Module 13.1: Spanning Trees

MCIT Online - CIT592 - Professor Val Tannen

Lecture Notes

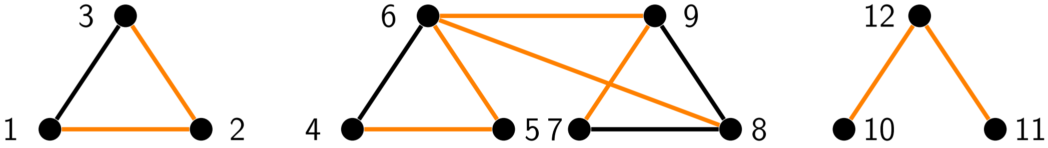
Spanning trees and forests

A **spanning subgraph** of the graph $G = (V, E)$ is a subgraph whose vertex set is the entire set V.

A **spanning tree** of a *connected* graph G is a spanning subgraph that is a tree.

A **spanning forest** of a graph G consists of a spanning tree for each of the connected components of G.

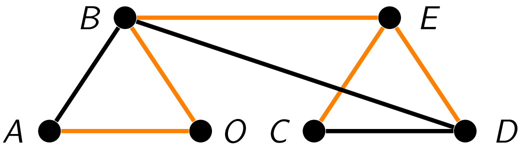
Example: Graph with spanning forest. Forest edges are in orange.



Penn Engineering

Laying cable

Problem. The graph below represents roads in a mountain village between houses. The cable company has already wired customer O and wants to lay cable to the other customers: A, B, C, D, E. Find a cheap way to do this.



Answer. If the cost of laying cable along each road segment is the same then the answer is **any spanning tree**. (Example with edges along which cable is laid is in orange above.)

The realistic version of this problem puts cost labels on the edges and asks for a **minimum spanning tree**.

Spanning trees always exist

Proposition. Every connected graph has a spanning tree. (Hence every graph has a spanning forest.)

Proof. Assume G is connected. Consider all the connected spanning subgraphs of G. There is at least one such subgraph, G itself.

By the Well-Ordering Principle at least one of the connected spanning subgraphs, call it S, must have the **smallest** number of edges.

Claim. S is acyclic.

By contradiction. Suppose S has a cycle. We can delete an edge, call it e, in this cycle and the subgraph remains connected. Indeed, any walk that uses e can instead use the rest of the cycle. This yields a connected subgraph with **strictly less** edges than S. Contradiction.

S is an acyclic and connected spanning subgraph, hence a spanning tree.

Cut edges

Recall that an edge is a **cut edge** in a graph if erasing it strictly increases the number of connected components.

We have proved in a previous segment:

Proposition. Removing a cut edge in a graph increases the number of connected components by exactly one.

We have also proved:

Proposition. An edge is a cut edge iff it does not belong to any cycle.

Corollary. A connected graph is a tree iff each one of its edges is a cut edge.

The edge-pruning algorithm

Input: A connected graph $G = (V, \{e_1, \ldots, e_m\})$ where $|V| \geq 2$.

Output: A spanning tree T of G.

(1) Let $T = G$

(2) For $k = 1, \ldots, m$ do:

(2a) if e_k is **not** a cut edge in T

(2b) then remove e_k from T

(2c) else keep e_k in T

(3) Output T

Note that every edge is examined **once** when a decision to **remove** or **keep** the edge is made.

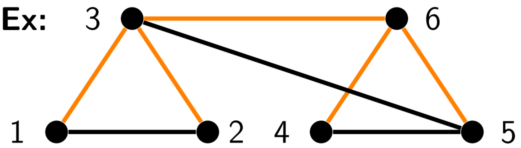
Why edge-pruning works

T starts as a spanning sugbraph and no vertices are removed. Therefore, the output is spanning too.

T starts connected and no edge removal disconnects it. Therefore, the output is connected too.

Suppose, toward a contradiction, that the algorithm outputs a graph with a cycle. The edge in that cycle that was **last examined** was not a cut edge when it was examined. Contradiction.

