Míchání karet a kombinatorické posloupnosti

Antonín Slavík

Abstrakt. V článku se věnujeme několika metodám míchání karet. Připomeneme klasický Fisherův–Yatesův algoritmus a popíšeme některé jeho modifikace. Tyto nestandardní metody míchání jsou z pohledu matematiky mnohem zajímavější a souvisejí s některými známými kombinatorickými posloupnostmi.

Matematikové milují hrací karty. Svědčí o tom rozsáhlá literatura věnovaná nejrůznějším matematickým problémům souvisejícím s kartami. Někdy čtenář vystačí s elementární matematikou, jako např. v knize [11] popisující karetní triky s matematickým jádrem. Jindy jsou zapotřebí hlubší znalosti teorie pravěpodobnosti, kombinatoriky, algebry a dalších disciplín, jako např. v knize [2], která pojednává o míchání karet.

Karetní hráči používají různé metody míchání (viz např. [13]). Pokud očíslujeme karty přirozenými čísly z množiny $\{1, \ldots, n\}$, kde n je počet karet, pak výsledkem míchání je permutace této množiny. V samotném procesu míchání bývá přítomen prvek náhody, jedná se tedy o náhodnou permutaci.

Jakmile zvolíme metodu míchání, nabízejí se dvě otázky: Může být výsledkem míchání libovolná permutace množiny $\{1, \ldots, n\}$? A pokud ano, je míchání spravedlivé v tom smyslu, že všechny permutace jsou stejně pravděpodobné? Algoritmus míchání s touto vlastností budeme označovat jako nestranný. Běžně používané metody nejsou nestranné, ale při dostatečně dlouhém míchání jsou pravděpodobnosti všech permutací téměř vyrovnané. Těmto aspektům se podrobně věnuje kniha [2].

V tomto příspěvku se zaměříme na některé metody míchání, které pro karetní hráče nejspíše nebudou příliš atraktivní, avšak z pohledu matematiky jsou velmi zajímavé. Začneme s klasickým Fisherovým–Yatesovým algoritmem, což je nejjednodušší nestranná metoda míchání, a poté prozkoumáme některé modifikace. Zjistíme, že výpočet pravděpodobností pro různé permutace vede na posloupnosti čísel, které jsou známé ze zcela jiných kombinatorických úloh! Výklad je z velké části založen na výsledcích z článku [12]; tam lze najít i důkazy tvrzení, které zde vesměs vynecháváme.

1. Fisherův–Yatesův algoritmus

Nejjednodušší postup nestranného míchání karet plyne přímo z definice permutace jakožto uspořádané n-tice bez opakování. Na první pozici vybereme libovolné z n čísel a vyškrtneme je z množiny $\{1, \ldots, n\}$. Poté postupujeme dále a na i-tou pozici vybíráme rovnoměrně náhodně z $n - i + 1$ dosud nevyškrtnutých čísel. Tento přímočarý algoritmus je pojmenován (viz např. [5]) po britských statisticích R. Fisherovi a F. Yatesovi, kteří jej popsali ve třetím vydání statistických tabulek [4], str. 26–27; o oblibě této knihy svědčí skutečnost, že se dočkala celkem 6 vydání a 4 dotisků.

Doc. RNDr. Antonín Slavík, Ph.D., Matematicko-fyzikální fakulta UK, Sokolovská 83, 186 75 Praha 8, e-mail: `slavik@karlin.mff.cuni.cz`

Obr. 1. František Tichý: *Kouzelník s kartami* (1934, Národní galerie Praha)

Chceme-li Fisherův–Yatesův algoritmus naprogramovat, je nutné postupovat šikovně. Při nevhodné implementaci by každá aktualizace seznamu dosud nevyškrtnutých čísel mohla zabrat lineární čas a celková časová složitost by byla kvadratická. Ve skutečnosti lze algoritmus jednoduše implementovat v lineárním čase. Postup popsal R. Durstenfeld [3] a následně jej zpopularizoval D. E. Knuth [9], s. 145: Použijeme pole délky n, v jehož levé části bude postupně vznikat hledaná permutace, a v pravé části budou dosud nepoužitá čísla. Začneme tím, že do pole umístíme čísla $1, \ldots, n$ (v tomto pořadí). Poté pole procházíme zleva doprava a v i-tém kroku, kde $i \in \{1, \ldots, n\}$, vybereme číslo, které bude na i-té pozici permutace; uděláme to tak, že i-tý prvek pole vyměníme s j-tým prvkem, kde j je náhodně vygenerované číslo z množiny $\{i, \ldots, n\}$. Po provedení tohoto kroku je na začátku pole prvních i čísel permutace a na dalších místech jsou dosud nepoužitá čísla (nikoliv nutně ve vzestupném pořadí, což ovšem nevadí). Na konci výpočtu pole obsahuje náhodnou permutaci množiny $\{1, \ldots, n\}$. Pokud máme k dispozici generátor náhodných čísel s rovnoměrným rozdělením, pak každá permutace se na výstupu objeví s pravděpodobností $1/n!$.

