Combinatorics Through Guided Discovery[1]

Kenneth P. Bogart

March 19, 2002

[1]Development of this project was supported by National Science Foundation Grant Number DUE-0087466

Contents

Preface

These notes are for the use of students in Math 28 in the winter term of 2002. The notes consist almost entirely of problems. Some of the problems are designed to lead you to think about a concept, others are designed to help you figure out a concept and state a theorem about it, while still others ask you to prove the theorem. From time to time there is a discussion that pulls together some of the things you have learned or introduces a new idea for you to work with. Many of the problems are designed to build up your intuition for how combinatorial mathematics works. There are problems that some people will solve quickly, and there are problems that will take days of thought for everyone. Probably the best way to use these notes is to work on a problem until you feel you are not making progress and then go on to the next one. The next chance you get, discuss the problem you are stymied on with other members of the class. Often you will all feel you've hit dead ends, but when you begin comparing notes and listening *carefully* to each other, you will see more than one approach to the problem and be able to make some progress. In fact, after comparing notes you may realize that there is more than one way to interpret the problem. In this case your first step should be to think together about what the problem is actually asking you to do. You may have learned in school that for every problem you are given, there is a method that has already been taught to you, and you are supposed to figure out which method applies and apply it. That is not the case here. Based on some simplified examples, you will discover the method for yourself. Later on, you may recognize a pattern that suggests you should try to use this method again. The point of learning from these notes is that you are learning how to discover ideas and methods for yourself, not that you are learning to apply methods that someone else has told you about.

One of the downsides of how we learn math in high school is that many of us come to believe that if we can't solve a problem in ten or twenty minutes, then we can't solve it at all. Many of these problems were first conceived and solved by professional mathematicians, and *they* spent days or weeks on them. How can you be expected to solve them at all then? You have a context in which to work, and even though some of the problems are so open ended that you go into them without any idea of the answer, the context and the leading examples that preceded them give you a structure to work with. That doesn't mean you'll get them right away, but you will find a real sense of satisfaction when you see what you can figure out with concentrated thought. Besides, you can get hints!

You should try to write up answers to all the problems that you work on. Then you should ask other members of the class who have written up their solutions to the same problems to read yours and really try to understand it. When you are reading someone else's solution, your goal should be to help them make their ideas clear to you. You do this by asking about things that you are having trouble reading. The point is not whether you can see how to do the problem after reading what your classmate has written, but whether you understand how your classmate has solved the problem, and how his or her words are explaining the solution to you. Everyone in the class is trying to learn to do combinatorial mathematics, and on tests and on problems your instructor chooses to grade, your explanations have to be clear. This means that the instructor has to be able to "get" your ideas and understand exactly what you are saying. One of the best ways of learning to write clearly is to have someone who is as easily confused as you are point out to you where it is hard to figure out what you mean. Your instructor is going to choose some of your solutions to read carefully and give you detailed feedback on. When you get this feedback, you should think it over carefully and then write the solution again! You may be asked not to have someone else read your solutions to some of these problems until your instructor has. This is so that the instructor can offer help which is aimed at your needs.

As you work on a problem, think about why you are doing what you are doing. Is it helping you? If your current approach doesn't feel right, try to see why. Is this a problem you can decompose into simpler problems? Can you see a way to make up a simple example, even a silly one, of what the problem is asking you to do? If a problem is asking you to do something for every value of an integer n, then what happens with simple values of n like 0, 1, and 2? Don't worry about making mistakes; it is often finding mistakes that leads mathematicians to their best insights. Above all, don't worry if you can't do a problem. Some problems are given as soon as there is one technique you've learned that might help do that problem. Later on there may be other techniques that you can bring back to that problem to try again. The notes have been designed this way on purpose. If you happen to get a hard problem with the bare minimum of tools, you will have accomplished much. As you go along, you will see your ideas appearing again later in other problems. On the other hand, if you don't get the problem the first time through, it will be nagging at you as you work on other things, and when you see the idea for an old problem in new work, you will know you are learning.

Above all, these notes are dedicated to the principle that doing mathematics is fun. As long as you know that some of the problems are going to require more than one attempt before you hit on the main idea, you can relax and enjoy your successes, knowing that as you work more and more problems and share more and more ideas, problems that seemed intractable at first become a source of satisfaction later on.

Chapter 1

What is Combinatorics?

Combinatorial mathematics arises from studying how we can *combine* objects into arrangements. For example, we might be combining sports teams into a tournament, samples of tires into plans to mount them on cars for testing, students into classes to compare approaches to teaching a subject, or members of a tennis club into pairs to play tennis. There are many questions one can ask about such arrangements of objects. Here we will focus on questions about how many ways we may combine the objects into arrangements of the desired type. Sometimes, though, we ask if an arrangement is possible (if we have ten baseball teams, and each team has to play each other team once, can we schedule all the games if we only have the fields available for forty games?). Sometimes we will ask if all the arrangements we might be able to make have a certain desirable property (Do all ways of testing 5 brands of tires on 5 different cars [with certain additional properties] compare each brand with each other brand on at least one common car?). Problems of these sorts come up throughout physics, biology, computer science, statistics, and many other subjects. However, to demonstrate all these relationships, we would have to take detours into all these subjects. While we will give some important applications, we will usually phrase our discussions around everyday experience and mathematical experience so that the student does not have to learn a new context before learning mathematics in context!

1.1 About These Notes

These notes are based on the philosophy that you learn the most about a subject when you are figuring it out directly for yourself, and learn the least when you are trying to figure out what someone else is saying about it. On the other hand, there is a subject called combinatorial mathematics, and that is what we are going to be studying, so we will have to tell you some basic facts. What we are going to try to do is to give you a chance to discover many of the interesting examples that usually appear as textbook examples and discover the principles that appear as textbook theorems. Your main activity will be solving problems designed to lead you to discover the basic principles of combinatorial mathematics. Some of the problems lead you through a new idea, some give you a chance to describe what you have learned in a sequence of problems, and some are quite challenging. When you find a problem challenging, don't give up on it, but don't let it stop you from going on with other problems. Frequently you will find an idea in a later problem that you can take back to the one you skipped over or only partly finished in order to finish it off. With that in mind, let's get started.

1.2 Basic Counting Principles

1. Five schools are going to send their baseball teams to a tournament, in which each team must play each other team exactly once. How many games are required?

2. Now some number n of schools are going to send their baseball teams to a tournament, and each team must play each other team exactly once. Let us think of the teams as numbered one through n.

   (a) How many games does team one have to play in?

   (b) How many games, other than the one with team one, does team two have to play in?

   (c) How many games, other than those with the first $i-1$ teams, does team i have to play in?

   (d) In terms of your answers to the previous parts of this problem, what is the total number of games that must be played?

3. One of the schools sending its team to the tournament has to send its players from some distance, and so it is making sandwiches for team members to eat along the way. There are three choices for the kind of bread and five choices for the kind of filling. How many different kinds of sandwiches are available?

4. An *ordered pair* (a, b) consists of two things we call a and b. We say a is the first member of the pair and b is the second member of the pair. If M is an m element set and N is an n-element set, how many ordered pairs are there whose first member is in M and whose second member is in N? Does this problem have anything to do with any of the previous problems?

5. Since a sandwich by itself is pretty boring, students from the school in Problem 3 are offered a choice of a drink (from among five different kinds), a sandwich, and a fruit (from among four different kinds). In how many ways may a student make a choice of the three items now?

6. The coach of the team in Problem 3 knows of an ice cream parlor along the way where she plans to stop to buy each team member a triple decker cone. There are 12 different flavors of ice cream, and triple decker cones are made in homemade waffle cones. Having chocolate ice cream as the bottom scoop is different from having chocolate ice cream as the top scoop. How many possible ice cream cones are going to be available to the team members? How many cones with three different kinds of ice cream will be available?

7. The idea of a function is ubiquitous in mathematics. A function f from a set S to a set T is a relationship between the two sets that associates exactly one member $f(x)$ of T with each element x in S. We will come back to the ideas of functions and relationships in more detail and from different points of view from time to time. However, the quick review above should let you answer these questions.

   (a) Using f, g, $\ldots$, to stand for the various functions, write down all the different functions you can from the set $\{1, 2\}$ to the set $\{a, b\}$. For example, you might start with $f(1) = a$, $f(2) = b$. How many functions are there from the set $\{1, 2\}$ to the set $\{a, b\}$?

(b) How many functions are there from the three element set $\{1, 2, 3\}$ to the two element set $\{a, b\}$?

(c) How many functions are there from the two element set $\{a, b\}$ to the three element set $\{1, 2, 3\}$?

(d) How many functions are there from a three element set to a 12 element set?

(e) The function f is called **one-to-one** or an *injection* if whenever x is different from y, $f(x)$ is different from $f(y)$. How many one-to-one functions are there from a three element set to a 12 element set?

(f) Explain the relationship between this problem and Problem 6.

8. A group of hungry team members in Problem 6 notices it would be cheaper to buy three pints of ice cream for them to split than than to buy a triple decker cone for each of them, and that way they would get more ice cream. They ask their coach if they can buy three pints of ice cream. In how many ways can they choose three pints of different flavors? In how many ways may they choose three pints if the flavors don't have to be different?

9. Two sets are said to be *disjoint* if they have no elements in common. For example, $\{1, 3, 12\}$ and $\{6, 4, 8, 2\}$ are disjoint, but $\{1, 3, 12\}$ and $\{3, 5, 7\}$ are not. Three or more sets are said to be *mutually disjoint* if no two of them have any elements in common. What can you say about the size of the union of a finite number of finite (mutually) disjoint sets? Does this have anything to do with any of the previous problems?

10. Disjoint subsets are defined in Problem 9. What can you say about the size of the union of m (mutually) disjoint sets, each of size n? Does this have anything to do with any of the previous problems?

1.2.1 The sum and product principles

These problems contain among them the kernels of many of the fundamental ideas of combinatorics. For example, with luck, you just stated the sum principle (illustrated in Figure 1.1), and product principle (illustrated in

Figure 1.1: The union of these two disjoint sets has size 17.

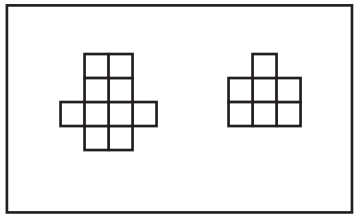


Figure 1.2: The union of four disjoint sets of size five.

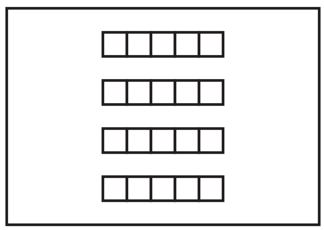


Figure 1.2) in Problems 9 and 10. These are two of the most basic principles of combinatorics.

You may have noticed some standard mathematical words and phrases such as *set*, *ordered pair*, *function* and so on creeping into the problems. One of our goals in these notes is to show how most counting problems can be recognized as counting all or some of the elements of a set of standard mathematical objects. For example Problem 4 is meant to suggest that the question we asked in Problem 3 was really a problem of counting all the ordered pairs consisting of a bread choice and a filling choice. We use $A \times B$ to stand for the set of all ordered pairs whose first element is in A and whose second element is in B and we call $A \times B$ the *Cartesian product* of A and B, so you can think of Problem 4 as asking you for the size of the Cartesian product of M and N.

When a set S is a union of disjoint sets $B_1, B_2, \ldots, B_m$ we say that the sets $B_1, B_2, \ldots, B_m$ are a **partition** of the set S. Thus a partition of S is a (special kind of) set of sets. So that we don't find ourselves getting confused between the set S and the sets B_i into which we have divided it, we often call the sets $B_1, B_2, \ldots, B_m$ the *blocks* of the partition. In this language, the **sum principle** says that

> if we have a partition of a set S, then the size of S is the sum of the sizes of the blocks of the partition.

The **product principle** says that

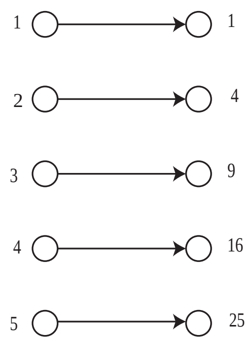
> if we have a partition of a set S into m blocks, each of size n, then S has size mn.

1.2.2 Functions

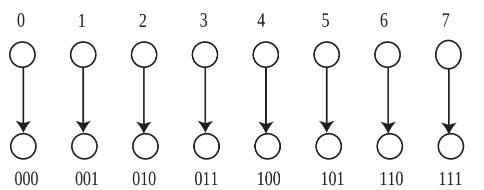
As another example, Problem 7 explicitly asked you to relate the idea of counting functions to the question of Problem 6. You have learned how to draw graphs in the Cartesian plane of functions from a set of numbers to a set of numbers. You may recall how we can determine whether a graph is a graph of a function by examining whether each vertical straight line crosses the graph at most one time. You might also recall how we can determine whether such a function is one-to-one by examining whether each horizontal straight line crosses the graph at most one time. The functions we deal with will often involve objects which are not numbers, and will often be functions from one finite set to another. Thus graphs in the cartesian plane will often not be available to us for visualizing functions. However, there is another kind of graph called a *directed graph or* digraph that is especially useful when dealing with functions between finite sets. In Figure 1.3 we show several examples. If we have a function f from a set S to a set T, we draw a line of dots or circles to represent the elements of S and another (usually parallel) line of circles or dots to represent the elements of T. We then draw an arrow from the circle for x to the circle for y if $f(x) = y$.

Notice that there is a simple test for whether a digraph whose vertices represent the elements of the sets S and T is the digraph of a function from S to T. There must be one and only one arrow leaving each vertex of the digraph representing an element of S. The fact that there is one arrow means that $f(x)$ is defined for each x in S. The fact that there is only one arrow means that each x in S is related to exactly one element of T. For further discussion of functions and digraphs see Sections A.1.1 and A.1.2 of Appendix A.

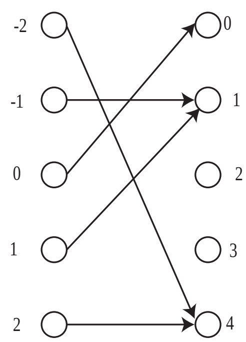
Figure 1.3: What is a digraph of a function?



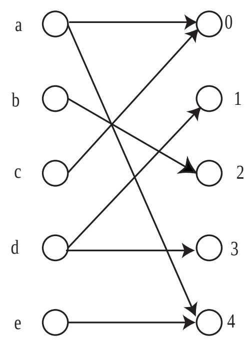
(a) The function given by $f(x) = x^2$ on the domain {1,2,3,4,5}.



(b) The function from the set {0,1,2,3,4,5,6,7} to the set of triples of zeros and ones given by $f(x)$ = the binary representation of x.



(c) The function from the set {-2,-1,0,1,2} to the set {0,1,2,3,4} given by $f(x) = x^2$.



(d) Not the digraph of a function.

11. Draw the digraph of the function from the set {Alice, Bob, Dawn, Bill} to the set {A, B, C, D, E} given by

$$f(X) = \text{the first letter of the name } X.$$

12. A function $f : S \to T$ is called an *onto function* or *surjection* if each element of T is $f(x)$ for some $x \in S$. Choose a set S and a set T so that you can draw the digraph of a function from S to T that is one-to-one but not onto, and draw the digraph of such a function

13. Choose a set S and a set T so that you can draw the digraph of a

function from S to T that is onto but not one-to-one, and draw the digraph of such a function.

14. What does the digraph of a one-to-one function (injection) from a finite set X to a finite set Y look like? (Look for a test somewhat similar to the one we described for when a digraph is the digraph of a function.) What does the digraph of an onto function look like? What does the digraph of a one-to-one and onto function from a finite set S to a set T look like?

1.2.3 The bijection principle

Another name for a one-to-one and onto function is **bijection**. The first two digraphs in Figure 1.3 are digraphs of bijections. The description in Problem 14 of the digraph of a bijection from X to Y illustrates one of the fundamental principles of combinatorial mathematics, the **bijection principle**;

> Two sets have the same size if and only if there is a bijection between them.

It is surprising how this innocent sounding principle guides us into finding insight into some otherwise very complicated proofs.

1.2.4 Counting subsets of a set

15. The *binary* representation of a number m is a list, or string, $a_1a_2\ldots a_k$ of zeros and ones such that $m = a_1 2^{k-1} + a_2 2^{k-2} + \cdots + a_k 2^0$. Describe a bijection between the binary representations of the integers between 0 and $2^n - 1$ and the subsets of an n-element set. What does this tell you about the number of subsets of an n-element set?

16. Notice that the first question in Problem 8 asked you for the number of ways to choose a three element subset from a 12 element subset. You may have seen a notation like $\binom{n}{k}$, $C(n, k)$, or ${}_nC_k$ which stands for the number of ways to choose a k-element subset from an n-element set. The number $\binom{n}{k}$ is read as "n choose k" and is called a **binomial coefficient** for reasons we will see later on. Another frequently used way to read the binomial coefficient notation is "the number of combinations of n things taken k at a time." You are going to be asked

to construct two bijections that relate to these numbers and figure out what famous formula they prove. We are going to think about subsets of the n-element set $[n] = \{1, 2, 3, \ldots, n\}$. As an example, the set of two-element subsets of [4] is

$$\{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}\}.$$

This example tells us that $\binom{4}{2} = 6$. Let C be the set of k-element subsets of $[n]$ that contain the number n, and let D be the set of k-element subsets of $[n]$ that don't contain n.

(a) Let C' be the set of $(k-1)$-element subsets of $[n-1]$. Describe a bijection from C to C'. (A verbal description is fine.)

(b) Let D' be the set of k-element subsets of $[n-1] = \{1, 2, \ldots n-1\}$. Describe a bijection from D to D'. (A verbal description is fine.)

(c) Based on the two previous parts, express the sizes of C and D in terms of binomial coefficients involving $n-1$ instead of n.

(d) Apply the sum principle to C and D and obtain a formula that expresses $\binom{n}{k}$ in terms of two binomial coefficients involving $n-1$. You have just derived the Pascal Equation that is the basis for the famous Pascal's Triangle.

1.2.5 Pascal's Triangle

The Pascal Equation that you derived in Problem 16 gives us the triangle in Figure 1.4. This figure has the number of k-element subsets of an n-element set as the kth number over in the nth row (we call the top row the zeroth row and the beginning entry of a row the zeroth number over). You'll see that your formula doesn't say anything about $\binom{n}{k}$ if $k = 0$ or $k = n$, but otherwise it says that each entry is the sum of the two that are above it and just to the left or right.

17. Just for practice, what is the next row of Pascal's triangle?

18. Without writing out the rows completely, write out enough of Pascal's triangle to get a numerical answer for the first question in Problem 8.

Figure 1.4: Pascal's Triangle

$$\begin{array}{ccccccccccccccc}
 & & & & & & & 1 & & & & & & & \\
 & & & & & & 1 & & 1 & & & & & & \\
 & & & & & 1 & & 2 & & 1 & & & & & \\
 & & & & 1 & & 3 & & 3 & & 1 & & & & \\
 & & & 1 & & 4 & & 6 & & 4 & & 1 & & & \\
 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & \\
 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & \\
1 & & 7 & & 21 & & 35 & & 35 & & 21 & & 7 & & 1
\end{array}$$

It is less common to see Pascal's triangle as a right triangle, but it actually makes your formula easier to interpret. In Pascal's Right Triangle, the element in row n and column k (with the convention that the first row is row zero and the first column is column zero) is $\binom{n}{k}$. In this case your formula says each entry in a row is the sum of the one above and the one above and to the left, except for the leftmost and right most entries of a row, for which that doesn't make sense. Since the leftmost entry is $\binom{n}{0}$ and the rightmost entry is $\binom{n}{n}$, these entries are both one (why is that?), and your formula then tells how to fill in the rest of the table.

