The MENSA Correctional Institute

Larry Carter and Stan Wagon

Abstract. We investigate the following puzzle and several variations, some of which have quite surprising answers. Alice and Bob are in prison under the care of warden Charlie. Alice will be brought into Charlie's office and shown 52 cards, face-up in a row in an arbitrary order. Alice can interchange two cards. Charlie then turns all cards face-down in their places and Alice leaves the room. Then Bob is brought in and Charlie calls out a random target card. Bob can turn over cards, one after another, at most 26 times as he searches for the target. Both prisoners are freed if Bob finds the target. Find a strategy that never fails.

1. THE MENSA CORRECTIONAL INSTITUTE. Problem 1, which we first saw in [**1**], is in the style of surprising hat or card problems that have become popular of late. The attraction of such problems is that clever algorithms guarantee success with a much higher probability than seems possible. The originator is unknown, though it is related to other problems of this type [**7**; **8**, p. 12].

Problem 1. Alice and Bob are imprisoned at the Mensa Correctional Institute under the care of warden Charlie. Alice will be brought into Charlie's office on Sunday and shown a standard deck of 52 cards, face-up in a row in some arbitrary order. Alice may, if she wishes, interchange two cards. She then leaves and Charlie turns each card face-down in its place. Bob is then brought in and Charlie calls out a random target card. Bob can turn over cards, one after another, at most 26 times as he searches for the target. If he finds it, both prisoners are freed; if he fails to find the target, they stay in prison. Find a strategy that guarantees success for the prisoners regardless of Charlie's choices.

As always in this type of puzzle, Alice and Bob know the rules and are allowed to plan a strategy in advance. On Sunday they cannot communicate; Alice has no idea what the target card will be.

Problem 1 makes sense for all deck sizes, so we will consider it and its variations as involving n cards; we take the cards as being the integers from 1 to n. An intriguing aspect of Problem 1 is that it appears as if Alice has to use her switch to somehow communicate where each card is, at least up to half the deck. That is, if she could somehow convey 51 bits of information, she could tell Bob which half of the deck contains each possible target. But by placing card i in position 1 (where Bob would know to look first) she has communicated only about 6 bits of information. So how can they possibly succeed in all cases? The solution is that she will communicate no information to Bob, but instead use her switch to transform the deck into a very particular type of Bob-friendly order. And, of course, they will devise a method whereby Bob uses what he sees to guide future moves (as opposed to, say, turning over the first 26 cards regardless of what he sees).

Solution to Problem 1. Charlie's linear arrangement determines a permutation (e.g., the array (3, 4, 6, 2, 5, 1) is $1 \to 3$, $2 \to 4$, and so on). Any permutation splits into cycles; the preceding example splits into $(1 \to 3 \to 6 \to 1)\,(2 \to 4 \to 2)\,(5)$. We'll use the traditional short form for cycles: $(1\,3\,6)\,(2\,4)\,(5)$. Alice examines the cycles for the initial arrangement. Perhaps she sees just the 52-cycle: $(1\,2\,3 \ldots 52)$. She finds the longest cycle that occurs—say $(c_1\, c_2 \ldots c_L)$—and then, in the linear array she faces, transposes cards c_1 and $c_{1+\lceil L/2 \rceil}$. In the example, she would transpose cards 1 and 27. This splits the long cycle into two equal or nearly-equal cycles. Because at most one cycle has length 27 or greater, her action means that Bob will face a permutation whose cycle decomposition has no cycle longer than 26.

It is useful to consider the cards as being hidden behind doors and view each flip of a card as the opening of a door. Bob enters the room and learns the target C. He opens door C and looks at card U. If he is lucky and $U = C$, they win immediately. Otherwise, Bob opens door U. If V is the card he then sees, either $V = C$ and they win (the permutation would have the 2-cycle $(C\,U)$), or he moves on to door V. He continues in this way until he is about to open door C, which must happen sometime; but that can happen only when he is looking at card C. Thanks to Alice, no cycle is longer than 26, so he will succeed in at most 26 steps. ■

We use *basic strategy* to refer to the method of Problem 1's solution. One thinks in terms of what Alice's strategy is, and then what Bob's strategy is, but in fact (and this holds for all the variations) only Bob's strategy matters because Alice's best move is always to optimize the objective according to what Bob will do. So the basic strategy, for Problem 1 and the variations to come, is just: Bob opens door C and then opens the door whose number is that of the most recently seen card. Given this, Alice will do the best thing, which, for Problem 1, is simple: halve the longest cycle. The solution given for Problem 1 works for general n, with success guaranteed in $\lceil n/2 \rceil$ flips. So when n is 3 or 4, success is guaranteed in 2 flips. When n is 2, only 1 flip is needed: Alice ensures that the permutation facing Bob is the identity. We immediately have an open question.

Open Question 1. Prove that there is no n for which Problem 1 has a 100% successful solution where Bob is allowed to examine only $\lceil n/2 \rceil - 1$ cards.

There are several interesting variations to Problem 1 that have surprising answers. We study two variants here: the "short game," where Bob gets to flip one card only, and the "long game," where the prisoners' goal is to minimize the expected number of flips before the target card is found.

2. THE SHORT GAME. Suppose Charlie wants a very short test and allows Bob only a single flip in his search for the target. From now on we assume that Charlie's choice of the initial shuffle and target card are random: they are selected randomly with uniform probability from the collection of all possible choices. The probability of success increases as the size of the deck decreases, so suppose there are five cards. If Alice were uninvolved, Bob's success rate would be 20% ($1/n$ in general). This next problem joins the list of problems that many professional mathematicians get wrong.

Problem 2 (The Short Game). The setup is as in Problem 1, but Alice and Bob are set free if and only if Bob finds the target C by looking at only one card. What is their best strategy for a 5-card deck?

The natural approach, given what we learned from Problem 1, is to have Bob open door C, exactly as before; this is still a "basic strategy," but with only a single door in play. To optimize things for Bob, Alice will examine the cycle structure. She checks whether the initial permutation has a 2-cycle. If it does, she chooses one such and switches the two cards. If not, either it is the identity, which she leaves untouched, or it has a longer cycle $(c_1 c_2 \ldots c_L)$, which she splits into a fixed point and an $(L-1)$-cycle using the move $(c_1\, c_2)$. The success probability for this basic strategy is $47\frac{1}{3}\%$, which is the same as saying that over all $5 \cdot 5!$ things that Charlie controls, the number of successes is 284. One can compute this value as follows. Let $N_{\text{basic}}(n)$ be the count of all successes using this strategy, and let $T(n)$ be the number of permutations of $\{1, \ldots, n\}$ having at least one transposition.