Pro účely následujícího výkladu algoritmus přeformulujeme v řeči karet. Na začátku máme n karet, které rozložíme na stole vedle sebe. Karty procházíme zleva doprava a v i-tém kroku, kde $i \in \{1, \ldots, n\}$, vyměníme i-tou a j-tou kartu, kde j vybíráme náhodně z množiny $\{i, \ldots, n\}$. Dochází tak pouze k výměnám mezi i-tou kartou a kartami, které jsou dále vpravo (resp. pro $j = i$ k žádné výměně nedojde).

2. Modifikované výměnné míchání

D. P. Robbins a E. D. Bolker si v článku [12] položili přirozenou otázku: Co kdybychom předchozí proces míchání karet upravili tak, že i-tou kartu bude možné vyměnit nejen s kartami, které leží vpravo, ale s jakoukoliv jinou kartou? Číslo j tedy budeme v každém kroku volit rovnoměrně náhodně z množiny $\{1, \ldots, n\}$.

Je zřejmé, že výstupem nového algoritmu může být libovolná permutace množiny $\{1, \ldots, n\}$, neboť totéž platilo i pro původní algoritmus. Na první pohled však nemusí být jasné, zda se každá permutace objeví na výstupu se stejnou pravděpodobností. Záporná odpověď plyne z jednoduchého pozorování: V každém kroku je n možností, jak zvolit číslo j. Celkem tedy existuje n^n možností, jak se algoritmus může zachovat. Počet všech permutací je $n!$ a pokud by byl algoritmus nestranný, musela by se každá z nich objevit na výstupu v právě $n^n/n!$ případech. Tento podíl však pro $n \geqq 3$ není celočíselný, což je spor. Zjistili jsme, že algoritmus upřednostňuje některé permutace před jinými. Které permutace jsou nejčastější, které naopak nejméně časté a jaké jsou jejich pravděpodobnosti? Uveďme hlavní výsledky z článku [12]:

- Pravděpodobnost, že se karty posunou cyklicky doprava, tj. výstupem bude permutace $(n, 1, 2, \dots, n-1)$, je $2^{n-1}/n^n$. Pravděpodobnost cyklického posunu doleva, tj. permutace $(2, 3, \dots, n, 1)$, je C_n/n^n, kde

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

- Pravděpodobnost identické permutace je T_n/n^n, kde $T_0 = T_1 = 1$ a $T_n = T_{n-1} + (n-1)T_{n-2}$ pro $n \geqq 2$; explicitní vzorec je

$$T_n = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} \frac{(2k)!}{2^k k!}.$$

Autoři dále ukázali, že žádná permutace nemá nižší pravděpodobnost než $2^{n-1}/n^n$, tj. cyklické posuny vpravo se na výstupu objevují nejméně často. V porovnání s průměrem je jejich pravděpodobnost dokonce výrazně nižší, neboť počet příznivých případů (z celkových n^n) připadajících na jednu permutaci je v průměru

$$\frac{n^n}{n!} \sim \frac{\mathrm{e}^n}{\sqrt{2\pi n}}$$

(použili jsme Stirlingův vzorec) a pro $n \to \infty$ tato čísla rostou asymptoticky rychleji než 2^{n-1}. Naopak cyklický posun vlevo má pravděpodobnost výrazně nadprůměrnou, neboť

$$C_n \sim \frac{4^n}{\sqrt{\pi n^3}}$$

roste asymptoticky rychleji než $n^n/n!$. Pravděpodobnost identické permutace je ještě vyšší, neboť

$$T_n \sim \frac{n^{n/2}}{\sqrt{2}\mathrm{e}^{n/2-\sqrt{n}+1/4}}$$

roste asymptoticky rychleji než C_n.

