CHAPTER 10

GRAPHS AND TREES

Graphs are used in many domains. Some examples are:

- network conception and management,
- routing in VLSI circuits, or more general routing problems (e.g. finding shortest paths),
- task scheduling in parallel systems.

And, from a more theoretical standpoint, many data structures can be modelled as graphs.

This chapter gives the basic definitions for directed or undirected graphs, using as an example the proof of the celebrated Euler theorem. As a special case of graphs this chapter introduces the concept of a 'tree', probably one of the most useful concepts for a computer scientist.

Claude Berge, *Graphs and Hypergraphs*, North-Holland, Amsterdam, (1976).
Kenneth Ross, Charles Wright, *Discrete Mathematics*, Prentice Hall, London (1988).

10.1 Graphs

10.1.1 Definitions

We define two types of graph, directed graphs and undirected graphs.

Definition 10.1 *A directed graph is a quadruple* (V, E, α, β) *where:*

- V *is a (possibly infinite) set of vertices (or nodes),*
- E *is a set of edges, disjoint from* V*,*
- α *and* β *are two mappings from* E *to* V *associating with each edge* e *its initial vertex (also called origin)* $\alpha(e)$ *and its terminal or end vertex (also called target)* $\beta(e)$*.* $\alpha(e)$ *and* $\beta(e)$ *are called the endpoints of edge* e*.*

The fact that an edge has an origin and a target enables us to orient edges. Normal transit will go from the origin towards the target. Several edges may have the same origin and the same target; such graphs are thus sometimes called *multigraphs*.

EXAMPLE 10.2

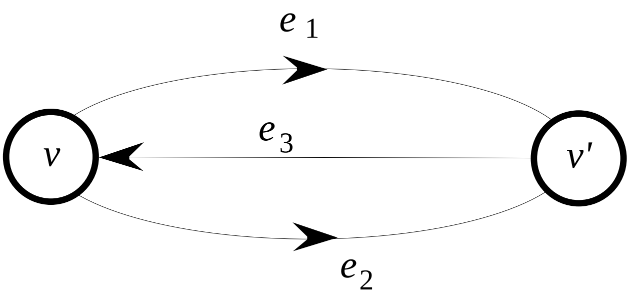

Figure 10.1

Figure 10.1 is a directed graph with two vertices, v and v', and three edges, e_1, e_2 and e_3, with

$$\begin{aligned} \alpha(e_1) = \alpha(e_2) &= v, \\ \beta(e_1) = \beta(e_2) &= v', \\ \alpha(e_3) &= v', \\ \beta(e_3) &= v. \end{aligned}$$

e_1 and e_2 have the same origin and the same target.

Definition 10.3 *An undirected graph is a graph in which we cannot distinguish the origin and the target of an edge. It is a triple (V, E, δ), where δ associates two not necessarily distinct vertices with each edge.*

EXAMPLE 10.4

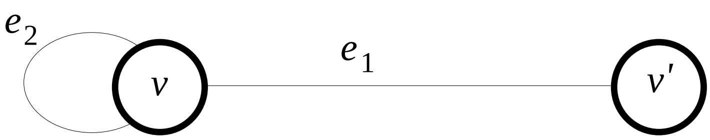

Figure 10.2

Figure 10.2 is an undirected graph with two vertices, v and v', and two edges, e_1 and e_2, with

$$\begin{aligned} \delta(e_1) &= \{v, v'\}, \\ \delta(e_2) &= \{v\}. \end{aligned}$$

A graph (directed or undirected) is said to be *finite* if it has a finite number of vertices and edges, i.e. if both sets V and E are finite.

A directed graph can always be transformed into an undirected graph by 'forgetting' the edge orientations. If $G = (V, E, \alpha, \beta)$ is a directed graph then (V, E, δ) is an undirected graph denoted by $\gamma(G)$, where $\delta(e) = \{\alpha(e), \beta(e)\}$.

Conversely, if a graph is undirected, it may be transformed into a directed graph by assigning an arbitrary direction to each edge:

- If $\delta(e) = \{v\}$, then $\alpha(e) = v = \beta(e)$.
- If $\delta(e) = \{v, v'\}$, with $v \neq v'$, then we may let ($\alpha(e) = v$ and $\beta(e) = v'$) or ($\alpha(e) = v'$ and $\beta(e) = v$).

Clearly, there are several ways of orienting a graph.

Definition 10.5 *If G is an undirected graph, then a directed graph G' is an orientation of G if $\gamma(G') = G$.*

10.1.2 Isomorphic graphs

Sometimes, the actual name of a vertex or an edge of a graph does not really matter and we may consider that two graphs differing only in the names of their vertices and their edges are in fact identical. The notion of *isomorphism* formally expresses this idea.

Two directed graphs $G = (V, E, \alpha, \beta)$ and $G' = (V', E', \alpha', \beta')$ are said to be *isomorphic* if there are two bijections $h_{vtx}: V \longrightarrow V'$ and $h_{edg}: E \longrightarrow E'$ such that

$$\forall e \in E, \qquad \alpha'(h_{edg}(e)) = h_{vtx}(\alpha(e)) \text{ and } \beta'(h_{edg}(e)) = h_{vtx}(\beta(e)).$$

Two undirected graphs $G = (V, E, \delta)$ and $G' = (V', E', \delta')$ are said to be *isomorphic* if there are two bijections $h_{vtx}: V \longrightarrow V'$ and $h_{edg}: E \longrightarrow E'$ such that

$$\forall e \in E, \qquad \delta'(h_{edg}(e)) = \begin{cases} \{h_{vtx}(v)\} & \text{if } \delta(e) = \{v\}, \\ \{h_{vtx}(v), h_{vtx}(v')\} & \text{if } \delta(e) = \{v, v'\}. \end{cases}$$

Given two (directed or undirected) graphs G and G' whose sets of vertices and of edges, respectively, are the pairwise disjoint sets V, E and V', E', the *disjoint union* of G and G' is the graph whose set of vertices is $V \cup V'$, whose set of edges is $E \cup E'$ and in which the edges and vertices are connected exactly as in graphs G and G'.

In some cases we may wish to construct the disjoint union of two graphs while their sets of vertices and edges are not disjoint; we will nevertheless be able to construct this disjoint union by substituting an isomorphic graph for one of the two graphs.

