Chapter 5

Sets

5.1 Naive Set Theory and Russell's Paradox

It can be a challenge to convince people that set theory is a profound and important branch of mathematics. That's because most students get a taste of sets in high school or even earlier, and at that level what's done can seem simpleminded and pointless. Don't be fooled. Set theory (like algebra, another subject that many people think is limited to the high school level) is full of fascinating and deep problems that have stumped many of the world's greatest mathematicians.

Naive Set Theory

The concept of sets in mathematics is quite recent, dating back only to about 1870. The essential idea of set theory is that any collection of objects of any sort that you can list or clearly describe may be considered to form a set. This principle is usually called the **axiom of comprehension.**

The other elementary principle of set theory is that a set has no other characteristics than being a collection of things. In other words, if two sets have exactly the same members, then they must be equal because there's nothing else that could distinguish them. This is called the **axiom of extensionality.**

Notation: You have probably already seen most of the basic notation for describing sets, but let's go over it anyway.

If an object x is in a set A, we say that x is an **element** or a **member** of A, which is written $x \in A$. (The symbol for set membership is a modified Greek epsilon.)

We use capital letters, usually from the beginning of the alphabet, to denote sets. Note that the expression on the right side of a membership statement must always denote a set, but the expression on the left side can denote any type of object.

The standard way of denoting a set is to show its members in braces: $\{ \ldots \}$. If we want to define a set that consists of a small, finite number of members, we can just list them inside the braces. This is called the **roster method** of denoting a set.

Example 1: $\{1, 3, 5, 7, 9\}$ denotes the set of all odd natural numbers less than 10.

A variant of the roster method can be used when a set has many elements. If we wanted to denote the set of all odd natural numbers up to 999, it would be absurd to list all those numbers. But if we write $\{1, 3, 5, 7, \ldots, 999\}$, the meaning is clear enough. This type of notation is acceptable. Furthermore, we could write $\{1, 3, 5, 7, \ldots\}$ as a perfectly clear notation for the set of *all* odd natural numbers. Thus even infinite sets can be shown with the roster method, *provided* that the meaning of the ellipsis is clear.

Note the reference to infinite sets. A rigorous definition of the words "finite" and "infinite" must wait until Section 7.5. In the meantime, a working definition may be helpful. Informally, we can say that a finite set is one that can be defined by the roster method, with no ellipsis. Alternatively, a finite set is one with no elements, or with n elements for some natural number n. "Infinite" simply means "not finite."

In the other method of denoting a set with braces, called **set-builder notation**, the elements are described instead of listed. The usual way to do this is to show a variable followed by a vertical line or a colon and then a proposition that shows what has to be true for something to be in the set.

Example 2: To denote the set of all odd natural numbers by the set-builder method, we can write

$$\{x \mid x \text{ is an odd natural number}\} \quad \text{or}$$

$$\{x \mid \exists n \in \mathbb{N}\,(x = 2n - 1)\}$$

We read the first notation "the set of all x such that x is an odd natural number." The second can be read "the set of all x such that $x = 2n - 1$, for some n in $\mathbb{N}$."

Set-builder notation can sometimes be shortened by restricting the variable to a set. So another way to write this same set would be

$$\{x \in \mathbb{N} \mid x \text{ is odd}\}$$

Of course, if we maintain our convention that the letter n is restricted to natural numbers, then the same set can be written simply as $\{n \mid n \text{ is odd}\}$.

Also, it's sometimes very convenient to use set-builder notation with the vertical line or colon preceded by an expression, rather than just a single letter. For example, the shortest and "neatest" notation for the set of all squares of natural numbers is

$$\{n^2 \mid n \in \mathbb{N}\}$$

Technically, this is an abbreviation for the more cumbersome notation

$$\{x \mid \exists n \in \mathbb{N}\,(x = n^2)\}$$

Every finite set can be described by either the roster method or set-builder notation. But for infinite sets, set-builder notation is more useful than the roster method. By the way, it's *only* in set-builder notation that a vertical line or colon is used as an abbreviation for "such that." Some mathematicians use the symbol ∍ as an abbreviation for these words in other contexts.

☞ Here is a *crucial* thing to learn about set-builder notation: suppose that a set has been defined by set-builder notation, say $A = \{x \mid P(x)\}$. Then for any x, it follows that $x \in A$ iff $P(x)$. This is simply the *definition* of this notation, and you don't need any theorems to justify this biconditional. Similarly, if $A = \{x \in B \mid P(x)\}$, then it's understood that $x \in A$ iff $x \in B$ and $P(x)$, for any x.

Example 3: Suppose that $A = \{x \in \mathbb{R} \mid \tan x > 5\}$, and we know that some number u is in A. Of course, we can then write $u \in \{x \in \mathbb{R} \mid \tan x > 5\}$, but this is rather cumbersome and can be stated much more simply by saying $\tan u > 5$. You should learn to make this translation *automatically.*

In the notation $\{x \mid P(x)\}$, the variable x should be considered bound, not free. That means it can be replaced by any other letter, so $\{x \mid P(x)\} = \{u \mid P(u)\}$, and so on. Such a replacement is often necessary to avoid having the same variable be free and bound at the same time.

Example 4: Suppose that for any real number y, we define A_y to be $\{x \mid x < y - 2\}$. So $A_3 = \{x \mid x < 1\}$, and so on. But what if we are considering a real number x? We can't say that $A_x = \{x \mid x < x - 2\}$. Rather, we have to change the dummy variable x; for example, we could write $A_x = \{z \mid z < x - 2\}$.

When set theory was invented, it was based completely on the axioms of comprehension and extensionality, and it's still based intuitively on the same two principles. However, as we soon see, this simple approach had some severe problems. For this reason, this early form of set theory is now called **naive set theory**. Here are the axioms of naive set theory, in symbols.

Extensionality: $A = B \leftrightarrow \forall x\,(x \in A \leftrightarrow x \in B)$

Comprehension: For any proposition $P(x)$, the set $\{x \mid P(x)\}$ exists.

Remarks: (1) In these axioms, and in this unit generally, the variables x, y, and z are not assumed to be restricted to real numbers; they can have any domain. In the extensionality axiom, x should be considered an unrestricted variable, ranging over *all* possible objects.

(2) Some authors call our extensionality axiom the *definition* of set equality. There's nothing wrong with this approach, and it does not differ from ours in any essential way.

(3) The comprehension axiom establishes the existence of any set of the form $\{x \in A \mid P(x)\}$, since that's the same as $\{x \mid x \in A \text{ and } P(x)\}$. This limited version of the comprehension axiom, in which the variable x must be restricted to a set, is called the **axiom of separation.**

Here are three definitions that illustrate the use of set-builder notation. The first two define important number systems that are intermediate between the systems $\mathbb{N}$ and $\mathbb{R}$. Note that we have *not* defined $\mathbb{N}$ and $\mathbb{R}$. Instead, we have taken them to be primitive and included axioms for them; this automatically means they must be undefined. Other approaches to the development of number systems are considered in Chapter 9.

Definition: A real number is called an **integer** iff it can be written as the difference of two natural numbers. The set of all integers is denoted $\mathbb{Z}$. In symbols,

$$\mathbb{Z} = \{n - m \mid m, n \in \mathbb{N}\}$$

An alternate way of describing $\mathbb{Z}$ is given in Exercise 13.

Definition: A real number is called **rational** if it can be written as a quotient of two integers. The set of all rational numbers is denoted $\mathbb{Q}$. In symbols,

$$\mathbb{Q} = \{a/b \mid a, b \in \mathbb{Z} \text{ and } b \neq 0\}$$

Note that the last two definitions illustrate the use of set-builder notation with an expression other than a variable before the vertical line, as discussed earlier in this section.

Definitions (intervals): In what follows, the letters a and b denote specific real numbers (normally with $a < b$), and x also stands for a real number:

$$\begin{aligned}(a, b) &= \{x \mid a < x < b\}\\ [a, b] &= \{x \mid a \leq x \leq b\}\\ (a, b] &= \{x \mid a < x \leq b\}\\ [a, b) &= \{x \mid a \leq x < b\}\end{aligned}$$

The first type of set is called an **open interval**, the second a **closed interval**, and the last two **half-open intervals** (less frequently, **half-closed intervals**). When it is important to make a distinction between intervals of the forms $(a, b]$ and $[a, b)$, the terms **open-closed interval** and **closed-open interval** are sometimes used.

The following notations are used to describe what are called **unbounded intervals** or **rays** (open or closed):

$$(a, \infty) = \{x \mid x > a\}$$

$$[a, \infty) = \{x \mid x \geq a\}$$
$$(-\infty, b) = \{x \mid x < b\}$$
$$(-\infty, b] = \{x \mid x \leq b\}$$

When using this notation, it is important to bear in mind that ∞ and $-\infty$ do *not* denote real numbers. The fact that these symbols always appear next to a parenthesis, never a square bracket, is intended to emphasize this fact. These symbols simply indicate that a certain set of real numbers has no end, either in the positive or the negative direction. The notation $(-\infty, \infty)$ is also used occasionally. But since this is just another way of denoting $\mathbb{R}$, it is not particularly useful.

Usually, intervals and interval notation refer to sets of real numbers. But they can be considered in any context where the inequality symbols have meaning.

The Paradoxes of Set Theory

What went wrong with naive set theory? You might guess that such a simpleminded theory, with these two very trivial-looking axioms, might have the drawback of not being powerful enough in the sense that it might not be possible to prove any interesting theorems from them. Surprisingly, the opposite is true. Naive set theory is *too* powerful; in fact it's **inconsistent**, meaning that it leads to contradictions. Several people discovered this about 1900, when the new subject was only a couple of decades old. The various forms of this contradiction in set theory are called paradoxes, but this word doesn't really convey the severity of the situation. A **paradox** usually refers to an *apparent* contradiction that can be straightened out with careful thought. The paradoxes of naive set theory have no solution, except to change the theory substantially.

Here is the simplest and most blatant paradox of set theory, usually credited to the great English philosopher and logician Bertrand Russell.

Theorem 5.1 (Russell's Paradox): Naive set theory is inconsistent; that is, it leads to a contradiction.

Proof: The proof is amazingly short; the core of it is a single use of the comprehension axiom, to form the set of all sets that are not members of themselves. In symbols, let

$$A = \{B \mid B \notin B\}$$

Then we simply ask whether A is a member of *itself!* If $A \in A$, then by the definition of A, $A \notin A$. On the other hand, if $A \notin A$, then since A is a set, we must have $A \in A$. So we have proved $A \in A \leftrightarrow A \notin A$, which is a contradiction. ■

A popularized version of Russell's paradox is known as the **Barber's paradox**. In a certain town there is a single barber, who is a man. The barber shaves all men in the town who do not shave themselves, and only those men. This sounds plausible, but the question is: Does the barber shave himself? There is no consistent answer to this question. If he does, he doesn't, and if he doesn't, he does.

The discovery of the paradoxes of naive set theory threw the foundations of mathematics for a loop. A mathematical theory that leads to contradictions is of no use. If a subject as simple looking as this could be inconsistent, what assurance is there that other branches of mathematics are consistent? And if there is no such assurance, who can guarantee the soundness of conclusions made in science and engineering on the basis of mathematics?

As a reaction to this development, many scholars in the early part of this century attempted to fix set theory, reformulating its axioms to achieve a consistent theory that would still be productive. A careful examination of the paradoxes of set theory leads to

Bertrand Russell (1872–1970) was not primarily a mathematician but continued an ancient tradition of philosophers making important contributions to the foundations of mathematics. He was born into a wealthy liberal family, orphaned by the age of four, and then raised by his grandmother and tutored privately.

There are two mathematical contributions for which Russell is remembered. One was his discovery of the inconsistency of naive set theory, as discussed in this section. The other was a monumental task that occupied him for over a decade: the three-volume *Principia Mathematica*, written with his former professor, Alfred North Whitehead, and completed in 1913. This work was the manifesto of the logicist school of thought, one response to the crisis in the foundations of mathematics that was precipitated by the paradoxes of set theory.

As a philosopher, Russell is considered one of the major figures in the modern analytical school. He was a prolific writer and wrote many books intended for the general public, notably the best-selling *A History of Western Philosophy* (1950), for which he won the Nobel Prize in Literature.

Outside academic circles, Russell is probably best known for his political and social activism. During World War I, his pacifism led to his dismissal from Trinity College and a six-month prison sentence. Many decades later, he vehemently opposed nuclear weapons, racial segregation, and the U. S. involvement in Vietnam. He advocated trial marriages and sexual freedom as early as the 1930s, a position that caused a court of law to nullify a faculty position that he had been offered by the City College of New York in 1940.

the conclusion that the full comprehension axiom is the culprit because it is just too general. Therefore, the revised set theory that is currently used keeps the original extensionality axiom, but replaces the comprehension axiom with about a half dozen more specific rules postulating the existence of various sets. The most widely used version of modern set theory was developed by Ernst Zermelo (1871–1953) and Abraham Fraenkel (1891–1965) and is called **Zermelo-Fraenkel (ZF) set theory**. We will not be discussing ZF set theory in this book. However, our set axioms (group IV in Appendix 1) are essentially the axioms of ZFC set theory (Zermelo-Fraenkel set theory plus the important axiom of choice, which is discussed in Section 7.7).

Did Zermelo and Fraenkel achieve their goal of creating a consistent version of set theory? Surprisingly, the answer to this is not known, and in a certain sense can *never* be known for sure! One of the reasons that set theory is important is that, with some esoteric possible exceptions, all of current mathematics can be carried out within the framework of ZFC set theory. Therefore, knowing the consistency of set theory would essentially be the same as knowing the consistency of mathematics. However, one of the most amazing and significant discoveries in the history of mathematics, known as **Gödel's incompleteness theorem**, states that the consistency of a "reasonable" mathematical theory cannot be proved without using postulates that go beyond that theory. Therefore, the consistency of set theory simply can't be proved using standard mathematical principles. The best that can be said is that nearly a century of experience with ZFC set theory has not produced any contradictions, and there is every reason to believe that it provides a consistent framework for mathematics.

Exercises 5.1

(1) Rewrite the following sets using the roster method.
- (a) $\{n \in \mathbb{N} \mid n^2 < 36\}$
- (b) $\{n^2 \mid n \in \mathbb{N} \text{ and } n < 6\}$
- (c) $\{x \in \mathrm{R} \mid \sin x = x\}$
- (d) $\{s \mid s \text{ is a New England state}\}$
- (e) $\{x \in \mathbb{Z} \mid |x| \text{ is prime and even}\}$

(2) Rewrite the following sets using set-builder notation.
- (a) $\{1, 4, 9, 16, 25, ..., 10{,}000\}$
- (b) $\{1, 4, 9, 16, 25, ...\}$
- (c) $\{-2, 4, -8, 16, ...\}$
- (d) {a, e, i, o, u, y}
- (e) $\{6, 17, 92\}$

(3) Which of the following sets are equal to each other?
- (a) $\{1, 2, 3\}$
- (b) $\{3, 2, 1\}$
- (c) $\{1, 2, 3, 1.0\}$
- (d) $\{x \in \mathbb{R} \mid 1 \le x \le 3\}$

(e) $\{n \in \mathbb{Z} \mid n^2 = 1 \text{ or } n^2 = 4 \text{ or } n^2 = 9\}$
(f) $\{n \in \mathbb{N} \mid n + 7 < 11\}$
(g) $\{\sqrt{1}, \sqrt{4}, \sqrt{9}\}$

(4) Let $A = \{2^x \mid x \in \mathbb{R} \text{ and } x^3 - x = 17\}$.
(a) Rewrite A in the form $\{y \mid ...\}$.
(b) Rewrite A in the form $\{x \mid ...\}$.
(c) Rewrite $\{2^x \mid x \in \mathbb{R} \text{ and } x^3 - x = 0\}$ using the roster method.

(5) Let $A = \{x + 3 \mid x \text{ is a real number that equals its tangent}\}$. Which of the following statements are true, if any? Explain your assertion.
(a) For any x in $\mathbb{R}$, $x \in A$ iff $x = \tan x$.
(b) For any x in $\mathbb{R}$, $x \in A$ iff $x + 3 = \tan(x + 3)$.
(c) For any x in $\mathbb{R}$, $x \in A$ iff $x - 3 = \tan(x - 3)$.

(6) Rewrite the following sets in interval notation:
(a) $\{x \in \mathbb{R} \mid x \leq 17\}$
(b) $\{x \in \mathbb{R} \mid x > -2 \text{ and } x \leq -1\}$
(c) $\{x \in \mathbb{R} \mid x \geq 5 \text{ or } -x < -2\}$
(d) $\{x \in \mathbb{R} \mid x^2 \leq 9\}$
(e) $\{x \in \mathbb{R} \mid |x + 3| < 4\}$

(7) Rewrite the following sets using set-builder notation *or* the roster method:
(a) $(-1, 7]$ (b) $(-\infty, 0)$ (c) $[5, 5]$

(8) True or false (with brief explanation):
(a) $3 \in [3, 7]$ (b) $3 \in (3, 7)$
(c) $3 \in [5, 2]$ (d) $3 \in [3, 3)$
(e) $-\infty \in (-\infty, 127)$ (f) $-\infty \in [-\infty, 127]$

(9) For each of the following statements, determine whether it's true or false in (i) $\mathbb{N}$, (ii) $\mathbb{Z}$, (iii) $\mathbb{Q}$, and (iv) $\mathbb{R}$. Give brief explanations. (All in all, this exercise has twelve true/false questions. For some guidelines for this type of problem, refer back to Example 6 of Section 3.3.)
(a) $\forall x,y\, \exists z\, (x - y = y^2 - z)$
(b) $\forall x,y\, [x \geq y \text{ or } \exists z\, (x < z < y)]$
(c) $\exists x,y\, (x^2 - y^2 = 2)$

(10) Using the extensionality axiom, prove that set equality satisfies axioms III-1, III-2, and III-3 (with the variables in those three axioms changed to set variables).

(11) Prove that the set form and the statement form of mathematical induction are equivalent to each other. You may use the axioms of naive set theory as well as all logic and equality axioms.

*(12) Here is an example of a "semantic paradox" known as Berry's paradox. Let A be the set of natural numbers that can be defined by English phrases less than sixty syllables long. (Examples of such phrases that define natural numbers are "the fifth smallest prime number," "the number of days in a week," and so on.) Since there are only a finite number of such phrases in English, A is finite. Therefore there are natural numbers that are not in A. Let n be the smallest natural number not in A.

Now consider the phrase "the smallest number that cannot be defined by an English phrase of fewer than sixty syllables." This phrase has fewer than sixty syllables and defines the number n. Therefore, $n \in A$, which is a contradiction.

Try to explain this paradox. That is, try to explain the flaw in the argument; there should be one since it should not be possible to prove a contradiction from scratch. The explanation is based on subtle philosophical considerations, rather than a technical point or trick. By the way, the problem does *not* lie with the last sentence of the first paragraph. As we will soon see (Theorem 5.6), if there is a natural number with a certain property, then there is a least one.