Theorem 1. $N_{basic}(n) = 2 \cdot n! - 1 + T(n)$.

Proof. Among all $n!$ possibilities for Charlie's shuffle, the number of fixed points is $n!$ (each i is a fixed point of $(n-1)!$ permutations, so $n\,(n-1)!$ in all); a fixed point leads to a success provided Alice does not disturb it. Also, for any Charlie shuffle that isn't the identity, Alice has a switch that introduces at least one more fixed point, so that is another $n! - 1$ successes. But when there is a transposition, its reversal adds 2 to the count instead of just 1; this gives the final $T(n)$ term. ■

The possible types (vector of cycle lengths) for permutations of $\{1, 2, 3, 4, 5\}$ having a transposition are (2, 3), (2, 2, 1), (2, 1, 1, 1), with frequencies $\binom{5}{2} 2 = 20$, $\binom{5}{2}\binom{3}{2}\frac{1}{2} = 15$, and $\binom{5}{2} = 10$, respectively, for a total of $T(5) = 45$. Therefore $N_{\text{basic}}(5) = 240 - 1 + 45 = 284$ as claimed. Somewhat surprisingly, the transposition count can be computed as $5! - \frac{5!}{2^2\, 2!} \left\lfloor 2^2 \frac{2!}{\sqrt{e}} \right\rceil = 15 \left\lfloor 8 \left(1 - \frac{1}{\sqrt{e}}\right) \right\rceil = 15 \lfloor 3.15 \rceil = 45$, where $\lfloor x \rceil$ is the integer nearest x, rounding up when necessary. We prove this in a moment.

The function $T(n)$ is well studied [**6**; **3**, A000266, A027616]. Notation: $T^c(n)$ denotes $n! - T(n)$ and m is $\left\lfloor \frac{n}{2} \right\rfloor$. Here are several facts about $T(n)$.

- Values: $T(1)$ through $T(10)$ are 0, 1, 3, 9, 45, 285, 1995, 15 855, 142 695, 1 427 895.

- Sum representation: $T(n) = n! \sum_{k=1}^{m} (-1)^{k+1} \frac{1}{2^k\, k!}$; $T^c(n) = n! \sum_{k=0}^{m} (-1)^k \frac{1}{2^k\, k!}$. This is proved later in this section using inclusion-exclusion.

- Exact formula: $T(n) = n! - \frac{n!}{2^m\, m!} \left\lfloor \frac{2^m\, m!}{\sqrt{e}} \right\rceil$; $T^c(n) = \frac{n!}{2^m\, m!} \left\lfloor \frac{2^m\, m!}{\sqrt{e}} \right\rceil$. This is due to S. Plouffe [**4**] (rediscovered by M. van Hoeij [**3**, A000266]) and is by far the fastest method of computation. Proof. For $T^c(n)$, use the sum representation and show that the integer $2^m\, m! \sum_{k=0}^{m} \left(\frac{-1}{2}\right)^k \frac{1}{k!}$ differs from $2^m\, m! \big/ \sqrt{e}$ by less than $1/2$ as follows. The alternating series error bound gives: $2^m\, m! \left| \frac{1}{\sqrt{e}} - \sum_{k=0}^{m} \left(\frac{-1}{2}\right)^k \frac{1}{k!} \right| \leq \frac{2^m\, m!}{2^{m+1}\, (m+1)!} < \frac{1}{2}$. This is one of many "round" formulas for integer sequences $n!\, b^n \sum_{k=0}^{n} \left(\frac{a}{b}\right)^k \frac{1}{k!}$, where a and b are coprime integers [**6**].

- Recurrence formula: $T^c(n) = n\, T^c(n-1) - (n-1)\, T^c(n-2) + \left(n^2 - 3\, n + 2\right) T^c(n-3)$, with $T^c(1) = T^c(2) = 1$, $T^c(3) = 3$; this can be derived from the sum representation.

- Integral representations: $T^c(n) = \frac{n!}{m!} \int_0^\infty e^{-t} \left(t - \frac{1}{2}\right)^m dt = \frac{n!}{m!\, \sqrt{e}} \int_{-1/2}^\infty e^{-t}\, t^m\, dt$. These arise from the general formula $\sum_{k=0}^{m} \frac{z^k}{k!} = \frac{e^z}{m!} \Gamma(m+1, z)$, where Γ is the incomplete gamma function defined by $\Gamma(m+1, z) = \int_z^\infty e^{-t}\, t^m\, dt$. Proof. The sum yields this recurrence: $T^c(n) = (-1)^m \frac{n!}{m!\, 2^m} + n(n-1)\, T^c(n-2)$; a single integration by parts on the integral defining Γ yields the same recurrence.

- Asymptotics: $T(n)$ is asymptotic to $n! \left(1 - \frac{1}{\sqrt{e}}\right)$. This follows from the exact formula.

By Theorem 1, the properties of $T(n)$ immediately give results about $N_{\text{basic}}(n)$. So we can quickly compute the success probability for a standard 52-card deck: it is somewhat larger than one might expect, at 4.603%. The asymptotic behavior of $T(n)$ means that Alice's switch improves the success rate from the no-Alice value of $\frac{1}{n}$ to $\left(3 - \frac{1}{\sqrt{e}}\right)\frac{1}{n}$; i.e., Alice's move increases the success probability by about 239%. We won't investigate the finer asymptotics—the error in the main asymptotic approximation—but the error is quite small: for $n = 8$ the success count is 96494, whereas $\left\lfloor \left(3 - \frac{1}{\sqrt{e}}\right) 8\,! \right\rfloor - 1 = 96504$.

