Discrete Structures

Lecture 25: Trees

Susan L. Epstein

HUNTER

1

Last time

Graph theory has important real-world applications

- Connectivity
- Euler and Hamilton paths and circuits
- Some proofs rely on computers to address their myriad cases

Trees have important real-world applications

2

Today's outline

- Basic ideas and examples
- Properties of trees
- Applications of trees

The material on graphs and trees in CSCI 150 supersedes your text

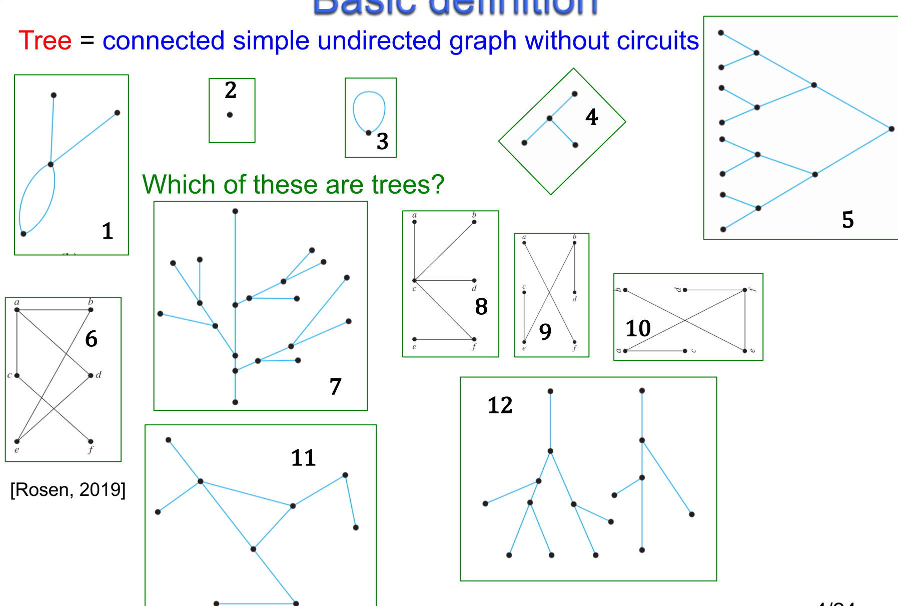
Fall 2023 CSCI 150 3/24

3

Basic definition

Tree = connected simple undirected graph without circuits

Which of these are trees?



[Rosen, 2019]

Fall 2023 CSCI 150 4/24

4

More on trees

Equivalent definitions:

Any simple undirected graph that

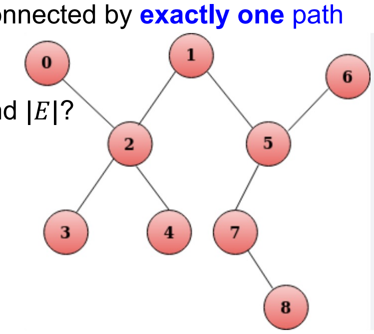
- Is connected and acyclic (contains no cycles/circuits)
- Is acyclic and would contain a simple cycle if any edge were added to it
- Is connected and would become disconnected when any edge is removed from it
- Where any pair of distinct vertices is connected by **exactly one** path

For a tree $G = (V, E)$

- What is the relationship between $|V|$ and $|E|$?

  $|V| - |E| = 1$
- What is the total degree of a tree?

  $2(|V| - 1)$



5

Vertices and edges

Theorem: $P(n)$: a tree with n vertices has $n - 1$ edges.

Proof by mathematical induction:

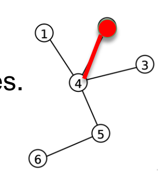
We must show that $P(n)$ is true for all $n \geq 1$.

Basis: $P(1)$: a tree with 1 vertex has $1 - 1$ edges. Such a tree has 0 edges, so $P(1)$ is true.

Inductive step: Assume for some k that $P(k)$ is true.