*(13) (a) Define the set $\mathbb{Z}'$ to consist of all natural numbers, negatives of natural numbers, and zero. In symbols,

$$\mathbb{Z}' = \{x \in \mathbb{R} \mid x \in \mathbb{N} \text{ or } (-x) \in \mathbb{N} \text{ or } x = 0\}$$

Prove that $\mathbb{Z}' = \mathbb{Z}$. Therefore, this is a correct alternate way of defining integers.

(b) Using part (a), deduce that $\mathbb{N}$ is the set of all positive integers.

(14) Let P(n) be a statement with a free *integer* variable n. Suppose that we are able to prove P(0) and $\forall n$ [P(n) implies P($n + 1$) and P($n - 1$)]. What would be the logical thing to conclude from this? Prove your claim. You may use the result of the previous exercise.

Critique the proofs in the remaining exercises. (If necessary, review the instructions for this type of problem in Exercises 4.2.)

(15) **Theorem:** If a, b, c and d are real numbers with $a < b$ and $c < d$, then $[a, b] = [c, d]$ iff $a = c$ and $b = d$.

Proof: Assume $a, b, c, d \in \mathbb{R}$, $a < b$, and $c < d$.

For the forward direction, assume $[a, b] = [c, d]$. Since $c \leq a \leq b$, the definition of intervals tells us that $a \in [a, b]$. So, by extensionality, $a \in [c, d]$. By definition of intervals, this implies that $c \leq a$. Similarly, we know that $c \in [c, d]$, hence $c \in [a, b]$, and therefore $a \leq c$. From the inequalities $c \leq a$ and $a \leq c$, it follows (as in Exercise 26, Section 4.4) that $a = c$.

A nearly identical argument shows that $b = d$.

For the reverse direction, assume $a = c$ and $b = d$. By Theorem 4.7 applied to $a = c$, we get $[a, b] = [c, b]$. The same theorem applied to $b = d$ yields $[c, b] = [c, d]$. Therefore, by transitivity of equality (axiom III-3), $[a, b] = [c, d]$.

(16) **Theorem:** Let A and B be any sets, P(x) any proposition, $C = \{x \in A \mid P(x)\}$ and $D = \{x \in B \mid P(x)\}$. Then, $A = B$ iff $C = D$.

Proof: Assume $A = B$. Then Theorem 4.7 immediately implies that $C = D$. By the same reasoning, if $C = D$, then $A = B$.

5.2 Basic Set Operations

Despite the fact that naive set theory is inconsistent, it turns out that as long as you avoid defining sets that are too big, naive set theory works quite well and does not seem to lead to any contradictions. And the more correct modern axiomatic set theory is much more complicated. So most mathematicians use naive set theory unless they are trying to be extremely careful and/or formal; they learn through experience how to use it safely. We take this approach, and you should feel free to do the same for most of your dealings with set theory.

It's impossible to give an ironclad set of guidelines for what to avoid when using naive set theory, but here is the most important one.

Rule of Thumb (When Using Naive Set Theory): Do not try to define the set of *all* sets or a set that involves all sets in its definition (such as the set defined in the proof of Russell's paradox). Any such definition will probably lead to a contradiction.

At times it is convenient to talk about the "class" of all sets. As long as classes of this sort are not allowed to be members of sets, paradoxes do not seem to arise. In contrast, there is no difficulty in defining the set of all real numbers, the set of all *sets* of real numbers, the set of all people who have ever lived, the set of all particles in the universe, and many other sets that might seem very big. If this restriction seems strange and esoteric, don't be concerned; it's not something you have to worry about very often. Instead, let's spend the rest of this section on some simple and familiar operations involving sets, which are shown in standard **Venn diagram** form in Figure 5.1.

Definitions: (a) The **union** of any two sets A and B, denoted $A \cup B$, is the set

$$\{x \mid x \in A \text{ or } x \in B\}$$

(b) The **intersection** of any two sets A and B, denoted $A \cap B$, is the set

$$\{x \mid x \in A \text{ and } x \in B\}$$

(c) The **relative complement of A in B** (or **the complement of A relative to B**), denoted $B - A$, is the set

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

(d) The **empty set** or **null set**, denoted $\varnothing$, is the set with no members.

(e) Two sets are called **disjoint** if their intersection is empty.

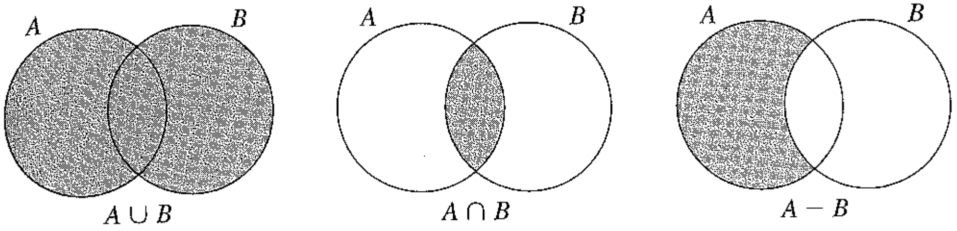


Figure 5.1 Venn diagrams illustrating basic set operations

Remarks: (1) The idea of a set with no members may seem odd at first, but it's a harmless and very useful concept. Also, unless we want to restrict the axioms further, there must be such a set. For example, $\{n \in \mathbb{N} \mid n < n\}$ is empty. Note that extensionality guarantees that any two empty sets are equal. In other words, there's only one of them, so mathematicians normally refer to *the* empty set.

(2) Notice that we've defined *relative* complement as opposed to just complement. That's because when people talk about the complement of a set, they always mean a relative complement—relative to some set that's understood. For example, suppose $A = [1, 3] = \{x \mid 1 \leq x \leq 3\}$, which is an interval on the real number line. If you then see a reference to the complement of A, it almost certainly means the complement of A relative to $\mathbb{R}$, which is $\{x \mid x < 1 \text{ or } x > 3\}$.

But now suppose there were a reference to the complement of the set $\{2, 17, 984\}$. Would that mean the complement of this set relative to $\mathbb{N}$? To $\mathbb{Z}$? To $\mathbb{R}$? Or perhaps relative to some other set? Unless the context made things clear, the reference would be quite ambiguous. Worse yet, suppose we were considering a set like $\{6, -2.7,$ Shakespeare, Canada$\}$. Would it make sense to talk about the complement of this set without saying what it is relative to? Not at all.

You might wonder why we can't define the absolute complement of a set, meaning simply the set of all objects (with no restriction) that are not in the set. The reason is that doing so quickly leads to a contradiction similar to Russell's paradox. In particular, the absolute complement of the null set would have to contain all sets. With these considerations in mind, we establish the following convention.

Convention: Suppose that, during a certain discussion, it is understood that all sets being considered are contained in some particular set U. Then it's permissible to write A' or $\overline{A}$ (called the **complement of A** or **A complement**) as an abbreviation for $U - A$ (see Figure 5.2). The set U may be called the **universal set** for the purposes of the discussion. But remember that the idea of a universal set is just a temporary convenience. There is no such thing as *the* universal set.

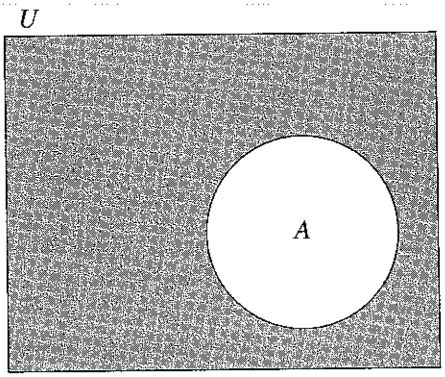


Figure 5.2 Venn diagram illustrating the complement of a set A

Here is a sample of the many elementary results that hold for these basic operations on sets. This subject is sometimes called the **algebra of sets.**

Theorem 5.2: For any sets A, B, C, and D,

(a) $A \cup B = B \cup A$
(b) $A \cap B = B \cap A$
(c) $A \cup (B \cup C) = (A \cup B) \cup C$
(d) $A \cap (B \cap C) = (A \cap B) \cap C$
(e) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
(f) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
(g) $A - (B \cap C) = (A - B) \cup (A - C)$
(h) $A - (B \cup C) = (A - B) \cap (A - C)$
(i) $A \cap (B - C) = (A \cap B) - (A \cap C)$

Proof: We just prove a couple of parts here. The rest are very similar and are left for the exercises.

(a) The usual way to prove two sets are equal is via the extensionality axiom. So we want to show

$$\forall x\,(x \in A \cup B \;\leftrightarrow\; x \in B \cup A)$$

To start, assume $x \in A \cup B$. By the definition of union, that means $x \in A$ or $x \in B$. But this is equivalent to $x \in B$ or $x \in A$. So, by the definition of union, $x \in B \cup A$. We have thus proved $x \in A \cup B \rightarrow x \in B \cup A$. The proof of the converse is similar. Since x is arbitrary, extensionality yields that $A \cup B = B \cup A$.

(e) As with part (a), the main step here is to prove a certain biconditional. But instead of doing that as two separate implications, let's do it as a single proof, a shortcut that was mentioned in the discussion of the biconditional rule:

$$\begin{aligned}
x \in A \cup (B \cap C) &\leftrightarrow x \in A \text{ or } x \in B \cap C && \text{Definition of } \cup \\
&\leftrightarrow x \in A \text{ or } (x \in B \text{ and } x \in C) && \text{Definition of } \cap \\
&\leftrightarrow (x \in A \text{ or } x \in B) \text{ and } (x \in A \text{ or } x \in C) && \text{Tautology 30} \\
&\leftrightarrow x \in A \cup B \text{ and } x \in A \cup C && \text{Definition of } \cup \\
&\leftrightarrow x \in (A \cup B) \cap (A \cup C) && \text{Definition of } \cap
\end{aligned}$$

So $x \in A \cup (B \cap C) \leftrightarrow x \in (A \cup B) \cap (A \cup C)$. Again applying UG and extensionality, the two sets must be equal. ■

The result of Theorem 5.2(e) is also illustrated in Figure 5.3. Although a picture can never constitute a rigorous proof, a careful Venn diagram can be a pretty reliable way to determine whether a statement of elementary set algebra is necessarily true.

The first six parts of Theorem 5.2 strongly resemble various field axioms (group V in our axiom system). Parts (a) and (b) say that $\cup$ and $\cap$ are commutative, (c) and (d) say that these set operations are associative, and (e) and (f) are distributive laws. Although there is some connection between set algebra and real number algebra, set algebra is more closely connected with propositional logic. The following theorem illustrates this connection further.

Theorem 5.3: Assume that A and B are both contained in some particular set U, and let A' and B' be abbreviations for $U - A$ and $U - B$. Then

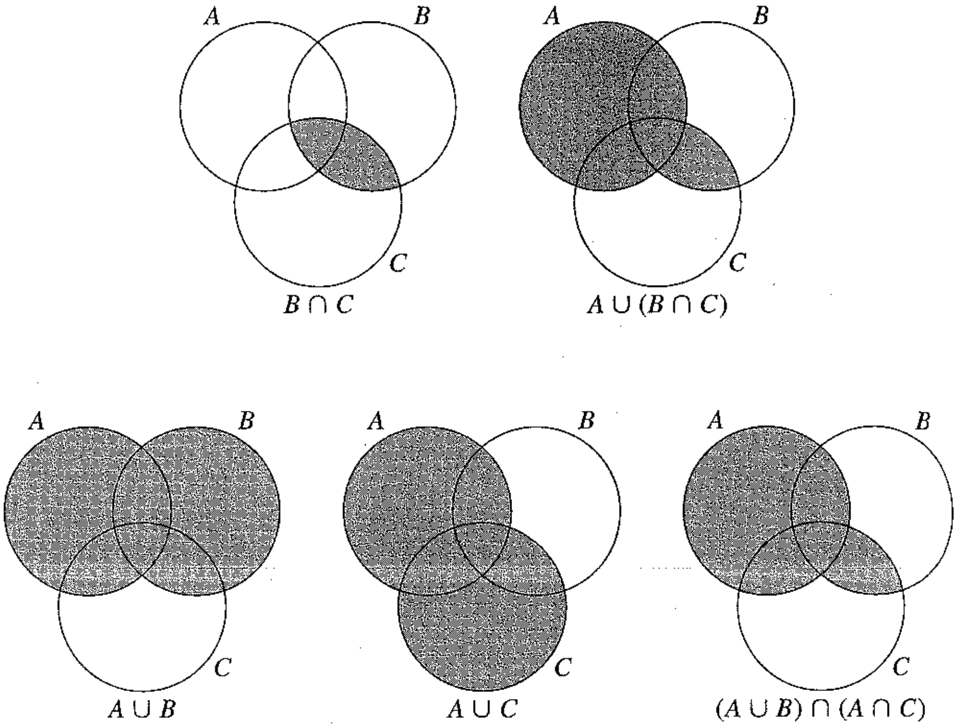
(a) $A \cap A' = \varnothing$
(b) $A \cup A' = U$
(c) $(A \cup B)' = A' \cap B'$ (De Morgan's law for sets)
(d) $(A \cap B)' = A' \cup B'$ (De Morgan's law for sets)
(e) $A - B = A \cap B'$
(f) $A' - B' = B - A$
(g) $(A')' = A$

Proof: We just prove part (c), leaving the rest for the exercises. Our proof is very similar to the proof of Theorem 5.2(e), except that we now let the variable x have the set U as its domain.

$$\begin{aligned}
x \in (A \cup B)' &\leftrightarrow x \notin A \cup B && \text{Definition of } ' \\
&\leftrightarrow \sim (x \in A \text{ or } x \in B) && \text{Definition of } \cup \\
&\leftrightarrow x \notin A \quad \text{and} \quad x \notin B && \text{De Morgan's law} \\
&\leftrightarrow x \in A' \quad \text{and} \quad x \in B' && \text{Definition of } ' \\
&\leftrightarrow x \in A' \cap B' && \text{Definition of } \cap
\end{aligned}$$

Applying UG and extensionality to this yields $(A \cup B)' = A' \cap B'$. ■

Remarks: (1) By now, you should be getting the idea of how to prove two sets are equal. In general, if you want to prove $A = B$ using the extensionality axiom, you must prove $x \in A \leftrightarrow x \in B$ (where x is arbitrary). In simple cases, this biconditional can often be proved without splitting it up, as we've done in the previous two cases (and



$B \cap C$ $A \cup (B \cap C)$

$A \cup B$ $A \cup C$ $(A \cup B) \cap (A \cap C)$

Figure 5.3 Venn diagrams illustrating Theorem 5.2(e)

could have done in Theorem 5.2(a)). In more complex cases, it's often necessary to split the biconditional up into two implications and prove each one separately.

(2) The proofs in this section don't have much content; they could be described as mostly definition unraveling, as discussed in Section 4.6. That is, if you look at the proofs so far in this section, each one simply rewrites the symbols $\cup$, $\cap$, $-$, and $'$ in terms of their definitions, and then uses one tautology to establish the desired biconditional. Of course, the extensionality axiom is also used, and in an indirect sense so is the comprehension axiom. But it's typical of basic set algebra that the proofs involve mostly definitions and propositional logic. The connection between set algebra and propositional logic can be made a bit more precise, as follows: every theorem of set algebra stating that two sets must be equal is, after unraveling the definitions of the set

symbols, equivalent to some tautology. This connection is the basis of a branch of mathematics called **boolean algebra** (see Exercise 16).

(3) Also, you may have noticed that many parts of Theorems 5.2 and 5.3 occur in pairs of similar-looking statements. This **duality**, as it's called, is also a basic aspect of boolean algebra (see Exercise 17).

Subsets, Proper and Otherwise

Definitions: We say that A is a **subset** of B (in symbols, $A \subseteq B$) iff every element of A is in B. Also, A is a **proper subset** of B ($A \subset B$) iff $A \subseteq B$ and $A \neq B$.

It is permissible to write $B \supseteq A$ and $B \supset A$ to mean $A \subseteq B$ and $A \subset B$, respectively. Normally, these reversed symbols and the associated word "superset" are used only when there is some specific reason to do so. For instance, it is simpler to say "every superset of A" than "every set of which A is a subset."

Here are a few simple results involving these concepts.

Theorem 5.4: (a) $A \subseteq A$

(b) $A \not\subset A$

(c) $\varnothing \subseteq A$

(d) $\varnothing \subset A \leftrightarrow A \neq \varnothing$

(e) $A \subseteq B$ and $B \subseteq \mathrm{C} \rightarrow A \subseteq \mathrm{C}$

(f) $A = B \leftrightarrow A \subseteq B$ and $B \subseteq A$

(g) $A \neq B \leftrightarrow (A - B) \cup (B - A) \neq \varnothing$

(h) $A \subset B \leftrightarrow A \subseteq B$ and $B \not\subseteq A$

Proof: Again, we prove only two parts and leave the rest for the exercises.

(c) By definition, $\varnothing \subseteq A$ means $\forall x\ (x \in \varnothing \rightarrow x \in A)$. Let x be arbitrary. By definition of $\varnothing$, $x \in \varnothing$ is automatically false, and therefore the conditional $x \in \varnothing \rightarrow x \in A$ is automatically true.

(e) This has already been proved in Proof Preview 2 (Section 2.3), so we do not repeat the argument. ■

The word "subset" might give the impression that a set would not be a subset of itself. But as you can see, the definition is written in such a way that *every* set is a subset of itself.On the other hand, *no* set is a *proper* subset of itself. For emphasis, mathematicians may say that every set is an *improper* subset of itself.

Be careful to keep the grammar of these symbols straight. The symbols $\in$ and $\subseteq$ are predicate symbols; that means that $x \in A$ and $A \subseteq B$ are complete statements that can stand alone or be combined using connectives and quantifiers. On the other hand, $\cup$, $\cap$, $-$, and $'$ are operator symbols. So expressions like $A \cup B$, $A \cap (B - C)$, and so on, are terms denoting sets, *not* statements. For example, it is *grammatically impossible* to have a line in a proof that just says $A \cup B$. Exercise 1 tests your understanding of the grammar of set theory. Also, proving some of the parts of Theorem 5.5 should provide good practice with these symbols.

Phrases like "contains" and "is contained in" are used ambiguously by mathematicians. We have already used the latter phrase to mean $\subseteq$, and this is probably how it is used most frequently. But it can also be used to mean $\in$. Fortunately, the ambiguity is not too serious because the context almost always makes it clear which meaning is intended.

Theorem 5.5: (a) $A \subseteq A \cup B$
(b) $A = A \cup B \leftrightarrow B \subseteq A$
(c) $A \cap B \subseteq A$
(d) $A = A \cap B \leftrightarrow A \subseteq B$
(e) $A \cap B = A \cup B \leftrightarrow A = B$
(f) $A \subseteq B \cap C \leftrightarrow A \subseteq B$ and $A \subseteq C$
(g) $A \cup B \subseteq C \leftrightarrow A \subseteq C$ and $B \subseteq C$

Proof: (b) For the forward direction, assume $A = A \cup B$. We want $B \subseteq A$, so assume $x \in B$. (Note we're doing a conditional proof within a conditional proof.) Therefore, $x \in A$ or $x \in B$, by propositional logic. So $x \in A \cup B$, by definition of $\cup$. Thus $x \in A$, by the first assumption. Since x was arbitrary, we've shown that $B \subseteq A$, as desired.