Otázka, která permutace má nejvyšší pravděpodobnost, zůstala v článku [12] nezodpovězena. Autoři však ukázali, že pro $n \to \infty$ pravděpodobnost žádné permutace není asymptoticky větší než pravděpodobnost identické permutace. Pro nízká čísla n použili počítač; např. pro $n = 6$ zjistili, že permutace $(3, 1, 2, 6, 4, 5)$ má nejvyšší pravděpodobnost $159/6^6$. Již víme, že nejnižší pravděpodobnost $32/6^6$ má cyklický posun vpravo. Je překvapivé, jak moc se pravděpodobnosti různých permutací liší!

Po více než 20 letech problém nalezení permutace s nejvyšší pravděpodobností úspěšně vyřešili D. Goldstein a D. Moews [7]: Ukázali, že identická permutace má nejvyšší pravděpodobnost pro všechna $n \geqq 18$, zatímco pro $n \in \{4, \ldots, 17\}$ je nejpravděpodobnější permutací složení jistých dvou cyklů.

Je pozoruhodné, že v právě popsaných výsledcích se objevují dvě klasické kombinatorické posloupnosti, Catalanova čísla C_n a telefonní čísla T_n, která na první pohled nemají s mícháním karet nic společného: C_n udává např. počet způsobů, jak rozdělit konvexní $(n+2)$-úhelník pomocí neprotínajících se úhlopříček na trojúhelníky, nebo též počet mřížových cest z bodu $(0, 0)$ do bodu (n, n), kde každý krok vede vpravo nebo nahoru a cesta nikdy nevystoupí nad diagonálu. Číslo T_n odpovídá např. počtu involucí n-prvkové množiny (permutací, kde každý nezávislý cyklus má délku maximálně 2) nebo počtu párování v úplném grafu na n vrcholech (podrobněji viz [8]).

Dále prozkoumáme jiný způsob míchání karet, který rovněž vede na známou kombinatorickou posloupnost.

3. Extrakční míchání

Další algoritmus míchání karet popsaný v článku [12] funguje následovně: Karty opět rozložíme do řady, jejich počet je $n \in \mathbb{N}$. Míchání pak sestává z $m \in \mathbb{N}$ kroků; v každém z nich náhodně vybereme libovolnou kartu (říkejme jí aktivní karta) a přesuneme ji na levý okraj řady.

Pokud $m = n$ (případně $m > n$), pak máme jistotu, že výsledkem míchání může být libovolná permutace. K získání libovolné permutace $\pi = (\pi(1), \ldots, \pi(n))$ stačí např. postupně volit aktivní karty s hodnotami $\pi(n), \pi(n-1), \ldots, \pi(1)$.

Proces míchání je jednoznačně určen volbou aktivní karty v každém z m kroků a celkový počet možností je n^m. Podobně jako u modifikovaného výměnného míchání vidíme, že n^m obecně není dělitelné číslem $n!$, tudíž extrakční míchání není nestranné a upřednostňuje některé permutace před jinými.

Určíme počet způsobů, jak získat libovolnou permutaci množiny $\{1, \ldots, n\}$. Uvědomíme si, že po skončení míchání jsou karty, které nebyly nikdy aktivní, uspořádány vzestupně na konci řady, zatímco na jejím začátku jsou aktivní karty uspořádané podle toho, kdy byly naposledy aktivní (jedna karta může být aktivní vícekrát).

Pokud bychom např. měli $n = 8$ karet a výsledkem míchání by byla permutace

$$\pi = (6, 7, 2, 8, 4, 1, 3, 5), \tag{1}$$

pak množina karet, které nebyly nikdy aktivní, může být buď $N = \{1, 3, 5\}$, nebo $N = \{3, 5\}$, nebo $N = \{1\}$, nebo $N = \emptyset$. V každém z těchto případů známe i množinu aktivních karet $A = \{1, \ldots, n\} \setminus N$ a její velikost $a = |A|$. Aby byl průběh míchání jednoznačně určen, potřebujeme vědět, ve kterých krocích byly jednotlivé karty aktivní. Pokud každé aktivní kartě přiřadíme čísla kroků, kdy byla aktivní, dostaneme rozklad množiny $\{1, \ldots, m\}$ na a neprázdných podmnožin.