By substitution $P(k + 1)$: a tree with $\mathrm{k} + 1$ vertex has k edges.

We must show that $P(k + 1)$ is true.



Let T be a tree with $|V| = k + 1$.

Because T is finite, there must be some $v \in V$ that is a leaf.

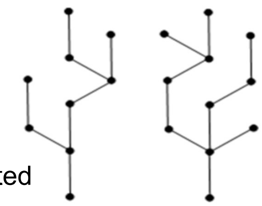
Let T' be the tree formed when both v and the edge from its parent are deleted from T.

By the inductive step, T' has $k - 1$ vertices, so T must have $k - 1 + 1$ vertices and $P(k + 1)$ is true.

Since we have proved the basis step and the inductive step, the theorem is true. **QED**

6

Forests



Equivalent definitions for forest

- Undirected graph in which any two vertices are connected by **at most** one path
- Unconnected acyclic graph all of whose connected components are trees
- A disjoint union of trees $G = \bigcup_i T_i\,(V_i, E_i)$ where $V_i \cap V_j = \emptyset$ for $i \neq j$

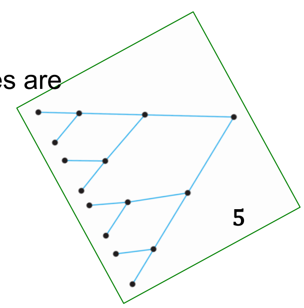
- Special cases
  - An empty forest
  - A single tree
  - An edgeless graph

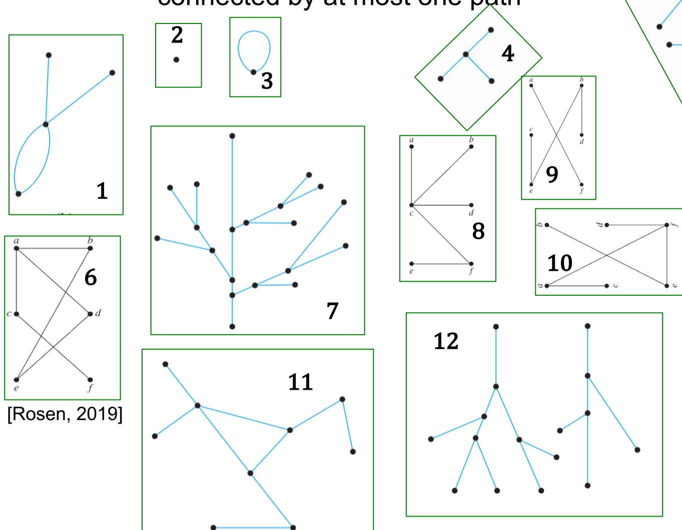
- How many trees in a forest $G = \bigcup_i T_i\,(V_i, E_i)$?
  In each tree $T_i(V_i, E_i)$, $|V_i| - |E_i| = 1$ so
  G has $|V| - |E|$ trees

7

Find the forests

Forest = undirected graph in which any two vertices are connected by at most one path





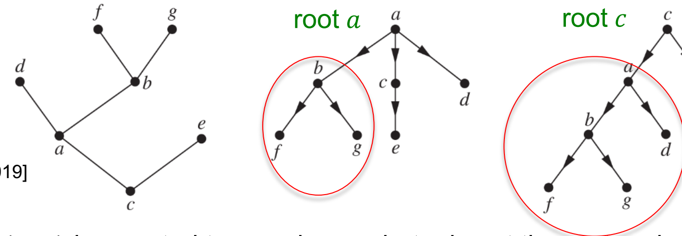
[Rosen, 2019]

Which of these are forests?

