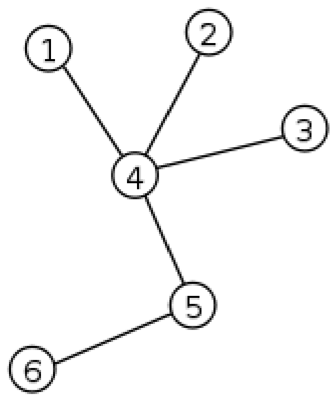
Tree (graph theory)

In graph theory, a **tree** is an undirected graph in which any two vertices are connected by *exactly one* path, or equivalently a connected acyclic undirected graph.[1] A **forest** is an undirected graph in which any two vertices are connected by *at most one* path, or equivalently an acyclic undirected graph, or equivalently a disjoint union of trees.[2]

Trees	
 A labeled tree with 6 vertices and 5 edges.	
Vertices	v
Edges	$v - 1$
Chromatic number	2 if $v > 1$
Table of graphs and parameters	

A **polytree**[3] (or **directed tree**[4] or **oriented tree**[5][6] or **singly connected network**[7]) is a directed acyclic graph (DAG) whose underlying undirected graph is a tree. A **polyforest** (or **directed forest** or **oriented forest**) is a directed acyclic graph whose underlying undirected graph is a forest.

The various kinds of data structures referred to as trees in computer science have underlying graphs that are trees in graph theory, although such data structures are generally **rooted trees**. A rooted tree may be directed, called a **directed rooted tree**,[8][9] either making all its edges point away from the root—in which case it is called an **arborescence**[4][10] or **out-tree**[11][12]—or making all its edges point towards the root—in which case it is called an **anti-arborescence**[13] or **in-tree**.[11][14] A rooted tree itself has been defined by some authors as a directed graph.[15][16][17] A **rooted forest** is a disjoint union of rooted trees. A rooted forest may be directed, called a **directed rooted forest**, either making all its edges point away from the root in each rooted tree—in which case it is called a

branching or **out-forest**—or making all its edges point towards the root in each rooted tree—in which case it is called an **anti-branching** or **in-forest**.

The term "tree" was coined in 1857 by the British mathematician Arthur Cayley.[18]

Definitions

Tree

A *tree* is an undirected graph G that satisfies any of the following equivalent conditions:

- G is connected and acyclic (contains no cycles).
- G is acyclic, and a simple cycle is formed if any edge is added to G.
- G is connected, but would become disconnected if any single edge is removed from G.
- G is connected and the 3-vertex complete graph K_3 is not a minor of G.
- Any two vertices in G can be connected by a unique simple path.

If G has finitely many vertices, say n of them, then the above statements are also equivalent to any of the following conditions:

- G is connected and has $n - 1$ edges.
- G is connected, and every subgraph of G includes at least one vertex with zero or one incident edges. (That is, G is connected and 1-degenerate.)
- G has no simple cycles and has $n - 1$ edges.

As elsewhere in graph theory, the order-zero graph (graph with no vertices) is generally not considered to be a tree: while it is vacuously connected as a graph (any two vertices can be connected by a path), it is not 0-connected (or even (−1)-connected) in algebraic topology, unlike non-empty trees, and violates the "one more vertex than edges" relation. It may, however, be considered as a forest consisting of zero trees.

An **internal vertex** (or **inner vertex** or **branch vertex**) is a vertex of degree at least 2. Similarly, an **external vertex** (or *outer vertex*, *terminal vertex* or *leaf*) is a vertex of degree 1.

An *irreducible tree* (or *series-reduced tree*) is a tree in which there is no vertex of degree 2 (enumerated at sequence A000014 in the OEIS).[19]

Forest

A *forest* is an undirected graph in which any two vertices are connected by at most one path. Equivalently, a forest is an undirected acyclic graph, all of whose connected components are trees; in other words, the graph consists of a disjoint union of trees. As special cases, the order-zero graph (a forest consisting of zero trees), a single tree, and an edgeless graph, are examples of forests. Since for every tree $V - E = 1$, we can easily count the number of trees that are within a forest by subtracting the difference between total vertices and total edges.
$TV - TE$ = number of trees in a forest.

