Chapter 5

Relations and Functions

Having developed some of the essentials about how proofs are planned and written, we are ready to begin exploring, from a more sophisticated perspective, certain fundamental mathematical ideas with which you are already familiar. In this chapter, we will broaden our understanding of the core concepts of *relation* and *function* to suit a more general mathematical outlook.

5.1 Relations

Questions to guide your reading of this section:

1. How is the notion of *ordered pair* defined? When are two ordered pairs considered to be equal?
2. What is meant by the *Cartesian product* $X \times Y$ of sets X and Y?
3. What is meant by a *relation*? a *relation on* $X \times Y$? a *relation on* X?
4. What is meant by the *domain* of a relation? the *range*?
5. What is an *interval* of real numbers? How is *interval notation* used to describe an interval?

Given a and b, the **ordered pair** (a, b) is formally defined as the set $\{\{a\}, \{a, b\}\}$ whose members are the sets $\{a\}$ and $\{a, b\}$. For the ordered pair (a, b), a is referred to as the **first** or **left coordinate**, and b is referred to as the **second** or **right coordinate**.

Example 5.1 The ordered pair $(2, 5)$ is, by definition, the set $\{\{2\}, \{2, 5\}\}$. The first coordinate of $(2, 5)$ is 2 and the second coordinate is 5. In contrast, the ordered pair $(5, 2)$ is the set $\{\{5\}, \{5, 2\}\}$ and has first coordinate 5 and second coordinate 2. Observe that $(2, 5) \neq (5, 2)$ because, as sets, the ordered pair $(2, 5)$ includes a member, namely the set $\{2\}$, that is not a member of the ordered pair (5, 2). Thus, the order in which coordinates are listed in an ordered pair really does make a difference. We also see how the ordered pair (2, 5) is different from the set {2, 5}, because the order in which elements are listed in a set is immaterial. For instance, whereas $(2, 5) \neq (5, 2)$, it is true that $\{2, 5\} = \{5, 2\}$.

> **Practice Problem 1** What ordered pair is represented by the set $\{\{0, -3\}, \{-3\}\}$?

Ordered pairs can be used to organize information. For instance, to convey the fact that a triangle has three sides and a rectangle has four sides, we might form the ordered pairs (*triangle*, 3) and (*rectangle*, 4). We also point out that the first coordinates of these ordered pairs are *not* numbers (there is no requirement that the coordinates be numbers).

The following theorem confirms our expectation that ordered pairs are the same exactly when their corresponding coordinates are identical. The proof relies heavily on the notion that sets are equal only when they have exactly the same members, but you will have to keep in mind that, formally, the members of an ordered pair are themselves sets.

Theorem 5.2 (*Equality of Ordered Pairs*) Two ordered pairs (x_1, y_1) and (x_2, y_2) are equal, that is, $(x_1, y_1) = (x_2, y_2)$, if and only if both $x_1 = x_2$ and $y_1 = y_2$.

Proof: (⇒) We will use contraposition. Suppose $x_1 \neq x_2$ or $y_1 \neq y_2$.

Case 1: $x_1 \neq x_2$.
Then $\{x_1\} \neq \{x_2\}$ and, because $x_2 \notin \{x_1\}$, $\{x_1\} \neq \{x_2, y_2\}$. Hence, $\{x_1\} \notin \{\{x_2\}, \{x_2, y_2\}\}$. Of course, $\{x_1\} \in \{\{x_1\}, \{x_1, y_1\}\}$. Therefore, it follows that $\{\{x_1\}, \{x_1, y_1\}\} \neq \{\{x_2\}, \{x_2, y_2\}\}$, which means $(x_1, y_1) \neq (x_2, y_2)$.

Case 2: $y_1 \neq y_2$.
We may assume that $x_1 = x_2$ (otherwise, *Case 1* applies). Let $x = x_1 = x_2$. Because $y_1 \neq y_2$, at least one of y_1 or y_2 is different from x. Without loss of generality, assume $y_1 \neq x$. Then, because $y_1 \notin \{x\}$, it follows that $\{x, y_1\} \neq \{x\}$. Also, because $y_1 \notin \{x, y_2\}$, it follows that $\{x, y_1\} \neq \{x, y_2\}$. Therefore, $\{x, y_1\} \notin \{\{x\}, \{x, y_2\}\}$. But, because $\{x, y_1\} \in \{\{x\}, \{x, y_1\}\}$, we may conclude that $\{\{x\}, \{x, y_1\}\} \neq \{\{x\}, \{x, y_2\}\}$, which means $(x_1, y_1) \neq (x_2, y_2)$.