It came as a big surprise to find that the basic strategy is *not* the best strategy for the given case, $n = 5$. For all n except 1, 2, 3, 4, 6, and 7, there is a better strategy (Theorem 8). A *strategy* here is simply a rule that tells Bob what to do upon hearing the target C. The basic strategy is just $\langle 1, 2, 3, \ldots \rangle$, where the Cth entry tells Bob which door to open on hearing C. The basic strategy (or any permutation of it) is the only one that actually opens all doors and one might think that a strategy that leaves some doors forever unopened would not be useful. Certainly such a closed-door approach would fail if we had to find the target in all cases; but for Problem 2 that is not required.

This leads one to look at the double-door strategy (DD_1), given by $\langle 1, 1, 3, 4, 5 \rangle$; this means Bob opens door C unless $C = 2$, in which case he opens door 1. And now the surprise: The success count $N_{DD_1}(5)$ is 286, two greater than the basic strategy's 284. The success probability is therefore $\frac{286}{5 \cdot 5!}$ or $47\frac{2}{3}\%$, slightly better than the basic strategy. One can verify the 286 by simple computation, but we will derive it from a more general method (Corollary 6(a)). It is not hard to program a check of all 5^5 strategies—functions from $\{1, 2, 3, 4, 5\}$ to itself—to confirm that DD_1 is the best strategy for $n = 5$ (and that the basic strategy is the best for $n \in \{1, 2, 3, 4, 6, 7\}$). Here is the formalization of all double-door strategies.

Definition. The *double-door strategy* DD_j uses j double doors with the rest being single doors. For example, if $n = 7$, DD_2 is the strategy $\langle 1, 1, 2, 2, 3, 4, 5 \rangle$. We use DD_∞ for the strategy with the maximum number of double doors; i.e., $DD_\infty = DD_{\lfloor n/2 \rfloor}$.

Theorem 1 extends to a general formula that applies to any strategy σ. Let I_σ be the number of permutations that are "ideal" in that they cannot be improved by Alice to increase the total count of good cases; let T_σ be the number of permutations for which some Alice switch will increase the good count by 2; we call permutations admitting such a switch "strong permutations."

Theorem 2. $N_\sigma = 2 \cdot n\,! - I_\sigma + T_\sigma$.

Proof. Call a card "good" if Bob will find it without the intervention by Alice; if Charlie's permutation is π, then i is good if and only if $\sigma(\pi(i)) = i$. Among all possibilities for π the number of good cards is $n\,!$ (each i is good in $(n-1)\,!$ permutations, so $n\,(n-1)\,!$ in all). For any non-ideal π, Alice has a switch that introduces at least one new good card, adding $n\,! - I_\sigma$ to the success count. But when π is strong—it contributes to the T_σ count—Alice's move adds 2 to the count instead of just 1. ■

For any strategy σ it is easy to compute I_σ. We assume throughout that the numbers in σ are an initial segment of $\{1, \ldots, n\}$ (so DD_1 is viewed as $\langle 1, 1, 2, 3, 4 \rangle$). Note that n is determined by σ; it is the number of entries in the strategy vector. Let $\tau(\sigma)$ denote the vector of frequencies in σ; so $\tau(\langle 1, 1, 2, 3, 4 \rangle) = (2, 1, 1, 1)$; let $\sigma_{\max} = \max(\sigma)$.

Proposition 3. $I_\sigma = (n - \sigma_{\max})\,! \prod \tau(\sigma)$.

Proof. A permutation is ideal if and only if it has $\sigma_{\max}$ cards in good positions; the other cards are unrestricted, yielding the factorial factor. The product is the number of ways of choosing a card to be in a good position for each of the doors appearing in σ. ■

We assume Charlie's choices to be purely random, but one can wonder whether he can accomplish anything by trying to make life difficult. Suppose, for instance, that there were a single strategy for the prisoners that got a better average score than any other strategy. The participants have unlimited computer access, so they would all discover this strategy. Charlie could then make his choices so that the optimum strategy would perform poorly. Of course, Alice and Bob would know that Charlie can act this way, so they would choose to use a nonoptimal strategy, perhaps the second-best strategy. Charlie can probably find a choice that defeats both the best and second-best strategies. And so on.

But Charlie cannot thwart the prisoners in this way. For any strategy, the prisoners can renumber the cards and doors to derive a disguised version of it that achieves the same average. That is, if the prisoners randomly relabel the cards and doors, then no matter what Charlie does, the average (for any of Charlie's choices, averaged over all possible renumberings) is the same. Knowing nothing about the new numbers, Charlie cannot affect the average. Thus Alice and Bob can force Charlie's choices to be essentially random.

THE INCLUSION-EXCLUSION METHOD. The classic Principle of Inclusion and Exclusion leads to an algorithm that computes N_σ exactly and efficiently for any strategy σ. For sets $A_1, \ldots, A_n$, the general principle states that

(IE) $$\left| \bigcup_{k=1}^{n} A_i \right| = \sum_{k=1}^{n} (-1)^{k+1} \left(\sum_{0 < i_1 < \cdots < i_k \leq n} \left| A_{i_1} \cap \cdots \cap A_{i_k} \right| \right).$$

A classic application is counting derangements: permutations that fix no point [**3**, A000166]. Let A_i be permutations of $\{1,\ldots, n\}$ that fix i. For each k, the cardinality of the intersection in (IE) is $(n-k)!$, because there is complete freedom in $n-k$ positions and the number of choices for $(i_1,\ldots, i_k)$ is $\binom{n}{k}$. So, by (IE), the number of fixers is $\sum_{k=1}^{n}(-1)^{k+1}\binom{n}{k}(n-k)! = n!\sum_{k=1}^{n}(-1)^{k+1}\frac{1}{k!}$. Thus the number of derangements is $n!\left(1-\sum_{k=1}^{n}(-1)^{k+1}\frac{1}{k!}\right) = n!\sum_{k=1}^{n}(-1)^{k}\frac{1}{k!}$. This is the same as $\Gamma(n+1,-1)/e$. It is also known as the subfactorial function and equals $\left\lfloor\frac{n!}{e}\right\rfloor$ [**6**; **3**, A000166].