For the reverse direction, assume $B \subseteq A$. We want $A = A \cup B$, which by Theorem 5.4(f) says $A \subseteq A \cup B$ and $A \cup B \subseteq A$. We have just shown $A \subseteq A \cup B$ in part (a). For the other part, assume $x \in A \cup B$. So $x \in A$ or $x \in B$. But since $B \subseteq A$, if $x \in B$, then $x \in A$. Therefore we can conclude that $x \in A$. Since x was arbitrary, we've shown $A \cup B \subseteq A$, and so we obtain $A = A \cup B$.

The proofs of the other parts are left for the exercises. ■

Now that we know more about sets, we can prove an important consequence of mathematical induction.

Theorem 5.6 : Every nonempty set of natural numbers has a *least* element.

Proof: We prove this by induction, but we must carefully phrase what we are proving. Let P(n) be the statement "Every set that contains a natural number less than or equal to n contains a least natural number." Note that P has a quantified set variable, but the natural number variable n is free in it. We wish to prove $\forall n$ P(n), by induction.

Let $n = 1$. By Theorem 4.16(a), 1 is the least natural number. So if a set contains a natural number less than or equal to 1, it contains 1, which is the least natural number in the set.

Now assume P(n), and let A be any set that contains a natural number less than or equal to $n + 1$. We must show that A contains a least natural number. In the case that $n + 1$ is the least number in A, we are of course done. If not, let $B = \{m \in A \mid m < n + 1\}$. The set B contains natural numbers that are less than $n + 1$, and therefore, by Theorem 4.16(c), equal to or less than n. Thus we can apply the induction hypothesis to B; so B contains a least natural number, which is also the least natural number in A.

This completes the proof of $\forall n\ P(n)$. The theorem easily follows: if A is a nonempty subset of ℕ, then there is some natural number in A. Let n be such a number (by ES). Using $P(n)$, we conclude that A contains a least natural number. ■

Theorem 5.6 can be used in conjunction with indirect proof to prove statements by what is sometimes called the "no least counterexample" method (see Exercise 8).

Theorem 5.6 states an important property of ℕ that fails for many other number systems and ordered structures. In particular, it fails with the word "natural" replaced by "real." (For instance, ℤ is a nonempty subset of ℝ with no least member; so is ℝ itself.) An ordering relation with the property described in this theorem is called a **well-ordering.** In other words, the theorem asserts that ℕ is **well ordered** and is sometimes called the **well-ordering property of** ℕ (see Exercises 16 through 18 of Section 6.3). Assuming a few basic properties of ℕ, the well-ordering property of ℕ is equivalent to mathematical induction, and is sometimes used as an axiom instead of induction.

The Sum Rule for Counting

Set theory can be extremely abstract, but it also deals with many problems that are very concrete. In particular, the study of finite sets is closely related to *counting problems,* problems whose goal is to determine the number of members in some finite set. The following counting formula is quite simple to understand (though not to prove rigorously), and yet is surprisingly useful.

Theorem 5.7 (Sum rule for counting): Let A and B be finite sets, with m and n members, respectively.

(a) If A and B are disjoint, then $A \cup B$ has $m + n$ members.

(b) More generally, if $A \cap B$ has k members, then $A \cup B$ has $m + n - k$ members.

Proof: Rather than providing an extremely informal proof of this result now, we prove it rigorously in Section 7.6, where we examine counting problems in depth. ■

Clearly, part (b) of this theorem makes part (a) superfluous. We state part (a) separately because it is such an important special case. Part (a) easily generalizes to three or more sets. It is more complicated to generalize part (b), as the next example illustrates.

Example 1: Mudville High has three varsity teams. The table tennis team has 13 members, the Ultimate Frisbee team has 21 members, and the boomerang team has 16 members. How many varsity athletes are there? The obvious answer is 50, but this is wrong if there is duplication on the teams. Suppose there are 11 students who are on more than one team. Can we conclude that there are 39 athletes? This would also be wrong if there are people who play on all three teams. Exercise 12 asks you to investigate this further and find the correct formula.

We conclude this section with an overview of the most direct ways to prove basic relationships between sets.

Table 5.1 Summary of How to Prove Statements about Sets

(1) To prove a statement of the form $A \subseteq B$, assume that $x \in A$ (where x is arbitrary) and show that $x \in B$.

(2) To prove a statement of the form $A = B$, prove both $A \subseteq B$ and $B \subseteq A$.

(3) To prove a statement of the form $A \not\subseteq B$, find a member of A that is not in B.

(4) To prove a statement of the form $A \subset B$, prove $A \subseteq B$ and $B \not\subseteq A$.

(5) To prove a statement of the form $A \neq B$, prove $A \not\subseteq B$ or $B \not\subseteq A$. That is, find an element of either set that is not in the other one.

Exercises 5.2

(1) Classify each of the following expressions as either (i) a grammatically correct *statement*, (ii) a grammatically correct *expression* denoting a set, or (iii) grammatically incorrect and therefore meaningless. Assume that A, B, and C are set variables and P and Q are propositional variables.

(a) $A \cup B \subseteq C$
(b) $A \cup (B \subseteq C)$
(c) $A \leftrightarrow B$
(d) $(A \cup B) = \mathrm{P}$
(e) $\mathrm{P} \cup \mathrm{Q} \rightarrow \mathrm{P}$
(f) $(x \in A) \cup (x \in B)$
(g) $\mathrm{P} \wedge A \cup B \subseteq C$
(h) $\{x \in A \mid B \cup \{x\} \subseteq C\} \cap C$
(i) $A \cup B \in A \cap C$

(2) Prove any two parts of Theorem 5.2 that were not proved in the text.

(3) Prove any two parts of Theorem 5.3 that were not proved in the text.

(4) Prove any two parts of Theorem 5.4 that were not proved in the text.

(5) Prove any two parts of Theorem 5.5 that were not proved in the text.

(6) For each of the following statements, either prove that it is true for all sets or find a counterexample to show that it is not. Also, in parts (a) and (b), if the statement is not always true, at least try to prove that one side must be a subset of the other side.

(a) $A \cup (B - C) = (A \cup B) - (A \cup C)$ (Note this resembles Theorem 5.2(i)).
(b) $(A - B) \cup (A - C) = A - (B \cup C)$
(c) $A \subseteq B$ iff $(A - B) = \varnothing$
(d) $A \subseteq (B \cup C)$ iff $A \subseteq B$ or $A \subseteq C$

(7) True or false (with brief explanations):

(a) $\mathbb{R} \cup [3, 7) \subseteq \mathbb{R}$

(b) $[1, 4] \cup (3, 6] = [1, 9) \cap [2, 6]$

(c) $[1, 6] - [2, 5] = [1, 2] \cup [5, 6]$

(d) $[3, 6] \cup [6, 8] = [3, 8]$

(e) $(3, 6) \cup (6, 8) = (3, 8)$

(f) $\mathbb{R} - (\mathbb{Q} - \mathbb{N}) = (\mathbb{R} \cup \mathbb{N}) - \mathbb{Q}$

(8) Suppose we want to prove a statement of the form $\forall n$ P(n). If we want to use indirect proof, what do we assume? From that assumption, what can we assert to exist? Then, using Theorem 5.6, how can we strengthen this assertion? Often, this last assertion easily leads to a contradiction.

(9) Reprove Theorem 4.9 by the method outlined in Exercise 8. You may use Theorem 4.16, which does not require Theorem 4.9 in its proof.

(10) (a) Prove that $\mathbb{N} \subset \mathbb{Z}$.

*(b) Prove that $\mathbb{Z} \subset \mathbb{Q}$. ***Hint:*** Show that 1/2 is not in $\mathbb{Z}$.

(11) At a meeting of the Swampscott Phrenology Club, 37 members are present. Of these, 13 are wearing glasses, 8 are wearing sandals, and 20 are wearing *neither* glasses nor sandals. According to the sum rule, how many must be wearing *both* glasses and sandals?

(12) (a) Investigate the situation described in Example 1. If you wish, stick to the numbers in that example, but try various possibilities for the number of athletes on each pair of teams and the number on all three teams. On the basis of your results, conjecture a formula for the total number of athletes in terms of the number on the individual teams, the number who are on more than one team, and the number who are on all three teams. You might find Venn diagrams helpful for your investigation.

(b) Assuming your formula from part (a) is correct, derive a formula for the number of elements in the union of any three finite sets A, B, and C, in terms of the number of members in $A, B, C, A \cap B, A \cap C, B \cap C$, and $A \cap B \cap C$.

*(c) Carefully compare Theorem 5.7(b) and the formula you found in part (b) of this problem. Then try to describe (in words) how to calculate the number of elements in the union of four or more finite sets, in terms of the number of elements in the individual sets and the number of elements in the intersections of combinations of those sets.

Critique the proofs in Exercises 13 and 14. (If necessary, review the instructions for this type of problem in Exercises 4.2.)

(13) **Theorem:** Let a, b, and c be real numbers with $a < b < c$. Then $(a, b) \cup (b, c) = (a, c)$.

Proof: For the forward direction, assume x is any member of $(a, b) \cup (b, c)$. Then x is either in (a, b) or in (b, c), and we may proceed by cases. Case 1: Assume $x \in (a, b)$. That means $a < x < b$. But from $x < b$ and $b < c$ we obtain $x < c$, by transitivity. Therefore, $a < x < c$; this says that $x \in (a, c)$. Case 2: Assume $x \in (b, c)$. That means $b < x < c$. But from $b < x$ and $a < b$, we obtain $a < x$. Thus, $a < x < c$.

For the reverse direction, assume x is any member of (a, c). We again proceed by cases. Case 1: Assume $x < b$. Since $x \in (a, c)$, we also have $a < x$. Thus $a < x < b$; so $x \in (a, b)$. This implies that $x \in (a, b) \cup (b, c)$. Case 2: Assume $x > b$. We also have $x < c$. Thus $b < x < c$; so $x \in (b, c)$. This implies $x \in (a, b) \cup (b, c)$.

(14) **Theorem:** If $A \subset B$ and $B \subseteq C$, then $A \subset C$.

Proof: Assume $A \subset B$ and $B \subseteq C$. By the definition of $\subset$, $A \subseteq B$ and $A \neq B$. From $A \subseteq B$ and $B \subseteq C$, we have $A \subseteq C$, by Theorem 5.4(e). We must also show $A \neq C$. So assume, on the contrary, that $A = C$. Then $B \subseteq C$ becomes $B \subseteq A$. So we have $A \subseteq B$ and $B \subseteq A$, which yield $A = B$, by Theorem 5.4(f). This contradicts the fact that $A \neq B$.

(15) For any sets A and B, define $A \Delta B$, the **symmetric difference** of A and B, to be the set $(A - B) \cup (B - A)$. Prove.

(a) Commutativity of Δ: $A \Delta B = B \Delta A$

(b) Associativity of Δ: $A \Delta (B \Delta C) = (A \Delta B) \Delta C$

(c) The empty set is an identity for Δ: $A \Delta \varnothing = A$

(d) Each set is its own inverse for Δ: $A \Delta A = \varnothing$

(e) $\cap$ distributes over Δ: $A \cap (B \Delta C) = (A \cap B) \Delta (A \cap C)$

*(16) Remark 2 after Theorem 5.3 mentions a strong connection between set algebra and propositional logic. We now make this more precise. Consider any statement of set algebra that does not contain quantifiers or the symbols $\in$, $-$, or $\subset$. (Most of the results in Theorems 5.2 through 5.5 fit this description.) Turn the statement into a statement of pure logic by making the following replacements: change $\cup$ to $\vee$, $\cap$ to $\wedge$, $=$ to $\leftrightarrow$, $\subseteq$ to $\rightarrow$, $\varnothing$ to any contradiction, U (the universal set if one is being used) to any tautology, $'$ to $\sim$, and finally, every set variable to a propositional variable. You might also need to put in some parentheses to keep things grouped the way they were originally. Then, the original statement is a valid theorem of set algebra iff the new statement is a tautology.

(a) Transform each of the following results in the manner just described, and verify that the new statement is a tautology: parts (a) and (f) of Theorem 5.2; parts (a), (b), and (d) of Theorem 5.3; parts (c), (e), and (f) of Theorem 5.4; and parts (b), (c), and (f) of Theorem 5.5.

(b) Using our list of tautologies (Appendix 3) and the transformation process of part (a) *in reverse*, find at least two correct theorems of set algebra that have not been given in Theorems 5.2 through 5.5. Be careful, because not every tautology corresponds to a grammatically meaningful statement of set algebra; for example, tautology 12 does not.

(c) Figure out how to extend the above transformation process to include statements of set algebra that also contain the symbols − and ⊂.

*(17) Remark 3 after Theorem 5.3 mentions the duality principle of Boolean Algebra. For set algebra, this may be stated as follows. Consider any statement of set algebra of the type described in Exercise 16, except that it may also contain the symbol ⊂. Form a new statement of set algebra in this way: change every ∪ to ∩ and vice versa; change every ∅ to U and vice versa; and wherever a ⊆ or ⊂ occurs, switch the *expressions* on the left and right sides of the symbol. The new statement is called the **dual** of the original. Then the original is a valid theorem iff its dual is valid.

(a) Explain why the dual of the dual of any statement is the original statement. In other words, if Q is the dual of P, then P is the dual of Q.

(b) Identify at least five dual pairs of statements in Theorems 5.2 through 5.5.

(c) It is possible for a statement to be its own dual. Find three such statements.

(d) Using the ideas of this exercise and the previous one, describe how to define the dual of any statement of propositional logic.

5.3 More Advanced Set Operations

Set theory gets much more complicated when it starts dealing with sets of sets. A set of sets is often called a **collection** or a **family** of sets. Collections of sets are the main topic of this section. One simple way to define a set of sets is to start with any set and then consider the set of all the subsets of the original set. This process is so important that it deserves a name:

Definition: The **power set** of any set A is the set of all subsets of A, denoted $\wp(A)$. In symbols,

$$\wp(A) = \{B \mid B \subseteq A\} .$$

Note that $\wp(A)$ is automatically a set of sets, no matter what kind of set A is. Working with power sets can take some care, as the following examples illustrate.

Example 1: Let's figure out the members of $\wp(A)$, where $A = \{4, 7\}$. Clearly, A has one subset with two elements (itself), two subsets with one element, and one subset with no elements. Therefore, $\wp(A) = \{\varnothing, \{4\}, \{7\}, \{4, 7\}\}$ (see Figure 5.4).

Braces within braces are tricky at first, but with some practice you will find them familiar and easy to work with.

Example 2: What would $\wp(\varnothing)$ be? Well, what are the subsets of ∅? Does it have any subsets? Yes, it has one—itself. Therefore, $\wp(\varnothing) = \{\varnothing\}$. It's important to see that $\{\varnothing\}$ is not the same as ∅. The set ∅ has no members, while $\{\varnothing\}$ has one member.

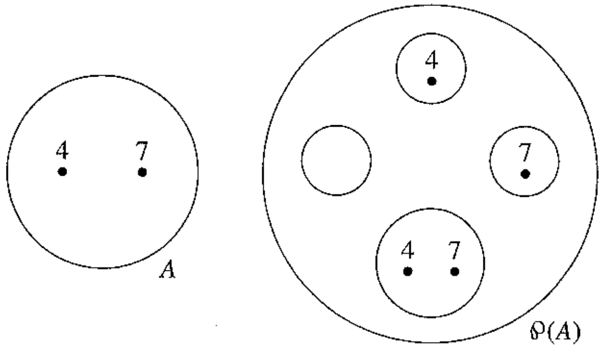


Figure 5.4 A simple set and its power set

Example 3: To carry this a step further, let's find the elements of $\wp(\wp(\varnothing))$. By Example 2, that has to be $\wp(\{\varnothing\})$. So what are the subsets of $\{\varnothing\}$? Well, this is a set with one element, so it must have two subsets: $\varnothing$ and the whole set. In other words, $\wp(\wp(\varnothing)) = \{\varnothing, \{\varnothing\}\}$.

The following theorem is helpful for keeping track of the number of elements in a power set and explains the origin of this term. It is also the first instance in this book of an induction proof that begins at $n = 0$ instead of $n = 1$. We mentioned in Section 4.5 that this is allowed, and Exercise 14 asks you to justify it.

Another complicating feature of this induction proof is that the theorem seems to be about sets more than it is about integers or natural numbers. Integers are mentioned in the statement of the theorem only to measure the number of elements in a set. It is fine to try to prove such a statement by induction, but it is usually necessary to have P(n) be the statement that the theorem is true for *all* sets with n elements. That will be our approach here. To put it another way, these theorems are basically about all *finite* sets. But rather than try to prove them for all finite sets at once, we prove them by induction on the number of elements in the set.

Theorem 5.8: If a set A has n elements, where n is any nonnegative integer, then $\wp(A)$ has 2^n elements.

Proof: We prove this by induction on n, starting at 0. For $n = 0$, the only set with 0 elements is $\varnothing$. We know that $\wp(\varnothing) = \{\varnothing\}$, which has one element. And since $2^0 = 1$, the theorem holds for $n = 0$.

For the induction step, assume the theorem holds for every set with n elements, and let A be any set with $n + 1$ elements. Pick any particular element of A, and call it c. *[This* *step is justified by ES.]* Let $B = A - \{c\}$. Note that B has n elements. Let's count the subsets of A. A subset of A either contains c as a member or it doesn't. The subsets of A that don't contain c are precisely the subsets of B, and so by the induction hypothesis there are 2^n of them. Furthermore, if D is any subset of B, then $D \cup \{c\}$ is a subset of A that contains c. It is also easy to see that every subset of A that contains c is of this form. Therefore, there are also 2^n subsets of A that contain c.

So the total number of subsets of A is $2^n + 2^n$, which equals 2^{n+1}, as desired. ■

Remarks: (1) This is probably the least formal proof in this book up to this point. You could try to prove this theorem more formally, but it is difficult to do so without using material from Chapter 7. Specifically, the argument in the third paragraph of the proof, and in fact the rigorous notion of what it means for a set to have n elements, are based on the concept of a one-to-one correspondence. Also, the definition of exponents is an inductive definition. But for most purposes, our proof is fine; it certainly conveys the main idea of why the theorem is true. A different approach to this proof is given in Theorem 7-17(d).

(2) Note that Examples 1 through 3 are of course consistent with Theorem 5.8: a set with 0, 1, or 2 elements must have (respectively) 1, 2, or 4 subsets.

(3) From Exercise 11 of Section 4.6, plus the fact that $0 < 2^0$, we know that $n < 2^n$ for any nonnegative whole number n. It follows that for every finite set A, $\wp(A)$ has more elements than A does. Theorem 7.26 shows, by a famous and ingenious argument, that this fact also holds for all infinite sets.

