Counting I

20480135385502964448038	31710048321735013941130 17	57632573310834796474093 98	82473310000429953116460 21
48944599186691567624099 2	32082344215973686470192 65	58009491235489891226286 63	84962439971234759227663 10
10826620324303796513709 81	34372546563551578648691 13	60429008011992802180260 01	85183991406760026607474 77
11784808947697061789949 93	35748833903058653923711 365	61161717891377378967014 05	85436912834701914523337 63
12531273516832396938513 27	36449099460404801899691 49	61448689730015823697235 12	86753092583741370924613 52
13015051292340778110690 11	37900441327370840944172 46	62473145938511692347461 52	86943211123639968672966 65
13115671111438664338821 94	38703321274379713553228 15	68144289442668749634882 74	87723212036084772458511 54
14700294527212035876862 14	40805058045778014513631 00	68708529455438868491478 81	87914221617225825463410 91
15782710472862574994338 86	41672834610257023481249 20	69149555081209500937323 97	90626280245921262839732 85
16382439218521762431923 54	42359968311237777882112 49	69496324513659871524235 41	91378455669255263498977 94
17635802191319859631023 65	46709394457494390421112 20	71282111436136198284156 50	91537629668031892919344 19
18262277956018422310296 94	48153793518653842796134 27	71739200836518623079253 94	92708801940776364069842 49
18439718626751020372014 20	48370529482129226044421 90	72156548742117556762205 87	93243014807221034903792 04
23969511937221345261772 37	51063894238550185506715 30	72569328471643910402330 50	94360908321466951471405 81
27813945682685998010963 54	51423681920047692180699 10	73328226570752354316203 17	94753081597345382490132 38
27966051967136104054080 19	51812340961301440840418 56	74264418295415734449641 39	94923766239174869749232 02
29310163947619752631903 47	51982673981256179943913 48	76321981265318093271863 21	95119725587798802882529 79
29334580582944051551972 96	53175929403162312197583 72	77121544322119128823105 11	96024134246191871125522 64
30755144104909759203153 48	53843581267717941283569 47	78589186642402623566100 10	96312171149061292194611 11
31114749852527934528600 17	54392117122489019954234 41	78981567867632129631786 79	99081898531027533359813 19
31456215879361201184387 01	56103798260928381927604 58	81475910170375733378486 16	99132374763417642998139 87
31489012556288811031985 49	56323175554652286776760 44	81494367168713711619320 35	
31576931053251112843219 93	56921683746370196174237 12	81760638316825365713067 91	

Two different subsets of the ninety 25-digit numbers shown above have the same sum. For example, maybe the sum of the numbers in the first column is equal to the sum of the numbers in the second column. Can you find two such subsets? This is a very difficult computational problem. But we'll prove that such subsets must exist! This is the sort of weird conclusion one can reach by tricky use of counting, the topic of this chapter.

Counting seems easy enough: 1, 2, 3, 4, etc. This explicit approach works well for counting simple things, like your toes, and for extremely complicated things for which there's no identifiable structure. However, subtler methods can help you count many things in the vast middle ground, such as:

- The number of different ways to select a dozen doughnuts when there are five varieties available.
- The number of 16-bit numbers with exactly 4 ones.

Counting is useful in computer science for several reasons:

- Determining the time and storage required to solve a computational problem— a central objective in computer science— often comes down to solving a counting problem.

- Counting is the basis of probability theory, which in turn is perhaps the most important topic this term.

- Two remarkable proof techniques, the "pigeonhole principle" and "combinatorial proof", rely on counting. These lead to a variety of interesting and useful insights.

We're going to present a lot of rules for counting. These rules are actually theorems, but we're generally not going to prove them. Our objective is to teach you counting as a practical skill, like integration. And most of the rules seem "obvious" anyway.

