Chapter 16 in volume 3 of *Winning Ways* by Berlekamp, Conway, and Guy, also published by A. K. Peters.