Sets

A **set** is a collection of things—any things—called its **elements** or **members**

- The set, call it A, whose elements are *Shakespeare, Tasmania, and Monday*
- The set, call it B, of all even positive numbers

Notation

$A = \{$*Shakespeare*, *Tasmania*, *Monday*$\}$ list all elements and enclose with $\{\ \}$

$B = \{2, 4, 6, 8, \dots\}$ **not entirely clear!**

$B = \{x \mid x$ is an even positive number$\}$ specify a **common property** of elements

Write $a \in S$ to say that a is an element of S (from Greek ε)

Write $a \notin S$ to say that a is not an element of S

$8 \in B$ $7 \notin B$ $C = \{x \mid x \in \mathbb{N}$ and $x \notin B\}$

where $\mathbb{N}$ is the set of all natural numbers. What is C?

Four fundamental features of sets

- **A set must be distinguished from its description**

  For instance, the following descriptions define the same set:

$$\{2,3,4\} \quad \{3,2,4\} \quad \{2,2,3,4,4\} \quad \{x \in \mathbb{N} \mid 2 \le x \le 4\} \quad \{y \in \mathbb{N} \mid 1 < y < 5\}$$

- **All elements of a set are distinct**

  In other words, no element may 'occur' more than once in a set

  We do not distinguish between $\{3,2,4\}$ and $\{2,2,3,4,4\}$

- **The elements of a set are not ordered in any way**

  We do not distinguish between $\{3,2,4\}$ and $\{2,3,4\}$

- **A set can be an element of another set**

  For example, $\{0,\{0\}\}$ has two elements: 0 and $\{0\}$

Important sets

$\emptyset = \{ \}$	**empty set**, the set with no elements
$\mathbb{N} = \{0, 1, 2, 3, \dots\}$	natural numbers
$\mathbb{N}^+ = \{1, 2, 3, \dots\}$	positive natural numbers
$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \dots\}$	integer numbers (from German Zahlen)
$\mathbb{Q} = \{x/y \mid x, y \in \mathbb{Z},\ y \neq 0\}$	rational numbers (from Italian quoziente—quotient)
$\mathbb{R} = \{\text{decimals}\}$	real numbers

Thus, for any x, we have $\boxed{x \notin \emptyset}$

On the other hand, the empty set can be an element of another set:

for example, $\boxed{\emptyset \in \{5, \textit{Zoe}, \emptyset\}}$

Describing sets by properties

Describing an infinite set by listing using ... might sometimes be possible (like we did with $\mathbb{N}$), but even then a bit imprecise.

But listing, say, rational numbers can be quite tricky. Instead of listing elements, we can describe a property that the elements of the set satisfy.

If $\mathcal{P}$ is a property, then the set S whose elements have the property $\mathcal{P}$ is denoted by

$$S = \{x \mid x \text{ has property } \mathcal{P}\}$$

We read this as

'S **is the set of all** x **such that** x **has property** $\mathcal{P}$'

Or, if we also know that all the elements in S come from a larger set A, then we can write

$$S = \{x \in A \mid x \text{ has property } \mathcal{P}\}$$

Describing sets by properties: an example

Let *Odd* be the set of all odd integers. Then we can describe *Odd* in several ways:

$$
\begin{aligned}
\textit{Odd} &= \{\ldots, -5, -3, -1, 1, 3, 5, \ldots\} \\
&= \{x \mid x \text{ is an odd integer}\} \\
&= \{x \in \mathbb{Z} \mid x \text{ is odd}\} \\
&= \{x \mid x = 2k + 1 \text{ for some integer } k\} \\
&= \{x \mid x = 2k + 1 \text{ for some } k \in \mathbb{Z}\}
\end{aligned}
$$

We can also use expressions for the elements on the left hand side:

$$
\begin{aligned}
\textit{Odd} &= \{2k + 1 \mid k \text{ is an integer}\} \\
&= \{2k + 1 \mid k \in \mathbb{Z}\}
\end{aligned}
$$

Subsets

A set B is a **subset** of a set A if every element of B is an element of A

Notation: $B \subseteq A$. Also say: B is **included** in A

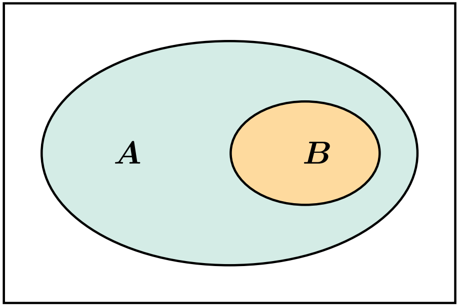


Figure 1: **Venn diagram** of $B \subseteq A$.

John Venn was a 19th-century British philosopher and mathematician who introduced the Venn diagram in 1881

$$\{3,4,5\} \subseteq \{1,5,4,2,1,3\} \qquad \{3,3,5\} \subseteq \{3,5\} \qquad \{5,3\} \subseteq \{3,5\}$$

Equal sets, proper subsets

Two sets A and B are **equal** if they have exactly the same elements

Notation: $A = B$

$$A = B \quad \text{iff} \quad A \subseteq B \ \text{ and } \ B \subseteq A$$

$\{1\} = \{1,1,1\}$ $\qquad$ $\{1\} \neq \{\{1\}\}$ $\qquad$ $\{0,2,8\} = \{\sqrt{4}, 0/5, 2^3\}$

B is a **proper subset** of A if $B \subseteq A$ and $A \neq B$

Notation: $B \subset A$ or $B \subsetneqq A$

Also say: B is **properly included** in A

$\{1\} \subset \{1,1,2\}$ $\qquad$ $\{1\} \not\subseteq \{\{1\}\}$

Exercise: Suppose $S = \{s, \emptyset\}$. Which of the following statements are true?

(a) $s \in S$ $\qquad$ (b) $\{s\} \in S$ $\qquad$ (c) $\emptyset \subseteq S$ $\qquad$ (d) $\emptyset \in S$

Set operations: union

The **union** of sets A and B is the set $A \cup B = \{x \mid x \in A \text{ or } x \in B\}$

$A \cup B$ is the set consisting of those elements that are in A or in B or both

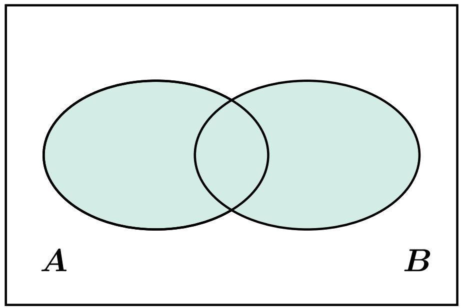


Figure 2: Venn diagram of $A \cup B$.

Suppose $A = \{4, 7, 8\}$ and $B = \{10, 4, 9\}$

Then $A \cup B = \{4, 7, 8, 9, 10\}$

Set operations: intersection

The **intersection** of sets A and B is the set $A \cap B = \{x \mid x \in A \text{ and } x \in B\}$

$A \cap B$ is the set consisting of all elements which are both in A and in B

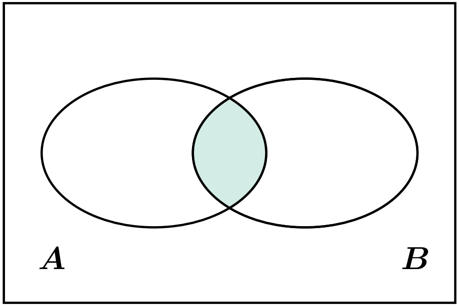


Figure 3: Venn diagram of $A \cap B$.

