CONTENTS

Chapter 1

The Pigeonhole Principle

How do we know that a computer program produces the right results? How do we know that a program will run to completion? If we know it will stop eventually, can we predict whether that will happen in a second, in an hour, or in a day? Intuition, testing, and "it has worked OK every time we tried it" should not be accepted as proof of a claim. Proving something requires formal reasoning, starting with things known to be true and connecting them together by incontestable logical inferences. This is a book about the mathematics that is used to reason about the behavior of computer programs.

The mathematics of computer science is not some special field. Computer scientists use almost every branch of mathematics, including some that were never thought to be useful until developments in computer science created applications for them. So this book includes sections on mathematical logic, graph theory, counting, number theory, and discrete probability theory, among other things. From the standpoint of a traditional mathematics curriculum, this list includes apples and oranges. One common feature of these topics is that all prove useful in computer science. Moreover, they are all *discrete mathematics*, which is to say that they involve quantities that change in steps, not continuously, or are expressed in symbols and structures rather than numbers. Of course, calculus is also important in computer science, because it assists in reasoning about continuous quantities. But in this book we will rarely use integrals and derivatives.

❈

One of the most important skills of mathematical thinking is the art of *generalization*. For example, the proposition

> *There is no triangle with sides of lengths* 1, 2, *and* 6

is true, but very specific (see Figure 1.1). The sides of lengths 1 and 2 would have to join the side of length 6 at its two ends, but the two short sides together aren't long enough to meet up at the third corner.

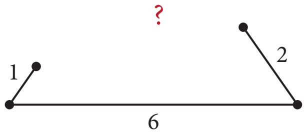

Figure 1.1. Can there be a triangle with sides of lengths 1, 2 and 6?

A more general statement might be (Figure 1.2)

> *There is no triangle with sides of lengths a, b, and c if a, b, c are any numbers such that $a + b \leq c$.*

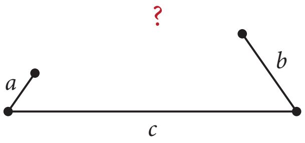

Figure 1.2. There is no triangle with sides of lengths a, b and c if $a + b \leq c$.

The second form is more general because we can infer the first from the second by letting $a = 1$, $b = 2$, and $c = 6$. It also covers a case that the picture doesn't show—when $a + b = c$, so the three "corners" fall on a straight line. Finally, the general rule has the advantage of not just stating what is impossible, but explaining it. There is no $1 - 2 - 6$ triangle because $1 + 2 \leq 6$.

So we state propositions in general form for two reasons. First, a proposition becomes more useful if it is more general; it can be applied with confidence in a greater variety of circumstances. Second, a general proposition makes it easier to grasp what is really going on, because it leaves out irrelevant, distracting detail.

❉

As another example, let's consider a simple scenario.

> *Annie, Batul, Charlie, Deja, Evelyn, Fawwaz, Gregoire, and Hoon talk to each other and discover that Deja and Gregoire were both born on Tuesdays.* (1.1)

Well, so what? Put two people together and they may or may not have been born on the same day of the week. Yet there is something going on here that can be generalized. As long as there are at least eight people, *some* two of them must have been born on the same day of the week, since a week has only seven days. *Some* statement like (1.1) must be true, perhaps with a different pair of names and a different day of the week. So here is a more general proposition.

> *In any group of eight people, some two of them were born on the same day of the week.*

But even that isn't really general. The duplication has nothing to do with properties of people or days of the week, except how many there are of each. For the same reason, if we put eight cups on seven saucers, some saucer would have two cups on it. In fact there is nothing magic about "eight" and "seven," except that the one is larger than the other. If a hotel has 1000 rooms and 1001 guests, some room must contain at least two guests. How can we state a general principle that covers all these cases, without mentioning the irrelevant specifics of any of them?

First, we need a new concept. A *set* is a collection of things, or *elements*. The elements that belong to the set are called its *members*. The members of a set must be *distinct*, which is another way of saying they are all different from each other. So the people mentioned in (1.1) form a set, and the days of the week form another set. Sometimes we write out the members of a set explicitly, as a list within curly braces {}:

$$P = \{\text{Annie, Batul, Charlie, Deja, Evelyn, Fawwaz, Gregoire, Hoon}\}$$
$$D = \{\text{Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, Saturday}\}.$$

When we write out the elements of a set, their order does not matter—in any order it is still the same set. We write $x \in X$ to indicate that the element x is a member of the set X. For example, Charlie $\in P$ and Thursday $\in D$.

