Jak znaleźć najlepszą żonę?

Andrzej GRZESIK[*]

Jak powszechnie wiadomo, znalezienie najlepszej żony (lub męża) nie należy do najprostszych zadań. Panuje też stereotypowe przekonanie, że dla matematyków to zadanie jest dużo trudniejsze niż dla „zwykłych ludzi". Jest to pogląd błędny – chociażby dlatego że, wbrew powszechnemu mniemaniu, matematyk to nie uwięziony w świecie książek samotnik. Co więcej, matematyk mógłby spróbować znaleźć jakąś naukową metodę pozwalającą wybrać najlepszą kandydatkę. Czy takie strategie w ogóle istnieją?

Warto znaleźć sobie sekretarkę

Poszukiwanie sekretarki jest znacznie łatwiejsze niż poszukiwanie żony, więc od tego zacznijmy.

Rozważmy takie zagadnienie – zgłasza się do nas wiele kandydatek na sekretarkę, będą po kolei przychodzić na rozmowę i musimy spośród nich wybrać tę najlepszą. Problem w tym, że w każdym momencie, gdy z którąś rozmawiamy, musimy się zdecydować, czy ją przyjmujemy, czy odrzucamy. Jeśli się na nią zdecydujemy, to z dalszymi już nie rozmawiamy, natomiast jeśli ją odrzucimy, do tej kandydatki wrócić już nie możemy. Jeżeli odrzucamy wszystkie, to zmuszeni jesteśmy do wybrania tej ostatniej. Czyli wiemy jedynie, ile kandydatek przyjdzie oraz jak przedstawia się aktualnie sprawdzana na tle tych, które już były. O tych, które pojawią się później, nie wiemy nic. Naszym zadaniem jest zmaksymalizowanie prawdopodobieństwa znalezienia najlepszej sekretarki.

Zauważmy najpierw, że opłaca się wybierać tylko spośród tych kandydatek, które okazały się lepsze od wszystkich poprzednich. To dlatego, że interesująca nas najlepsza sekretarka oczywiście spełnia ten warunek. Zauważmy też, że jeśli w pewnym momencie opłaca się nam zdecydować na jedną z takich kandydatek, to opłaca się także wybrać późniejszą kandydatkę, która jest lepsza od dotychczasowych (prawdopodobieństwo tego, że najlepsza sekretarka jest wśród dalszych kandydatek, maleje). Innymi słowy, najlepsza strategia to taka, w której najpierw odrzucamy k początkowych kandydatek spośród n, a następnie decydujemy się na pierwszą, która będzie lepsza od wszystkich dotychczasowych. Problemem jest teraz tylko znalezienie optymalnego progu.

Załóżmy, że najlepsza sekretarka jest na pozycji a w kolejce na rozmowę (tak jest z prawdopodobieństwem $\frac{1}{n}$). Jeśli $a \leqslant k$, to przy użyciu naszej strategii jej nie wybierzemy. Natomiast jeśli $a > k$, to wybierzemy ją wtedy, gdy najlepsza z kandydatek o numerach $1, 2, \ldots, a-1$ będzie wśród pierwszych k. W przeciwnym przypadku nie dotarlibyśmy do kandydatki numer a, tylko skończylibyśmy wcześniej. Czyli w tym przypadku prawdopodobieństwo naszej wygranej wynosi $\frac{k}{a-1}$. Zatem całkowite prawdopodobieństwo, że uda nam się znaleźć najlepszą sekretarkę, wynosi

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

Dla dużych n, czyli dużej liczby kandydatek, możemy przybliżyć sumę przez całkę, otrzymując z powyższego wzoru

$$\frac{k}{n}\int_k^n \frac{1}{x}\mathrm{d}x = \frac{k}{n}\ln\frac{n}{k} = -\frac{k}{n}\ln\frac{k}{n}.$$

Funkcja $-x\ln x$ przyjmuje swoje maksimum w punkcie $x=\frac{1}{\mathrm{e}}$, co można łatwo sprawdzić, obliczając pochodną. Zatem będziemy mieć największą szansę wygranej dla $\frac{k}{n}\approx\frac{1}{\mathrm{e}}$, czyli $k\approx\frac{n}{\mathrm{e}}\approx 0{,}37n$. Podstawiając tę wartość do wzoru, stwierdzamy, że prawdopodobieństwo wybrania najlepszej sekretarki wynosi $\frac{1}{\mathrm{e}}\approx 37\%$.