Suppose $A = \{4, 7, 8\}$ and $B = \{10, 4, 9\}$

Then $A \cap B = \{4\}$

If $A \cap B = \emptyset$ then A and B are called **disjoint**

Set operations: relative complement

The **complement** of a set B **relative** to a set A is the set

$$A - B = \{x \mid x \in A \text{ and } x \notin B\}$$

$A - B$ is also called the **difference** of A and B

$A - B$ is the set of all elements that belong to A but not to B

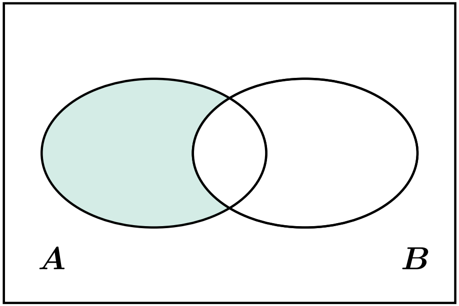


Figure 4: Venn diagram of $A - B$.

If $A = \{4, 7, 8\}$ and $B = \{10, 4, 9\}$ then $A - B = \{7, 8\}$

Set operations: (absolute) complement

In certain contexts we may regard all sets under consideration as being subsets of some given **universal set** U.

For instance, if we are investigating properties of the real numbers $\mathbb{R}$ (and subsets of $\mathbb{R}$), then we may take $\mathbb{R}$ as our universal set

Given a universal set U and $A \subseteq U$, the **complement** of A (in U) is the set

$$-A = U - A = \{x \in U \mid x \notin A\}$$

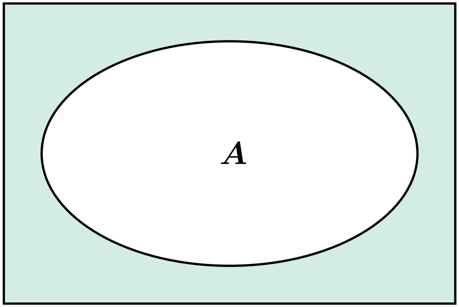


Figure 5: Venn diagram of $-A$

Set operations: powerset

The **powerset** of a set A is defined to be the set of **all** subsets of A

Notation: $\textbf{Pow}(A) \;=\; \{X \mid X \subseteq A\}$

Examples:

1. Let $A = \{2\}$. Then
$$\textbf{Pow}(A) \;=\; \{\emptyset,\; \{2\}\}$$

2. Let $B = \{1, 2, 3\}$. Then
$$\textbf{Pow}(B) \;=\; \{\emptyset,\; \{1\},\; \{2\},\; \{3\},\; \{1,2\},\; \{2,3\},\; \{1,3\},\; \{1,2,3\}\}$$

3. Let $C = \emptyset$. Then
$$\textbf{Pow}(C) \;=\; \{\emptyset\} \qquad (\neq \emptyset)$$

More common notation: $\textbf{Pow}(A) \;=\; 2^A$

Why?

Because when A has n elements, then $\textbf{Pow}(A)$ has 2^n elements

Important equalities

- Associative laws:

$$A \cup (B \cup C) = (A \cup B) \cup C, \quad A \cap (B \cap C) = (A \cap B) \cap C$$

- Commutative laws:

$$A \cup B = B \cup A, \quad A \cap B = B \cap A$$

- Identity laws (where U is the universal set):

$$A \cup \emptyset = A, \quad A \cup U = U, \quad A \cap U = A, \quad A \cap \emptyset = \emptyset$$

- Distributive laws:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

- Complement laws (where U is the universal set):

$$A \cup -A = U, \quad -U = \emptyset, \quad -(-A) = A, \quad A \cap -A = \emptyset, \quad -\emptyset = U$$

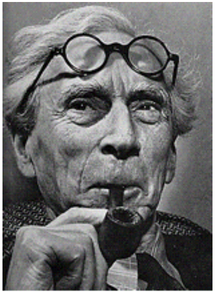
- De Morgan's laws:

$$-(A \cup B) = -A \cap -B, \quad -(A \cap B) = -A \cup -B$$

Russell's paradox (1901)

Russell's paradox shows that the 'object' $\{x \mid P(x)\}$ is not always meaningful.

Consider the set $A = \{B \mid B \notin B\}$



Give an example of an element of A

Problem: do we have $A \in A$?

For every set C, denote by $P(C)$ the statement $C \notin C$

Then $A = \{B \mid P(B)\}$.

- Suppose $A \in A$. Then not $P(A)$. Therefore, we must have $A \notin A$.
- But if $A \notin A$, then $P(A)$. Therefore, $A \in A$, which is a contradiction

Visit also http://plato.stanford.edu/entries/russell-paradox/

Popular version: the barber paradox

Suppose there is a town with just one male barber. According to law in this town, the barber shaves all and only those men in town who do not shave themselves. Who shaves the barber?

- if the barber does shave himself, then the barber (himself) must not shave himself
- if the barber does not shave himself, then the barber (himself) must shave himself

Sequences, tuples, and Cartesian products

A **sequence** of objects is a list of these objects **taken in a certain order**

The sequences $(1, 2, 3)$ $(2, 1, 3)$ $(3, 1, 2)$ are **different**

The sets $\{1, 2, 3\}$ $\{2, 1, 3\}$ $\{3, 1, 2\}$ are **the same**

Finite sequences are called **tuples**. A sequence with k elements is a **k-tuple**

A 2-tuple is also called a **pair**

The **Cartesian product** $A \times B$ of sets A and B is the *set* of all pairs (a, b) where $a \in A$ and $b \in B$

$$A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}$$

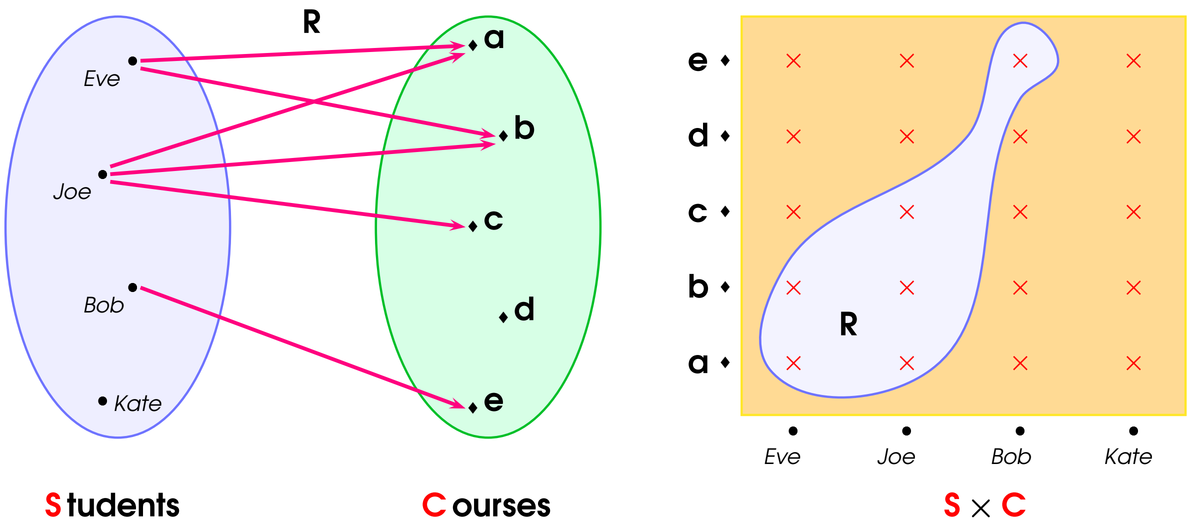
Example. Let $A = \{1, 2\}$ and $B = \{a, b, c\}$. Then

$$A \times B = \{(1, a), (2, a), (1, b), (2, b), (1, c), (2, c)\}$$

$\mathbb{R} \times \mathbb{R}$ is the **Euclidean plane**. What is $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$?

