

	<p style="text-align: center;">ASD 2016/2017</p> <p style="text-align: center;">Program D</p> <p style="text-align: center;">Elektroniczny słownik</p>	<p>Punkty [0,4] przyznaje Prowadzący Ćwiczenia</p>
---	---	--

Opis

Przez ponad 20 lat gromadzono informacje jakie hasła są najczęściej poszukiwane w Internecie. Wybrano zbiór tych haseł (nazwano je słowami kluczowymi) i postanowiono zbudować statyczny słownik wyszukujący hasła zapisane w bazie danych w postaci drzewa BST. Drzewo ma być takie, by średnia liczba porównań przy wyszukiwaniu haseł (również błędnych lub bardzo rzadko używanych słów nie istniejących w słowniku) była minimalna. Celem zadania jest utworzenie optymalnego drzewa statycznego zbioru słów kluczowych, wypisanie słów tego drzewa w sposób umożliwiający odczytanie jego struktury oraz obliczenie średniego kosztu operacji wyszukiwania w tym drzewie.

Dane

Pierwsza linia danych zawiera liczbę zestawów danych ($0 < z < 100$). Każda linia zestawu zawiera nazwę oraz dwie liczby całkowite (pozostałe znaki linii są pomijane). Z linii pierwszej odczytane są dwie liczby całkowite:

n = liczba słów ($1 < n < 1000$) oraz y_0 = liczba wyszukiwań słów t : $t < s_1$.

Kolejne n linii zawierają słowo s_i i dwie liczby całkowite x_i, y_i dla $i=1, \dots, n$.

Słowo s_i zawiera tylko litery ze zbioru $\{A, B, \dots, Z, a, b, \dots, z\}$

$1 \leq x_i \leq 999999$ to liczba wyszukiwań słowa s_i

$1 \leq y_i \leq 999999$ to liczba wyszukiwań słów t : $s_i < t < s_{i+1}$ (dla y_0 t : $t < s_1$; dla y_n t : $s_n < t$)

Wyniki

Niech $W = \sum_{k=1}^n x_k + \sum_{k=0}^n y_k$, wówczas $p_i = x_i/W$, $q_i = y_i/W$ oznaczają prawdopodob-

ieństwa odpowiednich słów. Dla podanego zbioru słów należy znaleźć optymalne drzewo BST realizujące minimalną średnią ważoną liczbę porównań w operacji SEARCH

wzór: $c = \sum_{i=1}^n (p_i * (DP(s_i) + 1)) + \sum_{i=0}^n (q_i * (DP(NIL_i)))$ gdzie DP oznacza głębokość węzła.

Słowa drzewa wypisywane są leksykograficznie, każde poprzedzone liczbą k kropek, ($k = 1 + 2 * DP(\text{węzła})$, korzeń ma głębokość 0). W ostatniej linii wypisana jest wartość c z dokładnością do 4-ch miejsc po kropce.

Dodatkowe wyjaśnienia oraz przykładowe dane i wyniki i znajdują się w pliku jawne.zip na platformie PEGAZ.