10.1.3 Simple graphs

In a graph, directed or undirected, a *loop* is an edge e, both endpoints of which are equal. This can be formally stated as follows: for a directed graph, $\alpha(e) = \beta(e)$ and, for an undirected graph, $\delta(e)$ is a singleton.

A directed or undirected graph is said to contain *multiple edges* when several edges are allowed between pairs of vertices, i.e. there may be edges e and e' with $\alpha(e) = \alpha(e')$ and $\beta(e) = \beta(e')$ for directed graphs, and $\delta(e) = \delta(e')$ for undirected graphs.

If a directed graph $G = (V, E, \alpha, \beta)$ has no multiple edges, we may identify the set E of its edges with a subset of the Cartesian product $V \times V$. Indeed, in this case, the mapping $(\alpha, \beta) : E \longrightarrow V \times V$ is injective.

A directed or undirected graph is said to be *simple* if it contains neither loops nor multiple edges. If an undirected graph $G = (V, E, \delta)$ is simple, we may identify the set of its edges with a subset of the set $\mathcal{P}_2(V)$ of two-element subsets of V. Indeed, since G has no loop, then for any edge e, $\delta(e)$ has two elements and, because G has no multiple edges, the mapping $\delta : E \longrightarrow \mathcal{P}_2(V)$ is injective.

10.1.4 Subgraphs and partial graphs

Intuitively, a partial graph is obtained from graph G by deleting some edges, and a subgraph is obtained from G by deleting some vertices, together with any edge whose origin or target is one of the deleted vertices.

Let $G = (V, E, \alpha, \beta)$ be a directed graph. The directed graph $G' = (V', E', \alpha', \beta')$ is a *partial graph* of G if

- $V' = V$,
- $E' \subseteq E$ and
- $\forall e \in E', \quad \alpha'(e) = \alpha(e)$ and $\beta'(e) = \beta(e)$.

It is a *subgraph* of G if

- $V' \subseteq V$,
- $E' = \{e \in E \,/\, \{\alpha(e), \beta(e)\} \subseteq V'\}$ and
- $\forall e \in E', \quad \alpha'(e) = \alpha(e)$ and $\beta'(e) = \beta(e)$.

It is a *subpartial graph* if it is a subgraph of a partial graph.

Let $G = (V, E, \delta)$ and $G' = (V', E', \delta')$ be undirected graphs. The undirected graph $G' = (V', E', \delta')$ is a partial graph of G if $V' = V$, $E' \subseteq E$, and $\forall e \in E'$, $\delta'(e) = \delta(e)$. It is a subgraph of G if $V' \subseteq V$, $E' = \{e \in E \,/\, \delta(e) \subseteq V'\}$, and $\forall e \in E', \delta'(e) = \delta(e)$. It is a subpartial graph if it is a subgraph of a partial graph.

EXERCISE 10.1 Show that graph G' is a subpartial graph of G if and only if it is a partial graph of a subgraph of G. Show that a subgraph and a partial graph of a subpartial graph of G are also subpartial graphs of G. $\diamondsuit$

10.1.5 Degree of a vertex

In the present subsection, we consider only graphs such that for any vertex the number of edges leading into that vertex or going out of that vertex is finite; all finite graphs clearly satisfy this requirement.

The *degree* $d(v)$ of a vertex v of an undirected graph $G = (V, E, \delta)$ is equal to the number of edges e such that $\delta(e) = \{v, v'\}$ with $v \neq v'$ plus twice the number of edges e such that $\delta(e) = \{v\}$. If $\delta(e) = \{v\}$, edge e will be counted twice in the degree of v! The degree $d(v)$ of a vertex v of a directed graph is equal to the number of edges e such that $v = \alpha(e)$ or $v = \beta(e)$ (if $\alpha(e) = \beta(e)$, edge e will be counted twice); in other words, the degree of a vertex is the sum of the number of ingoing edges and the number of outgoing edges. The *indegree* $d^-(v)$ of a vertex v of a directed graph G is equal to the number of edges e such that $\beta(e) = v$. The *outdegree* $d^+(v)$ is equal to the number of edges e such that $\alpha(e) = v$. We thus have, for a directed graph, $d(v) = d^+(v) + d^-(v)$.

Proposition 10.6 *The sum of the degrees of all vertices of a finite undirected graph is equal to twice the number of its edges.*

Proof. If $\delta(e) = \{v, v'\}$, with $v \neq v'$, then edge e is counted once in the degree of v and once in the degree of v'. If $\delta(e) = \{v\}$, then edge e is counted twice in the degree of v. Each edge is thus counted twice in the sum of the degrees of the vertices. □

Exercise 10.2 Show that in a finite directed graph, the sum of the indegrees of all vertices is equal to the sum of the outdegrees of these vertices. To what other number is this sum also equal? ♢

Exercise 10.3 Let G be a finite undirected graph with n vertices and m edges, where $n \geq 1$ and $m \geq 0$. For any integer $k \in \mathbb{N}$, let n_k be the number of vertices of degree k; let K be the maximum of the degrees of the vertices (i.e. $n_K > 0$ and $n_k = 0$ for $k > K$).

1. Show that

$$\sum_{k=0}^{K} k n_k = 2m \qquad \text{and} \qquad \sum_{k=0}^{K} n_k = n.$$

2. Show that $K \leq 2m$; give an example where equality holds.
3. Show that if G has neither loops nor multiple edges, then $K \leq n - 1$; give an example where equality holds. ♢

Exercise 10.4 Let G be a finite simple undirected graph with n vertices ($n > 1$).

1. Show that the degree of a vertex is always strictly less than n.
2. Prove that there cannot simultaneously be a vertex of degree 0 and a vertex of degree $n - 1$.
3. Deduce that there are at least two vertices having the same degree. ♢

EXERCISE 10.5 Let $G = (V, E, \delta)$ be an undirected simple finite graph. We assume that G contains no *triangles*: a triangle consists of three distinct vertices $v_1, v_2, v_3 \in V$ and three edges $e_1, e_2, e_3 \in E$ with $\delta(e_1) = \{v_2, v_3\}$, $\delta(e_2) = \{v_1, v_3\}$ and $\delta(e_3) = \{v_1, v_2\}$. Two vertices are said to be adjacent if they are connected by an edge.