Polytree

A *polytree*[3] (or *directed tree*[4] or *oriented tree*[5][6] or *singly connected network*[7]) is a directed acyclic graph (DAG) whose underlying undirected graph is a tree. In other words, if we replace its directed edges with undirected edges, we obtain an undirected graph that is both connected and acyclic.

Some authors restrict the phrase "directed tree" to the case where the edges are all directed towards a particular vertex, or all directed away from a particular vertex (see arborescence).

Polyforest

A *polyforest* (or *directed forest* or *oriented forest*) is a directed acyclic graph whose underlying undirected graph is a forest. In other words, if we replace its directed edges with undirected edges, we obtain an undirected graph that is acyclic.

Some authors restrict the phrase "directed forest" to the case where the edges of each connected component are all directed towards a particular vertex, or all directed away from a particular vertex (see branching).

Rooted tree

A *rooted tree* is a tree in which one vertex has been designated the *root*.[20] The edges of a rooted tree can be assigned a natural orientation, either *away from* or *towards* the root, in which case the structure becomes a *directed rooted tree*. When a directed rooted tree has an orientation away from the root, it is called an *arborescence*[4] or *out-tree*;[11] when it has an orientation towards the root, it is called an *anti-arborescence* or *in-tree*.[11] The *tree-order* is the partial ordering on the vertices of a tree with $u < v$ if and only if the unique path from the root to v passes through u. A rooted tree T which is a subgraph of some graph G is a normal tree if the ends of every T-path in G are comparable in this tree-order (Diestel 2005, p. 15). Rooted trees, often with additional structure such as ordering of the neighbors at each vertex, are a key data structure in computer science; see tree data structure.

In a context where trees are supposed to have a root, a tree without any designated root is called a *free tree*.

A *labeled tree* is a tree in which each vertex is given a unique label. The vertices of a labeled tree on n vertices are typically given the labels 1, 2, ..., n. A *recursive tree* is a labeled rooted tree where the vertex labels respect the tree order (i.e., if $u < v$ for two vertices u and v, then the label of u is smaller than the label of v).

In a rooted tree, the *parent* of a vertex v is the vertex connected to v on the path to the root; every vertex has a unique parent except the root which has no parent.[20] A *child* of a vertex v is a vertex of which v is the parent.[20] An *ascendant* of a vertex v is any vertex which is either the parent of v or is (recursively) the ascendant of the parent of v. A *descendant* of a vertex v is any vertex which is either the child of v or is (recursively) the descendant of any of the children of v. A *sibling* to a vertex v is any other vertex on the tree which has the same parent as v.[20] A *leaf* is a vertex with no children.[20] An *internal vertex* is a vertex that is not a leaf.[20]

The *height* of a vertex in a rooted tree is the length of the longest downward path to a leaf from that vertex. The *height* of the tree is the height of the root. The *depth* of a vertex is the length of the path to its root (*root path*). This is commonly needed in the manipulation of the various self-balancing trees, AVL trees in particular. The root has depth zero, leaves have height zero, and a tree with only a single vertex (hence both a root and leaf) has depth and height zero. Conventionally, an empty tree (a tree with no vertices, if such are allowed) has depth and height −1.

A *k-ary tree* is a rooted tree in which each vertex has at most k children.[21] 2-ary trees are often called *binary trees*, while 3-ary trees are sometimes called *ternary trees*.

Ordered tree

An *ordered tree* (or *plane tree*) is a rooted tree in which an ordering is specified for the children of each vertex.[20][22] This is called a "plane tree" because an ordering of the children is equivalent to an embedding of the tree in the plane, with the root at the top and the children of each vertex lower than that vertex. Given an embedding of a rooted tree in the plane, if one fixes a direction of children, say left to right, then an embedding gives an ordering of the children. Conversely, given an ordered tree, and conventionally drawing the root at the top, then the child vertices in an ordered tree can be drawn left-to-right, yielding an essentially unique planar embedding.

