

Algorytmy Tekstowe

Laboratorium 1 - Znajdowanie wzorca w tekście

10.03.2021

Magdalena Pastuła

Zadanie 2.

Znalezione wystąpienia wzorca "art" w ustawie w kolejności rosnącej:

```
[871, 1226, 4384, 4426, 4573, 4779, 4846, 5545, 5636, 6883, 7131, 7405, 7673, 7931, 8749, 9618, 9681, 9886, 10721, 10808, 11227, 12828, 14911, 14986, 15693, 15864, 15974, 16117, 16187, 16415, 16431, 23284, 23715, 23807, 24249, 24347, 24446, 24601, 25209, 25371, 26664, 26955, 27110, 27174, 27225, 27494, 27979, 28167, 28378, 30570, 30628, 30704, 30937, 31353, 32126, 32491, 32578, 32795, 33089, 34161, 34248, 35033, 35649, 36652, 37058, 37904, 38012, 38480, 38637, 38866, 39000, 39419, 40570, 41257, 41459, 41631, 41807, 41942, 42159, 42302, 42348, 42862, 42972, 43208, 44023, 44087, 44391, 44449, 44736, 44846, 46794, 46899, 48243, 48279, 48366, 48514, 48724, 48782, 48957, 49029, 49390, 49455, 49580, 49639, 50151, 50504, 50635, 51433, 51539, 51743, 52027, 52490, 52514, 52696, 53282, 53427, 53576, 53636, 54281, 54552, 54759, 54947, 55294, 55481, 56292, 56377, 56635, 57026, 57281, 57377, 57435, 57730, 57830, 58333, 58426, 58823, 58953, 59386, 59737, 59994, 60205, 60680, 61157, 61821, 61970, 62024, 62840, 63037, 63130, 63211, 64435, 64475, 64538, 65399, 65710, 66334, 66409, 66900, 67019, 67044, 67179, 67254, 67520, 68157, 68314, 68442, 68520, 68881, 68976, 69074, 69393, 69804, 69920, 69955, 70717, 70863, 71232, 71442, 72255, 73371, 74403, 74529, 74545, 74552, 74741, 74828, 74897, 75279, 76035, 76991, 77510, 77791, 77856, 77890, 78013, 78195, 78268, 78338, 78377, 78513, 78719, 78782, 79762, 82185, 83509, 83763, 84746, 85604, 85630, 85908, 85990, 86179, 86219, 86464, 86533, 86777, 86865, 86905, 86970, 87233, 87434, 87714, 87823, 88019, 88029, 88045, 88065, 88414, 88466, 88825, 89945, 90060, 91198, 91648, 91709, 91846, 92904, 93112, 94290, 94431, 95607, 97115, 98186, 101355, 102547, 103138, 104184, 104393, 108373, 112923, 113079, 114040, 139000, 152287, 153170, 155095, 156228, 157249, 161956, 162135, 171486, 178234, 193252, 193352, 195422, 199556, 201952, 204632, 205110, 210321, 210775, 216075, 216161]
```

Zadanie 3.

Czasy działania poszczególnych algorytmów dla wyszukiwania z p.2.

```
Algorytm naiwny: 0.039335 s  
Algorytm z wykorzystaniem automatu: 0.040905 s  
Algorytm KMP: 0.037907 s
```

Zadanie 4

Tekst oraz wzorzec, dla których czas działania algorytmu opartego o automat i KMP jest przynajmniej pięć razy krótszy niż naiwny. Nie udało mi się znaleźć ani wymyślić tekstu i wzorca, dla których zachodzi ta zależność. Próbowałam dla tekstu ustawy oraz wzorców: "cytryny", ponieważ nie występuje ono ani razu w tekście, wzorca "oraz", ponieważ występuje dosyć często oraz wzorca "ryczałt", które występuje dwa razy rzadziej niż "oraz". Otrzymane czasy poszczególnych algorytmów dla tych wzorców wyniosły:

Wzorzec "cytryny":

```
Czas algorytmu naiwnego: 0.041031 s
Czas algorytmu z wykorzystaniem automatu: 0.042091 s
Czas algorytmu KMP: 0.037119 s
```

Wzorzec "oraz":

```
Czas algorytmu naiwnego: 0.039446 s
Czas algorytmu z wykorzystaniem automatu: 0.039857 s
Czas algorytmu KMP: 0.034906 s
```

Wzorzec "ryczałt":

```
Czas algorytmu naiwnego: 0.045422 s
Czas algorytmu z wykorzystaniem automatu: 0.039858 s
Czas algorytmu KMP: 0.035910 s
```

Zadanie 5

Przykład wzorca, dla którego utworzenie funkcji przejść automatu jest przynajmniej pięć razy dłuższe niż utworzenie tabeli dopasowań do algorytmu KMP:

```
Wzorzec: "ABCD EFGH IJKL MNOP RSTV UWXY Z ABCD EFGH IJKL MNOP RSTV UWXYZ"
Czas tworzenia funkcji przejścia: 0.076611 s
Czas tworzenia tabeli dopasowań: 0.000297 s
```