Outwitting the Lying Oracle

ROBB T. KOETHER
JOHN K. OSOINACH, JR.
Hampden-Sydney College
Hampden-Sydney, VA 23943
rkoether@hsc.edu
josoinach@hsc.edu

Lamentations 2:14 "Your prophets have seen for you false and foolish visions; ... they have seen for you false and misleading oracles."
(*New American Standard Bible*)

At the Delphi Casino an oracle operates a table where gamblers place bets on coin flips. The gamblers win or lose the amounts they bet, depending on whether they correctly predict the outcomes of the coin flips. As you approach the table, the oracle says to you, "I know how the coin will land each time and I am willing to tell you, but I must warn you, I will try to win your bet by occasionally lying to you." This does not strike you as a very promising game, but after some negotiation, the oracle agrees to lie no more than once during the next three coin flips, provided that before each flip you first tell the oracle the amount of your wager.

The question is: How should you place your bets on the three coin tosses so that you win the greatest amount of money in the end, no matter what the oracle does and no matter what the coin tosses are? We assume that the oracle is always agreeable to any amount that you wish to wager, but you cannot wager more than you currently possess.

We first encountered this problem in an article in *Scientific American* [**3**]. A very similar, but more general, problem appeared as Problem 10801 in the *American Mathematical Monthly* [**2**], along with its solution [**1**]. We gave this problem as a "Problem of the Fortnight" at Hampden-Sydney College, where we assumed you began with \$100 and the coin was flipped three times. One student solved the problem in the following manner (slightly paraphrased):

"The greatest amount of money that you can be guaranteed to receive, regardless of what the oracle does and regardless of what the coin flips are, is \$200: You should bet \$50 on the first flip and agree with the oracle's prediction. If the oracle lies, then you will still have \$50 left, but will correctly guess the remaining two flips for \$200; if the oracle is truthful, then you will have \$150. On the next flip again bet \$50 and agree with the oracle's prediction. If the oracle lies, then you have \$100 with one flip remaining, which you will guess correctly for \$200; if the oracle is truthful, then you will still have \$200 and will bet \$0 on the final flip."

While this answer is not entirely rigorous, the key ideas are present: No matter what the oracle does and what the results of the tosses are, if you always agree with the oracle, you have a strategy that guarantees that you double your initial amount.

This answer raised some interesting questions: How does the solution change if we increase the number of flips and allow the oracle to lie more than once? Can you outwit the oracle by disagreeing with the oracle's prediction? Or, stated differently, is there a strategy by which you could expect to win more than the maximum *guaranteed* outcome? Furthermore, is it possible to use the size of the bet to influence the oracle either to lie or tell the truth?

In this article, we first analyze the original game, but with any number of flips, followed by a simple generalization where the oracle may lie more than once. We then investigate the problem of trying to outwit the oracle, that is, finding strategies that give you the best chance for a better expected outcome, if possible, as well as strategies that the oracle should employ to minimize your chances for a better outcome. The mathematics involves relatively straightforward applications of game theory and probability, leading to some interesting results.

Believing the oracle

Our initial strategy will be always to agree with the oracle's prediction and make our bets on the basis of that strategy. We will start by solving the basic problem, where the oracle may lie at most once, and then allow the oracle to lie multiple times.

Multiple flips, one lie We begin with a restatement of the basic problem.

The Lying Oracle Problem: The oracle agrees to flip the coin a specified number of times and to predict the outcome accurately, except for possibly one lie. Before each prediction, you may bet any amount up to your current holdings. The oracle will then announce the outcome, after which you must state the outcome on which you wish to bet. How should you place your bets for the coin tosses so that you win the greatest amount of money in the end, no matter what the oracle does and no matter what the coin tosses are?

Solution: Using the terminology of game theory, we will henceforth refer to you, the bettor, as "the player."

Let w_n represent the proportion of the player's current holdings that the player should wager when there are n flips remaining in order to optimize the final outcome, and let A_n be the ratio of the player's final winnings to the current holdings, when the player wagers the optimal amounts on the remaining n flips.

If the oracle tells the truth on the first of the remaining n flips, then the player has the proportion $1 + w_n$ of the player's current holdings. The player must continue to place bets cautiously since the oracle may still lie. Thus, the final proportion of the player's winnings would be $(1 + w_n)A_{n-1}$.