Binary relations

Let S be the set of students at Birkbeck, and C the set of available courses. The registration database or the relationship **'registered for'** can be represented as the set $\boxed{R = \{(s, c) \in S \times C \mid s \text{ registered for } c\}}$



Binary relations: definitions and examples

A **binary relation** between two sets A and B is a subset R of the Cartesian product $A \times B$. If $A = B$, then R is a **relation on A**

If $(x, y) \in R$ then we say that **x is R-related to y** and write xRy

– 'Smaller than' on $\mathbb{Z}$ $\quad \boxed{< \;=\; \{(x,y) \in \mathbb{Z} \times \mathbb{Z} \mid x < y\}}$

$$-1 < 0, \quad -1 < 1, \quad 0 < 2, \ldots \quad \text{but} \quad 1 \not< 0, \quad 1 \not< -1, \quad 2 \not< 0, \ldots$$

– 'Smaller than or equal to' on $\mathbb{Z}$ $\quad \boxed{\leq \;=\; \{(x,y) \in \mathbb{Z} \times \mathbb{Z} \mid x \leq y\}}$

$$-1 \leq -1, \quad -1 \leq 1, \quad 0 \leq 0, \ldots \quad \text{but} \quad 1 \not\leq 0, \quad 1 \not\leq -1, \quad 2 \not\leq 0, \ldots$$

Note that $\quad \boxed{< \;\subsetneq\; \leq}$

– 'Input–output' relation IO_P for a given computer program P

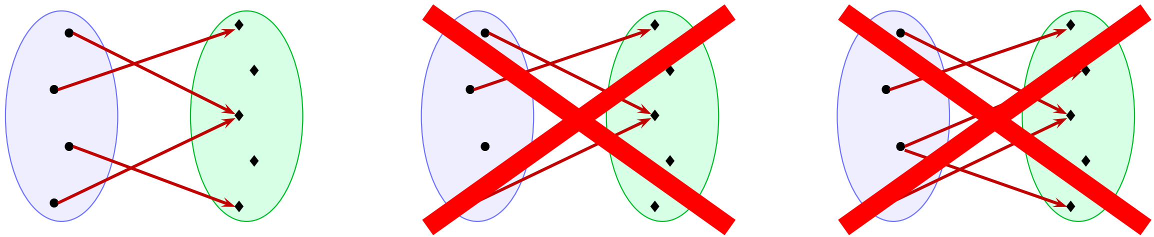
Let I be the set of possible inputs for P, and O the set of possible outputs

$$\boxed{\mathsf{IO}_P \;=\; \{(x,y) \in I \times O \mid P(x) = y\}}$$

$x \,\mathsf{IO}_P\, y$ iff given x as an input, P returns y

Functions

A **function** from a set A to a set B is a binary relation $R \subseteq A \times B$ in which **every element** of A is R-related to a **unique** element of B, or, in other words: for each $a \in A$, there is precisely one pair of the form (a, b) in R.

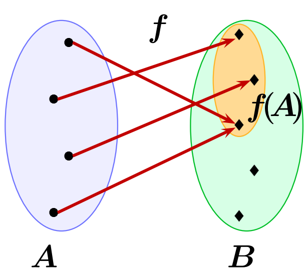


Which of the following relations are functions?

- $\{(x, y) \in \mathbb{N} \times \mathbb{N} \mid y = x - 1\}$
- $\{(x, y) \in \mathbb{Q} \times \mathbb{Q} \mid y = x - 1\}$
- $\{(x, y) \in \mathbb{N} \times \mathbb{R} \mid y^2 = x\}$

Notation

- Let f be a function from a set A to a set B. Since for each $x \in A$ there exists a uniquely determined $y \in B$ with $(x, y) \in f$, we write $\boxed{y = f(x)}$ and refer to $f(x)$ as the **image** of x under f
- We write $\boxed{f : A \to B}$ to indicate that f is a function from A to B
- A is called the **domain** of f. B is called the **codomain** of f
- The **range** of f is the set $\boxed{f(A) = \{f(x) \mid x \in A\}}$



- **every** element of the domain has to be mapped somewhere in the codomain
- but **not everything** in the codomain has to be a value of a domain element
- one element **cannot** be mapped to 2 different places
- but it **can** happen that 2 different elements are mapped to the same place

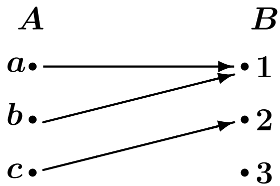
Different ways of describing functions

Example: Let $A = \{a, b, c\}$ and $B = \{1, 2, 3\}$.

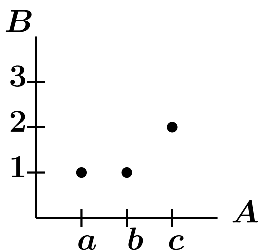
- We can describe a function $f : A \to B$ by listing all of its associations:

$$f(a) = 1, \qquad f(b) = 1, \qquad f(c) = 2.$$

- We can describe the same f by drawing points and arrows:



- We can describe the same f by drawing its **graph**:



Functions and not functions

Let H be the set of all humans (alive or dead).

Question: which of the following $H \to H$ associations are functions?

- **$f(x)$ is a parent of x**

  This f is NOT a function, because people have two parents.

- **$f(x)$ is the mother of x**

  This f is an $H \to H$ function, because each person has exactly one mother.

- **$f(x)$ is the oldest child of x**

  This f is NOT a $H \to H$ function, because some people have no children.

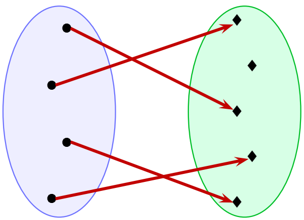
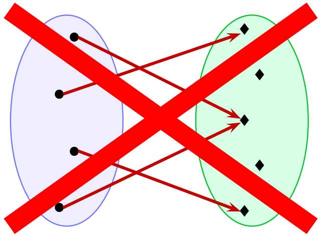
- **$f(x)$ is the set of all children of x**

  Though this f is a function, it is NOT a $H \to H$ function, because each person is associated with a **set** of people rather than one person.

  (This f is an $H \to \mathbf{Pow}(H)$ function.)

Injective functions

A function $f : A \to B$ is called an **injective** (or '**one-to-one**') function if
for all $a_1, a_2 \in A$, if $f(a_1) = f(a_2)$ then $a_1 = a_2$

This is logically equivalent to the implication $\boxed{a_1 \neq a_2 \;\Rightarrow\; f(a_1) \neq f(a_2)}$

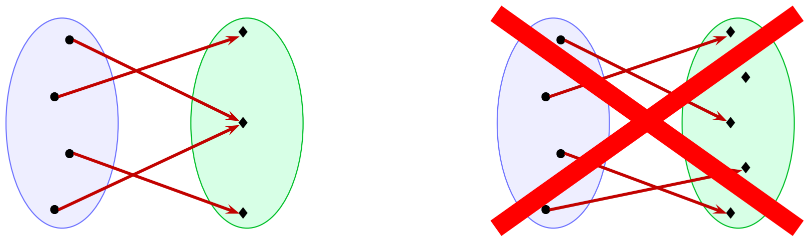
In other words, different inputs give different outputs.

