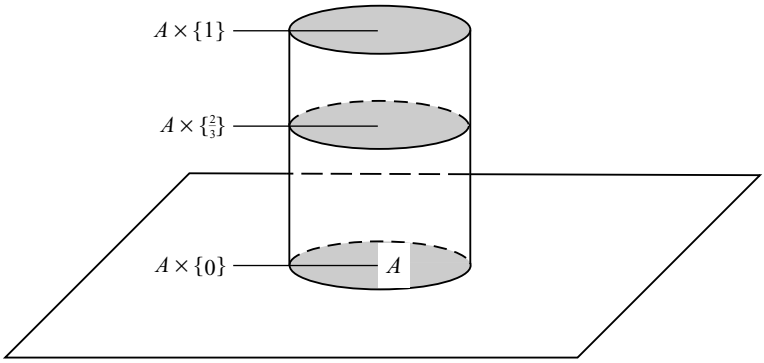

Fig. 2.7

Example 1 Let X be the set $\{0, 1, 2\}$ and let Y be the set $\{5, 6\}$. Then $X \times Y$ is a set with six elements: $X \times Y = \{(0, 5), (0, 6), (1, 5), (1, 6), (2, 5), (2, 6)\}$.

Example 2 Let $\mathbb{R}$ be the set of real numbers (conceived of as an infinite line) then $\mathbb{R} \times \mathbb{R}$ (also written $\mathbb{R}^2$) may be thought of as the real plane, where a point of this plane is identified with its ordered pair of coordinates (x, y) (with respect to the first and second coordinate axes $\mathbb{R} \times \{0\} = \{(a, 0): a \in \mathbb{R}\}$ and $\{0\} \times \mathbb{R} = \{(0, b): b \in \mathbb{R}\}$).

Example 3 Let X be any set. Then $X \times \emptyset$ is the empty set $\emptyset$ (since $\emptyset$ has no members).

Example 4 One may think of Euclidean 3-space $\mathbb{R}^3$ as being $\mathbb{R}^2 \times \mathbb{R}$ (that is, as the product of a plane with a line). Let A be a disc in the plane $\mathbb{R}^2$ and let $[0, 1]$ be the interval $\{x \in \mathbb{R} : 0 \leq x \leq 1\}$. Then $A \times [0, 1]$, regarded as a geometric object, is a vertical solid cylinder of height 1 and with base lying on the plane $\mathbb{R}^2 \times \{0\}$ (Fig. 2.7).

The ideas in this section go back mainly to Boole and Cantor. Cantor introduced the abstract notion of a set (in the context of infinite sets of real numbers). The 'algebra of sets' is due mainly to Boole – at least, in the equivalent form of the 'algebra of propositions' (for which, see Section 3.1).

In this section, we have introduced set theory simply to provide a convenient language in which to couch mathematical assertions. But there is much more to it than this: it can also be used as a foundation for mathematics. For this aspect, see discussion of the work of Cantor and Zermelo in the historical references.

Exercises 2.1

1. Which among the following sets are equal to one another?
   $X = \{x \in \mathbb{Z}\colon x^3 = x\}$;
   $Y = \{x \in \mathbb{Z}\colon x^2 = x\}$;
   $Z = \{x \in \mathbb{Z}\colon x^2 \leq 2\}$;
   $W = \{0, 1, -1\}$;
   $V = \{1, 0\}$.
2. List all the subsets of the set $X = \{\text{a, b, c}\}$. How many are there? Next, try with $X = \{\text{a, b, c, d}\}$. Now suppose that the set X has n elements: how many subsets does X have? Try to justify your answer.
3. Show that $X \backslash Y = X \cap Y^{\text{c}}$ (where the complement may be taken with respect to $X \cup Y$: that is, you may take the universal set U to be $X \cup Y$).
4. Define the **symmetric difference**, $A \Delta B$, of two sets A and B to be $A \Delta B = (A \backslash B) \cup (B \backslash A)$. Draw a Venn diagram showing the relation of this set to A and B. Show that this operation on sets is associative: for all sets A, B, C one has $(A \Delta B) \Delta C = A \Delta (B \Delta C)$.
5. Prove the parts of 2.1.1 that you have not yet checked.
6. List all the elements in the set $X \times Y$, where $X = \{0, 1\}$ and $Y = \{2, 3\}$. List all the subsets of $X \times Y$.
7. Let A, B, C, D be any sets. Are the following true? (In each case give a proof that the equality is true or a counterexample which shows that it is false.)
   (i) $(A \times C) \cap (B \times D) = (A \cap B) \times (C \cap D)$;
   (ii) $(A \times C) \cup (B \times D) = (A \cup B) \times (C \cup D)$.
8. Suppose that the set X has m members and the set Y has n members. How many members does the product set $X \times Y$ have?
   [Hint: try Exercise 2.1.6 first, then try with X having, say, three members and Y having four, ..., and so on – until you see the pattern. Then justify your answer.]
9. Give an example to show that if X is a subset of $A \times B$ then X does not need to be of the form $C \times D$ where C is a subset of A and D is a subset of B.

2.2 Functions

In this section we discuss functions and how they may be combined. The notion of a function is one of the most basic in mathematics, yet the way in which mathematicians have understood the term has changed considerably over the ages. In particular, history has shown that it is unwise to restrict the methods by which functions may be specified. Therefore the definition of function which we give may seem rather abstract since it concentrates on the end result – the

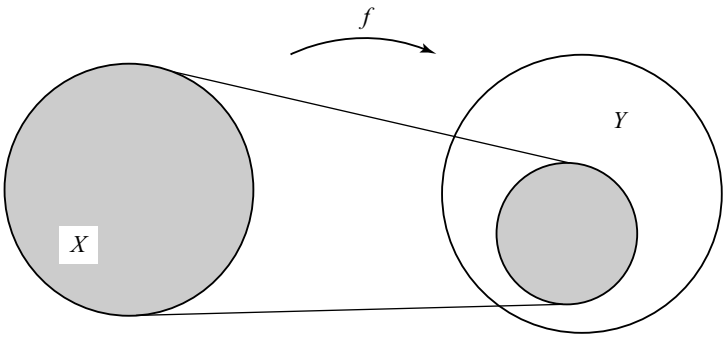

Fig. 2.8

function – rather than any way in which the function may be defined. For more on the development of the notion of function, see the notes at the end of the section.

As a first approximation, one may say that a function from the set X to the set Y is a rule which assigns to each element x of X an element of Y. Certainly something of the sort $f(x) = x^2 + 1$ serves to define a value $f(x)$ whenever the real number x is given and so this 'rule' defines a function from $\mathbb{R}$ to $\mathbb{R}$. But the term 'rule' is problematic since, in trying to specify what one means by a 'rule', one may exclude quite reasonable 'functions'.

Therefore, in order to bypass this difficulty, we will be rather less specific in our terminology and simply define a **function** from the set X to the set Y to be an assigment: to each element x of X is assigned an element of Y which is denoted by $f(x)$ and called the **image** of x. We refer to X as the **domain** of the function and Y is called the **codomain** of the function. The **image** of the function f is $\{f(x)\colon x \in X\}$, a subset of Y. The words **map** and **mapping** are also used instead of 'function'. The notation '$f\colon X \to Y$' indicates that f is a function from X to Y. A way of picturing this situation is shown in Fig. 2.8.

In the definition above, we replaced the word 'rule' by the rather more vague word 'assignment', in order to emphasise that a function need not be given by an explicit (or implicit) rule. In order to free our definition from the subtleties of the English language, we give a rigorous definition of function below.

You may find it helpful to think of a function $f\colon X \to Y$ as a 'black box' which takes inputs from X and yields outputs in Y and which, when fed with a particular value $x \in X$ outputs the value $f(x) \in Y$. Our rather free definition of 'function' means that we are saying nothing about how the 'black box' 'operates' (see Fig. 2.9).

Let us consider the following example.

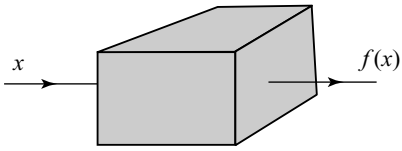

Fig. 2.9

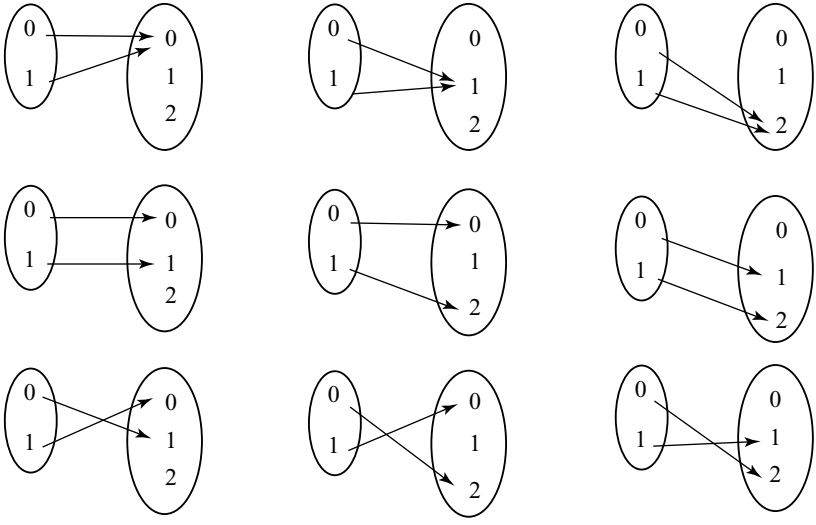

Fig. 2.10

Example We find all the functions from the set $\{0, 1\}$ to the set $\{0, 1, 2\}$. If one is used to functions given by rules such as $f(x) = x^2$, then one is tempted to spend rather a lot of time and energy in trying to describe the functions from X to Y by rules of that sort (e.g. $f(x) = x + 1$, $g(x) = 2 - x$, ...); but that is not what is asked for. All one needs to do to describe a particular function from X to Y is to specify, in an unambiguous way, for each element of X, an element of Y. So we can give an example of a function simply by saying that to the element 0 of X is assigned the element 2 of Y and to the element 1 of X is assigned the element 0 of Y. There is no need to explain this assignment in any other way.

Describing such functions in words is rather tedious: there are better ways. Figure 2.10 shows the nine ($=3^2$) possible functions from X to Y. (There are three choices for where to send $0 \in X$ and, for each of these three choices, three choices for the image of $1 \in X$: hence $3 \times 3 = 9$ in all.)

Another way of describing a function is simply to write down all pairs of the form $(x, f(x))$ with $x \in X$. So, the first function in the figure is completely

described under this convention by the set $\{(0, 0), (1, 0)\}$ and the second function is described by $\{(0, 0), (1, 1)\}$. The set of all such ordered pairs is called the graph of the function. We define this formally.

Definition The **graph** of a function $f\colon X \to Y$ is defined to be the following subset of $X \times Y$:

$$\mathrm{Gr}(f) = \{(x, y)\colon x \in X \text{ and } y = f(x)\}; \text{ that is}$$
$$\mathrm{Gr}(f) = \{(x, f(x))\colon x \in X\}.$$

Notice that, since a function takes only one value at each point of its domain, the graph of a function f has the property that for each x in X there is precisely one y in Y such that (x, y) is in $\mathrm{Gr}(f)$.

Since a function and its graph each determines the other, we may now give an entirely rigorous definition of 'function' by saying that a function is a subset, G, of $X \times Y$ which satisfies the condition that for each element x of X there exists exactly one y in Y such that (x, y) is in G. (That is, we identify a function with its graph.)

It follows that two functions f, g are equal if they have the same domain and codomain and if, for every x in the common domain, $f(x) = g(x)$.

Example Suppose $X = Y = \mathbb{R}$ and let $f\colon \mathbb{R} \to \mathbb{R}$ be the function which takes any real number x to x^2. Then $\mathrm{Gr}(f)$ is the set of those points in the real plane $\mathbb{R} \times \mathbb{R}$ which have the form (x, x^2) for some real number x. Think of this set geometrically to see why the term 'graph of a function' is appropriate for the notion defined above.

We may think of (the graph of) a function f as inducing a correspondence or relation from its domain X to its codomain Y. Since f is defined at each point of its domain, each element of X is related to at least one element of Y. Since the value of f at an element of X is uniquely defined, no element of X may be related to more than one element of Y. Therefore Fig. 2.11 *cannot* correspond to a function.

It is however possible that (a) there is some element of Y that is *not the image of any element* of X, (b) some element of Y is *the image of more than one element* of X. So Fig. 2.12 *may* arise from a function.

For instance, take $X = \mathbb{R} = Y$ and let $f(x) = x^2$. For an example of (a), consider -4 in Y (there is no $x \in \mathbb{R}$ with $x^2 = -4$); for an example of (b), consider 4 in Y (there are two values, $x = 2$, $x = -2$, from $\mathbb{R}$ with $x^2 = 4$). A function which avoids '(a)' is said to be surjective; one which avoids '(b)' is said to be injective; a function which avoids both '(a)' and '(b)' is said to be bijective. More formally, we have the following.