Figure 1.5: Pascal's Right Triangle

1							
1	1						
1	2	1					
1	3	3	1				
1	4	6	4	1			
1	5	10	10	5	1		
1	6	15	20	15	6	1	
1	7	21	35	35	21	7	1

Seeing this right triangle leads us to ask whether there is some natural way to extend the right triangle to a rectangle. If we did have a rectangular table of binomial coefficients, counting the first row as row zero (i.e., $n = 0$) and the first column as column zero (i.e., $k = 0$), the entries we don't yet have are values of $\binom{n}{k}$ for $k > n$. But how many k-element subsets does an n-element set have if $k > n$? The answer, of course, is zero, so all the other entries we would fill in would be zero, giving us the rectangular array in Figure 1.6. It is straightforward to check that Pascal's equation now works for all the entries in the rectangle that have an entry above them and an entry above and to the left.

Figure 1.6: Pascal's Rectangle

1	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0
1	2	1	0	0	0	0	0
1	3	3	1	0	0	0	0
1	4	6	4	1	0	0	0
1	5	10	10	5	1	0	0
1	6	15	20	15	6	1	0
1	7	21	35	35	21	7	1

19. (Optional) We defined $\binom{n}{k}$ to be 0 when $k > n$ in order to get a rectangular table of numbers that satisfies the Pascal Equation. Is there any other way to define $\binom{n}{k}$ when $k > n$ in order to get a rectangular table that agrees with Pascal's Right Triangle for $k \leq n$ and satisfies the Pascal Equation? Suppose we want to extend Pascal's Rectangle to the left and define $\binom{n}{-k}$ for $n \geq 0$ and $k > 0$ so that $-k < 0$. What should we put into row n and column $-k$ of Pascal's Rectangle in order for the Pascal Equation to hold true? What should we put into row $-n$ and column k or column $-k$ in order for the Pascal Equation to continue to hold? Do we have any freedom of choice?

20. There is a variant of the bijection we used to prove the Pascal Equation that can be used to give another proof of your formula in Problem 15 for the number of subsets of an n-element set using the Principle of Mathematical Induction. If you are familiar with Mathematical Induction, try to find the proof. If not, now is the time to visit the Appendix on Mathematical Induction (and work through the problems there).

1.2.6 The General Product Principle

21. Let us now return to Problem 7 and justify—or perhaps finish—our answer to the question about the number of functions from a three-element set to a 12-element set.

    (a) We begin with a question to which we can apply the product principle directly. Namely, how many functions f are there from the set $[2] = \{1, 2\}$ to the set $[12]$. How many functions are there with $f(2) = 1$? With $f(2) = 2$? With $f(2) = 3$? With $f(2) = i$ for any fixed i between 1 and 12? The set of functions from $[2]$ to $[12]$ is the union of 12 sets: the set of f with $f(2) = 1$, the set of f with $f(2) = 2$, $\ldots$, the set of f with $f(2) = 12$. How many functions does each of these sets have? From the product principle, what may you conclude about the number of functions in the union of these 12 sets?

    (b) Now consider the set of functions from $[3]$ to $[12]$. How many of these functions have $f(3) = 1$? What bijection (just describe it in words) are you using (implicitly, if not explicitly) to answer this question? How many of these functions have $f(3) = 2$? How many have $f(3) = i$ for any i between 1 and 12? For each i, let S_i be the set of functions f from $[3]$ to $[12]$ with $f(3) = i$. What is the size S_i? What is the size of the union of the sets S_i? How many functions are there from $[3]$ to $[12]$?

    (c) Based on the examples you've seen so far, make a conjecture about how many functions there are from $[m]$ to $[n]$.

    (d) Now suppose we are thinking about a set S of functions f from $[n]$ to some set X. Suppose there are k_1 choices for $f(1)$. Suppose that for each choice of $f(1)$ there are k_2 choices for $f(2)$. (For example, in counting one-to-one functions from $[3]$ to $[12]$, there are 12 choices for $f(1)$, and for each choice of $f(1)$ there are 11 choices for $f(2)$.) In general suppose that for each choice of $f(1)$, $f(2)$, $\ldots f(i-1)$, there are k_i choices for $f(i)$. (For example, in counting one-to-one functions from $[3]$ to $[12]$, for each choice of $f(1)$ and $f(2)$, there are 10 choices for $f(3)$.) How many functions do you think are in the set S? This is called the *product principle for functions* or the **general product principle**. If you can see how to prove the general product principle from the product principle, do so. If not, we will come back to this question later. How does it relate to counting functions from $[m]$ to $[n]$? How does it relate to counting one-to-one functions from $[m]$ to $[n]$?

(e) Prove the conjecture (about the number of functions in S) in Part 21c when $m = 2$ and when $m = 3$. Prove the conjecture for an arbitrary positive integer m.

22. How does the general product principle apply to Problem 6?

23. In how many ways can we pass out k distinct pieces of fruit to n children (with no restriction on how many pieces of fruit a child may get)?

24. Another name for a list of k distinct things chosen from a set S is a **k-element permutation of S** We can also think of a k-element permutation of S as a one-to-one function (or, in other words, injection) from $[k] = \{1, 2, \ldots, k\}$ to S. How many k-element permutations does an n-element set have? (For this problem it is natural to assume $k \leq n$. However the question makes sense even if $k > n$. What is the number of k-element permutations of an n-element set if $k > n$?

25. Assuming $k \leq n$, in how many ways can we pass out k distinct pieces of fruit to n children if each child may get at most one? What is the number if $k > n$? Assume for both questions that we pass out all the fruit.

26. The word *permutation* is actually used in two different ways in mathematics. A **permutation** of a set S is a bijection from S to S. How many permutations does an n-element set have?

Notice that there is a great deal of consistency between this use of the word permutation and the use in the previous problem. If we have some way $a_1, a_2, \ldots, a_n$ of listing our set, then any other list $b_1, b_2, \ldots, b_n$ gives us the bijection whose rule is $f(a_i) = b_i$, and any bijection, say the one given by $g(a_i) = c_i$ gives us a list $c_1, c_2, \ldots, c_n$ of S. Thus there is really very little difference between the idea of a permutation of S and an n-element permutation of S when n is the size of S.

There are a number of different notations for the number of k-element permutations of an n-element set. The one we shall use was introduced by

Don Knuth; namely $n^{\underline{k}}$, read "n to the k falling" or "n to the k down". It is standard to call $n^{\underline{k}}$ *the k-th falling factorial power of n*, which explains why we use exponential notation.

27. A tennis club has $2n$ members. We want to pair up the members by twos for singles matches. In how many ways may we pair up all the members of the club? Suppose that in addition to specifying who plays whom, for each pairing we say who serves first. Now in how many ways may we specify our pairs?

28. There is yet another bijection that lets us prove that a set of size n has 2^n subsets. Namely, for each subset S of $[n] = \{1, 2, \ldots, n\}$, define a function (traditionally denoted by χ_S) as follows.[1]

$$\chi_S(i) = \begin{cases} 1 \text{ if } i \in S \\ 0 \text{ if } i \notin S \end{cases}$$

The function χ_S is called the *characteristic function* of S. Notice that the characteristic function is a function from $[n]$ to $\{0, 1\}$.

   (a) For practice, consider the function $\chi_{\{1,3\}}$ for the subset $\{1, 3\}$ of the set $\{1, 2, 3, 4\}$. What are

      i. $\chi_{\{1,3\}}(1)$?
      ii. $\chi_{\{1,3\}}(2)$?
      iii. $\chi_{\{1,3\}}(3)$?
      iv. $\chi_{\{1,3\}}(4)$?

   (b) We define a function f from the set of subsets of $[n] = \{1, 2, \ldots, n\}$ to the set of functions from $[n]$ to $\{0, 1\}$ by $f(S) = \chi_S$. Explain why f is a bijection.

   (c) Why does the fact that f is a bijection prove that $[n]$ has 2^n subsets?

In Problems 15, 20, and 28 you gave three proofs of the following theorem.

Theorem 1 *The number of subsets of an n-element set is 2^n.*

The proofs in Problem 15 and 28 use essentially the same bijection, but they interpret sequences of zeros and ones differently, and so end up being different proofs.

[1]The symbol χ is the Greek letter chi that is pronounced Ki, with the i sounding like "eye."

1.2.7 The quotient principle

29. As we noted in Problem 16, the first question in Problem 8 asked us for the number of three-element subsets of a twelve-element set. We were able to use the Pascal Equation to get a numerical answer to that question. Had we had twenty or thirty flavors of ice cream to choose from, using the Pascal Equation to get our answer would have entailed a good bit more work. We have seen how the general product principle gives us an answer to Problem 6. Thus we might think that the number of ways to choose a three element set from 12 elements is the number of ways to choose the first element times the number of ways to choose the second element times the number of ways to choose the third element, which is $12 \cdot 11 \cdot 10 = 1320$. However, our result in Problem 16 shows that this is wrong. What is it that is different between the number of ways to stack ice cream in a triple decker cone with three different flavors of ice cream and the number of ways to simply choose three different flavors of ice cream? In particular, how many different triple decker cones use the same three flavors? Using this, compute the number of ways to choose three different flavors of ice cream (from 12 flavors) from the number of ways to choose a triple decker cone with three different flavors (from 12 flavors).

30. Based on what you observed in Problem 29, how many k-element subsets does an n-element set have?

31. In how many ways can we pass out k (identical) ping-pong balls to n children if each child may get at most one?

32. In how many ways may n people sit around a round table? (Assume that when people are sitting around a round table, all that really matters is what we might call their position relative to each other. For example, if we can get one arrangement of people around the table from another by having everyone get up and move to the right one place and sit back down, we get an equivalent arrangement of people. Notice that you can get a list from a seating arrangement by marking a place at the table, and then listing the people at the table, starting at that place and moving around to the right. We would say that two seating arrangements have everyone in the same relative position if everyone has the same person to the right in both arrangements.) There

are at least two different ways of doing this problem. Try to find them both.

33. A given k-element subset can be listed as a k-element permutation in $k!$ ways. We can partition the set of all k-element permutations of S up into blocks by letting B_K be the set of all k-element permutations of K for each k-element subset K of S. How many permutations are there in a block? If S has n elements, what does problem 24 tell you about the total number of k-element permutations of S? Describe a bijection between the set of blocks of the partition and the set of k-element subsets of S. What formula does this give you for the number $\binom{n}{k}$ of k-element subsets of an n-element set?

34. A basketball team has 12 players. However, only five players play at any given time during a game. In how may ways may the coach choose the five players? To be more realistic, the five players playing a game normally consist of two guards, two forwards, and one center. If there are five guards, four forwards, and three centers on the team, in how many ways can the coach choose two guards, two forwards, and one center? What if one of the centers is equally skilled at playing forward?

35. In Problem 32, describe a way to partition the n-element permutations of the n people into blocks so that there is a bijection between the set of blocks of the partition and the set of arrangements of the n people around a round table. What method of solution for Problem 32 does this correspond to?

36. In Problems 33 and 35, you have been using the product principle in a new way. One of the ways in which we previously stated the product principle was "If we partition a set into m blocks each of size n, then the set has size $m \cdot n$." In problems 33 and 35 we knew the size p of a set P of permutations of a set, and we knew we had partitioned P into some unknown number of blocks, each of a certain known size r. If we let q stand for the number of blocks, what does the product principle tell us about p, q, and r? What do we get when we solve for q?

The formula you found in the Problem 36 is so useful that we are going to single it out as another principle. The **quotient principle** says:

If we partition a set P into q blocks, each of size r, then $q = p/r$.

The quotient principle is really just a restatement of the product principle, but thinking about it as a principle in its own right often leads us to find solutions to problems. Notice that it does not always give us a formula for the number of blocks of a partition; it only works when all the blocks have the same size.

In Section A.1.3 of Appendix A we introduce the idea of an equivalence relation, see what equivalence relations have to do with partitions, and discuss the quotient principle from that point of view.

37. In how many ways may we string n beads on a necklace without a clasp? (Assume someone can pick up the necklace, move it around in space and put it back down, giving an apparently different way of stringing the beads that is equivalent to the first. How could we get a list of beads from a necklace?)

38. In how many ways may we attach two identical red beads and two identical blue beads to the corners of a square free to move around in (three-dimensional) space?

1.3 Some Applications of the Basic Principles

1.3.1 Lattice paths and Catalan Numbers

39. In a part of a city, all streets run either north-south or east-west, and there are no dead ends. Suppose we are standing on a street corner. In how many ways may we walk to a corner that is four blocks north and six blocks east, using as few blocks as possible?

40. Problem 39 has a geometric interpretation in a coordinate plane. A *lattice path* in the plane is a "curve" made up of line segments that either go from a point (i, j) to the point $(i + 1, j)$ or from a point (i, j) to the point $(i, j + 1)$, where i and j are integers. (Thus lattice paths always move either up or to the right.) The length of the path is the number of such line segments. What is the length of a lattice path from $(0, 0)$ to (m, n)? How many such lattice paths of that length are there? How many lattice paths are there from (i, j) to (m, n), assuming i, j, m, and n are integers?

41. Another kind of geometric path in the plane is a *diagonal lattice path*. Such a path is a path made up of line segments that go from a point (i, j) to $(i+1, j+1)$ (this is often called an *upstep*) or $(i+1, j-1)$ (this is often called a *downstep*), again where i and j are integers. (Thus diagonal lattice paths always move towards the right but may move up or down.) Describe which points are connected to $(0, 0)$ by diagonal lattice paths. What is the length of a diagonal lattice path from $(0, 0)$ to (m, n)? Assuming that (m, n) is such a point, how many diagonal lattice paths are there from $(0, 0)$ to (m, n)?

42. A school play requires a ten dollar donation per person; the donation goes into the student activity fund. Assume that each person who comes to the play pays with a ten dollar bill or a twenty dollar bill. The teacher who is collecting the money forgot to get change before the event. If there are always at least as many people who have paid with a ten as a twenty as they arrive the teacher won't have to give anyone an IOU for change. Suppose $2n$ people come to the play, and exactly half of them pay with ten dollar bills.

    (a) Describe a bijection between the set of sequences of tens and twenties people give the teacher and the set of lattice paths from $(0, 0)$ to (n, n).

    (b) Describe a bijection between the set of sequences of tens and twenties that people give the teacher and the set of diagonal lattice paths from $(0, 0)$ and $(0, 2n)$.

    (c) In each case, what is the geometric interpretation of a sequence that does not require the teacher to give any IOUs?

43. Notice that a lattice path from $(0, 0)$ to (n, n) stays inside (or on the edges of) the square whose sides are the x-axis, the y-axis, the line $x = n$ and the line $y = n$. In this problem we will compute the number of lattice paths from (0,0) to (n, n) that stay inside (or on the edges of) the triangle whose sides are the x-axis, the y-axis and the line $y = x$.

    (a) Explain why the number of lattice paths from $(0, 0)$ to (n, n) that go outside the triangle is the number of lattice paths from $(0, 0)$ to (n, n) that either touch or cross the line $y = x + 1$.

(b) Find a bijection between lattice paths from $(0,0)$ to (n,n) that touch (or cross) the line $y = x+1$ and lattice paths from $(-1,1)$ to (n,n).

(c) Find a formula for the number of lattice paths from $(0,0)$ to (n,n) that do not cross the line $y = x$. The number of such paths is called a *Catalan Number* and is usually denoted by C_n.

44. Your formula for the Catalan Number can be expressed as a binomial coefficient divided by an integer. Whenever we have a formula that calls for division by an integer, an ideal combinatorial explanation of the formula is one that uses the quotient principle. The purpose of this problem is to find such an explanation using diagonal lattice paths.[2] A diagonal lattice path that never goes below the y-coordinate of its first point is called a *Dyck Path*. We will call a Dyck Path from $(0,0)$ to $(2n,0)$ a *Catalan Path* of length $2n$. Thus the number of Catalan Paths of length $2n$ is the Catalan Number C_n.

(a) If a Dyck Path has n steps (each an upstep or downstep), why do the first k steps form a Dyck Path for each nonnegative $k \leq n$?

(b) Thought of as a curve in the plane, a lattice path can have many local maxima and minima, and can have several absolute maxima and minima, that is, several highest points and several lowest points. What is the y-coordinate of an absolute minimum point of a Dyck Path starting at $(0,0)$? Explain why a Dyck Path whose rightmost absolute minimum point is its last point is a Catalan Path.

(c) Let D be the set of all diagonal lattice paths from $(0,0)$ to $(2n,0)$. Suppose we partition D by letting B_i be the set of lattice paths in D that have i upsteps (perhaps mixed with some downsteps) following the last absolute minimum. How many blocks does this partition have? Give a succinct description of the block B_0.

(d) How many upsteps are in a Catalan Path?

(e) We are going to give a bijection between the set of Catalan Paths and the block B_i for each i between 1 and n. For now, suppose

[2]The result we will derive is called the Chung-Feller Theorem; this approach is based of a paper of Wen-jin Woan "Uniform Partitions of Lattice Paths and Chung-Feller Generalizations," **American Mathematics Monthly 58** June/July 2001, p556.

the value of i, while unknown, is fixed. We take a Catalan path and break it into three pieces. The piece F consists of all steps before the ith upstep in the Catalan path. The piece U consists of the ith upstep. The piece B is the portion of the path that follows the ith upstep. Thus we can think of the path as FUB. Show that the function that takes FUB to BUF is a bijection from the set of Catalan Paths onto the block B_i of the partition.

(f) Explain how you have just given another proof of the formula for the Catalan Numbers.

1.3.2 The Binomial Theorem

45. We know that $(x+y)^2 = x^2 + 2xy + y^2$. Multiply both sides by $(x+y)$ to get a formula for $(x+y)^3$ and repeat to get a formula for $(x+y)^4$. Do you see a pattern? If so, what is it? If not, repeat the process to get a formula for $(x+y)^5$ and look back at Figure 1.4 to see the pattern. Conjecture a formula for $(x+y)^n$.

46. When we apply the distributive law n times to $(x+y)^n$, we get a sum of terms of the form $x^i y^{n-i}$ for various values of the integer i.

    (a) If it is clear to you that each term of the form $x^i y^{n-i}$ that we get comes from choosing an x from i of the terms and a y from $n-i$ of the terms and multiplying these choices together, then answer this part of the problem and skip the next part. Otherwise, do the next part instead of this one. In how many ways can we choose an x from i terms and a y from $n-i$ terms?

    (b) Expand the product $(x_1+y_1)(x_2+y_2)(x_3+y_3)$. What do you get when you substitute x for each x_i and y for each y_i? Now imagine expanding
    $$(x_1+y_1)(x_2+y_2)\cdots(x_n+y_n).$$
    Once you apply the commutative law to the individual terms you get, you will have a sum of terms of the form
    $$x_{k_1}x_{k_2}\cdots x_{k_i}\cdot y_{j_1}y_{j_2}\cdots y_{j_{n-i}}.$$
    What is the set $\{k_1, k_2, \ldots, k_i\} \cup \{j_1, j_2, \ldots, j_{n-i}\}$? In how many ways can you choose the set $\{k_1, k_2, \ldots, k_i\}$? Once you have chosen

this set, how many choices do you have for $\{j_1, j_2, \ldots, j_{n-i}\}$? If you substitute x for each x_i and y for each y_i, how many terms of the form $x^i y^{n-i}$ will you have in the expanded product

$$(x_1 + y_1)(x_2 + y_2)\cdots(x_n + y_n) = (x + y)^n?$$

(c) Explain how you have just proved your conjecture from Problem 45. The theorem you have proved is called the **Binomial Theorem**.

47. What is $\sum_{i=1}^{n} \binom{10}{i} 3^i$?

48. What is $\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \cdots \pm \binom{n}{n}$ if n is an integer bigger than 0?

49. Explain why

$$\sum_{i=0}^{n} \binom{m}{i}\binom{n}{k-i} = \binom{m+n}{k}.$$

Try to find two different explanations.

50. From the symmetry of the binomial coefficients, it is not too hard to see that when n is an odd number, the number of subsets of $\{1, 2, \ldots, n\}$ of odd size equals the number of subsets of $\{1, 2, \ldots, n\}$ of even size. Is it true that when n is even the number of subsets of $\{1, 2, \ldots, n\}$ of even size equals the number of subsets of odd size? Why or why not?

1.3.3 The pigeonhole principle

51. American coins are all marked with the year in which they were made. How many coins do you need to have in your hand to guarantee that on two (at least) of them, the date has the same last digit?