Examples:

- $f : \mathbb{Z} \to \mathbb{Z}$ given by $f(x) = x^2$ is not injective
- $h : \mathbb{Z} \to \mathbb{Z}$ given by $h(x) = 2x$ is injective

Surjective and bijective functions

$f : A \to B$ is **surjective** (or '**onto**') if the range of f coincides with the codomain of f, that is, if for every $b \in B$ there exists $a \in A$ with $b = f(a)$.

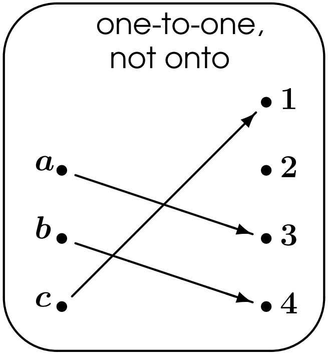


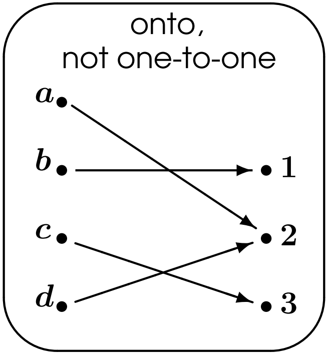
Examples:

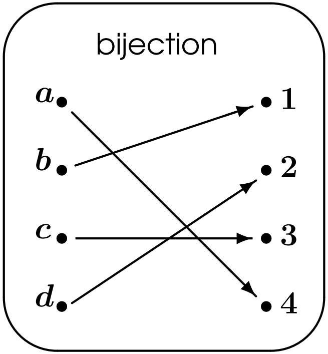
- $f : \mathbb{Z} \to \mathbb{Z}$ given by $f(x) = x^2$ is not surjective
- $h : \mathbb{Z} \to \mathbb{Z}$ given by $h(x) = 2x$ is not surjective
- $g : \mathbb{Q} \to \mathbb{Q}$ given by $g(x) = 2x$ is surjective and injective

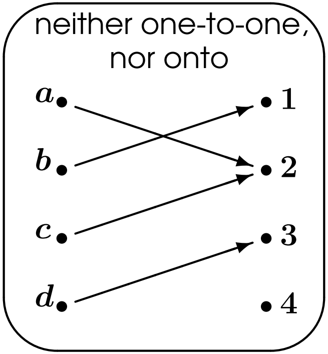
We call f **bijective** or a **one-to-one correspondence** if f is both injective and surjective

Examples


one-to-one,
not onto
a
b
c
1
2
3
4


onto,
not one-to-one
a
b
c
d
1
2
3


bijection
a
b
c
d
1
2
3
4


neither one-to-one,
nor onto
a
b
c
d
1
2
3
4

Bijections and cardinality

If A and B are **finite** sets, then there exists a bijection between A and B iff A and B have the same number of elements

Sets A and B have the same **cardinality** if there is a bijection from A to B

In this case we write $|A| = |B|$.

A has cardinality **strictly greater** than the cardinality of B if there is an injective function from B into A, but A and B do not have the same cardinality

In this case we write $|A| > |B|$.

According to a legend, Tamerlane (1336–1405), during one of his campaigns, ordered all of his warriors to put one stone onto the pile on their way to the battle and upon their return pick a stone off the pile. This allowed him to deduce how many warriors he had lost in the battle. A stone barrow as a tomb to the lost, which they themselves had constructed with their own hands.

Bijections and cardinality (cont.)

Examples:

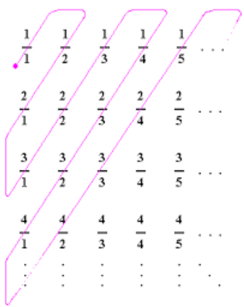
- $f : \mathbb{N} \to \mathbb{Z}$ given by $\quad f(x) = \begin{cases} -x/2 & \text{if } x \text{ is even} \\ (x+1)/2 & \text{if } x \text{ is odd} \end{cases}$

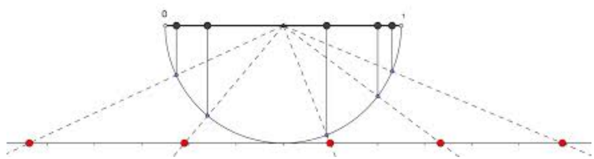
  is a bijection. So $\mathbb{N}$ and $\mathbb{Z}$ are of the same cardinality (or $|\mathbb{N}| = |\mathbb{Z}|$)

> An infinite set A is **countable** if it has the same cardinality as $\mathbb{N}$.
>
> An infinite set A is **uncountable** if it is not countable.

- Are the following sets countable?

  (a) $\mathbb{Q}$ (b) $(0, 1) = \{x \in \mathbb{R} \mid 0 < x < 1\}$ (c) $\mathbb{R}$





$\mathbb{R}$ is uncountable: Cantor's diagonal argument

Suppose, on the contrary, that the set of real numbers from $(0, 1)$ is **countable**

Then we can **enumerate** all the infinite decimal fractions of the form $0.d_1 d_2 \dots$

Let us write such an enumeration as an infinite table

$$
\begin{array}{l}
0.\ d_{11}\ d_{12}\ d_{13}\ d_{14}\ \dots \\
0.\ d_{21}\ d_{22}\ d_{23}\ d_{24}\ \dots \\
0.\ d_{31}\ d_{32}\ d_{33}\ d_{34}\ \dots \\
0.\ d_{41}\ d_{42}\ d_{43}\ d_{44}\ \dots \\
\dots
\end{array}
$$

For each $n = 1, 2, \dots$ we choose a digit c_n that is different from d_{nn} and not equal to 9, and consider the real number

$$0.\ c_1\ c_2\ c_3\ c_4\ \dots$$

By construction, this number is different from every member of the given table

$\longrightarrow$ $(0, 1)$ and $\mathbb{R}$ cannot be countable

Not examinable

Cantor's Theorem: $|A| \lneqq |2^A|$, for every set A

The set of all subsets of any set A has a strictly greater cardinality than A itself

Why? Let $f : A \to \mathbf{Pow}(A)$ be any function.

Then, for every $x \in A$, $f(x)$ is a subset of A.

Now take the following subset D of A: $\boxed{D = \{x \in A \mid x \notin f(x)\}}$

We show that **D is not in the range of f**, that is, for every $x \in A$, $D \neq f(x)$

thus, for $A = \{1, 2\}$ and $f(1) = \{2\}$, $f(2) = \emptyset$, we have $D = \{1, 2\}$

Indeed, x 'distinguishes' the sets D and $f(x)$:

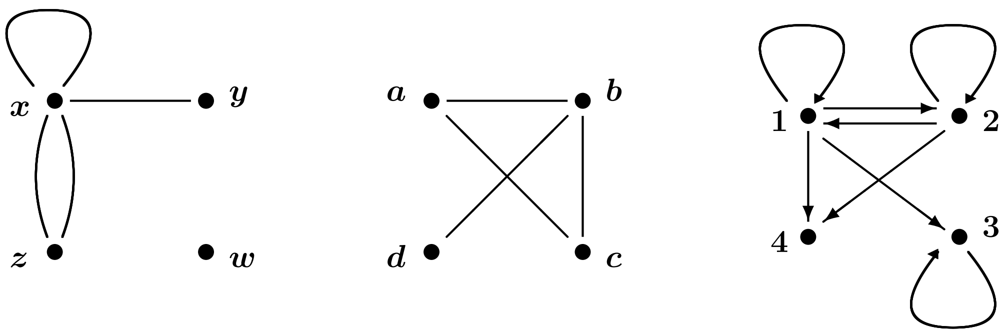
- If $x \in D$, then x should have the property describing D, so $x \notin f(x)$.
- If $x \in f(x)$, then the property describing D does not hold for x, so $x \notin D$.

