Lời giải bài CHAMELEON của PreVOI 2023

Nguyễn Ngọc Đăng Khoa

Tóm tắt đề bài

Cho một đồ thị liên thông gồm N đỉnh và M cạnh hai chiều. Gọi $F(X,Y)$ là số cạnh mà bắt buộc phải đi qua trên bất kì tuyến đường từ đỉnh X đến đỉnh Y. Tương tự như vậy, gọi $G(X,Y)$ là số đỉnh mà bắt buộc phải đi qua trên bất kì tuyến đường từ đỉnh X đến đỉnh Y (lưu ý là hai đỉnh X và Y đều được tính). Hãy tính tổng của $(F(X,Y) \times A + G(X,Y) \times B)^C$ với mọi cặp số (X,Y) thỏa mãn $1 \leq X < Y \leq N$, và đưa ra kết quả modulo $10^9 + 7$.

Giới hạn: $2 \leq N \leq 10^5$, $N - 1 \leq M \leq \min(\frac{N \times (N-1)}{2}, 2 \times 10^5)$, $0 \leq A \leq 10^9$, $0 \leq B \leq 10^9$, $1 \leq C \leq 2$.

Subtask 1 - $2 \leq N \leq 100$

Để làm được subtask này, thí sinh cần biết thuật toán BFS hoặc DFS trên đồ thị.

Để tính giá trị của $F(X,Y)$, ta lần lượt xét từng cạnh của đồ thị. Với mỗi cạnh, ta tạm thời xóa nó đi, rồi chạy thuật toán BFS hoặc DFS để tìm các thành phần liên thông của đồ thị. Sau đó, ta duyệt từng cặp đỉnh (X,Y). Khi hai đỉnh X và Y thuộc hai thành phần liên thông khác nhau, ta tăng giá trị hiện tại của $F(X,Y)$ lên 1. Ta có thể làm tương tự như vậy để tính giá trị của $G(X,Y)$, nhưng bây giờ ta lần lượt xóa đi từng đỉnh của đồ thị.

Ta phải chạy thuật toán BFS hoặc DFS một lần cho mỗi cạnh và đỉnh của đồ thị. Như vậy, độ phức tạp của thuật toán là $\mathcal{O}((N+M)^2)$.

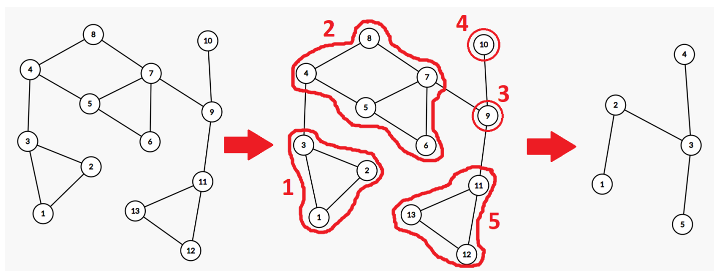
Subtask 4 và Subtask 5 - $B = 0$

Theo ý kiến của tác giả, subtask 4 và 5 dễ hơn subtask 2 và 3, nên hai subtask đó sẽ được chữa trước.

Để làm được hai subtask này, thí sinh cần biết khái niệm của cầu (bridge) trong đồ thị, bridge tree, và quy hoạch động (DP) trên cây. Có thể tham khảo thêm về bridge tree ở đây (phần 3.3).

Vì $B = 0$ nên ta chỉ cần quan tâm đến giá trị của $F(X, Y)$ với mỗi cặp đỉnh (X, Y). Đầu tiên, ta sẽ dựng bridge tree của đồ thị. Các bước để dựng bridge tree như sau:

1. Xác định các cầu (bridge) của đồ thị.
2. Tạm thời xóa đi các cầu khỏi đồ thị.
3. Nén lại mỗi thành phần liên thông thành một đỉnh.
4. Thêm lại các cầu vào đồ thị mới.



Có thể chứng minh rằng đồ thị mới thu được là một cây. Gọi K là số đỉnh của bridge tree. Mỗi đỉnh trong bridge tree tương ứng với một thành phần liên thông bị nén ở bước 3. Ta đánh số các đỉnh và các thành phần liên thông tương ứng từ 1 đến K.