(⇐) Suppose $x_1 = x_2$ and $y_1 = y_2$. It follows that

$$(x_1, y_1) = \{\{x_1\}, \{x_1, y_1\}\} = \{\{x_2\}, \{x_2, y_2\}\} = (x_2, y_2). \quad \square$$

Note the use of the phrase *without loss of generality* within the proof of Case 2 of our argument for (⇒). This phrase indicates that although there are really two possibilities to consider, $y_1 \neq x$ and $y_2 \neq x$, we understand that the argument for the second possibility $y_2 \neq x$ can be done by simply interchanging y_1 and y_2 in the argument we produced for the first possibility $y_1 \neq x$. Generally, when two or more possibilities must be considered at a certain point in a proof, but the argument supporting each possibility can be obtained by just interchanging names of objects, mathematicians typically incorporate the phrase *without loss of generality* and present the argument for only one of the possibilities.

Having established Theorem 5.2, we will rarely need to work with the formal definition of ordered pair. For example, although technically the ordered pair (2, 5) is defined to be the set $\{\{2\}, \{2, 5\}\}$, it will usually be sufficient for us simply to think of (2, 5) as the ordered pair whose first coordinate is 2 and whose second coordinate is 5.

Cartesian Products

The *xy*-coordinate system we use to plot ordered pairs of real numbers and graphs of equations is also called the *Cartesian coordinate system* after Rene Descartes, the mathematician and philosopher who invented it. More generally, the **Cartesian product** $X \times Y$ of sets X and Y is the set of all ordered pairs whose first coordinates are members of X and whose second coordinates are members of Y; that is, $X \times Y = \{(x, y) \mid x \in X, y \in Y\}$. Thus, the Cartesian product $\mathbf{R} \times \mathbf{R}$ is the set of all ordered pairs of real numbers; geometrically, we view this set as the familiar *xy*-plane.

Example 5.3 $\mathbf{R}^+ \times \mathbf{R}^-$ is the set of all ordered pairs whose first coordinates are positive real numbers and whose second coordinates are negative real numbers. Thus, $(2, -0.5) \in \mathbf{R}^+ \times \mathbf{R}^-$, but $(-2, 0.5) \notin \mathbf{R}^+ \times \mathbf{R}^-$ and $(2, 0) \notin \mathbf{R}^+ \times \mathbf{R}^-$.

Example 5.4 If $X = \{0, 1, 2\}$ and $Y = \{1, 0.5\}$, then

$$X \times Y = \{(0, 1), (0, 0.5), (1, 1), (1, 0.5), (2, 1), (2, 0.5)\},$$
$$Y \times X = \{(1, 0), (0.5, 0), (1, 1), (0.5, 1), (1, 2), (0.5, 2)\},$$
$$X \times X = \{(0, 0), (0, 1), (0, 2), (1, 0), (1, 1), (1, 2), (2, 0), (2, 1), (2, 2)\}, \text{ and}$$
$$Y \times Y = \{(1, 1), (1, 0.5), (0.5, 1), (0.5, 0.5)\}.$$

Example 5.5 Let S be the set of students majoring in mathematics at a certain college, and let P be the set of mathematics professors at this college. Then, $S \times P$ is the set of all ordered pairs for which the first coordinate is a student majoring in mathematics at this college, and the second coordinate is a mathematics professor at this college.

> **Practice Problem 2** If $A = \{\pi, -\pi, 0\}$ and $B = \{x \in \mathbf{N} \mid x < 3\}$, describe $A \times B$ and $B \times B$ by listing.

Relations

In mathematics, any set of ordered pairs is called a **relation**. This is because the *relationship* between two variables, say x and y, can be expressed by forming the set of all ordered pairs in which a value of the variable x occupies the first coordinate position and a corresponding value of the variable y occupies the second coordinate position.

Example 5.6 The following table gives the record high temperatures in degrees Fahrenheit recorded in the six New England states according to data collected by the National Climatic Data Center and reported by the U.S. Department of Commerce.

state	*ME*	*NH*	*VT*	*MA*	*CT*	*RI*
record high temperature	105	106	105	107	106	104

The relationship between the variable *state* and the variable *record high temperature* is expressed by the table, but it can easily be converted to a set of ordered pairs in which each New England state is paired with its record high temperature:

$$\{(ME, 105), (NH, 106), (VT, 105), (MA, 107), (CT, 106), (RI, 104)\}.$$

Example 5.7 The set of all ordered pairs of the form (P, V), where P is a former president of the United States and V is a vice president who served with that president, is a relation. The fact that Al Gore served as vice president under Bill Clinton is expressed via the membership of the ordered pair (*Bill Clinton*, *Al Gore*) in this relation.

Given sets X and Y, any subset of $X \times Y$ is called a **relation on** $X \times Y$. That is, a relation on $X \times Y$ is a set of ordered pairs whose first coordinates are members of X and whose second coordinates are members of Y. A relation on $X \times X$ is usually referred to as a **relation on** X; in other words, a relation on X is really just a set of ordered pairs for which all coordinates are members of X.

Example 5.8 If $X = \{0, 1, 2\}$ and $Y = \{1, 0.5\}$, then $\{(0, 0.5), (1, 1), (1, 0.5)\}$ is a relation on $X \times Y$, $\{(0.5, 0), (0.5, 2)\}$ is a relation on $Y \times X$, $\{(0, 2), (1, 0)\}$ is a relation on X, and $\{(0.5, 0.5), (0.5, 1), (1, 0.5)\}$ is a relation on Y.

