

Zadanie 4 z listy 7 - “Kompresja Danych”

Łukasz Klasinski

16 maja 2020

Zadanie 4

Tekst w został zakodowany metodą Burrowsa-Wheelera i uzyskaliśmy w wyniku parę (x, i) , gdzie x to ostatnia kolumna posortowanej tablicy przesunięć cyklicznych, a i to pozycja słowa w w posortowanej tablicy przesunięć cyklicznych. Jak wyglądać będzie zakodowana postać słowa w^m dla $m > 1$?

Jeśli będziemy chcieli zakodować takie słowo, to można zauważyć, że w przesunięciach cyklicznych nowego słowa w^m , powtórzą się prefiksy z tablicy przesunięć dla w - np jeśli wcześniej mieliśmy w niej:

$$w = x_1x_2x_3$$

To po zdublowaniu w (w^2) , w tablicy dla takiego słowa będzie odpowiadający wpis:

$$w^2 = x_1x_2x_3x_1x_2x_3$$

oraz jako, że to słowo jest symetryczne to będzie ono w tablicy dwukrotnie. Widzimy zatem, że dla dowolnego słowa z tablicy przesunięć w , w tablicy przesunięć w^m zostanie ono przedłużone m razy oraz powtórzone m razy. Jako że końcówki nowych cykli będą takie same jak końcówki cykli w , to otrzymamy w wyniku parę $(m * x, i * m)$, gdzie $m * x$ oznacza każdy element z x do potęgi m .