Here are a few basic results involving power sets:

Theorem 5.9: For any sets A and B:

(a) $\wp(A \cap B) = \wp(A) \cap \wp(B)$

(b) $A = B \leftrightarrow \wp(A) = \wp(B)$

(c) $A \subseteq B \leftrightarrow \wp(A) \subseteq \wp(B)$

(d) $A \subset B \leftrightarrow \wp(A) \subset \wp(B)$

Proof: (a)

$$\begin{aligned} C \in \wp(A \cap B) &\leftrightarrow C \subseteq A \cap B \\ &\leftrightarrow C \subseteq A \text{ and } C \subseteq B \qquad \text{By Theorem 5.5(f)} \\ &\leftrightarrow C \in \wp(A) \text{ and } C \in \wp(B) \\ &\leftrightarrow C \in \wp(A) \cap \wp(B) \end{aligned}$$

(b) If $A = B$, then $\wp(A) = \wp(B)$ by Theorem 4.7. For the reverse direction, assume $\wp(A) = \wp(B)$. Since $A \subseteq A$ and thus $A \in \wp(A)$, it follows by axiom III-4 that $A \in \wp(B)$. Therefore $A \subseteq B$. Similar reasoning shows $B \subseteq A$. By Theorem 5.4(f), $A = B$.

(c) and (d) See Exercise 6. ■

Indexed Families of Sets

Now let's consider more general sets of sets than just power sets. In theory, no special notation is needed to describe sets of sets. We could just begin a discussion or a proof with a statement such as "Let A be a collection of sets." However, certain notation for sets of sets has come into general use. For one thing, to distinguish them from ordinary sets, sets of sets are usually denoted by capital script letters. For the most part, we follow this practice.

It is also common, when describing a collection of sets, to denote the individual sets in the collection with a subscripted variable called an **index**. So a mathematician might define sets A_n for each natural number n, and then define the collection of all these sets A_n. In this type of situation, the set $\mathbb{N}$ (that is, the set over which the *subscript* ranges) is referred to as an **index set** for the collection of sets, and the collection itself is called an **indexed family of sets.**

Example 4: For each n in $\mathbb{N}$, let A_n be the closed interval $[n, n + 1/n]$. So $A_1 = [1, 2]$, $A_2 = [2, 2.5]$, and so on. Then we can define an indexed family of sets $\mathscr{A}$ by

$$\mathscr{A} = \{A_n \mid n \in \mathbb{N}\}$$

It is important to see that $\mathscr{A}$ is *not* a set of real numbers. It is a set of sets of real numbers. That is, $\mathscr{A}$ is not a subset of $\mathbb{R}$ or a member of $\wp(\mathbb{R})$; rather it's a subset of $\wp(\mathbb{R})$ and an element of $\wp(\wp(\mathbb{R}))$. You might be tempted to think of the collection $\mathscr{A}$ as an *infinite sequence* of sets rather than a set of sets. This is a plausible alternative view of the situation, and it becomes the preferable way to view indexed families when the *order* of the sets in the family is important or it is desirable to allow *repetition* of sets in the family. See Example 6 of Section 7.4.

$\mathbb{N}$ is not the only possible index set. Any set can be one. Here's an example where the set of real numbers is an index set.

Example 5: For each real number c, let

$$L_c = \{(x, y) \mid x \in \mathbb{R} \text{ and } y \in \mathbb{R} \text{ and } y = cx\}$$

Then we can define $\mathscr{A} = \{L_c \mid c \in \mathbb{R}\}$. Graphically, you can see that L_c is a straight line of slope c through the origin, in the xy plane. So $\mathscr{A}$ can be thought of as a set of lines in the plane (see Figure 5.5).

When an unspecified set is used as an index set, the letter I or J is usually used for the index set, and i or j (respectively) is usually used as the subscript. By the way, even though we are restricting our attention to indexed families of sets, it is permissible to define indexed families of any kind of mathematical object.

Unions and Intersections of Collections of Sets

We've discussed the familiar operations of forming the union and intersection of two sets. By repeating these, it's a simple matter to form the union and intersection of any

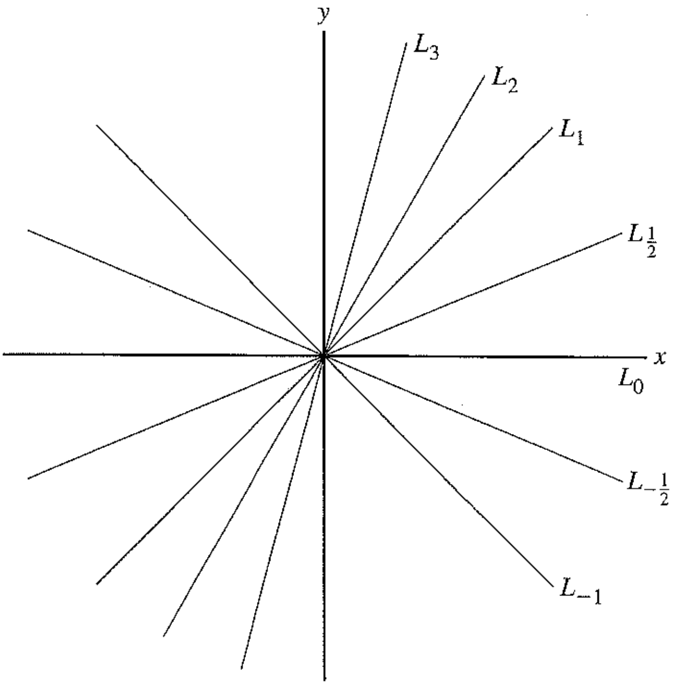


Figure 5.5 Illustration of Example 5: a family of sets indexed by $\mathbb{R}$

finite number of sets. We now define the important notions of the union and intersection of an arbitrary (and so, possibly infinite) collection of sets.

Definitions: Let $\mathscr{A}$ be any set of sets. Then the **union of** $\mathscr{A}$ or the **union over** $\mathscr{A}$, denoted $\bigcup_{A \in \mathscr{A}} A$ or simply $\bigcup\mathscr{A}$, is the set

$$\{x \mid \exists A \in \mathscr{A}(x \in A)\}$$

The **intersection of** $\mathscr{A}$ or the **intersection over** $\mathscr{A}$, denoted $\bigcap_{A \in \mathscr{A}} A$ or simply $\bigcap\mathscr{A}$, is the set

$$\{x \mid \forall A \in \mathscr{A}(x \in A)\}$$

Unless there is a universal set in the discussion, $\mathscr{A}$ must be nonempty for the second definition to make sense. The simple notations $\bigcup\mathscr{A}$ and $\bigcap\mathscr{A}$ introduced here are found in most set theory books but for some reason are rarely used by mathematicians other than set theorists. The longer notations are more common.

Notation: When working with an indexed family of sets, yet another notation is used for its union or intersection. If $\mathscr{A} = \{A_i \mid i \in I\}$, then the most common notation for the union of $\mathscr{A}$ is $\bigcup_{i \in I} A_i$, and similarly for the intersection of $\mathscr{A}$.

Take the time to see that these definitions say what they ought to. The union of $\mathscr{A}$ consists of all things that are in at least one of the sets in the collection $\mathscr{A}$, so it consists of all the sets in $\mathscr{A}$ "put together." Similarly, the intersection of $\mathscr{A}$ consist of all things that are in all the sets in the collection $\mathscr{A}$.

It was mentioned earlier that the quantifiers $\exists$ and $\forall$ are related to the connectives "or" and "and," respectively. The definitions of union and intersection of $\mathscr{A}$, together with the definitions of ordinary unions and intersections, are a good illustration.

Example 6: Let $\mathscr{A}$ be as defined in Example 4. Then the set $\cup\mathscr{A}$ could also be denoted $\bigcup_{n \in \mathbf{N}} A_n$. While $\mathscr{A}$ is not a set of real numbers, $\cup\mathscr{A}$ is; it's a set consisting of an infinite number of intervals. The set $\cap\mathscr{A}$ is also a set of real numbers, namely $\varnothing$.

Example 7: Let $\mathscr{A}$ be the collection defined in Example 5. Then $\cap\mathscr{A} = \{(0,0)\}$, since the origin is the one point common to all the lines L_c. Exercise 5 asks you to describe the set $\cup\mathscr{A}$.

Example 8: Here is the definition of the **Cantor set,** also known as **Cantor's discontinuum.** This set of real numbers is important in higher mathematics, and you will almost certainly encounter it again. To define the Cantor set, start with the closed unit interval [0, 1]. Then define sets A_n as follows:

Let A_1 be the open interval (1/3, 2/3).

Let $A_2 = (1/9, 2/9) \cup (7/9, 8/9)$.

Let $A_3 = (1/27, 2/27) \cup (7/27, 8/27) \cup (19/27, 20/27) \cup (25/27, 26/27)$.

To see the pattern here, note that A_1 is the middle third of the original interval, A_2 consists of the two middle thirds of what's left of the original interval after removing A_1, and so on (see Figure 5.6). Continuing in this way, we define sets A_n for every natural number n. (This definition can be made more algebraic and precise, if desired.) The Cantor set is then defined to be

$$[0, 1] - \bigcup_{n \in \mathbf{N}} A_n$$

Exercises 18 through 20 deal with some of the interesting features of the Cantor set.

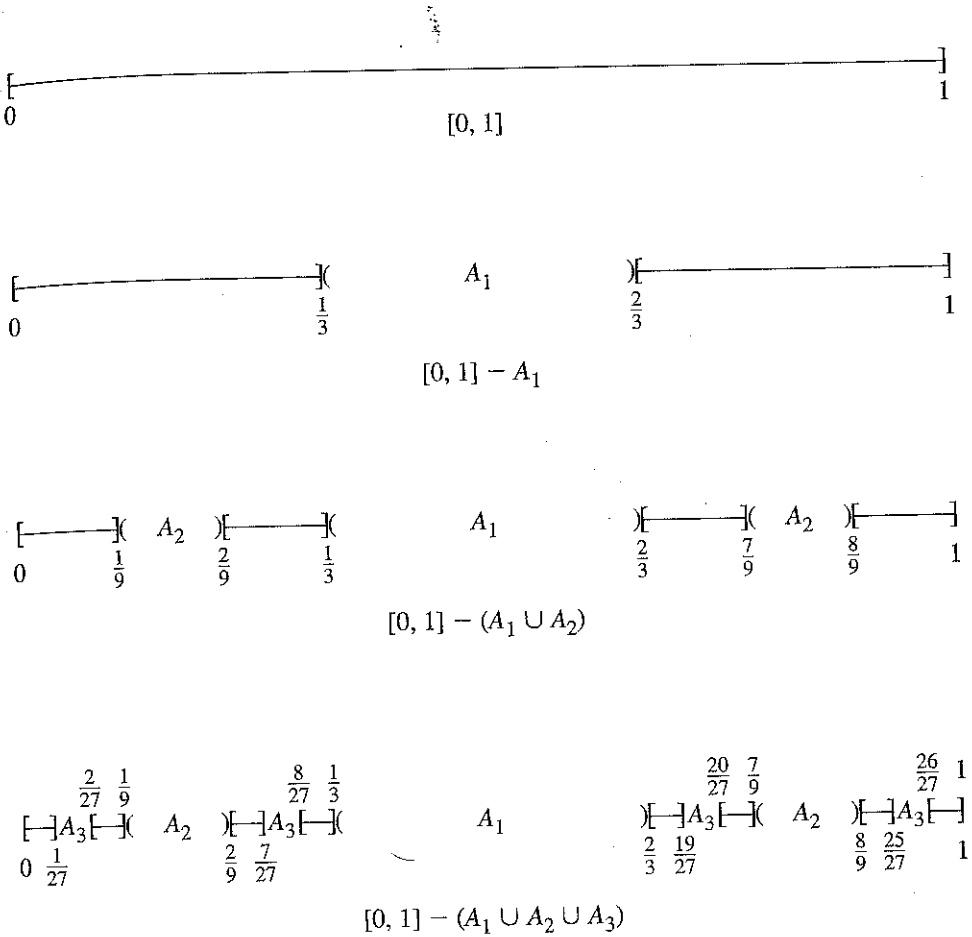


Figure 5.6 The first three or four stages in the construction of the Cantor set

Unions and intersections of infinite families of sets are sometimes called **infinitary** operations, as opposed to ordinary **finitary** operations. Most results about the finitary operations have infinitary analogs. Here are a few of these: see if you can figure out which parts of Theorems 5.2 and 5.3 the results of the following theorem are related to.

Theorem 5.10: Let B be any set, and $\mathscr{A} = \{A_i \mid i \in I\}$ any family of sets. As usual, when we use the symbol $'$, it means complement relative to some specified universal set.

(a) $B \cup \left(\bigcap_{i \in I} A_i\right) = \bigcap_{i \in I} (B \cup A_i)$

(b) $B \cap \left(\bigcup_{i \in I} A_i \right) = \bigcup_{i \in I} (B \cap A_i)$

(c) $(\bigcup_{i \in I} A_i)' = \bigcap_{i \in I} (A_i{}')$

(d) $(\bigcap_{i \in I} A_i)' = \bigcup_{i \in I} (A_i{}')$

(e) For each $i \in I, A_i \subseteq \bigcup_{i \in I} A_i$

(f) For each $i \in I,\ \bigcap_{i \in I} A_i \subseteq A_{\text{i}}$

Proof: (a) $x \in B \cup \left(\bigcap_{i \in I} A_i \right) \leftrightarrow x \in B$ or $x \in \left(\bigcap_{i \in I} A_i \right)$

$\leftrightarrow x \in B$ or $\forall i \in I\,(x \in A_i)$

$\leftrightarrow \forall i \in I\,(x \in B \text{ or } x \in A_i)$ By law of logic 17 (Figure 4.2)

$\leftrightarrow \forall i \in I\,(x \in B \cup A_i)$

$\leftrightarrow x \in \bigcap_{i \in I} (B \cup A_i)$

(e) Let $i \in I$ be arbitrary. If $x \in A_i$, then clearly $\exists i \in I\ (x \in A_i)$, so $x \in \bigcup_{i \in I} A_i$.

Therefore, $A_i \subseteq \bigcup_{i \in I} A_i$.

The proofs of the other parts are left for the exercises. ■

We have called parts (c) and (d) of Theorem 5.3 De Morgan's laws for sets. Similarly, parts (c) and (d) of Theorem 5.10 may be viewed as set versions of De Morgan's laws for quantifiers.

Exercises 5.3

(1) List all the members of:

(a) $\wp(\{1, 2, 3\})$

(b) $\wp(\{2, \{2\}\})$

(c) $\wp(\wp(\{\varnothing\}))$

(d) $\wp(\{1, 3, 5\}) \cap \wp(\{5, 6, 7\})$

(e) $\wp(\{1, 2, 3\}) - \wp(\{1, 3\})$

(2) True or false, with brief explanation:

(a) $3 \in \wp(\mathbb{N})$

(b) $\{3\} \in \wp(\mathbb{N})$

(c) $\{3\} \subseteq \wp(\mathbb{N})$

(d) $\{\varnothing\} \in \wp(\{\{\varnothing\}\})$

(e) $\wp(\mathbb{Z} \cap (2, 4)) = \{\varnothing, \{3\}\}$

(3) Characterize each of the following statements as always true, always false, or sometimes true and sometimes false. Explain briefly.

(a) $A \in \wp(A)$ (b) $A \subset \wp(A)$
(c) $A \cap B \in \wp(A \cup B)$ (d) $\wp(A) - \wp(B) = B - A$
(e) $\wp(\wp(A)) \in \wp(\wp(\wp(A)))$ (f) $\wp(A - B) \cap \wp(B - A) = \{\varnothing\}$

(4) Prove whichever of the following equations are true for all sets. For each one that's *not* always true, try to prove that one side is a subset of the other, and give a counterexample to the other direction. If neither side must be a subset of the other, give a counterexample to both directions.

(a) $\wp(A \cup B) = \wp(A) \cup \wp(B)$ (Compare with Theorem 5.9(a).)
(b) $\wp(A - B) = \wp(A) - \wp(B)$
(c) $\bigcup(\wp(A)) = A$
(d) $\wp(\bigcup\mathscr{A}) = \mathscr{A}$

(5) Referring to Examples 5 and 7, give a simple *geometric* description of $\bigcup\mathscr{A}$.

(6) Prove parts (c) and (d) of Theorem 5.9.

(7) Prove Theorem 5.10(b).

(8) Prove Theorem 5.10(c).

(9) Prove Theorem 5.10(d).

(10) Prove Theorem 5.10(f).

(11) For each $n \in \mathbb{N}$, let A_n be the interval $[2^{-n}, 2^{1-n})$. Give a simple description of the set $[0, 1] - \bigcup\limits_{n \in \mathbb{N}} A_n$. Explain briefly.

(12) Let $B = \{(x, y) \mid y = x\}$, where x and y are real variables. For each real number r, let $A_r = \{(x, y) \mid x^2 + y^2 = r^2\}$. Then $\{A_r \mid r \in \mathbb{R}\}$ is a family of circles indexed by $\mathbb{R}$.

(a) Note that using $\mathbb{R}$ as the index set causes most of the circles in this indexed family to be repeated. This is allowed but may cause needless confusion. Name two smaller index sets that can be used to define the same collection of circles without any repetition. On the other hand, it is occasionally useful to allow repetition of sets in an indexed family.

(b) What is the union of all the A_rs?
(c) Describe $B \cap A_r$.
(d) Verify that Theorem 5.10(b) holds in this case.

(13) Let $A = (0, 1)$ and $B = (0, 1]$. For any $C \subseteq \mathbb{R}$ and $y \in \mathbb{R}$, let $C + y$ be the set $\{x + y \mid x \in C\}$. That is, $C + y$ is the set C shifted or **translated** y units. Describe the following sets in words or in more concise mathematical notation.

(a) $\bigcup\limits_{n \in \mathbb{N}} (A + n)$ (b) $\bigcup\limits_{n \in \mathbb{N}} (B + n)$

(c) $\bigcup_{x \in A} (\mathbb{Z} + x)$ (d) $\bigcup_{x \in B} (\mathbb{Z} + x)$

(14) (a) Justify the use of induction proofs that start at 0 instead of at 1, as in Theorem 5.8. That is, prove

$$[\mathrm{P}(0) \wedge \forall n \geq 0\, (\mathrm{P}(n) \rightarrow \mathrm{P}(n+1))] \rightarrow \forall n \geq 0\, \mathrm{P}(n)$$

where n is an integer variable.

*(b) Generalizing part (a), show that an induction proof can start at any integer k. That is, prove

$$[\mathrm{P}(k) \wedge \forall n \geq k\, (\mathrm{P}(n) \rightarrow \mathrm{P}(n+1))] \rightarrow \forall n \geq k\, \mathrm{P}(n)$$

where n and k are both integer variables. ***Hint:*** Let $\mathrm{Q}(n)$ be the statement $\mathrm{P}(n+1-k)$. Also, you may use the result of Exercise 13 of Section 5.1.