Example 5.9 The set $\{(x, x^2) \mid x \in \mathbf{R}\}$ is a relation on $\mathbf{R}$ because all coordinates, both first and second, are real numbers. In a precalculus or calculus course, this relation might be described using the equation $y = x^2$.

Sometimes a letter or a symbol such as "~" (called a *tilde*) is used as the name of a relation. If ~ is a relation, it is common to write $x \sim y$ in place of $(x, y) \in \sim$. In part, this is because we can read $x \sim y$ as *x is related to y*, which is what we intend when (x, y) is a member of the relation ~.

Example 5.7 (continued) Let ~ be the relation consisting of those ordered pairs for which the first coordinate is a former president of the United States and the second coordinate is a vice president who served with that president. Because (*Bill Clinton*, *Al Gore*) ∈ ~, we may write *Bill Clinton* ~ *Al Gore*.

Example 5.9 (continued) If $S = \{(x, x^2) \mid x \in \mathbf{R}\}$, we may write $3S9$ because $(3, 9) \in S$.

> **Practice Problem 3** Let ∇ be the relation consisting of all ordered pairs of nonzero integers for which either coordinate, when squared, produces the same number.
>
> **(a)** Is it true that $-11\nabla 11$?
>
> **(b)** Write an equation relating a, b, and c if it is known that $-a\nabla bc$.
>
> **(c)** What ordered pairs must be members of ∇ if it is known that $-a\nabla bc$?

We must be careful, however, not to assume that just because x is related to y by a certain relation that it must also be the case that y is related to x.

Example 5.7 (continued) Whereas *Bill Clinton* ~ *Al Gore*, it does not follow that *Al Gore* ~ *Bill Clinton*, as this would signify that Bill Clinton served as a vice president under Al Gore, which is not true. Thus, *Al Gore* ≁ *Bill Clinton*.

The set of all first coordinates of the ordered pairs in a relation is called the **domain** of the relation. The set of all second coordinates is called the **range**.

Example 5.6 (continued) The relation

$$\{(ME, 105), (NH, 106), (VT, 105), (MA, 107), (CT, 106), (RI, 104)\}$$

in which a New England state is paired with its record high temperature has domain

$$\{ME, NH, VT, MA, CT, RI\},$$

the set consisting of the six New England states, and range

$$\{105, 106, 107, 104\},$$

the set of record high temperatures achieved in these states. Note that we should not refer to, for instance, NH as "a domain" of this relation or 106 as "a range." The domain is the *set of all first coordinates*, not a particular first coordinate, and the range is the *set of all second coordinates*, not a particular second coordinate.

Example 5.9 (continued) Because any real number can be the first coordinate of an ordered pair in $S = \{(x, x^2) \mid x \in \mathbf{R}\}$, the domain of S is $\mathbf{R}$.

However, as the square of a real number will never be negative, the range of S includes no negative numbers. Now for each nonnegative real number y, there is at least one real number x for which $x^2 = y$ (for instance, take $x = \sqrt{y}$). Thus, we may conclude that the range of S is $\mathbf{R}^+ \cup \{0\}$.

Example 5.7 (continued) The relation consisting of the set of all ordered pairs (P, V), where P is a former president of the United States and V is a vice president who served with that president, has the set of all former U.S. presidents as its domain and the set of all former U.S. vice presidents as its range.

> **Practice Problem 4** Find the domain and range of the relation $\{(x, y) \in \mathbf{N} \times \mathbf{Z} \mid y = -2x\}$.

The next few examples consider some of the most important relations in all of mathematics.

Example 5.10 For real numbers a and b, we write $a \le b$ to indicate that a is *less than or equal to* b. Formally, we may view $\le$ as the relation on $\mathbf{R}$ consisting of those ordered pairs of real numbers for which the first coordinate is less than or equal to the second coordinate. Thus, for any real numbers a and b, $(a, b) \in \le$ iff $a \le b$.

For instance, $(-2, \pi) \in \le$ because the first coordinate -2 is less than or equal to the second coordinate π. Also, $(\pi, \pi) \in \le$ because $\pi \le \pi$ is a true statement. But $(4, \pi) \notin \le$ because $4 \le \pi$ is not a true statement.

This example also illustrates that the idea of placing the symbol representing a relation between the objects being related is not completely unfamiliar. For example, we are much more used to writing $-2 \le \pi$ than the equivalent $(-2, \pi) \in \le$.

Example 5.11 Similarly, $<$ is the relation on $\mathbf{R}$ consisting of all ordered pairs of real numbers for which the first coordinate is *less than* the second coordinate. Thus, for instance, $(-2, \pi) \in <$ because $-2 < \pi$, but $(\pi, \pi) \notin <$ as $\pi \not< \pi$ and $(4, \pi) \notin <$ as $4 \not< \pi$.

