

## Basic Counting, Pigeonholing, Permutations

### 1 Counting by Matching

Counting is a theme throughout discrete mathematics: how many leaves in a tree, minimal colorings of a graph, trees with a given set of vertices, five-card hands in a deck of fifty-two, consistent rankings of players in a tournament, stable marriages given boy's and girl's preferences, and so on.

A good way to count things is to match up things to be counted with other things that we know how to count. We saw an example of this early in the term when we counted the size of a powerset of a set of size  $n$  by finding an exact matching between elements of the powerset and the  $2^n$  binary strings of length  $n$ .

The matching doesn't have to be exact, *i.e.*, a bijection, to be informative. For example, suppose we want to determine the cardinality of the set of watches in the 6.042 classroom on a typical day. The set of watches can be correlated with the set of people in the room; specifically, for each person there is at most one watch (at least, let's assume this). Now we know something about the cardinality of the set of students, since there are only 146 people signed up for 6.042. There are also three lecturers and eight TA's, and these would typically be the only nonstudents in the room. So we can conclude that there are *at most* 157 watches in the classroom on a typical day.

This type of argument is very simple, but also quite powerful. We will see how to use such simple arguments to prove results that are hard to obtain any other way.

### 2 Matchings as Bijections

The "matching up" we talked about more precisely refers to finding injections, surjections, and bijections between things we want to count and things to be counted. The following Theorem formally justifies this kind of counting:

**Theorem 2.1.** *Let  $A$  and  $B$  be finite sets and  $f : A \rightarrow B$  be a function. If*

1.  *$f$  is a bijection, then  $|A| = |B|$ ,*
2.  *$f$  is an injection, then  $|A| \leq |B|$ ,*
3.  *$f$  is a surjection, then  $|A| \geq |B|$ .*

This is one of those theorems that is so fundamental that it's not clear what simpler axioms are appropriate to use in proving it. In fact, we can't prove it yet, because we haven't defined the concept that it's all about, namely, the size or *cardinality*,  $|A|$ , of a finite set,  $A$ . Intuitively, a set,  $A$ , has  $n$  elements if it equals  $\{a_1, a_2, \dots, a_n\}$  where the  $a_i$  are all different. Now ellipsis is dangerous, so we should avoid it in a definition this basic. What is the notation " $a_1, a_2, \dots, a_n$ " intended to convey? It means that there is a first element,  $a_1$ , and a second element,  $a_2$ , and in general, given any  $i \leq n$ , there is an  $i$ th element  $a_i$ . Also, all the  $a_i$ 's for different  $i$ 's are different. This explains how we arrive at a rigorous definition:

**Definition 2.2.** A set  $A$  has *cardinality*  $n \in \mathbb{N}$ , in symbols,  $|A| = n$ , iff there is a **bijection** from  $\{1, 2, \dots, n\}$  to  $A$ . The special case when  $n = 0$  is that  $|\emptyset| = 0$ . A set is *finite* iff it has cardinality  $n$  for some  $n \in \mathbb{N}$ .

With this definition, we could prove Theorem 2.1 by appeal to basic properties of functions and natural numbers. For example, if  $f : A$  to  $B$  is a bijection, and  $|A| = n$ , then we can prove that  $|B| = n$  as follows: since  $|A| = n$ , there is a bijection  $g : \{1, 2, \dots, n\}$  to  $A$ . Then  $f \circ g$  is a bijection from  $\{1, 2, \dots, n\}$  to  $B$ , so by definition,  $|B| = n$ .

Here we used the fact that the composition of bijections is a bijection. This fact itself follows just from the logical properties of equality and the definition of a bijection; it does not even depend on any properties of numbers. So we can say that we proved part 1 of Theorem 2.1 from more fundamental mathematical concepts. The other two parts can be proved using similar properties of functions along with ordinary induction, but we'll skip them: the proofs are exercises in formal logic that are not very informative about counting.

Notice that the condition in Theorem 2.1 that  $A$  and  $B$  are *finite* sets is important. It's not even clear what the size of an infinite set ought to be, or whether it's possible for one infinite set to be "larger" than another. We'll avoid this issue in these notes by only counting the sizes of finite sets.

## 2.1 Counting Functions

The bijection between length  $n$  binary strings and a powerset can be generalized to help in counting the number of functions from one set to another:

**Question:** How many different functions are there from finite set  $A$  to finite set  $B$ ?

**Theorem 2.3.** If  $A$  and  $B$  are finite sets, with  $|A| = n$  and  $|B| = m$ , then the cardinality of the set of functions from  $A$  to  $B$  is  $m^n$ .