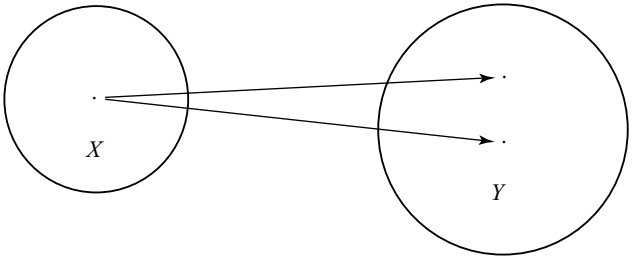

Fig. 2.11

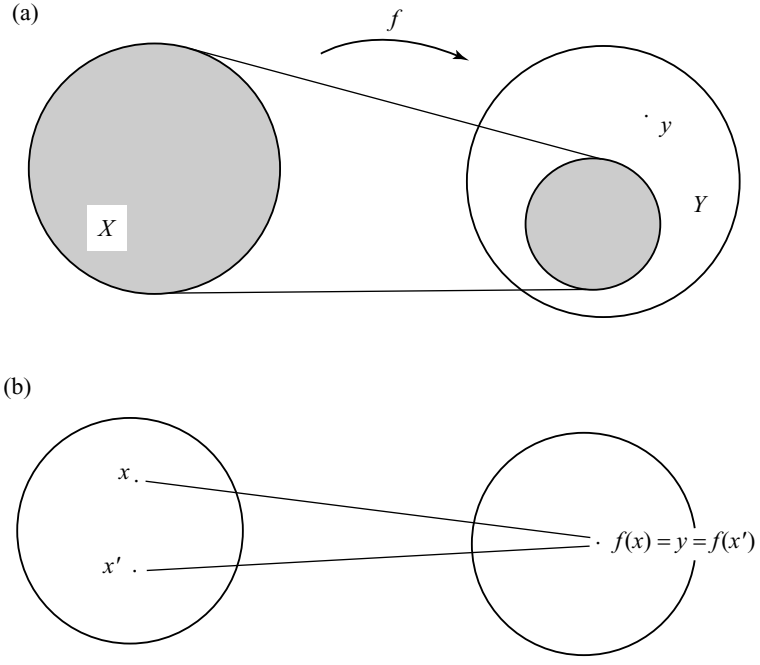

Fig. 2.12

Definition Let $f: X \to Y$ be a function. We say that f is **surjective** (or **onto**) if for each y in Y there exists (at least one) x in X such that $f(x) = y$ (that is, every element of Y is the image of element of X). The function f is **injective** (or **one-to-one**, also written '1-1') if for x, x' in X the equality $f(x) \equiv f(x')$ implies $x = x'$ (that is, distinct elements of X cannot have the same image in Y). Finally, f is **bijective** if f is both injective and surjective. A **surjection** is a function which is surjective; similarly with **injection** and **bijection**. A **permutation** of a set is a bijection from that set to itself. We will study the structure of permutations of finite sets in Chapter 4.

Example 1 The function $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^4$ is neither injective nor surjective. It is not injective since, for example, $f(2) = 16 = f(-2)$ but $2 \neq -2$. The fact that it is not surjective is shown by the fact that -1 (for example) is not in the image of f: it is not the fourth power of any real number.

Example 2 The function $s : \mathbb{P} \to \mathbb{P}$ defined by $s(n) = n + 1$ is injective but not surjective. It is not surjective because the equation $s(n) = 1$ has no solution in $\mathbb{P}$, that is, there is no $n \in \mathbb{P}$ with $n + 1 = 1$.

To show that s is injective, suppose that $s(n) = s(m)$, then $n + 1 = m + 1$. Then $n = m$. Turning this round (i.e. the 'contrapositive' statement – see p. 132) we have shown that if $n \neq m$ then $s(n) = s(m)$.

Example 3 The function $g : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^5$ is bijective. To prove this, one may proceed as follows.

Surjective. To say that this function g is surjective is precisely to say that every real number has a real fifth root – an assertion which is true and, we assume, known to you.

Injective. Suppose that $f(x) = f(y)$: that is, $x^5 = y^5$. Thus $x^5 - y^5 = 0$. Factorising this gives

$$(x - y) \cdot (x^4 + x^3y + x^2y^2 + xy^3 + y^4) = 0.$$

If we can show that the second factor $t = x^4 + x^3y + x^2y^2 + xy^3 + y^4$ is never zero except when $x = y = 0$ then it will follow that $x^5 - y^5$ equals 0 only in the case that $x = y$: in other words, it will follow that the function f is injective. Now, there are various ways of showing that the factor t is zero only if $x = y = 0$: perhaps the most elementary is the following.

We intend to use the fact that a sum of squares of real numbers is zero only if the terms in the sum are individually zero. Notice that the term $x^3y + xy^3$ equals $xy(x^2 + y^2)$. This suggests considering the term $(x + y)^2(x^2 + y^2)$ or at least, half of it. Following this up, we obtain:

$$t = \frac{1}{2}(x + y)^2(x^2 + y^2) + \frac{1}{2}x^4 + \frac{1}{2}y^4,$$

and this can be written as

$$\left(\left(\frac{1}{\sqrt{2}}\right) \cdot (x + y) \cdot x\right)^2 + \left(\left(\frac{1}{\sqrt{2}}\right) \cdot (x + y) \cdot y\right)^2 + \left(\left(\frac{1}{\sqrt{2}}\right) x^2\right)^2 + \left(\left(\frac{1}{\sqrt{2}}\right) y^2\right)^2.$$

Thus t is indeed a sum of squares and we can see that this sum is zero only if $x = y = 0$, as required.

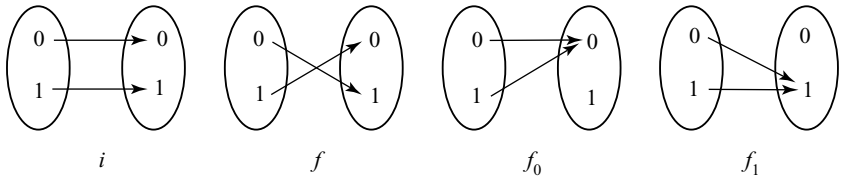

Fig. 2.13

In fact, for any function $f\colon \mathbb{R} \to \mathbb{R}$, $y = f(x)$, we can interpret the ideas of injective and surjective in terms of the graph of f ('graph' in the pictorial sense, drawn with the x-axis horizontal). Thus f is injective if and only if every horizontal line meets the graph in *at most one* point. Similarly, f is surjective if and only if every horizontal line meets the graph of f in *at least one* point. Using these ideas, it is easy to see that the function $f(x) = x^3$ is injective, that the function $h\colon \mathbb{R} \to \mathbb{R}$ given by $h(x) = x^3 - x$ is surjective but not injective, that the function $k\colon \mathbb{R} \to \mathbb{R}$ given by $k(x) = \mathrm{e}^x$ is injective but not surjective.

We can also express these ideas in terms of solvability of equations. To say that $f\colon X \to Y$ is surjective is to say that for every $b \in Y$ the equation $f(x) = b$ *has a solution*. To say that f is injective is to say that for every $b \in Y$ the equation $f(x) = b$ *has at most one solution*. To say that f is bijective is to say that for every $b \in Y$ the equation $f(x) = b$ *has exactly one solution*.

Example 1 Consider the possible functions from $\{0, 1\}$ to itself. There are four possible functions (two choices for the value of f at 0; then, for each of these, two choices for $f(1)$). Their actions can be shown as in Fig. 2.13. The functions i and f are bijections and f_0 and f_1 are neither injections nor surjections.

Example 2 Refer back to the first example of this section. There are no surjections from $\{0, 1\}$ to $\{0, 1, 2\}$ and hence there are no bijections. But the function f defined by $f(0) = 2$ and $f(1) = 0$ is an example of an injection (you may check that six of the nine functions are injective).

Definitions If X is a set then the function $\mathrm{id}_x\colon X \to X$ which takes every element to itself ($\mathrm{id}_x(x) = x$ for all x in X) is the **identity function** on X. If X and Y are sets and c is in Y then we may define the **constant function from X to Y with value** c by setting $f(x) = c$ for every x in X. (Do not confuse the identity function on a set with, when it makes sense, a function with constant value '1'.)

Definition Suppose that $f\colon X \to Y$ and $g\colon Y \to Z$ are functions. We can take any element x of X, apply f to it and then apply g to the result (since the result is in Y). Thus we end up with an element $g(f(x))$ of Z. What we have just done is

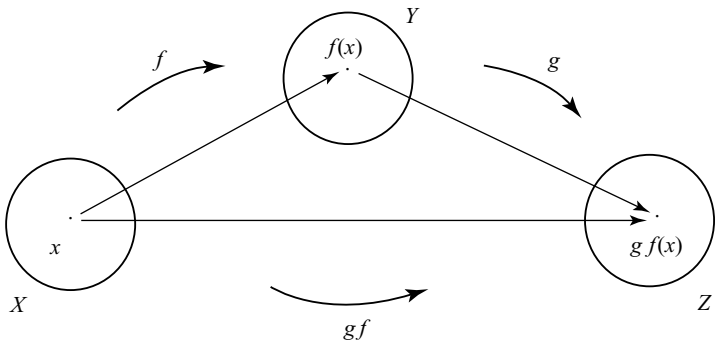

Fig. 2.14

to define a new function from X to Z: it is denoted by $gf\colon X \to Z$, is defined by $gf(x) = g(f(x))$, and is called the **composition** of f and g (note the reversal of order: gf means 'do f first and then apply g to the result'). See Fig. 2.14.

In the case that $X = Y = Z$ and $f = g$, the composition of f with itself is often denoted f^2 rather than ff (similarly for $f^3, \ldots$).

Example 1 Let f and g be the functions from $\mathbb{R}$ to $\mathbb{R}$ defined by $f(x) = x + 1$ and $g(x) = x^2$. The composite function fg is given by

$$fg(x) = f(g(x)) = f(x^2) = x^2 + 1.$$

Note that the composition gf is given by

$$gf(x) = g(x + 1) = (x + 1)^2 = x^2 + 2x + 1.$$

Thus, even if both functions gf and fg are defined, they need not be equal.

Example 2 Let $f, g\colon \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = 4x - 3$ and $g(x) = (x + 3)/4$.
Then

$$fg(x) = f((x + 3)/4) = 4((x + 3)/4) - 3 = x + 3 - 3 = x$$

and

$$gf(x) = g(4x - 3) = ((4x - 3) + 3)/4 = 4x/4 = x.$$

In this case, it does turn out that fg and gf are the same function, namely the identity function on $\mathbb{R}$.

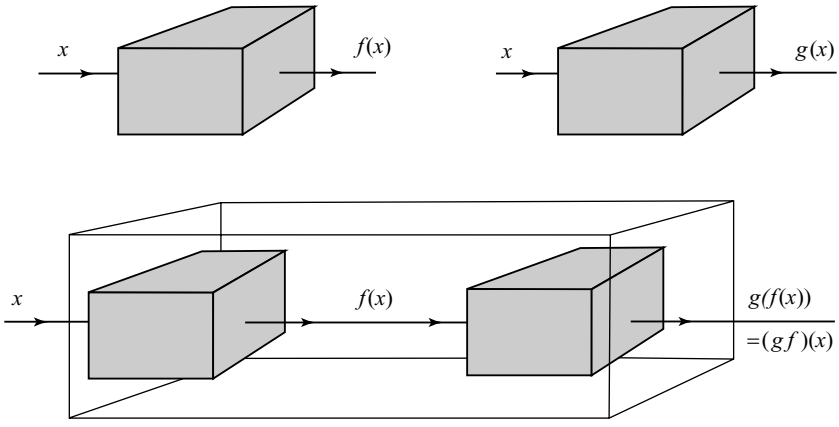

Fig. 2.15

Example 3 Suppose that F and G are computer programs, each of which takes integer inputs and produces integer outputs. To F we may associate the function f which is defined by $f(n) =$ that integer which is output by F if it is given input n. Similarly, let g be the function which associates to any integer n the output of G if G is given input n. We may connect these programs in series as shown in Fig. 2.15: thus the output of F becomes the input of G.

Regard this combination as a single program: the function which is associated to it is precisely the composition gf. If one thinks of a function as a 'black box', as indicated after the definition of function, then the picture above suggests a way of thinking about the composition of two functions.

Example 4 Let $f\colon X \to Y$ be any function. Then $f\mathrm{id}_x(x) = f(\mathrm{id}_x(x)) = f(x)$, so $f\mathrm{id}_x = f$. Similarly, $\mathrm{id}_Y f = f$.

Suppose now that we have functions $f\colon X \to Y$, $g\colon Y \to Z$ and $h\colon Z \to W$. Then we may form the composition gf and then compose this with h to get $h(gf)$. Alternatively we may form hg first and then apply this, having already applied f, to obtain $(hg)f$. The first result of this section says that the result is the same: $h(gf) = (hg)f$. This is the associative law for composition of functions.

Theorem 2.2.1 *If $f\colon X \to Y$, $g\colon Y \to Z$, $h\colon Z \to W$ are functions then $h(gf) = (hg)f$ and so this function from X to W may be denoted unambiguously by $hgf\colon X \to W$.*

Proof Consider Fig. 2.16. We see that the element x of X is sent to the same element w of W by the two routes. The first applies f and then the composite $hg\colon (hg)f$. The second applies the composite of gf and then $h\colon h(gf)$. □

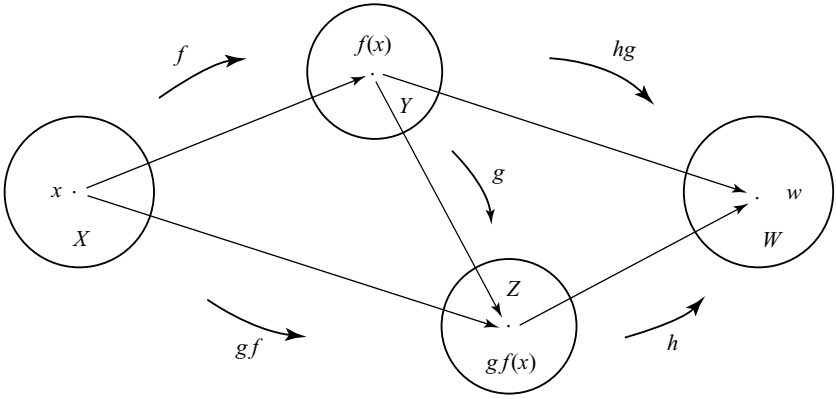

Fig. 2.16

Consider the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3$. If $g\colon \mathbb{R} \to \mathbb{R}$ is the function which takes each real number to its (unique!) real cube root then it makes sense to say that g reverses the action of f and, indeed, f reverses the action of g, since

$$\begin{array}{ccc} gf(x) = x = \mathrm{id}_{\mathbb{R}}(x) & \text{and} & fg(x) = x = \mathrm{id}_{\mathbb{R}}(x) \text{ for every } x \in \mathbb{R}. \\ \| \quad \| & & \| \quad \| \\ g(x^3) = (x^3)^{1/3} & & f(x^{1/3}) = (x^{1/3})^3 \end{array}$$

Definition Suppose that $f\colon X \to Y$ is a function: a function $g\colon Y \to X$ which goes back from Y to X and is such that the composition gf is id_X and the composition fg is id_Y is called an **inverse** function for (or of) f.

So, an inverse of f (if it exists!) reverses the effect of f. We show first that if an inverse for a function exists, then it is unique.