Example 5.12 Another important relation on $\mathbf{R}$ is *equality*, denoted $=$. For any real numbers a and b, $(a, b) \in =$ iff $a = b$. Hence, as $\pi = \pi$ we have $(\pi, \pi) \in =$, but as $-2 \ne \pi$ we have $(-2, \pi) \notin =$.

Example 5.13 Given a set A, $\subseteq$ is the relation on $\mathcal{P}(A)$, the power set of A (the set of all subsets of A), where, for any subsets B and C of A, $(B, C) \in \subseteq$ iff $B \subseteq C$. For instance, $(\{1,2\}, \mathbf{N}) \in \subseteq$ because $\{1,2\} \subseteq \mathbf{N}$, but $(\{1,2\}, \{1,3,5\}) \notin \subseteq$ because $\{1, 2\} \not\subseteq \{1, 3, 5\}$.

Intervals and Interval Notation

Because many of the relations we work with in mathematics involve real numbers, this is a good point at which to introduce the notion of an *interval* of real numbers and to review with you notations for intervals that you likely learned in your calculus courses.

A subset I of $\mathbf{R}$ is an **interval** provided that whenever $a, b \in I$ and $a < x < b$, it follows that $x \in I$. In other words, an *interval* of real numbers is a subset of $\mathbf{R}$ having the property that it includes every real number between any two of its members. Intuitively, intervals are "connected" in the sense that they have no "gaps" in them.

Example 5.14 The set $\mathbf{R}^+$ is an interval because between any two positive numbers there are only positive numbers.

But the set $A = \{1, 2, 3\}$ is *not* an interval because it does not include every real number between any two of its members. For instance, the real number 1.5 lies between the numbers 1 and 2, both of which are in A, yet 1.5 is not in A.

Given an interval:

- If there is a real number that is less than or equal to every member of the interval, the largest such real number is called the **left endpoint** of the interval.
- If there is a real number that is greater than or equal to every member of the interval, the smallest such real number is called the **right endpoint** of the interval.

Example 5.15 The interval $A = \{x \in \mathbf{R} \mid 1 \leq x < \pi\}$ has left endpoint 1 because 1 is the largest real number that is less than or equal to every member of A. The interval A has right endpoint π because π is the smallest real number that is greater than or equal to every member of A. Note that an endpoint of an interval may or may not be included as a member of the interval.

Example 5.16 The interval $\{x \in \mathbf{R} \mid x > 4\}$ has left endpoint 4, but no right endpoint. The interval $\{x \in \mathbf{R} \mid x \leq 4\}$ has right endpoint 4, but no left endpoint. And the interval $\mathbf{R}$ has neither a left endpoint nor a right endpoint.

Any interval of real numbers can be described using so-called **interval notation**. We simply write down the interval's left endpoint, or the symbol "$-\infty$" if the interval has no left endpoint, followed by a comma, then write down the interval's right endpoint, or the symbol "∞" if there is no right endpoint, and then use "square" or "round" brackets to indicate, respectively, whether or not an endpoint is included.

Example 5.16 (continued) The interval $\{x \in \mathbf{R} \mid x > 4\}$ can be expressed by writing $(4, \infty)$, and the interval $\{x \in \mathbf{R} \mid x \leq 4\}$ can be expressed by writing $(-\infty, 4]$. The set $\mathbf{R}$ of all real numbers may be expressed as $(-\infty, \infty)$.

Example 5.15 (continued) The interval $\{x \in \mathbf{R} \mid 1 \leq x < \pi\}$ can be expressed as $[1, \pi)$.

When using interval notation, a square bracket is never used next to either of the symbols "∞" or "$-\infty$" because neither of these symbols represents a real number. Rather, "∞" should be interpreted as *unbounded in the positive sense* and "$-\infty$" as *unbounded in the negative sense*.

Practice Problem 5 Express the given interval using interval notation and identify its endpoints (if any).

(a) $\{t \in \mathbf{R} \mid 0 \geq t > -2\}$. (b) $\{x \in \mathbf{R}^- \mid x < -3.7\}$.

The string of symbols (a, b) in isolation is ambiguous because it can refer to either the interval consisting of all real numbers between a and b (including neither a nor b) or the ordered pair having first coordinate a and second coordinate b. Context will dictate which interpretation is correct in a given setting. For example, if the range of a certain relation is $(3, 6)$, we would view $(3, 6)$ as an interval, whereas if $(3, 6)$ is a member of a certain relation, we would view $(3, 6)$ as an ordered pair.

Answers to Practice Problems

1. $(-3, 0)$.
2. $A \times B = \{(\pi, 1), (\pi, 2), (-\pi, 1), (-\pi, 2), (0, 1), (0, 2)\}$ and $B \times B = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$.
3. **(a)** yes. **(b)** $(-a)^2 = (bc)^2$. **(c)** $(-a, bc)$, $(-a, -bc)$, (a, bc), and $(a, -bc)$.
4. The domain is $\mathbf{N}$, and the range is the set of all negative even integers.
5. **(a)** $(-2, 0]$, which has left endpoint -2 and right endpoint 0. **(b)** $(-\infty, -3.7)$, which has no left endpoint and right endpoint -3.7.