Ex:



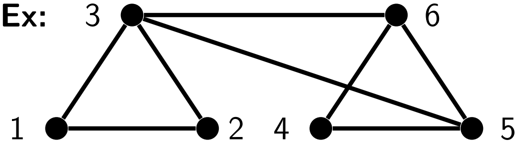
Edges in order of examination: 1–2, 2–3, 3–1, 3–5, 3–6, 4–5, 4–6, 5–6

Output spanning tree with edges: 2–3, 3–1, 3–6, 4–6, 5–6

Quiz

Run the edge-prunning algorithm on the same graph we saw but with a different order for examining the edges:

Ex:



Edges in order of examination:
3–6, 3–5, 4–6, 4–5, 5–6,
2–3, 1–2, 1–3

Does the edge 5–6 belong to the resulting spanning tree?

(A) Yes.

(B) No.

Answer

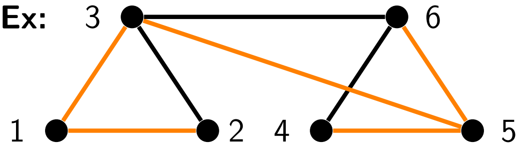
(A) Yes.
Correct. See more information for the resulting spanning tree.

(B) No.
Incorrect. See more information for the resulting spanning tree.

More Information

By applying the edge-prunning algorithm by considering edges in the given order the resulting spanning tree is the following:

Ex:



One can clearly see that the edge 5-6 belongs to the tree.

The edge-growing algorithm

Input: A connected graph $G = (V, \{e_1, \ldots, e_m\})$ where $|V| \geq 2$.

Output: A spanning tree T of G.

(1) Let $T = (V, \emptyset)$ (the edgeless graph on V)

(2) For $k = 1, \ldots, m$ do:

  (2a) if adding e_k to T does **not** form a cycle with edges already in T

  (2b) then add e_k from T

  (2c) else leave e_k out of T

(3) Output T

Note that every edge is examined **once** when a decision to **add** or **leave out** the edge is made.

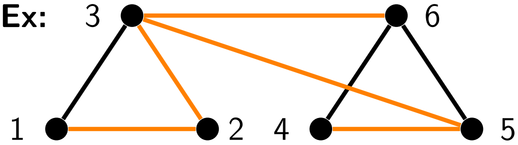
Why edge-growing works

T starts as a spanning sugbraph and no vertices are added.
Therefore, the output is spanning too.

T starts acyclic and no edge addition creates a cycle.
Therefore, the output is acyclic too.

It can be shown that the output is also connected.

Ex:



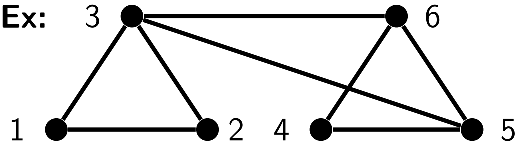
Edges in order of examination:
1–2, 2–3, 3–1, 3–5, 3–6,
4–5, 4–6, 5–6

Output spanning tree with edges: 1–2, 2–3, 3–5, 3–6, 4–5
(Shown in orange above.)

Quiz

Run the edge-growing algorithm on the same graph that we saw but with a different order for examining the edges:

Ex:



Edges in order of examination:
3–6, 3–5, 4–6, 4–5, 5–6,
2–3, 1–2, 1–3

Does the edge 5–6 belong to the resulting spanning tree?

(A) Yes.

(B) No.

Answer

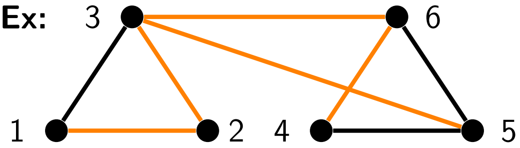
(A) Yes.
Incorrect. See more information for the resulting spanning tree.

(B) No.
Correct. See more information for the resulting spanning tree.

More Information

By applying the edge-growing algorithm by considering edges in the given order the resulting spanning tree is the following:

Ex:



One can clearly see that the edge 5-6 does not belong to the tree.

Module 13.2: Graph Coloring

MCIT Online - CIT592 - Professor Val Tannen

Lecture Notes

Graph coloring

Let $G = (V, E)$ be a graph and let k be a positive integer. A **k-coloring** of G is a function $f : V \to [1..k]$.

A coloring is called **proper** when for any edge u–v we have $f(u) \neq f(v)$.

A graph that admits a proper k-coloring is called **k-colorable**. Note that k-colorable implies j-colorable for any $j \geq k$.

Clearly, a graph with n vertices is n-colorable.

The smallest k such that G is k-colorable is called the **chromatic number** of G and is denoted by $\chi(G)$.

Prop. $\chi(G) = 1$ iff G is edgeless.