On the other hand, if the oracle lies, then the player has the proportion $1 - w_n$, but now the player is free to bet the maximum amount on all $n - 1$ remaining flips, thereby doubling the player's money each time. In that case, the final proportion would be $(1 - w_n)2^{n-1}$.

We wish to find the value of w_n that will make these two expressions equal, nullifying the effect of the oracle's lie. That is, we wish to find the value of w_n such that

$$(1 - w_n)2^{n-1} = (1 + w_n)A_{n-1}. \tag{1}$$

This common value will be the value of A_n. In particular,

$$A_n = (1 - w_n)2^{n-1}. \tag{2}$$

Now by applying the above reasoning again, we see that $A_{n-1} = (1 - w_{n-1})2^{n-2}$. Substituting this into (1) produces

$$(1 - w_n)2^{n-1} = (1 + w_n)(1 - w_{n-1})2^{n-2}.$$

Solving for w_n yields the recurrence relation

$$\begin{aligned} w_1 &= 0, \\ w_n &= \frac{1 + w_{n-1}}{3 - w_{n-1}}, \quad n \geq 2. \end{aligned}$$

(The condition $w_1 = 0$ follows from the observation that with one flip and one lie, the player cannot guarantee a correct prediction; hence, the player should wager nothing.) An easy induction shows that the solution to this relation is

$$w_n = \frac{n-1}{n+1}, \quad n \geq 1. \tag{3}$$

A formula for A_n is obtained by substituting the expression for w_n given in equation (3) into equation (2). We summarize everything we have learned in the following theorem.

THEOREM 1. *If the oracle has not yet lied and there are n coin tosses remaining, then the player should bet* $(n-1)/(n+1)$ *of the player's current amount of money. However, if the oracle has lied and therefore cannot lie again, the player should bet everything. In either case, the player's final amount will be exactly* $2^n/(n+1)$ *times the player's current holdings.*

Multiple flips, multiple lies A natural generalization of this problem is to allow the oracle to lie more than once, so suppose that the oracle may lie up to k times during the coin flips. If the player's strategy is to continue to agree with the oracle's prediction, how should the player place the bets now so that the player again gets the greatest amount of money in the end?

In this case, we are dealing with a family of games, $G_{n,k}$, where $G_{n,k}$ represents the game of n coin flips and at most k lies. Thus, if the oracle tells the truth, then play proceeds to game $G_{n-1,k}$, and if the oracle lies, then play proceeds to game $G_{n-1,k-1}$. FIGURE 1 shows how the games proceed, beginning with 4 flips and 2 possible lies by the oracle. Note that game $G_{i,i}$ is always followed by game $G_{i-1,i-1}$, for all $i \geq 1$.

$$\begin{array}{ccccccccc}
G_{4,2} & \xrightarrow{T} & G_{3,2} & \xrightarrow{T} & G_{2,2} & \xrightarrow{T} & G_{1,1} & \xrightarrow{T} & G_{0,0} \\
{\scriptstyle L}\downarrow & & {\scriptstyle L}\downarrow & & {\scriptstyle L}\downarrow & & {\scriptstyle L}\downarrow & & \\
G_{3,1} & \xrightarrow{T} & G_{2,1} & \xrightarrow{T} & G_{1,1} & \xrightarrow{T} & G_{0,0} & & \\
{\scriptstyle L}\downarrow & & {\scriptstyle L}\downarrow & & {\scriptstyle L}\downarrow & & & & \\
G_{2,0} & \xrightarrow{T} & G_{1,0} & \xrightarrow{T} & G_{0,0} & & & &
\end{array}$$

Figure 1 The game tree starting with 4 flips and 2 lies

Let $w_{n,k}$ represent the proportion of the player's current holdings that the player should wager in game $G_{n,k}$ in order to optimize the final winnings, and let $A_{n,k}$ represent the ratio of the final winnings to the player's current holdings, provided the player wagers the optimal amounts in game $G_{n,k}$ and all succeeding games. Let us call $w_{n,k}$ the *critical wager*.

If the player continues to believe the oracle's predictions, then whenever k equals n, the player should bet \$0 from that stage on, as the oracle could lie every time, giving the player no opportunity to recover from a loss. Notice also that the games $G_{n,0}$ simply double the player's money with each coin toss, and that the games $G_{n,1}$ were analyzed above.