Theorem 2.2.2 *If a function $f\colon X \to Y$ has an inverse, then this inverse is unique.*

Proof To see this, suppose that each of g and h is an inverse for f. Thus

$$fg = \mathrm{id}_Y = fh \quad \text{and} \quad gf = \mathrm{id}_X = hf.$$

Now consider the composition $(gf)h = (\mathrm{id}_X)h = h$ (cf. Example 4 on p. 94). By Theorem 2.2.1, this is equal to $g(fh) = g(\mathrm{id}_Y) = g$, so h and g are equal. □

Notation The inverse of f, if it exists, is usually denoted by f^{-1}. It should be emphasised that this is inverse with respect to composition, *not* with respect

to multiplication. Thus the inverse of the function $f(x) = x + 1$, which adds 1, is the function $g(x) = x - 1$, which subtracts 1, not the function $h(x) = 1/(x + 1)$.

Example 2 on p. 93 shows that the inverse of the function $4x - 3$ exists. On the other hand, the function $f\colon \mathbb{Z} \to \mathbb{N}$ given by $f(x) = x^2$ cannot have an inverse. That this is so can be seen in two ways. Either note that since f is not onto, there are natural numbers on which an inverse of f could not be defined (what would '$f^{-1}(3)$' be?). Alternatively, since f is not 1-1, its action cannot be reversed ('$f^{-1}(4)$' would have to be both -2 and 2 but then 'f^{-1}' would not be a well defined function).

The following result gives the precise criterion for a function to have an inverse. Although it is straightforward, the proof of this result may seem a little abstract. You should not be unduly disturbed if you do find it so: the purpose of the various parts of the proof will become clearer as you become more familiar with notions such as surjective and injective.

Theorem 2.2.3 *A function $f\colon X \to Y$ has an inverse if and only if f is a bijection.*

Proof For the first part of the proof, we suppose that f has an inverse and show that f is both injective and surjective. So let f^{-1} denote the inverse of f.

Suppose that $f(x_1) = f(x_2)$. Apply f^{-1} to both sides to obtain

$$f^{-1}(f(x_1)) = f^{-1}(f(x_2)).$$

Thus

$$f^{-1}f(x_1) = f^{-1}f(x_2).$$

Since $f^{-1}f$ is the identity function on X, we deduce that $x_1 = x_2$ and hence that f is injective.

To show that f is surjective, take any y in Y. The composite function ff^{-1} is the identity on Y so $y = f(f^{-1}(y))$. Thus y is of the form $f(x)$ where x is $f^{-1}(y) \in X$ and so f is indeed surjective.

Now we suppose, for the converse, that f is bijective and we define f^{-1} by

$$f^{-1}(y) = x \quad \text{if and only if} \quad f(x) = y.$$

The fact that f is injective means that f^{-1} is well defined (there cannot be more than one x associated to any given y) and the fact that f is surjective means that f^{-1} is defined on all of Y. It follows from the definition that $f^{-1}f$ is the identity on X and that ff^{-1} is the identity on Y. □

The above theorem may be regarded as an 'algebraic' characterisation of bijections. For related characterisations of injections and surjections, see Example 3 on p. 185.

Example Consider the (four) functions from $\{0, 1\}$ to itself, using the notation of Fig. 2.13. The theorem above tells us that, of these, i and f have inverses. In fact, each is its own inverse: i is the identity function $\mathrm{id}_{\{0,1\}}$; $f^2 = i$.

For (many) more examples of bijections, refer forward to Section 4.1.

Corollary 2.2.4 *Let $f: X \to Y$ and $g: Y \to Z$ be bijections. Then*

(i) *gf is a bijection from X to Z, with inverse $f^{-1}g^{-1}$, that is $(gf)^{-1} = f^{-1}g^{-1}$,*
(ii) *$f^{-1}: Y \to X$ is a bijection, with inverse f, that is $(f^{-1})^{-1} = f$. Also*
(iii) *id_X is a bijection (and is its own inverse!).*

Proof (i) By 2.2.3, there exist inverses $f^{-1}: Y \to X$ and $g^{-1}: Z \to Y$ for f and g. Then the composite function $(f^{-1}g^{-1})(gf)$ equals $f^{-1}(g^{-1}g)f = f^{-1}\mathrm{id}_Y f = f^{-1}f = \mathrm{id}_X$. Similarly $(gf)(f^{-1}g^{-1}) = \mathrm{id}_Z$. So the function $gf: X \to Z$ has an inverse $f^{-1}g^{-1}$ so, by 2.2.3, is a bijection.

(ii) We have $f^{-1}f = \mathrm{id}_X$ and $ff^{-1} = \mathrm{id}_Y$ since f^{-1} is the inverse of f. Hence the inverse of f^{-1} is f and, in particular (by 2.2.3), f^{-1} is a bijection.

(iii) It is immediate from the definition that the identity function is injective and surjective. □

This corollary will be of importance when we discuss permutations in Section 4.1.

Finally in this section, we discuss the cardinality of a (finite) set.

Suppose that we have two sets X and Y which have a finite number of elements n and m, respectively. If there is an injective map from X to Y then, since distinct elements of X are mapped to distinct elements of Y, there must be at least n different elements in Y. Thus $n \leq m$. If there is a surjective map from X to Y, there must exist, for each element of Y, at least one element of X to map to it, and so (since each element of X has just one image in Y) there must be at least as many elements in X as in Y: $n \geq m$. Putting together these observations, we deduce that if there is a bijection from X to Y then X and Y have the same number of elements. This observation forms the basis for the following definition (which is due to Cantor).

Definition We say that sets X and Y **have the same cardinality** (i.e. have the same 'number' of elements) and write $|X| = |Y|$ if there is a bijection from X to Y. If X is a non-empty set with a finite number of elements then there is a bijection from X to a set of the form $\{1, 2, \ldots, n\}$ for some integer n; we write $|X| = n$ and say that X has n elements. We also set $|\emptyset| = 0$.

In the above definition, we did not require the sets X and Y to be finite. So we have defined what it means for two, possibly infinite, sets to have the *same* number of elements, without having had to define what we mean by an infinite number (in fact, the above idea was used by Cantor as the basis of his definition of 'infinite numbers').

If you are tempted to think that one would never in practice use a bijection to show that two sets have the same number of elements, then consider the following example (with a little thought, you should be able to come up with further examples).

Example Suppose that a hall contains a large number of people and a large number of chairs. Someone claims that there are precisely the same number of people as chairs. How may this claim be tested? One way is to try (!) to count the number of people and then count the number of chairs, and see if the totals are equal. But there is an easier and more direct way to check this: simply ask everyone to sit down in a chair (one person to one chair!). If there are no people left over and no chairs left over then the function which associates to each person the chair on which they are sitting is a bijection from the set of people to the set of chairs in the hall, and so we conclude that there are indeed the same number of chairs as people. This method has tested the claim without counting either the number of people or the number of chairs.

We can return to explain more clearly a point which arose in the proof of Theorem 1.6.6. There we considered two relatively prime integers a and b, and wished to show that $\phi(ab) = \phi(a)\phi(b)$. The proof given consisted in defining a function f from the set G_{ab} to the set $G_a \times G_b$ by setting $f([t]_{ab}) = ([t]_a, [t]_b)$. It was then shown that f is injective and surjective, so bijective, and hence the result $\phi(ab) = \phi(a)\phi(b)$ followed, since $\phi(n)$ is the number of elements in G_n.

The next result shows how to compute the cardinality of the union of two sets with no intersection.

Theorem 2.2.5 *Let X and Y be finite sets which are disjoint (that is, $X \cap Y = \emptyset$). Then*

$$|X \cup Y| = |X| + |Y|.$$

Proof We include a proof of this (fairly obvious) fact so as to illustrate how one may use the definition of cardinality in proofs.

Suppose that X has n elements: so there is a bijection f from X to the set $\{1, 2, \ldots, n\}$. If Y has m elements then there is a bijection g from Y to the set $\{1, 2, \ldots, m\}$. Define a map h from $X \cup Y$ to the set $\{1, 2, \ldots, n + m\}$ as follows:

$$h(x) = \begin{cases} f(x) & \text{if } x \in X, \\ n + g(x) & \text{if } x \in Y. \end{cases}$$

Since there is no element x in both X and Y, there is no conflict in this two-clause definition of $h(x)$. The images of the elements of X are the integers in the range $\{1, 2, \ldots, n\}$ and the images of the elements of Y are those in the range $\{n + 1, \ldots, n + m\}$. It is easy to check that h is surjective (since both f and g are surjective) and that, since both f and g are injective, h is injective. Thus h is bijective as required. □

Corollary 2.2.6 *Let X and Y be finite sets. Then*

$$|X| + |Y| = |X \cap Y| + |X \cup Y|.$$

Proof The sets $X \cap Y$ and $X \backslash (X \cap Y)$ are disjoint and their union is X. So, by Theorem 2.2.5,

$$|X \backslash (X \cap Y)| + |X \cap Y| = |X|$$

and hence

$$|X \backslash (X \cap Y)| = |X| - |X \cap Y|.$$

Now consider $X \backslash (X \cap Y)$ and Y: these sets are disjoint since if $x \in X \backslash (X \cap Y)$ then x is not a member of $X \cap Y$ and hence is not a member of Y. The union of $X \backslash (X \cap Y)$ and Y is

$$Y \cup (X \cap Y^c) = (Y \cup X) \cap (Y \cup Y^c) = Y \cup X = X \cup Y.$$

So, applying Theorem 2.2.5 again gives

$$\begin{aligned} |X \cup Y| &= |X \backslash (X \cap Y)| + |Y| \\ &= |X| - |X \cap Y| + |Y| \quad \text{(by the above).} \end{aligned}$$

Rearranging gives the required result. □

Example A group of 50 people is tested for the presence of certain genes. Gene X confers the ability to yodel; gene Y endows its bearer with great skill

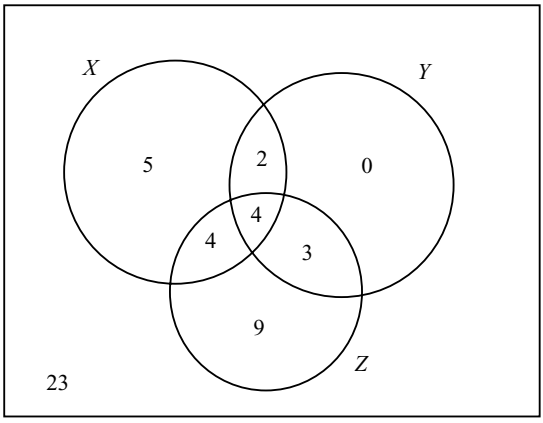

Fig. 2.17

at Monopoly; gene Z produces an allergy to television commercials. It is found that of this group, fifteen have gene X, nine have gene Y and twenty have gene Z. Of these, six have both genes X and Y, eight have genes X and Z and seven have genes Y and Z. Four people have all three genes. How many of this group lack all three of these genes? How many non-yodelling bad Monopoly players are there?

If we draw a Venn diagram as shown in Fig. 2.17, with X being the set of people with gene X and so on, then we may fill in the number of people in each 'minimal region', and so deduce the answers. For example, we are told that the centre region, which represents $X \cap Y \cap Z$, contains four elements. We are also told that the cardinality of $X \cap Y$ is 6. So it must be that $(X \cap Y) \cap Z^{\mathrm{c}}$ has $6 - 4 = 2$ elements. And so on (all the time using 2.2.5 implicitly).

The first question asks for the number of elements of $X^{\mathrm{c}} \cap Y^{\mathrm{c}} \cap Z^{\mathrm{c}}$: that is 23. The second question asks for the cardinality of the set $X^{\mathrm{c}} \cap Y^{\mathrm{c}} = (X \cup Y)^{\mathrm{c}}$: that is 32.

What a mathematician nowadays understands by the term 'function' is very different from what mathematicians of previous centuries understood by the term. Indeed, the way we may regard an expression such as $f(x) = x^2 + 1$ from a purely algebraic point of view would have been foreign to a mathematician of even the eighteenth century, to whom an algebraic expression of this sort would have had very strong geometric overtones. Mathematicians of those times considered that a function must be (implicitly or explicitly) given by a 'rule' of some sort which involves only well understood algebraic operations (addition, division, extraction of roots and so on) together with 'transcendental' functions (such as sine and exponential). Nevertheless, the development of

the calculus, independently by Leibniz and Newton, towards the end of the seventeenth century, raised a host of problems about the nature of functions and their behaviour. Resolution of these problems over the following two centuries necessitated a thorough examination of the foundations of analysis and this was one of the main forces involved in changing the face of mathematics during the nineteenth century to something resembling its present-day form.

The work of Euler was probably the most influential in separating the algebraic notion of a function from its geometric background. As for extending the notion of 'function' beyond what is given explicitly or implicitly by a single 'rule', the main impetus here was the development of what is now called Fourier Analysis. On a methods course, you will probably meet/have met the fact that many (physically defined) functions (waveforms, for example) can be represented as infinite sums of simple terms involving sine and cosine.

The idea of representing certain functions as infinite sums of simple functions, in particular, representation by power series ('infinitely long polynomials'), was well established by the late seventeenth century. The general method was stated by Brook Taylor in his *Methodus Incrementorum* of 1715 and 1717 (hence the term 'Taylor series'), although there were many precursors.

The physical problem whose analysis forced mathematicians to re-examine their ideas concerning functions was the problem of describing the motion of a vibrating string which is given an initial configuration and then released (considered by Johann Bernoulli, then d'Alembert, Euler and Daniel Bernoulli) and, somewhat later, Fourier's investigations on the propagation of heat. The analysis of these problems involved representing functions by trigonometric series.

What was new was the generality of those functions which can be represented by trigonometric series throughout their domain. In particular, such functions need not be given by a single 'rule' and they may have discontinuities (breaks) and 'spikes' – hardly in accordance with what most mathematicians of the day would have meant by a function.

A great deal of controversy was generated and this can be largely ascribed to the fact that the idea of 'function' was not at all rigorously defined (so different mathematicians had different ideas as to what was admissible as a function) and, indeed, was too restrictive.