1 Counting One Thing by Counting Another

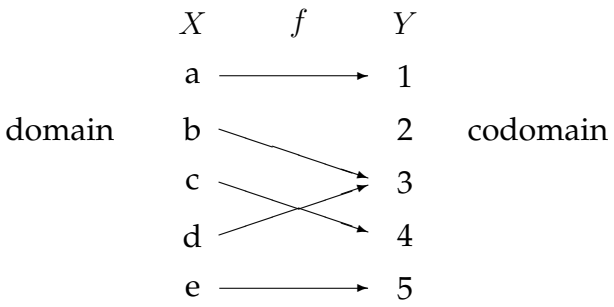
How do you count the number of people in a crowded room? We could count heads, since for each person there is exactly one head. Alternatively, we could count ears and divide by two. Of course, we might have to adjust the calculation if someone lost an ear in a pirate raid or someone was born with three ears. The point here is that we can often *count one thing by counting another*, though some fudge factors may be required. This is the central theme of counting, from the easiest problems to the hardest.

In more formal terms, every counting problem comes down to determining the size of some set. The ***size*** or ***cardinality*** of a set S is the number of elements in S and is denoted $|S|$. In these terms, we're claiming that we can often *find the size of one set S by finding the size of a related set T.* We already have a mathematical tool for relating one set to another: relations. Not surprisingly, a particular kind of relation is at the heart of counting.

1.1 Functions

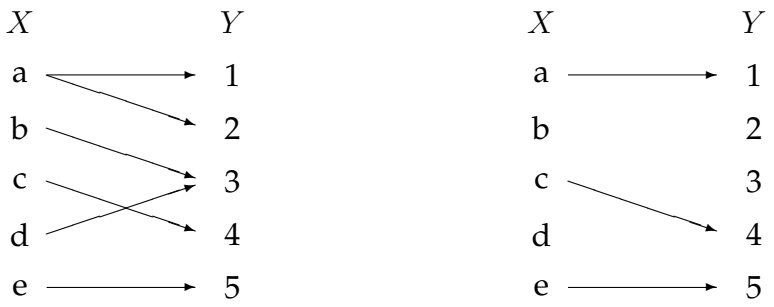
Functions like $f(x) = x^2 + 1$ and $g(x) = \tan^{-1}(x)$ are surely quite familiar from calculus. We're going to use functions for counting as well, but in a way that might not be familiar. Instead of using functions that map one real number to another, we'll use functions that relate elements of finite sets.

In order to count accurately, we need to carefully define the notion of a function. Formally, a ***function*** $f : X \to Y$ is a relation between two sets, X and Y, that relates *every* element of X to *exactly one* element of Y. The set X is called the ***domain*** of the function f, and Y is called the ***codomain***. Here is an illustration of a function:



In this example, the domain is the set $X = \{a, b, c, d, e\}$ and the range is the set $Y = \{1, 2, 3, 4, 5\}$. Related elements are joined by an arrow. This relation is a function because *every* element on the left is related to *exactly one* element on the right. In graph-theoretic terms, this is a function because every element on the left has degree *exactly 1*. If x is an element of the domain, then the related element in the codomain is denoted $f(x)$. We often say f ***maps*** x to $f(x)$. In this case, f maps a to 1, b to 3, c to 4, d to 3 and e to 5. Equivalently, $f(a) = 1$, $f(b) = 3$, $f(c) = 4$, $f(d) = 3$, and $f(e) = 5$.

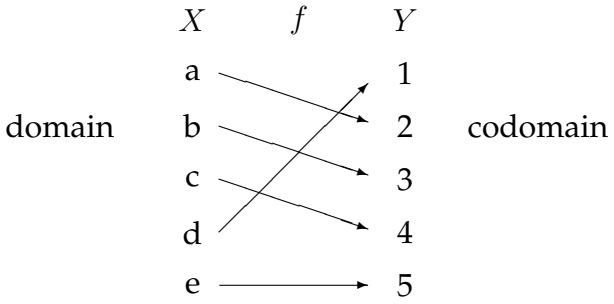
The relations shown below are *not* functions:



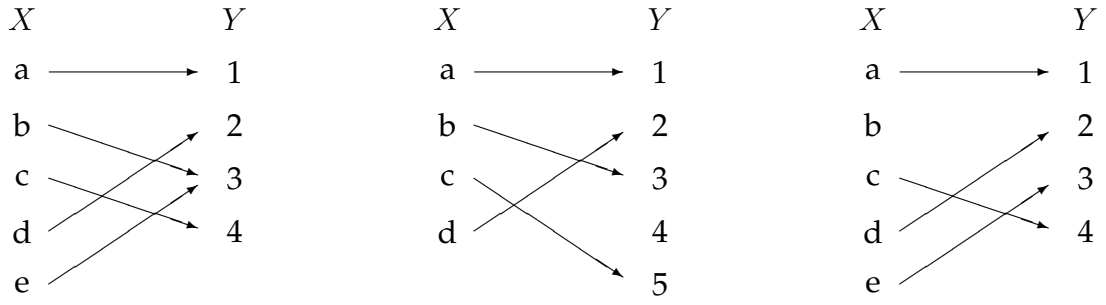
The relation on the left is not a function because a is mapped to *two* elements of the codomain. The relation on the right is not a function because b and d are mapped to *zero* elements of the codomain. Tsk-tsk!

1.2 Bijections

The definition of a function is rather asymmetric; there are restrictions on elements in the domain, but not on elements in the codomain. A ***bijection*** or ***bijective function*** is a function $f : X \to Y$ that maps exactly one element of the domain to each element of the codomain. In graph terms, *both* domain and codomain elements are required to have degree exactly 1 in a bijective function. Here is an example of a bijection:



In contrast, these relations that are *not* bijections:



The first function is not bijective because 3 is related to *two* elements of the domain. The second function is not bijective because 4 is related to *zero* elements of the domain. The last relation is not even a function, because b is associated with zero elements of the domain.

Bijective functions are also found in the more-familiar world of real-valued functions. For example, $f(x) = 6x+5$ is a bijective function with domain and codomain $\mathbb{R}$. For every real number y in the codomain, $f(x) = y$ for exactly one real number x in the domain. On the other hand, $f(x) = x^2$ is not a bijective function. The number 4 in the codomain is related to both 2 and -2 in the domain.

1.3 The Bijection Rule

If we can pair up all the girls at a dance with all the boys, then there must be an equal number of each. This simple observation generalizes to a powerful counting rule:

Rule 1 (Bijection Rule). *If there exists a bijection $f : A \to B$, then $|A| = |B|$.*

In the example, A is the set of boys, B is the set of girls, and the function f defines how they are paired.

The Bijection Rule acts as a magnifier of counting ability; if you figure out the size of one set, then you can immediately determine the sizes of many other sets via bijections.

For example, let's return to two sets mentioned earlier:

$A =$ all ways to select a dozen doughnuts when five varieties are available
$B =$ all 16-bit sequences with exactly 4 ones

Let's consider a particular element of set A:

$$\underbrace{0\,0}_{\text{chocolate}} \quad \underbrace{}_{\text{lemon-filled}} \quad \underbrace{0\,0\,0\,0\,0\,0}_{\text{sugar}} \quad \underbrace{0\,0}_{\text{glazed}} \quad \underbrace{0\,0}_{\text{plain}}$$

We've depicted each doughnut with a 0 and left a gap between the different varieties. Thus, the selection above contains two chocolate doughnuts, no lemon-filled, six sugar, two glazed, and two plain. Now let's put a 1 into each of the four gaps:

$$\underbrace{0\,0}_{\text{chocolate}} \;1\; \underbrace{}_{\text{lemon-filled}} \;1\; \underbrace{0\,0\,0\,0\,0\,0}_{\text{sugar}} \;1\; \underbrace{0\,0}_{\text{glazed}} \;1\; \underbrace{0\,0}_{\text{plain}}$$

We've just formed a 16-bit number with exactly 4 ones— an element of B!

This example suggests a bijection from set A to set B: map a dozen doughnuts consisting of:

c chocolate, l lemon-filled, s sugar, g glazed, and p plain

to the sequence:

$$\underbrace{0\ldots0}_{c}\;1\;\underbrace{0\ldots0}_{l}\;1\;\underbrace{0\ldots0}_{s}\;1\;\underbrace{0\ldots0}_{g}\;1\;\underbrace{0\ldots0}_{p}$$