As D is not in the range of f, we obtain that f is **not onto**,

and so **can't be a bijection**

Graphs

Graphs are drawings with dots and (not necessarily straight) lines or arrows.



The dots are called **vertices** (or **nodes**).

The lines or arrows are called **edges**.

Different kinds of graphs

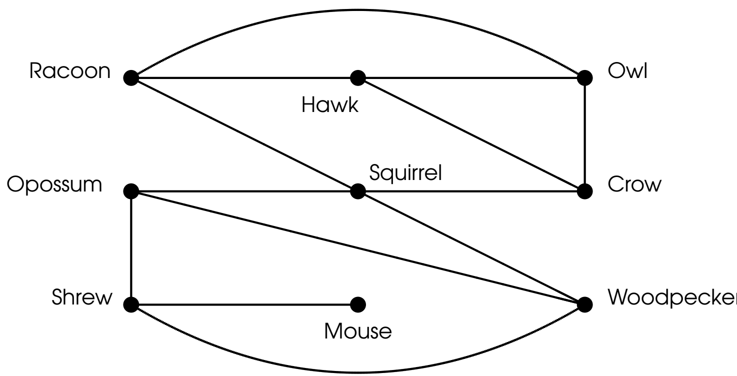
Type	Edges	Multiple edges	Loop edges
(simple) graph	undirected	no	no
multigraph	undirected	yes	yes
directed graph	directed	no	yes
...	...	...	...

Because graphs have applications in a variety of disciplines,
many different terminologies of graph theory have been introduced.

Example 1: Niche overlap graphs in ecology

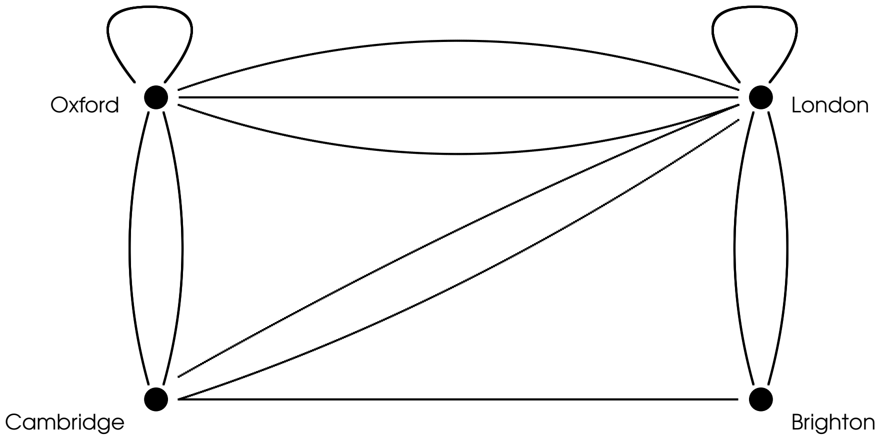
Competitions between species in an ecosystem can be modelled using a **niche overlap graph**:

Each species is represented by a vertex. An edge connects two vertices if the two species represented by these vertices compete (that is, some of the food resources they use are the same).



$\rightsquigarrow$ **simple graph**

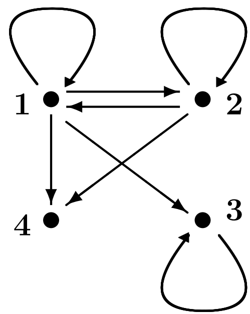
Example 2: Road networks



$\rightsquigarrow$ **multigraph**

Example 3: Representing binary relations

$$R = \{(1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,4), (3,3)\}.$$



$\leadsto$ **directed graph**

Undirected graphs: basic terminology

If there is an edge e between vertices u and v, we say that

- u and v are **adjacent**, and
- e is **incident with** u and v.

The **degree** of a vertex is the number of edges incident with it.

- A vertex of degree zero is called **isolated**.

  So an isolated vertex is not adjacent to any vertex.

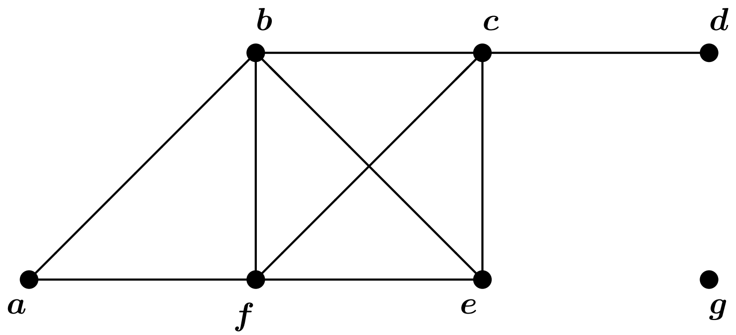
- A vertex of degree one is called **pendant**.

  So a pendant vertex is adjacent to exactly one other vertex.

Handshaking theorem:

$$\text{number of edges} = \frac{\text{sum of the degrees of vertices}}{2}$$

Degrees of vertices: example 1



- degree$(a) = 2$,
- degree$(b) =$ degree$(c) =$ degree$(f) = 4$,
- degree$(e) = 3$,
- degree$(d) = 1$, so d is pendant,
- degree$(g) = 0$, so g is isolated.

Directed graphs: basic terminology

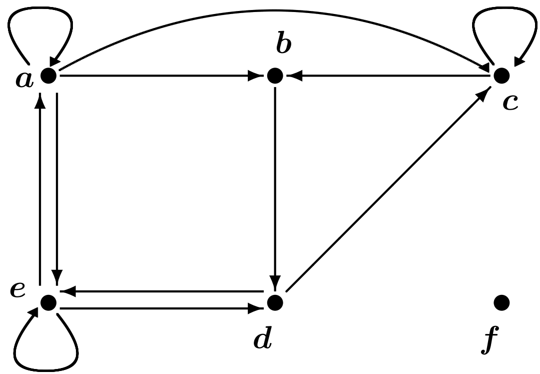
If there is an edge e going from vertex u to v, we say that

- u is **adjacent to** v,
- u is the **initial** or **start vertex** of e, and
- v is the **terminal** or **end vertex** of e.

The **in-degree** of a vertex v is the number of edges with v as their terminal vertex. The **out-degree** of a vertex v is the number of edges with v as their initial vertex. (A loop at a vertex contributes 1 to both the in-degree and the out-degree of this vertex.)

$$\begin{aligned}\text{number of edges} &= \text{sum of the in-degrees of vertices}\\ &= \text{sum of the out-degrees of vertices.}\end{aligned}$$

Degrees of vertices: example 2



- in-degree(a) $=$ in-degree(b) $=$ in-degree(d) $= 2$,

  in-degree(c) $=$ in-degree(e) $= 3$,

  in-degree(f) $= 0$,

- out-degree(a) $= 4$, out-degree(b) $= 1$,

  out-degree(c) $=$ out-degree(d) $= 2$, out-degree(e) $= 3$,

  out-degree(f) $= 0$.

Representing graphs: adjacency matrix

- List the vertices in some order horizontally from left to right.
- Then, using the same order, list them vertically from top to bottom.
- The entry in the i^{th} row and the j^{th} column is the

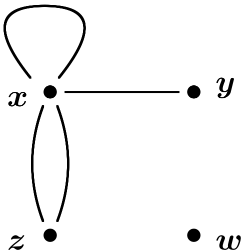
  number of edges going from vertex i to vertex j.

If the graph is undirected, then

the number in the i^{th} row and j^{th} column

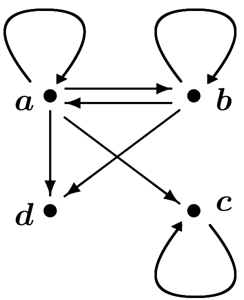
$=$ the number in the j^{th} row and i^{th} column.

