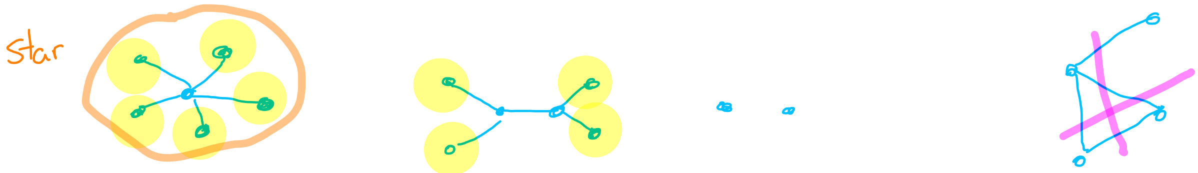
Nadia Lafrenière

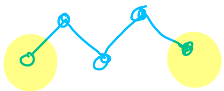
Trees

04/20/2020

We introduce the notion of trees, a very important type of graph. Over the next week or two, we will study the properties of trees and forests.

Definition
A graph with no cycle is acyclic.
An acyclic graph is a forest; a connected forest is a tree.
A leaf is a vertex of degree 1 in a tree.





Every component here (except the one with a purple X on it) is a tree, and the whole thing (without the one with the X) is a forest.

Leaves are highlighted. A star is the graph in which there is one vertex adjacent to every other.

Caveat: As graphs, trees don't need to have one specific root. We can always distinguish one root, but it is not needed. We will go back to this subject later.

Lemma
Every tree with at least two vertices has at least two leaves.
Deleting a leaf from an n-vertex tree produces a tree with $n-1$ vertices.

Proof
A tree is always connected so there is a path p between any two vertices $\{u,v\}$. Since there is no cycle, that path can only be extended finitely many times without returning to a previously visited vertex. The last time it can be extended in one direction, that vertex is a leaf, as there is no cycle.

When one deletes a leaf u from a tree, it does not disconnect it, since there is no path going through that vertex (not as an endpoint), i.e. for v,w in the graph, there is no path passing through u from v to w.

One consequence of that lemma is that we can build every tree with at least two vertices by "adding leaves". We will discuss that topic on Wednesday.

The following theorem gives multiple characterizations of trees:

Theorem

Let G be a graph with n vertices ($n \geq 1$). The following statements are equivalent:

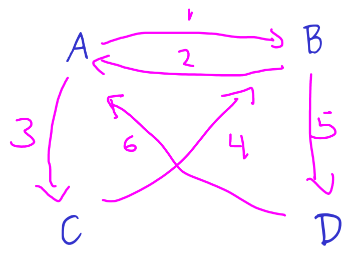
(A) G is connected and has no cycles.

(B) G is connected and has $n-1$ edges.

(C) G has no cycles and $n-1$ edges.

(D) G has no loop and has, for each pair of vertices $\{u,v\}$, exactly one uv-path.

The proof of such a statement is sort of a Hamiltonian closed walk within the complete digraph with vertices A, B, C and D:



Proof

(1 A=>B) We need to prove that if G is connected and has no cycle, it has $n-1$ edges.

By theorem from 13/04, it must have at least $n-1$ edges for it to be connected. To prove there is at most $n-1$ edges, we prove by induction on n (the number of vertices) that a graph with n edges has a cycle (that is a proof by contradiction):

If $n=1$, the edge is a loop and that is a cycle.

Assume a graph with $n=k$ vertices and k edges has a cycle.

We need to prove that a connected graph with $k+1$ vertices and $k+1$ edges has a cycle. If there is a leaf, remove it and delete the

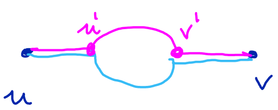
incident edge; we apply induction hypothesis to prove there is a cycle in the rest of the graph.
If there is no leaf, then the lemma from page 1 proves the graph is not a tree.
(2 B=>A) We need to show that if G is connected and has $n-1$ edges, it has no cycle. We prove the contraposition: if G is connected and has a cycle, there is more than $n-1$ edges.
Since G has a cycle, there is at least an edge that is not a cut-edge (by the theorem from 4/6). Deleting that edge would mean the graph has one fewer edge and is connected, which means, by theorem from 4/13, that the graph with one fewer edge has at least $n-1$ edges. So the original graph has at least n edges.

Until now, we proved A<=>B. They are equivalent, so we can use them together from now on.

(3 A=>C) Since A<=>B, we know that G is connected, has no cycle and has $n-1$ edges, which already proves C).
(4 C=>B) We want to show that if a graph has $n-1$ edges and no cycles, it is connected. We look at each of the k components.
In the component i $(1 \leq i \leq k)$, assume there are $n(i)$ vertices.
Since the component is connected and has no cycle, it has $n(i)-1$ edges (by A=>B). Hence, the total number of edges is $\Sigma(n(i)-1) = \Sigma(n(i)) - k = n-k$. However, the hypothesis of C is that the graph has $n-1$ edges. So there is exactly one component, and the graph is connected.