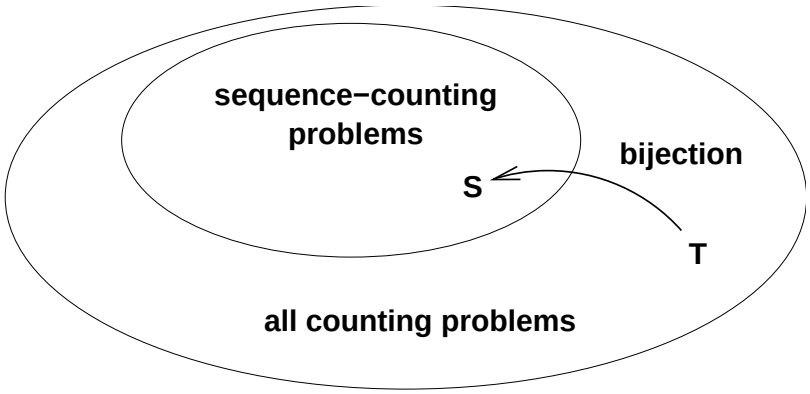
The resulting sequence always has 16 bits and exactly 4 ones, and thus is an element of B. Moreover, the mapping is a bijection; every such bit sequence is mapped to by exactly one order of a dozen doughnuts. Therefore, $|A| = |B|$ by the Bijection Rule!

This demonstrates the magnifying power of the bijection rule. We managed to prove that two very different sets are actually the same size— even though we don't know exactly how big either one is. But as soon as we figure out the size of one set, we'll immediately know the size of the other.

This particular bijection might seem frighteningly ingenious if you've not seen it before. But you'll use essentially this same argument over and over and over, and soon you'll consider it boringly routine.

1.4 Sequences

The Bijection Rule lets us count one thing by counting another. This suggests a general strategy: get really good at counting just a *few* things and then use bijections to count *everything else*:



This is precisely the strategy we'll follow. In particular, we'll get really good at counting *sequences*. When we want to determine the size of some other set T, we'll find a bijection from T to a set of sequences S. Then we'll use our super-ninja sequence-counting skills to determine $|S|$, which immediately gives us $|T|$. We'll need to hone this idea somewhat as we go along, but that's pretty much the plan!

In order to pull this off, we need to clarify some issues concerning sequences and sets. Recall that a ***set*** is an unordered collection of distinct elements. A set is often represented by listing its elements inside curly-braces. For example, $\{a, b, c\}$ is a set, and $\{c, b, a\}$ is another way of writing the same set. On the other hand, $\{a, b, a\}$ is not a set, because element a appears twice.

On the other hand, a ***sequence*** is an ordered collection of elements (called ***components*** or ***terms***) that are not necessarily distinct. A sequence is often written by listing the terms inside parentheses. For example, (a, b, c) is a sequence, and (c, b, a) is a different sequence. Furthermore (a, b, a) is a perfectly valid three-term sequence.

The distinction between sets and sequences is crucial for everything that follows. If you don't keep the distinction clear in your mind, you're doomed!

2 Two Basic Counting Rules

We'll harvest our first crop of counting problems with two basic rules.

2.1 The Sum Rule

Linus allocates his big sister Lucy a quota of 20 crabby days, 40 irritable days, and 60 generally surly days. On how many days can Lucy be out-of-sorts one way or another? Let set C be her crabby days, I be her irritable days, and S be the generally surly. In these terms, the answer to the question is $|C \cup I \cup S|$. Now assuming that she is permitted at most one bad quality each day, the size of this union of sets is given by the Sum Rule:

Rule 2 (Sum Rule). *If $A_1, A_2, \ldots, A_n$ are disjoint sets, then:*

$$|A_1 \cup A_2 \cup \ldots \cup A_n| = |A_1| + |A_2| + \ldots + |A_n|$$

Thus, according to Linus' budget, Lucy can be out-of-sorts for:

$$\begin{aligned} |C \cup I \cup S| &= |C| \cup |I| \cup |S| \\ &= 20 + 40 + 60 \\ &= 120 \text{ days} \end{aligned}$$

Notice that the Sum Rule holds only for a union of *disjoint* sets. Finding the size of a union of intersecting sets is a more complicated problem that we'll take up later.

2.2 The Product Rule

The product rule gives the size of a product of sets. Recall that if $P_1, P_2, \ldots, P_n$ are sets, then

$$P_1 \times P_2 \times \ldots \times P_n$$

is the set of all sequences whose first term is drawn from P_1, second term is drawn from P_2 and so forth.

