CS1231 Chapter 11

Graphs

11.1 Paths

Warning 11.1.1. There are several commonly used, conflicting sets of terminologies for graphs. Always check the definitions being used when looking into the literature.

Definition 11.1.2. Let G be an undirected graph, where $G = (V, E)$.

(1) Denote by $\mathrm{V}(G)$ and $\mathrm{E}(G)$ the set of all vertices and the set of all edges in G respectively, i.e.,

$$\mathrm{V}(G) = V \quad \text{and} \quad \mathrm{E}(G) = E.$$

(2) When there is no risk of ambiguity, we may write an edge $\{x, y\}$ as xy.

(3) The graph G is *finite* if $\mathrm{V}(G)$ is finite, else it is *infinite*.

Definition 11.1.3. Let G, H be undirected graphs,.

(1) We say that H is a *subgraph* of G, or *G contains H* (as a subgraph), if $\mathrm{V}(H) \subseteq \mathrm{V}(G)$ and $\mathrm{E}(H) \subseteq \mathrm{E}(G)$.

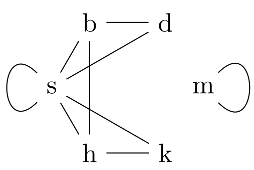
all the vertices and edges come from the bigger graph

(2) A *proper subgraph* of G is a subgraph H of G such that $H \neq G$.

Example 11.1.4. Consider the graph G, where

$$\mathrm{V}(G) = \{\mathrm{b}, \mathrm{d}, \mathrm{h}, \mathrm{k}, \mathrm{m}, \mathrm{s}\},$$
$$\mathrm{E}(G) = \{\mathrm{bd}, \mathrm{bh}, \mathrm{bs}, \mathrm{ds}, \mathrm{hk}, \mathrm{hs}, \mathrm{ks}, \mathrm{mm}, \mathrm{ss}\}.$$

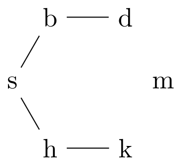
Here is a drawing of G.

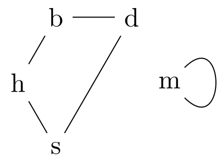


The graphs G_1, G_2, where

$$\mathrm{V}(G_1) = \{\mathrm{b}, \mathrm{d}, \mathrm{h}, \mathrm{k}, \mathrm{m}, \mathrm{s}\}, \qquad \mathrm{V}(G_2) = \{\mathrm{b}, \mathrm{d}, \mathrm{h}, \mathrm{m}, \mathrm{s}\},$$
$$\mathrm{E}(G_1) = \{\mathrm{bd}, \mathrm{bs}, \mathrm{hk}, \mathrm{hs}\}, \qquad \mathrm{E}(G_2) = \{\mathrm{bd}, \mathrm{bh}, \mathrm{ds}, \mathrm{hs}, \mathrm{mm}\},$$

are subgraphs of G. Here are drawings of G_1 and G_2 respectively.



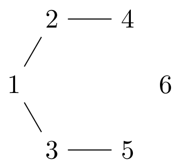


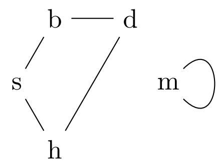
position of h and s are swapped,
if original position, it would look like part of original graph

The graph G_3, G_4, where

$$\mathrm{V}(G_3) = \{1, 2, 3, 4, 5, 6\}, \qquad \mathrm{V}(G_4) = \{\mathrm{b}, \mathrm{d}, \mathrm{h}, \mathrm{m}, \mathrm{s}\},$$
$$\mathrm{E}(G_3) = \{12, 13, 24, 35\}, \qquad \mathrm{E}(G_4) = \{\mathrm{bd}, \mathrm{bs}, \mathrm{dh}, \mathrm{hs}, \mathrm{mm}\},$$

are *not* subgraphs of G because $\mathrm{V}(G_3) \not\subseteq \mathrm{V}(G)$ and $\mathrm{E}(G_4) \not\subseteq \mathrm{E}(G)$. Here are drawings of G_3 and G_4 respectively.