Even for continuous functions ('functions without breaks') it is 1837 before one finds a definition, given by Dirichlet, of continuous function which casts aside the old restrictions. It is also worth noting that it is Dirichlet who in 1829 presents functions of the following sort: $f(x) = 0$ if x is rational; $f(x) = 1$ if x is irrational. By any standards this is a rather peculiar function (it is discontinuous everywhere): nevertheless it is a function.

The ramifications of all this have been of great significance in the development of mathematics: the reader is referred to [Grattan-Guinness], [Manheim] or one of the more general histories for more on the topic. What we take from all this is the point that it is probably unwise to try to restrict the methods by which a function may be defined. In particular, we have seen above that the modern definition of a function avoids all reference to how a certain function may be specified, but rather concentrates on the most basic conditions that a function must satisfy.

The examples that we have given of functions illustrate that our definition of 'function' is very 'free' and may allow in all kinds of functions which we do not want to consider. But in that case we may simply restrict attention to the kinds of function which are relevant for our particular purpose, whether they be continuous, differentiable, computable, given by a polynomial, or whatever.

Exercises 2.2

1. Describe all the functions from the set $X = \{0, 1, 2\}$ to the set $Y = \{0, 5\}$.
2. Decide which of the following functions are injective, which are surjective and which are bijective:
   (i) $f\colon \mathbb{Z} \to \mathbb{Z}$ defined by $f(x) = x - 1$;
   (ii) $f\colon \mathbb{R} \to \mathbb{R}^+$ defined by $f(x) = |x|$ (where $\mathbb{R}^+$ denotes the set of non-negative real numbers; here $|x|$ is defined to be x if $x \geq 0$ and to be $-x$ if $x < 0$)
   (iii) $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = |x|$;
   (iv) $f\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined by $f(x, y) = x$;
   (v) $f\colon \mathbb{Z} \to \mathbb{Z}$ defined by $f(x) = 2x$.
3. Draw the graphs of (a) the identity function on $\mathbb{R}$, (b) the constant function on $\mathbb{R}$ with value 1.
4. Let $f\colon \mathbb{R} \to \mathbb{R}$ and $g\colon \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x + 1$, $g(x) = x^2 - 2$.
   Find fg, gf, $f^2 (= ff)$ and g^2.
5. Find bijections
   (i) from the set of positive real numbers $\mathbb{R}^+$ to the set $\mathbb{R}$,
   (ii) from the open interval $(-\pi/2, \pi/2) = \{x \in \mathbb{R}\colon -\pi/2 < x < \pi/2\}$ to the set $\mathbb{R}$,
   (iii) from the set of natural numbers $\mathbb{N}$ to the set $\mathbb{Z}$ of integers.
6. Describe all the bijections from the set $X = \{0, 1, 2\}$ to itself.
7. Find the inverses of the following functions $f\colon \mathbb{R} \to \mathbb{R}$:
   (i) $f(x) = (4 - x)/3$;
   (ii) $f(x) = x^3 - 3x^2 + 3x - 1$.

8. Let A, B and C be sets with finite numbers of elements. Show that
$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.$$

9. Show that the following data are inconsistent: 'Of a group of 50 students, 23 take mathematics, 14 take chemistry and 17 take physics. 5 take mathematics and physics, 3 take mathematics and chemistry and 7 take chemistry and physics. Twelve students take none of mathematics, chemistry and physics'. [Hint: see the last example of the section.]

10. Let X be a set and let A be one of its subsets. The **characteristic function** of A is the function $\chi_A \colon X \to \{0, 1\}$ defined by
$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$
(a) Show that if A and B are subsets of X then they have the same characteristic function if and only if they are equal.
(b) Show that every function from X to $\{0, 1\}$ is the characteristic function of some subset of X.
The notation Y^Z is sometimes used for the set of all functions from the set Z to the set Y. Parts (a) and (b) above show that the map which takes a set to its characteristic function is a bijection from the set $P(X)$ of all subsets of X to the set $\{0, 1\}^X$. Since the notation '2' is sometimes used for the set $\{0, 1\}$, this explains why the notation 2^X, instead of $P(X)$, is sometimes used for the set of all subsets of X.

11. Let X be a finite set. Show that $P(X)$ has $2^{|X|}$ elements. That is, in the notation mentioned at the end of the previous example, show that $|2^X| = 2^{|X|}$.
[Hint: to give a rigorous proof, induct on the number of elements of X.]

2.3 Relations

Consider the function $f \colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$. Since this function is not injective (or surjective), it does not have an inverse. So we cannot say that associating to a number its square roots defines a function. On the other hand, there certainly is a relationship between a number and its square roots (if it has any). The mathematical definition of a relation allows us to encompass this more general situation.

Definition Let X, Y be sets. A **relation** R from X to Y is simply a subset of the Cartesian product: $R \subseteq X \times Y$. As an alternative to writing $(x,y) \in R$ we also

write xRy. We say that x is **related** (in the sense of R) to y if $(x,y) \in R$, that is, if xRy. If $X = Y$ we then talk of a relation **on** X.

This definition may seem very abstract and to be rather a long way from what we might normally term a relation. For, in the above definition, the relation R may be any subset of $X \times Y$: we do not insist on a 'material' connection between those elements x, y such that $(x,y) \in R$. We do find, however, that any normal use of the term relation may be covered by this definition. And, as with the case of functions, the advantage of making this wide, abstract definition is that we do not limit ourselves to a notion of 'relation' which might, with hindsight, be seen as overly restrictive.

The following examples give some idea of the variety and ubiquity of relations.

Example 1 Let $\mathbb{N}$ be the set of natural numbers and consider the relation '$\leq$' on $\mathbb{N}$. This is defined by the condition: $x \leq y$ if and only if x is less than or equal to y $(x, y \in \mathbb{N})$. Alternatively, it may be defined arithmetically by $x \leq y$ if and only if $y - x \in \mathbb{N}$. However one chooses to define it, it is a relation in the sense of the above definition: let $X = \mathbb{N} = Y$ and take the subset $R = \{(x, y)\colon x \leq y\}$ of $\mathbb{N} \times \mathbb{N}$. Then $(x,y) \in R$ if and only if $x \leq y$.

One may note that relations are often specified, not by directly defining a subset of $X \times Y$, but rather, as in this example, by specifying the condition which must be satisfied for elements x and y to be related.

Notation of the sort 'xRy' is more common than '$(x,y) \in R$': the relations of 'less than or equal to' and 'equals' are usually written $x \leq y$ and $x = y$.

Example 2 Any function $f\colon X \to Y$ determines a relation: namely its graph $\mathrm{Gr}(f) = \{(x, y)\colon x \in X \text{ and } y = f(x)\} \subseteq X \times Y$ (as introduced in Section 2.2). Thus, we may define the associated relation R either by saying that R is the set $\mathrm{Gr}(f)$ or by setting xRy if and only if $y = f(x)$. Thus a function $f\colon X \to Y$, when regarded as a subset of $X \times Y$, is just a special sort of relation (namely, a relation which satisfies the condition: every $x \in X$ is related to exactly one element of Y).

Example 3 We can define the relation R on the set of real numbers to be the set of all pairs (x,y) with $y^2 = x$. Thus xRy means 'y is a square root of x'. As we mentioned in the introduction to this section, this is not a function, but it is a relation in the sense that we have defined.

Example 4 Let X be the set of integers and let R be the relation: xRy if and only if $x - y$ is divisible by 3.

Thus $1R4$ and $1R7$ but not $2R3$.

Example 5 Let $X = \{1, 2, \ldots, 11, 12\}$ and let D be the relation 'divides' – so xDy if and only if x divides y. As a subset of $X \times X$,

$D = \{(m,n)\colon m \text{ divides } n\}$.

Thus, for example $(4,8) \in D$ but $(4,10) \notin D$. (As an exercise in this notation, list D as a subset of $X \times X$.)

Example 6 Let C be the set of all countries. Define the relation B on C by cBd if and only if the countries c,d have a common border.

Here are some rather more 'abstract' relations.

Example 7 Let X and Y be any sets. Then the empty set $\emptyset$, regarded as a subset of $X \times Y$, is a relation from X to Y (the 'empty relation', characterised by the condition that no element of X is related to any element of Y). Another relation is $X \times Y$ itself – this relation is characterised by the fact that every element of X is related to every element of Y.

Example 8 Let X be any set. Then the relation $R = \{(x,x)\colon x \in X\}$ is the 'identity relation' on X: that is, xRx if and only if $x = x$. In other words, this is the relation 'equals'.

Example 9 Given a relation R from a set X to a set Y, we may define the dual, or complementary, relation to be $R^{\mathrm{c}} = (X \times Y)\backslash R$. Thus $x R^{\mathrm{c}} y$ holds if and only if xRy does not hold. For instance, the dual, R^{c}, of the identity relation on a set is the relation of being unequal: $x R^{\mathrm{c}} y$ if and only if $x \neq y$.

Also, we may define the 'reverse' relation, R^{rev}, of R to be the relation from Y to X which is defined to be $R^{\mathrm{rev}} = \{(y,x)\colon (x,y) \in R\}$.

If $f\colon X \to Y$ is a function, then the reverse relation from Y to X is the subset $\{(f(x), x)\colon x \in X\}$ of $Y \times X$. This relation will be a function (namely f^{-1}) if and only if f is a bijection.

Observe that the complement and reverse are quite different. Take, for instance, X to be the set of all people (who are alive or have lived). Define the relation R by xRy if and only if x is an ancestor of y. Then the dual R^{c} of R is the relation defined by $x R^{\mathrm{c}} y$ if and only if x is not an ancestor of y. Whereas the reverse relation R^{rev} is defined by $x R^{\mathrm{rev}} y$ if and only if x is a descendant of y.

Definitions Let R be a relation on a set X. We say that R is

(1) **reflexive** if xRx for all x in X,
(2) **symmetric** if, for all $x, y \in X$, $x R y$ implies yRx,

(3) **weakly antisymmetric** if, for all $x,y \in X$, whenever xRy and yRx hold one has $x = y$,
(3′) **antisymmetric** if, for all $x,y \in X$, if xRy holds then yRx does not,
(4) **transitive** if, for all $x,y,z \in X$, if xRy and yRz hold then so does xRz.

We reconsider some of our examples in the light of these definitions.

Example 1 The relation '$\leq$' is reflexive ($x \leq x$), not symmetric (e.g. $4 \leq 6$ but not $6 \leq 4$), weakly antisymmetric ($x \leq y$ and $y \leq x$ implies $x = y$), transitive ($x \leq y$ and $y \leq z$ imply $x \leq z$).

Example 2 This example is not a relation on X unless $X = Y$: rather a relation between X and Y, so (note) the above definitions do not apply.

Example 3 This relation is not reflexive (x^2 is not equal to x in general), not symmetric, and not transitive ($4R16$ and $2R4$, but not $2R16$). Let us look at weak antisymmetry in more detail. Suppose xRy and yRx: so $x^2 = y$ and $y^2 = x$. Thus $x = x^4$ and, since x is real, x is either 0 or 1. Then y is also 0, respectively 1, since $y = x^2$. It follows that the relation is weakly antisymmetric.

Example 4 The relation is reflexive, symmetric and transitive but not (weakly) antisymmetric. First, consider reflexivity. Since 0 is divisible by 3 and $0 = x - x$, we see that, for every $x \in X$, xRx. For symmetry, note that if xRy then $x - y$ is divisible by 3 and so $y - x$ is divisible by 3; that is, yRx. For transitivity, suppose xRy and yRz, so $x - y$ is divisible by 3, as is $y - z$. Then $(x - y) + (y - z) = x - z$ is divisible by 3 and so xRz holds. Regarding antisymmetry: note that $3R6$ and $6R3$ yet 3 and 6 are not equal.

Example 5 This relation is reflexive, not symmetric (for example, 2 divides 4 but 4 does not divide 2), weakly antisymmetric and transitive.

As an exercise, you should examine Examples 6 to 8 in the light of these definitions.

In dealing with conditions such as those above, one should be careful over logic. For instance, a relation R on X is symmetric if and only if for every x and y in X, xRy implies yRx (exercise: is the empty relation $R = \emptyset \subseteq X \times X$ symmetric?). So, in order to show that a relation R is not symmetric, it is enough to find *one* pair of elements a,b such that aRb holds but bRa does not. For another example, to show that a relation is transitive, it must be shown that for *every* triple a,b,c, if aRb and bRc hold then so does aRc: it is not enough to check it for just some values of a, b and c.

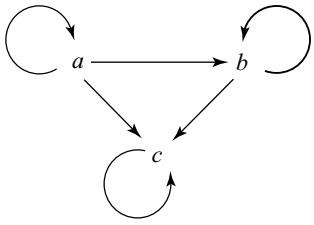

Fig. 2.18

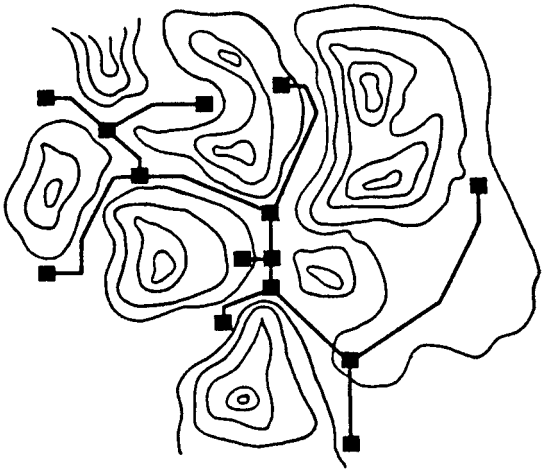

Fig. 2.19 Thin lines, contours; thick lines, railways; small squares, towns.

For more exercises in logic, see Exercises 2.3.2 and 2.3.3 at the end of this section as well as all those in Chapter 3.

Definition A useful pictorial way to represent a relation R on X is by its **digraph** (or **directed graph**) $\Gamma(R)$. To obtain this, we use the elements of X as the vertices of the graph $\Gamma(R)$ and join two of these vertices, x and y, by a directed edge (a directed arrow from x to y) whenever xRy.

Example Let X be the set $\{a,b,c\}$ and let R be the relation specified by

$$aRa, bRb, cRc, aRb, aRc, bRc.$$

The digraph of R is as shown in Fig. 2.18.