THEOREM 2. *For all $n \geq 1$ and for all k, $0 \leq k \leq n$, in the game $G_{n,k}$,*

$$A_{n,k} = \frac{2^n}{\sum_{i=0}^{k} \binom{n}{i}} \tag{4}$$

and

$$w_{n,k} = \frac{\binom{n-1}{k}}{\sum_{i=0}^{k} \binom{n}{i}}. \tag{5}$$

Proof. We will first establish a recurrence relation among the numbers $A_{n,k}$. Consider the first of n flips. If the oracle has told the truth, then the player would win $w_{n,k}$ of the player's current holdings on that flip. If the oracle has lied, then the player would lose the proportion $w_{n,k}$ on that flip. In the first case, the player's final winnings would be $(1 + w_{n,k})A_{n-1,k}$ and in the second case it would be $(1 - w_{n,k})A_{n-1,k-1}$.

In order to maximize the player's guaranteed winnings, these two amounts should be equal. Setting them equal and solving for $w_{n,k}$ yields

$$w_{n,k} = \frac{A_{n-1,k-1} - A_{n-1,k}}{A_{n-1,k-1} + A_{n-1,k}}. \tag{6}$$

It follows that

$$A_{n,k} = (1 + w_{n,k})A_{n-1,k} = \left(1 + \left(\frac{A_{n-1,k-1} - A_{n-1,k}}{A_{n-1,k-1} + A_{n-1,k}}\right)\right) A_{n-1,k}$$
$$= \frac{2A_{n-1,k-1}A_{n-1,k}}{A_{n-1,k-1} + A_{n-1,k}}.$$

We see from this equation that $A_{n,k}$ is the harmonic mean of $A_{n-1,k-1}$ and $A_{n-1,k}$, that is,

$$\frac{1}{A_{n,k}} = \frac{1}{2}\left(\frac{1}{A_{n-1,k-1}} + \frac{1}{A_{n-1,k}}\right). \tag{7}$$

We will use equation (7) to establish (4) by induction.

First, it is clear that $w_{n,0} = 1$, since the player will bet the full amount if the player knows that the oracle will not lie, and that $w_{n,n} = 0$, since the player will bet nothing if the oracle cannot be counted on to tell the truth at least once. It follows that $A_{n,0} = 2^n$ and $A_{n,n} = 1$, for all $n \geq 1$.

Thus, equation (4) holds for all n when $k = 0$ or $k = n$. In particular, it holds for all k, $0 \leq k \leq n$, when $n = 0$ or $n = 1$. We proceed by induction on n. Let us assume that equation (4) is correct for all k, $0 \leq k \leq n$, for some $n \geq 1$, and consider $A_{n+1,k}$, for some k where $0 < k < n + 1$. We complete the induction by computing

$$\frac{1}{A_{n+1,k}} = \frac{1}{2}\left(\frac{1}{A_{n,k-1}} + \frac{1}{A_{n,k}}\right) = \frac{1}{2}\left(\frac{\sum_{i=0}^{k-1}\binom{n}{i}}{2^n} + \frac{\sum_{i=0}^{k}\binom{n}{i}}{2^n}\right)$$
$$= \frac{\sum_{i=0}^{k-1}\binom{n}{i} + \sum_{i=0}^{k}\binom{n}{i}}{2^{n+1}} = \frac{1 + \sum_{i=1}^{k}\left(\binom{n}{i-1} + \binom{n}{i}\right)}{2^{n+1}}$$
$$= \frac{\sum_{i=0}^{k}\binom{n+1}{i}}{2^{n+1}}.$$

Now the reader can easily use (6) to verify that

$$w_{n,k} = \frac{\binom{n-1}{k}}{\sum_{i=0}^{k} \binom{n}{i}}. \qquad \blacksquare$$

TABLES 1 and 2 give the values of $A_{n,k}$ and $w_{n,k}$ for $1 \le n \le 7$ and $0 \le k \le 6$. It is interesting to note that the same solution was obtained by Pudaite [**2**], where the assumption was equivalent to the oracle's lying *exactly* k times in n coin flips.