Definition 11.1.5. Let G be an undirected graph, and u, v be vertices in G. A *path* between u and v in G is a subgraph of G of the form

$$(\{x_0, x_1, \ldots, x_\ell\}, \{x_0x_1, x_1x_2, \ldots, x_{\ell-1}x_\ell\}),$$

number of edges

where the x's are all different and $\ell \in \mathbb{N}$, satisfying $u = x_0$ and $v = x_\ell$. Here ℓ is called the *length* of the path. When there is no risk of ambiguity, we may denote the subgraph above by $x_0x_1 \ldots x_\ell$.

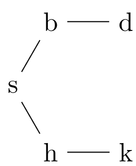
Remark 11.1.6. (1) Informally speaking, a path links two vertices in a graph via a sequence of edges, each joined to the next, that has no repeated vertex.

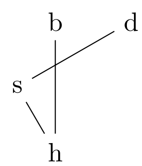
(2) Some consider paths of infinite length. We do not.

Example 11.1.7. Consider the graph G from Example 5.3.8. Define

$$P = \mathrm{dbshk} \quad \text{and} \quad Q = \mathrm{dshb}.$$

Then P is a path of length 4 between d and k in G, and Q is a path of length 3 between b and d in G. Here are drawings of P and Q respectively.



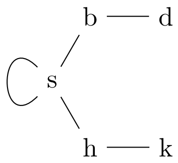


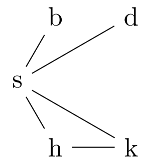
The graphs H_1 and H_2, where

$$H_1 = \mathrm{dbsshk} \quad \text{and} \quad H_2 = \mathrm{bshksd}$$

are *not* paths in G because s is in three edges in H_1, and four edges in H_2. (Note that each vertex in a path is in at most two edges in the path.) Here are drawings of H_1 and H_2

respectively.





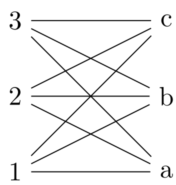
Remark 11.1.8. Consider again the graph G from Example 5.3.8.

(1) The subgraph $(\{s\}, \{\})$, which we may write as s, is a path of length 0 in G. It is essentially the vertex s with no edge, not even a loop.

(2) The subgraph $(\{s, h\}, \{sh\})$, which we may write as sh, is a path of length 1 in G. It is essentially the edge sh.

Figure 11.1: A path of length 0 (left) and a path of length 1 (right)

Exercise 11.1.9. How many paths are there between 1 and 3 in the undirected graph G with the following drawing? ✎ 11a



1 -> x -> 3 : 3 choices
1 -> x -> y -> 3 : not possible
1 -> x -> y -> z -> 3 : 6 choices

y can only be "2" (1 choice)
x can be any of a,b,c (3 choices)
z can then be any of remaining (2 choices)
3 * 1 * 2 = 6 choices

Briefly explain your answer.

Lemma 11.1.10. Let G be an undirected graph and u, v, w be vertices in G. Suppose there are a path P between u and v in G, and a path Q between v and w in G. Then there is a path between u and w in G.

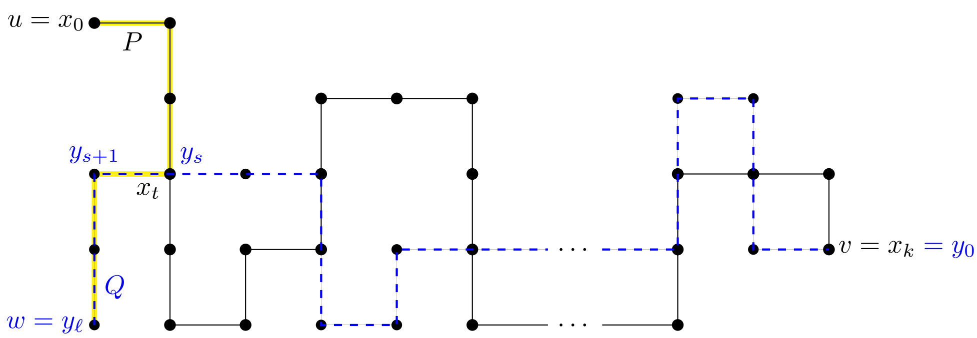


Figure 11.2: Combining paths

Proof. Let $P = x_0x_1 \dots x_k$ and $Q = y_0y_1 \dots y_\ell$, so that $k, \ell \in \mathbb{N}$ and