(15) This problem illustrates a natural situation where you might do an induction proof beginning at a number other than 0 or 1: suppose that you want to prove that some property holds for all *polygons*. It might be natural to attempt this by induction on the number of sides.

(a) In such a proof, what would be the initial value of n (represented by k in Exercise 12(b))?

(b) For what class of polygons would you first have to show the result?

(c) State carefully what the induction step would be.

(16) Mathematicians often say things like "Let $\mathscr{A}$ be a collection of nonempty disjoint sets." From this, which of the following can you conclude?

(a) Each set in $\mathscr{A}$ is nonempty.

(b) Each set in $\mathscr{A}$ is disjoint.

Carefully explain the difference between parts (a) and (b). To avoid this subtle linguistic confusion, it is more precise to say, "Let $\mathscr{A}$ be a collection of nonempty, *pairwise* disjoint sets."

(17) Critique the following well-known and entertaining proof. It is included here because it is similar to the proof of Theorem 5.8. (If necessary, review the instructions for this type of problem in Exercises 4.2.)

Theorem: All horses are the same color.

Proof: By induction on n, we prove that, given any set of n horses, they are all the same color. Clearly, this implies all horses are the same color. For $n = 1$ (or $n = 0$ if we choose to start there), the statement is trivial. So assume the statement holds for n, and let A be any set of $n + 1$ horses. Let c be any horse in A. By the induction hypothesis, all the horses in $A - \{c\}$ are the same color. Let k be any other horse in A. Again, all the horses in $A - \{k\}$ are the same color. So c and k are both the same color as all the other horses in A. Therefore, all the horses in A are the same color, as desired.

*(18) Prove that the Cantor set (Example 8) consists of all numbers in [0,1] that have a base 3 expansion with no 1's.

*(19) Prove that the Cantor set contains no intervals. You may use the result of Exercise 18.

*(20) Prove that the Cantor set contains no isolated points. (A number x is an **isolated point** of a set A of real numbers if $x \in A$ and for some $c > 0$, no other number between $x - c$ and $x + c$ is in A.)

Suggestions for Further Reading: For a more complete treatment of basic set theory, see Stoll (1979), Suppes (1960), or Vaught (1995). Devlin (1993) is a good text at a somewhat higher level. Most logic and set theory books discuss paradoxes, including the paradoxes of set theory; Kline (1982) does so in more detail than most. Many authors have attempted to capture the brilliant essence of Gödel's incompleteness theorem, including Nagel and Newman (1958), Hofstadter (1989), and Smullyan (1992). For more information about the subject of boolean algebra, see Pfleeger and Straight (1985), Rueff and Jeger (1970), or Stoll (1979). The first two of these cover important applications of boolean algebra such as switching theory.

Chapter 6

Relations

6.1 Ordered Pairs, Cartesian Products, and Relations

In this chapter, we study the important subject of binary relations. This section is devoted primarily to definitions of basic concepts, and Sections 6.2 and 6.3 discuss two useful types of relations. The single most important type of relation, functions, is covered in Chapter 7.

In Chapters 6 and 7, several concepts are defined in two different ways. The first version is always a rather nontechnical or intuitive one, and the second is a more rigorous one involving sets. Some books give only one of the two definitions in each case, but it's more educational to see both approaches.

Section 5.1 mentions that virtually all mathematical concepts can be defined and developed in terms of set theory. Even numbers (including natural numbers, rational numbers, real numbers, and so on) can be defined as special types of sets. All mathematicians are aware of this set-theoretic approach to mathematics but generally find it artificial and prefer not to think of relations and functions (let alone numbers!) as types of sets. Also, remember that set theory is only a hundred years old, whereas the study of numbers and functions is much older. So most of the fundamental ideas of mathematics are not based on set theory. On the other hand, the set-theoretic approach is very useful for proofs and theoretical work. So you can see why you should learn both approaches to these basic concepts.

Definition (intuitive): For any two objects a and b, the **ordered pair** (a, b) is a notation specifying the two objects a and b, in that order.

Definition (set-theoretic): For any two objects a and b, the **ordered pair** (a, b) is defined to be the set $\{\{a\}, \{a, b\}\}$.

Perhaps you can see why neither definition of ordered pairs is totally satisfactory. The intuitive definition is not a rigorous mathematical definition, any more than it would be to define a natural number as a sequence of digits. That sounds fine at first, but a number is definitely not the same thing as the *numeral* used to denote it. On the other hand, the set-theoretic definition is very strange looking and conveys none of the intuitive meaning of what an ordered pair is. For these reasons, many mathematicians view the concept of ordered pairs as an undefined, primitive notion.

There are just two important properties of ordered pairs. The first is that you can form the ordered pair of any two objects whatsoever. The second is the familiar condition for equality of ordered pairs:

$$(a, b) = (c, d) \text{ iff } a = c \text{ and } b = d$$

These properties appear as set axioms IV-4 and IV-5 in Appendix 1. However, if the set-theoretic definition of ordered pairs is used, these axioms are provable and therefore superfluous (see Exercise 7).

Definition: For any two sets A and B, their **cartesian product** is the set of all ordered pairs whose first member is in A and whose second member is in B; in symbols,

$$A \times B = \{(x, y) \mid x \in A \text{ and } y \in B\}$$

René Descartes (1596–1650), from whose name the word "cartesian" is derived, was an extremely important figure in the development of modern mathematics and philosophy. As a child his health was poor, and he developed a lifelong habit of spending his mornings in bed, thinking and writing. At the age of eighteen his life entered a less intellectual phase, including a short period of heavy gambling and several years of intermittent military service. Fortunately, inspired in part by three vivid dreams in 1619, Descartes quit the military and devoted the rest of his life to academic pursuits.

Descartes's major mathematical achievement was the invention of analytic geometry: the system whereby equations can be graphed and, conversely, geometric figures can be analyzed algebraically. The importance of this contribution to mathematics—a two-way link between symbolic entities (equations and inequalities) and pictorial entities (straight lines, circles, parabolas, and so on)—would be difficult to overestimate. This achievement also strongly influenced his philosophical views, notably that "pure reason," of a mathematical sort, was the correct path to truth and knowledge. His famous conclusion, "Cogito, ergo sum" ("I think, therefore I am"), also emphasized the importance of the individual and rationality.

Descartes was deeply religious, but his emphasis on the individual and reason was not consistent with the views of the Catholic Church. For this and other reasons, he spent the second half of his life away from his native France, mostly in Holland.

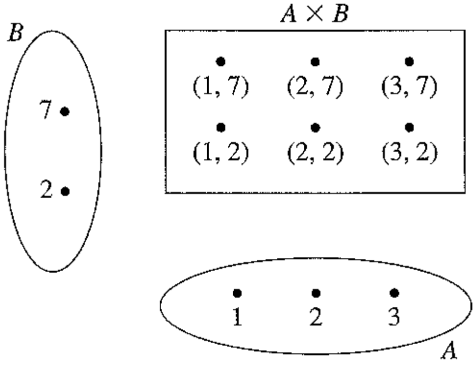


Figure 6.1 A simple cartesian product

You have been using cartesian products since eighth or ninth grade, whether or not you have called them that. Every time you draw a graph in an ordinary two-dimensional rectangular coordinate system, the coordinate system allows you to represent ordered pairs of real numbers on paper. In other words, the coordinate system turns your piece of paper into a picture of the cartesian product $\mathbb{R} \times \mathbb{R}$.

In the notation $A \times B$, the × is normally pronounced "cross" (not "times" or "ex"), and a cartesian product may also be called a cross product. But there is absolutely no connection between this notion and vector cross product. However, there is a definite connection between cartesian products and ordinary multiplication, as we now see.

Theorem 6.1: If A and B are sets, with m and n members respectively ($m, n \geq 0$), then $A \times B$ has mn members.

Proof: This proof can be done by a straightforward induction on either m or n (starting at 0), and we leave it for Exercise 9. ■

Example 1: Let $A = \{1, 2, 3\}$ and $B = \{7, 2\}$. Then $A \times B$ is the set $\{(1, 7), (1, 2), (2, 7), (2, 2), (3, 7), (3, 2)\}$. Note that $A \times B$ has six ordered pairs and that we have listed them in a systematic manner (see Figure 6.1).

Example 2: Suppose a certain T-shirt comes in sizes small, medium, large, and extra large and in colors red, blue, green, orange, and purple. If we write $A = \{\text{S, M, L, X}\}$ and $B = \{\text{R, B, G, O, P}\}$, then $A \times B$ consists of 20 ordered pairs. These ordered pairs may be viewed as representing all the possible choices of size and color for this type of shirt. Figure 6.2 is a tree diagram illustrating this situation.

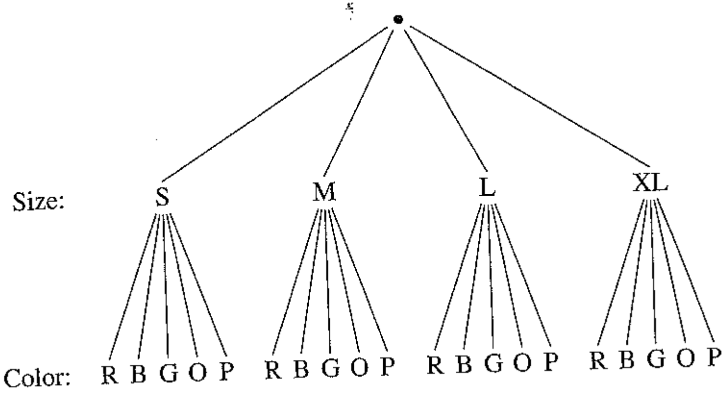


Figure 6.2 Tree diagram for Example 2

Is cartesian product a commutative operation? In other words, does $A \times B = B \times A$ in general? It's pretty clear that the answer is no. For instance, in Example 1, $A \times B$ contains $(2, 7)$ but not $(7, 2)$, whereas $B \times A$ contains $(7, 2)$ but not $(2, 7)$. And we know that $(2, 7) \neq (7, 2)$. So $A \times B$ and $B \times A$ are different as sets of ordered pairs. On the other hand, Theorem 6.1 at least guarantees that if A and B are finite sets, then $A \times B$ and $B \times A$ must have the same number of elements.

Is cartesian product an associative operation? In other words, does $(A \times B) \times C = A \times (B \times C)$, in general? Once again, the answer is no, but in this case the difference between the two cartesian products is often insignificant. For instance, let $A = B = C = \mathbb{N}$. Then $(A \times B) \times C$ contains $((2, 7), 9)$, while $A \times (B \times C)$ contains $(2, (7, 9))$. Does $((2, 7), 9) = (2, (7, 9))$? Technically, no, but in many situations, it might as well. This discussion leads us to the concepts of ordered triples, ordered n-tuples, and cartesian products of more than two sets.

Definitions: For any three objects a, b, and c, the **ordered triple** (a, b, c) is defined to be $((a, b), c)$.

For any sets A, B, and C, $A \times B \times C = \{(a, b, c) \mid a \in A,\ b \in B, \text{ and } c \in C\}$.

Remarks: (1) These definitions represent a somewhat arbitrary choice: technically, (a, b, c) means $((a, b), c)$, not $(a, (b, c))$. And $A \times B \times C$ means $(A \times B) \times C$, not $A \times (B \times C)$. But as the previous discussion indicates, this distinction can often be ignored.

(2) These definitions can be extended to define **ordered n-tuples** and cartesian products of any number of sets. These would be so-called **inductive** definitions. In other words, if we have already defined what $(a_1, a_2, \ldots, a_n)$ means, then we can define

$(a_1, a_2, ..., a_{n+1})$ to be $((a_1, a_2, ..., a_n), a_{n+1})$. And if we have already defined what $A_1 \times A_2 \times ... \times A_n$ means, we can then define the cartesian product $A_1 \times A_2 \times ... \times A_{n+1}$ to be $(A_1 \times A_2 \times ... \times A_n) \times A_{n+1}$. Inductive definitions are studied in Chapter 7. (One often writes A^2 for $A \times A$, A^3 for $A \times A \times A$, and so on.) Theorem 6.1 generalizes to cartesian products of three or more sets (see Exercise 10).

(3) Theorem 6.1 is the basis for an extremely useful formula called the **fundamental counting principle** or the **product rule for counting**. Recall Example 2. In that situation, there are four possibilities for the size of a shirt and five possibilities for the color. To select a shirt, a person must decide on both size and color. Theorem 6.1 says that the number of possible ways to make this *sequence* of decisions is 4 × 5, or 20. Like Theorem 6.1, this formula can be extended to a sequence of three or more decisions. In its general form, the product rule may be stated as follows:

☞ *Suppose that it is required to make a sequence of k decisions. If there are* n_1 *possible choices for the first decision,* n_2 *possible choices for the second decision, and so on, then the number of possible ways to make the whole sequence of decisions is* $n_1 n_2 ... n_k$.

Relations

We're now ready to define the major concept of this chapter. Once again, we give two definitions, one intuitive and one set-theoretic. Neither one is particularly complicated.

Definition (intuitive): A **binary relation** is a statement with two free variables (which are usually assumed to have specific sets as their domains, although the variables may be unrestricted).

Definition (set-theoretic): A **binary relation** is a set of ordered pairs.

The word "binary" means "pertaining to the number two"; in the term "binary relation" it refers to the fact that there are two free variables. Similarly, a **ternary relation** is a statement with three free variables, or a set of ordered triples. In general, an ***n*-ary relation** is a statement with *n* free variables or a set of ordered *n*-tuples. Also, the term **unary relation** is occasionally used, but this is essentially just a fancy term for a set or a subset of a specified universal set. Our primary interest is in binary relations, and so we normally omit the word "binary" when discussing them. Another (equivalent) way to give the set-theoretic definition of a relation is to say that it's a subset of some cartesian product.

Example 3: The intuitive definition of a binary relation describes what most people would call a relationship between two things. For example, the statement that one number is less than another is a binary relation. The statement that two numbers add up to 74 is a binary relation. The statement that one person is the father of another is a binary relation.

You can see that both forms of the definition of binary relations are quite simple. But it's not so easy to see the connection between the two or why they are two definitions of the same concept. Here is an example of how the two are connected.

Example 4: Let x and y be real variables. Then one simple type of relation (in the intuitive sense) would be an equation, like $x^2 + y^2 = 25$. The set-theoretic counterpart of this relation is then the set of all ordered pairs that satisfy the equation, or $\{(x, y) \mid x^2 + y^2 = 25\}$. So it would include ordered pairs such as (0, 5), (−5, 0), (4, −3), and so on. So when you *graph* an equation, you are really drawing a picture of the set of ordered pairs corresponding to that equation. Some mathematicians prefer to say that the equation is the relation, and the set of ordered pairs you get is not the relation but rather the graph of the relation.

Given any statement with two free variables, it's easy to form a set of ordered pairs in this manner. (You might have to make an arbitrary choice of which variable will be first and which second in the ordered pairs.)

Definitions: A subset of $A \times B$ is called a **relation between A and B**. Also, a subset of $A \times A$ is called a **relation on A**.

Notation: The letters R, S, and T denote binary relations. We sometimes write xRy as a shorthand for $(x, y) \in R$.

Definitions: Given a relation R, the **domain** (respectively, **range**) of R is the set of all objects that occur as a first (respectively, second) member of some ordered pair in R. In symbols,

$$\mathrm{Dom}(R) = \{x \mid \exists y\,(xRy)\}$$

$$\mathrm{Rng}(R) = \{y \mid \exists x\,(xRy)\}$$

Note that the concept of the domain of a *relation* is somewhat different from the concept of the domain of a *variable*, which we have been using since Chapter 3.

Example 5: Let $R = \{(3, 6), (8, 2), (1, 2), (0, 0), (-5, 3)\}$. Then R is clearly a relation, that is, a set of ordered pairs. By inspection, $\mathrm{Dom}(R) = \{3, 8, 1, 0, -5\}$, while $\mathrm{Rng}(R) = \{6, 2, 0, 3\}$.

Example 6: Again consider the relation on $\mathbb{R}$ defined by the equation $x^2 + y^2 = 25$. Whether we consider this relation to be an equation or a set of ordered pairs, its domain and range are both the interval $[-5, 5]$. How do we know this? The standard algebraic method is to solve the equation for each of the variables. If this equation is solved for y, it becomes $y = \pm\sqrt{25 - x^2}$. Since negative numbers do not have square roots, this is possible if and only if $x^2 \le 25$, which in turn means $-5 \le x \le 5$, or $x \in [-5, 5]$. (Here we are assuming that every nonnegative number has a square root.) Exercise 12 asks you to show this more rigorously.

Example 7: Consider the "less than" relation on $\mathbb{R}$, that is, $\{(x, y) \mid x, y \in \mathbb{R}$ and $x < y\}$. Its domain and range are both $\mathbb{R}$. What does the graph of this relation look like?

Example 8: Let A be the set of all people who have ever lived, and let R be the "parenthood" relation on A; that is, $R = \{(u, v) \mid u, v \in A$ and u is a parent of $v\}$. Then Dom(R) is the set of all people who have ever had a child, while Rng(R) might be A—might be because whether one believes in evolution or in divine creation, there is some question about whether the first humans are in Rng(R).

Inverse Relations

We conclude this section with a brief discussion of one of the most important ways of forming new relations. This concept is of primary importance in Chapter 7.

Definition: If R is any binary relation, the **inverse** of R, denoted R^{-1} and read R inverse, is the relation obtained by reversing the order of all the ordered pairs in R. In symbols,

$$R^{-1} = \{(y, x) \mid (x, y) \in R\}$$

This definition is given in terms of ordered pairs. If we are thinking of relations as propositions, we can pretty much say that "inverse" means "reverse," as the following examples illustrate.

Example 9: Recalling Example 7, the inverse of the "less than" relation is $\{(y, x) \mid y, x \in \mathbb{R}$ and $x < y\}$, which is the same as $\{(x, y) \mid x, y \in \mathbb{R}$ and $x > y\}$. In other words, the inverse of the "less than" relation is the "greater than" relation.

Example 10: Similarly, the inverse of the relation in Example 8 is $\{(u, v) \mid u, v \in A$ and u is a child of $v\}$. Thus, the inverse of the "parent of" relation is the "child of" relation.

Example 11: Figure 6.3 shows several pairs of graphs of relations and their inverses, involving real variables. Note that in each case the equation or other algebraic statement for the inverse relation is obtained simply by switching the variables. However, people often like to see graphs of equations solved for y, not x. For instance, it might be preferable to transform the first inverse equation in Figure 6.3 into the equivalent form $y = \pm\sqrt{x + 5}$, using elementary algebra.