Zatem najlepsza strategia to odrzucić pierwszych 37% kandydatek, a następnie wybrać tę, która będzie lepsza od wszystkich do tej pory. To daje nam aż 37% szansy wybrania najlepszej ze wszystkich. Czyli matematyk w ponad 1/3 przypadków bierze, co najlepsze!

W literaturze można spotkać to zagadnienie jako tzw. problem sekretarki. Warto zaznaczyć, że dotyczy on nie tylko rekrutacji pracowników. Jeśli jedziemy samochodem i chcemy po drodze zatankować, to mijamy kilka stacji, obserwując cenę benzyny, a następnie tankujemy na takiej, która ma najlepszą cenę. Ile wynosi to najlepsze „kilka"? Obliczyliśmy wcześniej – około 37%.

[*]student, Wydział Matematyki i Informatyki, Uniwersytet Jagielloński

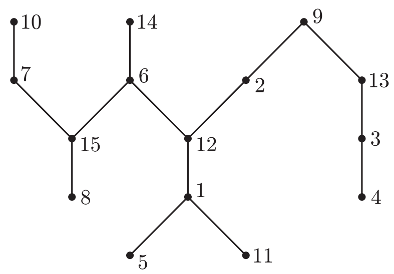
Nie można być wybrednym

Zauważmy, że jeśli będziemy mieć pecha i najlepsza kandydatka będzie wśród pierwszych 37% (a stanie się tak na... 37%), to dojdziemy do samego końca i weźmiemy ostatnią, która jest, krótko mówiąc, przeciętna. Zatem średnie miejsce, które zajmuje wśród wszystkich ta, którą wybierzemy, jest raczej dość odległe. Aby temu zapobiec, trzeba trochę zmodyfikować naszą strategię.

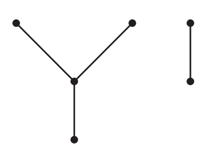
Jeśli przedostatnia testowana kandydatka zajmuje drugie miejsce wśród dotychczasowych, to lepiej zdecydować się na nią, a nie liczyć na to, że ostatnia będzie jeszcze lepsza. Ogólnie, łatwo zauważyć, że przedostatnia się „opłaca”, gdy jest w pierwszej połowie spośród dotychczasowych. Można obliczyć, kiedy warto wybrać trzecią od końca, czwartą, itd. W ten sposób dostaniemy strategię, która będzie zdefiniowana przez ciąg liczb określających, które miejsce wśród już widzianych może zajmować aktualna kandydatka, abyśmy się na nią zdecydowali.

Przykładowo, w przypadku 4 kandydatek taki ciąg to $(0, 1, 2, 4)$. To oznacza, że pierwszej nigdy nie przyjmiemy, drugą, gdy jest najlepsza do tej pory, trzecią, gdy jest 1 lub 2, a czwartą zawsze (jeśli do niej dojdziemy). W tym przypadku można obliczyć, że średnie miejsce tej, na którą się zdecydujemy, to 1,875. Nasuwa się pytanie: jak to wygląda dla większych wartości n? Okazuje się, że jest niesamowicie optymistycznie – dla 100 kandydatek taka strategia da nam sekretarkę z miejsca w przybliżeniu 3,6! Dla wartości n dążących do nieskończoności statystyczne miejsce wybranej kandydatki zbiega do około 3,83. Oznacza to, że nawet jeśli będzie 1 000 000 kandydatek, to i tak możemy przyjąć strategię, która da nam sekretarkę może nie najlepszą, ale na pewno ze ścisłej czołówki!

Trzeba pilnować porządku



Przykładowe uporządkowanie kandydatek na żonę. Krawędź pomiędzy wierzchołkami (kandydatkami) oznacza, że ta wyżej jest lepsza od tej niżej. Numery określają kolejność, w jakiej są sprawdzane.



Graf po sprawdzeniu 6 kandydatek.