1. Show that for two distinct adjacent vertices x and y, the number n_x of vertices of $V \setminus \{x, y\}$ adjacent to x and the number n_y of vertices of $V \setminus \{x, y\}$ adjacent to y satisfy the inequality

$$n_x + n_y \leq |V| - 2.$$

2. Deduce, by induction on the number $|V|$ of vertices, that the number $|E|$ of edges verifies

$$|E| \leq \frac{|V|^2}{4}. \qquad \diamondsuit$$

10.1.6 Paths

In a directed graph G, a *path* is a sequence $c = e_1, \ldots, e_n$ of edges such that $\forall i \in \{1, \ldots, n-1\}$, $\beta(e_i) = \alpha(e_{i+1})$. Vertex $\alpha(e_1)$ will be called the *origin* of path c and vertex $\beta(e_n)$ will be called the *target* of path c. A *circuit* is a path such that $\beta(e_n) = \alpha(e_1)$. A path (or a circuit) is *simple* if it does not contain the same edge twice. It is *elementary* if it does not contain two edges with the same origin or the same target (it is hence *a fortiori* simple). In a finite directed graph, a path or circuit is said to be an *Euler* path or circuit if it is simple and contains all edges. It is said to be a *Hamiltonian* path or circuit if it is elementary and goes through all vertices (i.e. $\forall v \in V$, $\exists i : v = \alpha(e_i)$ or $v = \beta(e_i)$).

For undirected graphs, a *chain* is a sequence

$$c = v_0, e_1, v_1, e_2, \ldots, v_{n-1}, e_n, v_n$$

of vertices and edges such that $\delta(e_i) = \{v_{i-1}, v_i\}$. This chain is said to connect vertices v_0 and v_n. A *cycle* is a chain such that $v_0 = v_n$. If the sequence $v_0, e_1, v_1, e_2, \ldots, v_{n-1}, e_n, v_n$ is a chain, the sequence $v_n, e_n, v_{n-1}, \ldots, e_2, v_1, e_1, v_0$ is also a chain and it connects v_n to v_0. A chain (or a cycle) is *simple* if it does not contain the same edge twice and *elementary* if it does not contain the same vertex twice, with the only exception being that v_0 and v_n may be equal. In a finite undirected graph, a chain (or a cycle) is an *Euler* chain (or a cycle) if it is simple and contains all edges; it is a *Hamiltonian* chain (or a cycle) if it is elementary and contains all vertices.

EXAMPLE 10.7

1. In the graph of Figure 10.3:
   - $e_1e_2e_3e_4$ is a simple circuit, and
   - $e_4e_1e_2e_3e_5$ is an Euler path but is not a Hamiltonian path.

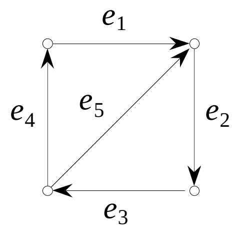

Figure 10.3

2. In the directed graph of Example 10.2:
   - $e_1 e_3$ is a simple circuit, a Hamiltonian circuit, but not an Euler circuit and
   - $e_1 e_3 e_2$ is an Euler path, non-elementary.
3. In the undirected graph of Example 10.4:
   - v, e_2, v is an elementary cycle,
   - v, e_2, v, e_1, v' is a simple chain, non-elementary, and an Euler chain and
   - $v', e_1, v, e_2, v, e_1, v'$ is a non-simple cycle.

The next result is an immediate consequence of the definitions.

Proposition 10.8 *If $e_1, \ldots, e_n$ is a simple (resp. elementary, Euler, Hamiltonian) path (circuit) of a directed graph G, then*

$$\alpha(e_1), e_1, \alpha(e_2), e_2, \ldots, \alpha(e_n), e_n, \beta(e_n)$$

is a simple (resp. elementary, Euler, Hamiltonian) chain (cycle) of $\gamma(G)$.

Conversely, if $v_0, e_1, \ldots, e_n, v_n$ is a simple chain, there is at least one orientation of G such that $e_1, \ldots, e_n$ is a path: since $\delta(e_i) = \{v_{i-1}, v_i\}$, it suffices to let $\alpha(e_i) = v_{i-1}$ and $\beta(e_i) = v_i$. The fact that the chain is simple implies that this construction is always possible because there is no index $j \neq i$ with $e_j = e_i$. Otherwise, if we had $e_i = e_j = e$, we would have $\delta(e) = \{v_{i-1}, v_i\} = \{v_{j-1}, v_j\}$; and it is easy to see that the above construction can be applied only when $\alpha(e) = v_{i-1} = v_{j-1}$ and $\beta(e) = v_i = v_j$.

Proposition 10.9 *If an undirected graph G contains two different simple chains connecting the same two distinct vertices, then it contains a simple cycle.*

Proof. Let $c = v_0, e_1, v_1, \ldots, v_{n-1}, e_n, v_n$ and $c' = v'_0, e'_1, v'_1, \ldots, v'_{n'-1}, e'_{n'}, v'_{n'}$ be two simple chains such that $v_0 = v'_0 \neq v_n = v'_{n'}$. We prove by induction on $n + n'$ that the existence of such chains implies the existence of a simple cycle:

1. Since $n \geq 1$ (because $v_0 \neq v_n$) and $n' \geq 1$ (for the same reasons), the least possible value of $n + n'$ is 2. In this case, $c = v, e, v'$, $c' = v, e', v'$, with $e \neq e'$, and v, e, v', e', v is a simple cycle.

2. (a) If the sets $E(c) = \{e_i \,/\, 1 \le i \le n\}$ and $E(c') = \{e'_j \,/\, 1 \le j \le n'\}$ are disjoint, then $v_0, e_1, v_1, \ldots, v_{n-1}, e_n, v_n = v'_{n'}, e'_{n'}, v'_{n'-1}, \ldots, v'_1, e'_1, v'_0$ is a simple cycle.
   (b) If $e_1 = e'_1$, then $v_1 = v'_1$, and two cases must be considered:
      (b.1) If $n' = 1$, then $v_n = v'_1 = v_1$ and $v_1, \ldots, v_{n-1}, e_n, v_n$ is a simple cycle. The case $n = 1$ is similar.
      (b.2) Otherwise the two chains $v_1, e_2, \ldots, v_{n-1}, e_n, v_n$ and $v'_1, e'_2, \ldots, v'_{n'-1}, e'_{n'}, v'_{n'}$ with respective lengths $n-1$ and $n'-1$ again verify the hypotheses, and thus G contains a simple cycle.
   (c) We are thus left with the case in which $E(c) \cap E(c') \ne \emptyset$ and in which $e_1 \ne e'_1$. Hence, there exists $i > 1$ and $j > 1$ such that $e_i = e'_j$, hence $v_{i-1} = v'_{j-1}$ or $v_{i-1} = v'_j$. In the first case, the chains $v_0, e_1, v_1, \ldots, v_{i-2}, e_{i-1}, v_{i-1}$ and $v'_0, e'_1, v'_1, \ldots, v'_{j-2}, e'_{j-1}, v'_{j-1}$, with respective lengths $i-1 < n$ and $j-1 < n'$, again verify the hypotheses, and G contains a simple cycle. In the second case, the chains $v_0, e_1, v_1, \ldots, v_{i-2}, e_{i-1}, v_{i-1}$ and $v'_0, e'_1, v'_1, \ldots, v'_{j-2}, e'_{j-1}, v'_{j-1}, e'_j, v'_j$, with respective lengths $i-1 < n$ and $j \le n$, also verify the hypotheses and G contains a simple cycle. □

