Demostración de correctitud de un algoritmo Greedy para resolver el Problema

1 Para el caso en que el grafo es un árbol

Recordemos que, en un árbol, al eliminar ciertas aristas se incrementa el número de componentes conexas.

Lemma 1 (Número de componentes en un árbol). *Si eliminamos todas las aristas de un árbol con n vértices, es decir, eliminamos $n-1$ aristas, obtenemos n componentes conexas.*

Proof. Procedemos por inducción en el número de vértices n.

Caso base: Si $n = 2$, un árbol tiene exactamente una arista. Al remover esta única arista, el grafo se divide en 2 componentes conexas, lo cual cumple la afirmación.

Hipótesis de inducción: Supongamos que para un árbol con $n = k-1$ vértices, removiendo $(k-2)$ aristas obtenemos $k-1$ componentes conexas.

Paso inductivo: Consideremos un árbol T con k vértices. Removamos una arista e del árbol. Al hacerlo, el árbol se divide en dos subárboles, digamos T_1 y T_2, que tienen n_1 y n_2 vértices respectivamente, donde $n_1 + n_2 = k$ y ambos $n_1, n_2 \geq 1$. Por la hipótesis de inducción, removiendo $n_1 - 1$ aristas de T_1 se obtienen n_1 componentes conexas y removiendo $n_2 - 1$ aristas de T_2 se obtienen n_2 componentes conexas. En total, removiendo $1 + (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 1 = k - 1$ aristas, obtenemos $n_1 + n_2 = k$ componentes conexas. □

Lemma 2 (Corte mínimo k en un árbol). *Sea $T = (V, E)$ un árbol con pesos no negativos en sus aristas, y sea la lista de aristas ordenada de forma no decreciente según sus pesos:*

$$w(e_1) \leq w(e_2) \leq \cdots \leq w(e_{n-1}).$$

Entonces, el conjunto de aristas de menor peso que separa el árbol en k componentes conexas está dado por:

$$C = \{e_1, e_2, \ldots, e_{k-1}\}.$$

Proof. Procedemos por contradicción. Supongamos que existe otro conjunto C' de $k-1$ aristas tal que el peso total $w(C')$ es menor que $w(C)$. Sin embargo, al remover $k-1$ aristas de T, por el resultado del lema anterior, el árbol se divide en k componentes conexas. Como C está formado por las $k-1$ aristas de menor peso, cualquier otro conjunto C' de $k-1$ aristas necesariamente tendrá un peso total mayor o igual a $w(C)$. Esto contradice la hipótesis de que $w(C') < w(C)$, por lo que C debe ser el conjunto deseado. □

Corollary 1. *Sobre árboles, el problema se puede resolver en tiempo polinomial, ya que basta con ordenar las aristas (lo cual se puede hacer en tiempo $O(n \log n)$) y seleccionar las $k-1$ aristas de menor peso.*

2 Para el caso en que el grafo es un bosque

Consideremos ahora un bosque, es decir, un grafo disconexo cuyos componentes son árboles.

Proposition 1. *Sea t el número de componentes conexas en un bosque $F = (V, E)$. Para cualquier $n \geq t$, si eliminamos $(n - t)$ aristas del bosque, obtenemos n componentes conexas.*

Proof. Procedemos por inducción en n.

Caso base: Si $n = t$, no eliminamos ninguna arista, y el bosque ya tiene t componentes, lo que cumple la afirmación.

Hipótesis de inducción: Supongamos que para $n = k-1$ (con $k-1 \geq t$), removiendo $(k-1-t)$ aristas se obtienen $k-1$ componentes conexas.

Paso inductivo: Consideremos un bosque F y supongamos que removiendo $(k-1-t)$ aristas se obtienen $k-1$ componentes conexas. Ahora, removamos una arista adicional de uno de los árboles resultantes. Por el lema aplicado a árboles, al remover una arista de ese árbol se incrementa el número de componentes en 1. Así, en total, removiendo $(k-1-t)+1 = k-t$ aristas, obtenemos k componentes conexas. □

Lemma 3. *Sea $F = (V, E)$ un bosque con pesos no negativos en sus aristas, y sean t el número de componentes conexas de F y las aristas ordenadas por la función de peso de forma no decreciente:*

$$w(e_1) \leq w(e_2) \leq \cdots \leq w(e_{|V|}).$$

Entonces, el conjunto de aristas de menor peso cuya eliminación divide el bosque en k componentes conexas está dado por:

$$C = \{e_1, e_2, \ldots, e_{k-t}\}.$$

Proof. Supongamos, para obtener una contradicción, que existe un conjunto C' de aristas con $|C'| < k - t$ y $w(C') < w(C)$ que, al removerlas, divide el bosque en k componentes conexas. Sin embargo, según la proposición anterior, remover menos de $k-t$ aristas en un bosque con t componentes produce menos de k componentes conexas. Esto contradice la definición de corte mínimo k, que debe separar el bosque en exactamente k componentes. Por lo tanto, C es el conjunto de aristas de menor peso cuya remoción divide F en k componentes conexas. □