8

Rooted tree

- Rooted tree = tree with exactly 1 designated vertex (its root) and every edge is directed away from that root



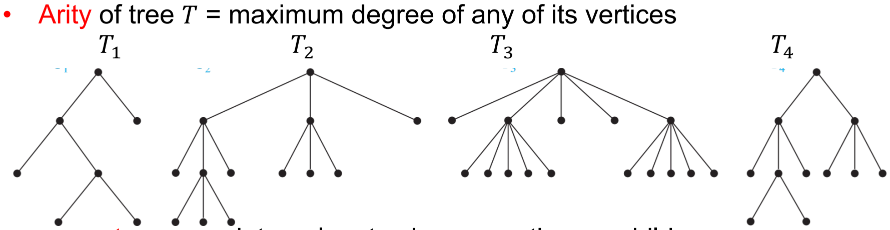
[Rosen, 2019]

- If $T = (V, E)$ is a rooted tree and $v \in V$ but v is not the root and $u \in V$ has an edge directly to v, then u is the parent of v and v is the child of u.
- Vertices with the same parents are siblings
- Ancestors of v are its parent and those on the path from the root to v
- Descendants of v are vertices that have v as an ancestor
- Leaf = node with no children
- Internal node = vertex that is neither the root nor a leaf
- Subtree rooted at vertex v = v, its descendants, and the edges incident to them

9

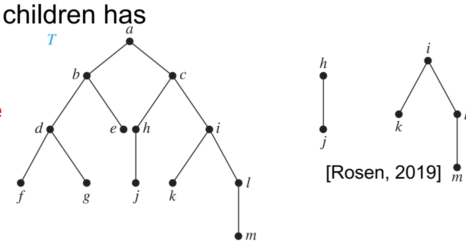
Arity of a tree

- Arity of tree T = maximum degree of any of its vertices



[Rosen, 2019]

- m-ary tree = no internal vertex has more than m children
- Binary tree = m-ary tree with $m = 2$
- Full m-ary tree = every non-leaf has exactly m children
- In a binary tree, a vertex with 2 children has
  - A left child and a right child
  - A left subtree rooted at its left child and a right subtree rooted at its right child



[Rosen, 2019]

10

Today's outline

- ✓ Basic ideas and examples
- Properties of trees
- Applications of trees

The material on graphs and trees in CSCI 150 supersedes your text

Another characterization of a tree

Theorem: An undirected graph is a tree iff there is a unique simple path between every pair of its vertices.

Proof:

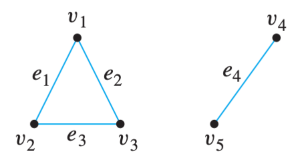
Part 1: Let $T = (V, E)$ be a tree and $x, y \in V$ be any pair of of distinct vertices in T. By definition of tree, T is connected and therefore has a path between x and y. Assume there are 2 distinct paths $x = pq \ldots rs = y$ and $x = ab \ldots cd = y$. We will show that this assumption leads to a contradiction. The path $x = pq \ldots ryc \ldots ba = x$ begins and ends at x and is a circuit. This contradicts that T is a tree. Because the assumption led to a contradiction, the path between x and y must be unique.

Part 2: Let $T = (V, E)$ be a graph with a unique simple path between every pair of its vertices. Then by definition, T is connected. Assume T has a circuit $ab \ldots xc \ldots yd \ldots a$. that includes some pair of vertices x and y. We will show that this assumption leads to a contradiction. The paths $xc \ldots y$ and $yd \ldots ab \ldots x$ are simple and distinct. Because the assumption led to a contradiction, T has no circuits and by definition is a tree. **QED**

Another characterization of a tree

- Consider a connected graph where $|V| = n$.
- How can you use $|E|$ to determine whether T is a tree?
- Theorem: For any $n \in \boldsymbol{Z}^{+}$, if $G = (V, E)$ is a connected graph with n vertices and $n - 1$ edges, G is a tree. (Proof is in your text.)

- What if $G = (V, E)$ is not connected; could it still be a tree?
  5 vertices and 4 edges?