Exercise 10.6 Show that in an undirected graph, the shortest chain between two distinct vertices is elementary. Does this also hold for the shortest path from a vertex to another one in a directed graph? ◇

In an undirected graph, the *distance* $d(v, v')$ between two vertices v and v' is defined as follows:

- If $v = v'$ then $d(v, v') = 0$.
- If $v \ne v'$ then $d(v, v')$ is equal to
  - the length of the shortest chain connecting these two vertices, if such a chain exists, or
  - ∞ otherwise.

The characteristic properties of a distance are indeed verified:

- $d(v, v') = 0$ if and only if $v = v'$,
- $d(v, v') = d(v', v)$,
- $d(v, v'') \le d(v, v') + d(v', v'')$.

Exercise 10.7 Prove the triangular inequality $d(v, v'') \le d(v, v') + d(v', v'')$. ◇

The *diameter* of an undirected graph is the maximum distance between two distinct vertices, i.e. $\sup\{d(v, v') \,/\, v, v' \in V\}$.

If a graph is directed, we may also define the distance between two vertices as the length of the shortest path going from the first one to the second one. But this is no longer a distance in the mathematical sense because, while it still satisfies the triangular inequality, it is no longer symmetrical: it may well occur that $d(v, v') \ne d(v', v)$.

EXERCISE 10.8 Let $X = \{0, 1, 2, 3, 4\}$. The *Petersen graph* is the undirected graph defined as follows: its vertices are the pairs of elements of X, and two vertices are connected by an edge if and only if they are two disjoint pairs of elements of X. For instance, if $\{0, 1\}$, $\{1, 2\}$ and $\{2, 3\}$ are vertices, there is an edge connecting $\{0, 1\}$ and $\{2, 3\}$, but there is no edge connecting $\{0, 1\}$ and $\{1, 2\}$:

1. Determine the number of vertices, the number of edges, the degree of vertices and the diameter of this graph.
2. Draw this graph and indicate for each vertex the two elements of X constituting the pair that it contains. ♢

EXERCISE 10.9
1. Let (V, E, α, β) be a directed graph with n vertices $v_1, \ldots, v_n$. It is associated with the matrix M defined by: each entry $M_{i,j}$ of M (for $1 \le i \le n$ and $1 \le j \le n$) is the number of edges of E with origin v_i and target v_j (i.e. $\alpha(e) = v_i$ and $\beta(e) = v_j$). Prove by induction that for any integer $k > 0$, the matrix M^k has as its (i, j)th entry the number of distinct paths of length k between v_i and v_j.
2. Can you generalize this result to an undirected graph? ♢

EXERCISE 10.10 In $\mathbb{Z}^2$ we define the *4-distance* d_4 and the *8-distance* d_8 as follows:

$$\begin{aligned} d_4\big((x, y), (x', y')\big) &= |x - x'| + |y - y'| \ , \\ d_8\big((x, y), (x', y')\big) &= \max\big(|x - x'|, |y - y'|\big) \ . \end{aligned}$$

1. Define the two undirected graphs G_4 and G_8 having $\mathbb{Z}^2$ as the set of vertices, and where the distances above defined in terms of lengths of chains coincide with d_4 and d_8 respectively.
2. Draw the subgraphs of G_4 and G_8 corresponding to the set of the sixteen vertices (x, y) where $0 \le x \le 3$ et $0 \le y \le 3$. ♢

10.1.7 Connectivity

A directed graph is said to be *strongly connected* if for any pair (v, v') of distinct vertices there is a path going from v to v'.

An undirected graph is *connected* if for any pair (v, v') of distinct vertices there exists a chain connecting v and v'.

The *connected component* $CC_G(v)$ of a vertex v of an undirected graph G is equal to $\{v\}$ together with the set of vertices v' of G such that there exists a chain connecting v to v'.

Proposition 10.10 *If $v' \in CC_G(v)$ then $CC_G(v) = CC_G(v')$.*

Proof. We assume that $v \neq v'$. (If $v = v'$, the result is trivial.) If $v' \in CC_G(v)$ there is a chain c connecting v and v' and a chain c' connecting v' and v. Let $v'' \in CC_G(v)$:

- If $v'' = v$, then since c' connects v' to v, we have that $v'' = v \in CC(v')$.

- Otherwise, there is a chain c'' connecting v to v''; then $c'c''$ is a chain connecting v' to v'', and $v'' \in CC_G(v')$.

Thus, $CC_G(v) \subseteq CC_G(v')$. The converse inclusion $CC_G(v') \subseteq CC_G(v)$ can be proved in the same way. □

The following proposition follows immediately from the definitions.

Proposition 10.11 *An undirected graph $G = (V, E, \delta)$ is connected if and only if $\forall v \in V$, $V = CC_G(v)$.*

An *isolated* vertex is a vertex of degree 0. If a vertex v of G is isolated, then $CC_G(v) = \{v\}$. If G is a graph without loops, a vertex v of G is isolated if and only if $CC_G(v) = \{v\}$.