We need some basic terminology about numbers in order to talk about sets. An *integer* is one of the numbers 0, 1, 2, …, or −1, −2, …. The *real* numbers are all the numbers on the number line, including all the integers and also all the numbers in between integers, such as $\frac{1}{2}$, $-\sqrt{2}$, and π. A number is *positive* if it is greater than 0, *negative* if it is less than 0, and *nonnegative* if it is greater than or equal to 0.

For the time being, we will be discussing finite sets. A *finite* set is a set that can (at least in principle) be listed in full. A finite set has a *size* or *cardinality*, which is a nonnegative integer. The cardinality of a set X is denoted $|X|$. For example, in the example of people and the days of the week on which they were born, $|P| = 8$ and $|D| = 7$, since eight people are listed and there are seven days in a week. A set that is not finite—the set of integers, for example—is said to be *infinite*. Infinite sets have sizes too—an interesting subject to which we will return in our discussion of infinite sets in Chapter 7.

Now, a *function* from one set to another is a rule that associates each member of the first set with exactly one member of the second set. If f is a function from X to Y and $x \in X$, then $f(x)$ is the member of Y that the function f associates with x. We refer to x as the *argument* of f and $f(x)$ as the *value* of f on that argument. We write $f : X \to Y$ to indicate that f is a function *from* set X *to* set Y. For example, we could write $b : P \to D$ to denote the function that associates each of the eight friends with the day of the week on which he or she was born; if Charlie was born on a Thursday, then $b(\text{Charlie}) = \text{Thursday}$.

A function $f : X \to Y$ is sometimes called a *mapping* from X to Y, and f is said to *map* an element $x \in X$ to the element $f(x) \in Y$. (In the same way, a real map associates a point on the surface of the earth with a point on a sheet of paper.)

Finally, we have a way to state the general principle that underlies the example of (1.1):

$$\text{If } f : X \to Y \text{ and } |X| > |Y|, \text{ then there are elements } x_1, x_2 \in X \text{ such that } x_1 \neq x_2 \text{ and } f(x_1) = f(x_2). \tag{1.2}$$

The statement (1.2) is known as the *Pigeonhole Principle*, as it captures in mathematical form this commonsense idea: if there are more pigeons than pigeonholes and every pigeon goes into a pigeonhole, then some pigeonhole must have more than one pigeon in it. The pigeons are the members of X and the pigeonholes are the members of Y (Figure 1.3).

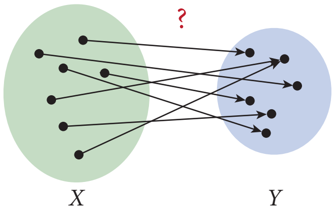

Figure 1.3. The Pigeonhole Principle. If $|X| > |Y|$ and f is any function from X to Y, then the values of f must be the same for some two distinct members of X.

We will provide a formal proof of the Pigeonhole Principle on page 34, once we have developed some of the basic machinery for doing proofs. For now, let's scrutinize the statement of the Pigeonhole Principle with an eye toward understanding mathematical language. Here are some questions we might ask:

1. What are X and Y?

   They are finite sets. To be absolutely clear, we might have begun the statement with the phrase, "For any finite sets X and Y," but the assertion that f is a function from X to Y makes sense only if X and Y are sets, and it is understood from context that the sets under discussion are finite—and we therefore know how to compare their sizes.

2. Why did we choose "x_1" and "x_2" for the names of elements of X?

   We could in principle have chosen any variables, "x" and "y" for example. But using variations on "X" to name elements of the set X suggests that x_1 and x_2 are members of the set X rather than the set Y. So using "x_1" and "x_2" just makes our statement easier to read.

3. Was the phrase "such that $x_1 \neq x_2$" really necessary? The sentence is simpler without it, and seems to say the same thing.

   Yes, the "$x_1 \neq x_2$" is necessary, and no, the sentence doesn't say the same thing without it! If we didn't say "$x_1 \neq x_2$," then "x_1" and "x_2" could have been two names for the same element. If we did not stipulate that x_1 and x_2 had to be different, the proposition would not have been false—only trivial! Obviously if $x_1 = x_2$, then $f(x_1) = f(x_2)$. That is like saying that the mass of Earth is equal to the mass of the third planet from the sun. Another way to state the Pigeonhole Principle would be to say, "there are distinct elements $x_1, x_2 \in X$ such that $f(x_1) = f(x_2)$."

One more thing is worth emphasizing here. A statement like "there are distinct elements $x_1, x_2 \in X$ with property blah" does *not* mean that there are *exactly* two elements with that property. It just means that at least two such elements exist for sure—maybe more, but definitely not less.