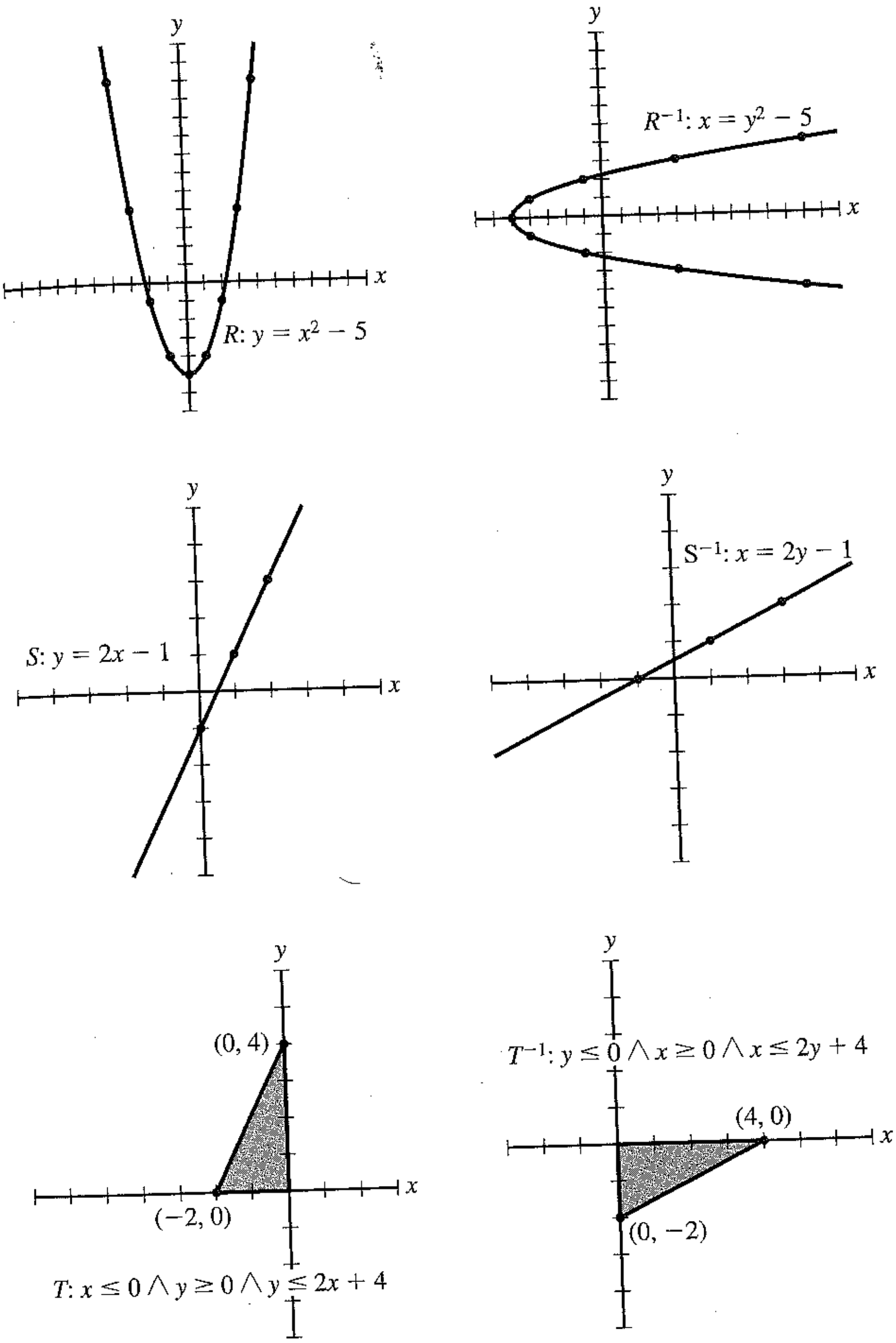


Figure 6.3 **Graphs of several relations and their inverses**

Observe (from Figure 6.3) that there is a basic geometric similarity between the graphs of R and R^{-1}: *For any relation R on $\mathbb{R}$, the graph of R^{-1} is obtained from the graph of R by reflecting it across the line $y = x$.* Note that we don't get the inverse graph from the original by a rotation; an actual reflection, or flip, is required.

Example 12: A relation can be its own inverse. Refer to Example 4.

We conclude this section with some simple facts:

Theorem 6.2: For any relation R:
(a) $(R^{-1})^{-1} = R$
(b) $\mathrm{Dom}(R^{-1}) = \mathrm{Rng}(R)$
(c) $\mathrm{Rng}(R^{-1}) = \mathrm{Dom}(R)$
Proof: See Exercise 11. ■

Another way of putting part (a) of Theorem 6.2 is that if S is the inverse of R, then R is also the inverse of S. In other words, the property that one relation is the inverse of the other is symmetric. So it is common to say simply, "R and S are inverses."

Exercises 6.1

(1) Let $R = \{(8, 4), (-3, 4), (2, 1)\}$ and $S = \{(8, 2), (2, 3), (8, 4), (\pi, \pi)\}$. Find:
(a) $\mathrm{Dom}(R)$ (b) $\mathrm{Rng}(R)$
(c) $\mathrm{Dom}(S)$ (d) $\mathrm{Rng}(S)$
(e) R^{-1} (f) $R \cup S^{-1}$
(g) $\mathrm{Dom}(R \cap S)$ (h) $\mathrm{Dom}(R) \cap \mathrm{Dom}(S)$
(i) $\mathrm{Dom}(R \cup S)$ (j) $\mathrm{Dom}(R) \cup \mathrm{Dom}(S)$

(2) Graph the following subsets of $\mathbb{R} \times \mathbb{R}$.
(a) The relation R of Exercise 1
(b) The relation S of Exercise 1
(c) $\{(x, y) \mid y = x^2 \text{ or } x = y^2\}$
(d) $\{(x, y) \mid y = x^2 \text{ and } x = y^2\}$
(e) $\{(x, y) \mid y \geq 2x \text{ and } x \leq 3y + 2\}$
(f) $\{(x, y) \mid y \geq 2x \text{ or } x \leq 3y + 2\}$
(g) $\{(x, y) \mid y = x^3 \text{ and } y \neq x\}$
(h) $\{(x, y) \mid |x| - |y| = 5\}$
(i) $\ln y = 3 \ln x$
(j) $\{(x, y) \mid x^2 + y^2 = 4$ or $x^2 + y^2 = 0$ or $(x - 1/2)^2 + (y - 1)^2 = 0$ or $(x + 1/2)^2 + (y - 1)^2 = 0$ or $y = -\sqrt{1 - x^2}\}$. ***Hint:*** Have a nice day!

(3) Find the domain and range of each of the following relations on $\mathbb{R}$. (The variables x and y denote real numbers). As much as possible, use the method shown in

Example 6. But you may need to resort to graphing or calculus. Justify your answers with more than a graph.

(a) $x^2 + y^3 = 7$ (b) $xy = 1$
(c) $|x| + |y| = 5$ (d) $|x + y| = 5$
(e) $y = \sin x + \cos x$ (f) $e^x + e^y = 0$

(4) Let $A = \{1, 2, 3\}$ and $B = \{2, 5\}$. How many elements are there in these sets?

(a) $\wp(A \times B)$ (b) $\wp(A) \times \wp(B)$
(c) $(A \times B) - (A \cup B)$ (d) $(A \cup B) - \wp(A \cup B)$
(e) $B \times B \times B \times B$

(5) In the country of Tannu Tuva, a valid license plate consists of any digit except 0, followed by any two letters of the alphabet, followed by any two digits.

(a) Let D be the set of all digits and L the set of all letters. With this notation, write the set of all possible license plates as a cartesian product.

(b) How many possible license plates are there?

(6) The Fishskill Numismatics Club has 10 members. The club must choose a slate of officers: president, treasurer, and secretary. How many possible slates are there, given that:

(a) Two or more positions may be filled by the same person.
(b) The officers must be three different people.

(7) Using the set-theoretic definition of ordered pairs (and naive set theory), derive axiom IV-5, the condition for the equality of ordered pairs.

(8) (a) Using the definition of ordered triples and the condition for the equality of ordered pairs, prove that $(x, y, z) = (u, v, w)$ iff $x = u, y = v$, and $z = w$.

(b) By induction, extend part (a) to the obvious condition for equality of two ordered n-tuples $(a_1, a_2, ..., a_n)$ and $(b_1, b_2, ..., b_n)$.

(9) Prove Theorem 6.1. Do *not* try to make the proof of the induction step very rigorous; just make sure you do a careful count. ***Hint:*** You may want to refer to the proof of Theorem 5.8, which has a similar flavor.

(10) Using Theorem 6.1 and induction, state and prove a generalization of Theorem 6.1 to cartesian products of any number of sets.

(11) Prove Theorem 6.2.

(12) Prove rigorously that the range of the relation $\{(x, y) \mid x, y \in \mathbb{R} \text{ and } x^2 + y^2 = 25\}$ is $[-5, 5]$. You may use the discussion in Example 6 as a starting point, and you may assume without proof that every nonnegative real number has a square root. You may also use Theorem A-11 of Appendix 2.

(13) Prove that for any relations R and S:
(a) $\mathrm{Dom}(R \cup S) = \mathrm{Dom}(R) \cup \mathrm{Dom}(S)$
(b) $\mathrm{Rng}(R \cup S) = \mathrm{Rng}(R) \cup \mathrm{Rng}(S)$

(14) Prove or disprove: for all relations R and S,
(a) $\mathrm{Dom}(R \cap S) = \mathrm{Dom}(R) \cap \mathrm{Dom}(S)$
(b) $\mathrm{Rng}(R \cap S) = \mathrm{Rng}(R) \cap \mathrm{Rng}(S)$
(c) $\mathrm{Dom}(R - S) = \mathrm{Dom}(R) - \mathrm{Dom}(S)$
(d) $\mathrm{Rng}(R - S) = \mathrm{Rng}(R) - \mathrm{Rng}(S)$
For any part that isn't true, try to prove that one side is a subset of the other.

(15) Prove or disprove: $\forall A, B, C \ (A \times B = A \times C \text{ iff } B = C)$

Critique the proofs in the remaining exercises in this section. (If necessary, refer to Exercises 4.2 for the instructions for this type of problem).

(16) **Theorem:** For any sets A and B, $\mathrm{Dom}(A \times B) = A$ and $\mathrm{Rng}(A \times B) = B$.
Proof: Let's first show $\mathrm{Dom}(A \times B) = A$. To do this, we must show each side is a subset of the other. So first assume $x \in \mathrm{Dom}(A \times B)$. Then, for some y, (x, y) is in $A \times B$. Thus $x \in A$. Conversely, assume $x \in A$. Then, for any $y \in B$, we have that (x, y) is in $A \times B$. Therefore, $x \in \mathrm{Dom}(A \times B)$. This completes the proof that $\mathrm{Dom}(A \times B) = A$. The proof that $\mathrm{Rng}(A \times B) = B$ is almost identical.

(17) **Theorem:** For any relations R and S:
(a) $(R \cup S)^{-1} = R^{-1} \cup S^{-1}$
(b) $(R \cap S)^{-1} = R^{-1} \cap S^{-1}$
(c) $(R - S)^{-1} = R^{-1} - S^{-1}$
Proof: (a) For any objects x and y,

$$\begin{aligned}(x, y) \in (R \cup S)^{-1} &\text{ iff } (y, x) \in (R \cup S)\\ &\text{ iff } (y, x) \in R \text{ or } (y, x) \in S\\ &\text{ iff } (x, y) \in R^{-1} \text{ or } (x, y) \in S^{-1}\\ &\text{ iff } (x, y) \in R^{-1} \cup S^{-1}\end{aligned}$$

By extensionality, we are done.
(b) and (c) These arguments are almost identical to the argument for part (a).

6.2 Equivalence Relations

Definitions: A relation R is called

- **reflexive on A** iff $\forall x \in A, xRx$.
- **symmetric** iff $\forall x, y$ (xRy implies yRx).
- **transitive** iff $\forall x, y, z$ (xRy and yRz implies xRz).

An **equivalence relation on** A is a relation on A that is reflexive on A, symmetric, and transitive.

It would be a good idea for you to take note of exactly where the phrase "on A" appears in these definitions and to try to see *why* it appears where it does. None of the instances of this phrase can be casually deleted. On the other hand, mathematicians sometimes call a relation R **reflexive**, with no reference to a set A. Technically, this means that R is reflexive on the set $\mathrm{Dom}(R) \cup \mathrm{Rng}(R)$. Similarly, if we simply say that R is an **equivalence relation**, we mean that R is an equivalence relation on $\mathrm{Dom}(R)$. Exercises 5 and 6 are intended to clarify some subtleties in the above terminology, as is the following:

Theorem 6.3: (a) R is reflexive iff, whenever (x, y) is in R, so are (x, x) and (y, y).

(b) If R is symmetric, then $\mathrm{Dom}(R) = \mathrm{Rng}(R)$.

(c) If R is symmetric and transitive, then R is an equivalence relation (on the set $\mathrm{Dom}(R)$).

Proof: (a) and (b) See Exercise 10.

(c) Assume R is symmetric and transitive. By part (b), R is a relation on $\mathrm{Dom}(R)$. We also must show that R is reflexive. By part (a), we just need to show that xRy implies xRx and yRy. This is also left for Exercise 10. ■

Theorem 6.3(c) can be a time-saver. If we just want to show that R is an equivalence relation, with no reference to a domain, then we don't need to prove it's reflexive. In practice, the more usual situation is that we know R is a relation on some set A, and we want to show it's an equivalence relation on A. Then, besides proving symmetry and transitivity, we must also show that $\mathrm{Dom}(R) = A$.

Notation: If R is known to be an equivalence relation, it is fairly common to write $x \equiv y$, $x \approx y$, or $x \sim y$ instead of xRy. In fact, mathematicians sometimes use these symbols to name equivalence relations, for example, "Let $\equiv$ be an equivalence relation."

Example 1: There's one equivalence relation that we've been working with since Chapter 3, namely the **identity,** or **equality, relation**. The first three of our equality axioms say that the relation $x = y$ is reflexive, symmetric, and transitive. Hopefully, it makes sense to you that equality is one way to define how things can be equivalent; in fact it can be argued that equality is the simplest possible equivalence relation.

The one thing that's a bit unusual about equality as a relation is that there's no particular domain for it. It can be thought of as a relation whose domain consists of all objects. Another approach is to say that given any set A, the equality relation, if restricted to $A \times A$, is an equivalence relation on A.

Notation: The identity relation on a set A is denoted id_A. In symbols,

$$\mathrm{id}_A = \{(x, x) \mid x \in A\}$$

Example 2: The relation $\{(1, 1), (2, 2), (2, 3), (3, 2), (3, 3), (4, 4)\}$ is an equivalence relation on $\{1, 2, 3, 4\}$, as is easily verified (see Exercise 7).

Example 3: Let's find all possible equivalence relations on the set $A = \{1, 2\}$. This is simpler than it sounds. To be reflexive on A, such a relation must include (1, 1) and (2, 2). The only other ordered pairs it could contain are (1, 2) and (2, 1). And, by symmetry, if it contains one of these, then it contains the other. So there are only two equivalence relations on $\{1, 2\}$: one is id_A and the other is $A \times A$. Exercise 19 asks you to carry out more investigations of this sort.

Example 4: The "siblinghood" relation on the set P of all people who have ever lived is an equivalence relation, provided that we agree that each person is to be considered his or her own sibling. That is, let $S = \{(x, y) \mid x$ and y have the same parents$\}$, where x and y are people variables. It is then quite simple to verify that S is reflexive on P, symmetric, and transitive. Similarly, let $F = \{(x, y) \mid x$ and y have the same father$\}$ and $M = \{(x, y) \mid x$ and y have the same mother$\}$. Then F and M are also equivalence relations on P.

Example 5: On the other hand, consider the "half-siblinghood" relation, that is, $\{(x, y) \mid x$ and y have at least one parent in common$\}$. This is not an equivalence relation. Can you see why?

Example 6: **Similarity** is an equivalence relation on the set of all triangles. Recall that two triangles are **similar** iff they have equal angles. Alternatively, two triangles are similar iff one is a scale model of the other (with a reflection, or flip, also allowed). Similarity is also an equivalence relation on the set of all polygons. For polygons with more than three sides, similarity is defined by the scale model idea, not by angles. The definition using angles would still yield an equivalence relation, but it's not called similarity.

Example 7: Another equivalence relation on the set of all triangles or on the set of all polygons is **congruence.** Recall that two polygons are **congruent** iff they have the same angles and sides, in the same order (except that one can be clockwise and the other counterclockwise). This is the same as saying that if one of the polygons were rigidly constructed out of sticks, you could turn it into the other one just by moving the first polygon and/or flipping it over.

Example 8: On the set $\mathbb{R}$, the relation $\{(x, y) \mid x - y$ is an integer$\}$ is an equivalence relation. More generally, let c be any fixed real number, and replace the words "is an integer" in this definition with the words "is an integer multiple of c." Then for each c, this is an equivalence relation, called **congruence modulo c**. The usual notation for congruence modulo c is $x \equiv y \pmod{c}$. This type of equivalence relation, with c being an integer, is extremely important in number theory. It is discussed further in Section 8.3. By the way, congruence modulo c has nothing to do with the geometric notion of congruence discussed in Example 7. But in general, when mathematicians refer to any type of congruence or similarity, it always denotes an equivalence relation.

Example 9: There are many other ways to define an equivalence relation on $\mathbb{R}$. One is $|x| = |y|$, which is the same as saying $x = \pm y$. A somewhat more complicated one would be $\sin x = \sin y$. These equivalence relations are examples of a very general sort, and even though we have not studied functions yet, here is the general way such equivalence relations are formed.

Theorem 6.4: If A is any set, and f is any function that is defined on all the elements of A, then the relation $\{(x, y) \mid f(x) = f(y)\}$ is an equivalence relation on A.

Proof: This theorem doesn't take much proving. Since $f(x) = f(x)$ in all cases, the relation is reflexive. If $f(x) = f(y)$, then $f(y) = f(x)$, so the relation is symmetric. Transitivity is equally simple to show. ■

Note that Theorem 6.4 is derived directly from the fact that equality is an equivalence relation. It says that any function defined on a set creates an equivalence relation on that set. It turns out that a sort of converse holds: given any equivalence relation R on a set A, there's a function f defined on A such that, for any $x, y \in A$, xRy holds if and only if $f(x) = f(y)$. (See Exercise 9 of Section 7.1.)

For instance, in Example 9, the first relation corresponds to the function $f(x) = |x|$ and the second to the function $f(x) = \sin x$. In Example 4, F corresponds to the function $f(x) = x$'s father, M corresponds to the function $f(x) = x$'s mother, while for the relation S, $f(x)$ could be the ordered pair (x's father, x's mother).

In Example 8 it's somewhat harder to find the appropriate functions, but it can be done. For the first equivalence relation, we could use $f(x) = x - \lfloor x \rfloor$, where $\lfloor x \rfloor$ means the greatest integer that is equal to or less than x. (So $\lfloor 5 \rfloor = 5$, $\lfloor 3.27 \rfloor = 3$, $\lfloor -3.27 \rfloor = -4$, and so on. Therefore, $f(5) = 0$, $f(3.27) = 0.27$, $f(-3.27) = 0.73$, and so on. For positive numbers, $f(x)$ is the decimal part of x, but for negative numbers, it's a bit different.)