Proposition 10.12 *Let G be a connected undirected graph. Let G' be the graph obtained by deleting an edge e with two distinct end vertices v' and v'' (i.e. $\delta(e) = \{v', v''\}$ with $v' \neq v''$). Then $V = CC_{G'}(v') \cup CC_{G'}(v'')$.*

Proof. Let v be any vertex of G. Because G is connected, there exists a chain $v_0, e_1, v_1, \ldots, e_n, v_n$ with $v' = v_0$ and $v_n = v$. If edge e does not occur in that chain, then $v \in CC_{G'}(v')$. Otherwise, let i be the largest index such that $e_i = a$. Then $v_i \in \{v', v''\}$ and $v_i, e_{i+1}, \ldots, e_n, v_n$ is a chain of G', and hence $v = v_n \in CC_{G'}(v'') \cup CC_{G'}(v')$. □

Exercise 10.11 Let C_n be an undirected graph with n vertices consisting of a single cycle, i.e.

$$V = \{v_1, \ldots, v_n\},$$
$$E = \{e_1, \ldots, e_n\},$$
$$\delta(e_i) = \begin{cases} \{v_{i-1}, v_i\} & \text{if } i > 1, \\ \{v_n, v_1\} & \text{if } i = 1. \end{cases}$$

1. Show that each vertex of C_n is of degree 2.
2. Show that if G is an undirected connected graph with n vertices all of which are of degree 2, then G is isomorphic to C_n.
3. Show that if G is an undirected graph with n vertices all of which are of degree 2, then G is a disjoint union of graphs $G_{n_1}, \ldots, G_{n_k}$, where G_{n_i} is isomorphic to C_{n_i}, with $n_1 + \cdots + n_k = n$. ♢

10.1.8 An historical example: Königsberg seven bridges

The old town of Königsberg in eastern Prussia (nowadays Kaliningrad, in Russia), contains an island connected to the 'mainland' by seven bridges as shown in Figure 10.4.

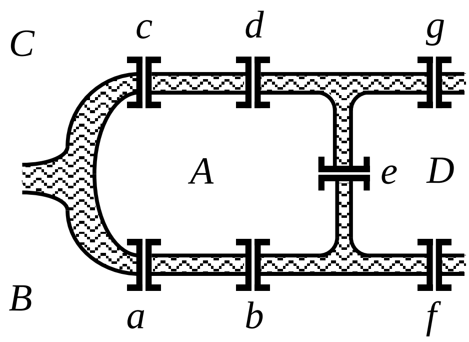

Figure 10.4

Is it possible to take a walk through Königsberg starting from some point and travelling across all the bridges exactly once?

The problem is that of finding an Euler chain in the undirected graph shown in Figure 10.5.

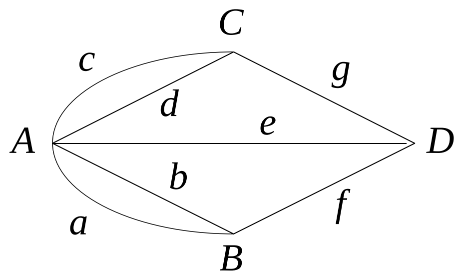

Figure 10.5

If, moreover, the walk ends at its starting point we will have an Euler cycle.

The answer to this problem can be obtained by the following theorem (proved by Euler in 1766).

Theorem 10.13 *An undirected finite graph G with no isolated vertex has an Euler chain if and only if*

(i) *it is connected*
(ii) *it has zero or two odd degree vertices.*

In the case in which there is no odd degree vertex, this Euler chain is a cycle. In the case in which there are two odd degree vertices, they are the origin and target of the chain.

Proof.
1. Assume that there is an Euler chain $v_0, e_1, v_1, e_2, \ldots, v_{n-1}, e_n, v_n$.
Because there is no isolated vertex,

$$V = \cup_{e \in E} \delta(e) \,,$$

and since $E = \{e_1, \ldots, e_n\}$,

$$v \in V \Longleftrightarrow \exists i : v = v_i.$$

G is thus connected.

Let v be any vertex; the number $d(v)$ of edges with endpoint v is equal to

$$2 \times |\{i \,/\, v_i = v, 1 \leq i \leq n-1\}| \quad + \begin{cases} +1 \text{ if } v = v_0, \\ +1 \text{ if } v = v_n. \end{cases}$$

Indeed, for each $i \in \{1, \ldots, n-1\}$, v_i is the endpoint of the two distinct edges e_i and e_{i+1}. Moreover, v_0 is the endpoint of edge e_1 and v_n is the endpoint of edge e_n. Note that if we have an edge e_i, both endpoints v_{i-1} and v_i of which are equal, this edge will be counted twice in the computation. We deduce the following from this characterization of the degrees of the vertices:

- If $v_0 = v_n$, i.e. if there is an Euler cycle, then all vertices are of even degree.
- If $v_0 \neq v_n$ then all vertices are of even degree except for the endpoints v_0 and v_n of the chain.

2. Consider a connected graph with zero or two vertices of odd degree. We show that it has an Euler chain.

Note first that we may add or delete on the graph any number of loops without modifying either the parity of the degrees of the vertices or the existence of Euler chains. If we add or if we delete on graph G an edge e with $\delta(e) = \{v\}$, we increase or decrease the degree of v by 2. Moreover, if we add to graph G an edge e with $\delta(e) = \{v\}$, thus obtaining a graph G', then $v_0, e_1, ..., e_{i-1}, v, e_{i+1}, ..., v_n$ is an Euler chain of G if and only if $v_0, e_1, ..., e_{i-1}, v, e, v, e_{i+1}, ..., v_n$ is an Euler chain of G'; and we have a similar result if $v = v_0$ or if $v = v_n$.

We may then reason by induction on the number of edges of a graph *without loops*.

If this number is 0, the property holds vacuously. If this number is 1, the unique edge e of this graph is such that $\delta(e) = \{v, v'\}$, with $v \neq v'$, and v, e, v' is an Euler chain.

Let G be a graph with $n+1$ edges:

(a) If all vertices have an even degree, we consider any edge e whose endpoints are v and v'. Let G' be the partial graph obtained by deleting this edge. Both vertices v and v' now have an odd degree. G' must be connected. Indeed, let $CC_{G'}(v)$ be the connected component of v in G'. Firstly, $v' \in CC_{G'}(v)$; otherwise $CC_{G'}(v)$ would be a graph with a single vertex of odd degree, which is impossible (because the sum of the degrees of a graph is always even). There thus exists a chain c connecting v and v' in G'. We now show that $CC_{G'}(v)$ contains all vertices of G'. Let v'' be any vertex. Because G is connected, there exists in G a chain connecting v and v''. If this chain uses edge e, we substitute the chain c for edge e. We thus have a chain of G' connecting v and v''. Because G' is connected and has two odd degree vertices v and v', there exists an Euler chain c' connecting v' and v, and v, e, c' is thus an Euler cycle of G.