Next we can apply (IE) to $T(n)$, the count of permutations having at least one 2-cycle. For K odd, the double factorial $K!!$ denotes $1\cdot 3\cdot 5\cdots K$. Let A_{ij} be the set of permutations of $\{1,\ldots, n\}$ that switch cards i and j. Then $\left|\bigcup_{ij} A_{ij}\right|$ is $T(n)$. The intersection in (IE) is empty unless the pairs are disjoint, and then the intersection cardinality is $(n-2k)!$. The number of times a set of k transpositions occurs is $\binom{n}{2k}(2k-1)!!$ where the first factor determines the entries that make up the transpositions and the second arises from the fact that the smallest entry has $2k-1$ potential partners, the smallest unused entry then has $2k-3$ potential partners, and so on. This is the same as $\frac{1}{2^k\,k!}\,\frac{n!}{(n-2k)!}$. So $T(n)$ is the following sum, where m denotes $\lfloor n/2\rfloor$ and can be used as the upper limit because the transpositions in each term of the inner sum must be disjoint. We use the exact formula of Section 2 for the final equality.

$$T(n) = n!\sum_{k=1}^{m}(-1)^{k+1}\frac{1}{2^k\,k!\,(n-2k)!}(n-2k)! = n!\sum_{k=1}^{m}(-1)^{k+1}\frac{1}{2^k\,k!} = \frac{n!}{2^m\,m!}\left\lfloor\frac{2^m\,m!}{\sqrt{e}}\right\rceil.$$

The IE method can be applied to *any* strategy. Given a strategy σ with frequency vector τ, let $D=\sigma_{\max}=|\tau|$ (the number of doors the strategy uses) and $d=\lfloor D/2\rfloor$. The number of cards is $n=|\sigma|$. Recall the elementary symmetric polynomials

$$e_q(X_1, X_2,\ldots, X_M) = \sum_{1\le j_1<j_2<\cdots<j_k\le M} X_{j_1}\cdots X_{j_q}.$$

For example, $e_4(\tau)=\tau_1\,\tau_2\,\tau_3\,\tau_4+\tau_1\,\tau_2\,\tau_3\,\tau_5+\ldots$ where summands arise from all 4-tuples from τ. These can be computed very quickly by the Pascal-triangle-like recursion $e_i(X_1,\ldots,X_j)=X_j\,e_{i-1}(X_1,\ldots,X_{j-1})+e_i(X_1,\ldots,X_{j-1})$ with initial conditions $e_0(\cdot)=1$ and $e_i(X_1,\ldots,X_j)=0$ if $j<i$. The next result presents a general and remarkably simple formula for the number of strong permutations.

Theorem 4. *For any strategy σ, the count of strong permutations, T_σ, is given by*

$$T_\sigma = \sum_{k=1}^{d}(-1)^{k+1}(n-2k)!\,(2k-1)!!\,e_{2k}(\tau).$$

Proof. Let A_{ij} be the set of permutations of the cards that are strong because of an $i\leftrightarrow j$ switch: permutations of $\{1,\ldots,n\}$ for which a switch of cards i and j improves the success count by 2. For example, suppose $\sigma=\langle 1, 1, 2, 3, 4\rangle$ and $\pi=\{3, 1, 2, 5, 4\}$. Then $\pi\in A_{13}$: a switch of cards 1 and 3 gives $\{1, 3, 2, 5, 4\}$ putting cards 1 and 3 in good positions. In general, $\pi\in A_{ij}$ if and only if $\pi(\sigma(i))=j$ and $\pi(\sigma(j))=i$.

The number of strong permutations is $\left|\bigcup_{1\le i<j\le n} A_{ij}\right|$. Let $q=\binom{n}{2}$. Use k as an index from 1 to q; think of k as indexing all possible pairs of cards, so that we can use A_k to mean A_{ij} where ij is the pair corresponding to k. So IE says

$$T_\sigma = \left|\bigcup_{k=1}^{q} A_k\right| = \sum_{k=1}^{q}(-1)^{k+1}\sum_{\substack{\text{distinct}\\ r_1,\ldots,r_k}}\left|A_{r_1}\cap\cdots\Box A_{r_k}\right|.$$

Claim. The intersection term is empty unless the involved pairs (i, j) are disjoint.

Proof of Claim. If $\pi\in A_{ij}$, then j is determined by i and vice versa: $j=\pi(\sigma(i))$ and $i=\pi(\sigma(j))$.

The claim means that the upper limit on k can be taken to be d. So we have

$$T_\sigma = \sum_{k=1}^{d}(-1)^{k+1}\sum_{\substack{\text{disjoint}\\ r_1,\ldots,r_k}}\left|A_{r_1}\cap\cdots\Box A_{r_k}\right|.$$

To finish, fix k and analyze the interior sum. Any permutation π in the intersection will transpose cards from k door-pairs, leaving the remaining $n-2k$ cards unconstrained; this gives $(n-2k)!$ as a multiplier for each such π. Next we need to work out how many π there are for a given k.

Group the summation into subcollections consisting of pairings from the same underlying door set. For example, the door set $\{1, 2, 3, 4\}$ leads to three possible pairings: (12) (34), (13) (24), and (14) (23). But this is only part of the classification since the r-indices refer to cards, not doors. So consider all card-switches for which the corresponding doors arise from the same set. Each pairing leads to a product of multiplicities τ_i for the overall count (in the example above, each of the three pairings leads to $\tau_1\tau_2\tau_3\tau_4$ choices for the cards being switched). Finally, we need the number of pairings once the door-set is chosen. That is just $(2k-1)!!$, as in the earlier proof about the transposition count.

Since each choice of door-set occurs in the same manner, the count we seek is $(2k-1)!!\,e_{2k}(\tau)$, which gives the claimed formula. ■

Corollary 5. *For any strategy σ,*

$$N_\sigma = 2 \cdot n! - (n - \sigma_{\max})! \left(\prod_{i=1}^{D} \tau_i \right) + \sum_{k=1}^{d} (-1)^{k+1} (n - 2k)! (2k - 1)!! \, e_{2k}(\tau).$$

Proof. By Theorems 2 and 4 and Proposition 3. ■

Corollary 6. *(a) For* $n \geq 4$, $N_{DD_1}(n) = 2 \cdot n! - 2 - (n-2)! \sum_{k=1}^{\lfloor (n-1)/2 \rfloor} (-1)^k \frac{1}{2^{k-1}\, k!} \left(\binom{n}{2} - (2k^2 - k) \right)$.