Properties

- Every tree is a bipartite graph. A graph is bipartite if and only if it contains no cycles of odd length. Since a tree contains no cycles at all, it is bipartite.
- Every tree is a median graph.
- Every tree with only countably many vertices is a planar graph.
- Every connected graph G admits a spanning tree, which is a tree that contains every vertex of G and whose edges are edges of G.
- Every connected graph with only countably many vertices admits a normal spanning tree (Diestel 2005, Prop. 8.2.4).
- There exist connected graphs with uncountably many vertices which do not admit a normal spanning tree (Diestel 2005, Prop. 8.5.2).
- Every finite tree with n vertices, with $n > 1$, has at least two terminal vertices (leaves). This minimal number of leaves is characteristic of path graphs; the maximal number, $n - 1$, is attained only by star graphs. The number of leaves is at least the maximum vertex degree.
- For any three vertices in a tree, the three paths between them have exactly one vertex in common (this vertex is called the *median* of these three vertices).
- Every tree has a center consisting of one vertex or two adjacent vertices. The center is the middle vertex or middle two vertices in every longest path. Similarly, every n-vertex tree has a centroid consisting of one vertex or two adjacent vertices. In the first case removal of the vertex splits the tree into subtrees of fewer than $n/2$ vertices. In the second case, removal of the edge between the two centroidal vertices splits the tree into two subtrees of exactly $n/2$ vertices.

Enumeration

Labeled trees

Cayley's formula states that there are n^{n-2} trees on n labeled vertices. A classic proof uses Prüfer sequences, which naturally show a stronger result: the number of trees with vertices 1, 2, ..., n of degrees d_1, d_2, ..., d_n respectively, is the multinomial coefficient

$$\binom{n-2}{d_1-1, d_2-1, \ldots, d_n-1}.$$

A more general problem is to count spanning trees in an undirected graph, which is addressed by the matrix tree theorem. (Cayley's formula is the special case of spanning trees in a complete graph.) The similar problem of counting all the subtrees regardless of size is #P-complete in the general case (Jerrum (1994)).

Unlabeled trees

Counting the number of unlabeled free trees is a harder problem. No closed formula for the number $t(n)$ of trees with n vertices up to graph isomorphism is known. The first few values of $t(n)$ are

1, 1, 1, 1, 2, 3, 6, 11, 23, 47, 106, 235, 551, 1301, 3159, … (sequence A000055 in the OEIS).

Otter (1948) proved the asymptotic estimate

$$t(n) \sim C\alpha^n n^{-5/2} \quad \text{as } n \to \infty,$$

with the values C and α known to be approximately 0.534949606... and 2.95576528565... (sequence A051491 in the OEIS), respectively. (Here, $f \sim g$ means that $\lim_{n \to \infty} f/g = 1$.) This is a consequence of his asymptotic estimate for the number $r(n)$ of unlabeled rooted trees with n vertices:

$$r(n) \sim D\alpha^n n^{-3/2} \quad \text{as } n \to \infty,$$

with D around 0.43992401257... and the same α as above (cf. Knuth (1997), chap. 2.3.4.4 and Flajolet & Sedgewick (2009), chap. VII.5, p. 475).

The first few values of $r(n)$ are[23]

1, 1, 2, 4, 9, 20, 48, 115, 286, 719, 1842, 4766, 12486, 32973, … (sequence A000081 in the OEIS)

Types of trees

- A *path graph* (or *linear graph*) consists of n vertices arranged in a line, so that vertices i and i+1 are connected by an edge for i=1,…,n−1.
- A *starlike tree* consists of a central vertex called *root* and several path graphs attached to it. More formally, a tree is starlike if it has exactly one vertex of degree greater than 2.
- A *star tree* is a tree which consists of a single internal vertex (and n−1 leaves). In other words, a star tree of order n is a tree of order n with as many leaves as possible.
- A *caterpillar tree* is a tree in which all vertices are within distance 1 of a central path subgraph.
- A *lobster tree* is a tree in which all vertices are within distance 2 of a central path subgraph.
- A *regular tree* of degree d is the infinite tree with d edges at each vertex. These arise as the Cayley graphs of free groups, and in the theory of Tits buildings.