The chromatic number of some graphs

Proposition. For $n \geq 2$ we have $\chi(P_n) = 2$.

Proposition.

$$\chi(C_n) = \begin{cases} 2 & \text{when } n \text{ is even} \\ 3 & \text{when } n \text{ is odd} \end{cases} .$$

Proposition. $\chi(K_n) = n$.

Proof. Obvious for $n = 1$. For $n \geq 2$, any k-coloring with $k < n$ cannot be proper. Indeed, by the Pigeonhole Principle, at least two vertices get the same color. But in K_n any two vertices are adjacent.

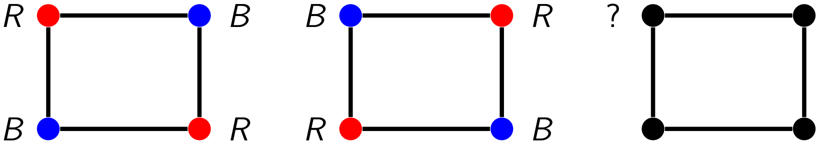
Bipartite graphs

2-colorable graphs are also called **bipartite**.

I will commonly refer to the two colors in bipartite graphs as **red** and **blue**.

Proposition. Every path graph is bipartite. A cycle graph is bipartite iff it has an even number of nodes.

In how many ways can we 2-color a graph? We try it on C_4:



In a connected graph: 2 ways! (Swap colors.) In general: 2 ways for each cc.

Every tree is bipartite

Proposition. Every tree is bipartite.

Proof. The edgeless tree is 1-colorable hence 2-colorable. For trees with 2 or more nodes the proof is by induction on the number of vertices.

(BC) $n = 2$. Color one of the vertices red and the other blue.

(IS) Let $k \geq 2$. Assume (IH) that any tree with k vertices is bipartite.
Let G have $k + 1$ vertices. As we saw in a previous segment. G has a leaf, u. Delete u and the one edge adjacent to u to obtain G_u. As we saw, G_u is also a tree. It has k vertices.

By IH, G_u is bipartite, hence it has a proper red/blue coloring. Let v be the vertex adjacent to u in G. W.l.o.g. assume v is red in this coloring. Now put back u and its adjacent edge and color u blue to extend the coloring to all of G.

Quiz

We saw that $\chi(K_5) = 5$. Now remove one edge from K_5. The chromatic number of the resulting graph is

(A) 3

(B) 4

(C) 5

Answer

We saw that $\chi(K_5) = 5$. Now remove one edge from K_5. The chromatic number of the resulting graph is

(A) 3

Incorrect. Even after we remove an edge the graph contains K_4 as a subgraph so its chromatic number must be at least 4.

(B) 4

Correct. We can reuse one of the 4 colors used on the K_4 subgraph. See complete explanation in More information.

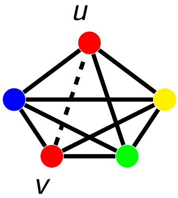
(C) 5

Incorrect. We can reuse one of the 4 colors used on the K_4 subgraph. See complete explanation in More information.

More information

Suppose we remove an edge, say u–v from K_5 and let G be the resulting graph. Let x, y, z be the other 3 vertices. The subgraph of G induced by v, x, y, z is isomorphic to K_4 so it has a proper 4-coloring. Now we extend this coloring to a 4-coloring of K_5 by giving u the same color as v.

G :



This coloring is proper for G because there is no edge between u and v.

On the other hand, there is no proper 3-coloring of G because that would also be a proper 3-coloring of its subgraph isomorphic to K_4, which is impossible.

DO NOT INCLUDE THIS ACTIVITY

ACTIVITY: Complete bipartite graphs

The **complete bipartite graph** $K_{m,n}$ has m red nodes, n blue nodes and an edge between every red node and every blue node.

Q1 How many edges?

Q2 What is the maximum degree? What is the minimum degree?

AT THE END WE SAY THAT By the way: $K_{m,n} \simeq K_{n,m}$. The isomorphism maps red nodes to blue nodes and vice versa. Keep it informal

Module 13.3: No Odd Cycles

MCIT Online - CIT592 - Professor Val Tannen

Lecture Notes

Characterization of bipartite graphs

Proposition. A graph is **bipartite** iff it does not contain a cycle of **odd** length.

From this proposition we obtain another proof of "every tree is bipartite".

We also obtain "every subgraph of a bipartite graph is also bipartite".

