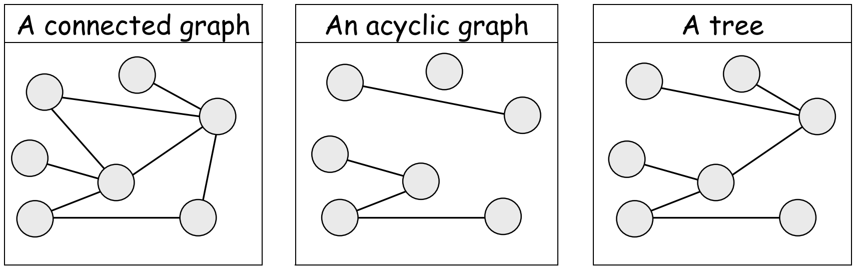
Minimal Spanning Trees

A graph is connected if, for every pair (u,v) of vertices, there is a path between u and v.

A graph is acyclic if it has no cycles.

A tree is a graph that is connected and acyclic.



Consider a finite graph G with n vertices.

i) If G is connected, then G has *at least* $n-1$ edges. It has exactly $n-1$ edges if and only if it is a tree.

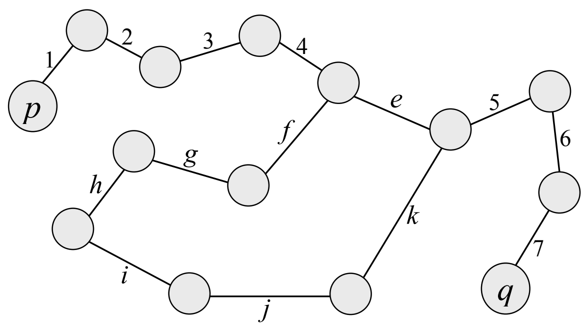
ii) If G is acyclic, it has *at most* $n-1$ edges. It has exactly $n-1$ edges if and only if it is a tree.

A spanning tree for a connected graph $G = (V, E)$ is a tree $T = (V, S)$, with $S \subseteq E$.

Every connected graph G contains a spanning tree T. In fact,

```
T = G;
while ( T contains a cycle )
    remove from T an edge on some cycle;
```

always terminates with T a spanning tree for G. The key is that removing an edge lying on a cycle of a connected graph cannot disconnect the graph.



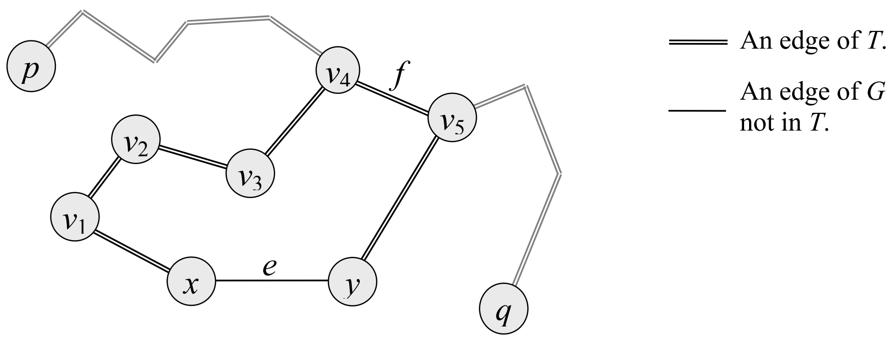
p, q = two vertices of G.
e = edge on a cycle of G.
1,2,3,4,e,5,6,7 = path from p to q, using edge e.
1,2,3,4,f,g,h,i,j,k,5,6,7 = alternate path from p to q, not using edge e.

If $G = (V, E, W)$ is a weighted connected graph, a minimal spanning tree (or MST) for G is a spanning tree whose total weight is minimal, among all spanning trees.

Every graph has an MST. The MST need not be unique, but we will see that it is unique if all the edge weights of the graph are distinct.

Lemma. Let $G = (V, E)$ be a connected graph, and let $T = (V, S)$ be a spanning tree. Let $e = xy$ be an edge of G not in T.

For any edge f on the path from x to y in T,
$T_{f\to e} = \big(V, (S \cup \{e\}) - \{f\}\big)$ is another spanning tree for G. (*In other words, we may substitute e for f and retain a spanning tree.*)



p, ..., v_4, v_5, ..., q is a path from p to q in T, using edge f.

p, ..., v_4, v_3, v_2, v_1, x, y, v_5, ..., q is a path from p to q in $T_{f\to e}$.

$T_{f\to e}$ is a connected graph on V, and $T_{f\to e}$ has the same number of nodes as T, implying $T_{f\to e}$ is a spanning tree.