Xét hai đỉnh X và Y trong đồ thị ban đầu. Gọi u và v lần lượt là chỉ số của hai thành phần liên thông chứa đỉnh X và đỉnh Y. Khi đó, giá trị của $F(X, Y)$ là số cạnh trên đường đi từ đỉnh u đến đỉnh v trong bridge tree, vì mỗi cạnh trong bridge tree tương ứng với một cầu trong đồ thị ban đầu. Ví dụ, trong hình ở trên, với $X = 3$ và $Y = 11$, ta có $u = 1$ và $v = 5$. Như vậy, $F(3, 11) = 3$, vì có 3 cạnh trên đường đi từ đỉnh 1 đến đỉnh 5 trong bridge tree.

Bây giờ ta đặt trọng số cho mỗi cạnh của bridge tree. Đối với subtask 4 và 5, ta sẽ đặt trọng số của tất cả các cạnh là A. Khi đó, giá trị của $F(X, Y) \times A$ là tổng trọng số của các cạnh trên đường đi từ đỉnh u đến đỉnh v, vì mỗi cạnh có trọng số là A.

Gọi S_i là số đỉnh thuộc thành phần liên thông thứ i, và $dist(u, v)$ là tổng trọng số của các cạnh trên đường đi từ đỉnh u đến đỉnh v. Khi đó, số cặp đỉnh (X, Y) trong đồ thị ban đầu mà tương ứng với cặp đỉnh (u, v) trong bridge tree là $S_u \times S_v$. Do đó, đáp án cuối cùng của bài toán là:

$$\sum_{1 \le u < v \le K} S_u \times S_v \times dist(u, v)^C$$

Như vậy, để làm được bài này, ta sẽ xử lí bài toán con ở trên.

Một bài toán con

Đề bài

Cho một cây gồm K đỉnh và $K-1$ cạnh, và một dãy số S gồm K phần tử. Mỗi cạnh của cây có một trọng số. Gọi $dist(u,v)$ là tổng trọng số của các cạnh trên đường đi từ đỉnh u đến đỉnh v. Hãy tính giá trị của tổng sau, modulo 10^9+7:

$$\sum_{1 \le u < v \le K} S_u \times S_v \times dist(u,v)^C$$

Giới hạn: $2 \le K \le 10^5$, $1 \le C \le 2$.

Xử lí $C = 1$

Đầu tiên ta chọn đỉnh 1 làm gốc của cây. Gọi P_i là đỉnh cha của đỉnh i, và W_i là trọng số của cạnh nối giữa đỉnh i và đỉnh P_i.

Với mỗi cạnh trong cây, ta tính số lần nó được đếm vào tổng. Xét cạnh nối giữa đỉnh i và đỉnh P_i. Nếu ta xóa cạnh này đi, thì đồ thị được tách ra thành hai cây con. Gọi hai cây con này lần lượt là T_1 và T_2, trong đó T_1 là cây con chứa đỉnh i. Có thể thấy rằng đường đi từ đỉnh u đến đỉnh v đi qua cạnh thứ i nếu một trong hai đỉnh đó thuộc cây con T_1, và đỉnh còn lại thuộc cây con T_2. Như vậy, số lần cạnh thứ i được đếm vào tổng là:

$$\sum_{u \in T_1, v \in T_2} S_u \times S_v = \sum_{u \in T_1} S_u \times \sum_{v \in T_2} S_v = \sum_{u \in T_1} S_u \times (\sum_{i=1}^{K} S_i - \sum_{u \in T_1} S_u)$$

Gọi sum_i là tổng của các giá trị S_u của cây con có gốc tại đỉnh i. Khi đó, tổng trên bằng $sum_i \times (sum_1 - sum_i)$.

Như vậy, đáp án của bài toán là:

$$\sum_{i=1}^{K} W_i \times sum_i \times (sum_1 - sum_i)$$

Có thể sử dụng DP trên cây để tính giá trị của sum_i cho mỗi đỉnh i. Độ phức tạp của thuật toán là $\mathcal{O}(K)$.

Xử lí $C = 2$

Ta có thể xử lí $C = 2$ một cách tương tự, nhưng bây giờ ta tính số lần mỗi cặp cạnh trong cây được đếm vào tổng. Xét cạnh nối giữa đỉnh i và đỉnh P_i và cạnh nối giữa đỉnh j và đỉnh P_j, trong đó $i \ne j$ (nếu $i = j$ thì ta xử lí như trên).

Có 3 trường hợp có thể xảy ra:

- TH1: i là tổ tiên của j. Khi đó cặp cạnh này được tính $sum_j \times (sum_1 - sum_i)$ lần.

