

# Algorytmy Tekstowe - laboratorium 1

Jakub Pinowski

## 1 Wyszukiwanie wzorca "art" w tekście ustawy o zryczałtowanym podatku dochodowym od niektórych przychodów osiąganych przez osoby fizyczne

### 1.1 Wyniki wyszukiwania wzorca

- Naive algorithm: !-273 found-! [1156, 1505, 4692, 4734, 4879, 5082, 5148, 5949, ...
- FA: !-273 found-! [1156, 1505, 4692, 4734, 4879, 5082, 5148, 5949, ...
- KMP: !-273 found-! [1156, 1505, 4692, 4734, 4879, 5082, 5148, 5949, ...

### 1.2 Porównanie czasów działania<sup>1</sup>

- Naive algorithm: 0.034662485122680664s
- FA: 0.030025243759155273s
- KMP: 0.03327298164367676s

## 2 "Kruszwil" we fragmencie Wikipedii

Znaleziono 13 wystąpień słowa "kruszwil" w podanym fragmencie Wikipedii.

### 2.1 Porównanie czasów działania

- Naive algorithm: 36.314714193344116s
- FA: 38.10478115081787s
- KMP: 43.138630390167236s

---

<sup>1</sup>Wszystkie czasy działania automatu skończonego i algorytmu KMP uwzględniają preprocessing związany z macierzą i funkcją przejścia

### 3 "Nieprzyjemne" dane dla algorytmu naiwnego

Z powodu złożoności  $O[m(n-m)]$  ( $n$  - długość tekstu;  $m$  - długość wzorca) najgorsze dane dla algorytmu naiwnego to takie, gdzie długość wzorca jest mniej więcej równa połowie długości tekstu. Stąd możemy stworzyć przykład:

- `text = "a" * 1000000`
- `pattern = "a" * 500000`

#### 3.1 Porównanie czasów działania

- Naive algorithm: 17.531678199768066s
- FA: 1.0810279846191406s
- KMP: 0.34475040435791016s

### 4 Porównanie czasu preprocessingu automatu skończonego i algorytmu KMP

- `pattern = "abcdefghijk" * 100000`

#### 4.1 Czasy preprocessingu

- FA preprocessing time: 2.7725579738616943
- KMP preprocessing time: 0.2004375457763672