Example Fig. 2.19 is a map showing a number of towns in mountain valleys, and the railway network that connects them. Define the relation R on the set

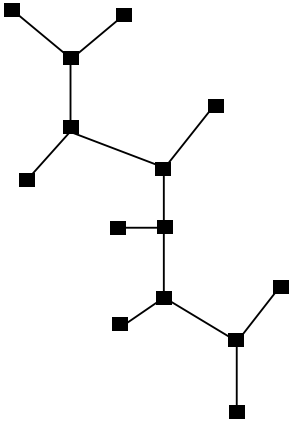

Fig. 2.20

of towns by aRb if and only if a and b are next to each other on the railway line.

The relation is symmetric so, if the directed graph of the relation has a directed edge going from a to b, then it also has one going from b to a. So we make the convention that an edge without any arrow stands for such a pair of directed edges. With this convention, the graph of the relation is as shown in Fig. 2.20.

A relation on a set may be specified by giving its digraph: the set X is recovered as the set of vertices of the digraph, and the pair (x,y) is in the relation if and only if there is a directed edge going from the vertex x to the vertex y.

Yet another way to specify a relation R on a set X is to give its **adjacency matrix**. This is a matrix with rows and columns indexed by the elements of X (listed in an arbitrary but fixed order). Each entry of the matrix is either 0 or 1. The entry at the intersection of the row indexed by x and the column indexed by y is 1 if xRy is true, and is 0 if xRy is false. For convenience, we present examples of adjacency matrices in tabular form.

Example Let $X = \{a, b, c\}$ and $R = \{(a, a),\ (b, b),\ (c, c),\ (a, b),\ (a, c),\ (b, c)\}$: so R is the relation with the digraph in Fig. 2.18. Its adjacency matrix is as shown:

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

It is possible to interpret some of the properties of relations in terms of their adjacency matrices. Thus a relation is reflexive if the entries down the main diagonal are all 1, it is symmetric if the matrix is symmetric (that is, if the entry at position (x,y) is equal to that at (y,x)) and it is weakly antisymmetric if the entries at (x,y) and (y,x) are never both 1 unless $x = y$. The transitivity of R can also be characterised in terms of the adjacency matrix but, since this is considerably more complicated, we omit this (see, for example [Kalmanson, p. 330]). We can immediately see that the above relation is reflexive, weakly antisymmetric and (if we check case by case) we can see that it is transitive.

Example For the set $\{1, 2, \ldots, 11, 12\}$ and the relation D (xDy if and only if x divides y) the adjacency matrix is

	1	2	3	4	5	6	7	8	9	10	11	12
1	1	1	1	1	1	1	1	1	1	1	1	1
2	0	1	0	1	0	1	0	1	0	1	0	1
3	0	0	1	0	0	1	0	0	1	0	0	1
4	0	0	0	1	0	0	0	1	0	0	0	1
5	0	0	0	0	1	0	0	0	0	1	0	0
6	0	0	0	0	0	1	0	0	0	0	0	1
7	0	0	0	0	0	0	1	0	0	0	0	0
8	0	0	0	0	0	0	0	1	0	0	0	0
9	0	0	0	0	0	0	0	0	1	0	0	0
10	0	0	0	0	0	0	0	0	0	1	0	0
11	0	0	0	0	0	0	0	0	0	0	1	0
12	0	0	0	0	0	0	0	0	0	0	0	1

Certain general types of relations, characterised by combinations of properties such as symmetry, reflexivity, ... frequently arise in mathematics and, indeed, in many spheres. We consider what are probably the two most important: partial orderings and equivalence relations.

Definition A relation R on a set X is a **partial order(ing)** if R is reflexive, weakly antisymmetric and transitive. Thus, for all $x \in X$ one has xRx; for all $x,y \in X$ if xRy and yRx then $x = y$; for all $x,y,z \in X$, if xRy and yRz then xRz.

Example 1 Define a relation R on the set of real numbers by xRy if and only if $x \leq y$. This relation is one of the most familiar examples of a partial order in mathematics. In many examples of partial orders which arise in practice, there will be some sense in which the relation 'xRy' can be read as 'x is smaller than or equal to y'.

Example 2 Both examples discussed above in connection with adjacency matrices are partial orders.

Example 3 Let A be the set $\{a, b, c\}$ and define X to be the set of all subsets of A: so X has the 8 elements

$$\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, A.$$

Define a relation R on X by $(U,V) \in R$ if and only if U is a subset of V. Then R is a partially ordered set whose adjacency matrix is as shown:

	$\emptyset$	$\{a\}$	$\{b\}$	$\{c\}$	$\{a,b\}$	$\{a,c\}$	$\{b,c\}$	A
$\emptyset$	1	1	1	1	1	1	1	1
$\{a\}$	0	1	0	0	1	1	0	1
$\{b\}$	0	0	1	0	1	0	1	1
$\{c\}$	0	0	0	1	0	1	1	1
$\{a,b\}$	0	0	0	0	1	0	0	1
$\{a,c\}$	0	0	0	0	0	1	0	1
$\{b,c\}$	0	0	0	0	0	0	1	1
A	0	0	0	0	0	0	0	1

One may define a similar partial order on the set of all subsets of any set. Thus, if X is the set of all subsets of a set A, the relation R on X defined by

$$(B, C) \in R \text{ if and only if } B \text{ is a subset of } C$$

is a partial order.

Example 4 The relation D on the set $\mathbb{Z}$ of integers which is given by xDy if and only if x divides y is another example of a partial order.

We may also define a **strict partial order** to be a set X with a relation R on it which is antisymmetric and transitive. For instance, the relation '$<$' on $\mathbb{N}$ given by $x < y$ if and only if $y - x$ is positive is a strict partial order. Another example, defined on the set of all people (past and present), is given by aRb if and only if a is an ancestor of b. (It is a strict partial order since a person cannot be an ancestor of himself or herself.)

It is straightforward to show that if R is a partial order on X then the relation S on X defined by xSy if and only if xRy and $x \neq y$ is a strict partial order. Conversely, if S is a strict partial order on the set X then the relation R defined by xRy if and only if xSy or $x = y$ is a partial order on X. The notation $(P, \leq)$ and term **partially ordered set** (or **poset** for short) are often used for a set P equipped with a partial order $\leq$.

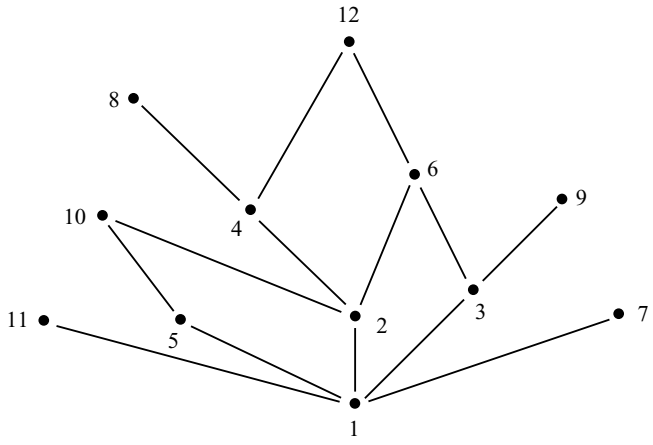

Fig. 2.21

There is a graphical way of representing partially ordered sets which have a finite number of elements: by use of Hasse diagrams.

Definition Let R be a strict partial order on a set X. If x, y are elements of X then y is an **immediate successor** of x (and x is an **immediate predecessor** of y) if xRy and if there is no z in X with xRz and zRy. Roughly, y is an immediate successor of x if y is 'greater than' x and if there is no element strictly between x and y.

In the case that R is a partial order (as opposed to a strict partial order), we modify the definition in the obvious way, saying that y is an immediate successor of x if xRy and $x \neq y$ and if, whenever $z \in X$ is such that xRz and zRy, then either $z = x$ or $z = y$.

The **Hasse diagram** of the (strict) partial order R on the set X is obtained as follows. Place one point on the plane for each element of X. The points must be placed in such a way that a line may be drawn going in a general upwards direction from each element x in X to each of its immediate successors. Draw in these lines.

Example 1 Let $X = \{1, 2, \ldots, 11, 12\}$ and let R be given by xRy if and only if x divides y. Note that 6 is an immediate successor of 2, but 12 is not, since $2R6$ and $6R12$. The Hasse diagram is as shown in Fig. 2.21.

Example 2 Let $X = \{1, 2, \ldots, 11, 12\}$ and let R be the usual ordering '$\leq$'. The Hasse diagram is then Fig. 2.22.

Example 3 Let X be the set of all subsets of $\{0, 1, 2\}$ and let R be the relation 'is a subset of': Fig. 2.23.

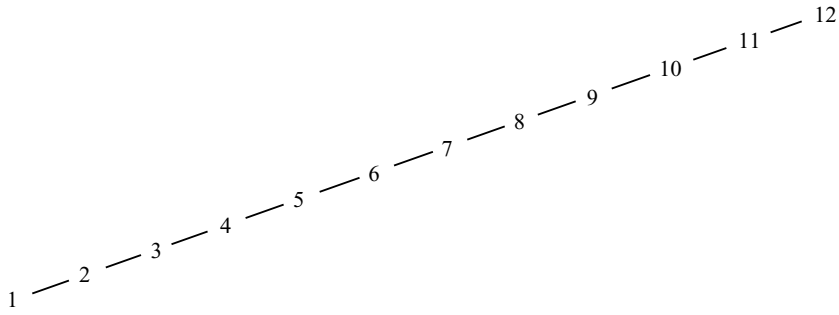

Fig. 2.22

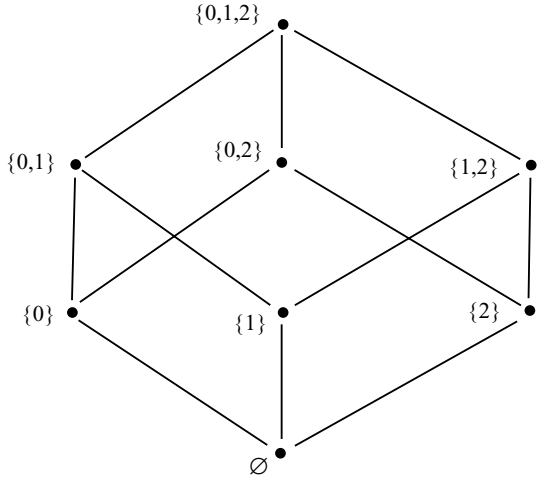

Fig. 2.23

Example 4 Let X be the set of integers $\{1, 2, 3, 6, 9, 18\}$ and let R be given by xRy if and only if x divides y. The Hasse diagram is as shown in Fig. 2.24.

We now come to our second special type of relation.

Definition A relation R on the set X is an **equivalence relation** if R is reflexive, symmetric and transitive.

Example 1 For any set X, the identity relation R on X, given by xRy, if and only if $x = y$, is an equivalence relation.

Example 2 Let X be the set of all integers and fix an integer $n \geq 2$. Define a relation E on X by aEb if and only if $a - b$ is divisible by n. Thus E is the relation

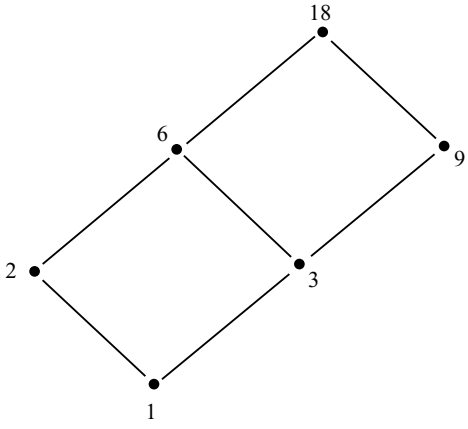

Fig. 2.24

of congruence modulo n. It is quickly checked that E satisfies the conditions for being an equivalence relation.

Example 3 The notion of logical equivalence (see Section 3.1 below) is, as its name suggests, an equivalence relation on the set of propositional terms.

Example 4 (For the reader who has met some linear algebra.) Let X be the set of $n \times n$ matrices with real entries. Matrices A and B in X are defined to be 'similar' if there are invertible matrices P and Q such that $B = P^{-1}AQ$. It is straightforward to check that similarity of matrices is an equivalence relation (i.e. if we define the relation S on X by $(A,B) \in S$ if and only if A is similar to B, then S is an equivalence relation). Matrices $A,B \in X$ are said to be 'equivalent' if there is an invertible matrix P such that $B = P^{-1}AP$. It is easy to verify that the relation E of equivalence in this sense is also an equivalence relation.

Example 5 Let $f\colon X \to Y$ be a function. Define a relation F on X by xFx_1 if and only if $f(x) = f(x_1)$. Then F is an equivalence relation.

Definition Let X be any set. By a partition of X we mean a particular way of dividing up the set X into 'blocks'. More precisely: a **partition** of X is a collection $\{X_i\colon i \in I\}$ of non-empty subsets of X which is

disjoint, in the sense that $X_i \cap X_j = \emptyset$ if i is different from j, and
covering, in the sense that each x in X belongs to one (and by disjointness, only one) X_i.

Example A group of people is divided up into separate teams to work on various projects: this 'division' determines a partition of the set of people involved, with each team constituting one member of the partition (so the 'X_i' are the various teams).

We now have on the one hand partitions, on the other hand equivalence relations. We will see that these amount to the same thing (this is not to say that the intuitive ideas coincide but rather that when the ideas are formalised mathematically we obtain 'equivalent' notions).

It may be worthwhile to warn the reader that the following proof may appear more 'abstract' than any in this book so far (mainly because the objects which the theorem refers to – equivalence relations and partitions – probably seem less concrete than numbers or even sets and functions). If you feel that the theorem and its proof do not make much sense to you, try picking a *particular* equivalence relation or partition and then going through the proof, working out what the details of the proof mean in the context of the example that you have chosen.

Theorem 2.3.1 *Let X be any set.*

(i) *Suppose that $\{X_i : i \in I\}$ is a partition of X. Then the relation R on X which is defined by xRy if and only if x and y belong to the same member of the partition is an equivalence relation.*

(ii) *If, conversely, E is an equivalence relation on X then E determines the partition whose 'blocks' X_i are the* **equivalence classes** *$[x]_E = \{y \in X : yEx\}$ of members x of X.*

Proof (i) Suppose that we are given a partition $\{X_i : i \in I\}$ of X. The relation R as defined is clearly reflexive. Also R is symmetric since if x and y are both in X_i then so are y and x. Finally R is transitive since if x and y are both in X_i (say) and if y and z are both in X_j (say) then, by disjointness, $i = j$ and so x and z lie in the same member of the partition.