5.2 Equivalence Relations and Partitions

Questions to guide your reading of this section:

1. What does it mean for a relation on a set to be *reflexive*? *symmetric*? *transitive*?
2. Under what circumstances is a relation an *equivalence relation*? What is meant by the *equivalence classes* created by an equivalence relation?
3. What is a *partition* of a set? What is meant by the *cells* created from a partition?
4. How are equivalence relations and partitions related to each other?

There are many properties a relation may possess.

Definition 5.17 A relation $\sim$ on a set A is

- **reflexive** if for every $a \in A$, $a \sim a$;
- **symmetric** if whenever $a \sim b$, it follows that $b \sim a$;
- **transitive** if whenever both $a \sim b$ and $b \sim c$, it follows that $a \sim c$.

The choice of terminology for the properties listed in Definition 5.17 makes some sense.

- When you look into a mirror, you *see yourself* via your *reflection*. *Reflexivity* of a relation requires that each member of the set on which the relation is defined be *related to itself*.

- We think of the letter

A

as being *symmetric* with respect to an imaginary vertical line we can draw straight down through the point at its top. This is because the portion of the letter lying to the left of this line can be *flipped* with the portion of the letter to the right of the line to re-create the original figure. *Symmetry* of a relation tells us that whenever a is related to b, we may *flip* a and b and conclude that b is related to a.

- The prefix *trans-* means *across*. We can think of *transitivity* of a relation as telling us that when there is a "bridge" b linking a to c, meaning a is related to b and b is related to c, we can move *across* the bridge to conclude that a is related to c.

Example 5.18 Let $A = \{1, 2, 3, 4\}$ and consider the relation

$$R_1 = \{(1,1), (2,2), (3,3), (2,3), (3,2), (2,4)\}$$

on A. Recalling that we may write xR_1y iff $(x, y) \in R_1$, we see that R_1 is *not* reflexive because, even though $1R_11$, $2R_12$, and $3R_13$, we do *not* have $4R_14$. Also, as $2R_14$, symmetry would also require that $4R_12$, but this is *not* the case as the ordered pair $(4, 2)$ is *not* in R_1. Thus, R_1 is *not* symmetric. (Even though $2R_13$ and $3R_12$, R_1 is not symmetric because the ability to flip coordinates and still have a member of the relation does not always hold.) Transitivity of R_1 would require, as $3R_12$ and $2R_14$, that $3R_14$. However, it is *not* true that $3R_14$ because the ordered pair $(3, 4)$ is *not* in R_1. Thus, R_1 is *not* transitive.

It is usually relatively straightforward to determine whether a relation is reflexive or symmetric. Determining whether a relation $\sim$ possesses the transitive property can, however, sometimes be challenging because we must consider *all* possible instances for which $a \sim b$ and $b \sim c$; that is, the right coordinate of an ordered pair in $\sim$ is the same as the left coordinate of an ordered pair in $\sim$. If we are able to reach the conclusion that $a \sim c$ in *each* such instance, we conclude that $\sim$ is transitive. But if we can find even one instance where we cannot reach this conclusion, $\sim$ is *not* transitive.

We point out that when looking for situations where $a \sim b$ and $b \sim c$, there is no need to consider the special cases $b = a$ and $b = c$, as we will find no violations of the transitive property in such situations. For example, if $b = a$, the requirement

$$(a \sim b \text{ and } b \sim c) \Rightarrow a \sim c$$

becomes

$$(a \sim a \text{ and } a \sim c) \Rightarrow a \sim c,$$

which is clearly true because the conclusion is actually part of the hypothesis. The situation is similar for the special case $b = c$.

Example 5.19 Define a relation R_2 on $A = \{1, 2, 3, 4\}$ so that

$$1R_21,\ 2R_22,\ 3R_23,\ 4R_24,\ 2R_23, \text{ and } 3R_22.$$

To determine whether R_2 is transitive, we consider all possible instances where the right coordinate of an ordered pair in R_2 is the same as the left coordinate of an ordered pair in R_2, realizing, based on the discussion immediately prior to this example, that we do not need to consider ordered pairs having identical coordinates. In each of these situations, we want to know if the appropriate conclusion can be reached. Knowing that $2R_23$ and $3R_22$, is it also true that $2R_22$? Knowing that $3R_22$ and $2R_23$, is it also true that $3R_23$? Because the answer to each of these questions is "yes," we may conclude that R_2 is transitive.

The relation R_2 is reflexive because *every* member of A is related to itself and is symmetric as interchanging the coordinates of any ordered pair in R_2 yields an ordered pair that is in R_2.

Practice Problem 1 Consider the relation

$$R_3 = \{(1, 1), (2, 2), (3, 3), (2, 3), (3, 2), (2, 4), (3, 4)\}$$

on the set $A = \{1, 2, 3, 4\}$. Show that R_3 is neither reflexive nor symmetric, but is transitive.

Although every relation is a set of ordered pairs, we must remember it is often the case that a relation is not directly specified in this way.