TABLE 1: Table of final winnings $A_{n,k}$

n \ k	0	1	2	3	4	5	6
1	2	1	1	1	1	1	1
2	4	$\frac{4}{3}$	1	1	1	1	1
3	8	$\frac{8}{4}$	$\frac{8}{7}$	1	1	1	1
4	16	$\frac{16}{5}$	$\frac{16}{11}$	$\frac{16}{15}$	1	1	1
5	32	$\frac{32}{6}$	$\frac{32}{16}$	$\frac{32}{26}$	$\frac{32}{31}$	1	1
6	64	$\frac{64}{7}$	$\frac{64}{22}$	$\frac{64}{42}$	$\frac{64}{57}$	$\frac{64}{63}$	1
7	128	$\frac{128}{8}$	$\frac{128}{29}$	$\frac{128}{64}$	$\frac{128}{99}$	$\frac{128}{120}$	$\frac{128}{127}$

TABLE 2: Table of critical wagers $w_{n,k}$

n \ k	0	1	2	3	4	5	6
1	1	0	0	0	0	0	0
2	1	$\frac{1}{3}$	0	0	0	0	0
3	1	$\frac{2}{4}$	$\frac{1}{7}$	0	0	0	0
4	1	$\frac{3}{5}$	$\frac{3}{11}$	$\frac{1}{15}$	0	0	0
5	1	$\frac{4}{6}$	$\frac{6}{16}$	$\frac{4}{26}$	$\frac{1}{31}$	0	0
6	1	$\frac{5}{7}$	$\frac{10}{22}$	$\frac{10}{42}$	$\frac{5}{57}$	$\frac{1}{63}$	0
7	1	$\frac{6}{8}$	$\frac{15}{29}$	$\frac{20}{64}$	$\frac{15}{99}$	$\frac{6}{120}$	$\frac{1}{127}$

Note that if the player does not bet the critical wager $w_{n,k}$ at each stage, then the oracle can follow a pure strategy that guarantees the player's final outcome to be less than if the player had bet the critical wager. If the player bets more than the critical wager, then the oracle will lie, reducing the amount the player has to start with for the next game by more than the critical wager. Likewise, if the player bets less than the critical wager, then the oracle will tell the truth, which will increase the amount the player has to start with for the next game by less than the critical wager. Thus, by betting an amount different from the critical wager, the player can induce the oracle to lie or be truthful, but always at a disadvantage to the player, provided the player continues to believe the oracle.

Outwitting the oracle

The above analysis makes two crucial assumptions: the player will always agree with the oracle, and the oracle knows that the player will always agree. These assumptions ensure that the player will never receive less than the guaranteed amount, no matter what the oracle does or how the coin is flipped, but they also guarantee that the player will never receive more than that amount. But what if the player suspects that the oracle is lying? Can the player expect to increase the final winnings by *not* agreeing with the oracle? Indeed, can the player induce the oracle to lie by betting a large amount, and then win that amount by disagreeing with the oracle? As we investigate this possibility, we will also assume that the oracle now suspects that the player may disagree.

A single flip Let's begin with a simple example. Suppose we have exactly one flip and the oracle has one lie. If the oracle knows that the player will always agree with the oracle's prediction, then the oracle will lie if the player bets any amount at all. However, if the player is unpredictable—the player may choose to disagree—is it to the player's advantage to bet some amount? Is there a strategy for betting a certain wager so that the player's *expected* payoff is more than the amount guaranteed by the previous analysis? After all, in this game the player following the previous strategy would bet nothing.

This game may be modeled by a simple two-by-two matrix, where the entries represent the payoffs for the player. The rows indicate the player's two strategies (Agree or Disagree), while the columns represent the oracle's two strategies (tell the Truth or tell a Lie). Hence, in this example, we have the following payoff matrix for the player:

$$\begin{array}{c} \\ \text{Agree} \\ \text{Disagree} \end{array} \begin{array}{c} \begin{array}{cc} \text{Truth} & \text{Lie} \end{array} \\ \begin{pmatrix} 1+w & 1-w \\ 1-w & 1+w \end{pmatrix} \end{array},$$

where w is the proportion wagered (whether optimal or not). Let p_T and p_L be the probabilities that the oracle will tell the truth or lie, respectively. Similarly, let p_A and p_D be the probabilities that the player will agree or disagree with the oracle, respectively. In order to decide which strategy to pursue, the player computes the expected payoff of each row of the payoff matrix; the player then chooses the strategy (row) whose expected payoff is the greater of the two. The player's expected payoff of agreeing with the oracle is $p_T(1+w) + p_L(1-w)$; likewise, the expected payoff of disagreeing with the oracle is $p_T(1-w) + p_L(1+w)$.