$$x_0 = u, \quad x_k = v = y_0, \quad w = y_\ell.$$

As $x_k = y_0 \in \mathrm{V}(Q)$, we know some $t \in \{0, 1, \dots, k\}$ satisfies $x_t \in \mathrm{V}(Q)$. Let t be the smallest element of $\{0, 1, \dots, k\}$ such that $x_t \in \mathrm{V}(Q)$. By the smallestness of t, none of $x_0, x_1, \dots, x_{t-1}$ is in $\mathrm{V}(Q)$. So if $s \in \{0, 1, \dots, \ell\}$ such that $x_t = y_s$, then

xt is the first intersection

basically, follow first path until intersection, then follow second path

$$x_0x_1 \dots x_t y_{s+1} y_{s+2} \dots y_\ell$$

is a path between u and w in G. $\square$

Remark 11.1.11. In the proof of Lemma 11.1.10 above, the existence of the required t is guaranteed by the Well-Ordering Principle. In more details, we know k is an element of

$$\{t \in \{0, 1, \dots, k\} : x_t \in \mathrm{V}(Q)\}.$$

This set is thus a nonempty subset of $\mathbb{N}$, and hence must have a smallest element by the Well-Ordering Principle. Similar applications of the Well-Ordering Principle can be found in the proof of Theorem 11.2.5 below.

11.2 Cycles

Definition 11.2.1. Let G be an undirected graph.

(1) A *cycle* in G is a subgraph of G of the form

$$\big(\{x_1, x_2, \dots, x_\ell\}, \{x_1x_2, x_2x_3, \dots, x_{\ell-1}x_\ell, x_\ell x_1\}\big),$$

where the x's are all different and $\ell \in \mathbb{N}_{\geqslant 3}$. Here ℓ is called the *length* of the cycle. When there is no risk of ambiguity, we may denote the subgraph above by $x_1x_2 \dots x_\ell x_1$.

(2) An undirected graph is *cyclic* if it has a loop or a cycle; else it is *acyclic*.

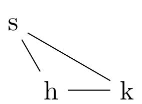
Remark 11.2.2. (1) Informally speaking, a cycle in a graph is a sequence of at least three edges, each joined to the next, and the last joined to the first, that has no repeated vertex.

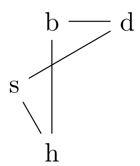
(2) By definition, a cycle has at least three vertices (and thus at least three edges). Therefore, in no sense can a loop be a cycle.

Example 11.2.3. Consider the graph G from Example 5.3.8. Define the graphs C, D by

$$C = \mathrm{shk} \quad \text{and} \quad D = \mathrm{sdbh}.$$

Then C is a cycle of length 3 in G, and D is a cycle of length 4 in G. Here are drawings of C and D respectively.

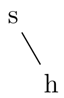


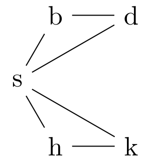


The graphs H_3 and H_4, where

$$H_3 = \mathrm{shs} \quad \text{and} \quad H_4 = \mathrm{bshksdb},$$

are *not* cycles in G because H_3 has only two vertices, and s is in four different edges in H_4. (Note that every cycle by definition has at least three vertices, and each vertex in a cycle is in exactly two edges in the cycle.) Here are drawings of H_3 and H_4 respectively.





Example 11.2.4. (1) The graphs H_1 and H_2 from Example 11.1.7 are both cyclic because the former has a loop {a} and the latter has a cycle C as defined in Example 11.2.3.

(2) The graph G_1 from Example 11.1.4 and the graph Q from Example 11.1.7 are acyclic.

Theorem 11.2.5. An undirected graph with no loop is cyclic if and only if it has two vertices between which there are two distinct paths.