Rule 3 (Product Rule). *If $P_1, P_2, \ldots P_n$ are sets, then:*

$$|P_1 \times P_2 \times \ldots \times P_n| = |P_1| \cdot |P_2| \cdots |P_n|$$

Unlike the sum rule, the product rule does not require the sets $P_1, \ldots, P_n$ to be disjoint. For example, suppose a *daily diet* consists of a breakfast selected from set B, a lunch from set L, and a dinner from set D:

$$\begin{aligned} B &= \{\text{pancakes, bacon and eggs, bagel, Doritos}\} \\ L &= \{\text{burger and fries, garden salad, Doritos}\} \\ D &= \{\text{macaroni, pizza, frozen burrito, pasta, Doritos}\} \end{aligned}$$

Then $B \times L \times D$ is the set of all possible daily diets. Here are some sample elements:

$$\begin{gathered} (\text{pancakes, burger and fries, pizza}) \\ (\text{bacon and eggs, garden salad, pasta}) \\ (\text{Doritos, Doritos, frozen burrito}) \end{gathered}$$

The Product Rule tells us how many different daily diets are possible:

$$\begin{aligned} |B \times L \times D| &= |B| \cdot |L| \cdot |D| \\ &= 4 \cdot 3 \cdot 5 \\ &= 60 \end{aligned}$$

2.3 Putting Rules Together

Few counting problems can be solved with a single rule. More often, a solution is a flury of sums, products, bijections, and other methods. Let's look at some examples that bring more than one rule into play.

Passwords

The sum and product rules together are useful for solving problems involving passwords, telephone numbers, and license plates. For example, on a certain computer system, a valid password is a sequence of between six and eight symbols. The first symbol must be a letter (which can be lowercase or uppercase), and the remaining symbols must be either letters or digits. How many different passwords are possible?

Let's define two sets, corresponding to valid symbols in the first and subsequent positions in the password.

$$F = \{a, b, \ldots, z, A, B, \ldots, Z\}$$
$$S = \{a, b, \ldots, z, A, B, \ldots, Z, 0, 1, \ldots, 9\}$$

In these terms, the set of all possible passwords is:

$$(F \times S^5) \cup (F \times S^6) \cup (F \times S^7)$$

Thus, the length-six passwords are in set $F \times S^5$, the length-seven passwords are in $F \times S^6$, and the length-eight passwords are in $F \times S^7$. Since these sets are disjoint, we can apply the Sum Rule and count the total number of possible passwords as follows:

$$\begin{aligned}
\left|(F \times S^5) \cup (F \times S^6) \cup (F \times S^7)\right| &= \left|F \times S^5\right| + \left|F \times S^6\right| + \left|F \times S^7\right| && \text{Sum Rule} \\
&= |F| \cdot |S|^5 + |F| \cdot |S|^6 + |F| \cdot |S|^7 && \text{Product Rule} \\
&= 52 \cdot 62^5 + 52 \cdot 62^6 + 52 \cdot 62^7 \\
&\approx 1.8 \cdot 10^{14} \text{ different passwords}
\end{aligned}$$

Subsets of an n-element Set

How many different subsets of an n element set X are there? For example, the set $X = \{x_1, x_2, x_3\}$ has eight different subsets:

$$\begin{array}{cccc} \{\} & \{x_1\} & \{x_2\} & \{x_1, x_2\} \\ \{x_3\} & \{x_1, x_3\} & \{x_2, x_3\} & \{x_1, x_2, x_3\} \end{array}$$

There is a natural bijection from subsets of X to n-bit sequences. Let $x_1, x_2, \ldots, x_n$ be the elements of X. Then a particular subset of X maps to the sequence $(b_1, \ldots, b_n)$

where $b_i = 1$ if and only if x_i is in that subset. For example, if $n = 10$, then the subset $\{x_2, x_3, x_5, x_7, x_{10}\}$ maps to a 10-bit sequence as follows:

$$\begin{array}{rccccccccccc} \text{subset:} & \{ & & x_2, & x_3, & & x_5, & & x_7, & & & x_{10} & \} \\ \text{sequence:} & (& 0, & 1, & 1, & 0, & 1, & 0, & 1, & 0, & 0, & 1 &) \end{array}$$

We just used a bijection to transform the original problem into a question about sequences—*exactly according to plan!* Now if we answer the sequence question, then we've solved our original problem as well.