Example 5.20 Let P be a set of people and define the relation $\sim$ on P so that $a \sim b$ iff b has the same father as a. Observe that:

- As any person has the same father as herself/himself, we may conclude that $a \sim a$ for every $a \in P$, thus making $\sim$ reflexive.
- Whenever $a \sim b$, meaning b has the same father as a, we may conclude that a has the same father as b; that is, $b \sim a$, so $\sim$ is symmetric.

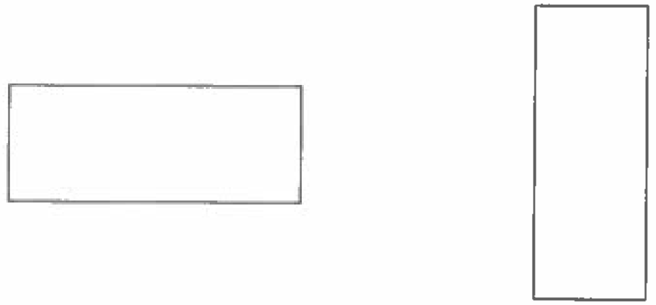
- Whenever $a \sim b$ and $b \sim c$, meaning b has the same father as a and c has the same father as b, we may conclude that c has the same father as a (a and c must have the same father, whoever is father to b); that is, $a \sim c$, so $\sim$ is transitive.

> **Practice Problem 2** Let P be a set of people and define the relation $\sim$ on P so that $a \sim b$ iff b has a grandfather in common with a. Verify that $\sim$ is reflexive and symmetric, but need not be transitive.

Equivalence Relations

A given relation attempts to capture a connection between the first coordinate and the second coordinate of each ordered pair that is a member of the relation. One type of connection a relation may express is a notion of *equivalence*; that is, a notion of one object *being the same as* another object in terms of one or more specified properties or attributes the objects may possess. Here are some examples of "equivalence" that come up in the study of mathematics:

- Numerical or algebraic expressions that have the *same value* are *equal*. For example, we may write $x^2 - 1 = (x - 1)(x + 1)$ because the expressions $x^2 - 1$ and $(x - 1)(x + 1)$ have the same value for each real number x.
- Geometric figures that have the *same shape and the same size* are equivalent in the sense of being *congruent* to one another. For example, the two rectangles shown below are congruent even though they are oriented differently; from the point of view of size and shape, they are equivalent.



- Statement forms that have the *same truth table* are *logically equivalent*. For example, one can easily check that the statement forms $P \vee Q$ and $\sim P \Rightarrow Q$ are logically equivalent.

We are interested in those properties that are shared by all relations that express a notion of equivalence. Because we would expect a to be "equivalent" to itself, b to be "equivalent" to a if a is "equivalent" to b, and a to be "equivalent" to c if a is "equivalent" to b and b is "equivalent" to c, we realize that reflexivity, symmetry, and transitivity must be among these properties. It turns out, in fact, that these three properties completely characterize the abstract notion of "equivalence."

Definition 5.21 A relation that is reflexive, symmetric, and transitive is called an **equivalence relation**. If $\sim$ is an equivalence relation on the set A, then:

- When $x \sim y$, we say that x and y are **equivalent**.
- The set $\{x \in A \mid x \sim a\}$ of all members of A that are equivalent to a specified $a \in A$ is the **equivalence class** of a and is denoted $[a]$.

Thus, to prove a relation is an equivalence relation, we must show it is reflexive, symmetric, and transitive.

Example 5.22 Two geometric figures A and B are **congruent**, denoted $A \cong B$, provided they have the same shape and the same size. Note that:

- Any geometric figure has the same shape and size as itself, thus making congruence reflexive.
- Whenever a geometric figure A has the same shape and size as a geometric figure B, it follows that B has the same shape and size as A, thus making congruence symmetric.
- Whenever a geometric figure A has the same shape and size as a geometric figure B, and B also has the same shape and size as a geometric figure C, it follows that A has the same shape and size as C, thus making congruence transitive.

Hence, if $\mathcal{G}$ is a collection of geometric figures, $\cong$ is an equivalence relation on $\mathcal{G}$. If $C \in \mathcal{G}$, the equivalence class of C under $\cong$ is $[C] = \{x \in \mathcal{G} \mid x \cong C\}$, the set of all figures in $\mathcal{G}$ that are congruent to the particular geometric figure C. So if C is a circle of radius 5 inches, then $[C]$ would be the set of all circles of radius 5 inches that are in $\mathcal{G}$.

Example 5.23 Let S be the set of students at a certain college who have declared a major and assume none of these students is pursuing more than one major. Define a relation $\sim$ on S so that $a \sim b$ iff a is pursuing the same major as b. Observe that:

- Because any student pursues the same major as himself/herself, $a \sim a$ for every $a \in S$, meaning $\sim$ is reflexive.
- Whenever a pursues the same major as b, it follows that b must be pursuing the same major as a, meaning $\sim$ is symmetric.
- Whenever a pursues the same major as b, and b pursues the same major as c, it follows that a must be pursuing the same major as c, meaning $\sim$ is transitive.