There are many ways in which you might explain your answer to Problem 51. For example, you can partition the coins according to the last digit of their date; that is, you put all the coins with a given last digit in a block together, and put no other coins in that block; repeating until all coins are in some block. Then you have a partition of your set of coins. If no two coins have the same last digit, then each block has exactly one coin. Since there are only ten digits, there are at most ten blocks and so by the sum principle there are at most ten coins. In fact with ten coins it is possible to have no two with the same last digit, but with 11 coins some block must have at least two coins in order for the sum of the sizes of at most ten blocks to be 11. This is one explanation of why we need 11 coins in Problem 51. This kind of situation arises often in combinatorial situations, and so rather than always using the sum principle to explain our reasoning , we enunciate another principle which we can think of as yet another variant of the sum principle. The **pigeonhole principle** states that

> If we partition a set with more than n elements into n parts, then at least one part has more than one element.

The pigeonhole principle gets its name from the idea of a grid of little boxes that might be used, for example, to sort mail, or as mailboxes for a group of people in an office. The boxes in such grids are sometimes called pigeonholes in analogy with stacks of boxes used to house homing pigeons when homing pigeons were used to carry messages. People will sometimes state the principle in a more colorful way as "if we put more than n pigeons into n pigeonholes, then some pigeonhole has more than one pigeon."

52. Show that if we have a function from a set of size n to a set of size less than n, then f is not one-to-one.

53. Show that if S and T are finite sets of the same size, then a function f from S to T is one-to-one if and only if it is onto.

54. There is a *generalized pigeonhole principle* which says that if we partition a set with more than kn elements into n blocks, then at least one block has at least $k+1$ elements. Prove the generalized pigeonhole principle.

55. All the powers of five end in a five, and all the powers of two are even. Show that for the first five powers of a prime other than two or five, one must have a last digit of one.

56. Show that in a set of six people, there is a set of at least three people who all know each other, or a set of at least three people none of whom know each other. (We assume that if person A knows person B, then person B knows person A.)

57. Draw five circles labeled Al, Sue, Don, Pam, and Jo. Find a way to draw red and green lines between people so that every pair of people is joined by a line and there is neither a triangle consisting entirely of red lines or a triangle consisting of green lines. What does Problem 56 tell you about the possibility of doing this with six people's names? What does this problem say about the conclusion of Problem 56 holding when there are five people in our set rather than six?

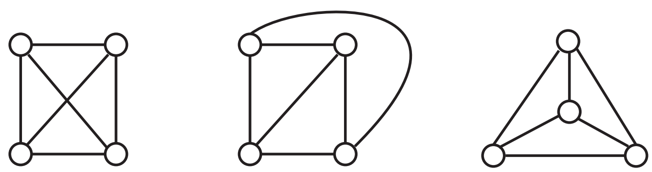
1.3.4 Ramsey Numbers

Problems 56 and 57 together show that six is the smallest number R with the property that if we have R people in a room, then there is either a set of (at least) three mutual acquaintances or a set of (at least) three mutual strangers. Another way to say the same thing is to say that six is the smallest number so that no matter how we connect 6 points in the plane (no three on a line) with red and green lines, we can find either a red triangle or a green triangle. There is a name for this property. The **Ramsey Number** $R(m, n)$ is the smallest number R so that if we have R people in a room, then there is a set of at least m mutual acquaintances or at least n mutual strangers. There is also a geometric description of Ramsey Numbers; it uses the idea of a *complete graph* on R vertices. A **complete graph** on R vertices consists of R points in the plane together with line segments (or curves) connecting each two of the R vertices.[3] The points are called *vertices* and the line segments are called *edges*. In Figure 1.7 we show three different ways to draw a complete graph on four vertices. We use K_n to stand for a complete graph on n vertices.

Our geometric description of $R(3, 3)$ may be translated into the language of graph theory (which is the subject that includes complete graphs) by saying $R(3, 3)$ is the smallest number R so that if we color the edges of a K_R with two colors, then we can find in our picture a K_3 all of whose edges have the same color. The graph theory description of $R(m, n)$ is that $R(m, n)$ is the smallest number R so that if we color the edges of a K_R with red and green, then we can find in our picture either a K_m all of whose edges are red or a K_n all of whose edges are green. Because we could have said our colors in the opposite order, we may conclude that $R(m, n) = R(n, m)$. In

[3]As you may have guessed, a *graph* is a collection of points called vertices and a collection of line segments or curves called edges and each joining two of the points. Thus a complete graph is just a graph in which *each* pair of vertices is joined by an edge.

Figure 1.7: Three ways to draw a complete graph on four vertices



particular $R(n, n)$ is the smallest number R such that if we color the edges of a K_R with two colors, then our picture contains a K_n all of whose edges have the same color.

58. Since $R(3,3) = 6$, an uneducated guess might be that $R(4,4) = 8$. Show that this is not the case.

59. Show that among ten people, there are either four mutual acquaintances or three mutual strangers. What does this say about $R(4,3)$?

60. Show that among an odd number of people there is at least one person who is an acquaintance of an even number of people and therefore also a stranger to an even number of people.

61. Find a way to color the edges of a K_8 with red and green so that there is no red K_4 and no green K_3.

62. Find $R(4,3)$.

As of this writing, relatively few Ramsey Numbers are known. $R(3, n)$ is known for $n < 10$, $R(4,4) = 18$, and $R(5,4) = R(4,5) = 25$.

1.4 Supplementary Chapter Problems

Note: The list of supplementary problems is far from complete, but this is a sampling of what I plan. I welcome suggestions.

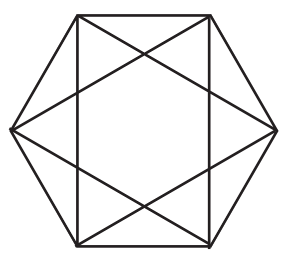
1. Remember that we can write n as a sum of n ones. How many plus signs do we use? In how many ways may we write n as a sum of a list of k positive numbers? Such a list is called a *composition* of n into k parts.

2. In Problem 1 we defined a composition of n into k parts. What is the total number of compositions of n (into any number of parts).

3. A list of parentheses is said to be balanced if there are the same number of left parentheses as right, and as we count from left to right we always find at least as many left parentheses as right parentheses. For example, (((()()))()) is balanced and ((()) and (()()))(() are not. How many balanced lists of n left and n right parentheses are there?

4. Suppose we plan to put six computers in a network as shown in Figure 1.8. The lines show which computers can communicate directly with which others. Consider two ways of assigning computers to the nodes of the network different if there are two computers that communicate directly in one assignment and that don't communicate directly in the other. In how many different ways can we assign computers to the network?

Figure 1.8: A computer network.



5. In a circular ice cream dish we are going to put four distinct scoops of ice cream chosen from among twelve flavors. Assuming we place four scoops of the same size as if they were at the corners of a square, and recognizing that moving the dish doesn't change the way in which we have put the ice cream into the dish, in how many ways may we choose the ice cream and put it into the dish?

6. In as many ways as you can, show that $\binom{n}{k}\binom{n-k}{m} = \binom{n}{m}\binom{n-m}{k}$.

7. A tennis club has $4n$ members. To specify a doubles match, we choose two teams of two people. In how many ways may we arrange the

members into doubles matches so that each player is in one doubles match? In how many ways may we do it if we specify in addition who serves first on each team?

Chapter 2

Combinatorial Applications of Induction

2.1 Some Examples of Mathematical Induction

In Chapter 1 (Problem 20), we used the principle of mathematical induction to prove that a set of size n has 2^n subsets. If you were unable to do that problem and you haven't yet read Appendix B, you should do so now.

2.1.1 Mathematical induction

The **principle of mathematical induction** states that

> In order to prove a statement about an integer n, if we can
>
> 1. Prove the statement when $n = b$, for some fixed integer b
> 2. Show that the truth of the statement for $n = k - 1$ implies the truth of the statement for $n = k$ whenever $k > b$,
>
> then we can conclude the statement is true for all integers $n \geq b$.

As an example, let us return to Problem 20. The statement we wish to prove is the statement that "A set of size n has 2^n subsets."

> Our statement is true when $n = 0$, because a set of size 0 is the empty set and the empty set has $1 = 2^0$ subsets. (This step of our proof is called a *base step*.)
>
> Now suppose that $k > 0$ and every set with $k-1$ elements has 2^{k-1} subsets. Suppose $S = \{a_1, a_2, \ldots a_k\}$ is a set with k elements. We partition the subsets of S into two blocks. Block B_1 consists of the subsets that do not contain a_n and block B_2 consists of the subsets that do contain a_n. Each set in B_1 is a subset of $\{a_1, a_2, \ldots a_{k-1}\}$, and each subset of $\{a_1, a_2, \ldots a_{k-1}\}$ is in B_1. Thus B_1 is the set of all subsets of $\{a_1, a_2, \ldots a_{k-1}\}$. Therefore by our assumption in the first sentence of this paragraph, the size of B_1 is 2^{k-1}. Consider the function from B_2 to B_1 which takes a subset of S including a_k and removes a_k from it. This function is defined on B_2, because every set in B_2 contains a_k. This function is onto, because if T is a set in B_1, then $T \cup \{a_k\}$ is a set in B_2 which the function sends to T. This function is one-to-one because if V and W are two different sets in B_2, then removing a_k from them gives two different sets in B_1. Thus we have a bijection between B_1 and B_2, so B_1 and B_2 have the same size. Therefore by the sum principle the size of $B_1 \cup B_2$ is $2^{k-1} + 2^{k-1} = 2^k$. Therefore S has 2^k subsets. This shows that if a set of size $k-1$ has 2^{k-1} subsets, then a set of size k has 2^k subsets. Therefore by the principle of mathematical induction, a set of size n has 2^n subsets for every nonnegative integer n.

The first sentence of the last paragraph is called the *inductive hypothesis*. In an inductive proof we always make an inductive hypothesis as part of proving that the truth of our statement when $n = k-1$ implies the truth of our statement when $n = k$. The last paragraph itself is called the *inductive step* of our proof. In an inductive step we derive the statement for $n = k$ from the statement for $n = k-1$, thus proving that the truth of our statement when $n = k-1$ implies the truth of our statement when $n = k$. The last sentence in the last paragraph is called the *inductive conclusion*. All inductive proofs should have a base step, an inductive hypothesis, an inductive step, and an inductive conclusion.

There are a couple details worth noticing. First, in this problem, our base step was the case $n = 0$, or in other words, we had $b = 0$. However, in other proofs, b could be any integer, positive, negative, or 0. Second, our proof that the truth of our statement for $n = k - 1$ implies the truth of our statement for $n = k$ required that k be at least 1, so that there would be an element a_k we could take away in order to describe our bijection. However, condition (2) of the principle of mathematical induction only requires that we be able to prove the implication for $k > 0$, so we were allowed to assume $k > 0$.

2.1.2 Binomial coefficients and the Binomial Theorem

63. When we studied the Pascal Equation and subsets in Chapter 1, it may have appeared that there is no connection between the Pascal relation $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ and the formula $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. Of course you probably realize you can prove the Pascal relation by substituting the values the formula gives you into the right-hand side of the equation and simplifying to give you the left hand side. In fact, from the Pascal Relation and the facts that $\binom{n}{0} = 1$ and $\binom{n}{n} = 1$, you can actually prove the formula for $\binom{n}{k}$ by induction. Do so.

64. Use the fact that $(x+y)^n = (x+y)(x+y)^{n-1}$ to give an inductive proof of the binomial theorem.

2.1.3 Inductive definition

You may have seen $n!$ described by the two equations $0! = 1$ and $n! = n(n-1)!$ for $n > 0$. By the principle of mathematical induction we know that this pair of equations defines $n!$ for all nonnegative numbers n. For this reason we call such a definition an **inductive definition**. An inductive definition is sometimes called a *recursive definition*. Often we can get very easy proofs of useful facts by using inductive definitions.

65. An inductive definition of a^n for nonnegative n is given by $a^0 = 1$ and $a^n = aa^{n-1}$.

    (a) Use this definition to prove the rule of exponents $a^{m+n} = a^m a^n$ for nonnegative m and n.

    (b) Use this definition to prove the rule of exponents $a^{mn} = (a^m)^n$.

66. Give an inductive definition of the summation notation $\sum_{i=1}^{n} a_i$. Use it to prove the distributive law

$$b\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} ba_i.$$

2.1.4 Proving the general product principle (Optional)

We stated the sum principle as

> If we have a partition of a set S, then the size of S is the sum of the sizes of the blocks of the partition.

In fact, the simplest form of the sum principle says that the size of the sum of two disjoint (finite) sets is the sum of their sizes.

67. Prove the sum principle we stated for partitions of a set from the simplest form of the sum principle.

We stated the simplest form of the product principle as

> If we have a partition of a set S into m blocks, each of size n, then S has size mn.

In Problem 21 we gave a more general form of the product principle which can be stated as

> Let S be a set of functions f from $[n]$ to some set X. Suppose there are k_1 choices for $f(1)$. Suppose that for each choice of $f(1)$ there are k_2 choices for $f(2)$. In general, suppose that for each choice of $f(1)$, $f(2)$, $\ldots f(i-1)$, there are k_i choices for $f(i)$. Then the number of functions in the set S is $k_1k_2\cdots k_n$.

68. If you weren't able to do so in Problem 21, prove the general form of the product principle from the simplest form of the product principle.

2.1.5 Double Induction and Ramsey Numbers

In Section 1.3.4 we gave two different descriptions of the Ramsey number $R(m,n)$. However if you look carefully, you will see that we never showed that Ramsey numbers actually exist; we merely described what they were and showed that $R(3,3)$ and $R(3,4)$ exist by computing them directly. As long as we can show that there is some number R such that when there are R people together, there are either m mutual acquaintances or n mutual strangers, this shows that the Ramsey Number $R(m,n)$ exists, because it is the smallest such R. Mathematical induction allows us to show that one such R is $\binom{m+n-2}{m-1}$. The question is, what should we induct on, m or n? In other words, do we use the fact that with $\binom{m+n-3}{m-2}$ people in a room there are at least $m-1$ mutual acquaintances or n mutual strangers, or do we use the fact that with at least $\binom{m+n-3}{n-2}$ people in a room there are at least m mutual acquaintances or at least $n-1$ mutual strangers? It turns out that we use both. Thus we want to be able to simultaneously induct on m and n. One way to do that is to use yet another variation on the principle of mathematical induction, the *Principle of Double Mathematical Induction.* This principle (which can be derived from one of our earlier ones) states that

> In order to prove a statement about integers m and n, if we can
>
> 1. Prove the statement when $m = a$ and $n = b$, for fixed integers a and b
> 2. Prove the statement when $m = a$ and $n > b$ and when $m > a$ and $n = b$ (for the same fixed integers a and b),
> 3. Show that the truth of the statement for $m = j$ and $n = k-1$ (with $j \geq a$ and $k > j$) and the truth of the statement for $m = j-1$ and $n = k$ (with $j > a$ and $k \geq b$) imply the truth of the statement for $m = j$ and $n = k$,
>
> then we can conclude the statement is true for all pairs of integers $m \geq a$ and $n \geq b$.

69. Prove that $R(m,n)$ exists by proving that if there are $\binom{m+n-2}{m-1}$ people in a room, then there are either at least m mutual acquaintances or at least n mutual strangers.

70. Prove that $R(m,n) \leq R(m-1,n) + R(m,n-1)$.

71. (a) What does the equation in Problem 70 tell us about $R(4,4)$?

    (b) Consider 17 people arranged in a circle such that each person is acquainted with the first, second, fourth, and eighth person to the right and the first, second, fourth, and eighth person to the left. can you find a set of four mutual acquaintances? Can you find a set of four mutual strangers?

(c) What is $R(4,4)$?

72. (Optional) Can you prove the equation of Problem 70 by induction on $m+n$? If so, do so, and if not, explain where there is a problem in trying to do so.

73. (Optional) Prove the Principle of Double Mathematical Induction from the Principle of Mathematical Induction.

2.2 Recurrence Relations

We have seen in Problem 20 (or Problem 21 in the Appendix on Induction) that the number of subsets of an n-element set is twice the number of subsets of an $n-1$-element set.

74. Explain why it is that the number of bijections from an n-element set to an n-element set is equal to n times the number of bijections from an $(n-1)$-element subset to an $(n-1)$-element set. What does this have to do with Problem 26?

We can summarize these observations as follows. If s_n stands for the number of subsets of an n-element set, then

$$s_n = 2s_{n-1}, \tag{2.1}$$

and if b_n stands for the number of bijections from an n-element set to an n-element set, then

$$b_n = nb_{n-1}. \tag{2.2}$$

Equations 2.1 and 2.2 are examples of *recurrence equations* or *recurrence relations*. A **recurrence relation** or simply a **recurrence** is an equation that expresses the nth term of a sequence a_n in terms of values of a_i for $i < n$. Thus Equations 2.1 and 2.2 are examples of recurrences.

2.2.1 Examples of recurrence relations

Other examples of recurrences are

$$a_n = a_{n-1} + 7, \tag{2.3}$$

$$a_n = 3a_{n-1} + 2^n, \tag{2.4}$$

$$a_n = a_{n-1} + 3a_{n-2}, \text{ and} \tag{2.5}$$

$$a_n = a_1 a_{n-1} + a_2 a_{n-2} + \cdots + a_{n-1} a_1. \tag{2.6}$$

A **solution** to a recurrence relation is a sequence that satisfies the recurrence relation. Thus a solution to Recurrence 2.1 is $s_n = 2^n$. Note that $s_n = 17 \cdot 2^n$ and $s_n = -13 \cdot 2^n$ are also solutions to Recurrence 2.1. What this shows is that a recurrence can have infinitely many solutions. In a given problem, there is generally one solution that is of interest to us. For example, if we are interested in the number of subsets of a set, then the solution to Recurrence 2.1 that we care about is $s_n = 2^n$. Notice this is the only solution we have mentioned that satisfies $s_0 = 1$.

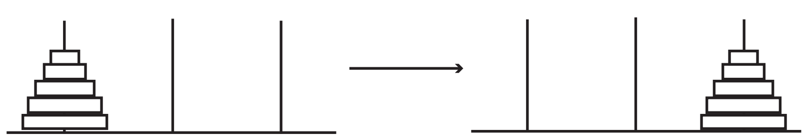
75. Show that there is only one solution to Recurrence 2.1 that satisfies $s_0 = 1$.

76. A first-order recurrence relation is one which expresses a_n in terms of a_{n-1} and other functions of n, but which does not include any of the terms a_i for $i < n - 1$ in the equation.

    (a) Which of the recurrences 2.1 through 2.6 are first order recurrences?

    (b) Show that there is one and only one sequence a_n that is defined for every nonnegative integer n, satisfies a first order recurrence, and satisfies $a_0 = a$ for some fixed constant a.

Figure 2.1: The Towers of Hanoi Puzzle



77. The "Towers of Hanoi" puzzle has three rods rising from a rectangular base with n rings of different sizes stacked in decreasing order of size on one rod. A legal move consists of moving a ring from one rod to another so that it does not land on top of a smaller ring. If m_n is the number of moves required to move all the rings from the initial rod to another rod that you choose, give a recurrence for m_n. (Hint: suppose you already knew the number of moves needed to solve the puzzle with $n-1$ rings.)

78. We draw n mutually intersecting circles in the plane so that each one crosses each other one exactly twice and no three intersect in the same point. (As examples, think of Venn diagrams with two or three mutually intersecting sets.) Find a recurrence for the number r_n of regions into which the plane is divided by n circles. (One circle divides the plane into two regions, the inside and the outside.) Find the number of regions with n circles. For what values of n can you draw a Venn diagram showing all the possible intersections of n sets using circles to represent each of the sets?

2.2.2 Arithmetic Series

79. A child puts away two dollars from her allowance each week. If she starts with twenty dollars, give a recurrence for the amount a_n of money she has after n weeks and find out how much money she has at the end of n weeks.

80. A sequence that satisfies a recurrence of the form $a_n = a_{n-1} + c$ is called an *arithmetic progression*. Find a formula in terms of the initial value a_0 and the common difference c for the term a_n in an arithmetic progression and prove you are right.