*Proof.* We will use a bijection from  $\{f \mid f : A \rightarrow B\}$  to  $\{s \mid s \text{ is a length } n \text{ string of elements from } B\}$ . The mapping is  $f \mapsto s_f$  where  $s_f ::= f(a_1)f(a_2) \cdots f(a_n)$ , i.e., the value of the  $i$ th position in  $s_f$  is equal to  $f(a_i)$ .

We will prove that this mapping is a bijection. First we prove that the mapping is injective (one-to-one) by contradiction. Suppose that  $f \neq g$  and  $s_f = s_g$ . But  $f \neq g$  implies that there exists an  $i \in \mathbb{N}$ , where  $1 \leq i \leq n$ , we have  $f(a_i) \neq g(a_i)$ . But that implies that the  $i$ th position in  $s_f$  and  $s_g$  are different, which is a contradiction.

Next we prove that the mapping is surjective (onto), i.e., that every length  $n$  string,  $s$ , of elements from  $B$  equals  $s_f$  for some function  $f : A \rightarrow B$ . Denote the  $i$ th position in such a string,  $s$ , by  $s[i]$ .

Now define a function  $f$  by the rule that  $f(a_i) ::= s[i]$ , for  $1 \leq i \leq n$ . This defines the required  $f : A \rightarrow B$  such that  $s_f = s$ .

Since this mapping is a bijection, we know that the number of functions from  $A$  to  $B$  is equal to the number of strings of length  $n$  from the elements of  $B$ . In section 5 we will see that the number of such strings is  $m^n$  by the Product Rule.  $\square$

### 3 The Pigeonhole Principle

Theorem 2.1 part 2 tells us that if there is an injection from  $A$  to  $B$ , then  $|A| \leq |B|$ . The contrapositive of this statement is that if  $|A| > |B|$ , and  $f$  is a function from  $A$  to  $B$ , then  $f$  is not an injection.

**Corollary 3.1.** *Let  $A$  and  $B$  be finite sets. If  $|A| > |B|$  and  $f : A \rightarrow B$ , then there exist distinct elements  $a$  and  $a'$  in  $A$  such that  $f(a) = f(a')$ .*

This Corollary is known as the *Pigeonhole Principle* because it can be paraphrased as:

**The Pigeonhole Principle:** If there are more pigeons than pigeonholes, then there must be at least two pigeons in one hole.

*Proof.* Let  $A$  be the set of “pigeons”, let  $B$  be the set of “holes”, and let the function  $f : A \rightarrow B$  define the assignment of pigeons to holes. Since  $|A| > |B|$ , Corollary 3.1 implies that there exist two distinct pigeons,  $a \neq a'$ , assigned to the same hole,  $f(a)$ .  $\square$

As a trivial application of the Pigeonhole Principle, suppose that there are three people in a room. The pigeonhole principle implies that two have the same gender. In this case, the “pigeons” are the three people and the “pigeonholes” are the two possible genders, male and female. Since there are more pigeons than holes, two pigeons must be in the same pigeonhole; that is, two people must have the same gender.

**Claim 3.2.** *In New York (City) there live at least two people with the same number of hairs.*

*Proof (found on the web).* I ran experiments with members of my family. My teenage son secured himself the highest marks sporting, in my estimate, about 900 hairs per square inch. Even assuming a pathological case of a 6 feet (two-sided) fellow 50 inch across, covered with hair head, neck, shoulders and so on down to the toes, the fellow would have somewhere in the vicinity of 7,000,000 hairs which is probably a very gross over-estimate to start with. The Hammond’s World Atlas I purchased some 15 years ago, estimates the population of the New York City between 7,500,000 and 9,000,000. The assertion therefore follows from the pigeonhole principle.  $\square$

The pigeonhole principle seems too obvious to be really useful, but the next two examples show how it gives short proofs of results that are difficult to obtain by other means.

### 3.1 Pigeonhole Principle Example: A Final Exam Question

A problem on an old final exam was to prove the following claim:

**Claim.** *In every set of 1000 integers, there are two integers  $x$  and  $y$  such that  $573 \mid (x - y)$ .*

At first glance, this looks very hard! Those 1000 numbers could be anything! Since there are no less than 1000 integer-valued variables here, even our old standby, induction, seems hopeless. Surprisingly, however, there is a short proof using the Pigeonhole Principle.

To apply the Pigeonhole Principle, we must identify two things: pigeons and holes. Furthermore, to prove anything with the Pigeonhole Principle, we must have more pigeons than holes. Since there are only two numbers mentioned in this problem, a natural thing to try is 1000 pigeons and 573 holes.