Problem znalezienia żony od problemu znalezienia sekretarki różni się tym, że kandydatek na żonę nie da się porównywać tak łatwo, jak kandydatek na sekretarkę. Wiadomo – jedna lepiej gotuje, inna lepiej wygląda, jeszcze inna interesuje się matematyką bardziej niż pozostałe... Gdy „sprawdzamy” kolejne kandydatki, to nie możemy powiedzieć „ma trzecie miejsce na liście dotychczasowych”, tylko coś w rodzaju „była lepsza od tej, gorsza od tamtej, a z tą się nie da porównać”. Czyli kandydatki tworzą nam pewien porządek częściowy, który można przedstawić na grafie – jak obok.

Problem znalezienia najlepszej żony polega więc na tym, że sprawdzamy kandydatki w pewnej kolejności, na bieżąco tworzymy graf porządku częściowego i decydujemy, czy daną kandydatkę wybrać, czy też bezpowrotnie odrzucić. Najważniejsze pytanie brzmi: czy da się znaleźć taką strategię, aby dla dowolnej liczby kandydatek, niezależnie od tego, jaki graf tworzą ani w jakiej kolejności przychodzą, mieć pewną niezerową szansę znalezienia najlepszej. Przez najlepszą rozumiemy tu dowolną, od której nie ma lepszej, czyli element maksymalny porządku.

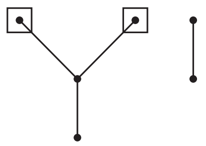
W 1980 roku Berezowski pokazał, że istnieje strategia, która daje nam zwycięstwo z prawdopodobieństwem równym co najmniej $a(1/w)^{w/(w-1)}$, gdzie a to liczba elementów maksymalnych, a w to szerokość grafu, czyli długość najdłuższego ciągu elementów nieporównywalnych. Ale to nie jest wynik, który nas satysfakcjonuje – wartość ta zależy od grafu, jaki tworzą kandydatki. Ponadto w pewnych przypadkach, na przykład gdy graf jest złożony z jednego elementu maksymalnego i wielu elementów mniejszych od niego, które są nieporównywalne między sobą, ta wartość może być dowolnie bliska zeru.

Najlepszy efekt daje wytrwałość i... rzut monetą

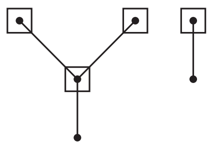
Przełom nastąpił w 1999 roku. Preater wykazał, że strategia, jakiej szukamy, istnieje i daje prawdopodobieństwo znalezienia najlepszej żony równe co najmniej 12,5%. Niedawno, bo w 2007 roku, ukazała się praca z udziałem polskich matematyków (Georgiou, Kuchta, Morayne, Niemiec), w której udowodniono, że strategia Preatera z drobną modyfikacją daje prawdopodobieństwo znalezienia najlepszej żony równe 25%!

Zgodnie z tą strategią, mając n kandydatek, należy po kolei:

- rzucić n razy monetą i odrzucić tyle początkowych kandydatek, ile wypadło orłów;
- rzucić raz monetą i jeśli wypadł orzeł, to zaznaczyć te wierzchołki, które mają największą wysokość, a jeśli wypadła reszka, to zaznaczyć wierzchołki, które mają największą lub o 1 mniejszą wysokość;
- wybrać pierwszą spośród następnych kandydatek, która będzie lepsza od jakiejś zaznaczonej i będzie najlepsza wśród dotychczasowych.



Graf po sprawdzeniu 6 kandydatek, gdy wypadł orzeł.



Graf po sprawdzeniu 6 kandydatek, gdy wypadła reszka.

Ta strategia daje nam prawdopodobieństwo 25%, że niezależnie od tego, jaki graf tworzą kandydatki na żonę i w jakiej kolejności przychodzą, wybierzemy właśnie tę najlepszą. Niesamowicie optymistyczny wynik – nawet jeśli bierzemy pod uwagę 3 miliardy kandydatek, to dzięki opisanej metodzie na 25% znajdziemy wśród nich najlepszą żonę!