Therefore, $\sim$ is an equivalence relation. If *Nancy* is a declared biology major at the college of interest, then $[Nancy] = \{x \in S \mid x \sim Nancy\}$ is the set of all declared majors at the college pursing the same major as Nancy; in other words, $[Nancy]$ is the set of all declared biology majors at the college. Thus, if *Joe* is also a declared biology major at the college, then $[Joe] = [Nancy]$.

Example 5.12 (continued) Because it is easily seen that equality of real numbers is reflexive, symmetric, and transitive, it follows that = is an equivalence relation on $\mathbf{R}$. Because a real number is equal to itself but to no other real number, for each $a \in \mathbf{R}$, the equivalence class of a under equality is $[a] = \{a\}$.

To demonstrate that a relation is *not* an equivalence relation, we need only show that a single one of the properties reflexivity, symmetry, or transitivity fails to hold.

Example 5.10 (continued) In Problem 5E.2 you will show the relation $\leq$ on $\mathbf{R}$ is not symmetric. Hence, $\leq$ is not an equivalence relation on $\mathbf{R}$.

Example 5.19 (continued) Recall that the relation R_2 was defined on the set $A = \{1, 2, 3, 4\}$ so that $R_2 = \{(1, 1), (2, 2), (3, 3), (4, 4), (2, 3), (3, 2)\}$. We previously showed R_2 is reflexive, symmetric, and transitive, so R_2 is an equivalence relation. Note that R_2 is identifying 2 and 3 as being equivalent to each other, but otherwise a member of A is only equivalent to itself. The equivalence classes produced by R_2 are $[1] = \{1\}$, $[2] = \{2, 3\} = [3]$, and $[4] = \{4\}$.

Practice Problem 3 Tell why the relation R_3 described in Practice Problem 1 is not an equivalence relation.

In terms of our *Know/Show* approach to developing proofs, observe that to establish that a relation $\sim$ on a set A is

- *reflexive*: we *know* $a \in A$ and we must *show* $a \sim a$;
- *symmetric*: we *know* $a \sim b$ and we must *show* $b \sim a$;
- *transitive*: we *know* $a \sim b$ and $b \sim c$ and we must *show* $a \sim c$.

In carrying out a proof that an unfamiliar relation is actually an equivalence relation, the relation's definition will be of prime importance in establishing these properties.

Example 5.24 Define a relation $\equiv$ on the set $\mathbf{Z}$ so that for any integers a and b, $a \equiv b$ iff there is an integer k such that $a - b = 4k$. We will prove that $\equiv$ is an equivalence relation. The definition of $\equiv$ will be fundamental to our task. Specifically, to show that $\equiv$ is

- *reflexive*: we *know* a is an integer and we must *show* $a \equiv a$, which means we must *show* there is an integer k such that $a - a = 4k$ (observe that the integer 0 will play the role of k here);
- *symmetric*: we *know* $a \equiv b$, meaning there is an integer k such that $a - b = 4k$, and we must *show* $b \equiv a$, meaning we must *show* there is an integer j such that $b - a = 4j$ (because $b - a$ is the opposite of $a - b$, it appears that the role of j will be taken by the integer $-k$);
- *transitive*: we *know* $a \equiv b$ and $b \equiv c$, meaning there exist integers j and k such that $a - b = 4j$ and $b - c = 4k$, and we must *show* $a \equiv c$, meaning we must *show* there is an integer l such that $a - c = 4l$ [observing that $a - c = (a - b) + (b - c)$ suggests the role of l will be played by $j + k$].

The following proof is constructed based on the plan we have just outlined.

Claim: The relation $\equiv$ is an equivalence relation on $\mathbf{Z}$.

Proof: We need only show that $\equiv$ is reflexive, symmetric, and transitive.

Subclaim 1: $\equiv$ is reflexive.
Consider $a \in \mathbf{Z}$. Because $a - a = 0 = 4 \cdot 0$ and 0 is an integer, by definition of $\equiv$, we may conclude that $a \equiv a$.

Subclaim 2: $\equiv$ is symmetric.
Suppose $a \equiv b$. Then, by the definition of $\equiv$, there is an integer k such that $a - b = 4k$. It follows that $b - a = -(a - b) = -4k = 4(-k)$, where $-k \in \mathbf{Z}$ as the opposite of an integer is also an integer. Thus, by the definition of $\equiv$, we may conclude that $b \equiv a$.

Subclaim 3: $\equiv$ is transitive.
Suppose $a \equiv b$ and $b \equiv c$. Then, by the definition of $\equiv$, there exist integers j and k such that $a - b = 4j$ and $b - c = 4k$. It follows that

$$a - c = (a - b) + (b - c) = 4j + 4k = 4(j + k),$$

where $j + k \in \mathbf{Z}$ as the sum of integers is also an integer. Thus, by the definition of $\equiv$, we may conclude that $a \equiv c$. □

Note that, as proving a relation is an equivalence relation involves proving three separate properties (reflexivity, symmetry, and transitivity), we have used *subclaims* within our proof in Example 5.24 to indicate where each property is being proved. The proof of each subclaim immediately follows its statement. Such a layout is fairly typical when a proof involves establishing multiple independent results.