But how many different n-bit sequences are there? For example, there are 8 different 3-bit sequences:

$$\begin{array}{cccc} (0,0,0) & (0,0,1) & (0,1,0) & (0,1,1) \\ (1,0,0) & (1,0,1) & (1,1,0) & (1,1,1) \end{array}$$

Well, we can write the set of all n-bit sequences as a product of sets:

$$\underbrace{\{0,1\} \times \{0,1\} \times \ldots \times \{0,1\}}_{n \text{ terms}} = \{0,1\}^n$$

Then Product Rule gives the answer:

$$\begin{aligned} |\{0,1\}^n| &= |\{0,1\}|^n \\ &= 2^n \end{aligned}$$

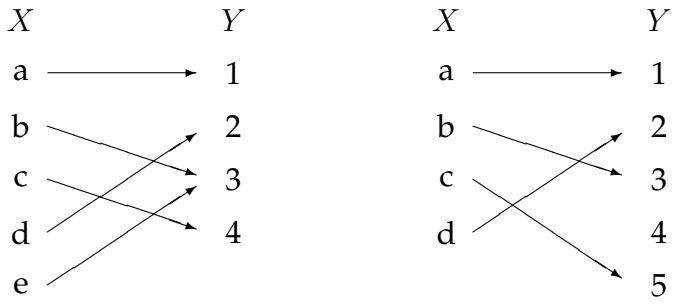
This means that the number of subsets of an n-element set X is also 2^n. We'll put this answer to use shortly.

3 More Functions: Injections and Surjections

Bijective functions are incredibly powerful counting tools. A few other kinds of functions are useful as well; we'll look at two now and one more next time. A function $f : X \to Y$ is:

- ***surjective*** if every element of Y is mapped to *at least once*
- ***injective*** if every element of Y is mapped to *at most once*
- ***bijective*** if every element of Y is mapped to *exactly once*

We've repeated the definition of a bijective function for comparison. Notice that these definitions immediately imply that a function is bijective if and only if it is both injective and surjective. Now the names "surjective" and "injective" are hopelessly unmemorable and nondescriptive. Some people prefer the terms ***onto*** and ***into***, respectively, perhaps on the grounds that these are hopelessly unmemorable and nondescriptive— but shorter. Anyway, here are a couple examples:



The function on the left is surjective (every element on the right is mapped to at least once), but not injective (element 3 is mapped to twice). The function on the right is injective (every element is mapped to at most once), but not surjective (element 4 is mapped to zero times).

Earlier, we observed that two sets are the same size if there is a bijection between them. Similarly, surjections and injections imply certain size relationships between sets.

Rule 4 (Mapping Rule).

1. *If $f : X \to Y$ is surjective, then $|X| \geq |Y|$.*
2. *If $f : X \to Y$ is injective, then $|X| \leq |Y|$.*
3. *If $f : X \to Y$ is bijective, then $|X| = |Y|$.*

3.1 The Pigeonhole Principle

Here is an old puzzle:

> A drawer in a dark room contains red socks, green socks, and blue socks. How many socks must you withdraw to be sure that you have a matching pair?

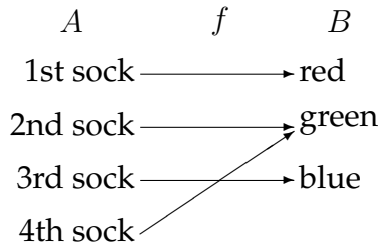
For example, picking out three socks is not enough; you might end up with one red, one green, and one blue. The solution relies on the Pigeonhole Principle, which is a friendly name for the contrapositive of part (2) of the Mapping Rule. Let's write it down:

> If $|X| > |Y|$, then no function $f : X \to Y$ is injective.