No, ale to nie koniec. Nikt nie twierdzi, że lepiej się nie da. Wiemy, że wskazana wyżej strategia nie da lepszego prawdopodobieństwa, ale przecież może być inna, lepsza metoda. W opisanej strategii odrzucaliśmy średnio połowę kandydatek, a następnie losowaliśmy, które wierzchołki zaznaczamy – może lepiej będzie, jeśli po 1/3 kandydatek zaznaczymy wierzchołki o największej wysokości, a po 2/3 także te o wysokości o jeden mniejszej? A może trzeba szukać całkiem innej metody? Mam powody przypuszczać, że istnieje strategia dająca, jak w przypadku poszukiwań sekretarki, szansę równą $1/e$, czyli około 37%. Czekam tylko na kogoś, kto znajdzie tę strategię i udowodni. Zachęcam do poszukiwań – zarówno teoretycznych, jak i praktycznych!

Logarytm w liczbach naturalnych

*Jan SZEJKO**

Zajmiemy się problemem znalezienia funkcji ze zbioru liczb całkowitych dodatnich w zbiór liczb naturalnych, możliwie wiernie oddającej podstawowe własności logarytmu. Następujące zadanie jest pierwszą próbą określenia własności, które szukana funkcja ma spełniać.

Przyjmujemy, że $0 \in \mathbb{N}$.

Problem 1. *Czy istnieje funkcja $f : \mathbb{Z}_+ \to \mathbb{N}$, która nie jest stale równa* 0 *i spełnia warunki:*

1) *f jest niemalejąca,*
2) *dla dowolnych $k, l \in \mathbb{Z}_+$ zachodzi $f(kl) = f(k) + f(l)$?*

Nie jest to trudny problem; pokażemy, że, niestety, nie ma takiej funkcji.

Dowód. Dla każdego $n > 1$ zachodzi $f(n) > 0$, gdyby bowiem dla pewnego $n > 1$ było $f(n) = 0$, to również $f(n^k) = kf(n) = 0$ dla każdego $k \in \mathbb{Z}_+$. A stąd byłoby $f(n) \equiv 0$ ze względu na monotoniczność f, czyli sprzeczność.

Pokażemy, że $a^{f(b)} \geqslant b^{f(a)}$ dla $a, b > 1$. Dla każdej liczby wymiernej $\frac{p}{q}$, korzystając kolejno z warunków 2) i 1), otrzymujemy:

$$\frac{p}{q} > \frac{f(b)}{f(a)} \Leftrightarrow pf(a) > qf(b) \Leftrightarrow f(a^p) > f(b^q) \Rightarrow a^p > b^q.$$

Aby skorzystać z ciągłości potęgowania, trzeba zauważyć, że dla ustalonych a i b relacja między a^p i b^q zależy tylko od wartości ułamka $\frac{p}{q}$ (bo $a^p > b^q \Leftrightarrow a^{\frac{p}{q}} > b$), a następnie wziąć malejący ciąg liczb wymiernych zbiegający do $\frac{f(b)}{f(a)}$.

Potęgowanie jest ciągłe, więc musi być również $a^p \geqslant b^q$ dla $\frac{p}{q} = \frac{f(b)}{f(a)}$, stąd $a^{f(b)} \geqslant b^{f(a)}$, czego należało dowieść. Zauważmy ciekawą rzecz: nigdzie nie założyliśmy nic więcej o a i b oprócz tego, że $a, b > 1$. Zatem analogiczna nierówność zachodzi, gdy zamienimy a i b miejscami: $b^{f(a)} \geqslant a^{f(b)}$. Mamy więc równość $a^{f(b)} = b^{f(a)}$ dla $a, b > 1$. To jednak nie może być prawda, ponieważ dla $a = 2$, $b = 3$ dostajemy równość liczby parzystej i nieparzystej. $\square$

Wiemy już, że nie ma funkcji, która spełnia warunki 1) i 2). Możemy jednak się zastanowić, co będzie, jeśli któryś z nich osłabimy. Zastąpmy warunek 2) słabszym warunkiem 2′).

Problem 2. *Czy istnieje funkcja $f : \mathbb{Z}_+ \to \mathbb{N}$, która nie jest stale równa* 0 *i spełnia warunki:*

1) *f jest niemalejąca,*
2′) *dla $k, l \in \mathbb{Z}_+$ względnie pierwszych zachodzi $f(kl) = f(k) + f(l)$?*

*student, Wydział Matematyki, Informatyki i Mechaniki Uniwersytetu Warszawskiego