Předpokládejme například, že počet kroků je $m = 10$ a posloupnost aktivních karet byla 1, 4, 7, 8, 6, 2, 2, 7, 6, 6. Čtenář může ověřit, že tento postup míchání vede k permutaci (1). Přitom karta 6 byla aktivní v časech 5, 9, 10, karta 7 v časech 3, 8, atd. Dostáváme tedy následující rozklad množiny $\{1, \ldots, 10\}$:

$$\{5, 9, 10\}, \quad \{3, 8\}, \quad \{6, 7\}, \quad \{4\}, \quad \{2\}, \quad \{1\}.$$

Rozklad je tvořen šesti množinami, protože aktivních karet je šest. Pořadí množin v rozkladu odpovídá pořadí aktivních karet v zápisu permutace (1). To má za následek, že maxima jednotlivých množin tvoří klesající posloupnost. Pořadí množin je důležité kvůli tomu, abychom mohli celý proces obrátit a uvědomit si, že každý rozklad $\{1, \ldots, m\}$ na a neprázdných podmnožin popisuje průběh míchání, jehož výsledkem je permutace (1). Skutečně, máme-li takový rozklad, seřadíme množiny tak, aby jejich maxima tvořila klesající posloupnost. Z počtu podmnožin zjistíme, kolik bylo aktivních karet. Ze zápisu (1) vyčteme, které karty to byly. A konečně ze samotného rozkladu plyne, kdy byly tyto karty aktivní.

Jaký je tedy počet případů vedoucích k permutaci (1), pokud má míchání např. $m = 10$ kroků? Nechť $\left\{ {r \atop j} \right\}$ značí počet rozkladů r-prvkové množiny na j neprázdných podmnožin. Jde o tzv. Stirlingova čísla 2. druhu, která jsou dobře známá z kombinatoriky a lze je počítat např. pomocí vztahu

$$\left\{ {r \atop j} \right\} = \frac{1}{j!} \sum_{l=0}^{j} (-1)^l \binom{j}{l} (j-l)^r.$$

Z výše uvedených úvah plynou následující počty způsobů, jak získat permutaci (1) v $m = 10$ krocích:

- Počet způsobů, kdy aktivní byly karty 6, 7, 2, 8, 4, je $\left\{ {10 \atop 5} \right\}$.
- Počet způsobů, kdy aktivní byly karty 6, 7, 2, 8, 4, 1, je $\left\{ {10 \atop 6} \right\}$.
- Počet způsobů, kdy aktivní byly karty 6, 7, 2, 8, 4, 1, 3, je $\left\{ {10 \atop 7} \right\}$.
- Počet způsobů, kdy aktivní byly karty 6, 7, 2, 8, 4, 1, 3, 5, je $\left\{ {10 \atop 8} \right\}$.

Celkový počet způsobů je tedy $\sum_{j=5}^{8} \left\{ {10 \atop j} \right\} = 71\,982$.

Zobecněním předchozích úvah dojdeme k následujícímu tvrzení, kde k značí minimální možný počet aktivních karet:

Nechť $\pi = (\pi(1), \ldots, \pi(n))$ je libovolná permutace množiny $\{1, \ldots, n\}$ a $m \in \mathbb{N}$. Pokud π je rostoucí, položme $k = 1$; v opačném případě nechť k je největší číslo takové, že $\pi(k) > \pi(k+1)$. Pak počet způsobů, jak získat π při extrakčním míchání karet s m kroky, je roven $\sum_{j=k}^{n} \left\{ {m \atop j} \right\}$.

Při extrakčním míchání s pevně zvoleným počtem kroků tedy pravděpodobnost libovolné permutace závisí pouze na hodnotě k. Hodnota $\sum_{j=k}^{n} \left\{ {m \atop j} \right\}$ je největší, když k je co nejmenší, tj. $k = 1$. To nastává pro permutace π splňující $\pi(2) \leqq \pi(3) \leqq \ldots \leqq \pi(n)$, mezi které patří například identická permutace. Naopak nejnižší pravděpodobnost mají permutace splňující $\pi(n-1) > \pi(n)$, pro které platí $k = n - 1$.

Např. pro $n = m = 6$ vychází nejvyšší možná pravděpodobnost

$$\frac{1}{6^6} \sum_{j=1}^{6} \left\{ {6 \atop j} \right\} = \frac{203}{6^6},$$

zatímco nejnižší pravděpodobnost je

$$\frac{1}{6^6}\sum_{j=5}^{6}\left\{{6 \atop j}\right\} = \frac{16}{6^6}.$$

Zjišťujeme, že extrakční míchání některé permutace značně zvýhodňuje.