13

Today's outline

- ✓ Basic ideas and examples
- ✓ Properties of trees
- Applications of trees

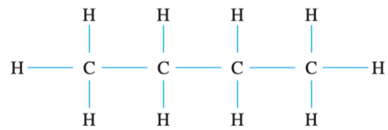
The material on graphs and trees in CSCI 150 supersedes your text

14

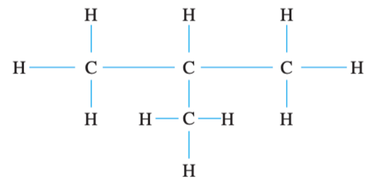
Trees as world models

Chemical molecule structures

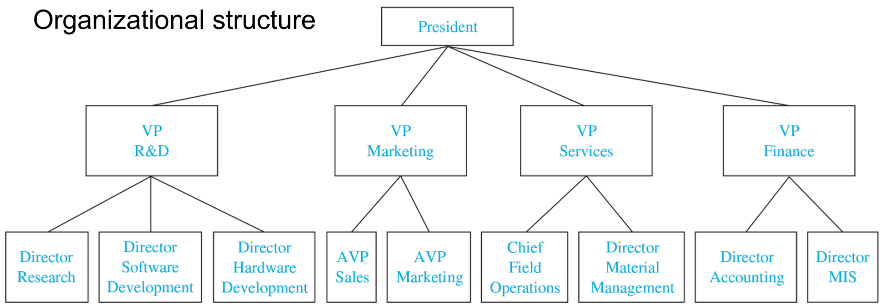
butane



isobutane



Organizational structure



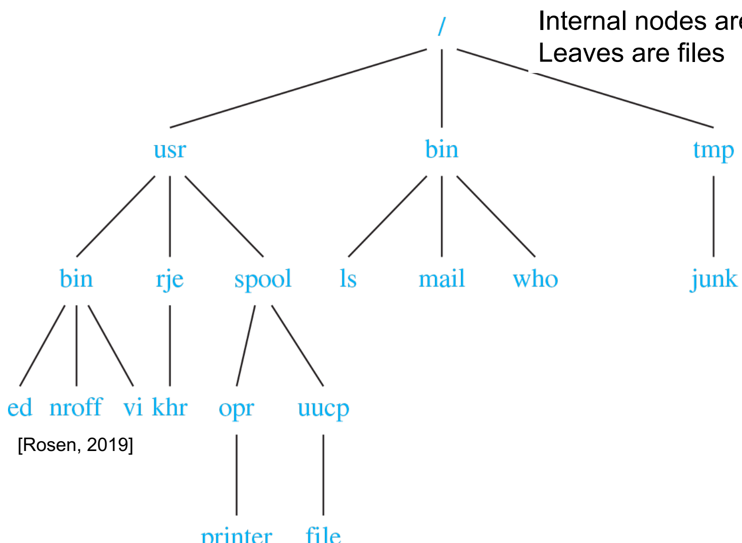
[Rosen, 2019]