❈

Mathematicians always search for the most general form of any principle, because it can then be used to explain more things. For example, it is equally obvious that we can't put 15 pigeons in 7 pigeonholes without putting at least 3 pigeons in some pigeonhole—but there is no way to derive that from the Pigeonhole Principle as we stated it. Here is a more general version:

Theorem 1.3. Extended Pigeonhole Principle. *For any finite sets X and Y and any positive integer k such that $|X| > k \cdot |Y|$, if $f : X \to Y$, then there are at least $k + 1$ distinct members $x_1, \ldots, x_{k+1} \in X$ such that $f(x_1) = \ldots = f(x_{k+1})$.*

The Pigeonhole Principle is the $k = 1$ case of the Extended Pigeonhole Principle.

We have used *sequence* notation here for the first time, using the same variable with numerical subscripts in a range. In this case the x_i, where $1 \leq i \leq k + 1$, form a sequence of length $k + 1$. This notation is very convenient since it makes it possible to use an algebraic expression such as $k + 1$ in a subscript. Similarly, we could refer to the $2i^{\text{th}}$ member of a sequence $y_1, y_2, \ldots$ as y_{2i}.

The minimum value of the parameter k in the Extended Pigeonhole Principle, as applied to particular sets X and Y, can be derived once the sizes of X and Y are known. It is helpful to introduce some notation to make this calculation precise.

An integer p *divides* another integer q, symbolically written as $p \mid q$, if the quotient $\frac{q}{p}$ is an integer—that is, dividing q by p leaves no remainder. We write $p \nmid q$ if p does not divide q—for example, $3 \nmid 7$. If x is any real number, we write $\lfloor x \rfloor$ for the greatest integer less than or equal to x (called the *floor* of x). For example, $\lfloor \frac{17}{3} \rfloor = 5$, and $\lfloor \frac{6}{2} \rfloor = 3$. We will also need the *ceiling* notation: $\lceil x \rceil$ is the smallest integer greater than or equal to x, so for example $\lceil 3.7 \rceil = 4$.

With the aid of these notations, we can restate the Extended Pigeonhole Principle in a way that determines the minimum size of the most heavily occupied pigeonhole for given numbers of pigeons and pigeonholes:

Theorem 1.4. Extended Pigeonhole Principle, Alternate Version. *Let X and Y be any finite sets and let $f : X \to Y$. Then there is some $y \in Y$ such that $f(x) = y$ for at least*

$$\left\lceil \frac{|X|}{|Y|} \right\rceil$$

values of x.

Proof. Let $m = |X|$ and $n = |Y|$. If $n \mid m$, then this is the Extended Pigeonhole Principle with $k = \frac{m}{n} - 1 = \lceil \frac{m}{n} \rceil - 1$. If $n \nmid m$, then again this is the Extended Pigeonhole Principle with $k = \lceil \frac{m}{n} \rceil - 1$, since that is the largest integer less than $\frac{|X|}{|Y|}$. ■

❉

Once stated in their general form, these versions of the Pigeonhole Principle seem to be fancy ways of saying something obvious. In spite of that, we can use them to explain a variety of different phenomena—once we figure out what are the "pigeons" and the "pigeonholes." Let's close with an application to *number theory*—the study of the properties of the integers. A few basics first.

If $p \mid q$, then p is said to be a *factor* or *divisor* of q.

A *prime* number is an integer greater than 1 that is divisible only by itself and 1. For example, 7 is prime, because it is divisible only by 7 and 1, but 6 is not prime, because $6 = 2 \cdot 3$. Note that 1 itself is not prime.

Theorem 1.5. The Fundamental Theorem of Arithmetic. *There is one and only one way to express an integer greater than* 1 *as a product of distinct prime numbers in increasing order and with positive integer exponents.*

We'll prove this theorem in Chapter 4, but make some use of it right now. The *prime decomposition* of a number n is that unique product

$$n = p_1^{e_1} \cdot \ldots \cdot p_k^{e_k}, \tag{1.6}$$

where the p_i are primes in increasing order and the e_i are positive integers. For example, $180 = 2^2 \cdot 3^2 \cdot 5^1$, and there is no other product $p_1^{e_1} \cdot \ldots \cdot p_k^{e_k}$ equal to 180, where $p_1 < p_2 < \ldots < p_k$, all the p_i are prime, and the e_i are integer exponents.

The prime decomposition of the product of two integers m and n combines the prime decompositions of m and of n—every prime factor of $m \cdot n$ is a prime factor of one or the other.