Čísla $B_n = \sum_{j=1}^{n} \left\{{n \atop j}\right\}$, $n \in \mathbb{N}$, udávají počet rozkladů n-prvkové množiny na libovolný počet podmnožin; nazývají se Bellova čísla. Míchání karet nás tedy dovedlo k další známé kombinatorické posloupnosti: V případě extrakčního míchání s $m = n$ kroky udává B_n počet případů vedoucích k identické permutaci, resp. jakékoliv permutaci splňující $\pi(2) \leqq \pi(3) \leqq \ldots \leqq \pi(n)$.

4. Kuriozita na závěr

Překvapivý výskyt Bellových čísel v souvislosti s mícháním karet zaujal i známého popularizátora matematiky M. Gardnera, který o něm píše v knize [6], kapitola 2. Čtenář se dozví, že čísla B_n jsou pojmenována podle amerického matematika skotského původu E. T. Bella, který objevil rozvoj

$$\mathrm{e}^{\mathrm{e}^x} = \mathrm{e} \cdot \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n. \tag{2}$$

Gardner píše, že Bell k tomuto výsledku dospěl poté, co v jisté příručce objevil chybný rozvoj funkce $\mathrm{e}^{\mathrm{e}^x}$. Nahlédnutím do Bellova článku [1] však zjistíme, že Gardnerovo tvrzení je nepřesné: Nejednalo se o chybný rozvoj $\mathrm{e}^{\mathrm{e}^x}$, ale $\mathrm{e}^{\sin x}$. Českého čtenáře potěší zajímavá a málo známá kuriozita: Onou příručkou, kde se chyba vyskytla, byly tabulky vzorců [10] pražského astronoma, fyzika a matematika Václava Lásky. Odtud se pak chyba rozšířila do dalších zdrojů. Správný rozvoj má tvar

$$\mathrm{e}^{\sin x} = 1 + x + \frac{x^2}{2!} - \frac{3x^4}{4!} - \frac{8x^5}{5!} - \frac{3x^6}{6!} + \ldots,$$

zatímco v Láskových tabulkách na str. 33 má člen s x^6 opačné znaménko. Bell si chyby všiml a ve snaze ji opravit hledal rozvoje funkcí ve tvaru $\mathrm{e}^{f(x)-f(0)}$ za předpokladu, že známe rozvoj f. Jedním ze vztahů, které přitom objevil, byla identita (2).

Bellova čísla souvisejí s pravděpodobností identické permutace ještě při jiném způsobu míchání karet, který spočívá v tom, že v i-tém kroku náhodně vybereme kartu a přesuneme ji na pozici i. Výpočty pravděpodobností jsou u takového míchání složitější a zvídavý čtenář je najde opět v článku [12].

L i t e r a t u r a

[1] Bell, E. T.: *Exponential numbers.* Amer. Math. Monthly 41 (1934), 411–419.

[2] Diaconis, P., Fulman, J.: *The mathematics of shuffling cards.* American Mathematical Society, 2023.

[3] Durstenfeld, R.: *Algorithm 235: Random permutation.* Commun. ACM 7 (1964), 420–420.

[4] Fisher, R. A., Yates, F.: *Statistical tables for biological, agricultural and medical research.* 3rd ed., Oliver & Boyd, 1948.

[5] Fisher–Yates shuffle [online].
https://en.wikipedia.org/wiki/Fisher%E2%80%93Yates_shuffle

[6] Gardner, M.: *Fractal music, hypercards and more... Mathematical recreations from Scientific American Magazine*, W. H. Freeman and Company, 1992.

[7] Goldstein, D., Moews, D.: *The identity is the most likely exchange shuffle for large n.* Aequationes Math. 65 (2003), 3–30.

[8] Hubač, D.: *Permutace s předepsanými délkami cyklů.* PMFA 69 (2024), 75–96.

[9] Knuth, D. E.: *The art of computer programming, Vol. 2.* Seminumerical algorithms. 3rd edition, Addison–Wesley, 1998.

[10] Láska, V.: *Sammlung von Formeln der reinen und angewandten Mathematik.* Friedrich Vieweg und Sohn, 1894.

[11] Mulcahy, C.: *Mathematical card magic. Fifty-two new effects.* CRC Press, 2013.

[12] Robbins, D. P., Bolker, E. D.: *The bias of three pseudo-random shuffles.* Aequationes Math. 22 (1981), 268–292.

[13] Shuffling [online]. https://en.wikipedia.org/wiki/Shuffling