Proposition 1. Let $G = (V, E, W)$ be a weighted graph, let $T = (V, S)$ be an MST for G, and let $e = xy$ be an edge of G not in T. Then

i) $w(e) \geq w(f)$ for any edge f on the path in T from x to y.

ii) If $w(e) = w(f)$, then we may obtain another MST $T_{f\to e}$ for G by replacing f by e in S.

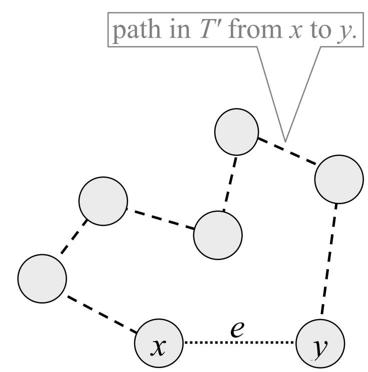
Proof. By the Lemma, if we replace f by e in T, we obtain a spanning tree $T_{f\to e}$ of cost $w(T_{f\to e}) = w(T) + w(e) - w(f)$.

If $w(e) < w(f)$, then $w(T_{f\to e}) < w(T)$, contrary to T being a *minimal* spanning tree.

If $w(e) = w(f)$, then $w(T_{f\to e}) = w(T)$, and so $T_{f\to e}$ also is an MST.

Proposition 2. If all the edge weights in G are distinct, then G has a unique MST.

Proof. If $T = (V, S)$ and $T' = (V, S')$ are two distinct MSTs for G, let $e = xy$ be the cheapest edge of G that is in one of T or T', but not both. (Since all the edge weights are distinct, there is a unique cheapest edge with this property.) Assume e is in T.



By Proposition 1, $w(e) \geq w(f)$ every edge f on the path in T' from x to y. But since edge weights are distinct, $w(e) > w(f)$. By the way e was chosen, every edge on the path in T' from x to y also lies in T. But these edges of T, plus the edge e of T, form a cycle, contrary to T being a tree.

We want to build up an MST for G, one edge at a time.

We know that $(V, \emptyset)$ is contained in *every* MST for G.

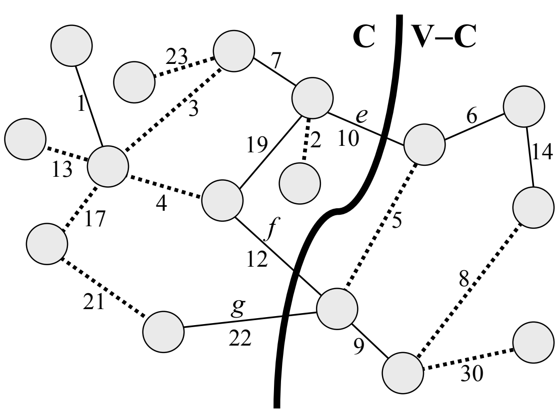
At some point, we will have a set A of edges of G, such that (V,A) is *contained in* an MST for G. We want to add another edge to A, so (V,A) is still contained in a (possibly different) MST for G.

A <u>cut</u> $(C, V-C)$ of the vertex set V is a partition of V into two disjoint subsets (C and $V-C$).

An edge of G <u>crosses the cut</u> if the edge has one endpoint in C and the other endpoint in $V-C$.

A cut $(C, V-C)$ respects a subset A of the edges of G (or respects a subgraph (V,A) of G) if no edge in A crosses the cut.

This is equivalent to saying that C is a union of connected components of (V,A).



.... edge of A.

——— edge of G not in A.

The cut $(C, V-C)$ respects the subgraph (V,A).

Edges e, f, and g of G cross the cut.

The edge of minimal weight crossing the cut is e.

Theorem. If (V,A) is a subgraph of G contained in an MST T for G, if $(C, V-C)$ is a cut respecting the subgraph (V,A), and if e is an edge of G of minimal weight subject to crossing the cut, then $(V, A \cup \{e\})$ is contained in an MST for G (not necessarily T).

Proof. If e is an edge of T, then $(V, A \cup \{e\})$ is contained T, and we are done.

Otherwise let x and y be the endpoints of e, with x in C and y in $V-C$.

There is a path in T from x to y (not containing e).

This path must cross the cut at least once, so there is an edge f on the path that does cross the cut. This means f cannot be in A. By the way e was chosen, $w(e) \le w(f)$.

On the other hand, Proposition 1 tells us that $w(e) \ge w(f)$. We must have $w(e) = w(f)$.

Proposition 1 then tells we may replace f by e in T, obtaining another MST, which we denoted $T_{f \to e}$. This MST contains $(V, A \cup \{e\})$ since the removed edge f was not in A.