Sun Bin's Legacy

Dana Mackenzie
scribe@danamackenzie.com

Introduction

Sun Bin was a legendary Chinese military strategist who lived more than 2000 years ago. Among other exploits, he is credited with helping his patron, general Tian Ji, defeat the King of Qi in a match consisting of three horse races. If Tian Ji had simply raced his top horse against the King's top horse, his second against the King's second, and his third against the King's third, he would have lost all three races.

But Sun Bin had an idea. He told Tian Ji to race his *worst* horse against the King's best, his best against the King's second-best, and his second-best against the King's worst. In this way, Tian Ji won two out of three races.

Of course, Tian Ji was a little bit lucky. If we denote his horses by A, B, and C and the king's horses by a, b, and c, and if we rank the horses by speed, they happened to fall in the order $a > A > b > B > c > C$. It's easy to see with 20/20 hindsight that Sun Bin's strategy works here, because $A > b$ and $B > c$. Had the horses been in a different order, say $a > b > c > A > B > C$, then Sun Bin's strategy would not have worked (but neither would any other strategy!).

A key point to realize, though, is that Sun Bin's strategy does not depend on knowing the relative speeds of *all six* horses in advance. We only need to know the rankings of each side's horses: that is, we only need to know that $A > B > C$ and $a > b > c$. Given only this information, Sun Bin's strategy of racing C against a, A against b, and B against c is the optimal strategy in two different ways: It gives him the best odds of winning the match, and gives him the largest number of expected races won. I will justify these claims in this paper.

Now, let's bring Sun Bin's problem into the twenty-first century. What happens if we have a match of N horses against N horses? Can we find an optimal strategy for winning the match? How about for winning the largest expected number of match races? The answer to both of these questions is yes, and the goal of this paper is to answer both questions. The second one is much harder, and (in my opinion) much more interesting mathematically.

To define the problem a little better, I prefer to phrase it in terms of a card game. The deck has $2N$ cards, with face values from 1 to $2N$ ($2N$ being high). Player 1 receives N cards, face down, and player 2 likewise. Neither is allowed to see the cards in his hand or the other player's hand. However, the cards are placed in rank order in front of each player, so that each player knows the relative ranking of his own cards and the opponent's cards. On the first trick, player 1 plays one card (still face down) and player 2 plays one against it. Play continues in this fashion (player 1 always going first!) until all the cards have been played, and then both players reveal their cards. The winner is the one who takes the most tricks. What is player 2's optimal strategy?

It is easy to see that this is equivalent to an "idealized" version of the horse problem, in which the faster horse always wins the race. I prefer the card version because (as is well known) horses do not always race according to form, and faster horses sometimes lose to slower horses. In the card version, there are no such ambiguities: card 7 always beats card 6, and that is that.

To me, the N-card (or N-horse) version of Sun Bin's problem is extremely natural, and it is a bit of a mystery why it seems to be nearly absent from the mathematical literature. The only reference I have been able to find is [AGY], written in 1979, and even that reference is cursory. The authors simply noted that the problem reduces to a linear assignment problem, and therefore there exist fast computer algorithms (the Hungarian Algorithm) to solve it.

A college friend of mine, Howard Stern, independently posed this problem in his first year of graduate school, in 1980. He made, in my opinion, some extremely impressive progress towards a solution, and arrived at a correct conjecture for the general strategy, but he was unable to prove it. In the three decades since then, he has showed the problem to a number of mathematicians and computer scientists, always thinking that somebody would know a solution or a general theorem that would solve the problem. However, none of them did. Finally, he asked me in 2012, and very soon I was hooked on Stern's problem! The proof of his conjecture turned out to be a fascinating mix of group theory, probability, and combinatorics.

What's more, I believe that the solution has some relevance to the general theory of linear assignment problems. Stern's original, unpublished work from 1980 appears to be a novel observation about what I call mixed-Monge optimization problems. And the first step of my proof of Howard's conjecture also involves a general result about symmetric mixed-Monge optimization problems.