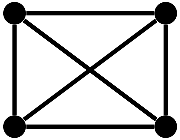
Lemma. If S is a subgraph of G then $\chi(S) \leq \chi(G)$.

Proof (of lemma). Indeed for any k, any k-coloring of G is also a k-coloring of S.

Proof (of proposition). The first implication is shown by contrapositive. If G contains an odd cycle then by the lemma, $\chi(G) \geq 3$, so G cannot be bipartite.

Quiz

Below you can see K_4, i.e. the complete graph with four vertices. If we remove an edge from K_4 is the resulting graph bipartite?

K_4 : 

(A) Yes.

(B) No.

Answer

(A) Yes.
Incorrect. Recall that a graph is bipartite iff it does not contain a cycle of odd length.

(B) No.
Correct. Since a graph is bipartite iff it does not contain a cycle of odd length, and the resulting graph still has cycles of length 3.

Distance in a connected graph

Let $G = (V, E)$ be a connected graph. The **distance** between two vertices $u, v \in V$, notation $d(u, v)$, is the length of a shortest path from u to v.

The distance can be defined because of the Well-Ordering Principle.

When $u = v$ we have $d(u, u) = 0$ given by the path of length 0. Moreover, $d(u, v) = d(v, u)$ since the reversal of a path has the same length.

Lemma (The Triangle Inequality). $d(u, v) \; \leq \; d(u, w) + d(w, v)$

Proof. Just the case when u, v, w are distinct. There is a path of length $d(u, w)$ from u to w and a path of length $d(w, v)$ from w to v. Concatenating these two paths gives us a walk of length $d(u, w) + d(w, v)$. But $d(u, v)$ is the length of a shortest path from u to v and hence it is $\leq$ than the length of any walk from u to v.

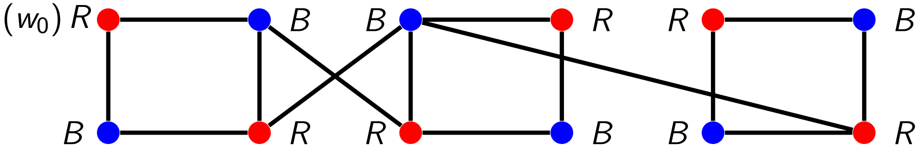
Define a proper 2-coloring

W.l.o.g. we can assume that G is **connected**. Indeed, a graph has a proper coloring iff each of its connected components have a proper coloring.

We define a 2-coloring using the concept of **distance** that we defined for any connected graph. Fix an arbitrary vertex w_0. Now color the vertices of G:

- w is colored **red** when $d(w_0, w)$ is **even** in particular, w_0 is colored red.
- w is colored **blue** when $d(w_0, w)$ is **odd**.

This is a 2-coloring, but is it **proper**? We have a proof of this in the segment entitled "Proof of 2-colorability". Here is an example of the proof "in action":



Module 13.4: Cliques and Independent Sets

MCIT Online - CIT592 - Professor Val Tannen

Lecture Notes

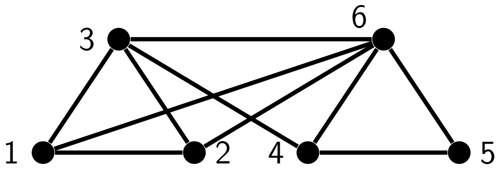
Cliques

A **complete** subgraph of a graph G is a subgraph that is isomorphic to the complete graph K_n for some n. The term **clique** (of/in G) is used both for such a complete subgraph of G and for a subset of the vertices of G that induces a subgraph that is complete.

An alternative definition of clique is a subset of vertices **any two** of which are **adjacent**.

The **size** of a clique is its number of vertices. Determining whether a graph has a clique of a given size is a well-known hard computational problem.

Example:



Cliques:

size 4: $\{1, 2, 3, 6\}$

size 3: $\{1, 2, 3\}$, $\{3, 4, 6\}$, $\{1, 3, 6\}$, $\{4, 5, 6\}$ etc.!

Cliques in practice

The word **clique** itself was first used by scientists studying **social networks**. I am sure you all remember "cliques" of friends in school... Not surprisingly, I belonged to a clique of "geeks" passionate about Math.

Let's call a clique in the Facebook graph **maximal** if it cannot be extended. It is believed that Facebook has maximal cliques with at least 200-250 people (www.quora.com).

