

Maksymalne powtórzenia w słowach Sturmia

Marcin Piątkowski

Standardowe słowa Sturmia to klasa silnie kompresowalnych słów nad alfabetem $\Sigma = \{a, b\}$. Definiowane są one rekurencyjnie. Dla ciągu liczb całkowitych $d = (d_0, d_1, \dots, d_{n-1})$, spełniającego warunki: $d_0 \geq 0$ oraz $d_i > 0$ dla $0 < i < n$, określamy n -te słowo standardowe Sturmia za pomocą rekurencji:

$$\begin{aligned}x_{-1} &= b \\x_0 &= a \\x_i &= x_{i-1}^{d_{i-1}} x_{i-2} \quad \text{dla } 0 < i < n\end{aligned}$$

Szczególnym przypadkiem standardowych słów Sturmia są słowa Fibonacciego otrzymywane dla ciągów postaci $d = (1, 1, \dots, 1)$.

Maksymalnym powtórzeniem w słowie s nazywamy podśłowo w słowa s postaci $w = u^k v$, gdzie $k \geq 2$, u jest minimalne o tej własności ($u \neq y^l$), zaś v jest właściwym prefiksem u , które nie może być rozszerzone w żadną stronę bez zmiany okresu (u).

Maksymalne powtórzenia pozwalają w zwarty sposób reprezentować wszystkie powtórzenia w słowie. Są ważnym pojęciem spotykanym w kombinatoryce słów oraz wielu zastosowaniach praktycznych: kompresji danych, biologii obliczeniowej, dopasowaniu wzorca (pattern matching), itp.

Symbolem $\rho(n)$ oznaczamy maksymalną liczbę maksymalnych powtórzeń we wszystkich słowach długości n . W ogólnym przypadku znane są szacowania $\rho(n)$ zarówno z góry jak i z dołu, jednak wyznaczenie dokładnej wartości $\rho(n)$ nadal jest problemem otwartym.

W czasie prezentacji pokazana zostanie konstrukcja zwanego wzoru na liczbę $\rho(n)$ dla klasy standardowych słów Sturmia oraz jej asymptotyczne oszacowanie.