See also

- Hypertree
- Tree structure
- Tree (data structure)
- Decision tree
- Pseudoforest
- Unrooted binary tree

Notes

1. Bender & Williamson 2010, p. 171.
2. Bender & Williamson 2010, p. 172.
3. See Dasgupta (1999).
4. Deo 1974, p. 206.
5. See Harary & Sumner (1980).
6. See Simion (1991).
7. See Kim & Pearl (1983).
8. Stanley Gill Williamson (1985). *Combinatorics for Computer Science*. Courier Dover Publications. p. 288. ISBN 978-0-486-42076-9.
9. Mehran Mesbahi; Magnus Egerstedt (2010). *Graph Theoretic Methods in Multiagent Networks*. Princeton University Press. p. 38. ISBN 978-1-4008-3535-5.
10. Ding-Zhu Du; Ker-I Ko; Xiaodong Hu (2011). *Design and Analysis of Approximation Algorithms*. Springer Science & Business Media. p. 108. ISBN 978-1-4614-1701-9.
11. Deo 1974, p. 207.
12. Jonathan L. Gross; Jay Yellen; Ping Zhang (2013). *Handbook of Graph Theory, Second Edition*. CRC Press. p. 116. ISBN 978-1-4398-8018-0.
13. Bernhard Korte; Jens Vygen (2012). *Combinatorial Optimization: Theory and Algorithms* (5th ed.). Springer Science & Business Media. p. 28. ISBN 978-3-642-24488-9.
14. Kurt Mehlhorn; Peter Sanders (2008). *Algorithms and Data Structures: The Basic Toolbox* (http://people.mpi-inf.mpg.de/~mehlhorn/ftp/Toolbox/Introduction.pdf) (PDF). Springer Science & Business Media. p. 52. ISBN 978-3-540-77978-0.

15. David Makinson (2012). *Sets, Logic and Maths for Computing*. Springer Science & Business Media. pp. 167–168. ISBN 978-1-4471-2499-3.

16. Kenneth Rosen (2011). *Discrete Mathematics and Its Applications, 7th edition*. McGraw-Hill Science. p. 747. ISBN 978-0-07-338309-5.

17. Alexander Schrijver (2003). *Combinatorial Optimization: Polyhedra and Efficiency*. Springer. p. 34. ISBN 3-540-44389-4.

18. Cayley (1857) "On the theory of the analytical forms called trees," (https://books.google.com/books?id=MlEEAAAAYAAJ&pg=PA172#v=onepage&q&f=false) *Philosophical Magazine*, 4th series, **13** : 172–176. However it should be mentioned that in 1847, K.G.C. von Staudt, in his book *Geometrie der Lage* (Nürnberg, (Germany): Bauer und Raspe, 1847), presented a proof of Euler's polyhedron theorem which relies on trees on pages 20–21 (https://books.google.com/books?id=MzQAAAAAQAAJ&pg=PA20#v=onepage&q&f=false) . Also in 1847, the German physicist Gustav Kirchhoff investigated electrical circuits and found a relation between the number (n) of wires/resistors (branches), the number (m) of junctions (vertices), and the number (μ) of loops (faces) in the circuit. He proved the relation via an argument relying on trees. See: Kirchhoff, G. R. (1847) "Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird" (https://books.google.com/books?id=gx4AAAAAMAAJ&vq=Kirchoff&pg=PA497#v=onepage&q&f=false) (On the solution of equations to which one is led by the investigation of the linear distribution of galvanic currents), *Annalen der Physik und Chemie*, **72** (12) : 497–508.

19. Harary & Prins 1959, p. 150.

20. Bender & Williamson 2010, p. 173.

21. See Black, Paul E. (4 May 2007). "k-ary tree" (http://xlinux.nist.gov/dads/HTML/karyTree.html) . U.S. National Institute of Standards and Technology. Retrieved 8 February 2015.