> **Practice Problem 4** Define the relation ~ on the set **Z** so that $a \sim b$ iff $a + b$ is even. Prove that ~ is an equivalence relation. You may freely use the fact that the sum and difference of even integers is even.

Partitions

We introduced the notion of *equivalence relation* in attempting to capture the general idea of objects *being the same as one another* with respect to a specified attribute or property. It would appear that another way to encapsulate this idea would be simply to group objects into sets in such a way that *equivalent* objects wind up being in the same set.

Example 5.25 Suppose P is a set consisting of two different triangles T_1 and T_2, four different quadrilaterals Q_1, Q_2, Q_3, and Q_4, and one hexagon H; that is, $P = \{T_1, T_2, Q_1, Q_2, Q_3, Q_4, H\}$. To capture the notion that figures in P having the same number of sides will be considered equivalent, we could sort them into three subsets $\{T_1, T_2\}$, $\{Q_1, Q_2, Q_3, Q_4\}$, and $\{H\}$.

Note that the sorting process illustrated in Example 5.25 puts each member of the original set P into exactly one of the three subsets formed through the consideration of a figure's number of sides.

Definition 5.26 A collection of nonempty subsets of a nonempty set A having the property that each element of A is a member of exactly one of the subsets in the collection is called a **partition** of A. Each of the subsets in a partition of A is called a **cell** of the partition. The cell including an element $a \in A$ is denoted $[a]$.

Example 5.25 (continued) Thus, $\{\{T_1, T_2\}, \{Q_1, Q_2, Q_3, Q_4\}, \{H\}\}$ is a partition of $P = \{T_1, T_2, Q_1, Q_2, Q_3, Q_4, H\}$. Note that the cell containing T_1 is $[T_1] = \{T_1, T_2\} = [T_2]$.

> **Practice Problem 5** Explain why $\{\mathbf{Q}^+, \mathbf{R}^+ - \mathbf{Q}^+, \{0\}, \mathbf{Q}^-, \mathbf{R}^- - \mathbf{Q}^-\}$ is a partition of the set **R** of all real numbers.

Given a nonempty set A and an equivalence relation ~ on A, we may note that:

- For each $a \in A$, the equivalence class $[a] = \{x \in A \mid x \sim a\}$ is nonempty, as it includes a (reflexivity of ~ implies that a is equivalent to itself).
- For each $a \in A$, the equivalence class $[a]$ is a subset of A ($[a]$ consists of those members of A that are equivalent to a).
- Each member of A is included in exactly one equivalence class, as $a \in [x]$ and $a \in [y]$ would tell us that $a \sim x$ and $a \sim y$, from which it follows, using the symmetry and transitivity of ~, that $x \sim y$, which itself implies that $[x] = [y]$.

Thus, the equivalence classes determined by an equivalence relation on a set always form a partition of the set.

Conversely, a partition of a set A naturally induces an equivalence relation on A in such a way that the resulting equivalence classes are precisely the subsets into which A had originally been partitioned. We simply declare that two members of A are equivalent iff they are members of the same cell of the partition.

Example 5.25 (continued) Define a relation ~ on $P = \{T_1, T_2, Q_1, Q_2, Q_3, Q_4, H\}$ so that $a \sim b$ iff a and b are in the same cell of the partition $\{\{T_1, T_2\}, \{Q_1, Q_2, Q_3, Q_4\}, \{H\}\}$ of P. Note that ~ consists of the ordered pairs (T_1, T_1), (T_1, T_2), (T_2, T_1), (T_2, T_2), (Q_1, Q_1), (Q_1, Q_2), (Q_2, Q_1), (Q_2, Q_2), (Q_1, Q_3), (Q_3, Q_1), (Q_3, Q_3), (Q_1, Q_4), (Q_4, Q_1), (Q_4, Q_4), (Q_2, Q_3), (Q_3, Q_2), (Q_2, Q_4), (Q_4, Q_2), (Q_3, Q_4), (Q_4, Q_3), and (H, H). It is easily verified that ~ is an equivalence relation (do so) with the equivalence classes being precisely the three cells $\{T_1, T_2\}$, $\{Q_1, Q_2, Q_3, Q_4\}$, and $\{H\}$ of the original partition.

Hence, we see that partitions and equivalence relations are simply two alternative methods for imposing a notion of "equivalence" on a set of objects. This observation further validates our choice of including all three of the reflexive, symmetric, and transitive properties when formulating the definition of *equivalence relation*.

> **Answers to Practice Problems**
>
> 1. R_3 is neither reflexive nor symmetric, for the same reasons, respectively, that the relation R_1 from Example 5.18 is neither reflexive nor symmetric. To verify that R_3 is transitive, we consider all possible situations where we have two ordered pairs in R_3 for which

(*continues*)