

Informatyczny kącik olimpijski (114): *Magic*

Tym razem omówimy zadanie z pierwszego dnia Pierwszej Olimpiady Informatycznej Juniorów (EJOI), która odbyła się w Sofii we wrześniu 2017 roku.

Zadanie: *Dane jest słowo $s = (s_1, s_2, \dots, s_n)$ o długości n , w którym występuje k różnych liter alfabetu angielskiego. Magicznym podslowem nazywamy niepuste podslowo (czyli spójny fragment słowa), które zawiera taką samą liczbę wystąpień każdej z k liter. Należy policzyć liczbę magicznych podslów w słowie s . Podslowa, które są takie same, ale znajdują się na różnych pozycjach, uznajemy za różne.*

Na potrzeby tego artykułu przyjmijmy dodatkowe założenie, że słowo s zawiera wyłącznie litery ze zbioru k -początkowych liter alfabetu angielskiego. Aby uzyskać to założenie, wystarczy po wczytaniu słowa nadać literom nowe identyfikatory będące kolejnymi małymi literami alfabetu angielskiego.

Rozwiązanie $O(n^3)$

Najprostsze rozwiązanie polega na rozpatrzeniu każdego podslowa niezależnie. Aby sprawdzić, czy dane podslowo jest magiczne, należy zliczyć liczbę wystąpień każdej litery w podslowie, a następnie sprawdzić, czy wszystkie otrzymane wartości są równe.

Liczba wszystkich podslów jest rzędu $O(n^2)$. Sprawdzenie, czy podslowo jest magiczne, wymaga $O(n)$ operacji. Zatem całe rozwiązanie wykonuje $O(n^3)$ operacji.

Rozwiązanie $O(n^2)$

Powyższe rozwiązanie można w łatwy sposób przyspieszyć. Każdą z możliwych początkowych pozycji $\{1, 2, \dots, n\}$ podslowa słowa s rozpatrzmy niezależnie. Załóżmy, że rozpatrujemy początkową pozycję x . Przeiterujemy wszystkie potencjalne końce podslów zaczynających się na tej pozycji. Niech s_x, s_{x+1}, \dots, s_y będzie aktualnie rozpatrywanym podslowem. Zauważmy, że aby policzyć liczbę wystąpień poszczególnych liter w tym podslowie, wystarczy skorzystać z wyniku dla podslowa $s_x, s_{x+1}, \dots, s_{y-1}$ oraz zaktualizować go o wystąpienie litery s_y . Zatem, dla ustalonego podslowa potrafimy wyznaczyć te wartości w stałej liczbie operacji $O(1)$. Natomiast sprawdzenie, czy dane podslowo jest magiczne, wymaga $O(k)$ operacji (należy sprawdzić, czy liczba wystąpień każdej z k liter jest taka sama). Zauważmy, że słowa, których długość nie jest podzielna przez k , na pewno nie są magiczne. Zatem dla tych słów możemy pominąć fazę sprawdzania licznosci poszczególnych liter. Słów, dla których wykonamy fazę sprawdzania licznosci liter, jest $O(\frac{n^2}{k})$. Sumarycznie wykonamy $O(n^2)$ operacji.

Przypadek dla $k = 1$

Przypadek dla $k = 1$ jest trywialny. Każde podslowo jest magiczne, zatem wynikiem jest $\frac{n(n+1)}{2}$.

Przypadek dla $k = 2$

Najpierw słowo s zamienimy na ciąg liczb a . Pierwszą literę słowa oraz wszystkie jej wystąpienia zamienimy na 1. Pozostałe litery zamienimy na -1 . Zauważmy teraz, że każde magiczne podslowo ma sumę 0.

Policzmy sumy prefiksowe p , gdzie $p(i) = \sum_{j=1}^i a_j$ dla $0 \leq i \leq n$. Zauważmy, że podslowo s_x, s_{x+1}, \dots, s_y jest magiczne, jeśli $a_x + a_{x+1} + \dots + a_y = p(y) - p(x-1) = 0$, czyli jeśli $p(y) = p(x-1)$. Wystarczy zatem policzyć liczbę par indeksów, dla których sumy prefiksowe są równe. W tym celu zliczmy wystąpienia elementów w ciągu p . Niech $w(x)$ oznacza liczbę wystąpień x w ciągu p . Wówczas wynikiem jest $\sum_{x \in [-n, n]} \frac{w(x)(w(x)-1)}{2}$. Rozwiązanie wykonuje $O(n)$ operacji.

Rozwiązanie $O(nk \cdot \log(n))$

Zacznijmy od przypisania każdej literze jej unikalnego identyfikatora ze zbioru $\{1, 2, \dots, k\}$. Następnie przekształcimy słowo s na ciąg liczb a , zamieniając każdą literę na odpowiadający jej identyfikator. Np. słowo **kajak** możemy zamienić na ciąg $a = 2, 1, 3, 1, 2$ (a otrzymało identyfikator 1, k otrzymało identyfikator 2, zaś j otrzymało identyfikator 3). Od teraz będziemy operowali na ciągu liczb naturalnych.

Niech:

- $w(x, i)$ będzie równe 1, jeśli na x -tej pozycji występuje wartość i oraz 0 w przeciwnym przypadku;
- $p(x, i) = \sum_{j=1}^x w(j, i)$ (liczba wystąpień wartości i na x -elementowym prefiksie ciągu a);
- $v(x) = [p(x, 1), p(x, 2), \dots, p(x, k)]$; (wektor reprezentuje liczbę wystąpień poszczególnych liter na x -elementowym prefiksie ciągu a);
- $v'(x) = v(x) - [p(x, 1), p(x, 1), \dots, p(x, 1)]$.

Podslowo a_x, a_{x+1}, \dots, a_y jest magiczne, jeśli:

$$v(y) - v(x-1) = [c, c, \dots, c]$$

dla pewnego $c \in \mathbb{Z}$. Po przekształceniach, które pozostawiam Czytelnikowi jako ćwiczenie, otrzymujemy równoważną definicję magicznego podslowa:

$$v'(y) - v'(x-1) = [0, 0, \dots, 0]$$

Zatem każda taka para $x, y \in \mathbb{N}$ ($1 \leq x \leq y \leq n$), że $v'(y) - v'(x-1) = [0, 0, \dots, 0]$ oznacza magiczne podslowo. Policzmy zatem, ile jest par takich samych wektorów w ciągu $v'(0), v'(1), \dots, v'(n)$. W tym celu posortujemy ten ciąg leksykograficznie. Wówczas wektory, które są równe, będą tworzyły spójny blok. Spośród bloku m takich samych wektorów można wybrać $\frac{m(m-1)}{2}$ par.

Obliczenie wektorów $v'(x)$ dla każdego $x \in \mathbb{N}$ ($1 \leq x \leq n$) wymaga $O(nk)$ operacji, zaś sortowanie $O(nk \cdot \log(n))$ operacji. Zatem całe rozwiązanie wykonuje $O(nk \cdot \log(n))$ operacji.

Bartosz ŁUKASIEWICZ