(b) For $n \geq 4$,

$$N_{DD_2}(n) = 2 \cdot n! - 8 - (n-4)! \sum_{k=1}^{\lfloor n/2 \rfloor - 1} \frac{(-1)^k}{2^k\, k!} (n - 2k - 1)(n - 2k) \left(4k^2 + 2k(2n - 7) + n^2 - 5n + 6 \right).$$

(c) If n *is even, let* $D = \frac{n}{2}$. *Then* $N_{DD_\infty}(n) = 2 \cdot n! - D!\, 2^D - \sum_{k=1}^{d} (-1)^k \frac{(n-2k)!\, (2k)!\, 2^k}{k!} \binom{D}{2k}$. *If* n *is odd, let* $D = \frac{n+1}{2}$. *Then* $N_{DD_\infty}(n) = 2 \cdot n! - (D-1)!\, 2^{D-1} - \sum_{k=1}^{d} (-1)^k \frac{(n-2k)!\, (2k)!\, 2^{k-1}}{k!} \left(\binom{D}{2k} + \binom{D-1}{2k} \right)$.

Proof. (a) Here $\sigma = \langle 1, 1, 2, 3, \ldots, n-1 \rangle$, $D = n - 1$; $d = \left\lfloor \frac{n-1}{2} \right\rfloor$; $\sigma_{\max} = n - 1$; $\tau = (2, 1, \ldots, 1)$, and $e_{2k}(\tau) = \binom{n-1}{2k} + \binom{n-2}{2k-1}$. The ideal count is 2 because only the identity and the transposition (12) are ideal. So $N_{DD_1} = 2 \cdot n! - 2 + \sum_{k=1}^{d} (-1)^{k+1} (n-2k)! (2k-1)!! \left(\binom{D}{2k} + \binom{D-1}{2k-1} \right)$. This simplifies to the more succinct form in the corollary.

(b) Here $\sigma = \langle 1, 1, 2, 2, 3, \ldots, n-2 \rangle$, $D = n - 2$; $d = \left\lfloor \frac{n}{2} \right\rfloor - 1$; $\sigma_{\max} = n - 2$; τ is $(2, 2, 1, \ldots, 1)$, and $e_{2k}(\tau) = 4\binom{n-4}{2k-2} + 4\binom{n-4}{2k-1} + \binom{n-4}{2k} = \binom{n-4}{2k} + 4\binom{n-3}{2k-1}$. The ideal count is $2! \cdot 2 \cdot 2 = 8$ (Proposition 3). Therefore $N_{DD_2} = 2 \cdot n! - 8 + \sum_{k=1}^{d} (-1)^{k+1} (n - 2k)! (2k-1)!! \left(\binom{n-4}{2k} + 4\binom{n-3}{2k-1} \right)$, which is the same as the form in the corollary.

(c) This is similar to (b), except that the symmetric polynomial computation depends on the parity of n; details omitted. ■

The formula for N_{DD_1} gives $N_{DD_1}(5) = 2 \cdot 5! - 2 - 6(-9 + 1) = 240 - 2 + 48 = 286$, larger than the 284 for the basic method. Recall that $N_{\text{basic}}(5) = 240 - 1 + 45 = 284$. So the DD_1 victory by 2 points is because the count of strong moves by Alice is 48 as opposed to 45, a gain of 3, while the ideal count loss increases by only 1.

The explicit DD_1 formula shows that the double-door strategy beats the basic strategy for a 52-card deck. The counts are

$$N_{\text{basic}}(52) = \underline{\mathbf{193\,052\,869\,315\,181\,895\,879\,191\,890\,783\,465\,074\,4}}19\,249\,071\,599\,532\,152\,229\,544\,643\,140\,624,$$

$$N_{DD_1}(52) = \underline{\mathbf{193\,052\,869\,315\,181\,895\,879\,191\,890\,783\,465\,074\,4}}77\,684\,913\,045\,479\,424\,283\,000\,117\,531\,248;$$

therefore DD_1 increases the success probability for a standard deck from 0.046… to a value 10^{-38} greater. But for other decks DD_1 can be worse: using the N-formulas for the two strategies one can show that DD_1 beats the basic strategy only when $n \equiv 0, 1 \pmod 4$. But one can do much better with DD_2 or DD_∞ (see Figure 1). Corollary 6(c) gives this exact value for $N_{DD_\infty}(52)$:

$$N_{DD_\infty}(52) = \underline{\mathbf{193\,05}}7\,905\,530\,197\,289\,909\,593\,674\,828\,337\,893\,152\,548\,821\,619\,003\,487\,812\,583\,424\,000\,000,$$

which means that $N_{DD_\infty}(52)$ exceeds $N_{\text{basic}}(52)$ by about $5 \cdot 10^{63}$. This increases the probability of success from 4.60283% to 4.60295%, a small improvement, but much better than the 10^{-38} gain from DD_1. To break it down further, the ideal count for DD_∞ when $n = 52$ causes a loss of about 10^{34}, but the number of strong moves yields a gain of about $5 \cdot 10^{63}$.

The symmetric polynomial approach leads to a computer implementation that allows extensive investigation into the question of the best strategy. Note that naively there are $52^{52} \sim 10^{89}$ strategies for $n = 52$, but they can be pruned down to 281589, the number of integer partitions of 52; this is because strategies corresponding to the same integer partition of n (such as $\langle 1, 1, 2, 3, 4 \rangle$ and $\langle 3, 3, 5, 2, 1 \rangle$ with $n = 5$) lead to the same count.

Corollary 7. *For* $9 \leq n \leq 60$, DD_∞ *is the best strategy for the short game.*

Proof. For this range one can compute N_σ for all strategies using Corollary 5 and the very fast recurrence for the symmetric polynomials. The number of strategies is the same as the number of integer partitions of n; for $n = 60$ this is just under one million. ■

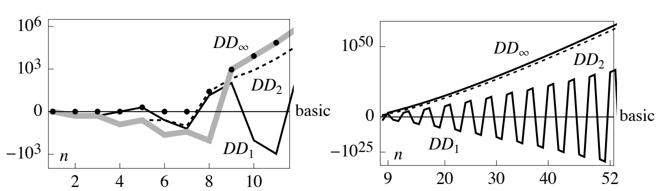

Figure 1. A signed logarithmic view of the success counts for DD_1, DD_2, and DD_∞, with $N_{\text{basic}}(n)$ subtracted. The dots in the left graph show the best strategy among these four. For $n \geq 9$, the full double-door strategy DD_∞ is apparently always the best.

