O rzucaniu monetą

Jacek Cichoń
jacek.cichon@pwr.edu.pl

Politechnika Wrocławska
Wydział Podstawowych Problemów Techniki
Katedra Podstaw Informatyki

Wrocław, 2 października, 2020

Układ dwójkowy

Przykłady

- $1 = 1 \cdot 2^0 = (1)_2$
- $2 = 1 \cdot 2^1 + 0 \cdot 2^0 = (10)_2$
- $3 = 1 \cdot 2^1 + 1 \cdot 2^0 = (11)_2$
- $\ldots$
- $20 = (10100)_2$
- $100 = (1100100)_2$

Układ dwójkowy

Przykłady

- $1 = 1 \cdot 2^0 = (1)_2$
- $2 = 1 \cdot 2^1 + 0 \cdot 2^0 = (10)_2$
- $3 = 1 \cdot 2^1 + 1 \cdot 2^0 = (11)_2$
- $\ldots$
- $20 = (10100)_2$
- $100 = (1100100)_2$

Pytanie

Ile bitów potrzebujemy do zapisania liczby n ?

Liczba niezbędnych bitów

Prosta obserwacja

$$(b_{k-1}b_k \ldots b_1 b_0)_2 = b_0 \cdot 2^0 + b_1 \cdot b_1 + \ldots + b_{k-1} \cdot 2^{k-1} \leqslant$$
$$1 + 2 + 2^2 + \ldots + 2^{k-1} = 2^k - 1 \ .$$

Zatem: za pomocą k bitów możemy przedstawić liczby $< 2^k$.
A z tego już łatwo wyprowadzamy następujący wniosek:

Niezbędna liczba bitów do dowolnej liczby z zakresu $\{1, \ldots, n\}$

$$b = \lfloor \log_2(n) \rfloor + 1 \approx \log_2(n)$$

Aplikacje Covidove

Bluetooth

Częstotliwość: 10 razy na sekundę

Ile zbieramy danych ?

- 1 doba $= 24 \cdot 60 \cdot 60 = 86400 \approx 10^5$ sekund
- 20 dni $\approx 2 \cdot 10^6$ sekund
- Ile danych: $2 \cdot 10^7$
- Liczba potrzebnych bitów: 21 bitów

Aplikacje Covidove

Bluetooth

Częstotliwość: 10 razy na sekundę

Ile zbieramy danych ?

- 1 doba $= 24 \cdot 60 \cdot 60 = 86400 \approx 10^5$ sekund
- 20 dni $\approx 2 \cdot 10^6$ sekund
- Ile danych: $2 \cdot 10^7$
- Liczba potrzebnych bitów: 21 bitów
- Próg bezpieczeństwa: 10 minut
- czyli: $10 \cdot 60 \cdot 10 = 6 \cdot 10^3$

Aplikacje Covidove

Bluetooth

Częstotliwość: 10 razy na sekundę

Ile zbieramy danych ?

- 1 doba $= 24 \cdot 60 \cdot 60 = 86400 \approx 10^5$ sekund
- 20 dni $\approx 2 \cdot 10^6$ sekund
- Ile danych: $2 \cdot 10^7$
- Liczba potrzebnych bitów: 21 bitów
- Próg bezpieczeństwa: 10 minut
- czyli: $10 \cdot 60 \cdot 10 = 6 \cdot 10^3$
- CZY POTRZEBNA JEST NAM ABSOLUTNA DOKŁADNOŚĆ ?

Aplikacje Covidove

Bluetooth

Częstotliwość: 10 razy na sekundę

Ile zbieramy danych ?

- 1 doba $= 24 \cdot 60 \cdot 60 = 86400 \approx 10^5$ sekund
- 20 dni $\approx 2 \cdot 10^6$ sekund
- Ile danych: $2 \cdot 10^7$
- Liczba potrzebnych bitów: 21 bitów
- Próg bezpieczeństwa: 10 minut
- czyli: $10 \cdot 60 \cdot 10 = 6 \cdot 10^3$
- CZY POTRZEBNA JEST NAM ABSOLUTNA DOKŁADNOŚĆ ?
- OCZYWIŚCIE, ŻE **NIE**

A może wykorzystać monetę?

Co wymagamy od ciągu kolejnych n rzutów ?