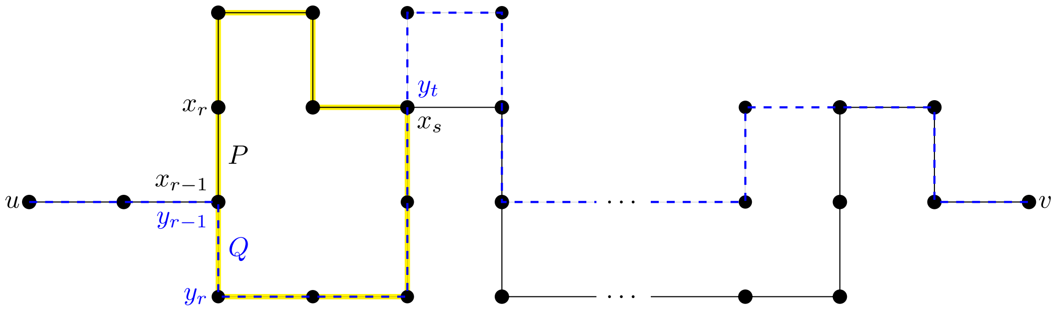


Figure 11.3: Constructing a cycle from two paths between u and v

Proof. Let G be an undirected graph with no loop.

($\Rightarrow$) Assume G is cyclic. According to the definition of cyclic graphs, as G has no loop, it must have a cycle, say,

$$x_1 x_2 \dots x_\ell x_1.$$

From this, we find two paths between x_1 and x_ℓ:

$$x_1 x_\ell \quad \text{and} \quad x_1 x_2 \dots x_\ell$$

These two paths are distinct because the first one has two vertices and the second one has at least three vertices as $\ell \geqslant 3$.

($\Leftarrow$) Let $u, v \in \mathrm{V}(G)$ with two distinct paths between them, say,

$$P = x_0 x_1 \dots x_k \quad \text{and} \quad Q = y_0 y_1 \dots y_\ell,$$

where $x_0 = u = y_0$ and $x_k = v = y_\ell$. If the two paths are of different lengths, then let us use the name Q for the longer of the two paths. This makes $k \leqslant \ell$.

As $P \neq Q$, we know $x_i \neq y_i$ for some $i \in \{0, 1, \dots, k\}$. Let r be the smallest element of $\{0, 1, \dots, k\}$ such that $x_r \neq y_r$. Here $r \neq 0$ because $x_0 = y_0$. So the smallest ness of r tells us $x_{r-1} = y_{r-1}$. It follows that amongst $x_{r-1}, y_{r-1}, x_r, y_r$, only x_{r-1} and y_{r-1} are equal because all the x's are different and all the y's are different.

r is the first position after the two paths diverge

r-1 is divergence point

Recall $x_k = y_\ell$. So some $s \in \{r, r+1, \dots, k\}$ makes $x_s \in \{y_r, y_{r+1}, \dots, y_\ell\}$. Let s be the smallest element of $\{r, r+1, \dots, k\}$ that makes $x_s \in \{y_r, y_{r+1}, \dots, y_\ell\}$. By the smallestness of s, none of $x_r, x_{r+1}, \dots, x_{s-1}$ is equal to any of $y_r, y_{r+1}, \dots, y_\ell$.

s is first convergerence point after divergence

Let $t \in \{r, r+1, \ldots, \ell\}$ that makes $x_s = y_t$. As $x_r \neq y_r$, either $s \neq r$ or $t \neq r$. So

$$x_{r-1}x_r \ldots x_s y_{t-1} y_{t-2} \ldots y_r y_{r-1}$$ Xr-1 = Yr-1

has at least three vertices, and thus is a cycle in G. □

Explanation of why $x_i \neq y_i$ for some $i \in \{0, 1, \ldots, k\}$ in the proof above (extra material). Suppose not. Then $x_i = y_i$ for all $i \in \{0, 1, \ldots, k\}$. In particular, this tells us $y_\ell = v = x_k = y_k$. So $k = \ell$ as the y's are all different. This implies $P = Q$, which contradicts our condition that P and Q are distinct, as required. ■

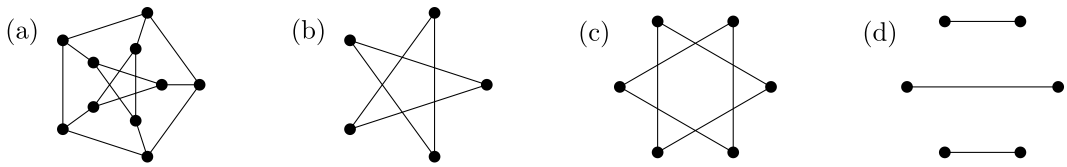
11.3 Connectedness

Definition 11.3.1. An undirected graph is *connected* if there is a path between any two vertices.