- TH2: j là tổ tiên của i. Xử lí tương tự TH1.

- TH3: i không phải tổ tiên của j và j không phải tổ tiên của i. Khi đó cặp cạnh này được tính $sum_i \times sum_j$ lần.

Tương tự như cách ta xử lí $C = 1$, ta có thể xử lí cả 3 trường hợp bằng DP trên cây. Độ phức tạp của thuật toán là $\mathcal{O}(K)$.

Subtask 2 và Subtask 3 - $A = 0$

Để làm được hai subtask này, thí sinh cần biết khái niệm của khớp trong đồ thị, block-cut tree, và quy hoạch động (DP) trên cây. Có thể tham khảo thêm về block-cut tree ở đây (phần 3.4).

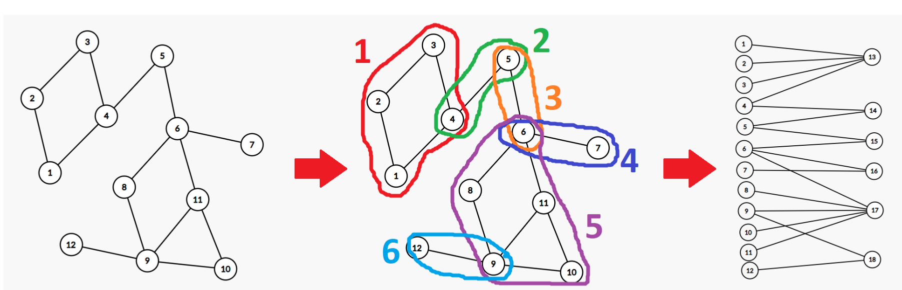
Vì $A = 0$ nên ta chỉ cần quan tâm đến giá trị của $G(X, Y)$ với mỗi cặp đỉnh (X, Y). Đầu tiên, ta sẽ dựng block-cut tree của đồ thị. Các bước để dựng block-cut tree như sau:

1. Xác định các BCC (biconnected component) của đồ thị. Gọi K là số BCC trong đồ thị. Đánh số các BCC từ 1 đến K.

2. Dựng một đồ thị hai phía gồm $N + K$ đỉnh, trong đó có N đỉnh được đánh số từ 1 đến N ở phía bên trái, và K đỉnh được đánh số từ $N + 1$ đến $N + K$ ở phía bên phải. Các đỉnh ở phía bên trái tương ứng với các đỉnh trong đồ thị ban đầu, và các đỉnh ở phía bên phải tương ứng với các BCC trong đồ thị ban đầu.

3. Với mỗi cặp số (i, j) thỏa mãn $1 \le i \le N$ và $1 \le j \le K$, nối một cạnh từ đỉnh i đến đỉnh $N + j$ nếu đỉnh i nằm trong BCC thứ j của đồ thị ban đầu.

(Lưu ý là thông thường phía bên trái chỉ bao gồm các đỉnh tương ứng với đỉnh khớp. Đối với bài toán này, nó sẽ tiện lợi hơn nếu phía bên trái bao gồm tất cả các đỉnh trong đồ thị ban đầu.)



Có thể chứng minh rằng đồ thị mới thu được là một cây. Xét hai đỉnh X và Y trong đồ thị ban đầu. Ta thấy rằng giá trị của $G(X, Y)$ là số đỉnh trên đường đi từ đỉnh X đến đỉnh Y của block-cut tree mà thuộc phía bên trái của đồ thị, vì mỗi đỉnh như vậy (ngoại trừ hai đỉnh X và Y) tương ứng với một khớp trong đồ thị ban đầu. Ví dụ, trong hình ở trên, $G(1, 12) = 6$, vì có 6 đỉnh trên đường đi từ đỉnh 1 đến đỉnh 12 trong block-cut tree mà thuộc phía bên trái của đồ thị, là các đỉnh 1, 4, 5, 6, 9, và 12.

Bây giờ ta đặt trọng số cho mỗi cạnh của block-cut tree. Đối với subtask 2 và 3, ta sẽ đặt trọng số của tất cả các cạnh là $\frac{B}{2}$. Lưu ý là nếu B là một số lẻ, thì ta có thể cộng giá trị của B lên $10^9 + 7$ rồi mới chia 2 - đáp án sẽ không thay đổi.