81. A person who is earning \$50,000 per year gets a raise of \$3000 a year for n years in a row. Find a recurrence for the amount a_n of money the person earns over $n+1$ years. What is the total amount of money that the person earns over a period of $n+1$ years? (In $n+1$ years, there are n raises.)

82. An *arithmetic series* is a sequence s_n equal to the sum of the terms a_0 through a_n of of an arithmetic progression. Find a recurrence for the

sum s_n of an arithmetic progression with initial value a_0 and common difference c (using the language of Problem 80). Find a formula for general term s_n of an arithmetic series.

2.2.3 First order linear recurrences

Recurrences such as those in Equations 2.1 through 2.5 are called *linear recurrences*, as are the recurrences of Problems 77 and 78. A **linear recurrence** is one in which a_n is expressed as a sum of functions of n times values of (some of the terms) a_i for $i < n$ plus (perhaps) another function (called the *driving function*) of n. A linear equation is called *homogeneous* if the driving function is zero (or, in other words, there is no driving function). It is called a constant coefficient linear recurrence if the functions that are multiplied by the a_i terms are all constants (but the driving function need not be constant).

83. Classify the recurrences in Equations 2.1 through 2.5 and Problems 77 and 78 according to whether or not they are constant coefficient, and whether or not they are homogeneous.

84. As you can see from Problem 83 some interesting sequences satisfy first order linear recurrences, including many that have constant coefficients, have constant driving term, or are homogeneous. Find a formula for the general term a_n of a sequence that satisfies a constant coefficient first order linear recurrence $a_n = ba_{n-1} + d$ in terms of b, d, a_0 and n and prove you are correct. If your formula involves a summation, try to replace the summation by a more compact expression.

2.2.4 Geometric Series

A sequence that satisfies a recurrence of the form $a_n = ba_{n-1}$ is called a *geometric progression*. Thus the sequence satisfying Equation 2.1, the recurrence for the number of subsets of an n-element set, is an example of a geometric progression. From your solution to Problem 84, a geometric progression has the form $a_n = a_0b^n$. In your solution to Problem 84 you may have had to deal with the sum of a geometric progression in just slightly different notation, namely $\sum_{i=0}^{n-1} db^i$. A sum of this form is called a **(finite) geometric series**.

85. Do this problem only if your final answer (so far) to Problem 84 contained the sum $\sum_{i=0}^{n-1} db^i$.

    (a) Expand $(1-x)(1+x)$? Expand $(1-x)(1+x+x^2)$. Expand $(1-x)(1+x+x^2+x^3)$.

    (b) What do you expect $(1-b)\sum_{i=0}^{n-1} db^i$ to be? What formula for $\sum_{i=0}^{n-1} db^i$ does this give you? Prove that you are correct.

In Problem 84 and perhaps 85 you proved an important theorem.

Theorem 2 *If $a_n = ba_{n-1} + d$, then $a_n = a_0 b^n + d\frac{1-b^n}{1-b}$.*

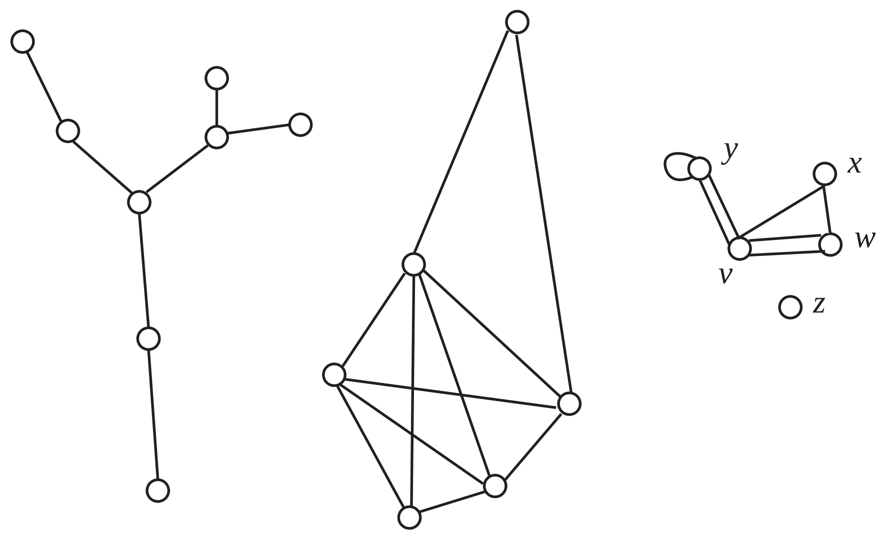
2.3 Trees

2.3.1 Undirected graphs

In Section 1.3.4 we introduced the idea of a directed graph. Graphs consist of vertices and edges. We describe vertices and edges in much the same way as we describe points and lines in geometry: we don't really say what vertices and edges are, but we say what they do. We just don't have a complicated axiom system the way we do in geometry. A **graph** consists of a set V called a vertex set and a set E called an edge set. Each member of V is called a *vertex* and each member of E is called an *edge*. Just as lines can connect points in geometry, edges can connect vertices in graph theory. We have one axiom like the axioms of geometry, namely, each edge connects two vertices. We draw pictures of graphs much like we draw pictures of geometric objects. In Figure 2.2 we show three pictures of graphs. Each circle in the figure represents a vertex; each line segment represents an edge. You will note that in the third graph we labelled the vertices; these labels are names we chose to give the vertices. We can choose names or not as we please. The third graph also shows that it is possible to have an edge that connects a vertex (like the one labelled y) to itself or it is possible to have two or more edges (like those between vertices v and y) between two vertices. The *degree* of a vertex is the number of times it appears as the endpoint of edges; thus the degree of y in the third graph in the figure is four.

86. The sum of the degrees of the vertices of a (finite) graph is related in a natural way to the number of edges. What is the relationship? Prove

Figure 2.2: Three different graphs



you are right. (Try to formulate your proof both with and without induction.)

2.3.2 Walks and paths in graphs

A *walk* in a graph is an alternating sequence $v_0e_1v_1 \ldots e_iv_i$ of vertices and edges such that edge e_i connects vertices v_{i-1} and v_i. A graph is called connected if, for any pair of vertices, there is a walk starting at one and ending at the other.

87. Which of the graphs in Figure 2.2 is connected?

88. A *path* in a graph is a walk with no repeated vertices. Find the longest path you can in the third graph of Figure 2.2.

89. A *cycle* in a graph is a walk whose first and last vertex are equal with no other repeated vertices. Which graphs in Figure 2.2 have cycles? What is the largest number of edges in a cycle in the second graph in Figure 2.2? What is the smallest number of edges in a cycle in the third graph in Figure 2.2?

90. A connected graph with no cycles is called a **tree**. Which graphs, if any, in Figure 2.2 are trees?

2.3.3 Counting vertices, edges, and paths in trees

91. Draw some trees and on the basis of your examples, make a conjecture about the relationship between the number of vertices and edges in a tree. Prove your conjecture. (Hint: what happens if you choose an edge and delete it, but not its endpoints?)

92. What is the minimum number of vertices of degree one in a finite tree? What is it if the number of vertices is bigger than one? Prove that you are correct.

93. In a tree, given two vertices, how many paths can you find between them? Prove that you are correct.

94. How many trees are there on the vertex set $\{1, 2\}$? On the vertex set $\{1, 2, 3\}$? When we label the vertices of our tree, we consider the tree which has edges between vertices 1 and 2 and between vertices 2 and 3 different from the tree that has edges between vertices 1 and 3 and between 2 and 3. See Figure 2.3. How many (labelled) trees are there

Figure 2.3: These two trees are different

on four vertices? You don't have a lot of data to guess from, but try to guess a formula for the number of trees with vertex set $\{1, 2, \cdots, n\}$. (If you organize carefully, you can figure out how many labelled trees there are with vertex set $\{1, 2, 3, 4, 5\}$ to help you make your conjecture.) Given a tree with two or more vertices, labelled with positive integers, define a sequence of integers inductively as follows: If the tree has two vertices, the sequence consists of one entry, namely the vertex with the larger label. Otherwise, let a_1 be the lowest numbered vertex of degree 1 in the tree. Let b_1 be the unique vertex in the tree adjacent to a_1

and write down b_1. Then write down the sequence corresponding to the tree you get when you delete a_1 from the tree. (If you are unfamiliar with inductive (recursive) definition, you might want to write down some labelled trees on, say, ten vertices and construct the sequence b.) How long will the sequence be if it is applied to a tree with n vertices (labelled with 1 through n)? What can you say about the last member of the sequence of b_is? Can you tell from the sequence of b_is what a_1 is? Find a bijection between labelled trees and something you can "count" that will tell you how many labelled trees there are on n labelled vertices.

2.3.4 Spanning trees

Many of the applications of trees arise from trying to find an efficient way to connect all the vertices of a graph. For example, in a telephone network, at any given time we have a certain number of wires (or microwave channels, or cellular channels) available for use. These wires or channels go from a specific place to a specific place. Thus the wires or channels may be thought of as edges of a graph and the places the wires connect may be thought of as vertices of that graph. A tree whose edges are some of the edges of a graph G and whose vertices are all of the vertices of the graph G is called a **spanning tree** of G. A spanning tree for a telephone network will give us a way to route calls between any two vertices in the network.

95. Show that every connected graph has a spanning tree. Can you give two essentially different proofs (they needn't be completely different, but should be different in at least one significant aspect)?

2.3.5 Minimum cost spanning trees

Our motivation for talking about spanning trees was the idea of finding a minimum number of edges we need to connect all the edges of a communication network together. In many cases edges of a communication network come with costs associated with them. For example, one cell-phone operator charges another one when a customer of the first uses an antenna of the other. Suppose a company has offices in a number of cities and wants to put together a communication network connecting its various locations with high-speed computer communications, but to do so at minimum cost. Then it wants to take a graph whose vertices are the cities in which it has offices and whose edges represent possible communications lines between the cities. Of course there will not necessarily be lines between each pair of cities, and the company will not want to pay for a line connecting city i and city j if it can already connect them indirectly by using other lines it has chosen. Thus it will want to choose a spanning tree of minimum cost among all spanning trees of the communications graph. For reasons of this application, if we have a graph with numbers assigned to its edges, the sum of the numbers on the edges of a spanning tree of G will be called the *cost* of the spanning tree.

96. Describe a method (or better, two methods different in at least one aspect) for finding a spanning tree of minimum cost in a graph whose edges are labelled with costs, the cost on an edge being the cost for including that edge in a spanning tree.

2.3.6 The deletion/contraction recurrence for spanning trees

There are two operations on graphs that we can apply to get a recurrence (though a more general kind than those we have studied for sequences) which will let us compute the number of spanning trees of a graph. The operations each apply to an edge e of a graph G. The first is called *deletion*; we *delete* the edge e from the graph by removing it from the edge set. The second operation is called *contraction*. Contractions of three different edges in the same graph are shown in Figure 2.4. We *contract* the edge e with endpoints v and w as follows:

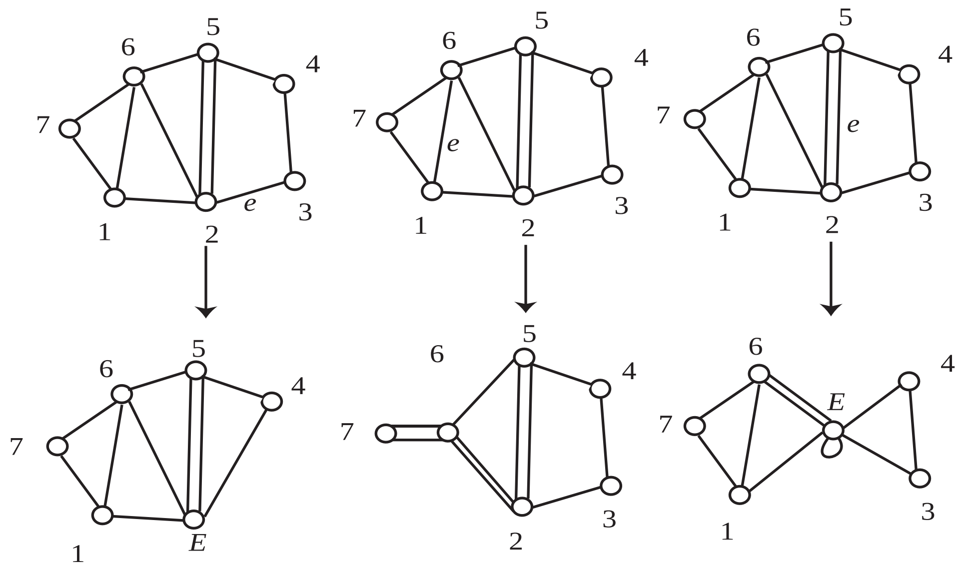
1. remove all edges having either v or w or both as an endpoint from the edge set,

2. remove v and w from the vertex set,

3. add a new vertex E to the vertex set,

4. add an edge from E to each remaining vertex that used to be an endpoint of an edge whose other endpoint was v or w, and add an edge from E to E for any edge other than e whose endpoints were in the set $\{v, w\}$.

Figure 2.4: The results of contracting three different edges in a graph.



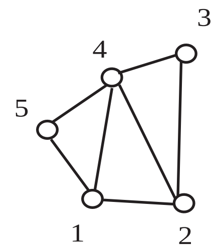
We use $G-e$ (read as G minus e) to stand for the result of deleting e from G, and we use G/E (read as G contract e) to stand for the result of contracting G from e.

97. How do the number of spanning trees of G not containing the edge e and the number of spanning trees of G containing e relate to the number of spanning trees of $G-e$ and G/e? Use $\#(G)$ to stand for the number of spanning trees of G (so that, for example, $\#(G/e)$ stands for the number of spanning trees of G/e. Find an expression for $\#(G)$ in terms of $\#(G/e)$ and $\#(G-e)$. This expression is called the *deletion-contraction recurrence*. Use it to compute the number of spanning trees of the graph in Figure 2.5.

2.3.7 Shortest paths in graphs

Suppose that a company has a main office in one city and regional offices in other cities. Most of the communication in the company is between the main office and the regional offices, so the company wants to find a spanning tree that minimizes not the total cost of all the edges, but rather the cost of

Figure 2.5: A graph.



communication between the main office and each of the regional offices. It is not clear that such a spanning tree even exists. This problem is a special case of the following. We have a connected graph with numbers assigned to its edges. (In this situation these numbers are often called weights.) The *(weighted) length* of a path in the graph is the sum of the weights of its edges. The *distance* between two vertices is the least (weighted) length of any path between the two vertices. Given a vertex v, we would like to know the distance between v and each other vertex, and we would like to know if there is a spanning tree in G such that the length of the path in the spanning tree from v to each vertex x is the distance from v to x in G.

98. Show that the following algorithm (known as Dijkstra's algorithm) applied to a weighted graph whose vertices are labelled 1 to n gives, for each i, the distance from vertex 1 to v as $d(i)$.

    (a) Let $d(1) = 0$. Let $d(i) = \infty$ for all other i. Let $v(1)$=1. Let $v(j) = 0$ for all other j. For each i and j, let $w(i,j)$ be the minimum weight of an edge between i and j, or ∞ if there are no such edges. Let $k = 1$. Let $t = 1$.

    (b) For each i, if $d(i) > d(k) + w(k,i)$ let $d(i) = d(k) + w(k,i)$.

    (c) Among those i with $v(i) = 0$, choose one with $d(i)$ a minimum, and let $k = i$. Increase t by 1.

    (d) Repeat the previous steps until $t = n$

99. Is there a spanning tree such that the distance from vertex 1 to vertex i given by the algorithm in Problem 98 is the distance for vertex 1 to vertex i in the tree (using the same weights on the edges, of course)?

2.4 Supplementary Problems

1. A hydrocarbon molecule is a molecule whose only atoms are either carbon atoms or hydrogen atoms. In a simple molecular model of a hydrocarbon, a carbon atom will bond to exactly four other atoms and hydrogen atom will bond to exactly one other atom. We represent a hydrocarbon compound with a graph whose vertices are labelled with C's and H's so that each C vertex has degree four and each H vertex has degree one. A hydrocarbon is called an "alkane" (common examples are methane (natural gas), propane, hexane (ordinary gasoline), octane (to make gasoline burn more slowly), etc.) if each C vertex is adjacent to four distinct vertices and the graph is a tree. How many different alkanes have exactly n vertices labelled C? (Here we say two trees are the same if we can make their drawings congruent by shortening and lengthening lines, or moving the vertices and edges around, making sure that after we move things around, the edges are attached to the same vertices as before.)

2. (a) Give a recurrence for the number of ways to divide $2n$ people into sets of two for tennis games. (Don't worry about who serves first.)

   (b) Give a recurrence for the number of ways to divide $4n$ people into sets of four for games of bridge. (Don't worry about how they sit around the bridge table or who is the first dealer.)

3. Use induction to prove your result in Supplementary Problem 2 at the end of Chapter 1.

Chapter 3

Distribution Problems

3.1 The idea of a distribution

Many of the problems we solved in Chapter 1 may be thought of as problems of distributing objects (such as pieces of fruit or ping-pong balls) to recipients (such as children). Some of the ways of viewing counting problems as distribution problems are somewhat indirect. For example, in Problem 31 you probably noticed that the number of ways to pass out k ping-pong balls to n children is the number of ways that we may choose a k-element subset of an n-element set. We think of the children as recipients and objects we are distributing as the identical ping-pong balls, distributed so that each recipient gets at most one ball. Those children who receive an object are in our set. It is helpful to have more than one way to think of solutions to problems. In the case of distribution problems, another popular model for distributions is to think of putting balls in boxes rather than distributing objects to recipients. Passing out identical objects is modeled by putting identical balls into boxes. Passing out distinct objects is modeled by putting distinct balls into boxes.

3.1.1 The twenty-fold way

When we are passing out objects to recipients, we may think of the objects as being either identical or distinct. We may also think of the recipients as being either identical (as in the case of putting fruit into plastic bags in the grocery store) or distinct (as in the case of passing fruit out to children). We may restrict the distributions to those that give at least one object to

Table 3.1: An incomplete table of the number of ways to distribute k objects to n recipients, with restrictions on how the objects are received

The Twentyfold Way: A Table of Distribution Problems		
k objects and conditions on how they are received	n recipients and mathematical model for distribution	
	Distinct	Identical
1. Distinct no conditions	n^k functions	? set partitions ($\leq n$ parts)
2. Distinct Each gets at most one	$n^{\underline{k}}$ k-element permutations	1 if $k \leq n$; 0 otherwise
3. Distinct Each gets at least one	? onto functions	? set partitions (n parts)
4. Distinct Each gets exactly one	$k! = n!$ bijections	1 if $k = n$; 0 otherwise
5. Distinct, order matters	? ?	? ?
6. Distinct, order matters Each gets at least one	? ?	? ?
7. Identical no conditions	? ?	? ?
8. Identical Each gets at most one	$\binom{n}{k}$ subsets	1 if $k \leq n$; 0 otherwise
9. Identical Each gets at least one	? ?	? ?
10. Identical Each gets exactly one	1 if $k = n$; 0 otherwise	1 if $k = n$; 0 otherwise

each recipient, or those that give exactly one object to each recipient, or those that give at most one object to each recipient, or we may have no such restrictions. If the objects are distinct, it may be that the order in which the objects are received is relevant (think about putting books onto the shelves in a bookcase) or that the order in which the objects are received is irrelevant (think about dropping a handful of candy into a child's trick or treat bag). If we ignore the possibility that the order in which objects are received matters, we have created $2 \cdot 2 \cdot 4 = 16$ distribution problems. In the cases where a recipient can receive more than one distinct object, we also have four more problems when the order objects are received matters. Thus we have 20 possible distribution problems.