(ii) Now suppose that E is an equivalence relation on X. Define $[x] = [x]_E$ to be the subset $\{y \in X : yEx\}$, containing x. It is claimed that the distinct sets of this kind form a partition of X.

First note that if $b \in [a]$ then $[b] = [a]$. For if $c \in [a]$ then we have cEa. Since $b \in [a]$ we also have bEa and so, by symmetry, aEb. Transitivity then implies cEb and so $c \in [b]$. Thus we have shown $[a] \subseteq [b]$. For the converse, suppose $d \in [b]$, so dEb. Since also bEa, transitivity yields dEa: that is, $d \in [a]$, as required.

Disjointness now follows quickly. Suppose that $[a]$ and $[b]$ have an element c (say) in common. Then by the above $[c] = [a]$ and also $[c] = [b]$. Hence $[a] = [b]$, as required.

Finally, the sets are covering since each element a is in some set of the form $[x]$: namely $[a]$. Thus we do have a partition of X. □

As one example of this essential equivalence between partitions and equivalence relations (via equivalence classes) consider the notion of congruence modulo n. The equivalence classes determined by the relation aRb if and only if $n|a-b$ are the (n) congruence classes modulo n. Conversely, given the partition of $\mathbb{Z}, \{nk: k \in \mathbb{Z}\}, \{nk+1: k \in \mathbb{Z}\}, \ldots, \{nk+(n-1): k \in \mathbb{Z}\}$, the corresponding equivalence relation is R.

For another example, take X to be the set of all the points on the real plane $\mathbb{R} \times \mathbb{R}$, apart from the origin (0,0): define an equivalence relation on X by setting x equivalent to y if and only if (0,0), x and y all lie on a straight line. By Theorem 2.3.1 this equivalence relation determines a partition of the plane into a disjoint covering family of subsets. These subsets are the equivalence classes of points and they are simply the straight lines (minus the origin) which pass through the origin.

De Morgan and C.S. Peirce studied relations in the abstract in the latter part of the nineteenth century.

Exercises 2.3

1. For each of the following relations R on the set X decide whether R is reflexive, symmetric, (weakly) antisymmetric or transitive:
   (a) $X = \mathbb{Z}$, aRb if and only if $a \leq b+1$;
   (b) $X = \mathbb{Z}$, aRb if and only if $a+b$ is even;
   (c) $X = \mathbb{P}$, aRb if and only if a and b are coprime:
   (d) $X = \mathbb{Z}$, aRb if and only if $a+b$ is divisible by 3;
   (e) $X = \mathbb{R}$, aRb if and only if $a^2 \leq b^2$;
   (f) $X = \mathbb{R} \times \mathbb{R}$, $(a,b)R(c,d)$ if and only if $a = c$;
   (g) $X = \mathbb{N} \times \mathbb{N}$, $(a,b)R(c,d)$ if and only if either $a < c$ or $(a = c$ and $b \leq d)$.
2. (a) Show that any relation which is both symmetric and antisymmetric must be the empty relation.
   (b) Show that if a relation is antisymmetric then it is weakly antisymmetric.
   (c) Give an example of a non-empty relation which is symmetric and weakly antisymmetric (!).
   (d) Show that if a relation is symmetric then so is its complement.
   (e) Show that if a relation is transitive then so is its reverse.
   (f) Give an example of a non-empty relation which is symmetric and transitive but which is not reflexive.

3. What is wrong with the following argument?
   '*Theorem*' Every transitive symmetric relation is reflexive.
   '*Proof*' Let R be a relation on the set X and suppose that R is transitive and symmetric. Let $x \in X$. From xRy we have, by symmetry, yRx, and so, by transitivity, we deduce xRx. Thus R is reflexive, as required.
   There is something wrong with the argument, since the 'Theorem' is false – if you look at the solution for Exercise 2.3.2 (f), you will find an example of a transitive symmetric relation which is not reflexive!
4. Let X be the set of all European countries (choose your own definitions of 'European' and 'country'): alternatively let X be the set of all states of the USA. Define the relation B on X by xBy if and only if x and y have a common border (let us make the convention that x does not have a common border with itself). Draw the digraph of this relation. (Since the relation is symmetric, it would be reasonable to use an edge without an arrow to stand for a pair of directed edges between two vertices.)
5. Let X be the set $\{a, b, c, d, e\}$ and R be the relation whose adjacency matrix is shown. Prove that R is a partial order on X and draw its Hasse diagram.

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

6. Let X be the set $\{1, 2, 3, 4\}$ and let

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

   Write down the adjacency matrix for R. Show that R is a partial order and draw its Hasse diagram.
7. Let X be the set $\{1, 2, 3, 4\}$ and let

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

   Show that R is an equivalence relation and write down its equivalence classes.
8. Let $S = \{1, 2, 3, 4\}$ and let X be the set $S \times S$. Define a relation R on X by

$$(a, b)R(c, d) \text{ if and only if } a + b = c + d.$$

   Show that R is an equivalence relation and list the equivalence classes.

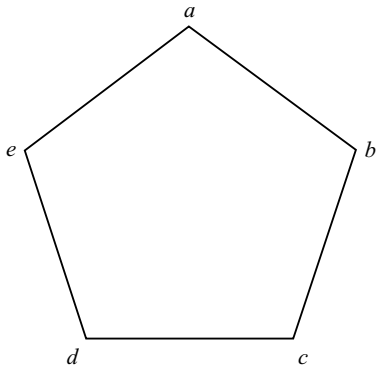

Fig. 2.25

9. Let a pentagon have vertices denoted a to e as shown in Fig. 2.25. Define a relation R on the set of vertices, by aRb if and only if a and b do not lie on the same edge of the pentagon. Decide whether or not R is transitive and draw the digraph of R.
10. Let X be the set of $n \times n$ matrices with real entries and let S and E be respectively the (equivalence) relations of similarity and equivalence as defined in Example 4 on p. 113. Show that the partition of X corresponding (in the sense of 2.3.1) to E **refines** that corresponding to S, in the sense that for all matrices A and B, $(A, B) \in E$ implies $(A, B) \in S$ (and hence every S-equivalence class is a disjoint union of E-equivalence classes).

2.4 Finite state machines

'Calculating machines' have a longer history than is often realised. The first mechanical digital calculator was built by Blaise Pascal sometime between 1642 and 1644. Pascal's machine was limited to addition and subtraction but, in 1673, Leibniz built a machine which could also multiply and divide (it is interesting to note, in connection with this, Leibniz' advocacy of the binary system of numeration and his dream of a 'logical calculus'). Calculating machines based on Leibniz' design were in general use until they were very recently supplanted by electronic calculators.

Charles Babbage (1792–1871) designed mechanical calculators which would carry out a sequence of computations. He proposed to the Royal Astronomical Society in 1822 that he build a giant 'Difference Engine' – a

machine which would compute and even set in type mathematical tables (the current hand-calculated ones were infested with errors). He had already built a small machine of this kind. The construction began well but foundered over financial and related difficulties, and the 'Difference Engine' was never completed.

Babbage went on to conceive a much more sophisticated calculating machine which would have a 'memory' and would be programmable by punched cards (such cards were already used to control weaving looms). But again financial difficulties and Babbage's seeming inability ever to stop tinkering and to finish a project, brought the construction to a standstill.

Babbage's work slipped into obscurity but advances were made, such as the development by Hollerith, towards the end of the century, of punched card systems for handling large masses of data.

Large analogue (as opposed to digital) computers – called 'differential analysers' – were built in the USA in the 1930s, and prototype digital computers were built by various scientists in the USA and UK in the late 1930s and early 1940s. By that stage, electrical and electronic components rather than mechanical ones were being used.

In a paper of 1937, Alan Turing described a theoretical computing machine which would be able to compute according to any rule or set of rules fed to it – a programmable computer. Such a theoretical machine is now called a Turing machine and could, if given the appropriate instructions and enough time and space, perform any computation which might be described as algorithmic (i.e. proceeding according to some rule or set of rules).

What was probably the first working electronic computer, named COLOSSUS, became operational in 1943 at Bletchley – a top-secret code-breaking centre in England. This machine was built by Turing and others who were engaged in deciphering German secret messages.

The first general-purpose electronic computer – ENIAC – became operational in 1944–5. It was built by a team at the Moore School (attached to the University of Pennsylvania) in Philadelphia.

The first stored-program digital computer, the 'Baby', was built by Tom Kilburn and Freddie Williams at the University of Manchester and became operational in 1948.

Of course, present-day computers are immensely more powerful and faster than those original ones, but all can be seen as realisations of Turing's idea.

A Turing machine can, in principle, model any computation. In this section, we will consider a restricted class of Turing machines. Although these do not have the flexibility of Turing machines (there are certain computations which they cannot perform), they are relatively easy to construct in practice and the

class of computations which they can perform has certain properties which are interesting from the point of view of formal language theory.

There are several, related, types of finite state machines. We first consider the most basic of these.

Definition A **finite state machine** M is a triple (S,A,μ) where S is the set of **states** of M and includes a distinguished **initial state** 0, A is an **alphabet** – its elements are called **letters** – and μ is a **transition function** $\mu\colon S \times A \to S$. Thus μ assigns a state to every pair of the form (s,a) where s is a state and a is a letter of A.

The picture to bear in mind is that of a machine with an input tape which it can read: the entries on the tape are letters of the alphabet A. It will operate in a discrete manner (as does any digital computer). When it is started up, its internal configuration or state is 0. It reads the first letter on the tape. Depending on what that letter is, its internal state may change, becoming i say. It then reads the next letter on the tape: that letter, together with its current internal state, determines what its new state is to be, ... and so on. Thus, at a given time, the machine is in a certain internal state; it reads a letter; it then moves to another (or remains in the same) state and begins to read the next letter on the tape. The transition function $\mu\colon S \times A \to S$ has the following interpretation: if the machine is in a state s and it reads the letter a on the tape then its internal state becomes $\mu((s,a))$.

We will usually denote the states of M by integers and 0 will always denote the initial state. The members of A will usually be denoted by letters of the Roman alphabet. We will make the convention that the machine begins to read the tape at its left-most end and always moves from left to right. Let us consider two examples.

Example 1 Let M be the machine with three states $\{0, 1, 2\}$, alphabet $A = \{\text{a}, \text{b}\}$ and μ given by the table shown (the entry at the intersection of the row labelled i and the column labelled x is the value $\mu(i,x)$).

	a	b
0	0	1
1	1	2
2	2	0

Consider the sequence baababa. Initially, the machine is in state 0. On reading the first b, the machine moves into state 1. It remains in 1 after reading the two a's and moves into state 2 after reading the next b. It stays there after reading

the a, but moves back to state 0 on reading the next b and then remains in state 0 on reading the final a.

Example 2 Let M be the machine with states $\{0, 1, 2, 3\}$, alphabet $A = \{\text{a,b,c}\}$ and with μ given by

	a	b	c
0	1	2	1
1	0	3	2
2	2	1	2
3	3	3	3

We consider what this machine does on reading the sequence abcab: starting in state 0, reading a takes the machine into state 1, reading b gives state 3, reading c leaves it in 3, as will all further transitions. Thus the machine reads abcab and moves from initial state 0 to final state 3.

An alternative way to give the transition function μ is to use the **state diagram**. This is a directed graph with labelled edges. The vertices (usually denoted by numbers in circles) of this graph are the elements of S. Two vertices i and j are joined by an arrow with label a (for instance) if when the machine is in state i and reads letter a it moves to state j. Thus the state diagrams of our two examples are as shown in Fig. 2.26.

We now discuss a type of finite state machine that is very important for applications. Such machines are components of cash machines, lifts and, indeed, they are used in computers to recognise, for instance, key words of a programming language and to respond appropriately. They are also important in theoretical computer science and in the theory of formal languages.

Definition A finite state **automaton** is a finite state machine $M = (S, A, \mu)$ together with a subset F of S, known as the set of **acceptance states** of M.

We may regard automata as being intended to *recognise* certain sequences of letters (words) in the alphabet A (a password or PIN, for example). Formally, we have the following definition.

Definition Let $M = (S, A, \mu)$ with set F of acceptance states be a finite state automation. A sequence σ of letters in the alphabet A is **accepted by** M if, after reading the sequence σ, the automaton is in state s for some s in F.

Example 1 We return to the first example above and let F consist of the state 2. Thus a sequence of a's and b's is accepted by the automaton if it is in the state

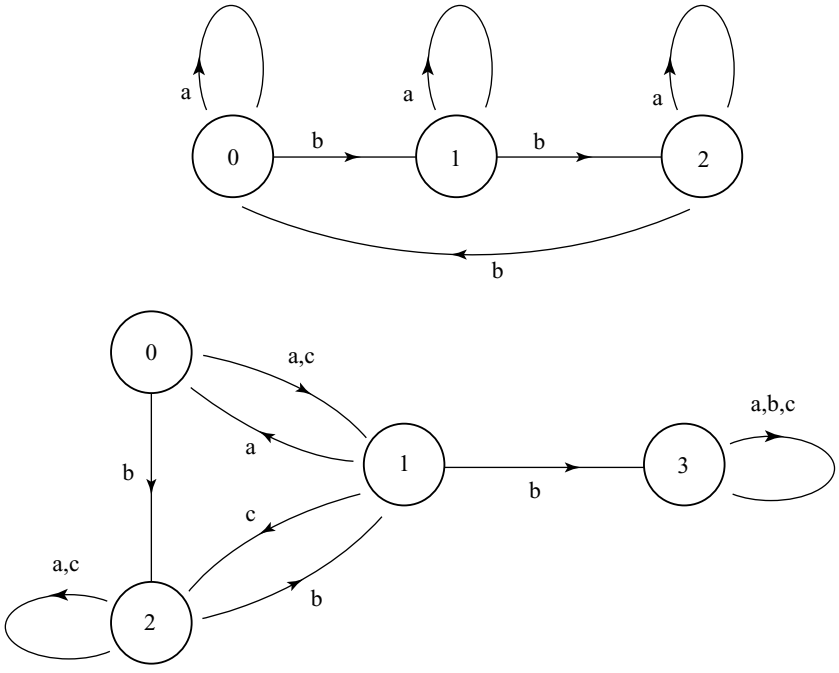

Fig. 2.26

2 after the sequence has been read. If a sequence consisting entirely of a's is read, the automaton stays in state 0, and so this sequence is not accepted. If our sequence has one b in it, the automaton arrives at state 1. For a sequence with two b's, such as aabaaba, the automaton is sent into state 1 when it encounters the first b and, on encountering the second, moves to state 2 where it remains: so such a sequence is accepted. Continuing in this way, we can see that sequences with three or four b's are not accepted and that, in general, a sequence is accepted if and only if the number of b's in the sequence is congruent to 2 modulo 3.