Fall 2023 CSCI 150 15/24

15

Computer file system as a tree

Root is root directory
Internal nodes are directories
Leaves are files

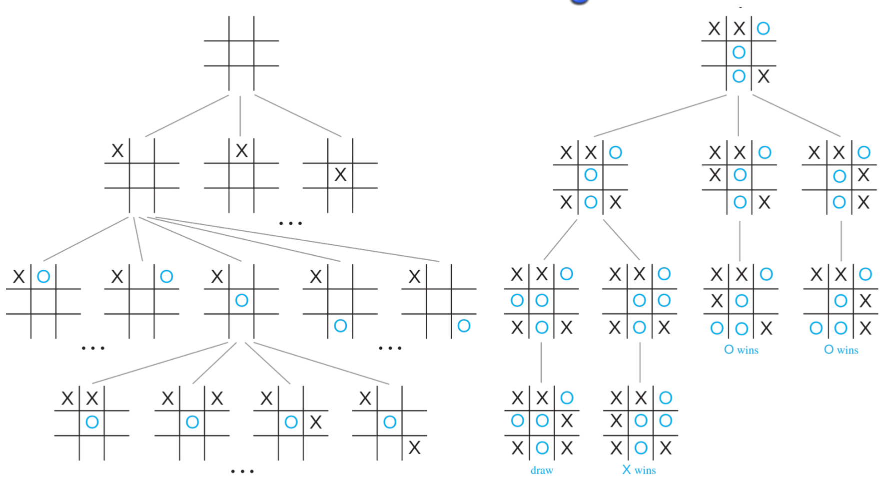


[Rosen, 2019]

Fall 2023 CSCI 150 16/24

16

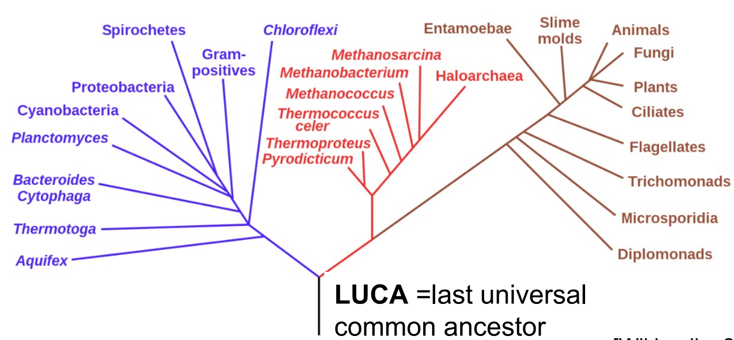
Parts of the tic-tac-toe game tree



17

The tree of life

A tree for rRNA genes with the 3 life domains: Bacteria, Archaea, and Eukaryota **Bacteria** **Archaea** **Eukaryota**

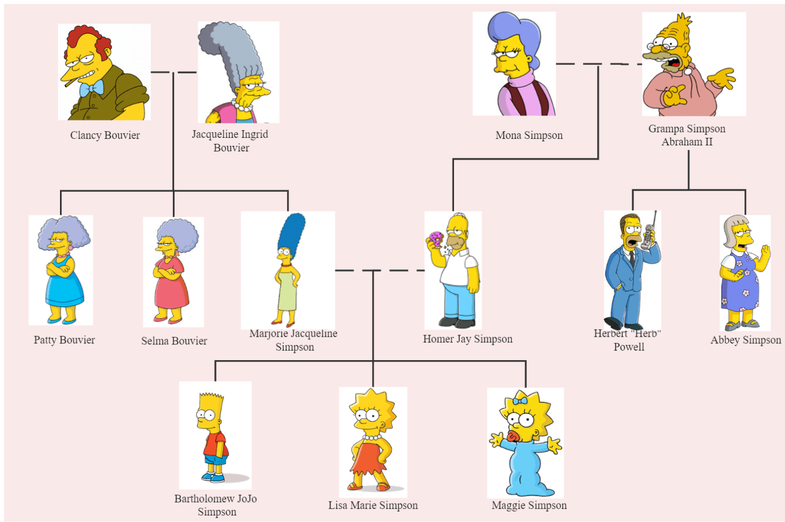


[Wikipedia, 2009]

Since 2009, genetic findings have incorporated many new species and alternative representations are not tree-like in the sense we use here