We describe these problems in Table 3.1. Since there are twenty possible distribution problems, we call the table the "Twentyfold Way," adapting terminology suggested by Joel Spencer for a more restricted class of distribution problems. In the first column of the table we state whether the objects are distinct (like people) or identical (like ping-pong balls) and then give any conditions on how the objects may be received. The conditions we consider are whether each recipient gets at most one object, whether each recipient gets at least one object, whether each recipient gets exactly one object, and whether the order in which the objects are received matters. In the second column we give the solution to the problem and the name of the mathematical model for this kind of distribution problem when the recipients are distinct, and in the third column we give the same information when the recipients are identical. We use question marks as the answers to problems we have not yet solved and models we have not yet studied. We give explicit answers to problems we solved in Chapter 1 and problems whose answers are immediate. The goal of this chapter is to develop methods that will allow us to fill in the table with formulas or at least quantities we know how to compute, and we will give a completed table at the end of the chapter. We will now justify the answers that are not question marks.

If we pass out k distinct objects (say pieces of fruit) to n distinct recipients (say children), we are saying for each object which recipient it goes to. Thus we are defining a function from the set of objects to the recipients. We saw the following theorem in Problem 21c.

Theorem 3 *There are n^k functions from a k-element set to an n-element set.*

We proved it in Problem 21e. If we pass out k distinct objects (say pieces of fruit) to n indistinguishable recipients (say identical paper bags) then we are dividing the objects up into disjoint sets; that is we are forming a partition of the objects into some number, certainly no more than the number k of objects, of parts. Later in this chapter (and again in the next chapter) we shall discuss how to compute the number of partitions of a k-element set into n parts. This explains the entries in row one of our table.

If we pass out k distinct objects to n recipients so that each gets at most one, we still determine a function, but the function must be one-to-one. The number of one-to-one functions from a k-element set to an n element set is the same as the number of one-to-one functions from the set $[k] = \{1, 2, \ldots, k\}$ to an n-element set. In Problem 24 we proved the following theorem.

Theorem 4 *If $0 \le k \le n$, then the number of k-element permutations of an*

n-element set is

$$n^{\underline{k}} = n(n-1)\cdots(n-k+1) = n!/(n-k)!.$$

If $k > n$ there are no one-to-one functions from a k element set to an n element, so we define $n^{\underline{k}}$ to be zero in this case. Notice that this is what the indicated product in the middle term of our formula gives us. If we are supposed distribute k distinct objects to n identical recipients so that each gets at most one, we cannot do so if $k > n$, so there are 0 ways to do so. On the other hand, if $k \leq n$, then it doesn't matter which recipient gets which object, so there is only one way to do so. This explains the entries in row two of our table.

If we distribute k distinct objects to n distinct recipients so that each recipient gets at least one, then we are counting functions again, but this time functions from a k-element set *onto* an n-element set. At present we do not know how to compute the number of such functions, but we will discuss how to do so later in this chapter and in the next chapter. If we distribute k identical objects to n recipients, we are again simply partitioning the objects, but the condition that each recipient gets at least one means that we are partitioning the objects into exactly n blocks. Again, we will discuss how compute the number of ways of partitioning a set of k objects into n blocks later in this chapter. This explains the entries in row three of our table.

If we pass out k distinct objects to n recipients so that each gets exactly one, then $k = n$ and the function that our distribution gives us is a bijection. The number of bijections from an n-element set to an n-element set is $n!$ by Theorem 4. If we pass out k distinct objects of n identical recipients so that each gets exactly 1, then in this case it doesn't matter which recipient gets which object, so the number of ways to do so is 1 if $k = n$. If $k \neq n$, then the number of such distributions is zero. This explains the entries in row four of our table.

We now jump to row eight of our table. We saw in Problem 31 that the number of ways to pass out k identical ping-pong balls to n children is simply the number of k-element subsets of a k-element set. In Problem 33 we proved the following theorem.

Theorem 5 *If $0 \leq k \leq n$, the number of k-element subsets of an n-element* *set is given by*

$$\binom{n}{k} = \frac{n^{\underline{k}}}{k!} = \frac{n!}{k!(n-k)!}.$$

We define $\binom{n}{k}$ to be 0 if $k > n$, because then there are no k-element subsets of an n-element set. Notice that this is what the middle term of the formula in the theorem gives us. This explains the entries of row 8 of our table.

In row 10 of our table, if we are passing out k identical objects to n recipients so that each gets exactly one, it doesn't matter whether the recipients are identical or not; there is only one way to pass out the objects if $k = n$ and otherwise it is impossible to make the distribution, so there are no ways of distributing the objects. This explains the entries of row 10 of our table. Several other rows of our table can be computed using the methods of Chapter 1.

3.1.2 Ordered functions

100. Suppose we wish to place k distinct books onto the shelves of a bookcase with n shelves. For simplicity, assume for now that all of the books would fit on any of the shelves. Since the books are distinct, we can think of them as the first book, the second book and so on, perhaps in a stack. How many places are there where we can place the first book? When we place the second book, if we decide to place it on the shelf that already has a book, does it matter if we place it to the left of the book that is already there? How many places are there where we can place the second book? Once we have some number of books placed, if we want to place a new book on a shelf that already has some books, does it matter whether we place the new book to the immediate left of a book already on the shelf? In how many ways may we place the ith book into the bookcase? In how many ways may we place all the books?

101. Suppose we wish to place the books in Problem 100 so that each shelf gets at least one book. Now in how many ways may we place the books? (Hint: how can you make sure that each shelf gets at least one book as you start out putting books on the shelves?)

The assignment of which books go to which shelves of a bookcase is simply a function from the books to the shelves. But a function doesn't determine which book sits to the left of which others on the shelf, and this information is part of how the books are arranged on the shelves. In other words, the order in which the shelves receive their books matters. Our function must thus assign an ordered list of books to each shelf. We will call such a function an ordered function. More precisely, an **ordered function** from a set S to a set T is a function that assigns an (ordered) list of elements of S to some, but not necessarily all, elements of T in such a way that each element of S appears on one and only one of the lists.[1] (Notice that although it is not the usual definition of a function from S to T, a function can be described as an assignment of subsets of S to some, but not necessarily all, elements of T so that each element of S is in one and only one of these subsets.) Thus the number of ways to place the books into the bookcase is the entry in the middle column of row 5 of our table. If in addition we require each shelf to get at least one book, we are discussing the entry in the middle column of row 6 of our table. An *ordered onto function* is one which assigns a list to each element of T.

3.1.3 Broken permutations and Lah numbers

102. In how many ways may we stack k distinct books into n identical boxes so that there is a stack in every box? (Hint: Imagine taking a stack of k books, and breaking it up into stacks to put into the boxes in the same order they were originally stacked. If you are going to use n boxes, in how many places will you have to break the stack up into smaller stacks, and how many ways can you do this?) (Alternate hint: How many different bookcase arrangements correspond to the same way of stacking k books into n boxes so that each box has at least one book?). The hints may suggest that you can do this problem in more than one way!

We can think of stacking books into identical boxes as partitioning the books and then ordering the blocks of the partition. This turns out not to be a useful computational way of visualizing the problem because the number of ways to order the books in the various stacks depends on the sizes of the stacks and not just the number of stacks. However this way of thinking actually led to the first hint in Problem 102. Instead of dividing a set up

[1] The phrase ordered function is not a standard one, because there is as yet no standard name for the result of an ordered distribution problem.

into nonoverlapping parts, we may think of dividing a *permutation* (thought of as a list) of our k objects up into n ordered blocks. We will say that a set of ordered lists of elements of a set S is a **broken permutation** of S if each element of S is in one and only one of these lists.[2] The number of broken permutations of a k-element set with n blocks is denoted by $L(n, k)$. The number $L(n, k)$ is called a *Lah Number*.

The Lah numbers are the solution to the question "In how many ways may we distribute k distinct objects to n identical recipients if order matters and each recipient must get at least one?" Thus they give the entry in row 6 and column 6 of our table. The entry in row 5 and column 6 of our table will be the number of broken permutations with less than or equal to n parts. Thus it is a sum of Lah numbers.

We have seen that ordered functions and broken permutations explain the entries in rows 5 and 6 or our table.

3.1.4 Compositions of integers

103. In how many ways may we put k identical books onto n shelves if each shelf must get at least one book?

104. A **composition** of the integer k into n parts is a list of n positive integers that add to k. How many compositions are there of an integer k into n parts?

105. Your answer in Problem 104 can be expressed as a binomial coefficient. This means it should be possible to interpret a composition as a subset of some set. Find a bijection between compositions of k into n parts and certain subsets of some set. Explain explicitly how to get the composition from the subset and the subset from the composition.

106. Explain the connection between compositions of k into n parts and the problem of distributing k identical objects to n recipients so that each recipient gets at least one.

The sequence of problems you just completed should explain the entry in the middle column of row 9 of our table of distribution problems.

---

[2]The phrase broken permutation is not standard, because there is no standard name for the solution to this kind of distribution problem.

3.1.5 Multisets

In the middle column of row 7 of our table, we are asking for the number of ways to distribute k identical objects (say ping-pong balls) to n distinct recipients (say children).

107. In how many ways may we distribute k identical books on the shelves of a bookcase with n shelves, assuming that any shelf can hold all the books?

108. A multiset chosen from a set S may be thought of as a subset with repeated elements allowed. For example the multiset of letters of the word Mississippi is $\{i, i, i, i, m, p, p, s, s, s, s\}$. To determine a multiset we must say how many times (including, perhaps, zero) each member of S appears in the multiset. The size of a multiset chosen from S is the total number of times any member of S appears. For example, the size of the multiset of letters of Mississippi is 11. What is the number of multisets of size k that can be chosen from an n-element set?

109. Your answer in the previous problem should be expressible as a binomial coefficient. Since a binomial coefficient counts subsets, find a bijection between subsets of something and multisets chosen from a set S.

110. How many solutions are there in nonnegative integers to the equation $x_1 + x_2 + \cdots + x_m = r$, where m and r are constants?

The sequence of problems you just completed should explain the entry in the middle column of row 7 of our table of distribution problems. In the next two sections we will give ways of computing the remaining entries.

3.2 Partitions and Stirling Numbers

We have seen how the number of partitions of a set of k objects into n blocks corresponds to the distribution of k distinct objects to n identical recipients. While there is a formula that we shall eventually learn for this number, it requires more machinery than we now have available. However there is a good method for computing this number that is similar to Pascal's equation. Now that we have studied recurrences in one variable, we will point out that

Pascal's equation is in fact a *recurrence in two variables*; that is it lets us compute $\binom{n}{k}$ in terms of values of $\binom{m}{i}$ in which either $m < n$ or $i < k$ or both. It was the fact that we had such a recurrence and knew $\binom{n}{0}$ and $\binom{n}{n}$ that let us create Pascal's triangle.

3.2.1 Stirling Numbers of the second kind

We use the notation $S(k, n)$ to stand for the number of partitions of a k element set with n blocks. For historical reasons, $S(k, n)$ is called a *Stirling number of the second kind*.

111. In a partition of the set $[k]$, the number k is either in a block by itself, or it is not. How does the number of partitions of $[k]$ with n parts in which k is in a block with other elements of $[k]$ compare to the number of partitions of $[k-1]$ into n blocks? Find a two variable recurrence for $S(n, k)$, valid for n and k larger than one.

112. What is $S(k, 1)$? What is $S(k, k)$? Create a table of values of $S(k, n)$ for k between 1 and 5 and n between 1 and k. This table is sometimes called *Stirling's Triangle (of the second kind)* How would you define $S(k, n)$ for the nonnegative values of k and n that are not both positive? Now for what values of k and n is your two variable recurrence valid?

113. Extend Stirling's triangle enough to allow you to answer the following question and answer it. (Don't fill in the rows all the way; the work becomes quite tedious if you do. Only fill in what you need to answer this question.) A caterer is preparing three bag lunches for hikers. The caterer has nine different sandwiches. In how many ways can these nine sandwiches be distributed into three identical lunch bags so that each bag gets at least one?

114. The question in Problem 113 naturally suggests a more realistic question; in how many ways may the caterer distribute the nine sandwich's into three identical bags so that each bag gets exactly three? Answer this question. (Hint, what if the question asked about six sandwiches and two distinct bags? How does having identical bags change the answer?)

115. In how many ways can we partition k items into n blocks so that we have k_i blocks of size i for each i? (Notice that $\sum_{i=1}^{k} k_i = n$ and $\sum_{i=1}^{k} ik_i = k$.)

116. Describe how to compute $S(n, k)$ in terms of quantities given by the formula you found in Problem 115.

117. Find a recurrence similar to the one in Problem 111 for the Lah numbers $L(n, k)$.

118. The total number of partitions of a k-element set is denoted by $B(k)$ and is called the k-th *Bell number*. Thus $B(1) = 1$ and $B(2) = 2$.

    (a) Show, by explicitly exhibiting the partitions, that $B(3) = 5$.

    (b) Find a recurrence that expresses $B(k)$ in terms of $B(n)$ for $n \leq k$ and prove your formula correct in as many ways as you can.

    (c) Find $B(k)$ for $k = 4, 5, 6$.

3.2.2 Stirling Numbers and onto functions

119. Given a function f from a k-element set K to an n-element set, we can define a partition of K by putting x and y in the same block of the partition if and only if $f(x) = f(y)$. How many blocks does the partition have if f is onto? How is the number of functions from a k-element set onto an n-element set relate to a Stirling number? Be as precise in your answer as you can.

120. Each function from a k-element set K to an n-element set N is a function from K onto *some* subset of N. If J is a subset of N of size j, you know how to compute the number of functions that map onto J in terms of Stirling numbers. Suppose you add the number of functions mapping onto J over all possible subsets J of N. What simple value should this sum equal? Write the equation this gives you.

121. In how many ways can the sandwiches of Problem 113 be placed into three distinct bags so that each bag gets at least one?

122. In how many ways can the sandwiches of Problem 114 be placed into distinct bags so that each bag gets exactly three?

123. (a) How many functions are there from a set K with k elements to a set N with n elements so that for each i from 1 to n, k_i elements of N are each the images of i different elements of K. (Said differently, we have k_1 elements of N that are images of one element of K, we have k_2 elements of N that are images of two elements of K, and in general k_i elements of N that are images of i elements of K.) (We say y is the image of x if $y = f(x)$.)

  (b) How many functions are there from a k-element set K to a set $N = \{y_1, y_2, \ldots y_n\}$ so that y_i is the image of j_i elements of K for each i from 1 to n. This number is called a *multinomial coefficient* and denoted by $\binom{k}{j_1, j_2, \ldots, j_n}$.

  (c) Explain how to compute the number of functions from a k-element set K onto an n-element set N by using multinomial coefficients.

  (d) What do multinomial coefficients have to do with expanding the kth power of a multinomial $x_1 + x_2 + \cdots + x_n$? This result is called the *multinomial theorem*

3.2.3 Stirling Numbers and bases for polynomials

124. Find a way to express n^k in terms of $k^{\underline{j}}$ for appropriate values j. Notice that $x^{\underline{j}}$ makes sense for a numerical variable x (that could range over the rational numbers, the real numbers, or even the complex numbers instead of only the nonnegative integers, as we are implicitly asuming n does), just as x^j does. Find a way to express the power x^k in terms of the polynomials $x^{\underline{j}}$ for appropriate values of j and explain why your formula is correct.

You showed in Problem 124 how to get each power of x in terms of the falling factorial powers $x^{\underline{j}}$. Therefore every polynomial in x is expressible in terms of a sum of numerical multiples of falling factorial powers. We say that the ordinary powers of x and the falling factorial powers of x are each bases for the "space" of polynomials, and that the numbers $S(k, n)$ are "change of basis coefficients."

125. Show that every power of $x + 1$ is expressible as a sum of powers of x. Now show that every power of x (and thus every polynomial in x) is a sum of numerical multiples (some of which could be negative) of powers

of $x+1$. This means that the powers of $x+1$ are a basis for the space of polynomials as well. Describe the change of basis coefficients that we use to express the binomial powers $(x+1)^n$ in terms of the ordinary x^j explicitly. Find the change of basis coefficients we use to express the ordinary powers x^n in terms of the binomial powers $(x+1)^k$.

126. By multiplication, we can see that every falling factorial polynomial can be expressed as a sum of numerical multiples of powers of x. In symbols, this means that there are numbers $s(k,n)$ (notice that this s is lower case, not upper case) such that we may write $x^{\underline{k}} = \sum_{n=0}^{k} s(k,n)x^n$. These numbers $s(k,n)$ are called Stirling Numbers of the first kind. By thinking algebraically about what the formula

$$x^{\underline{n}} = x^{\underline{n-1}}(x-n+1) \tag{3.1}$$

means, we can find a recurrence for Stirling numbers of the first kind that gives us another triangular array of numbers called Stirling's triangle of the first kind. Explain why Equation 3.1 is true and use it to derive a recurrence for $s(k,n)$ in terms of $s(k-1,n-1)$ and $s(k-1,n)$.

127. Write down the rows of Stirling's triangle of the first kind for $k=0$ to 6.

Notice that the Stirling numbers of the first kind are also change of basis coefficients. The Stirling numbers of the first and second kind are change of basis coefficients from the falling factorial powers of x to the ordinary factorial powers, and vice versa.

128. Explain why every rising factorial polynomial $x^{\overline{k}}$ can be expressed in terms of the falling factorial polynomials $x^{\underline{n}}$. Let $b(k,n)$ stand for the change of basis coefficients that allow us to express $x^{\overline{k}}$ in terms of the falling factorial polynomials $x^{\underline{n}}$; that is, define $b(k,n)$ by the equations

$$x^{\overline{k}} = \sum_{n=0}^{k} x^{\underline{n}}.$$

(a) Find a recurrence for $b(k,n)$.

(b) Find a formula for $b(k,n)$ and prove the correctness of what you say in as many ways as you can.

(c) Is $b(k,n)$ the same as any of the other families of numbers (binomial coefficients, Bell numbers, Stirling numbers, Lah numbers, etc.) we have studied?

(d) Say as much as you can (but say it precisely) about the change of basis coefficients for expressing $x^{\underline{k}}$ in terms of $x^{\overline{n}}$.

3.3 Partitions of Integers

We have now completed all our distribution problems except for those in which both the objects and the recipients are identical. For example, we might be putting identical apples into identical paper bags. In this case all that matters is how many bags get one apple (how many recipients get one object), how many get two, how many get three, and so on. Thus for each bag we have a number, and the multiset of numbers of apples in the various bags is what determines our distribution of apples into identical bags. A multiset of positive integers that add to n is called a **partition** of n. Thus the partitions of 3 are 1+1+1, 1+2 (which is the same as 2+1) and 3. The number of partitions of k is denoted by $P(k)$; in computing the partitions of 3 we showed that $P(3) = 3$.

129. Find all partitions of 4 and find all partitions of 5, thereby computing $P(4)$ and $P(5)$.

3.3.1 The number of partitions of k into n parts

130. A *partition of the integer k into n parts* is a multiset of n positive integers that add to k. We use $P(k,n)$ to denote the number of partitions of k into n parts. Thus $P(k,n)$ is the number of ways to distribute k identical objects to n identical recipients so that each gets at least one. Find $P(6,3)$ by finding all partitions of 6 into 3 parts. What does this say about the number of ways to put six identical apples into three identical bags so that each bag has at least one apple?

131. With the binomial coefficients, with Stirling numbers of the second kind, and with the Lah numbers, we were able to find a recurrence by asking what happens to our subset, partition, or broken permutation of a set S of numbers if we remove the largest element of S. Thus it

is natural to look for a recurrence to count the number of partitions of k into n parts by doing something similar. However since we are counting distributions in which all the objects are identical, there is no way for us to identify a largest element. However we can ask what happens to a partition of an integer when we remove its largest part, or if we remove one from every part.

(a) How many parts does the remaining partition have when we remove the largest part from a partition of k into n parts? What can you say about the number of parts of the remaining partition if we remove one from each part?

(b) If the largest part of a partition is j and we remove it, what integer is being partitioned by the remaining parts of the partition? If we remove one from each part of a partition of k into n parts, what integer is being partitioned by the remaining parts?

(c) Use the answers to the last two questions to find and describe a bijection between partitions of k into n parts and partitions of smaller integers into appropriate numbers of parts.

(d) Find a recurrence (which need not have just two terms on the right hand side) that describes how to compute $P(k, n)$ in terms of the number of partitions of smaller integers into a smaller number of parts.