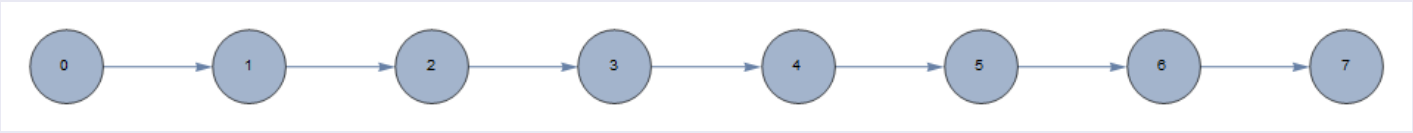
Rozważamy rodzinę $\mathcal{X}$ wszystkich możliwych ciągów długości n
$X = (X_1, X_2, X_3, \ldots, X_n), \quad X_i \in \{0, R\}$ rzutów .

- jeżeli $1 \leqslant i \leqslant n$, to $\frac{|\{X \in \mathcal{X} : X_i = O\}|}{2^n} = \frac{1}{2}$
- jeżeli $1 \leqslant i < j \leqslant n$, to $\frac{|\{X \in \mathcal{X} : (X_i, X_j) = (O, R)\}|}{2^n} = \frac{1}{4}$
- jeżeli $1 \leqslant i < j < k \leqslant n$, to $\frac{|\{X \in \mathcal{X} : (X_i, X_j, X_k) = (O, R, O)\}|}{2^n} = \frac{1}{8}$
- $\ldots$

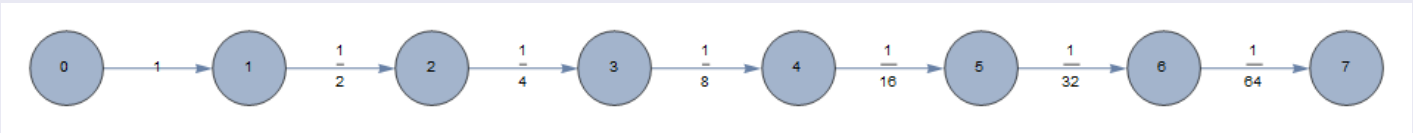
[fachowe terminy: jednostajność, niezależność, uczciwa moneta]

