

Zadanie 1 z listy 5 - “Kompresja Danych”

Łukasz Klasieński

26 kwietnia 2020

Zadanie 1

Przyjmijmy, że w algorytmie LZ77 rozmiar bufora słownikowego jest równy s , rozmiar bufora kodowania to t , a rozmiar alfabetu jest równy a . Dla danej liczby $k > 0$ podaj warunki, przy których możliwe jest skonstruowanie dowolnie długiego ciągu danych, w którym wszystkie dopasowania w LZ77 są krótsze niż k

Rozwiązanie

Warunki: $|t| \leq (k - 1) * |a| * 2$

Wtedy konstrukcja wygląda w taki sposób, że dodajemy do naszego ciągu danych po kolei wszystkie słowa z alfabetu $(k - 1)$ razy, następnie powtarzamy tę operację od końca.

Przykład:

· alfabet : $[a, b, c]$

· $t = 3$

· $k = 2$

· ciąg danych: $abc - cba - abc - cba \dots$

Przykład:

· alfabet : $[a, b, c]$

· $t = 6$

· $k = 3$

· ciąg danych: $aabbcc - ccbbaa - aabbcc \dots$

Dlaczego działa? Dzięki takiej konstrukcji, w oknie nie znajdziemy porównania $\geq k$ ponieważ kiedy będziemy chcieli znaleźć maksymalny prefix danego ciągu, to dany znak w buforze po sobie będzie mieć albo własne kopie, albo znak który jest następny/poprzedni w alfabecie (przeciwnie od szukanego). Zatem najlepsze dopasowanie będzie tylko dla $(k - 1)$ **sklonowanych** znaków.

Widać zatem, że przy takich założeniach możemy budować dowolnie długie ciągi spełniające k .