Under this interpretation, a pigeon is an integer, but what is a hole? Ideally, the existence of two pigeons and in the same hole should be equivalent to the existence of two numbers  $x$  and  $y$  such that  $573 \mid (x - y)$ . This suggests numbering the holes  $0, 1, \dots, 572$  and putting in hole  $n$  all integers congruent to  $n$  modulo 573. Now we can construct a proof:

*Proof.* Let  $S$  be a set of 1000 integers. Let  $M = \{0, 1, \dots, 572\}$ . Let  $f$  from  $S$  to  $M$  be the function defined by  $f(n) = n \bmod 573$ . Since  $|S| > |M|$ , Corollary 3.1 implies that there exist distinct elements  $x$  and  $y$  in  $S$  such that  $f(x) = f(y)$ . This means  $(x \bmod 573) = (y \bmod 573)$  and so  $573 \mid (x - y)$ .  $\square$

Really there was nothing special about the numbers 1000 and 573 other than the fact that  $1000 > 573$ . We could have made a stronger claim: if  $n > m$ , then in every set of  $n$  integers, there are two integers  $x$  and  $y$  such that  $m \mid (x - y)$ .

### 3.2 Example: Subsets of a List of Numbers

Show that any given 10 distinct positive numbers less than 100, that two completely different subsets sum to the same quantity.

The numbers all vary between 1 and 99. Therefore the maximum sum of any 10 chosen numbers is  $90 + 91 + 92 + \dots + 99 = 945$ . The number of different subsets of the 10 numbers is  $2^{10} - 1$  (excluding the null set) = 1023. We have 1023 pigeons and 945 holes. Using the pigeonhole principle, we can argue that two different subsets map to the same sum. If these subsets have a common number or numbers, we can always remove the common numbers to produce two completely different subsets that sum to the same quantity.

### 3.3 20 Questions and Binary Search

Here is a game. I think of an animal. You can ask me 20 questions that take a yes/no answer such as, "Is the animal bigger than a breadbox?" To win the game, you must ask a question like, "Is the animal a walrus?" or "Is the animal a zebra?" and receive a "yes" answer. In effect, you have 19 questions to determine which animal I am thinking of, and then you must use 1 question to confirm your guess.

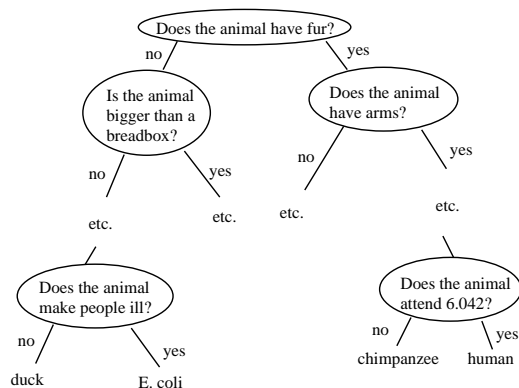


Figure 1: A strategy for the animal game can be represented by a depth-19 binary tree. Each internal node represents a yes/no question such as, “Does the animal have fur?” Each leaf node represents a final guess. A run of the algorithm corresponds to a path from the root to a leaf.

Suppose that I know a million animals. Can you always determine which animal I am thinking of? Any questioning strategy you use can be represented by a depth-19 binary tree as shown in Figure 1. Each internal node in the tree represents a question, and each leaf represents a final guess at my animal. A depth-19 binary tree can have at most  $2^{19} = 524,288$  leaves, and I can use any of a million animals. By the Pigeonhole Principle, at least two animals must be associated with some leaf in the tree; this implies that you cannot always determine which animal I am thinking of with only 19 questions. More generally, if I know  $n$  animals, then  $\lceil \log_2 n \rceil$  questions are necessary to always identify the one I’m thinking of; a binary tree of lower depth must have fewer than  $n$  leaves, and so some animals cannot be distinguished.

A similar argument applies to a binary search algorithm. In a binary search, we are looking for a particular item in a *sorted* list. We begin by comparing the middle element in the list to the item we are looking for. If our item precedes the middle element, then we continue the search recursively in the first half of the list. Similarly, if our item follows the middle element, then we search recursively in the second half of the list. For example, binary search could be applied to the animal game. You could sort the list of a million animals, pick out the middle one, and begin with a question like, “Does the animal alphabetically precede marmot?” Not surprisingly, given the similarity between the animal game and binary search, a Pigeonhole Principle argument shows that binary search requires at least  $\log n$  comparisons to find an item in an  $n$ -element list in the worst case.

### 3.4 Example: Weighing Coins

Now let’s consider the problem of identifying an off-weight counterfeit coin among a collection of coins using a balance scale. In this example, we’ll do a refined analysis using the Pigeonhole Principle.

Let’s consider 12 coins of which 11 have the same weight and a counterfeit one with a different weight. With three weighings on a balance scale, you must identify the counterfeit coin and determine whether it is heavier or lighter than the rest. (A balance scale has a left pan and a right pan. In a weighing you put some coins in each pan. The scale then reveals whether the left pan is heavier, the right pan is heavier, or the two are equal.)