Theorem 1.7. *If m, n, and p are integers greater than* 1*, p is prime, and $p \mid m \cdot n$, then either $p \mid m$ or $p \mid n$.*

Proof. By the Fundamental Theorem of Arithmetic (Theorem 1.5), there is one and only one way to write

$$m \cdot n = p_1^{e_1} \cdot \ldots \cdot p_k^{e_k},$$

where the p_i are prime. But then p must be one of the p_i, and each p_i must appear in the unique prime decomposition of either m or n. ■

The exponent of a prime p in the prime decomposition of $m \cdot n$ is the sum of its exponents in the prime decompositions of m and n (counting the exponent as 0 if p does not appear in the decomposition). For example, consider the product $18 \cdot 10 = 180$. We have

$$\begin{aligned}
18 &= 2^1 \cdot 3^2 && \text{(exponents of 2, 3, 5 are 1, 2, 0)}\\
10 &= 2^1 \cdot 5^1 && \text{(exponents of 2, 3, 5 are 1, 0, 1)}\\
180 &= 2^2 \cdot 3^2 \cdot 5^1 &&\\
&= 2^{1+1} \cdot 3^{2+0} \cdot 5^{0+1}. &&
\end{aligned}$$

We have color-coded the exponents to show how the exponents of 2, 3, and 5 in the product 180 are the sums of the exponents of those primes in the decompositions of the two factors 18 and 10.

Another important fact about prime numbers is that there are infinitely many of them.

Theorem 1.8. *There are arbitrarily large prime numbers.*

"Arbitrarily large" means that for every $n > 0$, there is a prime number greater than n.

Proof. Pick some value of k for which we know there are at least k primes, and let $p_1, \ldots, p_k$ be the first k primes in increasing order. (Since $p_1 = 2$, $p_2 = 3, p_3 = 5$, we could certainly take $k = 3$.) We'll show how to find a prime number greater than p_k. Since this process could be repeated indefinitely, there must be infinitely many primes.

Consider the number N that is one more than the product of the first k primes:

$$N = (p_1 \cdot p_2 \cdot \ldots \cdot p_k) + 1. \tag{1.9}$$

Dividing N by any of $p_1, \ldots, p_k$ would leave a remainder of 1. So N has no prime divisors less than or equal to p_k. Therefore, either N is not prime but has a prime factor greater than p_k, or else N is prime itself. ■

In the $k = 3$ case, for example, $N = 2 \cdot 3 \cdot 5 + 1 = 31$. Here N itself is prime; Problem 1.11 asks you to find an example of the case in which N is not prime.

A *common divisor* of two numbers is a number that divides both of them. For example, 21 and 36 have the common divisors 1 and 3, but 16 and 21 have no common divisor greater than 1.

With this by way of background, let's work a number theory example that uses the Pigeonhole Principle.

Example 1.10. *Choose m distinct numbers between* 2 *and* 40 *inclusive, where* $m \geq 13$. *Then at least two of the numbers have some common divisor greater than* 1.

"Between *a* and *b inclusive*" means including all numbers that are $\geq a$ and also $\leq b$—so including both 2 and 40 in this case.

Solution to example. Observe first that there are 12 prime numbers less than or equal to 40: 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, no two of which share a factor greater than 1. Let's define P to be this set of 12 prime numbers. (We needed to specify that $m \geq 13$, because the claim would be false with

$m = 12$ instead: the set P would be a counterexample.) Now consider a set X of m numbers in the range from 2 to 40 inclusive. We can think of the members of X as pigeons and the members of P as pigeonholes. To place pigeons in pigeonholes, use the function $f : X \to P$, where $f(x)$ is the smallest prime that divides x. For example, $f(16) = 2$, $f(17) = 17$, and $f(21) = 3$. By the Pigeonhole Principle, since $m > 12$, the values of f must be equal for two distinct members of X, and therefore at least two members of X have a common prime divisor. ■

Chapter Summary

- Mathematical thinking focuses on general principles, abstracted from the details of specific examples.
- A *set* is an unordered collection of *distinct* things, or *elements*. The elements of a set are its *members*.
- A set is *finite* if its members can be listed in full one by one. The number of members of a finite set X is called its *cardinality* or *size* and is denoted $|X|$. A set's size is always a *nonnegative integer*.
- A *function* or *mapping* between two sets is a rule associating each member of the first set with a unique member of the second.
- The *Pigeonhole Principle* states that if X is a set of pigeons and Y a set of pigeonholes, and $|X| > |Y|$, then any function mapping pigeons to pigeonholes assigns more than one pigeon to some pigeonhole.
- The *Extended Pigeonhole Principle* states that if X is a set of pigeons and Y a set of pigeonholes, and $|X| > k|Y|$, then any function mapping pigeons to pigeonholes assigns more than k pigeons to some pigeonhole.
- A *sequence* of terms can be denoted by a repeated variable with different numerical subscripts, such as $x_1, \ldots, x_n$. The subscript of a term may be an algebraic expression.
- The *Fundamental Theorem of Arithmetic* states that every positive integer has exactly one *prime decomposition*.