22. Stanley, Richard P. (2012), *Enumerative Combinatorics, Vol. I* (https://books.google.com/books?id=0wmJntp8IBQC&pg=PA573) , Cambridge Studies in Advanced Mathematics, **49**, Cambridge University Press, p. 573, ISBN 9781107015425

23. See Li (1996).

References

- Bender, Edward A.; Williamson, S. Gill (2010), *Lists, Decisions and Graphs. With an Introduction to Probability* (https://books.google.com/books?id=vaXv_yhefG8C)
- Dasgupta, Sanjoy (1999), "Learning polytrees", *in Proc. 15th Conference on Uncertainty in Artificial Intelligence (UAI 1999), Stockholm, Sweden, July–August 1999* (http://cseweb.ucsd.edu/~dasgupta/papers/poly.pdf) (PDF), pp. 134–141.
- Deo, Narsingh (1974), *Graph Theory with Applications to Engineering and Computer Science* (https://www.edutechlearners.com/download/Graphtheory.pdf) (PDF), Englewood, New Jersey:

Prentice-Hall, ISBN 0-13-363473-6

- Harary, Frank; Prins, Geert (1959), "The number of homeomorphically irreducible trees, and other species" (https://projecteuclid.org/euclid.acta/1485889061), *Acta Mathematica*, **101** (1–2): 141–162, doi:10.1007/BF02559543 (https://doi.org/10.1007%2FBF02559543), ISSN 0001-5962 (https://www.worldcat.org/issn/0001-5962)
- Harary, Frank; Sumner, David (1980), "The dichromatic number of an oriented tree", *Journal of Combinatorics, Information & System Sciences*, **5** (3): 184–187, MR 0603363 (https://www.ams.org/mathscinet-getitem?mr=0603363).
- Kim, Jin H.; Pearl, Judea (1983), "A computational model for causal and diagnostic reasoning in inference engines", *in Proc. 8th International Joint Conference on Artificial Intelligence (IJCAI 1983), Karlsruhe, Germany, August 1983* (https://www.ijcai.org/Proceedings/83-1/Papers/041.pdf) (PDF), pp. 190–193.
- Li, Gang (1996), "Generation of Rooted Trees and Free Trees", *M.S. Thesis, Dept. of Computer Science, University of Victoria, BC, Canada* (http://webhome.cs.uvic.ca/~ruskey/Theses/GangLiMScThesis.pdf) (PDF), p. 9.
- Simion, Rodica (1991), "Trees with 1-factors and oriented trees", *Discrete Mathematics*, **88** (1): 93–104, doi:10.1016/0012-365X(91)90061-6 (https://doi.org/10.1016%2F0012-365X%2891%2990061-6), MR 1099270 (https://www.ams.org/mathscinet-getitem?mr=1099270).

Further reading

Wikimedia Commons has media related to ***Tree (graph theory)***.

- Diestel, Reinhard (2005), *Graph Theory* (http://diestel-graph-theory.com/index.html) (3rd ed.), Berlin, New York: Springer-Verlag, ISBN 978-3-540-26183-4.
- Flajolet, Philippe; Sedgewick, Robert (2009), *Analytic Combinatorics*, Cambridge University Press, ISBN 978-0-521-89806-5
- "Tree" (https://www.encyclopediaofmath.org/index.php?title=Tree), *Encyclopedia of Mathematics*, EMS Press, 2001 [1994]
- Knuth, Donald E. (November 14, 1997), *The Art of Computer Programming Volume 1: Fundamental Algorithms* (3rd ed.), Addison-Wesley Professional
- Jerrum, Mark (1994), "Counting trees in a graph is #P-complete", *Information Processing Letters*, **51** (3): 111–116, doi:10.1016/0020-0190(94)00085-9 (https://doi.org/10.1016%2F0020-0190%2894%2900085-9), ISSN 0020-0190 (https://www.worldcat.org/issn/0020-0190).

- Otter, Richard (1948), "The Number of Trees", *Annals of Mathematics*, Second Series, **49** (3): 583–599, doi:10.2307/1969046 (https://doi.org/10.2307%2F1969046), JSTOR 1969046 (https://www.jstor.org/stable/1969046).