

Algorytmy i struktury danych z językiem Python

Algorytm Aho-Corasick

Angela Czubak

1 Teoria

1.1 Wprowadzenie

Algorytm Aho-Corasick został opracowany przez Alfreda V. Aho oraz Margaret J. Corasick. Jego celem jest znalezienie wzorców pochodzących z pewnego słownika w tekście.

Cechą charakterystyczną tego algorytmu jest to, że szukanie wystąpień zadanych słów następuje "na raz", dzięki czemu złożoność obliczeniowa tego algorytmu jest liniowa od długości tekstu, w którym wyszukujemy, liczby wystąpień wzorców oraz sumy długości tychże.

Algorytm ten jest stosowany np. w komendzie UNIX-a - **fgrep**.

Idea algorytmu opiera się na drzewach trie i automatach skończonych, których tworzenie omówię dalej.

2 Opis interfejsu

3 Kod źródłowy

4 Testy

5 Podsumowanie

6 Bibliografia

- http://pl.wikipedia.org/wiki/Algorytm_Aho-Corasick
- <http://www.cs.uku.fi/~kilpelai/BSA05/lectures/slides04.pdf>
- <https://wiki.python.org/moin/TimeComplexity>