(b) Let v and v' be the two vertices of odd degree of G, let e be an edge such that $v \in \delta(e)$ and let v'' be the vertex such that $\delta(e) = \{v, v''\}$. Let G' be the graph obtained by deleting this edge. The degree of v in G' is even.

(b.1) If the degree of v'' in G is odd (i.e. $v'' = v'$) then the degree of v'' in G' is even.

(b.1.1) If G' is connected, there exists an Euler cycle c going from v to v, and $c, e, v' = v''$ is an Euler chain of G.

(b.1.2) Otherwise, by Proposition 10.12, $CC_{G'}(v)$ and $CC_{G'}(v')$ are two connected graphs, and all their vertices have even degrees. There exists an Euler cycle c from v to v in $CC_{G'}(v)$ and an Euler cycle c' from v' to v' in $CC_{G'}(v')$. Then c, e, c' is an Euler chain in G.

(b.2) If the degree of v'' in G is even (i.e. $v'' \neq v'$) then the degree of v'' in G' is odd and $v'' \in CC_{G'}(v')$, because v'' and v' are the only two vertices of odd degree in G'.

(b.2.1) If G' is connected, there is an Euler chain c of G' connecting v'' and v', and v, e, c is an Euler chain in G.

(b.2.2) Otherwise, all vertices of $CC_{G'}(v)$ have an even degree. There exists an Euler cycle c of $CC_{G'}(v)$ connecting v and v and an Euler chain c' of $CC_{G'}(v')$ connecting v' and v'. Then c, e, c' is an Euler chain of G connecting v and v'. □

10.1.9 Graph colouring

A *colouring* of an undirected graph $G = (V, E, \delta)$ without a loop is a mapping $\gamma: V \longrightarrow C$, where C is a finite set of 'colours', such that

$$\forall e, \delta(e) = \{v, v'\} \Longrightarrow \gamma(v) \neq \gamma(v')$$

(i.e. two different vertices connected by an edge cannot have the same colour).

The *chromatic number* of a graph is the minimum number of colours needed to colour it.

EXAMPLE 10.14 Let E be a set of students and X be a set of exams. For each exam x in X, the set of students registered for that exam is $S(x)$. Each student can take at most one exam per day. What is the minimum length of the exam session?

Let G be the graph whose set of vertices is X. Two vertices x and x' are connected by an edge if and only if $S(x) \cap S(x') \neq \emptyset$. The minimal length of the session is the *chromatic number* of this graph.

10.1.10 Planar graphs

A (directed or undirected) graph is said to be *planar* if it can be drawn in the plane without any edges crossing.

REMARK 10.15 A graph, and even a planar graph, can be drawn in many ways. Figure 10.6 shows two possible drawings for the planar graph K_4.

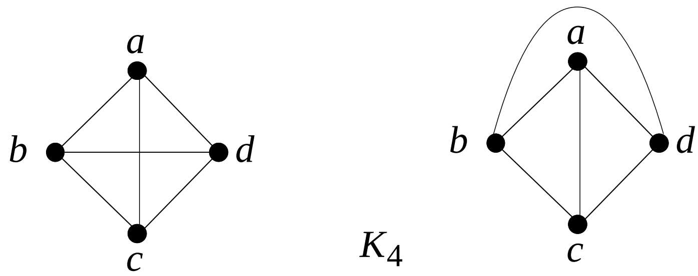

Figure 10.6

The two graphs in Figure 10.7 are not planar.

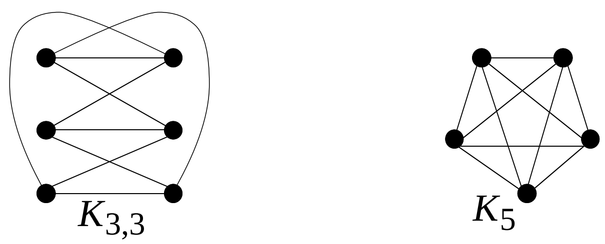

Figure 10.7

A famous problem that has been solved recently is the four-colour problem: is the chromatic number of a planar graph always less than or equal to 4? The answer is 'yes'.

EXERCISE 10.12 Find a planar graph with chromatic number 4. ♢

10.2 Trees and rooted trees

Trees and rooted trees are particular cases of graphs. Trees are usually undirected graphs, and rooted trees are usually directed graphs. However, in computer science, the term 'tree' often means rooted tree, and sometimes very particular rooted trees.

10.2.1 Trees

Definition 10.16 *A tree is an undirected connected graph without a simple cycle.*

If an undirected graph contains at least one edge e, it contains a cycle v, e, v', e, v, where $\delta(e) = \{v, v'\}$. But this cycle is not simple. This is why the definition of trees involves simple cycles.

Proposition 10.17 *Let G be a finite graph, let n be the number of its vertices, let m be the number if its edges and let p be the number of its connected components. Then G is without a simple cycle if and only if $m - n + p = 0$.*

Proof.
1. $m - n + p$ is always non-negative. We show this by induction on the number of edges:
   - It is true for a graph with no edges because if that graph has n vertices then it has n connected components.
   - Let e be an edge of G connecting v and v' (which are thus in the same connected component) and let us delete edge e from G. The graph G' thus obtained has $n' = n$ vertices, $m' = m-1$ edges and p' connected components with $p' = p$ or $p' = p + 1$; hence $p \geq p' - 1$. By the induction hypothesis, $m' - n' + p' \geq 0$, and $m - n + p \geq m - n + p' - 1 = m' - n' + p'$.
2. Assume that G contains a simple cycle c, and let us show that $m - n + p > 0$. Let e be any edge of this cycle, with $\delta(e) = \{v, v'\}$ (v' may be equal to v). There thus exists a chain c' connecting v' and v that does not use edge e.

Let G' be the graph obtained by deleting from G this edge e. We have for G' that $m' = m - 1$ and that $n' = n$. We also have that $p' = p$ because the number of connected components is not modified: if two vertices are connected by a chain of G using edge e, they are connected by the chain of G' obtained by substituting for each occurrence of e the chain c' which is in G'. Hence $m - n + p = m' + 1 - n' + p' > m' - n' + p' \geq 0$.

3. Conversely, we show that if $m - n + p > 0$, then G has a simple cycle. We again show this by induction on the number of edges:
   - If $m = 0$, we have shown in 1 that $m - n + p > 0$ could not occur.

