

Metody Probabilistyczne i Statystyka

ZADANIE DOMOWE 3

Termin wysyłania (MS Teams): **22 grudnia 2023 r. godz. 23:59**

Promocja świąteczno-noworoczna: spóźnienia będą naliczane dopiero od 01.01.2024 r.

Zadanie 1. *The Power of Two Choices / ballanced allocation*, patrz np. rozdział 17 w [MU17]

Rozszerz symulacje z zadania domowego 2 o eksperymentalne wyznaczanie maksymalnej liczby kul w urnie po wrzuceniu n kul do n urn (*maximum load*) w przypadku, gdy:

- (a) dla każdej kuli wybieramy niezależnie i jednostajnie losowo jedną z n urn, w której umieszczamy kulę,
- (b) dla każdej kuli wybieramy niezależnie i jednostajnie losowo z powtórzeniami d urn i umieszczamy kulę w najmniej zapełnionej z wybranych urn („remisy” rozstrzygamy w dowolny sposób).

Przeprowadź podobne eksperymenty jak w zadaniu domowym 2 celem wyznaczenia $L_n^{(d)}$ – maksymalnego zapełnienia pojedynczej urny po wrzuceniu n kul – dla $d = 1$ (punkt (a)) oraz $d = 2$ (punkt (b)). Przedstaw uzyskane wyniki jak w zadaniu domowym 2.

Podobnie jak w zadaniu domowym 2, wykonaj dodatkowo wykresy $\frac{l_n^{(1)}}{f_1(n)}$ oraz $\frac{l_n^{(2)}}{f_2(n)}$, gdzie $l_n^{(d)}$ oznacza średnią wartość $L_n^{(d)}$ dla danego n oraz $d \in \{1, 2\}$, a funkcje f_1 i f_2 zadane są wzorami $f_1(n) = \frac{\ln n}{\ln \ln n}$ oraz $f_2(n) = \frac{\ln \ln n}{\ln 2}$ (por. rozdział 17 w [MU17]).

W celu uzyskania większej czytelności wyników możesz zwiększyć zakres n wykonując symulacje np. dla $n \in \{10\,000, 20\,000, \dots, 1\,000\,000\}$ lub zastosować na wykresach skalę logarytmiczną.

Zadanie 2. *Sortowanie przez wstawianie losowych danych*¹

Zaimplementuj oraz przetestuj działanie algorytmu sortowania przez wstawianie INSERTIONSORT (patrz algorytm 1; jego szczegółowe omówienie można znaleźć np. w rozdziale 2.1 w [CLRS09]). W tym celu dla każdego $n \in \{100, 200, \dots, 10\,000\}$ wykonaj po $k = 50$ niezależnych powtórzeń:

- (a) generowania tablicy $A[1..n]$ będącej losową permutacją zbioru liczb $\{1, \dots, n\}$ (wszystkie permutacje powinny być jednakowo prawdopodobne; podobnie jak poprzednio, zadbaj o dobry generator liczb pseudolosowych),
- (b) sortowania wygenerowanej tablicy A ,
- (c) zapisywania statystyk obejmujących rozmiar danych n , liczbę wykonanych porównań między kluczami (elementami tablicy A) oraz liczbę przestawień kluczy.

¹Podobne zadanie – dla większej liczby algorytmów sortujących – będzie do wykonania na laboratorium do kursu *Algorytmy i struktury danych* (4 semestr).

Po zakończeniu eksperymentów, korzystając z zebranych danych, przedstaw na wykresach:

- (a) liczbę wykonanych porównań w poszczególnych powtórzeniach oraz średnią liczbę porównań $cmp(n)$ jako funkcję n ,
- (b) liczbę przestawień w poszczególnych powtórzeniach oraz średnią liczbę przestawień kluczy $s(n)$ jako funkcję n ,
- (c) iloraz $\frac{cmp(n)}{n}$ oraz $\frac{cmp(n)}{n^2}$ jako funkcje n ,
- (d) iloraz $\frac{s(n)}{n}$ oraz $\frac{s(n)}{n^2}$ jako funkcje n .

Algorytm 1 Algorytm sortowania przez wstawianie (por. rozdział 2.1 w [CLRS09]).

```
1: procedure INSERTIONSORT( $A[1..n]$ )
2:   for  $j = 2$  to  $n$  do
3:      $key \leftarrow A[j]$ 
4:     // Wstaw  $A[j]$  w posortowany podciąg  $A[1..j - 1]$ .
5:      $i \leftarrow j - 1$ 
6:     while  $i > 0$  and  $A[i] > key$  do
7:        $A[i + 1] \leftarrow A[i]$ 
8:        $i \leftarrow i - 1$ 
9:     end while
10:     $A[i + 1] \leftarrow key$ 
11:  end for
12: end procedure
```

Zadanie 3. Przedstaw wyniki eksperymentów przeprowadzonych w zadaniach 1 oraz 2 w formie krótkiego sprawozdania.

- (a) Zaprezentuj wykresy, zwięźle omów uzyskane rezultaty i przedstaw wnioski.
- (b) Na podstawie wykresów krótko scharakteryzuj koncentrację wyników uzyskanych w poszczególnych powtórzeniach wokół wartości średniej dla badanych wielkości.
- (c) Na podstawie wykresów postaw hipotezy odnośnie asymptotyki wartości średnich badanych wielkości i uzasadnij ich wybór (w razie potrzeby w niektórych wykresach możesz zastosować skalę logarytmiczną).

Rozwiązanie obejmujące implementację (kody źródłowe) oraz plik pdf ze sprawozdaniem należy przesłać przez platformę MS Teams. Nie należy dołączać żadnych zbędnych plików.

Literatura

- [CLRS09] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*. The MIT Press, 3rd edition, 2009.
- [MU17] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis*. Cambridge University Press, USA, 2nd edition, 2017.