Adjacency matrices: examples



	x	y	w	z
x	1	1	0	2
y	1	0	0	0
w	0	0	0	0
z	2	0	0	0

	y	x	w	z
y	0	1	0	0
x	1	1	0	2
w	0	0	0	0
z	0	2	0	0



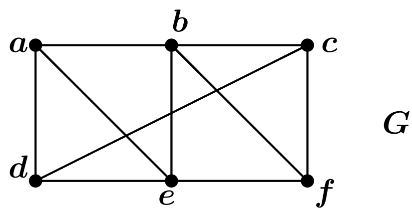
	a	b	c	d
a	1	1	1	1
b	1	1	0	1
c	0	0	1	0
d	0	0	0	0

	b	c	d	a
b	1	0	1	1
c	0	1	0	0
d	0	0	0	0
a	1	1	1	1

Paths in simple graphs

- A **path** is a sequence of vertices travelling along edges.
- The **length** of a path is the number of edges in it.
- A path is called **simple** if it does not contain the same *edge* twice.
- A **Hamiltonian path** is a *simple* path passing through *every vertex exactly once.*

Example:



(a, b, c, f, b, e) is a simple path in G of length 5

(d, c, a, e) is not a path in G — WHY? no edge between c and a

(a, b, e, d, a, b) is a path in G of length 5, but it is not simple
— WHY? contains the edge between a and b twice

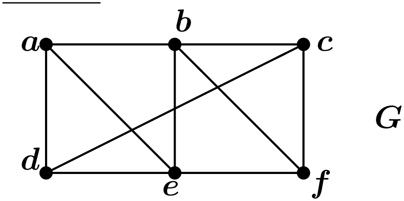
(d, a, e, b, f, c) is a Hamiltonian path in G

(a, b, c, f, b, e, d) is a simple path, but not a Hamiltonian path in G
— WHY? passes b more than once

Cycles in simple graphs

- A **cycle** is a path beginning and ending with the same vertex.
- The **length** of a cycle is the number of edges in it.
- A cycle is called **simple** if it does not contain the same *edge* twice.
- A **Hamiltonian cycle** is a *simple* cycle passing through *every vertex exactly once.*

Example



(a, d, c, b, a) is a simple cycle in G of length 4

(b, e, d, a, e, b) is a cycle in G of length 5, but it is not simple

— WHY? contains the edge between b and e twice

(c, f, e, d, a, b, c) is a Hamiltonian cycle in G

(e, d, c, f, e, b, a, e) is a simple cycle in G, but not Hamiltonian

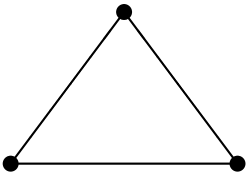
— WHY? passes e more than once

(a, d, e, f, b, a) is a simple cycle in G, but not Hamiltonian

— WHY? does not pass c

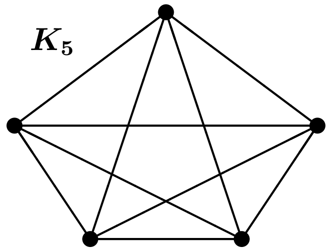
Special simple graphs

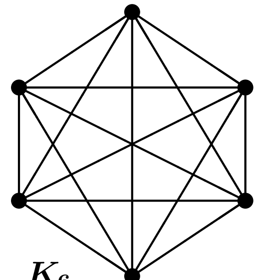
- The **complete graph on n vertices** (or **n-clique**), denoted $\boxed{K_n}$, is the simple graph that contains an edge between each pair of distinct vertices.



K_3

K_5

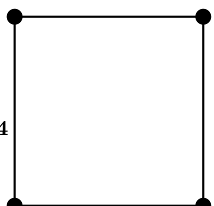


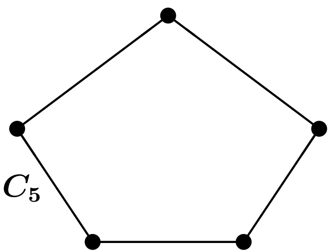


K_6

- The **n-cycle** is denoted by $\boxed{C_n}$, for $n \geq 4$.

C_4





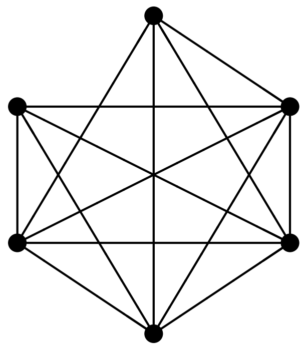
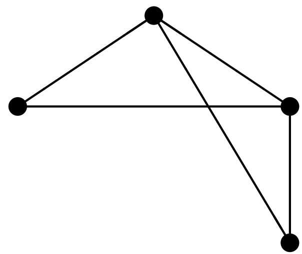
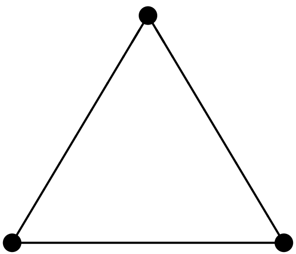
C_5

Subgraphs of graphs

When edges and vertices are removed from a graph, without removing endpoints of any remaining edges, a smaller graph is obtained.

Such a graph is called a **subgraph** of the original graph.

Example: Each of the following 3 graphs

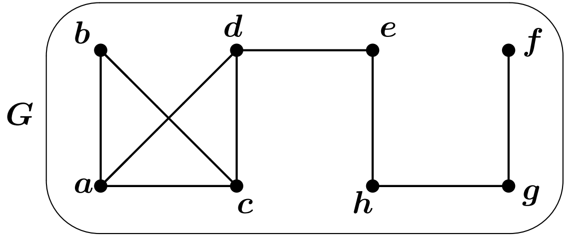




is a subgraph of $\boldsymbol{K_6}$.

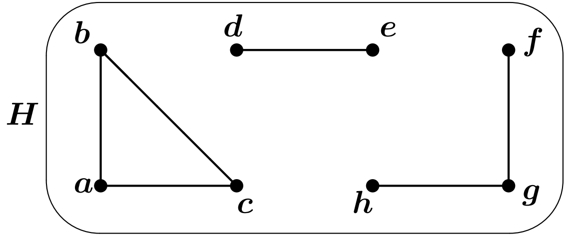
Connected graphs

A simple graph is called **connected** if there is a path between every pair of distinct vertices.

Example:



is connected

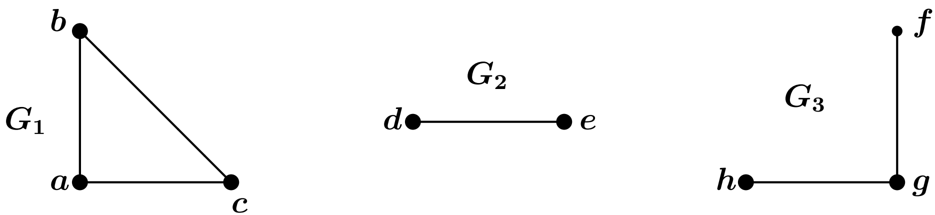


is not connected

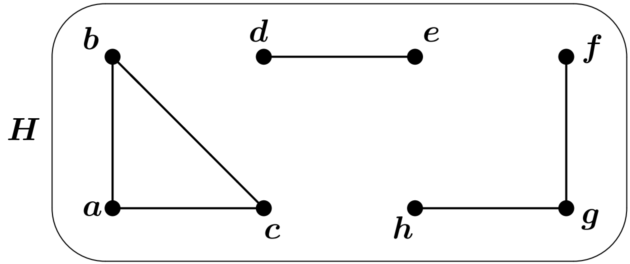
Connected components of graphs

A **connected component** of a graph is a maximal connected subgraph.