If the set F of acceptance states were changed to be $\{1\}$ then the set of sequences accepted would be those in which the number of b's is congruent to 1 modulo 3. Similarly, if F were $\{0\}$, the sequences accepted would be precisely those with k b's, where k is divisible by 3.

Example 2 Let M be the automaton with states $\{0, 1, 2, 3\}$ and alphabet {a,b}, for which F is $\{1\}$ and μ is given by the state diagram shown in Fig. 2.27. Consider the words accepted by this automaton. The a's it reads move the machine back and forth between states 0 and 1 and between states 2 and 3. Movement between states 0 and 3 and between 1 and 2 is governed by the bs

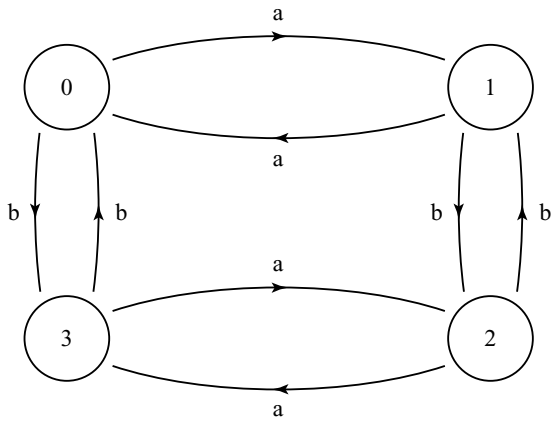

Fig. 2.27

read. If a word has an even number of as, the automaton must be in one of the states 0 or 3 after reading the word. For a word with an odd number of as, the automaton will be in either state 1 or state 2. Similarly, after reading a word with an even number of b's, the automaton is in one of the states 0 or 1 and it is in state 2 or 3 after reading a word with an odd number of bs. Since $F = \{1\}$, the accepted words are those with an odd number of as and an even number of bs.

Another special type of finite state machine arises by modifying the machine to produce output. This is done by considering A as the input alphabet, adding an output alphabet B and giving an output table ν: $S \times A \to B$ which is a rule assigning an output symbol $\nu(i,a)$ to each pair consisting of a state i and an input a. If the machine is in state i and reads the letter a then it outputs $\nu(i,a)$ before moving to state $\mu(i,a)$.

Example 1 We return to the example where M is the machine with states $\{0, 1, 2, 3\}$, alphabet $A = \{\text{a}, \text{b}, \text{c}\}$ and where μ is given by

	a	b	c
0	1	2	1
1	0	3	2
2	2	1	2
3	3	3	3

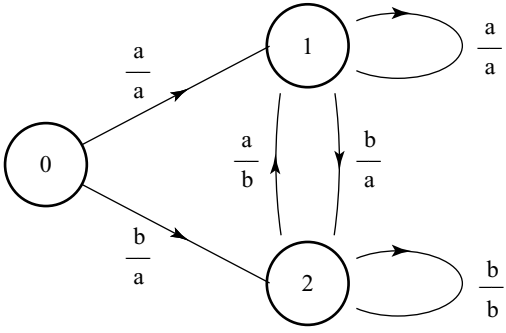

Fig. 2.28

Thinking of A as the input alphabet, we let B be the set $\{\alpha, \beta\}$ and define v by the table

	a	b	c
0	α	α	β
1	α	β	α
2	β	β	α
3	α	α	α

Thus on reading the sequence acbaab, the machine would go through the sequence of states 0,1,2,1,0,1,3 and would output $\alpha\alpha\beta\alpha\alpha\beta$.

Example 2 As another example of a machine with output, we consider the unit delay machine M. This has states $\{0, 1, 2\}$, input and output alphabets $\{a,b\}$ and the functions μ,v given by the tables as shown:

	μ Next state		v Output	
	a	b	a	b
0	1	2	a	a
1	1	2	a	a
2	1	2	b	b

The tables for μ and v have been combined in an obvious way. The state diagram of this machine is given in Fig. 2.28.

Each arrow in this diagram is labelled by two letters of A. The upper letter is the input required to move in the direction of the arrow and the lower is the corresponding output. Thus if the sequence abbaba is read, the machine goes through states 0,1,2,2,1,2 and then ends in state 1 and it outputs the sequence aabbab. The output starts with the letter a and then repeats the input sequence up

to its penultimate letter. A little thought will show that this is what the machine does to any sequence.

There is a strong connection between formal languages and the 'machines' which we have discussed above. A formal language consists of certain words, defined over a given alphabet. It is specified by a 'grammar' – a set of rules which determine how words of the language may be built from other words of the language. A measure of complexity of the language is obtained by asking what is the simplest kind of machine that will accept precisely the words of that language. There is a classification of these languages, depending on the kind of grammatical rules. It turns out that the simplest languages are precisely those that are accepted by a finite state machine. The most general languages are those that are accepted by a Turing machine. Between these extremes, we have the types of language accepted by other types of automata. For more on this topic, see [Salomaa] for instance.

Exercises 2.4

1. Draw state diagrams for the machines shown:
   (a) The machine M with $S = \{0, 1, 2\}$, $A = \{\text{a,b}\}$ and μ given by

	a	b
0	0	1
1	1	2
2	2	0

   (b) The machine M with $S = \{0, 1, 2\}$, $A = \{\text{a,b,c}\}$ and μ given by

	a	b	c
0	1	0	2
1	0	0	1
2	2	0	2

2. Construct the tables of transition functions for the finite state machines whose state diagrams are shown in Fig. 2.29.
3. Let M be the finite automaton as specified. Determine the words accepted by M:
   (a) the machine in 2.4.1(a) above with $F = \{1\}$;
   (b) the machine in 2.4.2(b) above with $F = \{1\}$;
   (c) the machine in 2.4.2(b) above with $F = \{2\}$;
   (d) the machine in 2.4.2(c) above with $F = \{3\}$.
4. Let M be the automaton with $F = \{1, 3\}$ and state diagram as shown in Fig. 2.30.

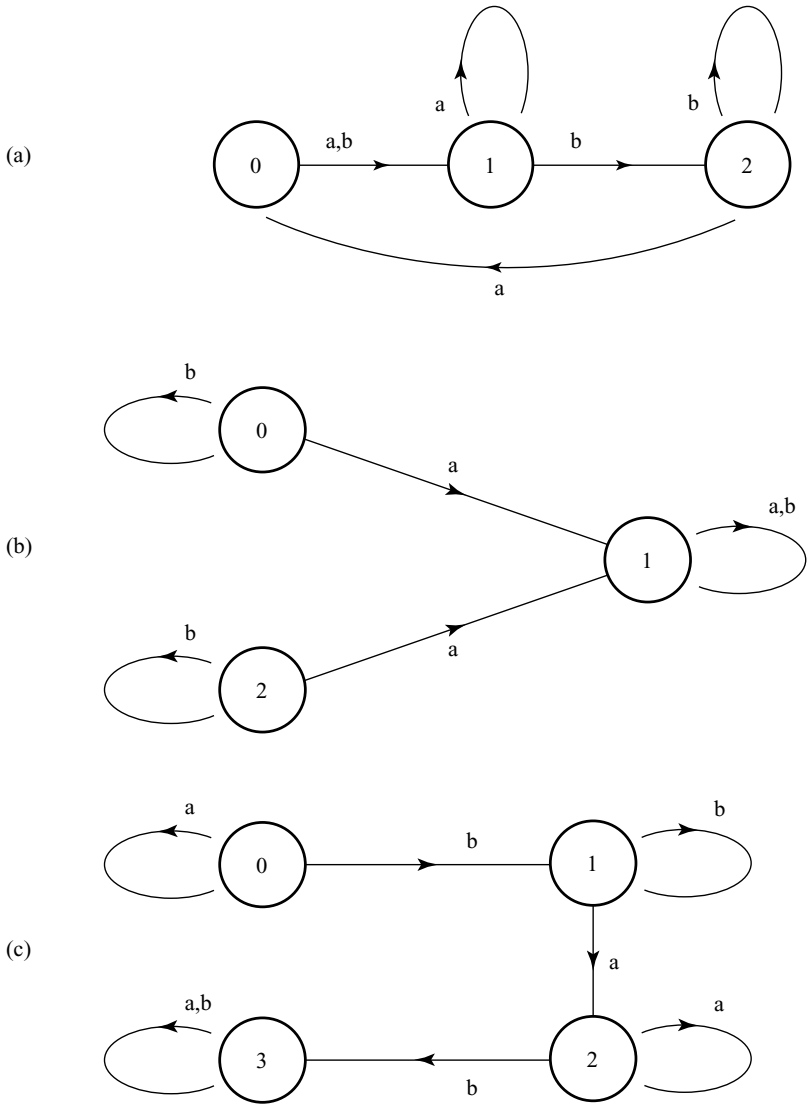

Fig. 2.29

Determine the words accepted by M and design a finite automaton with just two states which accepts the same words as those accepted by M.

5. Design finite state machines to meet the following specifications.
   (i) An automaton with alphabet $\{a,b\}$ that will only accept sequences composed entirely of the letter a.

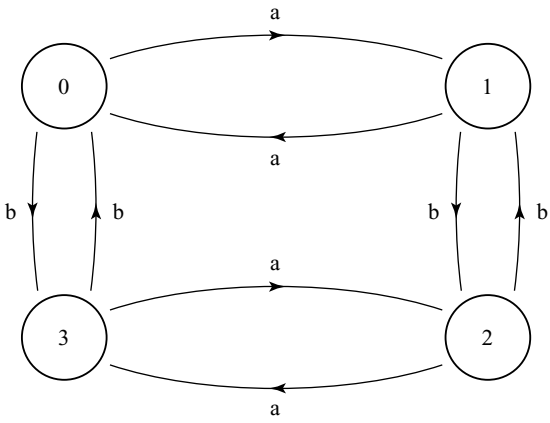

Fig. 2.30

(ii) A finite state machine with input alphabet {a,b} and output alphabet {α,β} that will output a sequence whose last term is β exactly when a word with an even number of b's is read.

(iii) A finite state machine which will read a word in {a, b, c} and output each occurrence of a and b but will replace every second occurrence of c by a.

6. A cautious millionaire has a home safe. Design an automaton, to be attached to the safe, which will read four-digit decimal numbers but will only accept the millionaire's personal number (which is 1357).

Summary of Chapter 2

We introduced sets and the operations of intersection, union and complement on sets and we gave a list of basic identities involving these operations. The product operation on sets was introduced. In the second section we gave an abstract, very general, definition of function, discussed composition of functions and special types of functions, in particular bijections, then used this to define what it means for sets to have the same cardinality. In Section 2.3 we considered relations and various properties which these may have. Partial orders and their Hasse diagrams were introduced, as were equivalence relations and the partitions which they induce. In the last section, we introduced finite state machines, both without and with output.

3 Logic and mathematical argument

Mathematics and mathematical reasoning is precise. This is in contrast to most discourse where deductions and chains of reasoning have gaps and imprecision. A chain of reasoning in mathematics should be compelling, in the sense that anyone who can follow the argument (which might, of course, be a very difficult task), should feel obliged to accept its correctness. Of course, in practice mathematical arguments can contain errors. But again, once an error has been pointed out, it should not be a matter of opinion whether or not there is an error. Over the centuries, indeed millennia, people have puzzled over what makes a chain of reasoning watertight. Especially over the past 150 years the logic of mathematics has been thoroughly investigated. A rather remarkable outcome of this mathematical investigation of logic is that it is, in principle, possible to formalise mathematical reasoning to the point where any purported mathematical argument could be checked by a computer. By this, we mean that the chain of reasoning can be checked (the correctness or applicability of the conclusions rather depends on the assumptions made at the outset).

It is interesting that, in practice, mathematicians do not always or even often produce proofs that are in a form which are easily computer checkable. Simple lemmas and computations are susceptible to such formal checking but, typically, proofs are quite (or very) complex and are written in a way which makes them comprehensible to human readers who have the necessary factual background and who have developed sufficient intuitive understanding that they do not need all the details to be written down. Of course this leaves open the possibility of error not being detected and it is by no means unusual for long and complex proofs in research papers to contain errors or 'gaps' which need to be filled (the proof of 'Fermat's Last Theorem' illustrates this). But, in finding and developing proofs, mathematicians are guided more by their intuitions and understanding than by precise logic and, as a result, typically these errors are corrected and gaps are filled easily enough once they have been noticed. Nevertheless sometimes

proofs contain not gaps, but vast chasms, which can take much time and effort to fill and which, very occasionally, simply cannot be filled: sometimes proofs do have to be withdrawn. It is also the case that errors can go unnoticed for some time. But, once a result has been proved, mathematicians will use it to draw further conclusions and this process usually points up any significant errors.

We spend the first section describing the most basic part of the logic that underlies mathematical reasoning: the logic of simple propositions. This is the logic of 'and', 'or', 'not' and 'implies': the logic of *combining* and *manipulating* statements. In order to *construct* mathematical statements, we need more, including the quantifiers 'for all' and 'there exists': these are discussed in the second section. In the final section, we review various of the proof strategies that we use in this book.

3.1 Propositional logic

Propositional logic enables us to handle the elementary logical connections between statements (or 'propositions'). By a **proposition** we mean an assertion which has a definite **truth value**: true (denoted by 't') or false (denoted by 'f'). For instance the following are propositions.

The sum of the first n positive integers is $n(n+1)/2$.
2 is an odd integer.
2 is a prime number.
Every even number greater than 2 is the sum of two prime numbers.

