5.1. **binary trees.** A binary tree is a rooted planar tree (one vertex is labelled as the root and the tree is embedded in the plane with root at the top) in which every node has two daughters, a left daughter and a right daughter.

A tree is a connected graph with $n$ vertices and $n-1$ edges. So, I stated with the definition of graph and connectedness. A graph is a set of points connected by edges. After talking about it for a while we decided that we shouldn't use the word "connected" in the definition since we are going to define that concept.

5.1.1. *graphs.*

**Definition 5.2.** A *graph* $\Gamma$ consists of two sets and one function.

(1) $\Gamma_0$ is the set of *vertices* of $\Gamma$.
(2) $\Gamma_1$ is the set of *edges* of $\Gamma$.
(3) Each edge $e$ has two endpoints $a, b$ which are vertices. Mathematically this is given by a function:

$$h(e) = \{a, b\}$$

This is a function

$$h : \Gamma_1 \to \mathcal{P}_2(\Gamma_0)$$

from the set of edges of $\Gamma$ to the set of all two-element subsets of $\Gamma_0$.

I insisted that "not connected" is an easier concept than "connected" since you don't need the concept of a "path" even though that is easy.

**Definition 5.3.** A graph $\Gamma$ is *not connected* if the set of vertices can be partitioned into a disjoint union of two nonempty sets:

$$\Gamma_0 = A \coprod B$$

so that there is no edge with one endpoint in $A$ and the other endpoint in $B$.

A graph is called *connected* if is it not not-connected.

5.1.2. *trees.*

**Definition 5.4.** A *tree* is a connected graph with $n$ vertices and $n-1$ edges.

The first problem is to show that trees do not have cycles.

**Definition 5.5.** A *cycle* is a sequence of vertices $v_1, v_2, \cdots, v_n$ $(n \geq 3)$ and edges $e_1, e_2, \cdots, e_n$ so that the endpoints of $e_i$ are $v_i$ and $v_{i+1}$ for $i = 1, 2, \cdots, n-1$ and $h(e_n) = \{v_n, v_1\}$.

To show that trees have no cycles we need the following lemma.

**Lemma 5.6.** *A connected graph with $n$ vertices has at least $n-1$ edges.*

We assume for today that this is true.

**Theorem 5.7.** *Trees do not have cycles.*

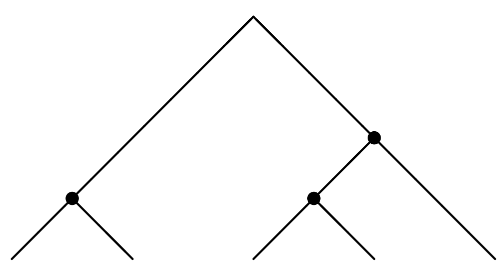
*Proof.* Suppose that we have a tree with $n$ vertices and $n-1$ edges. Suppose it has a cycle. Then we will get a contradiction.

If we remove one of the edges of the cycle, the graph will still be connected because the two endpoints of the edge are connected by a path (the rest of the cycle). This gives a connected graph with $n$ vertices and $n-2$ edges which is impossible. □

5.1.3. *binary trees.* Trees have nodes and leaves. A *leaf* is a vertex which is an endpoint of only one edge. The number of edges which are attached to a vertex $v$ (i.e., have $v$ as an endpoint) is called the *degree* of the vertex. Thus leaves have degree 1. "Degree" is a noun without a good adjective so we use the word "valence" which has adjective form "valent." Thus a leaf is *univalent* (instead of "one-degreed" which sounds stupid).

A *node* is a vertex of degree $\geq 2$. You can take any vertex of a tree and call it the *root.* Two vertices are called *adjacent* if they are the endpoints of one edge.

**Definition 5.8.** A *rooted binary tree* is a tree with a root which has degree 2 and all other nodes of degree 3. The tree should be *planar* which means it is embedded in the $xy$-plane with the root at the top and leaves at the bottom. Each trivalent node has three adjacent vertices, one above called the *parent* of the node and two below called the left and right *daughters.* The root has a left and right daughter but no parent.



We need to draw the picture in a special way. If the tree has $n$ nodes then it has $2n$ edges since each edge has a node at the top and each node has two edges below it. (This is a 2-1 correspondence.) This means it has $2n+1$ vertices. Since $n$ of these are nodes, our binary tree has $n+1$ leaves.

Place these leaves evenly spaced in a horizontal line. (Put them at the point $x = 0, 1, 2, \cdots, n$ on the $x$-axis.

Draw all lines with slope 1 or -1. (This puts the root at the point $(n/2, n/2)$ in the $xy$-plane. (See the drawing.)

**Theorem 5.9.** *The number of rooted binary trees with $n$ nodes is the Catalan number $C(n)$.*

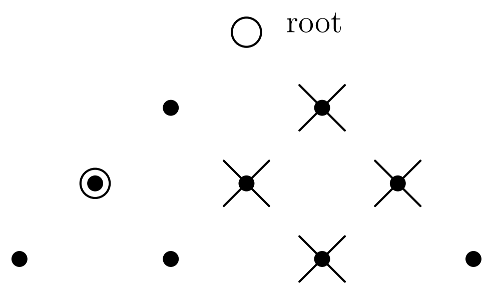
We will prove this next week. But for now, I want to assume this and try to get a bijection with the set of clusters.

5.1.4. *clusters.* We could define a cluster to be the set of points in the plane given by the trivalent nodes of a rooted binary tree. Since the root is always at the same place, I won't consider it to be a variable.

**Theorem 5.10.** *Each binary tree is determined by its set of nodes.*

*Proof.* Draw two lines with slope 1 and -1 from each node to the $x$-axis. This is the tree. □

Clusters are actually defined to be sets of $n$ points which are pairwise "compatible" which in our case means that they could be nodes in the same binary tree.




**Exercise 5.11.** : Show that, if there is a node at the circled position, there cannot be a node at any of the positions marked "X"

Liz was not satisfied with the geometric answer that I gave so I challenged the class to come up with a numerical answer. The vertices should be numbered as follows.

| | | | | | | |
|---|---|---|---|---|---|---|
| | | 24 | | 13 | | |
| | 34 | | 23 | | 12 | |
| 44 | | 33 | | 22 | | 11 |