Algoritmy a grafy 1 (BI-AG1), Cvičení č. 7

Randomizace, hešování

Paralelka 104, Úterý 16:15-17:45

Cvičící: Šimon Lomič
lomicsim@fit.cvut.cz

Informace: lomicsim.github.io

Fakulta informačních technologií
České vysoké učení technické v Praze
https://courses.fit.cvut.cz/BI-AG1

(Verze dokumentu: 13. 11. 2018 13:53)

7.1 Generování náhodných čísel

Mějme instrukci `randombit()`, která vrací hodnoty {0,1} s rovnoměrnou pravděpodobností nezávisle na výsledku předchozích volání.

(a) Jak za pomocí této instrukce vygenerovat náhodné číslo z intervalu $[0, 2^k)$ pro zadané $k > 0$.

(b) Jak vygenerovat náhodné číslo z intervalu $[0, n), n \in \mathbb{N}$.

- Jaký je maximální počet volání `randombit()`?

7.2 Pravděpodobnostní prostor

- **Diskrétní pravděpodobnostní prostor** je dvojice $(\Omega, \mathbf{P})$, kde
  - Ω je konečná nebo spočetně nekonečná množina **elementárních jevů**.
  - $\mathbf{P} : \Omega \to [0, 1]$ je funkce, která elementárním jevům přiřazuje jejich **pravděpodobnosti** taková, že součet pravděpodobností všech elementárních jevů je roven 1, čili $\sum_{\omega \in \Omega} \mathbf{P}(\omega) = 1$.
- **Jev** je nějaká množina $A \subseteq \Omega$ elementárních jevů.
- **Pravděpodobnost jevu** A je: $\mathbf{P}(A) = \sum_{\omega \in A} \mathbf{P}(\omega)$.

Cvičení: Mějme klasickou šestistěnnou kostku. Hodíme s ní, a pokud padne liché číslo, hodíme ještě jednou

(a) Určete pravděpodobnost, že padne číslo 2.

(b) V jakém pravděpodobnostním prostoru (Ω, P) se tento experiment odehrává?

(c) Uveďte příklady jevů $A \subseteq \Omega$ a určete jejich pravděpodobnost.

7.3 Náhodná veličina, střední hodnota

- **Náhodná veličina** (proměnná) je funkce $X : \Omega \to \mathbb{R}$. Přiřazuje tedy každému elementárnímu jevu $\omega \in \Omega$ nějakou číselnou hodnotu $X(\omega)$.
- **Střední hodnota** $\mathbf{E}[X] = \sum_{x_i} x_i \cdot \mathbf{P}[X = x_i]$ (průměr všech možných hodnot veličiny X vážený jejich pravděpodobnostmi).
- **Čekáme-li na náhodný jev**, který nastane s pravděpodobností p, dočkáme se ho ve střední hodnotě v $1/p$-tém pokusu.

Cvičení:

(a) *Mějme klasickou šestistěnnou kostku. Hodíme s ní, a pokud padne liché číslo, hodíme ještě jednou.*
Určete střední hodnotu čísla, které padne na kostce.

(b) Určete střední hodnotu náhodné veličiny $X =$ počet použití instrukce `randombit()`, při generování náhodného čísla z intervalu $[0, n)$.

7.4 Pravděpodobnost, čekání na jev

(a) Popište algoritmus, který v lineárním čase vygeneruje náhodnou permutaci množiny $\{1, 2, \ldots, n\}$.

(b) Mějme velmi velký soubor (nevejde se do paměti). Navrhněte algoritmus, který vypíše rovnoměrně náhodný řádek a projde celý soubor pouze jednou (dopředu neznáme počet řádků).
  - (Za domácí úkol vyberte rovnoměrně náhodnou k-tici.)

(c) Uvažujte následující experiment: mějme n přihrádek a postupně do nich umisťujme čísla $\{1, 2, \ldots, n\}$ tak, že vždy vybereme rovnoměrně náhodně přihrádku a pokud v ní již něco je, vybíráme znovu. Určete střední počet pokusů volby přihrádky.

7.5 Domácí úkol

(a) Mějme velmi velký soubor (nevejde se do paměti). Navrhněte algoritmus, který vybere náhodnou k-tici řádků a projde celý soubor pouze jednou (každá k-tice musí mít stejnou pravděpodobnost). **(0.5 b)**

(b) Mějme zakořeněný strom T o n vrcholech. Navrhněte lineární algoritmus, který vypíše délku nejdelší cesty v T. **(0.5 b)**

Úkol odevzdejte na příštím cvičení (případně e-mailem).