(e) What is $P(k, 1)$ for a positive integer k?

(f) What is $P(k, k)$ for a positive integer k?

(g) Use your recurrence to compute a table with the values of $P(k, n)$ for values of k between 1 and 7.

It is remarkable that there is no known formula for $P(k, n)$, nor is there one for $P(k)$. This section and some future parts of these notes are devoted to developing methods for computing values of $p(n, k)$ and finding properties of $P(n, k)$ that we can prove even without knowing a formula.

3.3.2 Representations of partitions

132. How many solutions are there in the positive integers to the equation $x_1 + x_2 + x_3 = 7$ with $x_1 \geq x_2 \geq x_3$?

133. Explain the relationship between partitions of k into n parts and lists $x_1, x_2, \ldots, x_n$ of positive integers with $x_1 \geq x_2 \geq \ldots \geq x_n$. Such a representation of a partition is called a *decreasing list* representation of the partition.

134. Describe the relationship between partitions of k and lists or vectors $(x_1, x_2, \ldots, x_n)$ such that $x_1 + 2x_2 + \ldots kx_k = k$. Such a representation of a partition is called a *type vector* representation of a partition, and it is typical to leave the trailing zeros out of such a representation; for example $(2, 1)$ stands for the same partition as $(2, 1, 0, 0)$. What is the decreasing list representation for this partition, and what number does it partition?

135. How does the number of partitions of k relate to the number of partitions of $k + 1$ whose smallest part is one?

3.3.3 Ferrers and Young Diagrams and the conjugate of a partition

The decreasing list representation of partitions leads us to a handy way to visualize partitions. Given a decreasing list $(k_1, k_2, \ldots k_n)$, we draw a figure made up of rows of dots that has k_1 equally spaced dots in the first row, k_2 equally spaced dots in the second row, starting out right below the beginning of the first row and so on. Equivalently, instead of dots, we may use identical squares, drawn so that a square touches each one to its immediate right or immediately below it along an edge. See Figure 3.1 for examples. The figure we draw with dots is called the Ferrers diagram of the partition; sometimes the figure with squares is also called a Ferrers diagram; sometimes it is called a Young diagram. At this stage it is irrelevant which name we choose and which kind of figure we draw; in more advanced work the boxes are handy because we can put things like numbers or variables into them. From now on we will use boxes and call the diagrams Young diagrams.

136. Draw the Young diagram of the partition (4,4,3,1,1). Describe the geometric relationship between the Young diagram of (5,3,3,2) and the Young diagram of (4,4,3,1,1).

137. The partition $(k_1, k_2, \ldots, k_n)$ is called the *conjugate* of the partition $(j_1, j_2, \ldots, j_m)$ if we obtain the Young diagram of one from the Young

Figure 3.1: The Ferrers and Young diagrams of the partition (5,3,3,2)

diagram of the other by flipping one around the line with slope -1 that extends the diagonal of the top left box. What is the conjugate of (4,4,3,1,1)? How is the largest part of a partition related to the number of parts of its conjugate? What does this tell you about the number of partitions of a positive integer k with largest part m?

138. A partition is called *self-conjugate* if it is equal to its conjugate. Find a relationship between the number of self-conjugate partitions of k and the number of partitions of k into distinct odd parts.

139. Explain the relationship between the number of partitions of k into even parts and the number of partitions of k into parts of even multiplicity, i.e. parts which are each used an even number of times as in (3,3,3,3,2,2,1,1).

140. Show that $P(k,n)$ is at least $\frac{1}{n!}\binom{k-1}{n-1}$.

We have seen that the number of partitions of k into n parts is equal to the number of ways to distribute k identical objects to n recipients so that each receives at least one. If we relax the condition that each recipient receives at least one, then we see that the number of distributions of k identical objects to n recipients is $\sum_{i=1}^{n} P(k,i)$ because if some recipients receive nothing, it does not matter which recipients these are. This completes rows 7 and 8 of our table of distribution problems. The completed table is shown in Figure 3.2. There are quite a few theorems that you have proved which are summarized by Table 3.2. It would be worthwhile to try to write them all down!

Table 3.2: The number of ways to distribute k objects to n recipients, with restrictions on how the objects are received

The Twentyfold Way: A Table of Distribution Problems

k objects and conditions on how they are received	n recipients and mathematical model for distribution	
	Distinct	Identical
1. Distinct no conditions	n^k functions	$\sum_{i=1}^{k} S(n,i)$ set partitions ($\leq n$ parts)
2. Distinct Each gets at most one	$n^{\underline{k}}$ k-element permutations	1 if $k \leq n$; 0 otherwise
3. Distinct Each gets at least one	$S(k,n)n!$ onto functions	$S(k,n)$ set partitions (n parts)
4. Distinct Each gets exactly one	$k! = n!$ permutations	1 if $k = n$; 0 otherwise
5. Distinct, order matters	$(k+n-1)^{\underline{k}}$ ordered functions	$\sum_{i=1}^{n} L(k,i)$ broken permutations ($\leq n$ parts)
6. Distinct, order matters Each gets at least one	$(k)^{\underline{n}}(k-1)^{\underline{k-n}}$ ordered onto functions	$L(k,n) = \binom{k}{n}(k-1)^{\underline{k-n}}$ broken permutations (n parts)
7. Identical no conditions	$\binom{n+k-1}{k}$ multisets	$\sum_{i=1}^{n} P(k,i)$ number partitions ($\leq n$ parts)
8. Identical Each gets at most one	$\binom{n}{k}$ subsets	1 if $k \leq n$; 0 otherwise
9. Identical Each gets at least one	$\binom{k-1}{n-1}$ compositions (n parts)	$P(k,n)$ number partitions (n parts)
10. Identical Each gets exactly one	1 if $k = n$; 0 otherwise	1 if $k = n$; 0 otherwise

Chapter 4

Algebraic Counting Techniques

4.1 The Principle of Inclusion and Exclusion

4.1.1 The size of a union of sets

One of our very first counting principles was the sum principle which says that the size of a union of disjoint sets is the sum of their sizes. Computing the size of overlapping sets requires, quite naturally, information about how they overlap. Taking such information into account will allow us to develop a powerful extension of the sum principle known as the "principle of inclusion and exclusion."

141. In a biology lab study of the effects of basic fertilizer ingredients on plants, 16 plants are treated with potash, 16 plants are treated with phosphate, and among these plants, eight are treated with both phosphate and potash. No other treatments are used. How many plants receive at least one treatment? If 32 plants are studied, how many receive no treatment?

142. Give a formula for the size of the union $A \cup B$ of two sets A in terms of the sizes $|A|$ of A, $|B|$ of B, and $|A \cap B|$ of $A \cap B$. If A and B are subsets of some "universal" set U, express the size of the complement $U - (A \cup B)$ in terms of the sizes $|U|$ of U, $|A|$ of A, $|B|$ of B, and $|A \cap B|$ of $A \cap B$.

143. In Problem 141, there were just two fertilizers used to treat the sample plants. Now suppose there are three fertilizer treatments, and 15 plants

are treated with nitrates, 16 with potash, 16 with phosphate, 7 with nitrate and potash, 9 with nitrate and phosphate, 8 with potash and phosphate and 4 with all three. Now how many plants have been treated? If 32 plants were studied, how many received no treatment at all?

144. Give a formula for the size of $A_1 \cup A_2 \cup A_3$ in terms of the sizes of A_1, A_2, A_3 and the intersections of these sets.

145. Conjecture a formula for the size of a union of sets

$$A_1 \cup A_2 \cup \cdots \cup A_n = \bigcup_{i=1}^{n} A_i$$

in terms of the sizes of the sets A_i and their intersections.

The difficulty of generalizing Problem 144 to Problem 145 is not likely to be one of being able to see what the right conjecture is, but of finding a good notation to express your conjecture. In fact, it would be easier for some people to express the conjecture in words than to express it in a notation. Here is some notation that will make your task easier. Let us define

$$\bigcap_{i:i\in I} A_i$$

to mean the intersection over all elements i in the set I of A_i. Thus

$$\bigcap_{i:i\in\{1,3,4,6\}} = A_1 \cap A_3 \cap A_4 \cap A_6. \tag{4.1}$$

This kind of notation, consisting of an operator with a description underneath of the values of a dummy variable of interest to us, can be extended in many ways. For example

$$\begin{aligned}\sum_{I:I\subseteq\{1,2,3,4\},\ |I|=2} |\cap_{i\in I} A_i| &= |A_1 \cap A_2| + |A_1 \cap A_3| + |A_1 \cap A_4| \\ &+ |A_2 \cap A_3| + |A_2 \cap A_4| + |A_3 \cap A_4|. \end{aligned} \tag{4.2}$$

146. Use notation something like that of Equation 4.1 and Equation 4.2 to express the answer to Problem 145. Note there are many different correct ways to do this problem. Try to write down more than one and

choose the nicest one you can. Say why you chose it (because your view of what makes a formula nice may be different from somebody else's). The nicest formula won't necessarily involve all the elements of Equations 4.1 and 4.2.

147. A group of n students goes to a restaurant carrying backpacks. The manager invites everyone to check their backpack at the check desk and everyone does. While they are eating, a child playing in the check room randomly moves around the claim check stubs on the backpacks. What is the probability that, at the end of the meal, at least one student receives his or her own backpack? In other words, in what fraction of the total number of ways to pass the backpacks back does at least one student get his or her own backpack back? (It might be a good idea to first consider cases with $n = 3$, 4, and 5. Hint: For each student, how big is the set of backpack distributions in which that student gets the correct backpack?) What is the probability that no student gets his or her own backpack?

148. As the number of students becomes large, what does the probability that no student gets the correct backpack approach?

The formula you have given in Problem 146 is often called **the principle of inclusion and exclusion** for unions of sets. The reason is the pattern in which the formula first adds (includes) all the sizes of the sets, then subtracts (excludes) all the sizes of the intersections of two sets, then adds (includes) all the sizes of the intersections of three sets, and so on. Notice that we haven't yet proved the principle. We will first describe the principle in an apparently more general situation that doesn't require us to translate each application into the language of sets. While this new version of the principle might seem more general than the principle for unions of sets; it is equivalent. However once one understands the notation used to express it, it is more convenient to apply.

Problem 147 is sometimes called the *hatcheck problem*; the name comes from substituting hats for backpacks. If is also sometimes called the *derangement problem*. A *derangement* of an n-element set is a permutation of that set (thought of as a bijection) that maps no element of the set to itself. One can think of a way of handing back the backpacks as a permutation f of the students: $f(i)$ is the owner of the backpack that student i receives. Then a derangement is a way to pass back the backpacks so that no student gets his or her own.

4.1.2 The hatcheck problem restated

The last question in Problem 147 requires that we compute the number of ways to hand back the backpacks so that nobody gets his or her own backpack. We can think of the set of ways to hand back the backpacks so that student i gets the correct one as the set of permutations of the backpacks with the property that student i gets his or her own backpack. Since there are $n-1$ other students and they can receive any of the remaining $n-1$ backpacks in $(n-1)$ ways, the number of permutations with the property that student i gets the correct backpack is $(n-1)!$. How many permutations are there with the properties that student i gets the correct backpack *and* student j gets the correct backpack? (Let's call these properties i and j for short.) Since there are $n-2$ remaining students and $n-2$ remaining backpacks, the number of permutations with properties i and j is $(n-2)!$. Similarly, the number of permutations with properties $i_1, i_2, \ldots, i_k$ is $(n-k)!$. Thus when we compute the size of the union of the sets

$$S_i = \{f : f \text{ is a permutation with property } i\},$$

we are computing the number of ways to pass back the backpacks so that at least one student gets the correct backpack. This answers the first question in Problem 147. The last question in Problem 147 is asking us for the number of ways to pass back the backpacks that have *none* of the properties. To say this in a different way, the question is asking us to compute the number of ways of passing back the backpacks that have exactly the *empty set*, $\emptyset$, of properties.

4.1.3 Basic counting functions: $N_{\textbf{at least}}$ and $N_{\textbf{exactly}}$

Notice that the quantities that we were able to count easily were the number of ways to pass back the backpacks so that we satisfy a certain subset $K = \{i_1, i_2, \ldots, i_k\}$ of our properties. In fact, among the $(n-k)$ ways to pass back the backpacks with this particular set K of properties is the permutation that gives each student the correct backpack, and has not just the properties in K, but the whole set of properties. Similarly, for any set M of properties with $K \subseteq M$, the permutations that have all the properties in M are among the $(n-k)!$ permutations that have the properties in the set K. Thus we can think of $(n-k)!$ as counting the number of permutations that have *at least* the properties in K. In particular, $n!$ is the number of ways to pass back the backpacks that have at least the empty set of properties. We thus write $N_{\text{at least}}(\emptyset) = n!$, or $N_{\text{a}}(\emptyset) = n!$ for short. For a k-element subset K of the properties, we write $N_{\text{at least}}(K) = (n-k)!$ or $N_{\text{a}}(K) = (n-k)!$ for short.

The question we are trying to answer is "How many of the distributions of backpacks have exactly the empty set of properties?" For this purpose we introduce one more piece of notation. We use $N_{\text{exactly}}(\emptyset)$ or $N_{\text{e}}(\emptyset)$ to stand for the number of backpack distributions with exactly the empty set of properties, and for any set K of properties we use $N_{\text{exactly}}(K)$ or $N_{\text{e}}(K)$ to stand for the number of backpack distributions with exactly the set K of properties. Thus $N_{\text{e}}(K)$ is the number of distribution in which the students represented by the set K of properties get the correct backpacks back and no other students do.

4.1.4 The principle of inclusion and exclusion for properties

For the principle of inclusion and exclusion for properties, suppose we have a set of arrangements (like backpack distributions) and a set P of properties (like student i gets the correct backpack) that the arrangements might or might not have. We suppose that we know (or can easily compute) the numbers $N_{\text{a}}(K)$ for every subset K of P. We are most interested in computing $N_{\text{e}}(\emptyset)$, the number of arrangements with none of the properties, but it will turn out that with no more work we can compute $N_{\text{e}}(K)$ for every subset K of P. Based on our answer to Problem 147 we expect that

$$N_{\text{e}}(\emptyset) = \sum_{S:S\subseteq P} (-1)^{|S|} N_{\text{a}}(S) \tag{4.3}$$

and it is a natural guess that

$$N_{\text{e}}(K) = \sum_{S:K\subseteq S\subseteq P} (-1)^{|S|-|K|} N_{\text{a}}(S). \tag{4.4}$$

Equations 4.3 and 4.4 are called **the principle of inclusion and exclusion for properties**.

149. Verify that the formula for the number of ways to pass back the backpacks in Problem 147 so that nobody gets the correct backpack has the form of Equation 4.3.

150. Find a way to express $N_{\mathrm{a}}(S)$ in terms of $N_{\mathrm{e}}(J)$ for subsets J of P containing S. In particular, what is the equation that expresses $N_{\mathrm{a}}(\emptyset)$ in terms of $N_{\mathrm{e}}(J)$ for subsets J of P?

151. Substitute the formula for N_{a} from Problem 150 into the right hand sides of the formulas of Equations 4.3 and 4.4 and simplify what you get to show that for Equations 4.3 and 4.4 the right-hand sides are indeed equal to the left-hand sides. This will prove that those equations are true. (Hint: You will get a double sum. If you can figure out how to reverse the order of the two summations, the binomial theorem may help you simplify the formulas you get.)

152. In how many ways may we distribute k identical apples to n children so that no child gets more than three apples?

4.1.5 Counting onto functions

153. Given a function f from the k-element set K to the n-element set $[n]$, we say f has property i if $f(x) \neq i$ for every x in K. How many of these properties does an onto function have? What is the number of functions from a k-element set onto an n-element set?

154. Find a formula for the Stirling number (of the second kind) $S(k, n)$.

4.1.6 The chromatic polynomial of a graph

We defined a graph to consist of set V of elements called vertices and a set E of elements called edges such that each edge joins two vertices. A *coloring* of a graph by the elements of a set C (of colors) is an assignment of an element of C to each vertex of the graph; that is, a function from the vertex set V of the graph to C. A coloring is called *proper* if for each edge joining two distinct vertices[1], the two vertices it joins have different colors. You may have heard

[1]If a graph had a loop connecting a vertex to itself, that loop would connect a vertex to a vertex of the same color. it is because of this that we only consider edges with two distinct vertices in our definition

of the famous four color theorem of graph theory that says if a graph may be drawn in the plane so that no two edges cross (though they may touch at a vertex), then the graph has a proper coloring with four colors. Here we are interested in a different, though related, problem: namely, in how many ways may we properly color a graph (regardless of whether it can be drawn in the plane or not) using k or fewer colors? When we studied trees, we restricted ourselves to connected graphs. (Recall that a graph is connected if, for each pair of vertices, there is a walk between them.) Here, disconnected graphs will also be important to us. Given a graph which might or might not be connected, we partition its vertices into blocks called *connected components* as follows. For each vertex v we put all vertices connected to it by a walk into a block together. Clearly each vertex is in at least one block, because vertex v is connected to vertex v by the trivial walk consisting of the single vertex v and no edges. To have a partition, each vertex must be in one and only one block. To prove that we have defined a partition, suppose that vertex v is in the blocks B_1 and B_2. Then B_1 is the set of all vertices connected by walks to some vertex v_1 and B_2 is the set of all vertices connected by walks to some vertex v_2.

155. Show that $B_1 = B_2$.

Since $B_1 = B_2$, these two sets are the same block, and thus all blocks containing v are identical, so v is in only one block. Thus we have a partition of the vertex set, and the blocks of the partition are the connected components of the graph. Notice that the connected components depend on the edge set of the graph. If we have a graph on the vertex set V with edge set E and another graph on the vertex set V with edge set E', then these two graphs could have different connected components. It is traditional to use the Greek letter γ (gamma)[2] to stand for the number of connected components of a graph; in particular, $\gamma(V, E)$ stands for the number of connected components of the graph with vertex set V and edge set E. We are going to show how the principle of inclusion and exclusion may be used to compute the number of ways to properly color a graph using colors from a set C of c colors.

156. Suppose we have a graph G with vertex set V and edge set E. Suppose F is a subset of E. Suppose we have a set C of c colors with which to color the vertices.

[2]The greek letter gamma is pronounced gam-uh, where gam rhymes with ham.

(a) In terms of $\gamma(V, F)$, in how many ways may we color the vertices of g so that each edge in F connects two vertices of the same color?

(b) Given a coloring of G, for each edge e in E, let us consider the property that the endpoints of e are colored the same color. Let us call this property "property e." In this way each set of properties can be thought of as a subset of E. What set of properties does a proper coloring have?

(c) Find a formula (which may involve summing over all subsets F of the edge set of the graph and using the number $\gamma(V, F)$ of connected components of the graph with vertex set V and edge set F) for the number of proper colorings of G using colors in the set C.

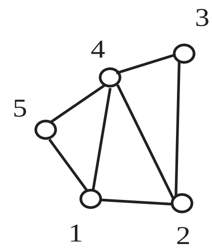
The formula you found in Problem 156c is a formula that involves powers of c, and so it is a polynomial function of c. Thus it is called the *chromatic polynomial* of the graph G. Since we like to think about polynomials as having a variable x and we like to think of c as standing for some constant, people often use x as the notation for the number of colors we are using to color G. Frequently people will use $\chi_G(x)$ to stand for the number of way to color G with x colors, and call $\chi_G(x)$ the chromatic polynomial of G.

157. In Chapter 2 we introduced the deletion-contraction recurrence for counting spanning trees of a graph. Figure out how the chromatic polynomial of a graph is related to those resulting from deletion of an edge e and from contraction of that same edge e. Try to find a recurrence like the one for counting spanning trees that expresses the chromatic polynomial of a graph in terms of the chromatic polynomials of $G - e$ and G/e for an arbitrary edge e. Use this recurrence to give another proof that the number of ways to color a graph with x colors is a polynomial function of x.

158. Use the deletion-contraction recurrence to compute the chromatic polynomials of the graph in Figure 4.1. (You can simplify your computations by thinking about the effect on the chromatic polynomial of deleting an edge that is a loop, or deleting one of several edges between the same two vertices.)

