

Algorytmy Tekstowe

Lista zadań nr 5

$MaxSuf(w)$ dla słowa w , to maksymalny leksykograficznie sufix słowa w . Słowo w nazwiemy *sobiemaksymalnym*, jeżeli $w = MaxSuf(w)$.

Algorytm 1

function NP(m)

```
{  int period = 1;
    for (int i = 2; i <= m; i++)
        if (pat[i] != pat[i - period]) period = i;
    return period;
}
```

1. Udowodnij, że Algorytm 1 oblicza okres słowa $pat[1..m]$, o ile jest to słowo sobiemaksymalne.
2. Pokaż przykład wykazujący, że dla słów, które nie są sobiemaksymalne algorytm 1 nie koniecznie znajduje okres słowa.
3. Opracuj algorytm, który dla słowa sobiemaksymalnego $pat[1..m]$ znajdzie wszystkie wystąpienia pat w tekście $tekst[1..n]$ w czasie $O(n+m)$ wykorzystując dodatkową pamięć (oprócz dostępnych tylko do odczytu pat i $tekst$) $O(1)$. (*Wskazówka:* Zauważ, że algorytm 1 obliczający na bieżąco okres słowa sobiemaksymalnego może z powodzeniem zastąpić funkcję π czy $Bord$ w algorytmie KMP.)
4. Niech dla słowa $x = uv$, $v = MaxSuf(x)$. Udowodnij, że jeżeli v występuje w tekście T na pozycji i , to w T nie ma żadnego wystąpienia x na pozycjach $i = |u| + 1, i + |u| + 2, \dots, i$.
5. Zakładając, że potrafisz znaleźć dla słowa $pat[1..m]$ jego maksymalny sufix w czasie $O(m)$ i pamięci roboczej $O(1)$, opracuj algorytm znajdujący wszystkie wystąpienia pat w tekście $tekst[1..n]$ w czasie $O(n+m)$ i pamięci roboczej $O(1)$. Oszacuj w miarę dokładnie stałą w przypadku czasu.
6. Napisz procedurę znajdującą najdłuższy, sobiemaksymalny prefiks słowa w oraz jego okres (prefiksu) w czasie $O(|w|)$ i pamięci roboczej $O(1)$. (*Wskazówka:* wystarczy bardzo niewielka przeróbka w Algorytmie 1, uwzględniająca mniejszość/większość niezgodnych znaków.)
7. Wykorzystaj rezultat poprzedniego zadania, by napisać procedurę znajdującą maksymalny sufix słowa w w czasie $O(|w|)$ i pamięci roboczej $O(1)$.
8. Zauważ, że wykonanie wszystkich powyższych zadań pozwala skonstruować algorytm, który w czasie $O(n+m)$ i pamięci roboczej $O(1)$ znajduje wszystkie wystąpienia wzorca w tekście. (Za to zadanie nie ma punktów, należy je zrobić dla własnej satysfkcji.)