On the other hand, the oracle's optimal strategy occurs when these two expected payoffs are equal. Setting the expected payoffs from the two rows equal yields

$$p_T(1+w) + p_L(1-w) = p_T(1-w) + p_L(1+w).$$

Solving for p_T and p_L gives $p_T = p_L$ (assuming that $w > 0$). Since $p_T + p_L = 1$, we have that $p_T = p_L = 1/2$. Substituting these values into the player's expected payoff from the first row gives us an expected payoff of 1. A similar calculation yields $p_A = p_D = 1/2$. Thus, the player cannot expect to do any better in this case than in the original scenario.

Multiple flips A similar analysis works in general. Let $E_{n,k}$ denote the expected payoff of the game $G_{n,k}$ when the player and the oracle employ their optimal strategies. In this game, the payoff matrix is

$$\begin{array}{c} \\ \text{Agree} \\ \text{Disagree} \end{array} \begin{array}{c} \begin{array}{cc} \quad\text{Truth} & \qquad\quad \text{Lie} \end{array} \\ \begin{pmatrix} (1+w)E_{n-1,k} & (1-w)E_{n-1,k-1} \\ (1-w)E_{n-1,k} & (1+w)E_{n-1,k-1} \end{pmatrix}, \end{array}$$

where w is the proportion of the wager. The strategies adopted by the players depend, of course, on the values of w, $E_{n-1,k}$, and $E_{n-1,k-1}$.

LEMMA 1. *In the game $G_{n,k}$, for all $n \geq 1$ and for all k, $1 \leq k \leq n$,*

$$E_{n,0} = 2^n, \tag{8}$$

$$E_{n,n} = 1, \tag{9}$$

$$E_{n,k} < E_{n,k-1}, \tag{10}$$

$$E_{n,k} = \frac{2E_{n-1,k-1}E_{n-1,k}}{E_{n-1,k-1} + E_{n-1,k}}. \tag{11}$$

Proof. We will establish (8) and (9) first. The game $G_{1,0}$ is trivial. The oracle must tell the truth and the player will agree. Therefore, $E_{1,0} = 2$. Notice also that we have already analyzed the game $G_{1,1}$ and found that $E_{1,1} = 1$.

In the game $G_{n,0}$, the oracle must always tell the truth, which gives the player a pure strategy of agreeing with the oracle each time. Thus, the player's optimal strategy is to wager the entire amount and will therefore double the amount wagered each time. Hence we have $E_{n,0} = 2^n$.

The game $G_{n,n}$ has payoff matrix

$$\begin{array}{c} \\ \text{Agree} \\ \text{Disagree} \end{array} \begin{array}{c} \begin{array}{cc} \quad\text{Truth} & \qquad\quad \text{Lie} \end{array} \\ \begin{pmatrix} (1+w)E_{n-1,n-1} & (1-w)E_{n-1,n-1} \\ (1-w)E_{n-1,n-1} & (1+w)E_{n-1,n-1} \end{pmatrix}. \end{array}$$

(Recall that whether or not the oracle lies, the next game is $G_{n-1,n-1}$.) Again, a straightforward analysis shows that $E_{n,n} = E_{n-1,n-1}$ and it follows that $E_{n,n} = 1$ for all $n \geq 1$.

We now establish parts (10) and (11) of the lemma. First, note that we have already shown that $E_{1,1} < E_{1,0}$. Also, if we define $E_{0,1} = E_{0,0} = 1$, then we see that

$$E_{1,1} = \frac{2E_{0,0}E_{0,1}}{E_{0,0} + E_{0,1}} = 1.$$

We will proceed by induction on n. Suppose that (10) and (11) hold for some $n \geq 1$ and for all k, $1 \leq k \leq n$. Consider, for some such k, the payoff matrix of game $G_{n+1,k}$:

$$\begin{array}{c} \\ \text{Agree} \\ \text{Disagree} \end{array} \begin{array}{c} \begin{array}{cc} \quad\text{Truth} & \qquad\quad \text{Lie} \end{array} \\ \begin{pmatrix} (1+w)E_{n,k} & (1-w)E_{n,k-1} \\ (1-w)E_{n,k} & (1+w)E_{n,k-1} \end{pmatrix}. \end{array}$$

We say that a row is a *dominated row* if its entries are never greater than the corresponding entries of the other row in the payoff matrix. On the other hand, we say that a column is a *dominated column* if its entries are never less than the corresponding entries of the other column in the payoff matrix. This difference reflects the fact that the oracle's goal is to reduce the amount that the player wins; hence, the oracle always seeks the smallest possible payoff for the player.

