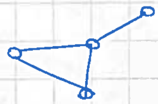
Lecture Plan Monday April 4: Trees

Definition: (page 641) A loop-free, undirected graph $G=(V,E)$ is called a tree if it is connected and has no cycles.

Example:

tree



not a tree, since it has a cycle

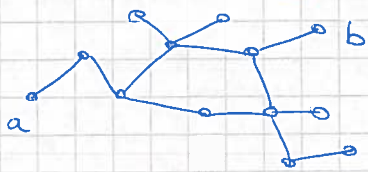
not a tree, since it is not connected
this graph can be seen as a collection of trees, since its connected components are trees

Definition: A loop-free undirected graph without cycles is called a forest. (Its connected components will be trees.)

Definition: If $G=(V,E)$ is a connected undirected graph, a spanning tree for G is a spanning subgraph which is a tree. That is: a subgraph containing all vertices of G, which is connected and has no cycles.

Theorem 1: (page 642) If a,b are distinct vertices in a tree $T=(V,E)$, then there is a unique path in T connecting a and b.

Proof: Since T is connected, there is at least one path in T between a and b. If there were multiple, different paths from a to b in T, part of these paths would form a cycle:



(example)

But since T is a tree, it cannot have cycles. Hence the path between a and b must be unique.

Theorem 2: (page 642) If $G=(V,E)$ is an undirected graph, then G is connected if and only if G has a spanning tree.

Proof: If G has a spanning tree T, then for each $a,b\in V$ with $a\neq b$, there is a path in T between a and b since T is connected. This path is also a path in G (since T is a subgraph of G). So G is connected.
Now assume that G is connected.
If G is a tree, we're done.
If G is not a tree, then G has a cycle or G has some loops. Remove all loops from G and call this new graph G_1. This G_1 must still be connected.
If G_1 is a tree, we're done.
If G_1 is not a tree, it must have a cycle. Pick any edge e_1 in this cycle and set $G_2=G_1-e_1$. Then G_2 is still connected.

If G_2 is a tree, we're done.
If not, G_2 must have a cycle. Pick an edge e_2 in this cycle and set $G_3 = G_2 - e_2$.
Then G_3 is still connected.
If G_3 is a tree, we're done.
If not, ----.

go on like this, breaking up cycle after cycle until none are left and we have a spanning tree. (We cannot keep breaking up cycles for forever, since our graph has only a finite number of edges).

Theorem 3: If $T=(V,E)$ is a tree, then $|V| = |E|+1$

page 642

Proof: We proof the result by the stronger induction version, on the number of edges.
Basic step: If $|E|=0$ and $T=(V,E)$ is a tree, then T can have only one vertex, since T must be connected. (T cannot have zero vertices, since any graph has at least one vertex).
Hence $|V|=1$ and so $|V|=|E|+1$.
Now let $k \in \mathbb{Z}_{\geq 0}$ and assume as an induction hypothesis that for all trees with at most k edges, the formula $|V|=|E|+1$ is true.
Now let $T=(V,E)$ be a tree with $k+1$ edges. Notice that $k+1 \geq 0+1=1$, so we can pick an edge $e \in E$.
Then $T-e$ has two connected components, $T_1=(V_1,E_1)$ and $T_2=(V_2,E_2)$ which are both trees with at most k edges. So by the induction hypothesis:
$|V_1|=|E_1|+1$ and $|V_2|=|E_2|+1$.
Now $|V|=|V_1|+|V_2| = |E_1|+|E_2|+2 = (|E_1|+|E_2|+1)+1 = |E|+1$.
This finishes our proof by induction.

Theorem 4: If $T=(V,E)$ is a tree, and $|V| \geq 2$, then T has at least 2 pendant vertices.