159. In how many ways may you properly color the vertices of a path on n vertices with x colors? Describe any dependence of the chromatic

Figure 4.1: A graph.



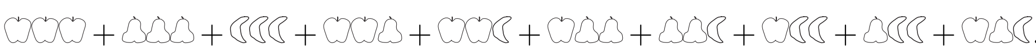
polynomial of a path on the number of vertices. In how many ways may you properly color the vertices of a cycle on n vertices with x colors? Describe any dependence of the chromatic polynomial of a cycle on the number of vertices.

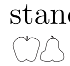
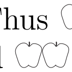
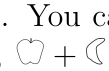
160. In how many ways may you properly color the vertices of a tree on n vertices with x colors?

161. What do you observe about the signs of the coefficients of the chromatic polynomial of the graph in Figure 4.1? What about the signs of the coefficients of the chromatic polynomial of a path? Of a cycle? Of a tree? Make a conjecture about the signs of the coefficients of a chromatic polynomial and prove it.

4.2 The Idea of Generating Functions

Suppose you are going to choose three pieces of fruit from among apples, pears and bananas for a snack. We can symbolically represent all your choices as



Here we are using a picture of a piece of fruit to stand for taking a piece of that fruit. Thus  stands for taking an apple,  for taking an apple and a pear, and  for taking two apples. You can think of the plus sign as standing for the "exclusive or," that is,  would stand for "I take an apple or a banana but not both." To say "I take both an apple and a banana," we would write 🍎🍌. We can extend the analogy to mathematical notation by condensing our statement that we take three pieces of fruit to

$$\text{🍎}^3 + \text{🍐}^3 + \text{🍌}^3 + \text{🍎}^2\text{🍐} + \text{🍎}^2\text{🍌} + \text{🍎}\text{🍐}^2 + \text{🍐}^2\text{🍌} + \text{🍎}\text{🍌}^2 + \text{🍐}\text{🍌}^2 + \text{🍎}\text{🍐}\text{🍌}.$$

In this notation 🍎^3 stands for taking a multiset of three apples, while $\text{🍎}^2\text{🍌}$ stands for taking a multiset of two apples and a banana, and so on. What our notation is really doing is giving us a convenient way to list all three element multisets chosen from the set $\{\text{🍎}, \text{🍐}, \text{🍌}\}$.

Suppose now that we plan to choose between one and three apples, between one and two pears, and between one and two bananas. In a somewhat clumsy way we could describe our fruit selections as

$$\text{🍎}\text{🍐}\text{🍌}+\text{🍎}^2\text{🍐}\text{🍌}+\cdots+\text{🍎}^2\text{🍐}^2\text{🍌}+\cdots+\text{🍎}^2\text{🍐}^2\text{🍌}^2+\text{🍎}^3\text{🍐}\text{🍌}+\cdots+\text{🍎}^3\text{🍐}^2\text{🍌}+\cdots+\text{🍎}^3\text{🍐}^2\text{🍌}^2. \tag{4.5}$$

162. Using an A in place of the picture of an apple, a P in place of the picture of a pear, and a B in place of the picture of a banana, write out the formula similar to Formula 4.5 without any dots for left out terms. (You may use pictures instead of letters if you prefer, but it gets tedious quite quickly!) Now expand the product $(A+A^2+A^3)(P+P^2)(B+B^2)$ and compare the result with your formula.

163. Substitute x for all of A, P and B (or for the corresponding pictures) in the formula you got in Problem 162 and expand the result in powers of x. Give an interpretation of the coefficient of x^n.

If we were to expand the formula

$$(\text{🍎} + \text{🍎}^2 + \text{🍎}^3)(\text{🍐} + \text{🍐}^2)(\text{🍌} + \text{🍌}^2). \tag{4.6}$$

we would get Formula 4.5. Thus formula 4.5 and formula 4.6 each describe the number of multisets we can choose from the set 🍎, 🍐, 🍌 in which 🍎 appears between 1 and three times and 🍐, and 🍌 each appear once or twice. We interpret Formula 4.5 as describing each individual multiset we can choose, and we interpret Formula 4.6 as saying that we first decide how many apples to take, and then decide how many pears to take, and then decide how many bananas to take. At this stage it might seem a bit magical that doing ordinary algebra with the second formula yields the first, but in fact we could define addition and multiplication with these pictures more formally so we could explain in detail why things work out. However since the pictures are for motivation, and are actually difficult to write out on paper, it doesn't make much sense to work out these details. We will see an explanation in another context later on.

4.2.1 Picture functions

As you've seen, in our descriptions of ways of choosing fruits, we've treated the pictures of the fruit as if they are variables. You've also likely noticed that it is much easier to do algebraic manipulations with letters rather than pictures, simply because it is time consuming to draw the same picture over and over again, while we are used to writing letters quickly. In the theory of generating functions, we associate variables or polynomials or even power series with members of a set. There is no standard language describing how we associate variables with members of a set, so we shall invent some. By a *picture* of a member of a set we will mean a variable, or perhaps a product of powers of variables (or even a sum of products of powers of variables). A function that assigns a picture $P(s)$ to each member s of a set S will be called a *picture function*. The **picture enumerator** for a picture function P defined on a set S will be

$$E_P(S) = \sum_{s:s\in S} P(s).$$

We choose this language because the picture enumerator lists, or enumerates, all the elements of S according to their pictures. Thus Formula 4.5 is the picture enumerator the set of all multisets of fruit with between one and three apples, one and two pears, and one and two bananas.

164. How would you write down a polynomial in the variable A that says you should take between zero and three apples?

165. How would you write down a picture enumerator that says we take between zero and three apples, between zero and three pears, and between zero and three bananas?

166. Notice that if we use A^2 to stand for taking two apples, and P^3 to stand for taking three pears, then we have used the product A^2P^3 to stand for taking two apples and three pears. Thus we have chosen the picture of the ordered pair (2 apples,3 pears) to be the product of the

pictures of a multiset of two apples and a multiset of three pears. Show that if S_1 and S_2 are sets with picture functions P_1 and P_2 defined on them, and if we define the picture of an ordered pair $(x_1, x_2) \in S_1 \times S_2$ to be $P((x_1, x_2)) = P_1(x_1)P_2(x_2)$, then the picture enumerator of P on the set $S_1 \times S_2$ is $E_{P_1}(S_1)E_{P_2}(S_2)$. We call this the **product principle for picture enumerators**.

4.2.2 Generating functions

167. Suppose you are going to choose a snack of between zero and three apples, between zero and three pears, and between zero and three bananas. Write down a polynomial in one variable x such that the coefficient of x^n is the number of ways to choose a snack with n pieces of fruit? Hint: substitute something for A, P and B in your formula from Problem 165.

168. Suppose an apple costs 20 cents, a banana costs 25 cents, and a pear costs 30 cents. What should you substitute for A, P, and B in Problem 165 in order to get a polynomial in which the coefficient of x^n is the number of ways to choose a selection of fruit that costs n cents?

169. Suppose an apple has 40 calories, a pear has 60 calories, and a banana has 80 calories. What should you substitute for A, P, and B in Problem 165 in order to get a polynomial in which the coefficient of x^n is the number of ways to choose a selection of fruit that has n calories?

170. We are going to choose a subset of the set $[n] = \{1, 2, \ldots, n\}$. Suppose we use x_1 to be the picture of choosing 1 to be in our subset. What is the picture enumerator for either choosing 1 or not choosing 1? Suppose that for each i between 1 and n, we use x_i to be the picture of choosing i to be in our subset. What is the picture enumerator for either choosing i or not choosing i to be in our subset? What is the picture enumerator for all possible choices of subsets of $[n]$? What should we substitute for x_i in order to get a polynomial in x such that the coefficient of x^k is the number of ways to choose a k-element subset of n? What theorem have we just reproved (a special case of)?

In Problem 170 we see that we can think of the process of expanding the polynomial $(1+x)^n$ as a way of "generating" the binomial coefficients $\binom{n}{k}$ as the coefficients of x^k in the expansion of $(1+x)^n$. For this reason, we say that $(1+x)^n$ is the "generating function" for the binomial coefficients $\binom{n}{k}$. More generally, the **generating function** for a sequence a_i, defined for i with $0 \le i \le n$ is the expression $\sum_{i=0}^{n} a_i x^i$, and the **generating function** for the sequence a_i with $i \ge 0$ is the expression $\sum_{i=0}^{\infty} a_i x^i$. This last expression is an example of a power series. In calculus it is important to think about whether a power series converges in order to determine whether or not it represents a function. In a nice twist of language, even though we use the phrase generating function as the name of a power series in combinatorics, we don't require the power series to actually represent a function in the usual sense, and so we don't have to worry about convergence.[3]

4.2.3 Power series

For now, most of our uses of power series will involve just simple algebra. Since we use power series in a different way in combinatorics than we do in calculus, we should review a bit of the algebra of power series.

171. In the polynomial $(a_0 + a_1x + a_2x^2)(b_0 + b_1x + b_2x^2 + b_3x^3)$, what is the coefficient of x^2? What is the coefficient of x^4?

172. In Problem 171 why is there a b_0 and a b_1 in your expression for the coefficient of x^2 but there is not a b_0 or a b_1 in your expression for the coefficient of x^4? What is the coefficient of x^4 in

$$(a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4)(b_0 + b_1x + b_2x^2 + b_3x^3 + b_4x^4)?$$

Express this coefficient in the form

$$\sum_{i=0}^{4} \text{something},$$

where the something is an expression you need to figure out. Now suppose that $a_3 = 0$, $a_4 = 0$ and $b_4 = 0$. To what is your expression equal after you substitute these values? In particular, what does this have to do with Problem 171?

[3] In the evolution of our current mathematical terminology, the word function evolved through several meanings, starting with very imprecise meanings and ending with our current rather precise meaning. The terminology "generating function" may be though of as an example of one of the earlier usages of the term function.

173. The point of the Problems 171 and 172 is that so long as we are willing to assume $a_i = 0$ for $i > n$ and $b_j = 0$ for $j > m$, then there is a very nice formula for the coefficient of x^k in the product

$$(\sum_{i=0}^{n} a_i x^i)(\sum_{j=0}^{m} b_j x^j).$$

Write down this formula explicitly.

174. Assuming that the rules you use to do arithmetic with polynomials apply to power series, write down a formula for the coefficient of x^k in the product

$$(\sum_{i=0}^{\infty} a_i x^i)(\sum_{j=0}^{\infty} b_j x^j).$$

We use the expression you obtained in Problem 174 to *define* the product of power series. That is, we define the product

$$(\sum_{i=0}^{\infty} a_i x^i)(\sum_{j=0}^{\infty} b_j x^j)$$

to be the power series $\sum_{k=0}^{\infty} c_k x^k$, where c_k is the expression you found in Problem 174. Since you derived this expression by using the usual rules of algebra for polynomials, it should not be surprising that the product of power series satisfies these rules.[4]

4.2.4 Product principle for generating functions

Each time that we converted a picture function to a generating function by substituting x or some power of x for each picture, the coefficient of x had a meaning that was significant to us. For example, with the picture enumerator for selecting between zero and three each of apples, pears, and bananas, when we substituted x for each of our pictures, the exponent i in the power x^i is the number of pieces of fruit in the fruit selection that led us to x^i. After we simplify our product by collecting together all like powers of x, the coefficient of x^i is the number of fruit selections that use i pieces of fruit. In the same way, if we substitute x^c for a picture, where c is the number of calories

[4] Technically we should explicitly state these rules and prove that they are all valid for power series multiplication, but it seems like overkill at this point to do so!

in that particular kind of fruit, then the i in an x^i term in our generating function stands for the number of calories in a fruit selection that gave rise to x^i, and the coefficient of x^i in our generating function is the number of fruit selections with i calories. The product principle of picture enumerators translates directly into a product principle for generating functions.

175. Suppose that we have two sets S_1 and S_2. Let v_1 (v stands for value) be a function from S_1 to the nonnegative integers and let v_2 be a function from S_2 to the nonnegative integers. Define a new function v on the set $S_1 \times S_2$ by $v(x_1, x_2) = v_1(x_1) + v_2(x_2)$. Suppose further that $\sum_{i=0}^{\infty} a_i x^i$ is the generating function for the number of elements x_1 of S_1 of value i, that is with $v_1(x_1) = i$. Suppose also that $\sum_{j=0}^{\infty} b_j x^j$ is the generating function for the number of elements of x_2 of S_2 of value j, that is with $v_2(x_2) = j$. Prove that the coefficient of x^k in

$$(\sum_{i=0}^{\infty} a_i x^i)(\sum_{j=0}^{\infty} b_j x^j)$$

is the number of ordered pairs (x_1, x_2) in $S_1 \times S_2$ with total value k, that is with $v_1(x_1) + v_2(x_2) = k$. This is called the **product principle for generating functions**.

176. Let i denote an integer between 1 and n.

    (a) What is the generating function for the number of subsets of $\{i\}$ of each possible size? (Notice that the only subsets of $\{i\}$ are $\emptyset$ and $\{i\}$.)

    (b) Use the product principle for generating functions to prove the binomial theorem.

4.2.5 The extended binomial theorem and multisets

177. Suppose once again that i is an integer between 1 and n.

    (a) What is the generating function in which the coefficient of x^k is the number of multisets of size k chosen from $\{i\}$? This series is an example of what is called an *infinite geometric series*.

(b) Express generating function in which the coefficient of x^k is the number of multisets chosen from $[n]$ as a power of a power series. What does Problem 108 (in which your answer could be expressed as a binomial coefficient) tell you about what this generating function equals?

178. What is the product $(1-x)\sum_{k=0}^{n} x^k$? What is the product

$$(1-x)\sum_{k=0}^{\infty} x^k?$$

179. Express the generating function for the number of multisets of size k chosen from $[n]$ (where n is fixed but k can be any nonnegative integer) as a 1 over something relatively simple.

180. Find a formula for $(1+x)^{-n}$ as a power series whose coefficients involve binomial coefficients. What does this formula tell you about how we should define $\binom{-n}{k}$ when n is positive?

181. If you define $\binom{-n}{k}$ in the way you described in Problem 180, you can write down a version of the binomial theorem for $(x+y)^n$ that is valid for both nonnegative and negative values of n. Do so. This is called the *extended binomial theorem*.

182. Write down the generating function for the number of ways to distribute identical pieces of candy to three children so that no child gets more than 4 pieces. Write this generating function as a quotient of polynomials. Using both the extended binomial theorem and the original binomial theorem, find out in how many ways we can pass out exactly ten pieces. Use one of our earlier counting techniques to verify your answer.

183. What is the generating function for the number of multisets chosen from an n-element set so that each element appears at least j times and less than m times. Write this generating function as a quotient of polynomials, then as a product of a polynomial and a power series.

4.2.6 Generating functions for integer partitions

184. If we have five identical pennies, five identical nickels, five identical dimes, and five identical quarters, give the picture enumerator for the combinations of coins we can form and convert it to a generating function for the number of ways to make k cents with the coins we have. Do the same thing assuming we have an unlimited supply of pennies, nickels, dimes, and quarters.

185. Recall that a partition of an integer n is a multiset of numbers that adds to n. In Problem 184 we found the generating function for the number of partitions of an integer into parts of size 1, 5, 10, and 25. Give the generating function for the number partitions of an integer into parts of size one through ten. Give the generating function for the number of partitions of an integer into parts of any size. This last generating function involves an infinite product. Describe the kinds of terms you actually multiply and add together to get the last generating function. Rewrite any power series that appear in your product as quotients of polynomials or as integers divided by polynomials.

186. What is the generating function for the number of partitions of an integer in which each part is even?

187. What is the generating function for the number of partitions of an integer into distinct parts, that is, in which each part is used at most once?

188. Use generating functions to explain why the number of partitions of an integer in which each part is used an even number of times equals the generating function for the number of partitions of an integer in which each part is even.

189. Use the fact that
$$\frac{1-x^{2i}}{1-x^i} = 1 + x^i$$
and the generating function for the number of partitions of an integer into distinct parts to show how the number of partitions of an integer n into distinct parts is related to the number of partitions of an integer n into odd parts.

190. Write down the generating function for the number of ways to partition an integer into parts of size no more than m, each used an odd number of times. Write down the generating function for the number of partitions of an integer into parts of size no more than m, each used an even number of times. Use these two generating functions to get a relationship between the two sequences for which you wrote down the generating functions.

4.3 Generating Functions and Recurrence Relations

Recall that a recurrence relation for a sequence a_n expresses a_n in terms of values a_i for $i < n$. For example, the equation $a_i = 3a_{i-1} + 2^i$ is a first order linear constant coefficient recurrence.

4.3.1 How generating functions are relevant

Algebraic manipulations with generating functions can sometimes reveal the solutions to a recurrence relation.

191. Suppose that $a_i = 3a_{i-1} + 3^i$.

    (a) Multiply both sides by x^i and sum both the left hand side and right hand side from $i = 1$ to infinity. In the left-hand side use the fact that
    $$\sum_{i=1}^{\infty} a_i x^i = (\sum_{i=0}^{\infty} x^i) - a_0$$
    and in the right hand side, use the fact that
    $$\sum_{i=1}^{\infty} a_{i-1} x^i = x \sum_{i=0}^{\infty} a_i x^i$$
    to rewrite the equation in terms of the power series $\sum_{i=0}^{\infty} a_i x^i$. Solve the resulting equation for the power series $\sum_{i=0}^{\infty} a_i x^i$.

    (b) Use the previous part to get a formula for a_i in terms of a_0.

(c) Now suppose that $a_i = 3a_{i-1} + 2^i$. Repeat the previous two steps for this recurrence relation. (There is a way to do this part using what you already know. Later on we shall introduce yet another way to deal with the kind of generating function that arises here.)

192. Suppose we deposit \$5000 in a savings certificate that pays ten percent interest and also participate in a program to add \$1000 to the certificate at the end of each year (from the end of the first year on) that follows (also subject to interest.) Assuming we make the \$5000 deposit at the end of year 0, and letting a_i be the amount of money in the account at the end of year i, write a recurrence for the amount of money the certificate is worth at the end of year n. Solve this recurrence. How much money do we have in the account (after our year-end deposit) at the end of ten years? At the end of 20 years?

4.3.2 Fibonacci Numbers

The sequence of problems that follows describes a number of hypotheses we might make about a fictional population of rabbits. We use the example of a rabbit population for historic reasons; our goal is a classical sequence of numbers called Fibonacci numbers. When Fibonacci introduced them, he did so with a fictional population of rabbits.

4.3.3 Second order linear recurrence relations

193. Suppose we start with 10 pairs of baby rabbits, and that after baby rabbits mature for one month they begin to reproduce, with each pair producing two new pairs at the end of each month afterwards. Suppose further that over the time we observe the rabbits, none die. Show that $a_n = a_{n-1} + 2a_{n-2}$. This is an example of a *second order linear* recurrence with constant coefficients. Using the method of Problem 191, show that

$$\sum_{i=0}^{\infty} a_i x^i = \frac{10}{1 - x - 2x^2}.$$

This gives us the generating function for the sequence a_i giving the population in month i; shortly we shall see a method for converting this to a solution to the recurrence.

194. In Fibonacci's original problem, each pair of mature rabbits produces one new pair at the end of each month, but otherwise the situation is the same as in Problem 193. Assuming that we start with one pair of baby rabbits (at the end of month 0), find the generating function for the number of pairs of rabbits we have at the end on n months.

195. Find the generating function for the solutions to the recurrence

$$a_i = 5a_{i-1} - 6a_{i-2} + 2^i.$$

The recurrence relations we have seen in this section are called *second order* because they specify a_i in terms of a_{i-1} and a_{i-2}, they are called *linear* because a_{i-1} and a_{i-2} each appear to the first power, and they are called *constant coefficient recurrences* because the coefficients in front of a_{i-1} and a_{i-2} are constants.

4.3.4 Partial fractions

The generating functions you found in the previous section all can be expressed in terms of the reciprocal of a quadratic polynomial. However without a power series representation, the generating function doesn't tell us what the sequence is. It turns out that whenever you can factor a polynomial into linear factors (and over the complex numbers such a factorization always exists) you can use that factorization to express the reciprocal in terms of power series.