Clearly, neither row is dominated by the other (assuming that $w > 0$). It is also clear, from the assumption that $E_{n,k} < E_{n,k-1}$, that column 1 cannot be dominated by column 2. However, column 2 will be dominated by column 1 if

$$(1+w)E_{n,k} \leq (1-w)E_{n,k-1}.$$

This occurs when

$$w \leq \frac{E_{n,k-1} - E_{n,k}}{E_{n,k-1} + E_{n,k}}. \tag{12}$$

In this case, the oracle has a pure strategy: always tell the truth, in which case the player also has a pure strategy: always agree. This produces a payoff of $(1+w)E_{n,k}$. Subject to the inequality (12), this expression reaches a maximum value of

$$\frac{2E_{n,k-1}E_{n,k}}{E_{n,k-1} + E_{n,k}}$$

when

$$w = \frac{E_{n,k-1} - E_{n,k}}{E_{n,k-1} + E_{n,k}}.$$

On the other hand, when $w > (E_{n,k-1} - E_{n,k})/(E_{n,k-1} + E_{n,k})$, neither column is dominated by the other, in which case the oracle has a mixed strategy. The oracle's optimal strategy (p_T, p_L) will make the expected payoff of row 1 equal to the expected payoff of row 2. That is,

$$p_T(1+w)E_{n,k} + p_L(1-w)E_{n,k-1} = p_T(1-w)E_{n,k} + p_L(1+w)E_{n,k-1}. \tag{13}$$

This simplifies, since $w > 0$, to

$$p_T E_{n,k} = p_L E_{n,k-1}.$$

Using the fact that $p_L = 1 - p_T$, we may solve for p_T and p_L:

$$p_T = \frac{E_{n,k-1}}{E_{n,k-1} + E_{n,k}}, \tag{14a}$$

$$p_L = \frac{E_{n,k}}{E_{n,k-1} + E_{n,k}}. \tag{14b}$$

Now, by substituting these expressions into either side of (13), we compute the expected payoff to be

$$\frac{2E_{n,k-1}E_{n,k}}{E_{n,k-1} + E_{n,k}}.$$

This establishes that the optimal payoff occurs when

$$w \geq \frac{E_{n,k-1} - E_{n,k}}{E_{n,k-1} + E_{n,k}},$$

in which case the oracle will utilize the optimal strategy given by (14a) and (14b). Thus

$$E_{n+1,k} = \frac{2E_{n,k-1}E_{n,k}}{E_{n,k-1} + E_{n,k}}, \tag{15}$$

which establishes (11).

As was remarked in (7), equation (15) implies that $E_{n+1,k}$ is the *harmonic mean* of $E_{n,k-1}$ and $E_{n,k}$; that is,

$$\frac{1}{E_{n+1,k}} = \frac{1}{2}\left(\frac{1}{E_{n,k-1}} + \frac{1}{E_{n,k}}\right).$$

Therefore,

$$E_{n,k} < E_{n+1,k} < E_{n,k-1}.$$

Since these inequalities hold for all k, $1 \leq k \leq n$, it follows that

$$1 = E_{n,n} < E_{n+1,n} < E_{n,n-1} < \cdots < E_{n,1} < E_{n+1,1} < E_{n,0} = 2^n,$$

establishing that

$$E_{n+1,k} < E_{n+1,k-1}$$

for all k, $2 \leq k \leq n$. As special cases, we have already shown that $E_{n+1,n+1} = 1$ and $E_{n+1,0} = 2^{n+1}$, so we may conclude that inequality (10) of the lemma holds in general. ■

COROLLARY 1. *If the player and the oracle follow their optimal strategies in the game $G_{n,k}$, then an optimal wager is any amount w such that*

$$\frac{E_{n-1,k-1} - E_{n-1,k}}{E_{n-1,k-1} + E_{n-1,k}} \leq w \leq 1.$$

As before, we will call the value

$$w_{n,k} = \frac{E_{n-1,k-1} - E_{n-1,k}}{E_{n-1,k-1} + E_{n-1,k}} \tag{16}$$

the *critical wager*.