- Let G' be the graph obtained by deleting from G an arbitrary edge e. We have shown in 1 that in this case $m - n + p \geq m' - n' + p'$ and that $p' = p$ or $p' = p + 1$.
  - If $m' - n' + p' > 0$, then, by the induction hypothesis, G' contains a simple cycle, and hence so does G.
  - If $m' - n' + p' = 0$, then, because $m - n + p > 0$, we have that $m - n + p > m' - n' + p'$, which implies $p' = p$. In other words, deleting edge e does not modify connected components. As the endpoints v and v' of e are in the same connected component of G, they are in the same connected component of G'. There thus exists a simple chain c connecting v and v' (see Exercise 10.6) and not using edge e. Adding edge e to this chain, we have a simple cycle. $\square$

Theorem 10.18 *Let G be an undirected graph with n vertices ($n \geq 2$). The following properties are equivalent:*

1. *G is connected and has no simple cycles.*
2. *G has no simple cycles and has $n - 1$ edges.*
3. *G is connected and has $n - 1$ edges.*
4. *G has no simple cycles; if we add an edge to it we form a simple cycle.*
5. *G is connected; if we delete an edge from it, it is no longer connected.*
6. *$\forall v, v' \in S$, ($v \neq v'$), there exists a unique simple chain connecting v and v'.*

Proof.
($1 \Longrightarrow 2$) Because G is connected, $p = 1$. If G has no simple cycles then $m - n + 1 = 0$ and the number of edges m of G is $n - 1$.
($2 \Longrightarrow 3$) If G has no simple cycles, $m - n + p = 0$, and if $m = n - 1$ then $p = 1$ and thus G is connected.
($3 \Longrightarrow 4$) If G is connected and has $n - 1$ edges then $m - n + p = 0$ and thus G has no simple cycles. If we add an edge it remains connected and $(m' - n' + p') = (m + 1 - n + p) > 0$. We have thus exhibited a simple cycle.
($4 \Longrightarrow 5$) If G were not connected we might add an edge to it without creating any cycles: it suffices to add an edge connecting two vertices of two distinct connected components. If deleting one edge yields another connected graph, we have $m' - n' + p' = m - 1 - n + p = 0$, and hence $m - n + p > 0$, and G would have a simple cycle.
($5 \Longrightarrow 6$) Let v and v' be two vertices of G. Because G is connected there exists a chain connecting v and v'. If there were two such chains, by Proposition 10.9 the graph would contain a simple cycle and we might thus delete one edge without destroying connectedness.

($6 \Longrightarrow 1$) G is connected. If it contained a simple cycle, we might find two distinct chains connecting two vertices. □

EXERCISE 10.13 Show that if we delete an edge from a tree then the remaining two connected components are trees. ◇

Proposition 10.19 *A tree with n vertices ($n \geq 2$) has at least two vertices with degree 1.*

Proof. In a graph, the sum of the degrees of the vertices is equal to twice the number of edges. In a tree, the number of edges is $n-1$ and the sum of the degrees of the vertices is thus equal to $2n-2$. Since a tree is a connected graph, there are no vertices of degree 0. Let k be the number of vertices of degree 1. There are thus $n-k$ vertices of degree at least equal to 2, and the sum of the degrees of the vertices is greater than or equal to $k+2(n-k) = 2n-k$. As $2n-2 \geq 2n-k$, we have $k \geq 2$. □

EXERCISE 10.14 Show that if a tree G has exactly two vertices of degree 1, then all other vertices are of degree 2. Deduce that G consists of a single elementary chain. ◇

Proposition 10.20 *A finite graph has a partial graph which is a tree if and only if it is connected.*

Proof. If a graph has a connected partial graph, it is connected. If a graph is connected and if we may delete an edge while preserving connectedness, we delete this edge. When no more edges can be deleted, we will have obtained a tree, by point 5 of Theorem 10.18. □

EXERCISE 10.15 Prove (by induction) that a tree can always be drawn in the plane in such a way that the edges form linear segments without cross-sections, except at the endpoints; in particular, it is planar. ◇

10.2.2 Rooted trees

Definition 10.21 *A rooted tree is a directed graph G such that $\gamma(G)$ is a tree and all vertices of G have indegree 1, except for a single vertex, called the root, whose indegree is 0.*

Proposition 10.22 *If G is a finite tree, then for any vertex v of G, there exists an orientation of G which is a rooted tree with root v.*

Proof. By induction on the number of vertices of G. The result is clear if the tree G has only one vertex. Otherwise, let v be any vertex of a tree G and let e be an edge with endpoints v and v'. Deleting this edge yields two connected graphs $G(v)$ and $G(v')$, which are still trees. We assign an orientation to these two trees in such a way that their respective roots are v and v', and we orient edge e from v to v'. We thus have a rooted tree with root v. In fact, for any vertex v'' different from v and v' of $G(v) \cup G(v')$, the indegree of v'' is equal to its indegree in the rooted tree constructed from $G(v)$ or from $G(v')$, i.e. 1. The indegree of v' is 1 and the indegree of v is 0. □

Proposition 10.23 *Let G be a rooted tree and let v be one of its vertices. Let $G(v)$ be the subgraph of G whose set of vertices is $\{v\}$ augmented by all the vertices that are the target of a path with origin v. Then $G(v)$ is a rooted tree with root v.*

Proof. By the construction of $G(v)$, $\gamma(G(v))$ is connected. Moreover, $\gamma(G(v))$ contains no simple cycle; otherwise, this simple cycle would also belong to $\gamma(G)$, which is impossible because $\gamma(G)$ is a tree. $\gamma(G(v))$ is thus a tree.

By the construction of $G(v)$, any vertex v' of $G(v)$ different from v is the target of an edge whose origin is in $G(v)$. This edge is the only edge of G with endpoint v'. The indegree of v' in $G(v)$ is thus 1. If the indegree of v were also equal to 1 in $G(v)$, there would exist an edge with target v and with origin in $G(v)$. There would also exist in G a simple circuit going through v, and there would thus exist in G a simple cycle, which is impossible. □

Theorem 10.24 (König's lemma) *Let G be an infinite rooted tree all of whose vertices have a finite outdegree. Then G has an infinite path originating at the root.*

Proof. Let $e_1, \ldots, e_n$ be the edges of G whose common origin is the root of G. One among the rooted trees $G(\beta(e_i))$ for $i = 1, \ldots, n$ is thus infinite. (These rooted trees are defined in Proposition 10.23.) Let i be such that $G(\beta(e_i))$ is infinite, and let $e'_1 = e_i$. Assume now that we have defined a path $e'_1 e'_2 ... e'_n$ such that $G(\beta(e'_n))$ is infinite and let $e''_1, \ldots, e''_k$ be the edges with origin $\beta(e'_n)$. Again, we will find a j such that $G(\beta(e''_j))$ is infinite and we will let $e'_{n+1} = e''_j$. Since we may repeat this construction indefinitely it indeed yields an infinite path in the rooted tree G. □

König's lemma is often applied in different forms that are consequences of the above theorem. Two such consequences are given in Proposition 10.25 and Proposition 10.26 below.