18

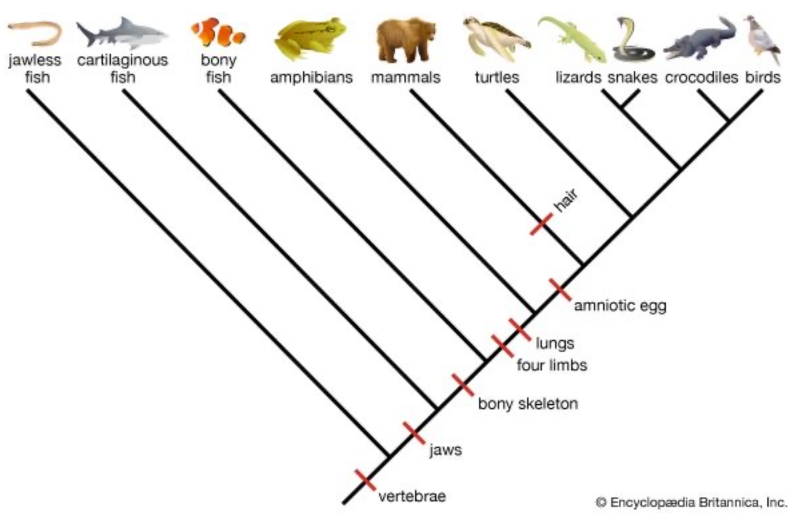
Genealogical tree



Fall 2023 CSCI 150 19/24

19

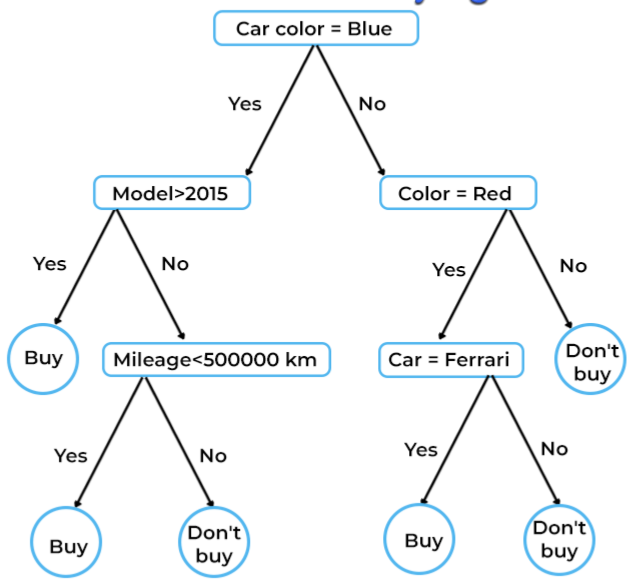
Vertebrate phylogenetic tree



Fall 2023 CSCI 150 20/24

20

Decision tree for buying a car



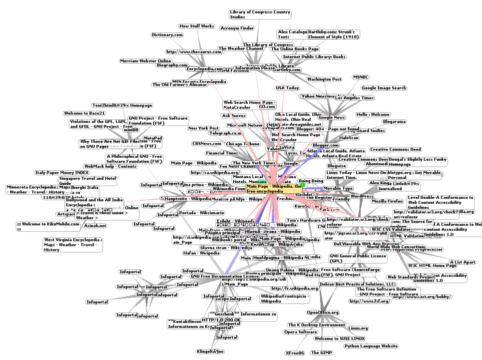
And many more…

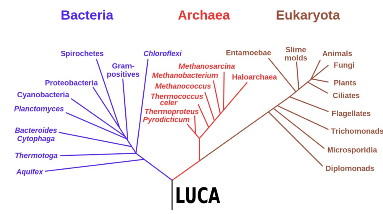
- Parse trees address grammatical sentence structure and support natural language processing (NLP)
- Balanced trees, B-trees, and B+-trees support efficient storage, retrieval and sorting
- Spanning trees support minimal path finding
- Trees support the design of sorting algorithms and universal address systems
- Trees support Huffman coding for optimal coding of a set of symbols

Working with trees

- Many operations that are difficult on graphs are far easier on trees

search traversal





- Most of those algorithms use recursion

Fall 2023 CSCI 150 23/24

23

What you should know

★ **Trees have important real-world applications**

- Trees and forests are subclasses of graphs with special properties

Any questions?

Next up: *Ethics*

Time to finish up that Opening sheet!

Problem set 25,26 is due on Wednesday, December 13 at 11PM

Fall 2023 CSCI 150 24/24

24