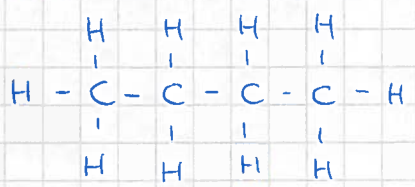
Proof: Write $|V|=n$. By Theorem 3, we know that $|E|=n-1$.
By Theorem 11.2 (from the chapter on graphs): $2(n-1) = \sum_{v \in V} \deg(v)$

Since T is connected and $|V| \geq 2$, $\deg(v) \geq 1$ $\forall v \in V$.
If $\deg(v)=1$, then v is a pendant vertex. (and all pendant vertices have degree 1)
If v is not a pendant vertex, then $\deg(v) \geq 2$.
So if T has k pendant vertices, and $n-k$ other vertices, we see:

$$2(n-1) = \sum_{v \in V} \deg(v) \geq k \cdot 1 + (n-k) \cdot 2 = 2n-k$$

Hence $2n-2 \geq 2n-k$, so $-2 \geq -k$ and hence $k \geq 2$.

Example: Molecules can often (though not always) be regarded as trees.



Here, a hydrogen atom must have degree 1, so it will always be a pendant vertex.
A carbon atom must have degree 4.

If you build a molecule with n carbon atoms, no double bonds (so no $C=C$) and no cycles, you can calculate how many H atoms you need.

If we have n C-atoms and k H-atoms, no cycles, and no double bonds, this means we have a tree with n+k vertices.
There are n vertices of degree 4
and k vertices of degree 1.

$$\text{Now } \underbrace{2\cdot|E|}_{\substack{\text{Thm 12.3} \\ 2(n+k-1)}} \overset{\text{Thm 11.2}}{=} \underbrace{\sum_{v\in V} \deg(v)}_{n\cdot 4+k\cdot 1}$$

So $2n+2k-2 = 4n+k$, so $k-2=2n$ and $k=2n+2$.
So our molecule needs $2n+2$ H-atoms.

__Theorem 5:__ If $G=(V,E)$ is an undirected loop-free graph, the following are equivalent:

a) G is a tree

b) G is connected, but removing any edge will disconnect G into two subgraphs which are trees.

c) G has no cycles and $|V| = |E|+1$

d) G is connected and $|V|=|E|+1$.

e) G has no cycles, and if $a,b\in V$ with $\{a,b\}\notin E$, then adding edge $\{a,b\}$ to G creates exactly one cycle.

Proof: (a) ⇒ (b):
We assume that G is a tree. By definition, G is connected. Let $e=\{a,b\}$ be any edge of G. If $G-\{a,b\}$ is still connected, then it has a path from a to b, and this will be a cycle in G when we add e to it. But as G is a tree, it has no cycles.
So $G-\{a,b\}$ cannot be connected. It has two components: the component with a and the component of things attached to b. Both these components are connected and have no cycles (since G has no cycles) or loops, so they must be trees.

(b) ⇒ (c).
We assume that G satisfies (b). If G had a cycle, removing an edge from this cycle would not disconnect the graph, which contradicts (b). So G cannot have a cycle. We now know that G is a loop-free undirected connected graph without cycles, so G is a tree. By Theorem 12.3: $|V|=|E|+1$.
This proves that (b) implies (c).

(c) ⇒ (d):
We assume that (c) holds. To prove (d), we must show that G is connected.
Let $\kappa(G)=r$ and let $G_1, G_2, \dots, G_r$ be the connected components of G.
Each G_i is a connected, loop-free undirected graph without cycles, hence a tree. If $G_i=(V_i,E_i)$, then by Theorem 12.3: $|V_i|=|E_i|+1$.

$$\text{Now}\quad |V| = \sum_{i=1}^{r} |V_i| = \sum_{i=1}^{r}(|E_i|+1) = \left(\sum_{i=1}^{r}|E_i|\right) + r = |E|+r.$$

But we know from (c) that $|V|=|E|+1$, so $r=1$. Hence G is connected.
So (c) implies (d).

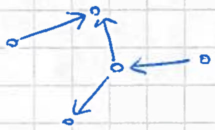
(d) $\Rightarrow$ (e):
We assume that (d) holds. Since G is connected, it has a spanning tree T by Theorem 12.2. By Theorem 12.3, T must have $|V|-1$ edges. As $|V| = |E| + 1$, T must now have all the edges of G, so $T = G$.
Hence G is a tree, so G has no cycles.
If $\{a,b\} \notin E$ for $a,b \in V$, then adding $\{a,b\}$ to G will give a graph $G' = (V, E')$. It is a connected, undirected graph and $|V| = |E|+1 = |E'|$ so G' is not a tree. Therefore, it must have a cycle.
Can G' have multiple cycles? Any cycle created in G' must use edge $\{a,b\}$. It must therefore consist of edge $\{a,b\}$ and a path between a and b in G. Since G is a tree, Theorem 12.1 shows that there is only one such path. So G' has exactly one cycle.