COROLLARY 2. *Let w be the amount of the wager in the game $G_{n,k}$. Then the oracle's optimal strategy is given by*

$$p_T = \begin{cases} 1 & \text{if } 0 \leq w \leq w_{n,k} \\ \frac{1}{2} + \frac{1}{2}w_{n,k} & \text{if } w_{n,k} < w \leq 1 \end{cases} \tag{17}$$

and the player's optimal strategy is given by

$$p_A = \begin{cases} 1 & \text{if } 0 \leq w \leq w_{n,k} \\ \frac{1}{2} + \frac{1}{2}\left(\frac{w_{n,k}}{w}\right) & \text{if } w_{n,k} < w \leq 1. \end{cases}$$

Proof. By using formulae (14a), (14b), and (16), we see that

$$p_T - p_L = w_{n,k}$$

when the oracle has a mixed strategy, from which the formula for the oracle's strategy follows. We will now compute the player's optimal strategy (p_A, p_D) for the game $G_{n,k}$. This strategy occurs when the expected values of the two columns of the payoff matrix are equal, giving the equation

$$p_A(1+w)E_{n-1,k} + p_D(1-w)E_{n-1,k} = p_A(1-w)E_{n-1,k-1} + p_D(1+w)E_{n-1,k-1}.$$

This simplifies to

$$w(p_A - p_D) = w_{n,k},$$

from which the formula for the player's strategy follows. ■

If the player bets any amount up to the critical wager $w_{n,k}$, then Corollary 2 prescribes a pure strategy: always agree. On the other hand, it is now rational for the player to bet more than the critical wager. Indeed, it is rational for the player to bet even the full amount ($w = 1$), provided the player is willing to disagree with the oracle occasionally. We will pursue this possibility further in the next section.

It is interesting to note that, if the player bets more than the critical wager $w_{n,k}$, then the oracle's mixed strategy is *not* dependent on the size of the wager, even though the player's mixed strategy is.

THEOREM 3. *For all $n \geq 1$ and for all k, $0 \leq k \leq n$, in the game $G_{n,k}$,*

$$E_{n,k} = \frac{2^n}{\sum_{i=0}^{k} \binom{n}{i}}$$

and

$$w_{n,k} = \frac{\binom{n-1}{k}}{\sum_{i=0}^{k} \binom{n}{i}}.$$

Proof. Lemma 1 establishes the same recurrence relation for $E_{n,k}$ that was earlier established for $A_{n,k}$. Thus, the solution for $E_{n,k}$ is the same as the solution for $A_{n,k}$. Furthermore, equation (16) is of the same form as equation (6), so $w_{n,k}$ will be the same as before. ■

We see that the expected payoff in this case is the same as the guaranteed payoff in the earlier case where the player always agreed with the oracle. Therefore, *you can't outwit the oracle (in the long run) by disagreeing with the oracle!* You might as well agree with the oracle, even though you know the oracle might lie.

Probability of a given sequence

We have now seen that by betting a sufficiently large amount and occasionally disagreeing with the oracle, we can induce the oracle to follow a predictable mixed strategy; that is, the oracle will tell the truth with known probability p_T. That makes it possible to calculate the probability of any particular sequence of truths and lies.

Note the significance of the denominator in Theorems 2 and 3. The term $\binom{n}{i}$ represents the number of ways in which the oracle can lie *exactly* i times with n flips remaining. Hence, the denominator $\sum_{i=0}^{k}\binom{n}{i}$ represents the total number of ways in which the oracle can lie with n flips and up to k lies. It turns out, as shown in the following theorem, that these different sequences of truths and lies are all equally likely. This seems reasonable, since this gives the player the least amount of information on which to choose whether to agree or disagree.

THEOREM 4. *Beginning with game $G_{n,k}$, the probability of any given sequence of truths and lies for the n coin tosses is* $1/\sum_{i=0}^{k}\binom{n}{i}$.