Gary Antonick wrote a post about Stern's problem in his "Numberplay" blog for the *New York Times* on January 13, 2014. It became his most-commented-on blog post in more than a year. I am indebted to one of his readers, a reader known to me only as Lee from London, who pointed out the ancient Chinese legend of Sun Bin. This is surely the first appearance of the problem in recorded history, so I think it is only appropriate to call it "Sun Bin's Legacy."

At this point I would strongly encourage readers to play the card game and see if they can figure out the strategy themselves, before reading on. Antonick's post [A] includes a wonderful applet by Gary Hewitt that will enable you to play against the computer online with 3 to 7 cards.

The main theorem of this paper is as follows:

Main Theorem. For sufficiently large N, the optimal strategy for player 2 is to play his cards in the order $(1, 2, \ldots, k, N, N\text{-}1, \ldots, k+1)$ for some k. In other words, he plays his lowest card (1) against the opponent's highest, his second-worst card against player 1's second-highest, etc. Note that k represents the number of tricks that player 2 should (try to) "throw" or lose on purpose. This strategy is optimal in the sense of maximizing the *expected number of tricks won*. The optimal number $k = k^*(N)$ is given by the following formula:

$$k^*(N) = \sup\left\{ k : \sum_{j=0}^{k-1} \binom{N}{j}^2 + \sum_{j=0}^{N-k} \binom{2N}{j} \geq \binom{2N}{N} \right\}.$$

Comments on the Main Theorem:

(1) There is also a somewhat simpler "approximate" formula $k \approx k(N)$:

$$k(N) = \sup\left\{k : \sum_{j=0}^{N-k} \binom{2N}{j} \geq \binom{2N}{N}\right\}.$$

It is approximate in the sense that $k^*(N) - k(N) = 0$ or 1. In fact, I do not know a single value of N for which $k(N) \neq k^*(N)$.

(2) Stern conjectured the general form of the optimal strategy in 1980, but did not make a conjecture for the optimal number $k^*(N)$ of tricks to throw. At that point there was not enough data to make a conjecture, and it is highly unlikely that anyone would have come up with the formula above anyway. It was a complete shock to me that I was able to derive an exact formula.

Here is a table of the optimal number of tricks to throw for small values of N:

N	k	N	k
2	n/a	9	2
3	1	10	2
4	1	11	2
5	1	12	2
6	1	13	3
7	2	14	3
8	2	15	3

And here are two "worked examples," showing the first two jumps in $k^*(N)$.

Example 1: $N = 7$. Here the approximate formula tells us to look up the 14-th row of Pascal's triangle and add the terms until we get a sum that is greater than the central element. We find that

$$1 + 14 + 91 + 364 + 1001 + 2002 = 3473 > 3432.$$

Then the approximate number of tricks to throw is (N+1) minus the number of terms added: in this case $(7 + 1) - 6 = 2$.

For the exact computation, we add a couple of squared terms from the 7-th row:

$$1 + 14 + 91 +364 + 1001 + 2002 + 1^2 + 7^2 = 3523 > 3432.$$

Thus the exact number of tricks to throw is at least 2. On the other hand, if we try throwing one more, we get

$$1 + 14 + 91 + 364 + 1001 + 1^2 + 7^2 + 21^2 = 1962 < 3432.$$

Thus the exact number of tricks to throw is at most 2, and hence the exact formula agrees with the approximate formula.

Example 2: $N = 13$. Now we look up the 26th row of Pascal's triangle and add up the terms until we get a sum that is greater than the center element. We find that

$$1 + 26 + 325 + 2600 + 14950 + 65780 + 230230 + 657800 + 1562275 + 3124550 + 5311735 = 10970722 > 10400600.$$

Therefore the approximate number of tricks to throw is $(N + 1)$ minus the number of terms added, i.e. $(13 + 1) - 11 = 3$.