Proposition 10.25 *Let $G = (V, E, \alpha, \beta)$ be a directed graph such that any vertex v has a finite outdegree $d^+(v)$.*

If a vertex v of this graph is the origin of infinitely many finite paths, it must also be the origin of an infinite path.

Proof. Let C be the infinite set of the paths with origin v. For $n > 0$, let C_n be the set of paths with origin v and length n so that $C = \bigcup_{n>0} C_n$. Also let $C_0 = \{v\}$ and $C' = C_0 \cup C$. We define the set $R \subseteq C' \times C'$ by $(c, c') \in R$ if and only if

- either $c = v \in C_0$ and $c' \in C_1$,
- or there exists $n > 0$ and $e \in E$ such that $c \in C_n$ and $c' = ce$.

We then show that the graph $G' = (C', R, \alpha', \beta')$ with $\alpha'(c, c') = c$ and $\beta'(c, c') = c'$ is a rooted tree. The indegree of v is 0 and its outdegree is $d^+(v)$. The indegree of $c \in C$ is 1 and its outdegree is equal to $d^+(v')$, where v' is the target of the path c.

By König's lemma, this rooted tree has an infinite path $v, c_1, c_2, \ldots, c_n \ldots$ with $c_1 = e_1$, $c_{i+1} = c_i e_{i+1}$. Hence we deduce that $e_1 e_2 \cdots e_n \cdots$ is an infinite path with origin v in graph G. □

EXERCISE 10.16 Let A be a finite alphabet and let L be a subset of A^*. Show that if L is infinite, there exists at least one infinite string $a_0 a_1 a_2 \cdots a_n \cdots$ of letters of A such that $\forall n \geq 0$, $\exists w_n \in A^* : a_0 \cdots a_n w_n \in L$. ♢

Proposition 10.26 *Let E_n be a finite non-empty set, for any integer $n \geq 0$, and assume that*

(i) $n \neq m \Longrightarrow E_n \cap E_m = \emptyset$.

Let R be a binary relation on $E = \cup_{n\geq 0} E_n$ such that:

(ii) *If $e \; R \; e'$ then there exists n such that $e \in E_n$ and $e' \in E_{n+1}$.*

(iii) $\forall n \geq 0$, $\forall e' \in E_{n+1}$, $\exists e \in E_n : e \; R \; e'$.

Then there exists an infinite sequence $e_0, e_1, \ldots, e_n, \ldots$ such that $\forall n \geq 0$, $e_n \in E_n$ and $e_n \; R \; e_{n+1}$.

Proof. Consider (E, R) as a directed graph. By (ii), the outdegree of an element of E_n is at most equal to the number of elements of E_{n+1}. It is thus finite. By (i), $\bigcup_{n>0} E_n$ is infinite. By (iii), each element of E_{n+1} is the target of a path whose origin is in E_0. There thus exist infinitely many paths whose origin is in E_0, and since E_0 is finite there exists an element e_0 of E_0 that is the origin of infinitely many finite paths. By Proposition 10.25, it also is the origin of an infinite path going through the sequence of vertices $e_0, e_1, \ldots, e_n, \ldots$, which is the required sequence. □

EXERCISE 10.17 Let V be a subset of $\mathbb{N} \times \mathbb{N}$ with the following properties:

(i) For any $n \geq 0$, the set $\{m \in \mathbb{N} \,/\, (n, m) \in S\}$ is finite.

(ii) For any $n \geq 0$, there exists an injection $f_n : \{0, 1, \ldots, n\} \to \mathbb{N}$ such that $\forall i \in \{0, 1, \ldots, n\}$, $(i, f_n(i)) \in V$.

Prove that there exists an injection $f : \mathbb{N} \to \mathbb{N}$ such that $\forall i \in \mathbb{N}$, $(i, f(i)) \in V$. ♢

10.2.3 Ordered rooted trees

Let G be a rooted tree and let v be one of its vertices. Vertex v' is said to be a *child* of v if there is an edge with origin v and target v'. If v is not the root of the rooted tree, v has indegree 1: there thus exists exactly one edge e with v as target. The origin vertex of this edge e will be called the *parent* of v. It is easy to show that any vertex v is the parent of its children.

A rooted tree is said to be *ordered* if, for any vertex v, the set of children of v is endowed with a total ordering. When drawing such a rooted tree in the plane (usually, counterintuitively, with the root at the top of the graph and the children below their parent), this total ordering on the children of a same parent will be materialized by writing them from left to right. This is why such rooted trees are called 'ordered'. The trees that we have studied in Chapter 3 are ordered rooted trees which may be empty, i.e. they may have empty sets of vertices and edges.

A *complete binary tree* is an ordered rooted tree in which each vertex either has two children (respectively called the *left child* and the *right child*) or none. A vertex without a child is called a *leaf*.

EXAMPLE 10.27 Two different ordered binary trees that represent the same rooted tree are shown in Figure 10.8.

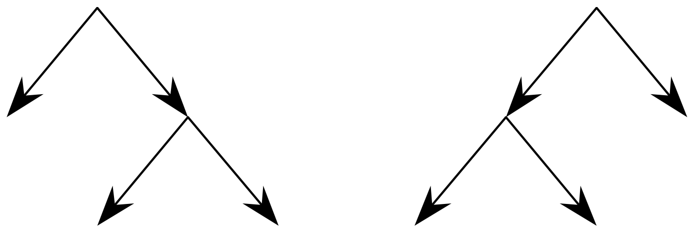

Figure 10.8

EXERCISE 10.18

1. Show that a finite complete binary tree has an odd number of vertices.
2. Show that a complete binary tree with $2n-1$ vertices has n leaves. ♢