Liczniki

Klasyczny licznik



Licznik probabilistyczny Morrisa

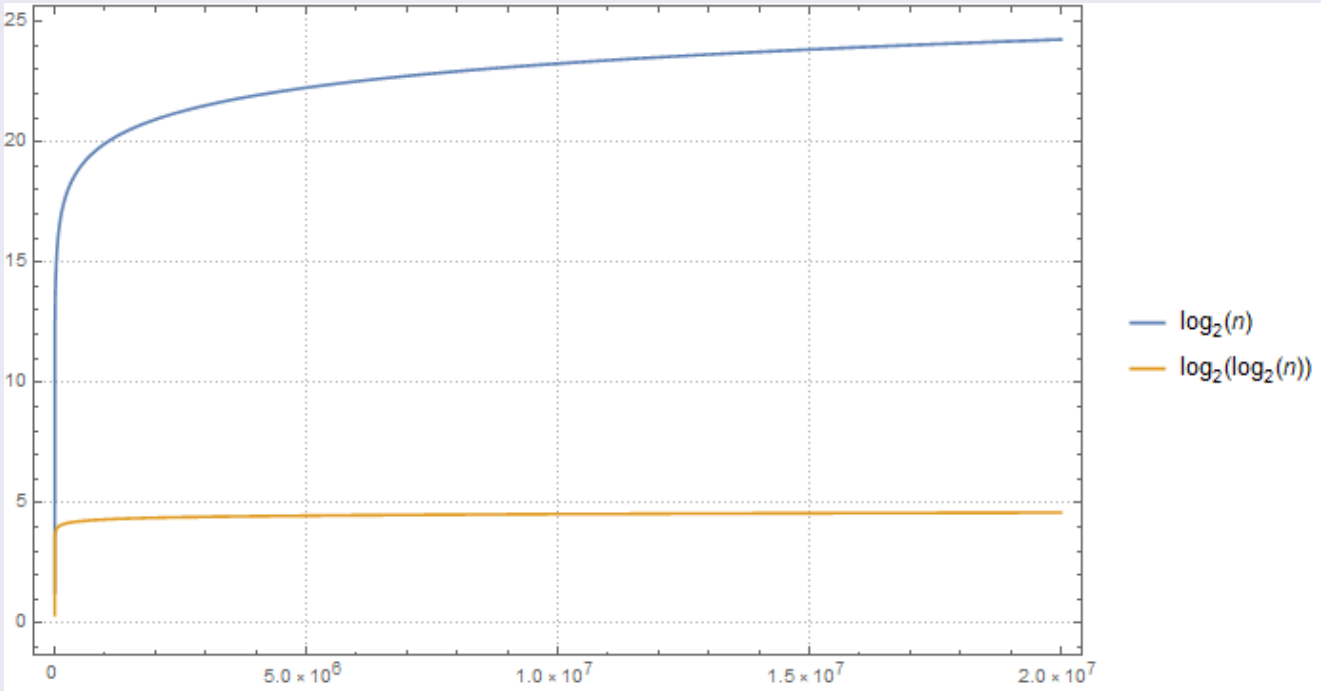


- Aby dojść do pozycji 4 potrzebujemy (średnio) $1 + 2 + 2^2 + 2^3 = 2^4 - 1$ kroków.
- Tw. Po n krokach wartość licznika wynosi $\approx \log_2(n)$
- Tw. Jeśli wartość licznika wynosi C, to wykonano $\approx 2^C - 1$ kroków.

Liczba potrzebnych bitów

- Wiemy, że po n rzutach wartość licznika C jest równa $\approx \log_2(n)$
- Do tego potrzebujemy $\lfloor \log_2(\log_2(n)) \rfloor + 1 \approx \log_2(\log_2(n))$ bitów !!!

Porównanie dwóch funkcji



$[\log_2(\log_2(6.022 \cdot 10^{23})) = 6.30368\ldots;\ \log_2(\log_2(10^{80})) = 8.05395\ldots]$

Jeszcze jeden trik

1. oznaczamy orła (O) przez 0; oznaczamy reszkę (R) przez 1
2. ciągowi rzutów $(b_1, b_2, b_3, b_4, \ldots)$ przyporządkowujemy liczbę

$$(0.b_1\,b_2\,b_3\,b_4\,b_5\,b_6\,b_7\,b_8\,b_9\,b_{10}\,b_{11}\,b_{12}\,b_{13}\ldots)_2$$

3. otrzymaną liczbę oznaczamy przez $rand()$

Twierdzenie

Jeśli $0 \leqslant a \leqslant b \leqslant 1$, to

$$\Pr(rand() \in [a, b]) = b - a$$

1. $\Pr(rand() < \frac{1}{2}) = \frac{1}{2}$
2. $\Pr(rand() < \frac{1}{4}) = \frac{1}{4}$
3. $\ldots$

Licznik klasyczny/Licznik probabilistyczny

Licznik klasyczny

```
INIT: L = 0

onIncrease()
{
  L = L+1
}

onGet()
{
  return L
}
```

Licznik probabilistyczny

```
INIT: L = 0

onIncrease()
{
  if random() < 1/(2^L) then
    L = L+1
}

onGet()
{
  return 2^L - 1
}
```

Algorytmy zrandomizowane

- Algorytmy wykorzystujące mechanizmy losowe są znane od dawna (np. klasa algorytmów Monte Carlo)
- Nowe, piękne, bardzo pomysłowe algorytmy (podobne do licznika probabilistycznego) powstają w ostatnich latach.
- Jest tu wiele do zrobienia.

Algorytmy zrandomizowane

- Algorytmy wykorzystujące mechanizmy losowe są znane od dawna (np. klasa algorytmów Monte Carlo)
- Nowe, piękne, bardzo pomysłowe algorytmy (podobne do licznika probabilistycznego) powstają w ostatnich latach.
- Jest tu wiele do zrobienia.
- Być może to WY będziecie ich autorami.
- Jest wiele wyzwań !!!!

WYZWANIA

- Oczywiste wyzwania dla informatyków: algorytmy służące do analizy bardzo dużych zestawów danych (Big Data, Internet of Things)
- Dla fizyków: jak zbudować tanie generatory liczb losowych (może: kropki kwantowe?)
- Dla bioinżynierów: ile i jak dokładne dane przechowywać w indywidualnych urządzeniach monitorujących stan zdrowia
- Dla bioinżynierów: jak powinny działać aplikacje monitorujące przyszłe zjawiska epidemiczne?
- ...

WIELKIE WYZWANIE

Informacja

- podejście informatyczne (przepustowość kanałów informacyjnych; teoria kodowania)
- podejście termodynamiczne (entropia, temperatura)

CZY JEST INFORMACJA ???

To już koniec