The exact computation involves adding some squared terms from the 13th row. Because

$$1 + 26 + 325 + 2600 + 14950 + 65780 + 230230 + 657800 + 1562275 + 3124550 + 5311735 + 1^2 + 13^2 + 78^2 > 10400600$$

we can be certain that the exact number of tricks to throw is at least 3. And because

$$1 + 26 + 325 + 2600 + 14950 + 65780 + 230230 + 657800 + 1562275 + 3124550 + 1^2 + 13^2 + 78^2 + 286^2 < 10400600$$

the exact number of tricks to throw is less than 4. Hence the exact number of tricks to throw is 3, which agrees with the approximate computation.

These two examples are completely typical. The additional "nuisance terms" from the N-th row of Pascal's triangle, even though they are squared, are dwarfed by the largest terms from the $(2N)$-th row. This is why the "approximate" formula agrees with the exact formula in every case I know of.

(3) It is also of interest to derive upper and lower bounds for the exact number of tricks to throw. After all, if you are playing the game with $N = 200$ cards, it may not be so easy to look up the 400^{th} row of Pascal's triangle! I prove the following estimate in this paper:

Theorem: If $N > 400$, then the optimal number of tricks to throw satisfies the inequalities

$$\sqrt{N \ln N / 4} < k^*(N) < \sqrt{N \ln N / 2}.$$

Computer calculations by Stern show that these inequalities hold for $400 \geq N \geq 91$ as well. The left-hand inequality is false for $N = 90$.

As $N \to \infty$, $k^*(N) \sim \sqrt{N \ln N / 2}$. It is interesting that this asymptotic limit is approached extremely slowly. Stern's computer calculations show that all the way up to $N = 500$, the ratio $k^*(N)/\sqrt{N \ln N}$ is closer to 0.5 than it is to 0.7071…, its eventual limit. (For $N = 500$, the ratio is 0.559…)

Finally, as mentioned briefly above, there is a second version of Sun Bin's Legacy, which is to find the strategy that guarantees the highest probability of winning a majority of tricks, regardless of the number of tricks won. Curiously, neither Stern nor I worked seriously on this question. In my case, this was because I expected the majority-of-tricks problem to be harder, because the objective function is nonlinear.

Imagine my astonishment when, within one day of Gary Antonick's post going up on the "Numberplay" blog, one of his readers found the optimal strategy for the majority-of-tricks problem! Here I assume $N = 2n+1$ is odd. Reader Bill Courtney showed that the optimal strategy is to throw n tricks. Thus player 2 pairs his top $(n+1)$ cards against player 1's bottom $(n+1)$ cards, in order. It is easy to see that if there is *any way at all* to win $(n+1)$ tricks, then this strategy will do so. The proof is left to the reader (or see Courtney's comment to [A]).

While Courtney's strategy maximizes the probability of winning a simple majority, it is extravagantly wasteful on the level of tricks. It will on average lose nearly half the tricks. By contrast, the "Pascal's triangle" strategy described above will on average lose only about $\sqrt{N \ln N / 2}$ tricks. By playing with a large enough deck, you can win as close as you want to 100 percent of the tricks!

The outline of the rest of the paper is as follows:

I. Basic Results, Mixed Monge Matrices and the Shape Theorem.

This section sets up the problem as a linear assignment problem, shows that the objective function is given by a "mixed Monge matrix," and derives a weak form of the optimal strategy.

In particular, I show that the optimal strategy always involves throwing some tricks in reverse order, and playing the rest of the tricks in normal order. However, there may be "gaps" in the thrown tricks. Most of the work in this section is due to Stern (unpublished).

II. The Symmetry Lemma.

The objective function in section I leads to a mixed Monge matrix that is symmetric about the "anti-main diagonal," and skew-symmetric (after subtracting a constant from each entry) about the main diagonal. I exploit this symmetry to prove that *if* you have decided which tricks to throw (say tricks 1, 3, and 7) then the optimal strategy for these tricks is to play your i-th worst card against your opponent's i-th best card. Still, there may be gaps in the thrown tricks.