- If a graph is connected, then it has only 1 connected component, *itself.*
- But if it is not connected, it can have more:



are the 3 connected components of

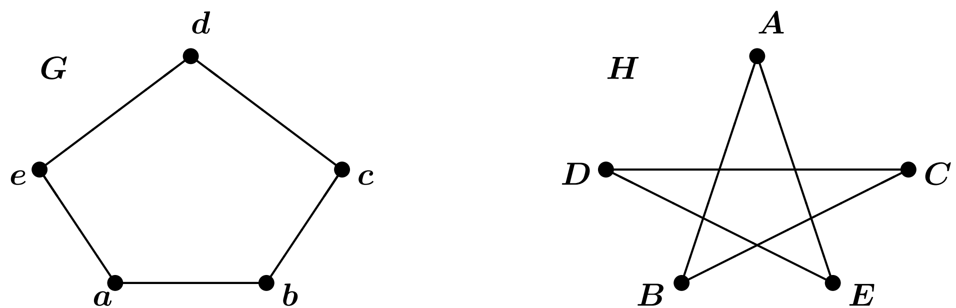


Isomorphism of graphs: an example

The following instructions were given to two persons:

> "Draw and label five vertices with a, b, c, d, and e.
> Connect a and b, b and c, c and d, d and e, and a and e."

They drew the graphs:



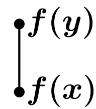
Surely, these drawings describe the same situation, though the graphs G and H appear dissimilar.

Isomorphism of graphs

Graphs G and H are **isomorphic** if there is an **isomorphism** between them:
a function f from the vertices of G to the vertices of H such that

- f is a **bijection** (one-to-one and onto)
- f **'takes' edges to edges**:

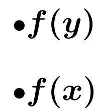
  for all vertices x, y in G, if 

then 

in H

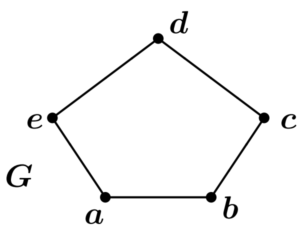
- f **'takes' non-edges to non-edges**:

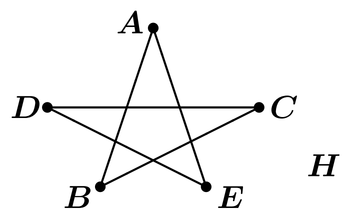
  for all vertices x, y in G, if 

then 

in H

Example:





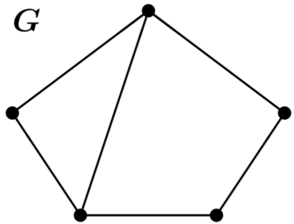
The function f defined by taking

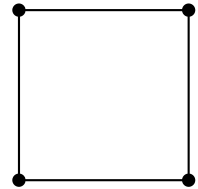
$$f(a) = A, \qquad f(b) = B, \qquad f(c) = C, \qquad f(d) = D, \qquad f(e) = E$$

is an isomorphism, showing that graphs G and H are isomorphic.

Isomorphic or not — how can we decide?

G





H

Task: Determine whether two graphs G and H are isomorphic or not.

isomorphic = there is an isomorphism

not isomorphic = there is no isomorphism

- We can try **all possible functions** from G to H and check whether any of them is an isomorphism.
- But this might take a lot of time: there are $4^5 = 1024$ possible functions even for this simple example. So for larger graphs it is hopeless.

Is there a quicker way?

Isomorphism of graphs: invariants

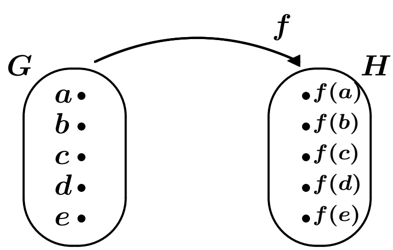
A property $\mathcal{P}$ of graphs is called an **invariant** if it is 'preserved under isomorphisms': **If G and H are isomorphic graphs, and G has property $\mathcal{P}$, then H has property $\mathcal{P}$ as well.**

Example: "*Having* 5 *vertices*" is an invariant:

If G and H are isomorphic graphs, and G has 5 vertices, then H has 5 vertices as well.

WHY? If G and H are isomorphic, then there is an isomorphism f between them.

– f is a *bijection* from the vertices of G (domain) to the vertices of H (codomain).

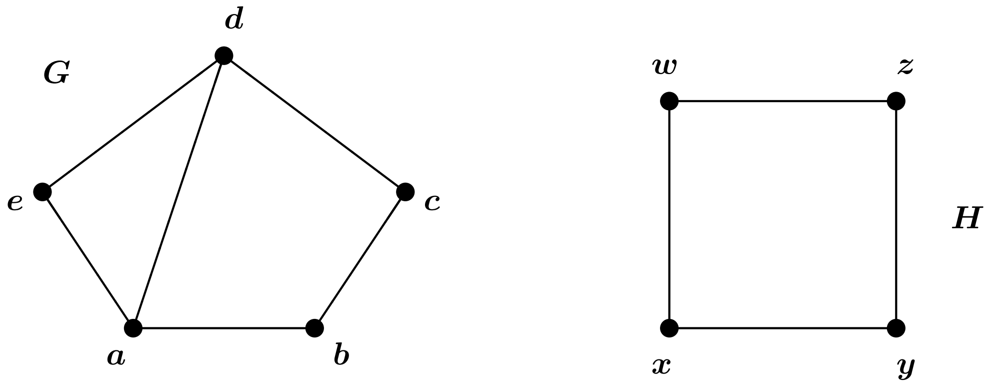


– As f is *one-to-one,* H has at least 5 vertices.

– And as f is *onto,* H has at most 5 vertices.

Exercise

Task: Determine whether G and H are isomorphic or not:



Solution: We've just seen that the property "having 5 vertices" is an invariant. This property holds for G, but not for H. Therefore, G and H are not isomorphic.

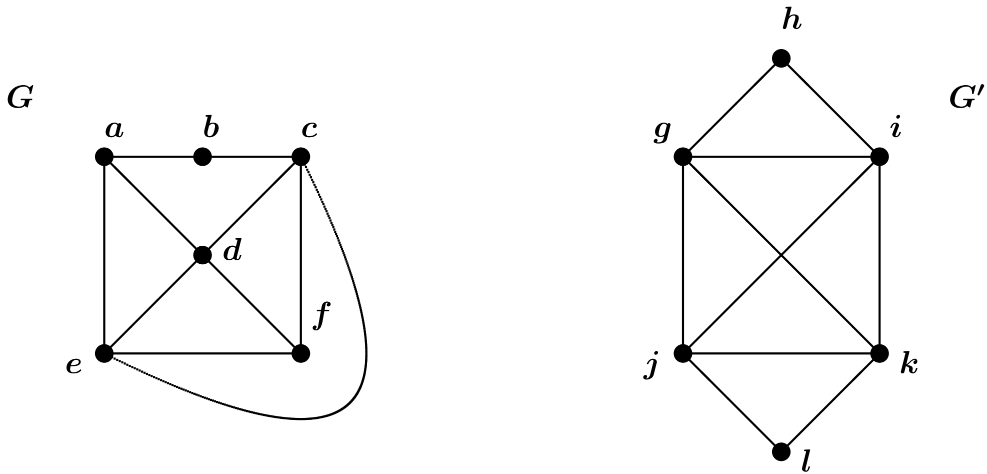
Some more examples of invariants

if in one of my graph there's one vertex with 17 degrees, then in the other graph there must be another vertex with 17 degrees.
Degree = how many edges a vertex has.