Figure 1 compares the basic strategy, DD_1, DD_2, and DD_∞. One might guess that, as n gets very large, the double-door method will lose out to one that uses triple doors. But we checked cases up to $n = 1000$ and this never happens. Thus the evidence is strong that DD_∞ is best for all

$n \geq 9$.

Open Question 2. Is it true that for $n \geq 9$ there is no strategy that outperforms the full double-door strategy DD_∞ for the short game?

While DD_2 is not the best algorithm, the fact that its ideal count is constant as n varies allows a general comparison with the basic method, yielding a proof that it is better than the basic method whenever $n \geq 8$.

Theorem 8. *For all $n \geq 8$, the double double-door strategy DD_2 beats the basic strategy.*

Proof. Let Δ be the difference $T_{DD_2}(n) - T_{\text{basic}}(n)$, divided by $(n-4)!$. Let $m = \lfloor n/2 \rfloor$. Combining terms in the formula of Corollary 6(b) and the sum representation of the transposition count $T(n)$ yields $\Delta = \sum_{k=1}^{m} \frac{(-1)^k}{2^k\, k!} 4k\,(2k-1)\left(n^2 - 5n + 3 + 5k - 2k^2\right)$. Evaluating this symbolically (with *Mathematica*'s help) and simplifying the complicated expression that results using the reduction $\Gamma\left(b, -\frac{1}{2}\right) = (b-1)\,\Gamma\left(b-1, -\frac{1}{2}\right) + \left(\frac{-1}{2}\right)^{b-1} \sqrt{e}$ yields

$$\Delta = \frac{m\,(m^2+3m+2)}{2^{m-1}\sqrt{e}\;(m+2)!}\left(2^m\,(m-1)\,\Gamma\left(m-1, -\frac{1}{2}\right) + (-1)^m\, 2\,\sqrt{e}\,\left(n^2 - 5n - 2m^2 + m + 7\right)\right).$$

Claim. If $m \in \mathbb{N}$ and $m \geq 2$, then $\Gamma\left(m-1, -\frac{1}{2}\right) > \frac{1}{2}\,(m-2)!$.

Proof of claim. The first two cases can be individually checked, so assume $m \geq 4$. Then $\Gamma\left(m-1, -\frac{1}{2}\right) = \int_0^\infty e^{-t}\, t^{m-2}\, dt + \int_{-1/2}^{0} e^{-t}\, t^{m-2}\, dt \geq (m-2)! - \frac{\sqrt{e}}{2^{m-1}} > (m-2)! - 1 \geq \frac{1}{2}(m-2)!$ where the integral is bounded by a rectangular area.

The claim allows the elimination of the Γ term in Δ to get Δ^-, a lower bound on Δ:

$$\Delta^- = \frac{1}{\sqrt{e}} + \left(\frac{-1}{2}\right)^{m-2} \frac{1}{(m-1)!}\left(n^2 - 5n - 2m^2 + m + 7\right).$$

Reintroducing the factorial, it is easy to see that $(n-4)!\,\Delta^- > 17$ for $n \geq 8$. The floors of the first four values are 17, 147, 391, 2321. For $n \geq 12$, the second term of the sum is easily seen to be under $9/16$ in absolute value, so $(n-4)!\,\Delta^-$ is at least 1775. Therefore $(n-4)!\,\Delta \geq 17$ when $n \geq 8$ and the gain in strong moves outweighs the loss in the ideal count, which is $8 - 1 = 7$. This means $N_{DD_2}(n) > N_{\text{basic}}(n)$ for such n. ■

3. THE LONG GAME. The setup is as before, but now Bob can flip over as many cards as needed in his search for the target and the goal is to minimize the number of flips. Perhaps Charlie will free U prisoners where U counts the unexamined cards when the target is found; the prisoners therefore want to maximize the expected value of U.

Problem 3 (The Long Game). Same setup as in Problem 1, but Bob turns over cards until he finds the target. What strategy minimizes the expected number of flips until the target C is found?

The basic cycle-chasing strategy is a natural starting point. Bob opens door C and thereafter opens the door corresponding to the last card he has seen. If $p = (p_i)$ is the cycle-type of the permutation Bob faces, then $\frac{1}{n}\sum p_i^2$, or $\frac{1}{n}\,\|p\|^2$, is the expected number of flips because p_i/n is the chance he lands in the ith cycle and p_i is the flip count when he does. It is easy to see from this that Alice's best move is exactly as in Problem 1: bisect the longest cycle. For a type p, let $\hat{p}$ be the post-Alice type: the largest entry in p is replaced by its two equal, or nearly equal, halves.

For any strategy σ, let $E_\sigma(n)$ be the expected number of flips for Bob to find the target when he uses σ. To compute $E_{\text{basic}}(n)$ one must first consider all permutations that Charlie might choose. The collection of types of permutations of $\{1, \ldots, n\}$ is identical to the set of integer partitions of n. For any type p, let ν_p be the number of permutations having that type and let τ be the frequency vector of p.

Lemma 9. *For any permutation type $p = (p_i)$, let $n = \sum p_i$. Then $\nu_p = \frac{n!}{\prod p_i\, \prod(\tau_k!)}$.*

Proof. Turn any of the $n!$ permutations into disjoint cycles by partitioning the permutation list from left to right into sequences of length p_i. Each p_i-cycle occurs p_i times, hence the division by the p product. And the τ_k cycles of the same length will appear $\tau_k!$ many times, requiring division by the $\tau!$ product. ■

Now one gets the expected flip count as a weighted average; here $\hat{p}$ refers to the type after Alice's bisection.

Proposition 10. $E_{\text{basic}}(n) = \frac{1}{n \cdot n!} \sum_{all\ types\ p} \nu_p \ \|\hat{p}\|^2$.

