Zadanie: XRS
XOR-Ścieżki

Warsztaty ILO 2017-2018, grupa olimpijska, dzień 16. Dostępna pamięć: 128 MB.

Rozwiązanie wzorcowe $O(n \cdot log(max(c_i)))$

Jako $d_\oplus(x \rightsquigarrow y \rightsquigarrow z)$ oznaczmy $\oplus$-*długość* najkrótszej krawędziowo ścieżki przechodzącej przez wierzchołki x, y oraz z.

Obserwacja.1. *Dla dowolnych wierzchołków a, b, c, d zachodzi równość:*

$$d_\oplus(a \rightsquigarrow b \rightsquigarrow c) \ \oplus \ d_\oplus(a \rightsquigarrow b \rightsquigarrow d) = d_\oplus(c \rightsquigarrow b \rightsquigarrow d)$$

Dowód powyższej obserwacji opiera się na przemienności oraz odwrotności operacji **xor**:

$$\begin{gathered} d_\oplus(a \rightsquigarrow b \rightsquigarrow c) \ \oplus \ d_\oplus(a \rightsquigarrow b \rightsquigarrow d) = \\ = d_\oplus(a \rightsquigarrow b) \ \oplus \ d_\oplus(b \rightsquigarrow c) \ \oplus \ d_\oplus(a \rightsquigarrow b) \ \oplus \ d_\oplus(b \rightsquigarrow d) = \\ = \underbrace{\Big(d_\oplus(a \rightsquigarrow b) \ \oplus \ d_\oplus(a \rightsquigarrow b)\Big)}_{0} \oplus d_\oplus(b \rightsquigarrow c) \ \oplus \ d_\oplus(b \rightsquigarrow d) = \\ = d_\oplus(b \rightsquigarrow c) \ \oplus \ d_\oplus(b \rightsquigarrow d) = d_\oplus(c \rightsquigarrow b \rightsquigarrow d) \end{gathered}$$

Obserwacja.2. *Dla dowolnych dwóch wierzchołków a, b zachodzi:*

$$d_\oplus(a \rightsquigarrow lca(a,b) \rightsquigarrow b) = d_\oplus(a \rightsquigarrow b)$$

Obserwacja ta jest oczywista w przypadku drzew i raczej nie wymaga tłumaczenia.

Zauważmy, że jeśli ustalimy sobie pewien wierzchołek źródłowy s i dla każdego wierzchołka x obliczymy wartości $d_\oplus(s \rightsquigarrow x)$, to posiłkując się przedstawionymi obserwacjami, gdyż:

$$d_\oplus(s \rightsquigarrow x_i) \ \oplus \ d_\oplus(s \rightsquigarrow x_j) = d_\oplus(x_i \rightsquigarrow x_j)$$

Wszystkie wartości $d_\oplus(s \rightsquigarrow x)$ można z łatwością obliczyć za pomocą algorytmu DFS.

W naszym rozwiązaniu, będziemy obliczać ile w naszym grafie znajduje się ścieżek, których wartość ich $\oplus$-*długości* ma zapalony konkretny bit. Otóż zauważmy, że końcowy wynik możemy opisać następującą formułą:

$$\sum_{i=1}^{32} 2^{i-1} \cdot cnt(i)$$

Przy czym $cnt(i)$ to liczba ścieżek w grafie, dla których ich $\oplus$-*długość* ma zapalony i-ty bit od końca. Wartość $cnt(i)$ możemy obliczyć w względnie prosty sposób. Wystarczy zauważyć, że:

$$cnt(i) = cnt_{d_\oplus}(i,0) \cdot cnt_{d_\oplus}(i,1)$$

Gdzie $cnt_{d_\oplus}(i,0)$ to liczba ścieżek dla których ich $\oplus$-*długość* posiada zapalony i-ty bit, zaś $cnt_{d_\oplus}(i,1)$ liczbę ścieżek dla których i-ty bit nie jest zapalony.

Posiłkując się obserwacją 2, możemy zauważyć, że odpowiednie wartości $cnt_{d_\oplus}$ są równe liczbie ścieżek wychodzących wierzchołka źródłowego s dla, których ich wartość $d_\oplus$ ma zapalony konkretny bit lub nie. Wystarczy, więc dla wszystkich wierzchołków x grafu obliczyć wartości $d_\oplus(s \rightsquigarrow x)$, i na podstawie zapalonych w nich bitów wyznaczyć wartości $cnt_{d_\oplus}$, z których to pomocą możemy obliczyć cnt zatem też i końcowy wynik.