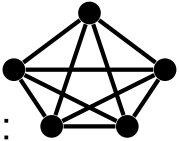
Cliques are used in **computational biology**: e.g., in gene expression clustering, evolutionary trees, ecological niches, and protein-protein interaction networks.

They are also used in **computational chemistry** and in **communication network** analysis (`wikipedia.org/wiki/Clique`).

Quiz

WE ARE NOT MAKING THIS QUIZ INTO AN IN-VIDEO ITEM. WE MIGHT PUT SOMETHING SIMILAR IN THE Module 13 quiz prep

How many cliques of size k does K_n have? Below you can see K_5 in order to think about the solution with a concrete example:

K_5 : 

(A) $k!$

(B) $\frac{n!}{k!\,(n-k)!}$

(C) $\frac{n!}{(n-k)!}$

Answer

(A) $k!$

Incorrect. Recall that in a complete graph every two vertices are adjacent. Use this fact coupled with the definition of a clique to find the correct answer.

(B) $\frac{n!}{k!\,(n-k)!}$

Correct. Since in a complete graph every two vertices are adjacent, every subset of k vertices forms a complete sub-graph. Therefore, there are $\binom{n}{k}$ cliques of size k in K_n.

(C) $\frac{n!}{(n-k)!}$

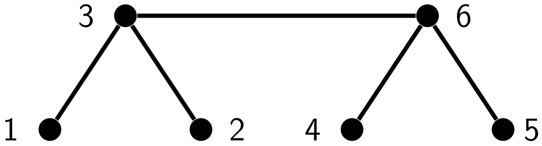
Incorrect. Recall that the order we pick the vertices in a subgraph does not matter.

Independent set

Let $G = (V, E)$ be a graph. A subset of vertices $S \subseteq V$ is called an **independent set** when *no two* vertices in S are adjacent. Observe that an alternative definition of independent set is that the induced subgraph is edgeless.

Proposition. A graph is k-colorable iff its set of vertices can be partitioned into k independent sets.

Example:



Independent sets:
size 4: $\{1, 2, 4, 5\}$
size 3: $\{1, 4, 5\}$, $\{1, 2, 6\}$, etc.
size 2: $\{1, 6\}$, $\{4, 5\}$, $\{2, 4\}$,etc.

Complement of a graph

Let $G = (V, E)$ be a graph. The **complement** of G is the graph $\overline{G} = (V, \overline{E})$ where

$$\overline{E} = \{\{u, v\} \mid u, v \in V \wedge u \neq v \wedge \{u, v\} \notin E\}.$$

Proposition. Let $G = (V, E)$ be a graph and let $S \subseteq V$ be a set of vertices. Then, S is a clique in G iff it is an independent set in $\overline{G}$.

Proof. In short, because for any $u, v \in S$, we have u–v in G iff $\neg(u$–$v)$ in $\overline{G}$.

This shows that the concepts of clique and of independent set are **dual**.

Mathematicians use the word **duality** for such situations. This is relevant when results about one concept transfer into results about its dual.

Quiz

How many edges does the complement of P_n have:

(A) $\frac{n(n-1)}{2}$

(B) $\frac{n(n-2)}{2}$

(C) $\frac{(n-1)(n-2)}{2}$

Answer

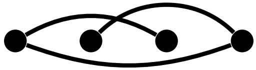
(A) $\frac{n(n-1)}{2}$
Incorrect. You should think how many edges K_n has and subtract the number of edges P_n has.

(B) $\frac{n(n-2)}{2}$
Incorrect. You should think how many edges K_n has and subtract the number of edges P_n has.

(C) $\frac{(n-1)(n-2)}{2}$
Correct. We have that K_n has $\frac{n(n-1)}{2}$ edges, and P_n has $n-1$ edges. Thus the complement of P_n has $\frac{n(n-1)}{2} - (n-1) = \frac{n(n-1)-2(n-1)}{2} = \frac{(n-2)(n-1)}{2}$

More Information

You can check your answer by looking at the complement of P_4 below:

P_4 : 

We can see that the complement of P_4 has $\frac{(4-1)(4-2)}{2} = \frac{6}{2} = 3$ edges.

ACTIVITY: Cliques and independent sets in cycles

WE DO NOT KEEP THIS IN VIDEO WE CAN USE IN HOMEWORKS SOME OF ITS STUFF

CALCULATING THE SIZE OF a maximum clique and the size of a maximum independent set in C_n (cases for n, use floor notation for odd cycles).