Vì mỗi cạnh nối giữa hai phía khác nhau, nên một đường đi bắt đầu từ phía bên trái gồm $2k$ cạnh sẽ đi qua $k + 1$ đỉnh nằm ở phía bên trái (tính cả hai đỉnh đầu và cuối). Do đó, đường đi từ đỉnh X đến đỉnh Y bao gồm $(G(X, Y) - 1) \times 2$ cạnh. Gọi $dist(X, Y)$ là tổng trọng số của các cạnh trên đường đi từ đỉnh X đến đỉnh Y. Khi đó, ta có $G(X, Y) \times B = (G(X, Y) - 1) \times 2 \times \frac{B}{2} + B = dist(X, Y) + B$.

Như vậy đáp án cuối cùng của bài toán là:

$$\sum_{1 \leq X < Y \leq N} (dist(X, Y) + B)^C$$

Khi $C = 1$, ta có:

$$\sum_{1 \leq X < Y \leq N} (dist(X, Y) + B)^1$$

$$= \sum_{1 \leq X < Y \leq N} dist(X, Y) + \frac{N \times (N - 1)}{2} \times B$$

Ta có thể sử dụng bài toán con ở subtask 4 và subtask 5 để tính giá trị của tổng này. Cây của ta sẽ gồm $N + K$ đỉnh, và dãy S thỏa mãn $S_i = 1$ với $1 \leq i \leq N$, và $S_i = 0$ với $N + 1 \leq i \leq N + K$.

Khi $C = 2$, ta có:

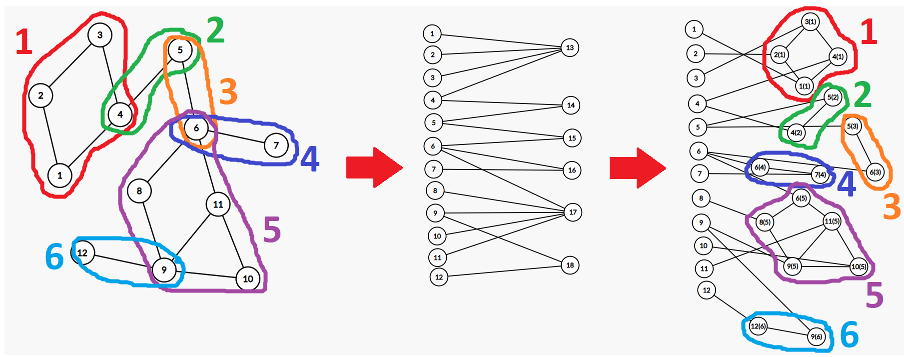
$$\sum_{1 \leq X < Y \leq N} (dist(X, Y) + B)^2$$

$$= \sum_{1 \leq X < Y \leq N} dist(X, Y)^2 + B \times 2 \times \sum_{1 \leq X < Y \leq N} dist(X, Y) + \frac{N \times (N - 1)}{2} \times B^2$$

Tương tự như trên, ta có thể sử dụng bài toán con để tính giá trị của tổng này.

Subtask 6 - Không có ràng buộc gì thêm

Đầu tiên, ta dựng block-cut tree của đồ thị giống subtask 2 và subtask 3. Sau đó, ta mở rộng mỗi đỉnh ở bên phải thành BCC tương ứng của nó trong đồ thị ban đầu. Một cạnh nối giữa đỉnh i và đỉnh $N + j$ trong block-cut tree ban đầu được thay thế bởi một cạnh nối giữa đỉnh i và đỉnh tương ứng với đỉnh i trong BCC thứ j.



Bây giờ ta đặt trọng số cho mỗi cạnh của đồ thị này. Như subtask 2 và subtask 3, các cạnh nối giữa một đỉnh và một BCC có trọng số là $\frac{B}{2}$, và như subtask 4 và subtask 5, các cạnh trong một BCC có trọng số là A.

Gọi $F'(X, Y)$ là tổng trọng số của các cạnh mà bắt buộc phải đi qua trên bất kì tuyến đường từ đỉnh X đến đỉnh Y trong đồ thị mới này. Khi đó, ta có $F(X, Y) \times A + G(X, Y) \times B = F'(X, Y) + B$. Do đó, đáp án cuối cùng của bài toán là:

$$\sum_{1 \leq X < Y \leq N} (F'(X, Y) + B)^C$$

Đến đây, ta có thể biến đổi tổng trên giống subtask 2 và subtask 3, rồi sử dụng thuật toán của subtask 4 và subtask 5 để xử lí phần còn lại. Độ phức tạp của thuật toán cuối cùng là $\mathcal{O}(N + M)$.