Example 11.3.2. (1) The graphs H_1, H_2 from Example 11.1.7 are connected, as one can verify exhaustively.

(2) The graphs G_1, G_2 from Example 11.1.4 are not connected, because there is no path between d and m in G_1 and in G_2.

Exercise 11.3.3. Which of the following are drawings of connected graphs? ✎ 11b



not connected not connected

Definition 11.3.4. Let G be an undirected graph. A *connected component* of G is a maximal connected subgraph of G, i.e., it is a connected subgraph H of G such that no connected subgraph of G contains H as a proper subgraph. ∀ connected H*⊆g (H*⊇H → H*=H)

if subgraph H* is bigger than H , then H* = H

Example 11.3.5. Consider the graph G from Example 11.1.4. The following are respectively drawings of two connected components H_s and H_m of G.

The graph G_1 from Example 11.1.4 is not a connected component of G because it is not connected. The graph D from Example 11.2.3 is not a connected component of G because it is a subgraph of the connected subgraph H_s of G above with two fewer edges ss and bh.

making it not maximal!

Proposition 11.3.6. Let G be an undirected graph. Then every vertex v in G is in some connected component of G.

Proof. Define the subgraph H of G by setting

$$\begin{aligned} \mathrm{V}(H) &= \{x \in \mathrm{V}(G) : \text{there is a path between } v \text{ and } x \text{ in } G\} \quad \text{and} \\ \mathrm{E}(H) &= \{xy \in \mathrm{E}(G) : x, y \in \mathrm{V}(H)\}. \end{aligned}$$

path of length 0

We know $v \in \mathrm{V}(H)$ because $(\{v\}, \{\})$ is a path between v and v. To finish the proof, we show that H is a connected component of G.

For connectedness, let $a, b \in \mathrm{V}(H)$. By Tutorial Exercise 11.6, in H, there is a path between v and a, and there is a path between v and b. Lemma 11.1.10 then gives us a path between a and b in H.

Let H^+ be a connected subgraph of G which contains H as a subgraph. We want to show that $H^+ = H$. The definition of subgraphs tells us already $\mathrm{V}(H) \subseteq \mathrm{V}(H^+)$ and $\mathrm{E}(H) \subseteq \mathrm{E}(H^+)$. So it remains to show $\mathrm{V}(H^+) \subseteq \mathrm{V}(H)$ and $\mathrm{E}(H^+) \subseteq \mathrm{E}(H)$.

Take any $x \in \mathrm{V}(H^+)$. As $v \in \mathrm{V}(H) \subseteq \mathrm{V}(H^+)$ and H^+ is connected, there is a path between v and x in H^+, hence in G. So $x \in \mathrm{V}(H)$ by the definition of $\mathrm{V}(H)$. shows that any vertex in H+ must be in H

Take any $xy \in \mathrm{E}(H^+)$. Then $x, y \in \mathrm{V}(H^+) \subseteq \mathrm{V}(H)$ by the previous paragraph. As $xy \in \mathrm{E}(H^+) \subseteq \mathrm{E}(G)$, the definition of $\mathrm{E}(H)$ tells us $xy \in \mathrm{E}(H)$. $\square$

Theorem 11.3.7. Let u, v be vertices in an undirected graph G. Then there is a path between u and v in G if and only if there is a connected component of G that has both u and v in it.