Problems

1.1. What are each of the following?

(a) $|\{0, 1, 2, 3, 4, 5, 6\}|$.

(b) $\left\lceil \frac{111}{5} \right\rceil$.

(c) $\left\lfloor \frac{5}{111} \right\rfloor$.

(d) The set of divisors of 100.

(e) The set of prime divisors of 100.

1.2. Let $f(n)$ be the largest prime divisor of n. Can it happen that $x < y$ but $f(x) > f(y)$? Give an example or explain why it is impossible.

1.3. Under what circumstances is $\lfloor x \rfloor = \lceil x \rceil - 1$?

1.4. Imagine a 9×9 square array of pigeonholes, with one pigeon in each pigeonhole. (So 81 pigeons in 81 pigeonholes—see Figure 1.4.) Suppose that all at once, all the pigeons move up, down, left, or right by one hole. (The pigeons on the edges are not allowed to move out of the array.) Show that some pigeonhole winds up with two pigeons in it. *Hint:* The number 9 is a distraction. Try some smaller numbers to see what is going on.

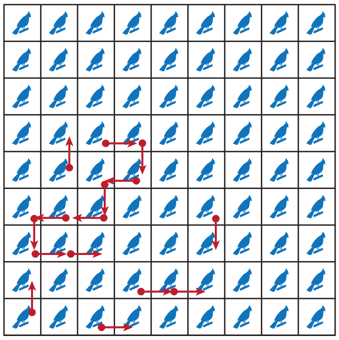

Figure 1.4. Each pigeonhole in a 9×9 array has one pigeon. All simultaneously move to another pigeonhole that is immediately above, below, to the left, or to the right of its current hole. Must some pigeonhole wind up with two pigeons?

1.5. Show that in any group of people, two of them have the same number of friends in the group. (Some important assumptions here: no one is a friend of him- or herself, and friendship is *symmetrical*—if A is a friend of B, then B is a friend of A.)

1.6. Given any five points on a sphere, show that four of them must lie within a closed hemisphere, where "closed" means that the hemisphere includes the circle that divides it from the other half of the sphere. *Hint:* Given any two points on a sphere, one can always draw a "great circle" between them, which has the same circumference as the equator of the sphere.

1.7. Show that in any group of 25 people, some three of them must have birthdays in the same month.

1.8. A collection of coins contains six different denominations: pennies, nickels, dimes, quarters, half-dollars, and dollars. How many coins must the collection contain to guarantee that at least 100 of the coins are of the same denomination?

1.9. Twenty-five people go to daily yoga classes at the same gym, which offers eight classes every day. Each attendee wears either a blue, red, or green shirt to class. Show that on a given day, there is at least one class in which two people are wearing the same color shirt.

1.10. Show that if four distinct integers are chosen between 1 and 60 inclusive, some two of them must differ by at most 19.

1.11. Find a k such that the product of the first k primes, plus 1, is not prime, but has a prime factor larger than any of the first k primes. (There is no trick for solving this. You just have to try various possibilities!)

1.12. Show that in any set of 9 positive integers, some two of them share all of their prime factors that are less than or equal to 5.

1.13. A *hash function* from strings to numbers derives a numerical hash value $h(s)$ from a text string s; for example, by adding up the numerical codes for the characters in s, dividing by a prime number p, and keeping just the remainder. The point of a hash function is to yield a reproducible result (calculating $h(s)$ twice for the same string s yields the same numerical value) and to make it likely that the hash values for different strings will be spread out evenly across the possible hash values (from 0 to $p - 1$). If the hash function has identical hash values for two different strings, then these two strings are said to *collide* on that hash value. We count the number of *collisions* on a hash value as 1 less than the number of strings that have that hash value, so if 2 strings have the same hash value there is 1 collision on that hash value. If there are m strings and p possible hash values, what is the minimum number of collisions that must occur on the hash value with the most collisions? The maximum number of collisions that might occur on some hash value?

INDEX