Proof. The average is over $n \cdot n!$ items because Charlie controls the permutation and the target. Weighting the types according to how often each occurs gives the formula. ■

Applying the proposition to $n = 3$ gives $11/9$ as the expected flip count. For $n = 52$, there are 281 589 partitions and we have

$$E_{\text{basic}}(52) = \frac{1}{52 \cdot 52!} \, 637403126188754477757115794980410374714731524507812313146920938021683 22,$$

about 15.197.

Proposition 11. *If Bob is acting alone then, regardless of his strategy, $E(n) = (n+1)/2$.*

Proof. Every door covers a random card: that card is C with probability $1/L$, where L is the number of unturned cards at that point. It follows easily that the expected flip count is $\frac{1}{n} \sum_{i=1}^{n} i$ as claimed. ■

Computations indicate that $E_{\text{basic}}(n)$ is asymptotic to $(1 - 0.4267\ldots)(n+1)$, so that Alice's switch leads to a 43% reduction in the expected flip count. In fact, this 0.4267 constant is $G_{1,2}$, the second moment version of the Golomb–Dickman constant [**5**, eqn. 12]. The derivation that follows was pointed out to us by Richard Stong.

Suppose Charlie's permutation has type $p = (p_1, \ldots, p_k)$ in nonincreasing order. Using the basic strategy but with no help from Alice, Bob's expected number of flips, is $\frac{1}{n} \|p\|^2$, which, averaged over all of Charlie's choices, must equal $(n+1)/2$ (by Proposition 11). But Alice finds the longest cycle-length, p_1, and splits that cycle into two pieces of size roughly $p_1/2$; this reduces the expected flip-count to $\frac{n+1}{2} - \frac{p_1^2}{n} + \frac{p_1^2}{2n} = \frac{n+1}{2} - \frac{n}{2}\left(\frac{p_1}{n}\right)^2$ where, if p_1 is odd, we ignore a $1/(2n)$ term, which is asymptotically 0.

The asymptotic distribution of $(p_1/n)^2$ was worked out by Shepp and Lloyd [**5**, eqn. 12], who proved that $E\left(\left(\frac{p_1}{n}\right)^2\right) \sim \frac{1}{2}\int_0^\infty x\, e^{\text{Ei}(-x)-x}\, dx = \frac{1}{2}\int_1^0 \ln(x)\, e^{\text{li}(x)} dx$, where $\text{Ei}(x)$ is the exponential integral $-\int_{-x}^\infty e^{-y} y^{-1}\, dy$ and $\text{li}(x)$ is the logarithmic integral $\int_0^x 1/\ln x\, dx$. (Note that $\text{Ei}(-x) = -\Gamma(0, x)$.) This limiting expected value is $G_{1,2} = 0.4267\ldots$. The first-order version of this constant is $G_{1,1} = E\left(\frac{L}{n}\right) = \int_0^1 e^{\text{li}(x)}\, dx \sim 0.624$; it has interesting connections to number theory (see [**2**; §3.10]). We now get Bob's expected count as

$$\frac{1}{2}(n+1) - \frac{n}{2} G_{1,2} \sim \frac{1}{2}(n+1) - \frac{1}{2}(n+1)\, G_{1,2} = 0.57\ldots \frac{n+1}{2}.$$

To get an idea of the speed of convergence: $E_{\text{basic}}(10) = 3.177$ whereas the asymptotic prediction is 3.153; for $n = 52$, the true expectation is 15.197 and the prediction is 15.193.

A Signaling Strategy. As in the short game, the elegant basic strategy is not the best; a subtle signaling strategy is far superior. The heart of the *signal method* is that Alice can use door 1 to signal Bob to use a certain permutation ρ when doing his basic cycle-chasing. Bob will start with door 1 to learn the signal; if he sees the target there, he is successful in one flip. Otherwise, he follows the basic strategy except that whenever he is supposed to open door i he instead opens door $\rho(i)$. Ignoring the signal-reading move, Bob is replacing the permutation $\hat{\pi}$ that he faces with $\rho \circ \hat{\pi}$, the point being that the cycle structure of $\rho \circ \hat{\pi}$ can be much more friendly to cycle-chasing than that of $\hat{\pi}$. Further, the opening of door 1 is not always a wasted move: the target might be right there; and door 1, already open, might show up in the cycle chase. Here are the exact details; we use π for Charlie's permutation so $\hat{\pi}$ is the permutation facing Bob.

The Signal Method

1. Alice and Bob preselect n permutations $S = \{\rho_i\}$ of $\{1, \ldots, n\}$.

2. Bob's instructions: He hears the target C and opens door 1 to learn $j = \hat{\pi}(1)$. If $j = C$, he stops. Otherwise, he opens door $\rho_j(C)$. If he sees card U there, he opens door $\rho_j(U)$. He repeats this process until he finds C, which he must because that is the only way to close the cycle at door $\rho_j(C)$.

3. Alice makes the switch that helps Bob the most. She considers her $\binom{n}{2} + 1$ possible moves (the +1 is because she might do nothing) and chooses the one that minimizes Bob's expected flip-count over all possible targets.

Alice will typically place a more favorable card in the signal position, door 1; but the card behind door 1 might already be the best, taking into account that Alice can use her switch on the other cards to make Bob's use of $\eta_{\pi(1)}$ more efficient.

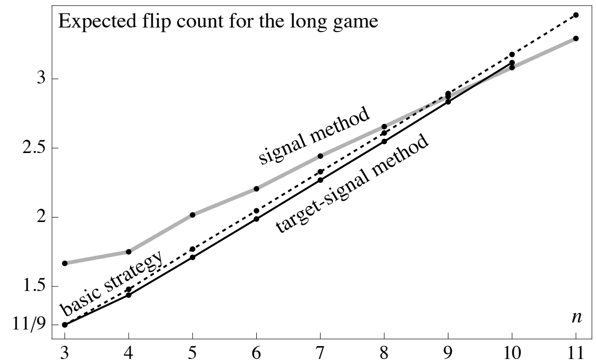

Figure 2. The signal method, using the best S-choice we found, becomes better than the basic strategy at $n = 9$. For $n = 4$, 10, and 11, S in the signal method is the cycle-power set S_n. The target-signal method uses the set S consisting of transpositions (2 3), (1 3), (1 2), and $n - 3$ copies of the identity and is the best of the three for $n \leq 9$.