The point of this discussion and these examples is the following rule of thumb:

☞ *An equivalence relation almost always expresses some way in which two things are the same or alike.*

This follows from Theorem 6.4: To say that an equivalence relation must be definable by a statement of the form $f(x) = f(y)$ is to say that the relation must express that some characteristic or property of x and y is the same. A binary relation that does not express some kind of alikeness is probably not an equivalence relation.

Example 10: Consider the relation on $\mathbb{R}$ defined by $R = \{(x, y) \mid x - y < 1\}$. This is not an equivalence relation; for one thing it's not symmetric. If we instead define $S = \{(x, y) \mid |x - y| < 1\}$, then S is reflexive and symmetric, but not transitive. If you try to find a function that corresponds to R or S, in the sense of Theorem 6.2, you won't be able to.

Example 11: Referring to Examples 4 and 5, note that "x and y have the same mother *and* father" defines an equivalence relation, whereas "x and y have the same mother *or* father" does not. We just saw how to define a function that corresponds to the former relation; there's no way to define one for the latter relation.

This is part of a general phenomenon. The intersection of two equivalence relations is always an equivalence relation, but their union is usually not (see Exercises 12 and 13). Here is another example of this.

Example 12: On the set of all people, let

$$R = \{(x, y) \mid x \text{ and } y \text{ have the same hair color}\}$$

$$S = \{(x, y) \mid x \text{ and } y \text{ have the same eye color}\}$$

Let's assume (perhaps unrealistically) that hair and eye color are defined precisely enough so that R and S are equivalence relations. Then $R \cap S = \{(x, y) \mid x \text{ and } y \text{ have the same hair and eye color}\}$, another equivalence relation. But $R \cup S$ is the same relation with "and" replaced by "or," which is not transitive.

We've been discussing the connection between equivalence relations and functions. There is an even more direct connection between equivalence relations and **partitions.**

Definition: Let A be a set. A **partition** of A is a collection of nonempty sets such that any two of them are disjoint, and the union of all of them is A. (See Exercise 21 for an alternative definition of partitions.)

Example 13: Let A = the set of all male people and B = the set of all female people. Then $\{A, B\}$ is a partition of the set of all people (assuming every person is male or female, exclusively).

Example 14: For each integer i, let A_i be the half-open interval $[i, i + 1)$. Then the collection of intervals $\{A_i \mid i \in \mathbb{Z}\}$ is a partition of $\mathbb{R}$. Note that this only works with half-open intervals. If we used closed intervals, they would not be disjoint from each other. If we instead used open intervals, the union of all the A_is would not be all of $\mathbb{R}$ (although these open intervals would form a partition of $\mathbb{R} - \mathbb{Z}$). Of course, we would still get a partition of $\mathbb{R}$ by letting $A_i = (i, i + 1]$.

Definitions: If R is an equivalence relation on a set A and $x \in A$, the **equivalence class** of x, denoted $[x]_R$, is the set $\{y \in A \mid xRy\}$. The collection of all these equivalence classes is called A **modulo** R, denoted A/R.

If there is no possibility of confusion (that is, if only one equivalence relation is being discussed), we just write $[x]$ instead of $[x]_R$.

Example 15: Let A be the set of all people, and let the equivalence relation R on A be defined by "x and y are the same age (in years)." If Lucian is a nine-year-old, then $[\text{Lucian}]_R$ is the set of all nine-year-olds. Any person's equivalence class (with respect to R) is the "club" consisting of all people of the same age as that person. Clearly, there is no overlap between different clubs, and each person is in exactly one of these clubs. The following lemma and theorem generalize these observations.

Lemma 6.5: Let R be an equivalence relation on a set A, and let $x, y \in A$. Then

(a) $x \in [x]$
(b) xRy iff $[x] = [y]$
(c) $\sim(xRy)$ iff $[x]$ and $[y]$ are disjoint.

Proof: We prove part (b) and leave the other parts for Exercise 11. For the forward direction, assume xRy. To show $[x] = [y]$, we need to show that, given any z, $z \in x$ iff $z \in y$, or equivalently, xRz iff yRz. *[Note how we must introduce the new variable z here.]* So now assume xRz. Since xRy we also have yRx, by symmetry. So yRz by transitivity. The converse is analogous. For the reverse direction, assume $[x] = [y]$. Then, since $y \in [y]$ by part (a), we have $y \in [x]$. So xRy by definition of equivalence classes. ■

We can now prove what is probably the most important single result about equivalence relations.

Theorem 6.6: For any equivalence relation, its equivalence classes form a partition of its domain.

Proof: Say R is an equivalence relation on A. By Lemma 6.5(a), each element of A is in some equivalence class. So each equivalence class is a nonempty subset of A, and the union of all the equivalence classes is A. And, by Lemma 6.5(b) and (c), any two distinct equivalence classes are disjoint. ■

Theorem 6.6 provides yet another way of understanding equivalence relations. An equivalence relation on a set partitions, or breaks up, the set into disjoint subsets (the equivalence classes). Each class is formed as a set of things that are alike in whatever sense corresponds to that equivalence relation. A converse to Theorem 6.6 holds, rather trivially: if $\mathscr{A}$ is any partition of a set B, then the relation $\{(x, y) \mid x \text{ and } y \text{ are in the same member of } \mathscr{A}\}$ is an equivalence relation on B (see Exercise 20).

Example 16: Consider the equality relation on any set A. For each x, $[x] = \{x\}$. So id_A partitions A into one-element sets. This is called the **finest** possible partition on A.

Example 17: For any set A, let $R = A \times A$. So xRy holds for all $x, y \in A$. It follows that R is an equivalence relation; the only equivalence class is A itself. This is called the **coarsest** possible partition on A.

Example 18: Consider the "siblinghood" relation of Example 4. Then for any person x, $[x]$ consists of x and all his or her siblings. If x is an only child, then $[x] = \{x\}$. So this relation partitions the set of all people who have ever lived into sibling classes.

Example 19: Referring to Example 8, consider congruence modulo 2 on the set $\mathbb{Z}$. Then if n is even, $[n]$ consists of all the even integers, whereas if n is odd, $[n]$ consists of all the odd integers. So $\mathbb{Z}$ is partitioned into two subsets by this relation.

Exercises 6.2

(1) Let A be the set of all people who have ever lived. For each of the following, state whether it's an equivalence relation on A and justify your assertion. If it is an equivalence relation, describe the equivalence classes and give an example of an equivalence class.

(a) x and y were born in the same year.
(b) x and y were born less than a week apart.
(c) x and y have the same maternal grandfather.
(d) x and y have the same four grandparents.
(e) x and y are first cousins or $x = y$.
(f) The set of all of x's children equals the set of all of y's children.

(2) For each of the following, state whether it's an equivalence relation on the specified set and justify your assertion.

(a) The relation $A \subseteq B$, on $\wp(\mathbb{N})$
(b) The relation m and n have the same digit in the 100's place, on $\mathbb{N}$
(c) The relation x and y differ by a rational number, on $\mathbb{R}$
(d) The relation A and B are not disjoint, on $\wp(\mathbb{N})$
(e) The relation A and B have the same smallest member, on $\wp(\mathbb{N}) - \{\varnothing\}$
(f) The relation A and B have the same smallest member, on $\wp(\mathbb{R}) - \{\varnothing\}$

(3) For each of the following relations on $\mathbb{R}$, state whether it's an equivalence relation (on whatever its domain is) and justify your assertion. If it is, describe the equivalence classes. In particular, describe $[3]$ and $[\pi]$, provided that these numbers are in the domain of the relation.

(a) $x^2 - 5 = y^2 - 5$
(b) $\sin x = \sin y$
(c) $\tan x = \tan y$
(d) $x + y$ is an integer.
(e) $|x| - |y|$ is an integer.
(f) $x - y$ is irrational.
(g) x/y is an integer.
*(h) $x/y = 2^i$, for some $i \in \mathbb{Z}$
*(i) $x - y = a + b\pi$, for some $a, b \in \mathbb{Q}$.

(4) Give an example of a relation R on $\mathbb{N}$ satisfying each of the following or explain why it is not possible to have one.

(a) R is reflexive on $\mathbb{N}$ and symmetric but not transitive.
(b) R is reflexive on $\mathbb{N}$ and transitive but not symmetric.

(c) R is symmetric and transitive but not reflexive on $\mathbb{N}$.
(d) R is reflexive on $\mathbb{N}$ but neither symmetric nor transitive.
(e) R is symmetric but not reflexive on $\mathbb{N}$ or transitive.
(f) R is transitive but not reflexive on $\mathbb{N}$ or symmetric.
(g) R is not reflexive on $\mathbb{N}$, symmetric, or transitive.

(5) Let R be the relation $\{(1, 1), (2, 2), (3, 3)\}$.
(a) Is R reflexive?
(b) Is R an equivalence relation?
(c) Is R a relation on $\mathbb{N}$?
(d) Is R reflexive on $\mathbb{N}$?
(e) Is R an equivalence relation on $\mathbb{N}$?
(f) Explain why your answers to parts (a) through (e) are not contradictory.

(6) Are the following statements true or false? If true, explain why; if not, find a counterexample.
(a) Whenever R is reflexive on A, then R is a relation on A.
(b) Whenever R is a relation on A and R is reflexive, then R is reflexive on A.
(c) An equivalence relation on A is precisely an equivalence relation whose domain is A.

(7) Show that the relation defined in Example 2 is an equivalence relation.

(8) Let $A = \mathbb{Z} \times (\mathbb{Z} - \{0\})$. In other words, A is the set of all ordered pairs of integers in which the second integer is nonzero. Define a relation $\sim$ on A by

$$(a, b) \sim (c, d) \text{ iff } ad = bc$$

Prove that $\sim$ is an equivalence relation on A. (The idea behind this is that the ordered pair (a, b) may be used to represent the fraction a/b, in which case $\sim$ becomes the standard cross-multiplication condition for the equality of fractions. This equivalence relation is important in the construction of the rationals from the integers; see Sections 8.3 and 9.5.)

(9) For any $A, B \subseteq \mathbb{R}$, A is said to be a **translate** of B iff there is a real number k such that $B = \{x + k \mid x \in A\}$.
(a) Prove that this relation is an equivalence relation on $\mathcal{P}(\mathbb{R})$.
(b) Describe the equivalence class of the set of negative real numbers.
(c) Prove that every translate of a closed interval is also a closed interval.
(d) Find two other phrases that could replace the words "closed interval" in part (b) and still yield a true statement. Prove these statements.
(e) Find two other phrases that could replace the words "closed interval" in part (b) and yield a *false* statement. Give counterexamples to verify that the statements are false.

(10) (a) Prove Theorem 6.3(a).
(b) Prove Theorem 6.3(b).
(c) Complete the proof of Theorem 6.3(c).

(11) Prove parts (a) and (c) of Lemma 6.5.

(12) (a) Prove that the intersection of any two equivalence relations on A is an equivalence relation on A.
*(b) More generally, prove that the intersection of *any collection* of equivalence relations on A is an equivalence relation on A.

(13) Give two examples to show that the union of two equivalence relations is not necessarily an equivalence relation.

(14) This exercise continues ideas introduced in Examples 16 and 17. If $\mathscr{B}$ and $\mathscr{C}$ are two partitions of a set A, we say $\mathscr{B}$ is a **refinement** of $\mathscr{C}$ or a **finer partition** than $\mathscr{C}$ iff every set in $\mathscr{B}$ is a subset of some set in $\mathscr{C}$. Prove that if R and S are equivalence relations on A, then the partition created by R is a refinement of the partition created by S if and only if $\forall x, y\ (xRy \rightarrow xSy)$.

(15) Refer to Example 8 and Exercise 14. Let A be any one of the sets $\mathbb{R}$, $\mathbb{Q}$, or $\mathbb{Z}$. If m and n are natural numbers, under what condition is the partition created by congruence modulo m a refinement of the partition created by congruence modulo n? Prove your assertion.

*(16) This exercise generalizes the ideas introduced in Example 8 and Exercise 2(c). Let A be as in Exercise 15. Let B be a nonempty subset of A that is closed under subtraction; that is, $\forall x, y \in B\ (x - y \in B)$. Such a set B is called a **subgroup of A under addition**. Now define a relation R on A by xRy iff $x - y \in B$. Prove that R is an equivalence relation on A, and describe the equivalence classes. (R is called **congruence modulo B**.)

(17) How many equivalence classes are there under the congruence modulo 3 relation on $\mathbb{Z}$? Describe them.

(18) Let $R = \{(m, n) \mid m, n \in \mathbb{Z} \text{ and } 3m + 4n \text{ is a multiple of } 7\}$.
(a) Is R an equivalence relation on $\mathbb{Z}$? Prove your claim.
*(b) R is a relation of the type discussed in Example 8. By experimentation, determine exactly which relation R is. You needn't prove your conclusion.

(19) (a) Find all equivalence relations on the set $\{1, 2, 3\}$.
*(b) Find all equivalence relations on the set $\{1, 2, 3, 4\}$.

(20) Prove the converse of Theorem 6.6 mentioned in the text.

(21) Suppose A is a set and $\mathscr{B}$ is a collection of sets. Prove that $\mathscr{B}$ is a partition of A iff $\mathscr{B}$ is a collection of nonempty subsets of A and every member of A is in exactly one member of $\mathscr{B}$.

*6.3 Ordering Relations

Ordering relations are just about as important in mathematics as equivalence relations. They are also a good deal more familiar to most students than equivalence relations, and so you will probably find this section conceptually simpler than the previous one. On the other hand, there are more details and minor variations involved with ordering relations than with equivalence relations, so you have to be careful to keep things straight.

There are two general types of orderings. This section is primarily devoted to reflexive orderings. At the end of the section we briefly discuss irreflexive orderings and show the simple, close connection between these two types of orderings. Irreflexive orderings are also used in Appendices 1 and 2.

Definitions: A relation R is **antisymmetric** iff whenever xRy and yRx, then $x = y$.

A **partial ordering** is a relation that is reflexive, antisymmetric, and transitive.

A **total ordering** is a partial ordering R in which xRy or yRx holds for every x and y in the domain of R.

As with equivalence relations, when we refer to an ordering (partial or total) on a set A, we mean an ordering whose domain is A.

Note that the only difference between the definitions of equivalence relations and partial orderings is symmetry versus antisymmetry. But this is a crucial difference. Whereas an equivalence relation expresses some sort of alikeness of elements and groups them together (in equivalence classes), an ordering sets elements apart by putting them in a hierarchy, or order.

If R is an ordering on A, we also say that A is **partially** (or **totally**) **ordered** by R. Total orderings may also be called **linear orderings** or **simple orderings**.

Notation: If R is an ordering, we may use the more common notation $x \leq y$ or $x \geq y$ to mean xRy. When this is done, standard abbreviations are automatically assumed: $x \geq y$ means the same as $y \leq x$, and $x < y$ means $x \leq y$ and $x \neq y$ (as does $y > x$).

Example 1: The set of real numbers $\mathbb{R}$ has a standard ordering on it. Both $\leq$ and $\geq$ are total orderings on $\mathbb{R}$ (and therefore also on any subset of $\mathbb{R}$, such as $\mathbb{N}$, $\mathbb{Z}$, or $\mathbb{Q}$). The relations $<$ and $>$ on $\mathbb{R}$ are not total orderings in the sense defined here, since they are not reflexive. They are *irreflexive* total orderings.

Example 2: It's instructive to consider some orderings on small sets. Let $A = \{1, 2, 3\}$. As a subset of $\mathbb{R}$, A inherits a total ordering, as mentioned in the previous example. If we use $\leq$ as the basis of this relation, what ordered pairs would it contain?

It can't contain just (1, 2) and (2, 3), because this set of ordered pairs is neither reflexive nor transitive. Exercises 1 and 9 ask you to answer this question and some similar ones. There are many ways to totally order this set or almost any set.

Definition: If R is a partial ordering, x and y are called **comparable** (with respect to R) iff xRy or yRx holds. If x and y are in Dom(R) but are not comparable, then they are called **incomparable.**

So a partial ordering is total iff it has no incomparable pairs of elements.

Example 3: Many important orderings are not total. Let $\mathscr{A}$ be any collection of sets. It is simple to show that the relation $A \subseteq B$ is a partial ordering on $\mathscr{A}$. In fact, all three of the conditions for this are proved in Chapter 5. This relation is often considered with $\mathscr{A} = \wp(C)$, for some set C. Choose some small set for C. Is the relation $\subseteq$ on $\mathscr{A}$ total? (See Exercise 5.)

Example 4: Let $C = \mathbb{R} \times \mathbb{R}$. We can use the total ordering on $\mathbb{R}$ to define a *partial* ordering on C, as follows: let's say that one ordered pair is related to another if the first pair $\leq$ the second in *both* coordinates. That is, define the relation S on C by

$$(a, b)\, S\, (c, d) \quad \text{iff} \quad a \leq c \ \text{ and } \ b \leq d$$

It's not hard to show that S is a partial ordering on C. It's also clear that S is not total. For example, consider (3, 7) and (9, 4). You can see that these ordered pairs are incomparable.

This example is a specific case of the following definition.

Definition: Let R and S be partial orderings on sets A and B, respectively. Then the **product ordering** $R \times S$ is the relation on $A \times B$ defined by

$$R \times S = \{((a, b), (a', b')) \mid aRa' \text{ and } bSb'\}$$

This definition may seem confusing at first because it involves ordered pairs of ordered pairs. But if you understand Example 4, you should see that this definition just formalizes the idea of that example. By the way, this notation is somewhat sloppy, in that $R \times S$ is not literally the cartesian product of R and S.

Theorem 6.7: Let R and S be partial orderings on sets A and B, respectively. Then the product ordering $R \times S$ is a partial ordering on $A \times B$. However, if both A and B have more than one element, then $R \times S$ is not total.

Proof: The proof that $R \times S$ is a partial ordering is left for Exercise 3(a). To prove the second claim, assume that a and a' are distinct elements of A, and b and b' are distinct elements of B. By antisymmetry, we can't have both aRa' and $a'Ra$, so without loss of generality let's say ~ (aRa'). (So either $a'Ra$, or a and a' are incomparable—it doesn't matter.) Similarly, without loss of generality let's say ~ (bRb').

Now consider the ordered pairs (a, b') and (a', b). It's a simple matter to show that these are incomparable in the ordering $R \times S$ (see Exercise 3(b)). ■

Example 5: Here's another example of product orderings. Suppose we want to rank mixed-doubles tennis teams (that is, teams of one male and one female player). Let M be the set of all male tennis players and F the set of all female ones. On each of these sets, we have the ordering defined by one player's being at least as good as another. For simplicity, let's assume these orderings are total; this means that for any two players, it's possible to say which one is better. (This assumption is somewhat unrealistic in tennis and most other sports, but let's not worry about that.)

We can then define the product ordering on $M \times F$, the set of all possible mixed-doubles teams. Under the product ordering, one team is at least as good as another iff its male and female players are both at least as good as the corresponding players of the other team. But this is not a total ordering. For example, since Pete Sampras is at least as good as Bob Wolf and Martina Hingis is at least as good as Roseanne Barr, the product ordering would say the team of Sampras and Hingis is at least as good as the team of Wolf and Barr. But the product ordering would not settle which is better, Sampras-Barr or Wolf-Hingis. These two teams would be incomparable under the product ordering.

