Card Shuffling

By Evita Nestoridi
Associate Professor
Stony Brook University

Photo: Evita Nestoridi

1 Thinking of a Good Problem

When interacting with people who love math I cannot help but notice a certain pattern. It's that feeling when someone gives you a puzzle that you find interesting; you are ready to completely focus on it and spend the next hours, days, months or years thinking about it. Usually a good problem resists and it turns into a life long companion. Other times, one finds the right piece of the puzzle.

Most of the times, one cannot solve the open question right away. In fact it is very common to put the question aside, and do something else. But the question is still on the back of one's head and will resurface many times as one learns new techniques and tools. And many times, the more one learns, the more they realize how far away from the truth they are.

In this article, I thought I would share with you a card shuffling problem that I love and was able to make some progress working jointly with Megan Bernstein.

2 The Mathematical Setup for Card Shuffling

I have been studying card shuffling for ten years now and I am still fascinated by how broad the subject is. Depending on the card shuffling model, one might see tools such as representation theory, or FKG and censoring inequalities, or entropy, or curvature showing up towards studying the most common question: the mixing time of the card shuffle.

We start with a deck of n distinct cards. First of all, let's establish that the different configurations of the deck correspond to permutations in the symmetric group S_n. For example, when the cards are in order (just like when we buy them at the store) then this gives the identity. If we transpose the top two cards then this gives rise to the transposition $(1, 2)$ and so on.

One of the most famous shuffling models is random transpositions, according to which we pick two cards uniformly at random with repetition and we swap them. So how long does it take to shuffle a deck of cards this way? Diaconis and Shahshahani [DS81] proved that it takes $2n \log n$ steps to shuffle the deck. And while this is not the fastest way to shuffle the deck, it has a unique significance in probability theory. It is the first time that representation theory of the symmetric group was used to study card shuffling. At the same time, it is the best toy model for studying the way that genes are organized in DNA that is well understood.

3 Mixing Times

To define the mixing time of a card shuffle, let's first talk about the transition matrix. Let $x, y \in S_n$ and let

$P(x,y)$ denote the probability that after one shuffle the configuration of the deck is y, given that initially the configuration of the deck was x. For the case of random transpositions, we have

$$P(x,y) = \begin{cases} \frac{2}{n^2}, & \text{if } y = x(a,b) \text{ for some } a \neq b \\ \frac{1}{n}, & \text{if } y = x \\ 0, & \text{otherwise.} \end{cases}$$

Taking powers of t, we notice that $P^t(x,y)$ gives the probability that after t shuffles the configuration of the deck is y, given that initially the configuration of the deck was x. When the shuffle satisfies certain natural conditions, then

$$P_x^t(y) := P^t(x,y) \to \frac{1}{n!} \text{ as } t \to \infty,$$

for every $x, y \in S_n$. We usually study this convergence to the uniform measure U with respect to the total variation distance:

$$\|P_x^t - U\|_{\text{T.V.}} := \frac{1}{2} \sum_{y \in S_n} \left| P^t(x,y) - \frac{1}{n!} \right|.$$

The mixing time is the first time that this distance becomes small enough. Rigorously, let $\varepsilon > 0$, then

$$t_{\text{mix}}(\varepsilon) = \inf\{t > 0 : \max_{x \in S_n} \{\|P_x^t - U\|_{\text{T.V.}}\} \leq \varepsilon\}.$$

One way to study the mixing time is to pass to the ℓ_2 distance via Cauchy-Swartz. In the case of random transpositions, the transition matrix turns out to be symmetric. Let $\{\beta_i\}_{i=1}^{n!}$ be the real eigenvalues of P. It turns out that for random transpositions (and other nice shuffles),

$$-1 < \beta_{n!} \leq \ldots < \beta_1 = 1,$$

and

(1) $$4\|P_x^t - U\|_{\text{T.V.}}^2 \leq \sum_{i=2}^{n!} \beta_i^{2t},$$

for every $x \in S_n$. And in the case of random-transpositions, Diaconis and Shahshahani [DS81] determined the β_i using the character values of the irreducible representations of the symmetric group , evaluated at transpositions. Another model where representation theory was key to diagonalizing the transition matrix is the random-to-random card shuffle.