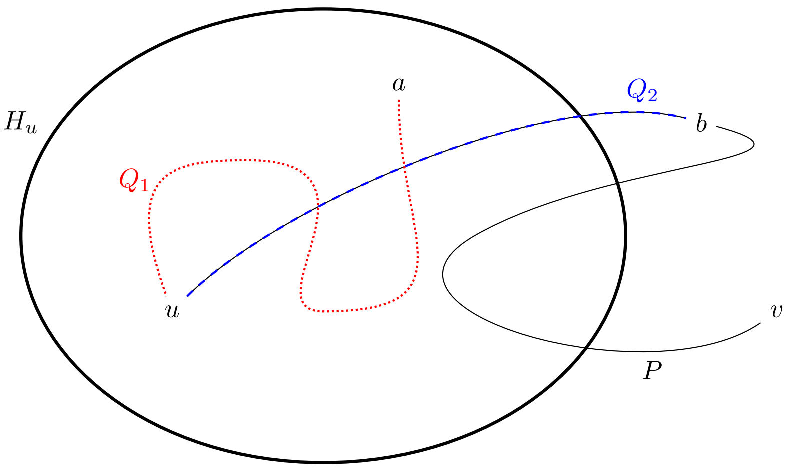


Figure 11.4: A connected component plus a path

Proof. ($\Rightarrow$) Assume there is a path between u and v in G, say,

$$P = x_0 x_1 \ldots x_\ell,$$

where $u = x_0$ and $v = x_\ell$. Use Proposition 11.3.6 to find a connected component H_u of G that has u in it. Define $H = \big(\mathrm{V}(H_u) \cup \mathrm{V}(P), \mathrm{E}(H_u) \cup \mathrm{E}(P)\big)$. We claim that H is a connected subgraph of G. From this, we will deduce $H = H_u$ using the maximality of the connected component H_u; thus $v \in \mathrm{V}(P) \subseteq \mathrm{V}(H) = \mathrm{V}(H_u)$.

Take $a, b \in \mathrm{V}(H)$. As $\mathrm{V}(H) = \mathrm{V}(H_u) \cup \mathrm{V}(P)$, we have three cases.

Case 1: suppose $a, b \in \mathrm{V}(H_u)$. As H_u is connected, there is a path between a and b in H_u, hence in H. both points in connected component

Case 2: suppose $a, b \in \mathrm{V}(P)$. Say $a = x_r$ and $b = x_s$, where $r \leqslant s$. Then a path between a and b in H is $x_r x_{r+1} \ldots x_s$. both points on path

Case 3: suppose one of a, b is in H_u and another is in P. Say $a \in \mathrm{V}(H_u)$ and $b \in \mathrm{V}(P)$. On the one hand, as H_u is connected, and $a, u \in \mathrm{V}(H_u)$, one can find a path, say Q_1, between a and u in H_u, hence in H. On the other hand, if $b = x_t$ where $t \in \{0, 1, \ldots, \ell\}$, then $Q_2 = x_0 x_1 \ldots x_t$ a path between u and b in H. Combining Q_1 and Q_2 using Lemma 11.1.10, we get a path between a and b in H.

one point on connected component, one point on path

(⇐) Assume there is a connected component, say H, of G that has both u and v in it. As H is connected, there is a path between u and v in H, hence in G. □

Tutorial exercises

An asterisk (*) indicates a more challenging exercise.

11.1. Let us look into the notion of complete graphs. Fix $n \in \mathbb{N}$.

Definition. A *complete graph* on n vertices, denoted K_n, is an undirected graph, with exactly n vertices and with no loop, in which there is an edge between any pair of distinct vertices.

(a) Draw K_1, K_2, K_3, and K_4. There is no need to label the vertices in your drawings.

(b) How many edges are there in K_n? Explain your answer briefly.

11.2. Consider the undirected graphs G with $\mathrm{V}(G) = \{\mathrm{a}, \mathrm{b}, \mathrm{c}\}$.

(a) How many such graphs are there?

(b) How many of them have no loop?

(c) How many of them have a cycle?

(d) How many of them are cyclic?

Briefly explain your answers.

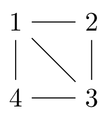
11.3. Two graphs of the same shape are considered unequal if their vertices are labelled differently. This is sometimes undesirable. Sometimes we want to consider two graphs as the same if (and only if) they are equal after a relabelling of the vertices. This notion of sameness is called *isomorphism* in mathematics.

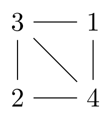
Definition. An *isomorphism* from an undirected graph G_1 to an undirected graph G_2 is a bijection $f\colon \mathrm{V}(G_1) \to \mathrm{V}(G_2)$ such that for all $x, y \in \mathrm{V}(G_1)$,

$$\{x, y\} \in \mathrm{E}(G_1) \quad \Leftrightarrow \quad \{f(x), f(y)\} \in \mathrm{E}(G_2).$$

An undirected graph G_1 is said to be *isomorphic* to an undirected graph G_2 if there is an isomorphism from G_1 to G_2.