(e) $\Rightarrow$ (a):
We assume that (e) holds. To prove (a), we must show that G is connected. But if G were not connected, we could add an edge to G that connects two components and hence does not create a cycle. This contradicts (e), so G must be connected.

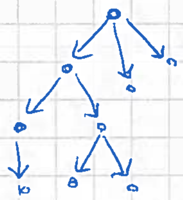
Definition: If G is a digraph, then G is called a directed tree if and only if its associated undirected graph is a tree.

A directed tree is called a rooted tree if there is a unique vertex r (called the root) for which $id(r) = 0$ (so no edges are incident to r) and for all other vertices v we have $id(v) = 1$.

Example:



Directed tree which is not rooted

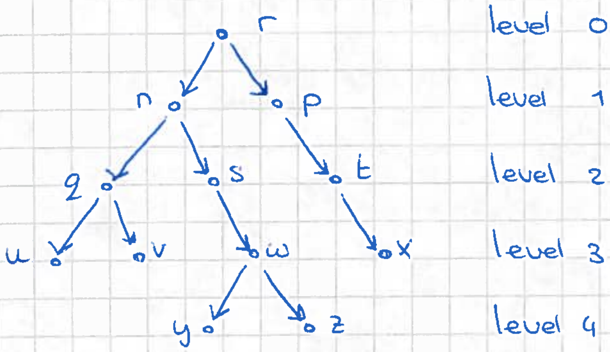


Rooted tree

Conventions: We draw rooted trees with the root on top and the arrows pointing down. That way, we can also leave out the arrows, since we know their direction.

Leaf = a vertex with out-degree 0 = a terminal vertex
Branch node = a vertex which is not a leaf = an internal vertex

We can order the vertices of a rooted tree into levels, depending on how many edges away they are from the root



vertices at level k have a path of length k from the root to the vertex.

s is a child of n, x is a child of t

n is the parent of s, t is the parent of x

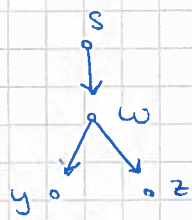
s, w, y and z are descendants of n (and also of r)

r, n and s are ancestors of w (and also of y and z)

u and v are siblings since they have the same parent

The subtree at a vertex is the subgraph induced by this vertex and all of its descendants.

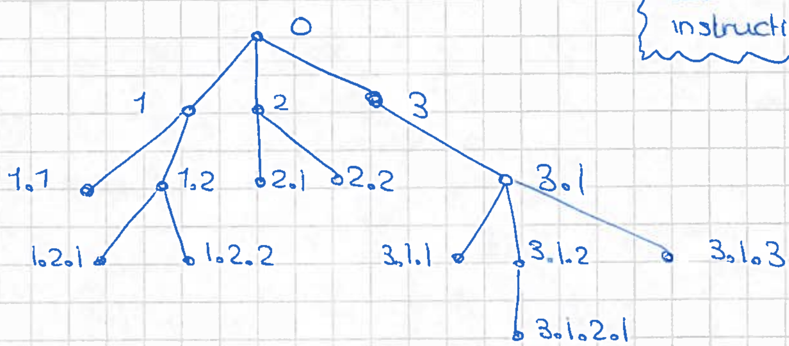
Subtree at s:



So in this subtree, s is now the root.

Definition: The universal address system

see book for step-by-step instruction



We can now order these labeled vertices as in a dictionary: (or a table of contents in a book)
the first digit is most important, then the next one, and so on:

0, 1, 1.1, 1.2, 1.2.1, 1.2.2, 2, 2.1, 2.2, 3, 3.1, 3.1.1, 3.1.2, 3.1.2.1, 3.1.3