

# Algorytmy tekstowe - ćwiczenia 1

Tomasz Lamża

## 1 Porównaj szybkość działania algorytmów dla problemu z p. 3.

Algorytm naiwny	0.06797 s
Algorytm automatu skończonego	0.04575 s
Algorytm KMP	0.05759 s

Tabela 1: Czasy wykonywania się poszczególnych algorytmów

Widzimy, że algorytm naiwny działa najwolniej, a najszybciej algorytm automatu skończonego.

## 2 Porównaj szybkość działania algorytmów poprzez wyszukanie słowa "kruszwil" we fragmencie polskiej Wikipedii

Algorytm naiwny	93.278 s
Tworzenie tablicy przejścia	0.0009636 s
Algorytm automatu skończonczonego	63.039 s
Tworzenie tablicy prefixów	0.00002503 s
Algorytm KMP	76.173 s

Tabela 2: Czasy wykonywania się poszczególnych algorytmów

Widzimy, że czas tworzenia table jest bez znaczenia. Jest to spowodowane tym, że wzór jest krótki.

### 3 Zaproponuj tekst oraz wzorzec, dla którego zmierzony czas działania algorytmów 2 oraz 3 będzie co najmniej 2 krotnie krótszy niż dla algorytmu naiwnego.

Patrząc na złożoności tych trzech algorytmów, widzimy czas dopasowania jedynie w algorytmie naiwnym zależy od długości wzorca, zatem aby czas 2 pozostałych algorytmów był znacząco krótszy należy wziąć długi wzorzec. Np. wiele literek 'a' oraz tekst z wikipedii. Otrzymane wyniki czasowe:

Algorytm naiwny	540 s
Algorytm automatu skończonego	82 s
Algorytm KMP	88 s

Tabela 3: Czasy wykonywania się poszczególnych algorytmów

Widzimy, że czas wykonania algorytmu naiwnego jest prawie 7 razy dłuższy. Pozostałe algorytmy wykonują się w podobnym czasie.

### 4 Zaproponuj wzorzec, dla którego zmierzony czas obliczenia tablicy przejścia automatu skończonego będzie co najmniej dwukrotnie dłuższy, niż czas potrzebny na utworzenie funkcji przejścia w algorytmie KMP.

Aby czas preprocessingu dla automatu skończonego był ponad 2 razy dłuższy, wystarczy wziąć pattern postaci 100x'a' oraz tekst z wikipedii.

Otrzymane czasy wyglądają następująco :

Tworzenie tablicy przejścia	0.00289 s
Tworzenie tablicy prefixów	0.00017 s

Tabela 4: Czasy wykonywania preprocesingu

Widzimy, że czas tworzenia tablicy przejścia zajmuje ponad 20 razy więcej czasu, niż stworzenie tablicy prefixów.