Example. The following are drawings of isomorphic graphs:



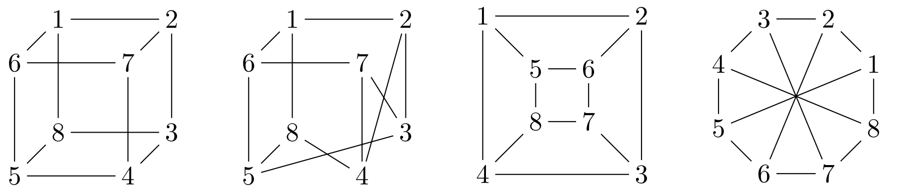


One isomorphism from the left graph to the right graph is the function $f\colon \{1, 2, 3, 4\} \to \{1, 2, 3, 4\}$ satisfying

$$f(1) = 3, \quad f(2) = 1, \quad f(3) = 4, \quad f(4) = 2.$$

It can be verified as in Tutorial Exercise 8.3 that the isomorphism relation on a set of graphs is an equivalence relation.

(a) Which of the graphs drawn below are isomorphic? Which are not isomorphic?



(i) (ii) (iii) (iv)

Briefly explain your answer.

(b) Let $n \in \{2, 3, 4\}$. How many undirected graphs are there with exactly n vertices and with no loop if we count isomorphic graphs as one?

11.4. The aim of this exercise is to investigate in what sense the connected components form a partition of an undirected graph.

(a) Is it true that, for all undirected graphs G with at least one vertex, the set $\{\mathrm{V}(H) : H \text{ is a connected component of } G\}$ is a partition of $\mathrm{V}(G)$?

(b) Is it true that, for all undirected graphs G with at least one edge, the set $\{\mathrm{E}(H) : H \text{ is a connected component of } G\}$ is a partition of $\mathrm{E}(G)$?

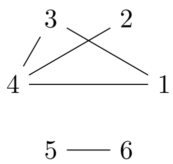
Prove that your answers are correct.

11.5. In this exercise, we look into the relationship between an undirected graph with no loop and its complement, in the sense defined below.

Definition. The *complement* of an undirected graph G with no loop, denoted $\overline{G}$, is the undirected graph with no loop defined by $\mathrm{V}(\overline{G}) = \mathrm{V}(G)$ and, for all distinct $x, y \in \mathrm{V}(\overline{G})$,

$$xy \in \mathrm{E}(\overline{G}) \quad \Leftrightarrow \quad xy \notin \mathrm{E}(G).$$

(a) Consider the graph whose drawing is as follows.



Draw the complement of this graph.

(b)* Prove that, for all undirected graphs G with at least one vertex but with no loop, either G is connected or $\overline{G}$ is connected.

11.6. The aim of this exercise is to fill a small gap in the proof of Proposition 11.3.6: when showing connectedness there, it is clear that we get the desired paths in the bigger graph, but it is less clear why we get the same in the subgraph. Here we analyze the situation by splitting the subgraph into layers.

(Induction corner) Fix an undirected graph G and $v \in \mathrm{V}(G)$. For each $n \in \mathbb{N}$, define a subgraph H_n of G by setting

$\mathrm{V}(H_n) = \{x \in \mathrm{V}(G) : \text{there is a path of length at most } n \text{ between } v \text{ and } x \text{ in } G\}$, and
$\mathrm{E}(H_n) = \{xy \in \mathrm{E}(G) : x, y \in \mathrm{V}(H_n)\}$.

Prove by induction that, for every $n \in \mathbb{N}$ and every $x \in \mathrm{V}(H_n)$, there is a path between v and x in H_n.

Extra exercises

11.7. We verify in this exercise that a connected component of an undirected graph must inherit all the possible edges from that graph.

Let G be an undirected graph, and H be a connected component of G. Prove that

$$\mathrm{E}(H) = \{xy \in \mathrm{E}(G) : x, y \in \mathrm{E}(H)\}.$$

11.8. Intuitively, undirected graphs that have many edges relative to the number of vertices must be connected. In this exercise, we investigate what "many" may mean here.

(a)* Let $n \in \mathbb{N}$ and G be an undirected graph with exactly n vertices and with no loop. Prove that, if $|\mathrm{E}(G)| > \binom{n-1}{2}$, then G is connected.

(b) How many connected undirected graphs are there with exactly 4 vertices and with no loop in which the number of edges is at most $\binom{4-1}{2}$ if we count isomorphic graphs as one?