III. The No-Gaps Theorem.

In this section, which is the most technical one, I show that if N is large enough (at least 10 million) then the optimal strategy has no gaps. That is, you should throw tricks 1, 2, …, k for some k. Although I do not derive the best strategy in this section, the proof depends on knowing that a *very good* strategy is to sacrifice the first $\sqrt{N \ln N / 2}$ tricks. Roughly speaking, this strategy beats any strategy with gaps in it.

IV. The Number of Tricks to Throw.

In the last section, I explain the wonderful and totally unexpected connection between the expected number of tricks won and the $2N$-th row of Pascal's triangle. I derive the exact number of tricks to throw, $k^*(N)$, the approximate number $k(N)$, and the asymptotic limit for both of them.

Bibliography.

[A] G. Antonick, "Stern-Mackenzie One-Round War," *New York Times* ("Numberplay" blog), Jan. 13, 2014, at http://wordplay.blogs.nytimes.com/2014/01/13/war/.
[AGY] A. Assad, Golden, B., and Yee, J. Scheduling players in team competitions. *Proceedings of the Second International Conference on Mathematical Modelling (St. Louis, MO, 1979),* Vol I, pp. 369-379. Rolla, MO: Univ. of Missouri-Rolla, 1980.

Addendum.

A much improved version of the proof (valid for all N, not just for $N > 10^7$) was published in R. Chatwin and D. Mackenzie, "How to Win at (One-Round) War," *College Math. Jour.* Vol. 46, No. 4, Sept. 2015, 242-253.

The story behind this paper may be of interest to Gathering for Gardner readers. One of the people who attended my G4G11 talk was Brian Hopkins, the editor of the *College Mathematics Journal* (published by the Mathematical Association of America). He invited me to submit a paper on the Sun Bin problem to the *College Mathematics Journal*.

In the meantime, I had started collaborating with Richard Chatwin, who had read about the Sun Bin problem in Gary Antonick's *New York Times* article referenced above. Chatwin, unlike me, is an expert on linear assignment problems. (He wrote his Ph.D. dissertation on airline overbooking, which involves this type of problem.) He asked the natural question, "What if we use the Hungarian algorithm?"

In the end, it turned out that the Hungarian algorithm *per se* does not solve Sun Bin's problem for all N, although it certainly can solve it for any individual value of N. The fundamental reason is that the algorithm repeatedly involves the step of finding the smallest element in a given row or column of a matrix and "pivoting" about that element. Identifying a particular element as the smallest amounts to proving a whole set of inequalities. The algorithm tells you *what inequalities you need to prove*, but not how to prove them! In fact, Chatwin found repeatedly that the inequalities he needed were precisely the ones already proved in my paper.

Nevertheless, Chatwin did make major improvements to certain parts of my proof, especially to Part III described above. The final (and definitive) result reduced the "at least 10 million cards" requirement to a much more manageable "at least 41 cards." That is, we can prove analytically that the Pascal's triangle strategy is optimal, provided that $N \geq 41$. For $3 \leq N \leq 40$, the analytic estimates are too inexact and we have to resort to a case-by-case analysis on the computer. Chatwin did the computer calculations necessary to prove that the Pascal's triangle strategy remains optimal. (Stern had already checked this for $N \leq 60$, but it was nice to have an independent verification.)

My original contribution to the G4G11 gift exchange was a 56-page paper with the proof that the Pascal's triangle strategy is optimal for $N > 10^7$. Because we now have a much better proof of a more complete result, it no longer seems necessary to me to have the entire 56-page paper reproduced in this volume. However, it does appear in the online version for any readers who might be interested in seeing the not quite fully-baked version of the proof.

I think that Martin Gardner would have approved of the way that a column written for the public (Gary Antonick's blog) put two mathematicians together who never would have been able to find each other otherwise; and the way that a conference in his honor put us in contact with the editor who eventually published our manuscript. So my main gift to the G4G11 exchange is to tell you that Martin Gardner's legacy (as well as Sun Bin's legacy) is alive and well!