

Lemma:

Let $G = (V, E)$ be an undirected graph, and let $f : E \rightarrow \mathcal{R}^+$ be a cost (a.k.a. weight) function for G .

Let $F = \{ (V_1, E_1), (V_2, E_2), \dots, (V_k, E_k) \}$ be a minimal spanning forest for G . Let

$$\hat{E} = \cup_{i=1}^k E_i$$

be the set of all edges in the spanning forest F .

Let $e = (u, v)$ be the edge of lowest cost chosen by Kruskal's algorithm. Then, there exists a spanning tree T which includes $\hat{E} \cup \{e\}$ which is as low cost as any spanning tree which includes \hat{E} .

Proof: (by contradiction)

Suppose there exists a spanning tree $T' = (V, E')$ which includes \hat{E} but not e , and $\text{cost}(T')$ is (strictly) less than the cost of any spanning tree that includes \hat{E} .

We reach a contradiction by constructing a new spanning tree $T'' = (V, E'')$ that includes \hat{E} and:

$$\text{cost}(T'') \leq \text{cost}(T')$$

Construction of T'' :

The vertices u and v must be connected in T' (because T' is a tree). The tree T' does not include e , therefore if we add e to T' , a cycle results. Let e' denote an edge on that cycle such that $e' \neq e$ and $e' \notin \hat{E}$. We know such an edge exists because u and v are connected in T' , but u and v are not connected in F .

Let $T'' = (V, E'')$ where

$$E'' = E' - \{e'\} \cup \{e\} .$$

The conditions under which e is chosen by Kruskal's algorithm guarantee that

$$f(e) \leq f(e')$$

Therefore:

$$\text{cost}(T'') \leq \text{cost}(T') \tag{1}$$

Inequality (1) is a contradiction, because the proof hypothesis requires:

$$\text{cost}(T'') > \text{cost}(T')$$