- the number of edges
- the number of vertices of each degree
- containing a triangle (K_3) as a subgraph
- containing two K_4s as disjoint subgraphs
- containing a simple cycle of length 4
- having a path of length 2 between two vertices of degree 2
- $\ldots$

Exercise

Task: Determine whether G and G' are isomorphic or not:

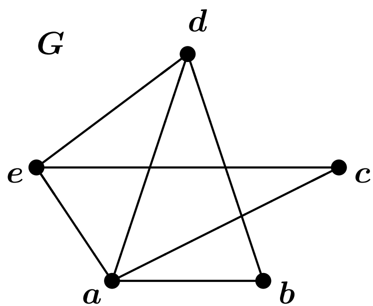


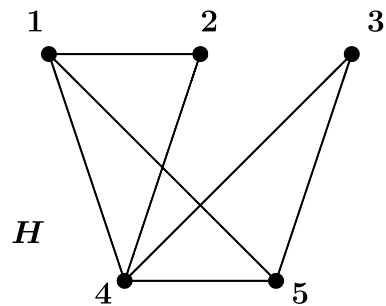
Solution: Now it is a bit harder to find an invariant. Both graphs have 6 vertices and 10 edges. But we have just seen that "having a vertex of degree 3" is an **invariant.**

This property holds in G (say, $\text{degree}(a) = 3$), but does not hold in G', showing that G and G' are not isomorphic.

Exercise

Task: Determine whether G and H are isomorphic or not:





Solution: The function f defined by taking

$$f(a) = 4, \qquad f(b) = 2, \qquad f(c) = 3, \qquad f(d) = 1, \qquad f(e) = 5$$

is an isomorphism (because f is a bijection, takes edges to edges, and non-edges to non-edges). This shows that graphs G and H are <u>isomorphic.</u>

Isomorphic or not — so how can we decide?

This will be in exam.

Task: Determine whether two graphs G and H are isomorphic or not.

Solution: There is no easy way. We have to try in parallel:

- To describe a bijection between the vertices of G and H that 'takes' edges to edges, non-edges to non-edges.
  $\rightsquigarrow$ If we succeed, the answer is YES.

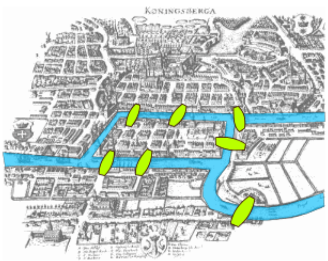
- To find an invariant $\mathcal{P}$
  and show that G has $\mathcal{P}$ but H doesn't, or the other way round.
  $\rightsquigarrow$ If we succeed, the answer is NO.

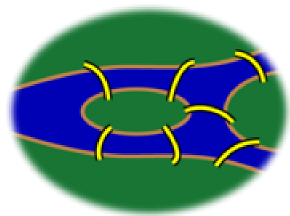
It would be nice to have a list of easily checkable invariants that isomorphic graphs and *only* isomorphic graphs share. Then we would just have to check those.

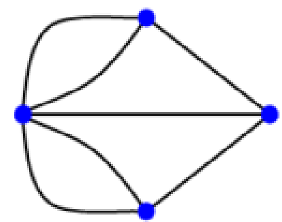
Unfortunately, no one has yet succeeded in finding such a list of invariants, so determining whether two graphs are *not* isomorphic might require some creative thinking. https://en.wikipedia.org/wiki/Graph_isomorphism_problem

Graph theory: a bit of history

Leonhard Euler (1707-1783) is considered to be one of the greatest mathematicians who ever lived. In 1736 he considered the following problem. The city of Königsberg, Prussia (now Kaliningrad, Russia) was set on the river Pregel, and included two large islands which were connected to each other and the mainland by seven bridges. The question was whether it is possible to walk with a route that crosses each bridge exactly once, and return to the starting point. Euler proved that it was not possible.

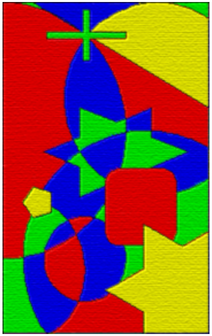






In proving the result, Euler formulated the problem in terms of graph theory, by abstracting the case of Königsberg — first, by eliminating all features except the landmasses and the bridges connecting them; second, by replacing each landmass with a vertex and each bridge with an edge.

Four colour theorem



The four colour theorem states that any plane separated into regions, such as a political map of the counties of a state, can be coloured using no more than **four colours** in such a way that no two adjacent regions receive the same colour. Two regions are adjacent if they share a border segment, not just a point. Each region must be contiguous — that is, it may not be partitioned as are Michigan and Russia.
(Three colours are clearly not enough: take the map with one region surrounded by three other regions.)

This theorem was the first major theorem to be proved using a **computer**

The conjecture was first proposed in 1852 when Francis Guthrie, while trying to colour the map of counties of England, noticed that only four different colours were needed. At the time, Guthrie was a student of Augustus De Morgan at University College.

How to formulate this problem in graph-theoretic terms?

Every planar graph is four-colourable https://en.wikipedia.org/wiki/Four_color_theorem

Alphabets and words

An **alphabet** is a finite set Σ of symbols

Examples:
- $\Sigma_1 = \{a, b, c, \ldots, z\}$, the set of all lower-case letters
- $\Sigma_2 = \{0, 1\}$, the binary alphabet
- $\Sigma_3 = \{\square, \diamond, \heartsuit\}$

A **word** or **string** (over an alphabet Σ) is a finite sequence of symbols from Σ

Examples:
- *abracadabra*, *azwzax* (over Σ_1)
- 111111111110000000000, 000110 (over Σ_2)
- $\heartsuit\heartsuit\square$, $\square\diamond\square\square\diamond\heartsuit$ (over Σ_3)
- the **empty word** ε is a word over **any** alphabet Σ (but we may assume that ε is *NOT a symbol* of any of our alphabets)

Words (cont.)

- $\boxed{\Sigma^*}$ is the set of all words over Σ (always contains ε)
- The **length** $|w|$ of a word w is the number of symbols in w

$$\boxed{|w| \;=\; \text{the number of occurrences of symbols in } w}$$

e.g., $|azwza| = 5$, $|\heartsuit\heartsuit\square| = 3$, $|\varepsilon| = 0$

- The **concatenation** of words x and y (notation: xy) is

  the word x followed by the word y

  - $w^n \;=\; \underbrace{ww\dots w}_{n}$ (e.g., $(\heartsuit\square)^0 = \varepsilon$, $(01)^3 = 010101$, $a^4 = aaaa$)
  - $x\varepsilon \;=\; \varepsilon x = x$, for every word x

- if $w = xy$ then x is a **prefix** of w, and y a **suffix** of w

  e.g., <u>*tor*</u> is a prefix and <u>*se*</u> is a suffix of <u>*tortoise*</u>

Languages

A **language** over an alphabet Σ is a set of words over Σ,
that is, a **language** is a subset of Σ^*

Examples:

(1) $\Sigma = \{a, b, c, \ldots, z\}$

- $L_1 =$ all English words
- $L_2 =$ all Latin words
- $L_3 = \{kdpekvg, leih, hkiiw, wowiszk\}$

(2) $\Sigma = \{0, 1\}$

- $L_4 = \{001, 101010, 111, 1001\}$
- $L_5 = \{0^n 1^m \mid n \text{ is an even, } m \text{ is an odd number}\}$