Now let's rewrite this a second time to eliminate the word "injective" since, by now, there's not a ghost of a chance that you remember what that means:

Rule 5 (Pigeonhole Principle). *If $|X| > |Y|$, then for every function $f : X \to Y$ there exist two different elements of X that are mapped to the same element of Y.*

Perhaps the relevance of this abstract mathematical statement to selecting footwear under poor lighting conditions is not obvious. However, let A be the set of socks you pick out, let B be the set of colors available, and let f map each sock to its color. The Pigeonhole Principle says that if $|A| > |B| = 3$, then at least two elements of A (that is, at least two socks) must be mapped to the same element of B (that is, the same color). For example, one possible mapping of four socks to three colors is shown below.



Therefore, four socks are enough to ensure a matched pair.

Not surprisingly, the pigeonhole principle is often described in terms of pigeons: if more than n pigeons fly into n pigeonholes, then at least two pigeons must fly into some hole. In this case, the pigeons form set A, the pigeonholes are set B, and f describes the assignment of pigeons to pigeonholes.

Mathematicians have come up with many ingenious applications for the pigeonhole principle. If there were a cookbook procedure for generating such arguments, we'd give it to you. Unfortunately, there isn't one. One helpful tip, though: when you try to solve a problem with the pigeonhole principle, the key is to clearly identify three things:

1. The set A (the pigeons).
2. The set B (the pigeonholes).
3. The function f (the rule for assigning pigeons to pigeonholes).

Hairs on Heads

There are a number of generalizations of the pigeonhole principle. For example:

Rule 6 (Generalized Pigeonhole Principle). *If $|X| > k \cdot |Y|$, then every function $f : X \to Y$ maps at least $k + 1$ different elements of X to the same element of Y.*

For example, if you pick two people at random, surely they are extremely unlikely to have *exactly* the same number of hairs on their heads. However, in the remarkable city of Boston, Massachusetts there are actually *three* people who have exactly the same number of hairs! Of course, there are many bald people in Boston, and they all have zero hairs. But I'm talking about non-bald people.

Boston has about 500,000 non-bald peole, and the number of hairs on a person's head is at most 200,000. Let A be the set of non-bald people in Boston, let $B = \{1, \ldots, 200,000\}$, and let f map a person to the number of hairs on his or her head. Since $|A| > 2\,|B|$, the Generalized Pigeonhole Principle implies that at least three people have exactly the same number of hairs. I don't know who they are, but I know they exist!

Subsets with the Same Sum

We asserted that two different subsets of the ninety 25-digit numbers listed on the first page have the same sum. This actually follows from the Pigeonhole Principle. Let A be the collection of all subsets of the 90 numbers in the list. Now the sum of any subset of numbers is at most $90 \cdot 10^{25}$, since there are only 90 numbers and every 25-digit number is less than 10^{25}. So let B be the set of integers $\{0, 1, \ldots, 90 \cdot 10^{25}\}$, and let f map each subset of numbers (in A) to its sum (in B).

We proved that an n-element set has 2^n different subsets. Therefore:

$$\begin{aligned} |A| &= 2^{90} \\ &\geq 1.237 \times 10^{27} \end{aligned}$$

On the other hand:

$$\begin{aligned} |B| &= 90 \cdot 10^{25} + 1 \\ &\leq 0.901 \times 10^{27} \end{aligned}$$

Both quantities are enormous, but $|A|$ is a bit greater than $|B|$. This means that f maps at least two elements of A to the same element of B. In other words, by the Pigeonhole Principle, two different subsets must have the same sum!

Notice that this proof gives no indication *which* two sets of numbers have the same sum. This frustrating variety of argument is called a ***nonconstructive proof***.

Sets with Distinct Subset Sums

How can we construct a set of n positive integers such that all its subsets have *distinct* sums? One way is to use powers of two:

$$\{1, 2, 4, 8, 16\}$$

This approach is so natural that one suspects all other such sets must involve larger numbers. (For example, we could safely replace 16 by 17, but not by 15.) Remarkably, there are examples involving *smaller* numbers. Here is one:

$$\{6, 9, 11, 12, 13\}$$

One of the top mathematicans of the century, Paul Erdős, conjectured in 1931 that there are no such sets involving *significantly* smaller numbers. More precisely, he conjectured that the largest number must be $\Omega(2^n)$. He offered $500 to anyone who could prove or disprove his conjecture, but the problem remains unsolved.