Now, A, B and C are equivalent. That means that two characteristics among connectedness, no cycles and $n-1$ edges are sufficient to show a graph is a tree.

(5 B=>D) Since B=>A, the graph has no cycle; in particular, it has no loop. It is also connected, so there is a path p between any pair of two vertices $\{u,v\}$. To show uniqueness of that path, we use the hypothesis that there is no cycle, and contradiction: Assume there exist 2 paths p and q between u to v.



Let u' be the first vertex in p and q whose next edges differ, and let v' be the next vertices that appear both in p and q. Then, the part of p between u' and v' and the part of q between u' and v' are paths with no common vertices that have the same endpoints; gluing them together creates a cycle.

(6 D=>A) Since there is a path between every pair of vertices, the graph is connected. The uniqueness of the path means there is no cycles, proving A.

Corollary

a) Every edge of a tree is a cut-edge. (by A)

b) Adding one edge to a tree forms exactly one cycle (corollary of A=>B).

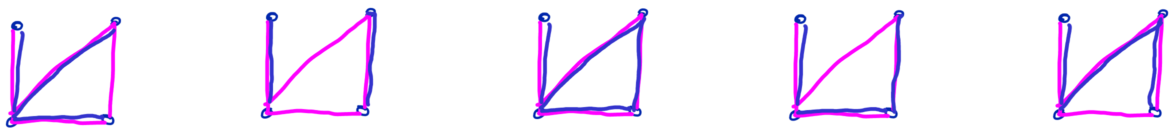
Spanning trees

Let $G=(V,E)$ be a graph.

A graph is a spanning subgraph of G if it has vertex set V.

A spanning tree is a spanning subgraph that is a tree.

Example



In blue, five spanning subgraphs of the graph in pink. Only the first, fourth and fifth ones are spanning trees.

Theorem

Every connected graph has a spanning tree.

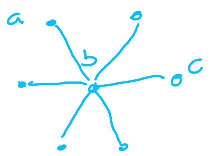
Proof

Every connected graph has a connected spanning subgraph. To remove the cycles from it, delete one after the edges that are in cycles. Once there are 1 edge fewer than vertices, the graph will be a tree.

Distance in trees and graphs

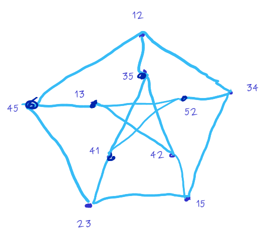
If G has a uv-path, the distance between u and v, noted $d(u,v)$, is the least length of a uv-path. If G has no such path, then $d(u,v)=\infty$.
The diameter of G is $\mathrm{diam}(G)$ is the maximum distance between two vertices.
The eccentricity of a vertex u is the distance to the furthest vertex.
The radius is the minimal eccentricity.

Example



Distance: $d(a,b)=d(b,c)=1$, $d(a,c)=2$
Diameter 2. A star always has radius 1, since the central vertex has eccentricity 1. The diameter, for all graph, is the maximal eccentricity.

Example

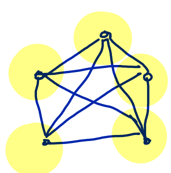


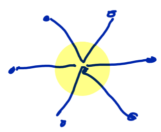
The Petersen graph has radius and diameter 2.
Recall there is an edge between two vertices if they represent disjoint 2-sets of $\{1,2,3,4,5\}$.
If two vertices ij and jk and not adjacent, they must share an element (as sets). Then, lm is disjoint from ij and jk, so $ij - lm - jk$ is a path of length 2.
Here, $\{i,j,k,l,m\}$ represents $\{1,2,3,4,5\}$.

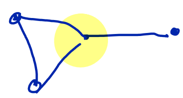
Theorem
If G is a simple graph, $\mathrm{diam}(G)\geq 3 \Rightarrow \mathrm{diam}(\bar{G})\leq 3$.

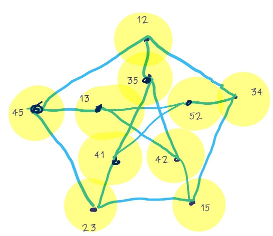
Proof: Read and understand as homework. In the book, that is Theorem 2.1.11, p.71.

The center of a graph is the induced subgraph with vertices of minimum eccentricity.





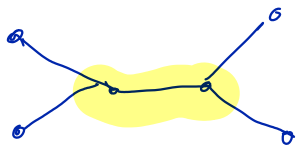




The Petersen graph and complete graph have center the whole graph. The star has, as center, only the central vertex.

Theorem (Jordan, 1869)
The center of a tree is a vertex or an edge.

That means it cannot be a set of vertices, whenever the graph is a tree. Of course, the examples above show it is not true for graphs in general.



Reference: Douglas B. West. Introduction to graph theory, 2nd edition, 2001. Section 2.1