Proof. Let T and L represent truths and lies, respectively, in a sequence of flips. We proceed by induction. In the game $G_{1,1}$ there are only two possible sequences: T or L. As we have already seen, the probability of each is $1/2$. Now suppose that for some $n \geq 1$, the likelihood of any particular sequence of truths and lies beginning with the game $G_{n,k}$ is $1/\sum_{i=0}^{k}\binom{n}{i}$, for all k, $0 \leq k \leq n$. Consider a sequence beginning with the game $G_{n+1,k}$ for some k, $0 \leq k \leq n+1$. The first term of the sequence is either T or L. By substituting the expressions in Theorem 3 into formula (17) and simplifying, we find that the probability that the first term is T is

$$p_T = \frac{\sum_{i=0}^{k}\binom{n}{i}}{\sum_{i=0}^{k}\binom{n+1}{i}},$$

and the probability that the first term is L is

$$p_L = \frac{\sum_{i=0}^{k-1}\binom{n}{i}}{\sum_{i=0}^{k}\binom{n+1}{i}}.$$

By hypothesis, the probability of the remaining n terms of the sequence is either $1/\sum_{i=0}^{k}\binom{n}{i}$ or $1/\sum_{i=0}^{k-1}\binom{n}{i}$, depending on the number of lies remaining. Therefore, if the first term is T, then the probability of the full sequence is

$$p_T\left(\frac{1}{\sum_{i=0}^{k}\binom{n}{i}}\right) = \frac{1}{\sum_{i=0}^{k}\binom{n+1}{i}}$$

and if the first term is L, then the probability of the full sequence is

$$p_L\left(\frac{1}{\sum_{i=0}^{k-1}\binom{n}{i}}\right) = \frac{1}{\sum_{i=0}^{k}\binom{n+1}{i}}.$$

Thus, regardless of the first term, the probability of every sequence beginning with game $G_{n+1,k}$ is $1/\sum_{i=0}^{k}\binom{n+1}{i}$. This completes the induction. ■

The significance of the equation

$$E_{n,k} = \frac{2^n}{\sum_{i=0}^{k}\binom{n}{i}}$$

now becomes more apparent. The expected payoff of game $G_{n,k}$ does not depend on the size of the wager w, provided $w \geq w_{n,k}$. Therefore, consider the simple case where $w = 1$ in every game. If the oracle tells so much as a single lie, then the player loses everything, ending up with \$0. However, if the player succeeds (by chance) in outwitting the oracle every time, then the player ends up with $\$2^n$. This happens with probability $1/\sum_{i=0}^{k}\binom{n}{i}$. Therefore, the expected payoff is

$$E_{n,k} = \frac{2^n}{\sum_{i=0}^{k}\binom{n}{i}},$$

just as we calculated earlier.

Final thoughts

Playing the game with the lying oracle is best suited for those who are averse to risk. After all, if you try to outwit the oracle, you can't expect to do any better than if you simply believe the oracle each time. Furthermore, by betting large sums of money you cannot tempt the oracle into trying to outwit you, provided the oracle suspects that you may disagree. Indeed, the oracle simply mixes the predictions with lies and truths in a fixed fashion, aloof to the amount you bet, unless you are too cautious with your bet.

Listed below are some variants of the game which may make for some interesting further investigation. In each case, what is the player's optimal strategy and expected payoff?

- In the above analysis, the oracle need not lie at all during the course of the coin flips. Suppose there is a minimum number of lies that the oracle must tell.
- The oracle might also require you to place *all* your wagers before the first coin flip, expressed as a proportion of the amount you'd have before each coin flip.
- Similarly, the oracle might require you to place all your wagers before the first coin flip, but expressed as *absolute* amounts. If your holdings ever drop below your next wager, then you lose everything [**3**].
- Suppose the oracle improves the payoff for guessing the coin flip correctly (say, a correct guess pays 3:1). (See the editorial comment in [**1**].)
- What if the probability distribution of the possible outcomes isn't uniform (say, the coin is weighted)?
- What if, instead of a coin, the oracle uses a die (or any other object where the number of possible outcomes is greater than 2)?
- Consider a k-gullible oracle, that is, an oracle that continues to believe that the player will agree until the player has disagreed k times. From that point on, the oracle suspects that the player may disagree.

Acknowledgment. We would like to thank the referees for their excellent suggestions, which led to many improvements of this paper. We would also like to thank Bill Gasarch for drawing our attention to Problem 10801 in the *Monthly*.

REFERENCES

1. Owen Byer and Deirdre Smeltzer, A gambler urns his money, *Amer. Math. Monthly* **109** (2002), 394–395.
2. Paul R. Pudaite, Problem 10801, *Amer. Math. Monthly* **107** (2000), 368.
3. Dennis E. Shasha, The Delphi flip, *Scientific American* **285**:2 (2001), 94.