Step 1 is the crux because it is not clear how to find a good choice for S. The desired behavior is that for many of Charlie's shuffles π, there should be some $\rho_i \in S$ and some move for Alice so that $\rho_i \circ \hat{\pi}$ has many small cycles. Perhaps this heuristic can lead to a precise method for finding good sets S. We have not found a way to do that, but in fact even random choices of S work reasonably well. By using some randomness and various search methods we found workable choices for S, but for definiteness consider powers of a cycle. Let $S_n = \{\gamma^j : j = 0, 1, \ldots, n-1\}$, where γ is the cycle $n \to n-1 \to n-2 \to \cdots \to 1 \to n$. For $n = 10$, using S_{10} against all $10 \cdot 10!$ choices of initial permutations and targets leads to $E_{\text{signal}, S_{10}} = 3.081\ldots$, a 3% improvement on $E_{\text{basic}}(10) = 3.1767$. Moreover, if we use S_{10} but only 10000 random initial shuffles (and all targets) we get 3.079, a reasonable approximation to the truth; this suggests that such a simulation might indicate the truth for $n = 52$. Doing this with S_{52} and 10^6 trials gives 10.73 as an estimate for $E_{\text{signal}, S_{52}}(52)$, a huge improvement over the 15.2 of $E_{\text{basic}}(52)$.

Figure 2 shows how the signal method using the best choice for S we found improves on the basic method when n is 9, 10, or 11. For $n \leq 9$ the data in Figure 2 are based on somewhat randomly found choices for S. But for $n = 5$ we know that the S we have is optimal, thanks to some exhaustive searching by Mark Rickert; in cycle notation, {identity, (3 4 5), (3 5 4), (2 4 3), (2 3 5 4)} is the best choice. All this leads to an extremely difficult question, since there are $n!^n$ choices for S.

Open Question 3. What is the best choice of S in the signal strategy? Can one improve on our expected values for $6 \leq n \leq 11$?

We can ask if there are shortcuts to the computation that avoid looking at all 52! choices for Charlie and give a proved value or good estimate for the result of using S_{52}.

Open Question 4. *Prove that for $n = 52$, the signal method for some choice of S is better than the basic strategy.*

The Target-Signal Method. It seemed that the basic strategy might be optimal for $n \leq 8$, but a small variation to the signal method led to a strategy that beats it handily for small n. This proves that the basic strategy is not optimal for $4 \leq n \leq 11$; the values in Figure 2 indicate that this is likely to be true for all n. The variation is the *target-signal strategy:* instead of using door 1 for the signal, Bob assumes that the signal is behind door C, where C is the target. Of course, Alice then won't know which door will signal the permutation ρ_i that Bob will use; but she can look at all possibilities and make the move that is the best choice on average.

We found a pattern for S that yields good results. Let S consist of the transpositions (2 3), (1 3), and (1 2) together with $n - 3$ copies of the identity. Figure 2 shows the experimental results. For $n = 4$ this choice of S leads to an expected flip count of $23/16 \sim 1.44$, compared to 1.48 for the basic method and 1.67 for the signal method. The target-signal method is better than any other method we know for $3 \leq n \leq 9$.

We have presented three strategies here, but perhaps (as asked in Problem 3) there is a completely different strategy that yields small expected values. For $n = 3$ one can examine all possible strategies ($3^3\, 2^6 = 1728$ of them) and conclude that the basic strategy's 11/9 cannot be lowered, solving Problem 3 in this simple case.

4 CONCLUSION. The basic problem seems simple, but a detailed investigation led to many strands: the use of randomness, the challenge of programming various methods so as to get good data efficiently, and the connection to classical combinatorial and numerical analysis ideas such as the inclusion-exclusion principle and new formulas for values of the incomplete gamma function. Of course, there are additional variations one might consider. For example, what happens in the short game when Bob can open two doors? What happens in either game when Alice can make two transpositions? One wonders if there are connections to the structure of the symmetric group and its Cayley graph.

Acknowledgements. We thank Mark Rickert for his enthusiasm in helping with some of the computer searches, Richard Stong for sharing his expertise in the probabilistic study of permutations, and Veit Elser for discovering the name of the prison.

REFERENCES

1. Joe Buhler, Elwyn Berlekamp, Emissary Puzzle Column, Fall 2015, p. 11, Problem 5, http://www.msri.org/system/cms/files/168/files/original/Emissary-Fall-2015-Web-Corrected.pdf.
2. J. Lagarias, Euler's constant: Euler's work and modern developments, *Bull. Amer. Math. Soc.* **50** (2013) 527–628.
3. *The On-Line Encyclopedia of Integer Sequences*, http://oeis.org.
4. S. Plouffe, Exact formulas for integer sequences, Notes, 1993, http://vixra.org/abs/1409.0044.
5. L. A. Shepp, S. P. Lloyd, Ordered cycle lengths in a random permutation, *Trans. Amer. Math. Soc.* **121** (1966) 340–357, http://www.ams.org/journals/tran/1966-121-02.
6. S. Wagon, Round formulas for exponential polynomials and the incomplete gamma function, *INTEGERS* **16** (2016) #A79.
7. Wikipedia, 100 Prisoners Problem, https://en.m.wikipedia.org/wiki/100_prisoners_problem.
8. P. Winkler, *Mathematical Mind-Benders*, Wellesley, Mass., A K Peters/CRC Pr., 2007.

LARRY CARTER was a professor in the Computer Science and Engineering department at the University of California at San Diego, where his research interests included high-performance computing and computer architecture. Before UCSD, he was a Research Staff Member at IBM's Watson Research Center. In retirement, Larry enjoys having time for cooking, biking, music, and recreational mathematics.
La Jolla CA 92037
carter@cs.ucsd.edu

STAN WAGON is recently retired from Macalester College. His main interest is in using the power of modern software to visualize abstract mathematical concepts. His latest book is a second edition, with G. Tomkowicz, of *The Banach–Tarski Paradox*. Other interests include ski mountaineering, climbing, mushroom hunting, and playing the piano. He is one of the the founding editors of *Ultrarunning* magazine, but now finds that covering long distances is much easier on skis than in running shoes.
Macalester College, St. Paul, MN 55105
wagon@macalester.edu