4 The Random-to-random Card Shuffle

When I was a graduate student, I was desperately trying to come up with a question of interest. And this question turned out to be a card shuffling problem. Consider a deck of n cards. Pick a card uniformly at random and remove it from the deck. Choose a position of the deck uniformly at random and place the card there. That's the random-to-random card shuffle. And one of the most common and natural question to ask is how many iterations are sufficient to shuffle the deck well enough.

I could have felt disappointed when my PhD adviser told me that this was a famous problem and that he conjectured that $\frac{3}{4}n\log n$ steps were sufficient and necessary for the deck to mix. On the contrary, I felt intrigued. I read its rich history and tried to understand what are the main difficulties. It was introduced by Diaconis and Saloff-Coste [DSC93], who proved a first upper bound for the mixing time of order $n\log n$. Uyemura-Reyes [UR] proved the mixing time is at most $4n\log n$ and at least $\frac{1}{2}n\log n$. Among his findings, Uyemura-Reyes found that there were at least $\sqrt{n}$ eigenvalues of the transition matrix that are basically equal to $1-\frac{1}{n}$. In combination with (1), this finding justifies the conjecture of Diaconis [Dia96] that the random to random exhibits cutoff at $\frac{3}{4}n\log n$. Saloff-Coste and Zúñiga [SCZ08] further proved that $2n\log n$ steps are enough to mix. Morris and Qin [QM17] improved this bound to $1.5324n\log n$. Subag [Sub13] proved the sharp bound for the mixing time of the form $\frac{3}{4}n\log n - \frac{1}{4}n\log n - O(n)$ by looking at the cards that haven't been selected yet.

I spent hours staring at my ceiling until I came up with the following probabilistic argument, which can also be found in my thesis:

Mark the first card that was selected. From now on, you mark an umarked card whenever you select it and place it directly above the top marked card or directly below any marked card. The main tool gives that given that all cards have been marked, then the deck is shuffled, i.e any configuration of the cards is equally likely.

Unfortunately, this was not good enough. It takes $2n\log n$ steps to mark all the cards. Around the same time, Dieker and Saliola made an important discovery, namely they found the eigenvalues of the transition matrix. In joint work with Megan Bernstein, we analyzed the behavior of eigenvalues and (1) bound to give the sharp upper bound $\frac{3}{4}n\log n - \frac{1}{4}n\log n$ for the mixing time.

On one hand, this was indeed great progress; on the other hand we still lack the full probabilistic motivation. Subag's work suggests that keeping track of the cards yet to be removed is enough to see that we need at least $\frac{3}{4}n\log n$ steps to shuffle the deck. But we don't have a similar statistic that guarantees that $\frac{3}{4}n\log n$ steps of the shuffle are also enough. This means that there is more work to do in order to actually understand the shuffling behavior of random to random. And of course, there are many other questions one could study, such as the limit profile or the separation distance cutoff.

REFERENCES

[Dia96] P Diaconis, *The cutoff phenomenon in finite markov chains*, 1659–64.

[DS81] Persi Diaconis and Mehrdad Shahshahani, *Generating a random permutation with random transpositions*, Z. Wahrsch. Verw. Gebiete **57** (1981), no. 2, 159–179. MR 626813

[DSC93] Persi Diaconis and Laurent Saloff-Coste, *Comparison techniques for random walk on finite groups*, Ann. Probab. **21** (1993), no. 4, 2131–2156. MR 1245303 (95a:60009)

[QM17] Chuan Qin and Ben Morris, *Improved bounds for the mixing time of the random-to-random shuffle*, Electron. Commun. Probab. **22** (2017), Paper No. 22, 7. MR 3635695

[SCZ08] L. Saloff-Coste and J. Zúñiga, *Refined estimates for some basic random walks on the symmetric and alternating groups*, ALEA Lat. Am. J. Probab. Math. Stat. **4** (2008), 359–392. MR 2461789 (2010a:60019)

[Sub13] Eliran Subag, *A lower bound for the mixing time of the random-to-random insertions shuffle*, Electron. J. Probab. **18** (2013), no. 20, 20. MR 3035748

[UR] J.-C. Uyemura-Reyes, *Random walks, semi-direct products and card shuffling*, ProQuest LLC, Ann Arbor, MI, 2002, Thesis

(Ph.D.)-Stanford University. MR-2703300.