Each of these statements is either true or false (in the case of the last no-one knows which), but not both! On the other hand the following are not propositions.

Look here!
Is every even number greater than 2 the sum of two prime numbers?
n is a prime number.

Of course, in the last example, if the context were such that 'n' denoted a *particular* natural number, then the sentence would be a proposition but, in the absence of such a context, the statement is neither true nor false since we have left open what 'n' denotes. Notice also that a question is not a proposition, since a question cannot itself be true or false.

Even these few examples may well have raised in the reader's mind a number of questions of a sort which we do not deal with here: for a more extensive discussion of formal logic and its relation to natural language, the reader should consult, for instance, [Hodges].

We use letters such as p, q and r to stand for propositions. For example, p might be the proposition '2 is an odd number' and q might be the statement '2 is a prime number'.

Each proposition p has a **negation**, which is itself a proposition, denoted by $\neg p$ and read 'not p'. For instance, the negation of the statement '2 is an odd number' is '2 is not an odd number'. The proposition p is true exactly if $\neg p$ is false. The relationship between a proposition p and its negation $\neg p$ may be expressed in a 'truth table'.

p	$\neg p$
t	f
f	t

This table says (read along the rows) that if the proposition p is true then $\neg p$ is false, and if p is false then $\neg p$ is true.

Given two propositions p and q, we can form new propositions from them: their **disjunction** $p \vee q$, and their **conjuction** $p \wedge q$.

The proposition $p \vee q$, read as 'p or q', is true exactly if *at least one* of p, q is true, while $p \wedge q$, read as 'p and q', is true exactly if *both* p and q are true. Our statement means that as well as explaining the truth values taken by the propositions $p \vee q$ and $p \wedge q$, we have (by implication) made the standard English words 'or' and 'and' into functions on truth values, with two truth values as input and one as output. Note that, as is usual in mathematics, we use 'or' in the *inclusive* sense that $p \vee q$ is true if either or both p and q are true. For example the proposition '2 is a prime number or 2 is an even number' is true. (When Boole introduced his calculus of propositions (see [Boole]), he actually used 'or' in the *exclusive* sense – 'p or q but not both' – but the inclusive sense turns out to be the more convenient in mathematics.)

As an example of disjunction and conjunction of propositions, if p is the proposition '2 is an odd number' and q is the statement '2 is a prime number' then $p \vee q$ is the statement '2 is an odd number or 2 is a prime number' (which is true) and $p \wedge q$ is the statement '2 is an odd number and 2 is a prime number' (which is false).

In general we do not know the truth values of p and q but we do know that each is either true or false: so there are four possible ways to assign truth values to p, q. Namely:

both p and q are true (in which case $p \vee q$ is true);
p is true and q is false (in which case $p \vee q$ is true);
p is false and q is true (in which case $p \vee q$ is true);
both p and q are false (in which case $p \vee q$ is false).

These four possibilities give the four rows (under the heading line) of the first table below. The reader should satisfy himself or herself that the rows of the second truth table correctly express the relationship between the truth value of $p \wedge q$ and the truth values of its constituents p and q.

p	q	$p \vee q$
t	t	t
t	f	t
f	t	t
f	f	f

p	q	$p \wedge q$
t	t	t
t	f	f
f	t	f
f	f	f

Note that forming the proposition $p \wedge q$ may be thought of as building up a compound proposition from two simpler ones: we say that $p \wedge q$ is a (**propositional**) **term in** p and q. We define 'term in' to be a transitive relation. Thus, for example, if p itself is of the form $r \wedge s$ and if q is of the form $s \vee (\neg t)$, where r, s and t are propositions, then $p \wedge q$, that is $(r \wedge s) \wedge (s \vee (\neg t))$, as well as being a term in p and q, is also a term in r, s and t. The expressions '**Boolean combination of**' and '**Boolean expression in**' are also used instead of 'term in'.

In order to avoid ambiguity when reading propositions, we make the convention that $\neg$ has priority over $\wedge$ and $\vee$. So $\neg p \wedge q$ will mean $(\neg p) \wedge q$ rather than $\neg(p \wedge q)$.

Suppose now that the proposition p is a term in (that is, built up, using $\vee$, $\wedge$ and $\neg$, from) the propositions $q_1, \ldots, q_n$. Then we may draw up a **truth table** for p which shows how the truth value (true 't' or false 'f') of p depends on the truth values of $q_1, q_2, \ldots, q_n$. The first n columns of the truth table are labelled by $q_1, q_2, \ldots, q_n$, and the last column is labelled by p: we may insert additional columns to ease the actual computations (as in the table just below, also see other examples below). There will be, apart from the heading row, 2^n rows of the table, corresponding to the 2^n different ways of assigning truth values to $q_1, q_2, \ldots, q_n$.

Example Let p be the proposition $(q \wedge r) \vee (r \wedge \neg s)$. This is a term in the three propositions q, r and s, so the truth table for p will have $2^3 = 8$ rows.

q	r	s	$q \wedge r$	$r \wedge \neg s$	p
t	t	t	t	f	t
t	t	f	t	t	t
t	f	t	f	f	f
t	f	f	f	f	f
f	t	t	f	f	f
f	t	f	f	t	t
f	f	t	f	f	f
f	f	f	f	f	f

Given two propositions p and q, we write $p \to q$ for the proposition which is read as 'p implies q'. It is also read as 'if p then q'. This proposition is defined to have truth value 't' except when p is true and q is false, when it is 'f'. Again, in doing this, we have implicitly made the ordinary English sentence construction 'if ... then ...' into a function on truth values.

Our definition implies that if p is false, then $p \to q$ is true, no matter whether q is true or false. In particular, the truth of $p \to q$ does not imply that there is any real connection between p and q. Thus, for example, because '$2 = 3$' is false, the following proposition is true: '$2 = 3$ implies 6 is a prime number'. This, no doubt, seems peculiar, but it turns out to be the only sensible way to assign a truth value 't' or 'f' to such statements. Also, if we rephrase the proposition as 'if $2 = 3$ then 6 is a prime number', perhaps it seems a little less strange that this is a true proposition: all it says is that *if* $2 = 3$ then 6 is prime so, since $2 = 3$ is false, the proposition is true 'by default'.

The truth table for '$p \to q$' follows.

p	q	$p \to q$
t	t	t
t	f	f
f	t	t
f	f	t

Notice that this is the 'same' truth table as that for $\neg p \vee q$ (which is shown below; we have added a column for $\neg p$ as a computational aid). That is, given any assignment of truth values to p and q, the propositions $p \to q$ and $\neg p \vee q$ have the same truth value. Indeed, quite commonly '$p \to q$' is introduced simply as a shorthand for '$\neg p \vee q$'.

p	q	$\neg p$	$\neg p \vee q$
t	t	f	t
t	f	f	f
f	t	t	t
f	f	t	t

There are many ways in the English language of asserting (the truth of) an implication $p \to q$: here are some of them. You should take the time to think why these really are all equivalent to $p \to q$.

'p is a sufficient condition for q'
'if p then q'
'p only if q'
'q if p'

'q is a necessary condition for p' (to hold) (because if p holds then necessarily so does q)
(and various other phrases such as 'q follows from p', 'since p holds so does q' etc.)

Other terms used in mathematics which can be explained using propositional logic include the following.

An assertion $p \to q$ is an **implication**. Its **converse** is the implication $q \to p$. These are different: take p to be 'n is a multiple of 6' (think of n as fixed by a context) and q to be 'n is a multiple of 2'. Then certainly $p \to q$ ('if n is a multiple of 6 then n is a multiple of 2') is true but the converse, $q \to p$ ('if n is a multiple of 2 then n is a multiple of 6') is not.

On the other hand, the **contrapositive** of the implication $p \to q$ is the implication $\neg q \to \neg p$ and this *is* logically equivalent to $p \to q$ (compute the truth tables to check this). In our example $\neg q \to \neg p$ reads 'if n is not a multiple of 2 then n is not a multiple of 6', which *is* equivalent to $p \to q$ ('if n is a multiple of 6 then n is a multiple of 2').

As an exercise in the use of truth tables, the reader might like to check that for any propositions p and q, the propositions $\neg(\neg p \wedge \neg q)$ and $p \vee q$ have the same truth tables. In terminology that we will define later in this section, we can therefore say that 'p or q' is logically equivalent to 'not (not p and not q)' (think about this to see if you agree that these come to the same thing). It follows that, if we wished, we could define 'or' in terms of 'and' and 'not', replacing $p \vee q$ by $\neg(\neg p \wedge \neg q)$.

We write $p \leftrightarrow q$ for the statement '$(p \to q) \wedge (q \to p)$'. This is read as '$p$ is equivalent to q' or 'p if and only if q' (and 'if and only if' is often abbreviated in mathematics to 'iff'). Since we have a conjunction sign between the two implications, it follows that $p \leftrightarrow q$ is only true when *both* $p \to q$ and $q \to p$ are true. Another way to say this is that $p \leftrightarrow q$ is only true when $p \to q$ and its converse are true. As we have already noted it is quite possible for a statement to be true but its converse to be false. We also express $p \leftrightarrow q$ by saying that 'p is a necessary and sufficient condition for q'. The statement $p \leftrightarrow q$ has the following truth table

p	q	$p \leftrightarrow q$
t	t	t
t	f	f
f	t	f
f	f	t

The computation for this is shown below.

p	q	$p \to q$	$q \to p$	$p \leftrightarrow q$
t	t	t	t	t
t	f	f	t	f
f	t	t	f	f
f	f	t	t	t

So $p \leftrightarrow q$ is true exactly when p and q have the same truth values. Therefore both '$2 = 2 \leftrightarrow 7$ is a prime number' and '$2 = 3 \leftrightarrow 6$ is a prime number' are true.

A Boolean term is said to be a **tautology** if it is true no matter what truth assignments are given to its component propositions. Two Boolean terms, built from the same propositions, which have the same truth tables (that is, which take the same truth value for each assignment of truth values to their component propositions) are said to be **logically equivalent**. We can tell whether or not a Boolean term is a tautology by calculating its truth table: it is a tautology if and only if every row of its truth table ends with 't'. Also two terms are logically equivalent if and only if they have 'the same' truth tables (corresponding rows of their truth tables end with the same truth value).

Example To decide whether either of the Boolean expressions $(p \to q) \to (q \to p)$ or $\neg(p \wedge q) \leftrightarrow (\neg p \vee \neg q)$ is a tautology, we calculate their truth tables as follows.

p	q	$p \to q$	$q \to p$	$(p \to q) \to (q \to p)$
t	t	t	t	t
t	f	f	t	t
f	t	t	f	f
f	f	t	t	t

p	q	$\neg(p \wedge q)$	$\neg p \vee \neg q$	$\neg(p \wedge q) \leftrightarrow (\neg p \vee \neg q)$
t	t	f	f	t
t	f	t	t	t
f	t	t	t	t
f	f	t	t	t

Therefore the first Boolean term is not a tautology, but the second is a tautology.

Although somewhat tedious, this method enables us to determine whether a Boolean term is a tautology. The second term above is a tautology of the form $r \leftrightarrow s$ (with $r = \neg(p \wedge q)$ and $s = \neg p \vee \neg q$): tautologies of this form are referred to as **logical identities** since such a term being a tautology means that

the terms r and s always take the same truth values as each other and so are logically equivalent. For example $\neg(\neg p) \leftrightarrow p$ is a tautology and hence is a logical identity, reflecting the fact that $\neg(\neg p)$ and p are logically equivalent.

Further examples of logical identities are the following, in which T denotes any proposition which is always true (any tautology, such as $p \leftrightarrow p$) and F one which is always false (any **contradiction**, such as $\neg p \leftrightarrow p$):

Theorem 3.1.1 *The following are logical identities*:

$(p \wedge p) \leftrightarrow p$ *and*	
$(p \vee p) \leftrightarrow p$	*idempotence;*
$(p \wedge \neg p) \leftrightarrow F$	*consistency;*
$(p \vee \neg p) \leftrightarrow T$	*law of the excluded middle;*
$(p \wedge q) \leftrightarrow (q \wedge p)$ *and*	
$(p \vee q) \leftrightarrow (q \vee p)$	*commutativity;*
$p \wedge (q \wedge r) \leftrightarrow (p \wedge q) \wedge r$ *and*	
$p \vee (q \vee r) \leftrightarrow (p \vee q) \vee r$	*associativity;*
$\neg(p \wedge q) \leftrightarrow \neg p \vee \neg q$ *and*	
$\neg(p \vee q) \leftrightarrow \neg p \wedge \neg q$	*De Morgan laws;*
$p \wedge (q \vee r) \leftrightarrow (p \wedge q) \vee (p \wedge r)$ *and*	
$p \vee (q \wedge r) \leftrightarrow (p \vee q) \wedge (p \vee r)$	*distributivity;*
$\neg(\neg p) \leftrightarrow p$	*double negative;*
$(p \rightarrow q) \leftrightarrow (\neg q \rightarrow \neg p)$	*contrapositive law;*
$p \wedge T \leftrightarrow p$ *and*	
$p \vee T \leftrightarrow T$	*properties of T;*
$p \wedge F \leftrightarrow F$ *and*	
$p \vee F \leftrightarrow p$	*properties of F;*
$p \wedge (p \vee q) \leftrightarrow p$ *and*	
$p \vee (p \wedge q) \leftrightarrow p$	*absorption laws.*

Each of the above identities may be established using truth tables; you should try some as exercises. But to prove them all this way would be inefficient, for some of them may be deduced easily from others. For instance, by the law of the excluded middle, $p \vee T$ is equivalent to $p \vee (p \vee \neg p)$ which, by associativity, is equivalent to $(p \vee p) \vee \neg p$ which, by idempotence, is equivalent to $p \vee \neg p$ then, by another appeal to the law of the excluded middle, this is equivalent to T. Thus, the logical identity $p \vee T \leftrightarrow T$ follows from certain of the others.

Note that in Theorem 3.1.1 p, q and r may themselves be compound propositions: so a particular case of De Morgan's Law $\neg(p \wedge q) \leftrightarrow (\neg p \vee \neg q)$ is

$$\neg((p \wedge \neg r) \wedge q) \leftrightarrow (\neg(p \wedge \neg r) \vee \neg q).$$