Product orderings have some practical value. But, as you can imagine, they're not likely to be completely accurate in a situation like the ranking of sports teams.

Orderings (especially nontotal ones) on small sets can often be clearly shown using **lattice diagrams**. In a lattice diagram for a relation R, each point of the domain is represented as a dot, and the fact that aRb is represented by an upward (but not necessarily vertical) path from a to b. Such a path from a to b may pass through other points; transitivity requires us to interpret things this way. A lattice diagram for a total ordering is a single line, or chain, so it isn't very interesting. Figure 6.4 shows some lattice diagrams, and Exercise 4 asks you to construct several others.

Example 6: This example is not of a product ordering, but it illustrates a related idea. Let A be the set of all people. For simplicity, let's assume that it's possible to accurately measure everyone's height and weight and that no two people have exactly the same height or weight. Then we get two different *total* orderings on A, defined by height and weight, respectively. That is, let

$$R = \{(a, b) \mid a, b \in A \text{ and } a \text{ is at least as tall as } b\}$$

$$S = \{(a, b) \mid a, b \in A \text{ and } a \text{ is at least as heavy as } b\}$$

A simple way to use these two orderings to form a new one is to take their intersection $R \cap S$ as sets of ordered pairs. Since intersection is defined by the word "and," it's easy to understand $R \cap S$: it consists of all ordered pairs (a, b) such that a is taller *and* heavier

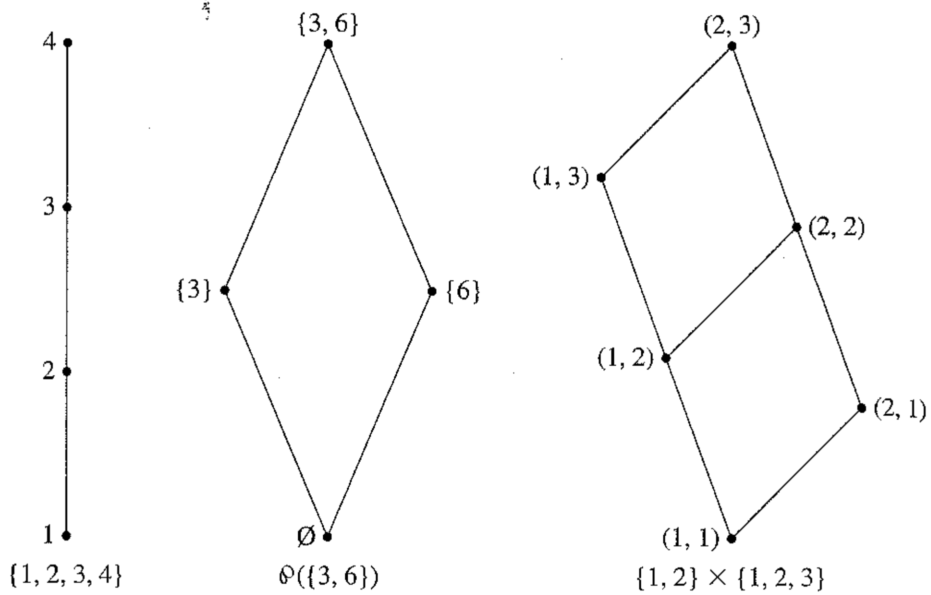


Figure 6.4 Lattice diagrams for three partial orderings

than b. This new relation must still be a partial ordering (see Exercise 6), but it's not total, since one person could be shorter and heavier than another. Two such people would be incomparable in the intersection ordering.

What makes intersections of orderings similar to product orderings? The simplest way to answer that is to say that both are defined with the word "and." This simple logical similarity is the reason why both products and intersections of partial orderings must be partial orderings, but they are rarely total.

Example 7: In the previous example, what would happen if we looked at $R \cup S$ instead of $R \cap S$? This union would consist of all ordered pairs (a, b) of people such that a is at least as tall *or* at least as heavy as b. This is not even a partial ordering (see Exercise 7). Compare this to Exercises 12 and 13 of Section 6.2.

Example 8: Let's reconsider Example 5. Recall that the product ordering on mixed-doubles teams is not total. Is there a way to set up a *total* ordering on $M \times F$, defined directly in terms of the separate orderings on M and F?

There are actually several ways to do this. Perhaps the simplest way is to give one sex or the other precedence for ranking teams. For example, it could be decided to give precedence to male players; this would mean that if one team had a better male player than another, that team would be automatically ranked higher than the other, no matter how the two female players stacked up. Only if two teams had the same male player (or, in a more practical setting, male players of equal ability) would the relative abilities of the teams' female players be taken into account.

Like some of our previous examples, this one is not completely realistic. If one team has a slightly better male player than another but a vastly inferior female player, it is unlikely that the first team should be considered the better one. But the way we are defining this ordering of teams, it would be ranked as the better one.

Let's now formalize this example and give it a name.

Definition: Let R and S be partial orderings on sets A and B, respectively. Then the **lexicographic ordering** on $A \times B$ (associated with R and S) is

$$\{((a, b), (a', b')) \mid (aRa' \text{ and } a \neq a') \text{ or } (a = a' \text{and } bSb')\}$$

That is, the first coordinates take precedence for ordering ordered pairs. Only in the case that the first coordinates are equal are the second coordinates used.

It would also be possible to define a type of lexicographic ordering that favors the second coordinate instead of the first, and this would give a very different ordering on $A \times B$ from the one defined above (see Exercise 10(c)). In contrast, there's only one way to define a product ordering $R \times S$, since that notion does not involve choosing which coordinate to give precedence.

Theorem 6.8: (a) If R and S are partial orderings on sets A and B, then the associated lexicographic ordering is a partial ordering on $A \times B$.

(b) If, in addition, R and S are total, then the associated lexicographic ordering is also total.

Proof: (a) Let L be the lexicographic ordering on $A \times B$. We must show L is reflexive on $A \times B$, antisymmetric, and transitive. To prove reflexivity, let $(a, b) \in A \times B$. We know that $a = a$, and bSb because S is reflexive. Therefore, $((a, b), (a, b))$ is in L, because it satisfies the second disjunct in the definition of lexicographic order.

To prove antisymmetry, assume $(a, b) L (a', b')$ and $(a', b') L (a, b)$. Both these assumptions are disjunctions, so there are four possible ways to make them both true. Three of these directly say that $a = a'$, and the fourth requires that aRa', $a \neq a'$, and $a'Ra$, which violates the antisymmetry of R. So $a = a'$. But in that case, the two assumptions require, respectively, that bSb' and $b'Sb$. But then $b = b'$, by the antisymmetry of S. So we have shown that $(a, b) = (a', b')$, as desired.

Exercise 11 asks you to prove the transitivity of L, as well as part (b). ■

Preorderings

We have seen several examples of orderings that are not total because they have incomparable pairs of elements. The next two examples illustrate another phenomenon that prevents some relations from being even partial orderings.

Example 9: For integers m and n, we say m **divides** n, denoted $m|n$, iff for some integer k, $n = km$. It is not difficult to show that $|$ is a partial ordering on $\mathbb{N}$ (see Exercise 12). (Specifically, this means that m and n are restricted to be positive; it doesn't matter whether k is similarly restricted.) This ordering is certainly not total; for instance, 2 and 3 are incomparable.

Does $|$ define a partial ordering on $\mathbb{Z}$? Perhaps surprisingly, it does not, because it's not antisymmetric. For example, $5|(-5)$ and $(-5)|5$, but $5 \neq -5$. Note that this is very different from saying that 5 and -5 are incomparable. Relations like this fall into a different category called preorderings. The relation $|$ is discussed further in Section 8.2.

Definition: A **preordering** is a relation that is reflexive and transitive.

Obviously, every partial ordering is a preordering. Example 9 shows that the converse of this does not hold. However, every preordering naturally gives rise to both an equivalence relation and a partial ordering (see Exercise 13). Here is a simple relation from real life that can be viewed as a preordering.

Example 10: As in Example 6, let R be the height relation on the set of all people. Again, if we assume that no two people are exactly the same height, then R is a total ordering. But what if there are people of the same height? (Or we could guarantee this by letting xRy mean that y is at least as tall as x, with height measured to the nearest centimeter.) If x and y are two people of the same height, then xRy and yRx, so R is not antisymmetric. It would make sense to say that such people are tied in this relation. Note that all people are comparable in height, so R may be called a **total preordering**. In contrast, the relation $|$ on $\mathbb{Z}$ has both incomparable elements (such as 2 and 3) and tied elements (such as 5 and -5).

Irreflexive Orderings

We conclude this section by showing that the difference between reflexive orderings and irreflexive orderings is not very profound. The definition and some simple results involving irreflexive orderings appear in Appendices 1 and 2, because the order axioms of $\mathbb{R}$ are usually stated in terms of $<$ rather than $\leq$.

Recall that $\mathrm{id}_A = \{(x, x) \mid x \in A\}$.

Definitions: A relation R is called **irreflexive** iff xRx is *not* true for any x.

A relation that is irreflexive and transitive is called an **irreflexive partial ordering**.

An **irreflexive total ordering** is an irreflexive partial ordering that also satisfies **trichotomy:** for every x and y in the domain of R, xRy, yRx, or $x = y$.

Note that "irreflexive" means more than merely "not reflexive," just as "antisymmetric" means more than merely "not symmetric." Also note that, in the context of irreflexivity, we must add the disjunct $x = y$ to the condition for being total.

Antisymmetry does not need to be included in the definition of an irreflexive partial ordering because it follows from the other two conditions. Furthermore, if R is irreflexive, then xRy, yRx, and $x = y$ cannot all be true. Therefore, an irreflexive partial ordering must satisfy **strong antisymmetry**: (xRy and yRx) is always false. This conclusion is proved as Theorem A-8 of Appendix 2.

The standard symbols < and > are typically used to denote irreflexive orderings. If < is an irreflexive ordering, it would seem sensible that we could obtain a reflexive ordering $\leq$ by defining $x \leq y$ to mean $x < y$ or $x = y$, as usual. Conversely, if $\leq$ is a reflexive ordering, it would seem sensible that we could obtain an irreflexive ordering < by defining $x < y$ to mean $x \leq y$ and $x \neq y$. This is indeed the simple link between the two types of orderings and is stated more precisely in the following result.

Theorem 6.9: (a) Let R be an irreflexive partial ordering on a subset of A. Then $R \cup \mathrm{id}_A$ is a reflexive partial ordering on A. The ordering R is total on A iff $R \cup \mathrm{id}_A$ is.

(b) Let S be a reflexive partial ordering on a set A. Then $S - \mathrm{id}_A$ is an irreflexive partial ordering on a subset of A. The ordering S is total on A iff $S - \mathrm{id}_A$ is.

Proof: (a) Assume R is as described, and let $S = R \cup \mathrm{id}_A$. Since the domain and range of R are subsets of A, the domain and range of S are both equal to A. Since id_A is reflexive on A, so is S. To show that S is transitive, assume aSb and bSc. Then either aRb or $a = b$; similarly, either bRc or $b = c$. We get a total of four possible cases, all of which imply that aSc. We also need to show that S is antisymmetric. So assume aSb and bSa. We need to show $a = b$. But $a \neq b$ would imply aRb and bRa, which would contradict the strong antisymmetry of R.

To prove the second statement, note that

$$\begin{aligned} R \text{ is total on } A \ &\text{iff } (xRy \text{ or } yRx \text{ or } x = y), \text{ for all } x, y \text{ in } A \\ &\text{iff } [(xRy \text{ or } x = y) \text{ or } (yRx \text{ or } y = x)], \text{ for all } x, y \text{ in } A \\ &\text{iff } (xSy \text{ or } ySx), \text{ for all } x, y \text{ in } A \\ &\text{iff } S \text{ is total on A} \end{aligned}$$

The proof of part (b) is similar and is left for Exercise 14. ■

Exercises 6.3

(1) (a) As discussed in Example 2, list all the ordered pairs in the ordering $\leq$ on the set $\{1, 2, 3\}$. How many ordered pairs are in this relation?

(b) Repeat part (a) for the sets $\{1\}$, $\{1, 2\}$ and $\{1, 2, 3, 4\}$.

(2) Draw the cartesian plane $\mathbb{R} \times \mathbb{R}$, and choose an arbitrary point (a, b). Now consider the product ordering on $\mathbb{R} \times \mathbb{R}$ discussed in Example 4.

(a) Relative to (a, b), where are the points that are greater than it?

(b) Relative to (a, b), where are the points that are less than it?

(c) Relative to (a, b), where are the points that are incomparable to it?

(d) Repeat parts (a) and (b) for the lexicographic ordering on $\mathbb{R} \times \mathbb{R}$.

(3) (a) Prove the first part of Theorem 6.7.
(b) Complete the proof of the second part of Theorem 6.7.

(4) (a) Draw a lattice diagram for the product ordering on {1, 2, 3} × {1, 2, 3}.
(b) Draw a lattice diagram for the ordering ⊆ on ℘({1, 3, 8}). (This can not be done without having paths cross each other.)
(c) Pick at least a half dozen people you know, and for this set of people, draw a lattice diagram for the ordering of Example 6.

(5) Find and prove a necessary and sufficient condition on a set A for the ordering ⊆ on ℘(A), discussed in Example 3, to be total.

(6) Prove that the intersection of two partial orderings on A is again a partial ordering on A, as stated in Example 6.

(7) Show that the relation R ∪ S defined in Example 7 is not a partial ordering.

(8) Prove that R is a partial (respectively, total) ordering on A iff R^{-1} is.

(9) (a) On the basis of Exercise 1, make a conjecture of the form "The number of ordered pairs in a total ordering on an $(n + 1)$ element set is ____ more than the number of ordered pairs in a total ordering on an n element set."
(b) By induction (and not too formally), prove your conjecture from part (a).
*(c) Find and prove a formula for the number of ordered pairs in any total ordering on a set with n elements.

(10) (a) Look up "lexicographic" in a dictionary, and then explain our use of the term "lexicographic ordering."
(b) Draw a graph showing the elements of ℕ × ℕ, as a subset of ℝ × ℝ. Then give a simple pictorial description of the lexicographic ordering on ℕ × ℕ that is based on the standard ordering on ℕ.
(c) Give a precise definition of the lexicographic ordering on $A \times B$, associated with R and S, in which the second coordinate takes precedence.
(d) Give a precise definition of the lexicographic ordering on $A \times B \times C$, associated with partial orderings R, S, and T (on A, B, and C, respectively).
*(e) A language such as English does not consist of just two-letter words or three-letter words. The ordering of words in a dictionary must allow words of any length. In analogy to this, give an informal but clear definition of the lexicographic ordering on the set of *all finite sequences* of elements of a set A, based on an ordering R on A.

(11) (a) Complete the proof of Theorem 6.9(a).
(b) Prove Theorem 6.9(b).

(12) Prove that the relation $|$, defined in Example 9, defines a partial ordering on $\mathbb{N}$.

(13) (a) Prove that if R is a preordering on A, then $R \cap R^{-1}$ is an equivalence relation on A.

(b) Applying part (a) to Example 10, describe the equivalence classes.

(c) A corollary to part (a) states that the original preordering R defines a "natural" partial ordering whose domain is the set of *equivalence classes* of $R \cap R^{-1}$. Explain this in the context of Example 10.

(14) Prove Theorem 6.9(b).

(15) Let A be any set with one element, perhaps $\{5\}$.

(a) How many reflexive orderings are there on A? Describe all of them.

(b) Are there any irreflexive orderings whose domain is A?

(c) Does this situation contradict Theorem 6.9? Explain.

The next three exercises concern topics introduced after Theorem 5.6.

(16) Say we have a partially ordered set A. On the basis of the last part of this section, it doesn't matter whether this ordering is reflexive or irreflexive. For any $x \in A$, we say x is:

Minimal iff $\sim \exists y \in A\ (y < x)$

Maximal iff $\sim \exists y \in A\ (y > x)$

A **least element** iff $\forall y \in A\ (x \leq y)$

A **greatest element** iff $\forall y \in A\ (x \geq y)$

Prove: (a) There cannot be more than one least element or more than one greatest element. (Hence, one usually refers to *the* least and *the* greatest element, if they exist.)

(b) A least (respectively, greatest) element must be minimal (respectively, maximal).

(c) There can be more than one minimal (respectively, maximal) element. Therefore, the converses of part (b) fail.

(17) A partial ordering on A is called a **well-ordering** on A iff every nonempty subset of A has a least element. A set with a well-ordering defined on it is said to be **well ordered**. Prove:

(a) A well-ordering on A must be total.

(b) Every subset of a well-ordered set is well ordered.

(c) Every subset of $\mathbb{N}$ is well ordered by $\leq$. (Recall Theorem 5.6.)

(d) The sets $\mathbb{Z}$, $\mathbb{Q}$, and $\{x \in \mathbb{R} \mid x \geq 0\}$ are *not* well ordered by $\leq$.

(18) By induction on n, prove that every subset of $\mathbb{R}$ with n members is well ordered. (You may begin this proof with either $n = 1$ or $n = 0$, as you wish. Essentially, what you are proving is that every *finite* subset of $\mathbb{R}$ is well ordered.)

(19) Prove or find a counterexample: if R is a partial ordering on A and S is a partial ordering on B, then $R \cup S$ is a partial ordering on $A \cup B$.

Critique the proofs in the remaining exercises in this section. (If necessary, refer to Exercises 4.2 for the instructions for this type of problem.)

(20) **Theorem:** If R is a partial ordering on A and S is a partial ordering on B, where A and B are *disjoint* sets, then $R \cup S$ is a partial ordering on $A \cup B$.

Proof: First, assume $x \in A \cup B$. If $x \in A$, then xRx. If $x \in B$, then xSx. In either case, $(x, x) \in R \cup S$, so $R \cup S$ is reflexive. To prove antisymmetry, assume (x, y) and (y, x) are both in $R \cup S$. If they are both in R, then $x = y$, by the antisymmetry of R. If they are both in S, we similarly have $x = y$. And it's not possible that $(x, y) \in R$ and $(y, x) \in S$, because A and B are disjoint. To prove transitivity, assume (x, y) and (y, z) are both in $R \cup S$. Again using the disjointness of A and B, the only way this can occur is that (xRy and yRz) or (xSy and ySz). This implies xRz or xSz, as desired.

(21) **Theorem:** If R is a total ordering on A and S is a total ordering on B, where A and B are *disjoint* sets, then $R \cup S$ is a total ordering on $A \cup B$.

Proof: Assume the givens. The proof that $R \cup S$ is a partial ordering on $A \cup B$ is as in the previous problem. And since R is total on A and S is total on B, $R \cup S$ is total on $A \cup B$.

Suggestions for Further Reading: The material in this chapter is covered in most books on set theory, such as the first four references given at the end of Chapter 5. See also Feferman (1989), Pfleeger and Straight (1985), or Ross and Wright (1985).