196. Express $\frac{1}{x-3} + \frac{2}{x-2}$ as a single fraction.

197. In Problem 196 you see that when we added numerical multiples of the reciprocals of first degree polynomials we got a fraction in which the denominator is a quadratic polynomial. This will always happen unless the two denominators are multiples of each other, because their least common multiple will simply be their product, a quadratic polynomial. This leads us to ask whether a fraction whose denominator is a quadratic polynomial can always be expressed as a sum of fractions whose denominators are first degree polynomials. Find numbers c and d so that

$$\frac{5x+1}{(x-3)(x+5)} = \frac{c}{x-3} + \frac{d}{x-5}.$$

198. In Problem 197 you may have simply guessed at values of c and d, or you may have solved a system of equations in the two unknowns c and c. Given constants a, b, r_1, and r_2 (with $r_1 \neq r_2$, write down a system of equations we can solve for c and d to write

$$\frac{ax+b}{(x-r_1)(x-r_2)} = \frac{c}{x-r_1} + \frac{d}{x-r_2}.$$

Writing down the equations in Problem 198 and solving them is called the *method of partial fractions*. This method will let you find power series expansions for generating functions of the type you found in Problems 193 to 195. However you have to be able to factor the quadratic polynomials that are in the denominators of your generating functions.

199. Use the method of partial fractions to convert the generating function of Problem 193 into the form

$$\frac{c}{x-r_1} + \frac{d}{x-r_2}.$$

Use this to find a formula for a_n.

200. Use the quadratic formula to find the solutions to $x^2 - x - 1 = 0$, and use that information to factor $x^2 - x - 1$.

201. Use the factors you found in Problem 200 to write

$$\frac{1}{x^2 - x - 1}$$

in the form

$$\frac{c}{x-r_1} + \frac{d}{x-r_2}.$$

Hint: You can save yourself a tremendous amount of frustrating algebra if you arbitrarily choose one of the solutions and call it r_1 and call the other solution r_2 and solve the problem using these algebraic symbols in place of the actual roots.[5] Not only will you save yourself some work, but you will get a formula you could use in other problems. When you are done, substitute in the actual values of the solutions and simplify.

[5] We use the words roots and solutions interchangeably.

202. Use the partial fractions decomposition you found in Problem 200 to write the generating function you found in Problem 194 in the form

$$\sum_{i=0}^{\infty} a_i x^i$$

and use this to give an explicit formula for a_n. (Hint: once again it will save a lot of tedious algebra if you use the symbols r_1 and r_2 for the solutions as in Problem 201 and substitute the actual values of the solutions once you have a formula for a_n in terms of r_1 and r_2.) When we have $a_0 = 1$ and $a_1 = 1$, i.e. when we start with one pair of baby rabbits, the numbers a_n are called *Fibonacci Numbers*. Use either the recurrence or your final formula to find a_2 through a_8. Are you amazed that your general formula produces integers, or for that matter produces rational numbers? Why does the recurrence equation tell you that the Fibonacci numbers are all integers? Try to find an algebraic explanation (not using the recurrence equation) of why the formula has to do so. Explain why there is a real number b such that the value of the nth Fibonacci number is almost exactly (but not quite) some constant times b^n. (Find b and the constant.)

203. Solve the recurrence $a_n = 4a_{n-1} - 4a_{n-2}$.

4.3.5 Catalan Numbers

204. Using either lattice paths or diagonal lattice paths, explain why the Catalan Number c_n satisfies the recurrence

$$c_n = \sum_{i=1}^{n-1} c_{i-1}c_{n-i}.$$

Show that if we use y to stand for the power series $\sum_{i=0}^{\infty} c_n x^n$, then we can find y by solving a quadratic equation. Solve for y. Taylor's theorem from calculus tells us that the extended binomial theorem

$$(1+x)^r = \sum_{i=0}^{\infty} \binom{r}{i} x^i$$

holds for any number real number r, where $\binom{r}{i}$ is defined to be

$$\frac{r^{\underline{i}}}{i!} = \frac{r(r-1)\cdots(r-i+1)}{i!}.$$

Use this and your solution for y (note that of the two possible values for y that you get from the quadratic formula, only one gives an actual power series) to get a formula for the Catalan numbers.

4.4 Supplementary Problems

1. Each person attending a party has been asked to bring a prize. The person planning the party has arranged to give out exactly as many prizes as there are guests, but any person may win any number of prizes. If there are n guests, in how many ways may the prizes be given out so that nobody gets the prize that he or she brought?

2. There are m students attending a seminar in a room with n seats. The seminar is a long one, and in the middle the group takes a break. In how many ways may the students return to the room and sit down so that nobody is in the same seat as before?

3. In how many ways may k distinct books be arranged on n shelves so that no shelf gets more than m books?

4. A group of n married couples comes to a group discussion session where they all sit around a round table. In how many ways can they sit so that no person is next to his or her spouse?

5. A group of n married couples comes to a group discussion session where they all sit around a round table. In how many ways can they sit so that no person is next to his or her spouse or a person of the same sex? This problem is called the *menage problem*. (Hint: Reason as you did in Problem 4, noting that if the set of couples who do sit side-by-side is nonempty, then the sex of the person at each place at the table is determined once we seat one couple in that set.)

6. What is the generating function for the number of ways to pass out k pieces of candy from an unlimited supply of identical candy to n

children (where n is fixed) so that each child gets between three and six pieces of candy (inclusive)? Use the fact that $(1 + x + x + x^3)(1 - x) = 1 - x^4$ to find a formula for the number of ways to pass out the candy. Reformulate this problem as an inclusion-exclusion problem and describe what you would need to do to solve it.

7. Find a recurrence relation for the number of ways to divide a convex n-gon into triangles by means of non-intersecting diagonals. How do these numbers relate to the Catalan numbers?

8. How does $\sum_{k=0}^{n} \binom{n-k}{k}$ relate to the Fibonacci Numbers?

9. Let m and n be fixed. Express the generating function for the number of k-element multisets of an n-element set such that no element appears more than m times as a quotient of two polynomials. Use this expression to get a formula for the number of k-element multisets of an n-element set such that no element appears more than m times.

10. One natural but oversimplified model for the growth of a tree is that all new wood grows from the previous year's growth and is proportional to it in amount. To be more precise, assume that the (total) length of the new growth in a given year is the constant c times the (total) length of new growth in the previous year. Write down a recurrence for the total length a_n of all the branches of the tree at the end of growing season n. Find the general solution to your recurrence relation. Assume that we begin with a one meter cutting of new wood which branches out and grows a total of two meters of new wood in the first year. What will the total length of all the branches of the tree be at the end of n years?

Appendix A

Relations

A.1 Relations as sets of Ordered Pairs

A.1.1 The relation of a function

1. Consider the functions from $S = \{-2, -1, 0, 1, 2\}$ to $T = \{1, 2, 3, 4, 5\}$ defined by $f(x) = x + 3$, and $g(x) = x^5 - 5x^3 + 5x + 3$. Write down the set of ordered pairs $(x, f(x))$ for $x \in S$ and the set of ordered pairs $(x, g(x))$ for $x \in S$. Are the two functions the same or different?

Problem 1 points out how two functions which appear to be different are actually the same on some domain of interest to us. Most of the time when we are thinking about functions it is fine to think of a function casually as a relationship between two sets. In Problem 1 the set of ordered pairs you wrote down for each function is called the *relation* of the function. When we want to distinguish between the casual and the careful in talking about relationships, our casual term will be "relationship" and our careful term will be "relation." So *relation* is a technical word in mathematics, and as such it has a technical definition. A *relation* from a set S to a set T is a set of ordered pairs whose first elements are in S and whose second elements are in T. Another way to say this is that a *relation* from S to T is a subset of $S \times T$.

2. Here are some questions that will help you get used to the formal idea of a relation and the related formal idea of a function. S will stand for a set of size s and T will stand for a set of size t.

(a) What is the size of the largest relation from S to T?

(b) What is the size of the smallest relation from S to T?

(c) The relation of a function $f : S \to T$ is the set of all ordered pairs $(x, f(x))$ with $x \in S$. What is the size of the relation of a function from S to T? That is, how many ordered pairs are in the relation of a function from S to T?

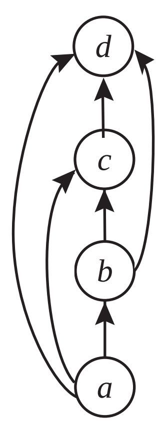
(d) How many different elements must appear as second elements of the ordered pairs in the relation of a one-to-one function (injection) from S to T? (See Problem 7 for a definition of a one-to-one function.)

(e) A function $f : S \to T$ is called an *onto function* or *surjection* if each element of T is $f(x)$ for some $x \in S$ What is the minimum size that S can have if there is a surjection from S to T?

A.1.2 Directed graphs

We visualize numerical functions like $f(x) = x^2$ with their graphs in Cartesian coordinate systems. We will call these kinds of graphs *coordinate graphs* to distinguish them from other kinds of graphs used to visualize relations that are non-numerical. In Figure A.1 we illustrate another kind of graph,

Figure A.1: The alphabet digraph.



a "directed graph" or "digraph" of the "comes before in alphabetical order" relation on the letters a, b, c, and d. To draw a *directed graph* of a relation on a set S, we draw a circle (or dot, if we prefer), which we call a *vertex*, for each element of the set, we usually label the vertex with the set element it corresponds to, and we draw an arrow from the vertex for a to that for b if a is related to b, that is, if the ordered pair (a, b) is in our relation. We call such an arrow an *edge* or a *directed edge*. We draw the arrow from a to b, for example, because a comes before b in alphabetical order. We try to choose the locations where we draw our vertices so that the arrows capture what we are trying to illustrate as well as possible. Sometimes this entails redrawing our directed graph several times until we think the arrows capture the relationship well.

We also draw digraphs for relations from a set S to a set T; we simply draw vertices for the elements of S (usually in a row) and vertices for the elements of T (usually in a parallel row) draw an arrow from x in S to y in T if x is related to Y. Notice that instead of referring to the vertex representing x, we simply referred to x. This is a common shorthand. Here are some exercises just to practice drawing digraphs.

3. Draw the digraph of the "is a proper subset of" relation on the set of subsets of a two element set. How many arrows would you have had to draw if this problem asked you to draw the digraph for the subsets of a three element set?

4. Draw the digraph of the relation from the set {A, M, P, S} to the set {Sam, Mary, Pat, Ann, Polly, Sarah} given by "is the first letter of."

A.1.3 Equivalence relations

So far we've used relations primarily to talk about functions. There is another kind of relation, called an equivalence relation, that comes up in the counting problems with which we began. In Problem 8 with three distinct flavors, it was probably tempting to say there are 12 flavors for the first pint, 11 for the second, and 10 for the third, so there are $12 \cdot 11 \cdot 10$ ways to choose the pints of ice cream. However, once the pints have been chosen, bought, and put into a bag, there is no way to tell which is first, which is second and which is third. What we just counted instead is lists of three distinct flavors—one to one functions from the set $\{1, 2, 3\}$ in to the set of ice cream flavors. Two of those lists become equivalent once the ice cream once the purchase is made if they list the same ice cream. In other words, two of those lists become equivalent (are related) if they list same subset of the set of ice cream flavors. To visualize this relation with a digraph, we would need one vertex for each of the $12 \cdot 11 \cdot 10$ lists. Even with five flavors of ice cream, we would need one vertex for each of $5 \cdot 4 \cdot 3 = 60$ lists. So for now we will work with the easier to draw question of choosing three pints of ice cream of different flavors from four flavors of ice cream.

5. Suppose we have four flavors of ice cream, V(anilla), C(hocolate), S(trawberry) and P(each). Draw the directed graph whose vertices consist of all lists of three distinct flavors of the ice cream, and whose edges connect two lists if they list the same three flavors. This graph makes it pretty clear in how many ways we may choose 3 flavors out of four. How many is it?

6. Now suppose again we are choosing three flavors of ice cream out of four, but instead of putting scoops in a cone or choosing pints, we are going to have the three scoops arranged symmetrically in a circular dish. Similarly to choosing three pints, we can describe a selection of ice cream in terms of which one goes in the dish first, which one goes in second (say to the right of the first), and which one goes in third (say to the right of the second scoop, which makes it to the left of the first scoop). But again, two of these lists will sometimes be equivalent. Once they are in the dish, we can't tell which one went in first. However, there is a subtle difference between putting each flavor in its own small dish and putting all three flavors in a circle in a larger dish. Think about what makes the lists of flavors equivalent, and draw the directed graph whose vertices consist of all lists of three of the flavors of ice cream and whose edges connect two lists that we cannot tell the difference between as dishes of ice cream. How many dishes of ice cream can we distinguish from one another?

7. Draw the digraph for Problem 32 in the special case where we have four people sitting around the table.

In Problems 5, 6, and 7 (as well as Problems 29, 32, and 33) we can begin with a set of lists, and say when two lists are equivalent as representations of the objects we are trying to count. In particular, in Problems 5, 6, and 7 you drew the directed graph for this relation of equivalence. Technically, you should have had an arrow from each vertex (list) to itself. This is what we mean when we say a relation is *reflexive*. Whenever you had an arrow from one vertex to a second, you had an arrow back to the first. This is what we mean when we say a relation is *symmetric*.

When people sit around a round table, each list is equivalent to itself: certainly if List1 and List2 are identical, then each person is in the same relative position to everyone else in the seating arrangement corresponding to List1 and the seating arrangement corresponding to List 2. To see the symmetric property of the equivalence of seating arrangements, if List1 and List2 are different, but everyone is in the same relative position when they sit according to List2 as when they sit according to List1, then everybody better be in the same relative position when they sit according to List1 as when they sit according to List2. What we are really saying here is that whatever we mean by "the same relative position," if A is the same as B, then B better be the same as A. Thus we are merely making an argument that the word same is only used in a symmetric situation. However, we can be more precise.

If everyone is in the same relative position in arrangements that come from two different lists, then everyone has the same person to the right in both arrangements. Everyone has the same person two people to the right in both arrangements. The same holds for three people to the right and so on. (Note that if we have n people, then the person $n-1$ places to my right around a round table is actually directly to my left!)

In Problems 5, 6 and 7 there is another property of those relations you may have noticed from the directed graph. Whenever you had an arrow from L_1 to L_2 and an arrow from L_2 to L_3, then there was an arrow from L_1 to L_3. This is what we mean when we say a relation is *transitive*. You also undoubtedly noticed how the directed graph divides up into clumps of mutually connected vertices. This is what equivalence relations are all about. Let's be a bit more precise in our description of what it means for a relation to be reflexive, symmetric or transitive.

- If R is a relation on a set X, we say R is *reflexive* if $(x, x) \in R$ for every $x \in X$.

- If R is a relation on a set X, we say R is *symmetric* if (x, y) is in R whenever (y, x) is in R.

- If R is a relation on a set X, we say R is *transitive* if whenever (x, y) is in R and (y, z) is in R, then (x, z) is in R as well.

Each of the relations of equivalence you worked with in the Problem 5, 6 and 7 had these three properties. Can you visualize the same three properties in the relations of equivalence that you would use in problems 29, 32, and 33? We call a relation an **equivalence relation** if it is reflexive, symmetric and transitive.

After some more examples, we will see how to show that equivalence relations have the kind of clumping property you saw in the directed graphs. In our first example, using the notation $(a, b) \in R$ to say that a is related to B is going to get in the way. It is really more common to write aRb to mean that a is related to b. For example, if our relation is the less than relation on $\{1, 2, 3\}$, you are much more likely to use $x < y$ than you are $(x, y) \in <$, aren't you? The reflexive law then says xRx for every x in X, the symmetric law says that if xRy, then yRx, and the transitive law says that if xRy and yRz, then xRz.

8. For the necklace problem, Problem 37, our lists are lists of beads. What makes two lists equivalent for the purpose of describing a necklace? Verify explicitly that this relationship of equivalence is reflexive, symmetric, and transitive.

9. Which of the reflexive, symmetric and transitive properties does the $<$ relation on the integers have?

10. A relation R on the set of ordered pairs of positive integers that you learned about in grade school in another notation is the relation that says (m, n) is related to (h, k) if $mk = hn$. Show that this relation is an equivalence relation. In what context did you learn about this relation in grade school?

11. Another relation that you learned about in school, perhaps first in the guise of "clock arithmetic," is the relation of equivalence modulo n. For integers (positive, negative, or zero) a and b, we write $a \equiv b \pmod{n}$ to mean that $a - b$ is an integer multiple of n, and in this case, we say that a is *congruent to b modulo n*. Show that the relation of congruence modulo n is an equivalence relation.

12. Define a relation on the set of all lists of n distinct integers chosen from $\{1, 2, \ldots, n\}$, by saying two lists are related if they have the same elements (though perhaps in a different order) in the first k places, and the same elements (though perhaps in a different order) in the last $n-k$ places. Show this relation is an equivalence relation.

13. Suppose that R is an equivalence relation on a set X and for each $x \in X$, let $C_x = \{y|y \in X \text{ and } yRx\}$. If C_x and C_z have an element y in common, what can you conclude about C_x and C_z (besides the fact that they have an element in common!)? Be explicit about what property(ies) of equivalence relations justify your answer. Why is every element of X in some set C_x? Be explicit about what property(ies) of equivalence relations you are using to answer this question. Notice that we might simultaneously denote a set by C_x and C_y. Explain why the union of the sets C_x is X. Explain why two distinct sets C_x and C_z are disjoint. What do these sets have to do with the "clumping" you saw in the digraph of Problem 5 and 6?

In Problem 13 the sets C_x are called *equivalence classes* of the equivalence relation R. You have just proved that if R is an equivalence relation of the set X, then each element of X is in exactly one equivalence class of R. Recall that a *partition* of a set X is a set of disjoint sets whose union is X. For example, $\{1, 3\}$, $\{2, 4, 6\}$, $\{5\}$ is a partition of the set $\{1, 2, 3, 4, 5, 6\}$. Thus another way to describe what you proved in Problem 13 is that if R is an equivalence relation on X, then the set of equivalence classes of R is a partition of X. Since a partition of S is a set of subsets of S, it is common to call the subsets into which we partition S the *blocks* of the partition so that we don't find ourselves in the uncomfortable position or referring to a set and not being sure whether it is the set being partitioned or one of the blocks of the partition.

14. In each of Problems 32, 33, 37, 5, and 6, what does an equivalence class correspond to? (Five answers are expected here.)

15. Given the partition $\{1, 3\}$, $\{2, 4, 6\}$, $\{5\}$ of the set $\{1, 2, 3, 4, 5, 6\}$, define two elements of $\{1, 2, 3, 4, 5, 6\}$ to be related if they are in the same part of the partition. That is, define 1 to be related to 3 (and 1 and 3 each related to itself), define 2 and 4, 2 and 6, and 4 and 6 to be related (and each of 2, 4, and 6 to be related to itself), and define 5 to be related to itself. Show that this relation is an equivalence relation.

16. Suppose $P = \{S_1, S_2, S_3, \ldots, S_k\}$ is a partition of S. Define two elements of S to be related if they are in the same set S_i, and otherwise not to be related. Show that this relation is an equivalence relation. Show that the equivalence classes of the equivalence relation are the sets S_i.

In Problem 16 you just proved that each partition of a set gives rise to an equivalence relation whose classes are just the parts of the partition. Thus in Problem 13 and Problem 16 you proved the following Theorem.

Theorem 6 *A relation R is an equivalence relation on a set S if and only if S may be partitioned into sets S_1, S_2, ..., S_n in such a way that x and y are related by R if and only if they are in the same block S_i of the partition.*

In Problems 5, 6, 32 and 37 what we were doing in each case was counting equivalence classes of an equivalence relation. There was a special structure to the problems that made this somewhat easier to do. For example, in 5, we had $4 \cdot 3 \cdot 2 = 24$ lists of three distinct flavors chosen from V, C, S, and P. Each list was equivalent to $3 \cdot 2 \cdot 1 = 3! = 6$ lists, including itself, from the point of view of serving 3 small dishes of ice cream. The order in which we selected the three flavors was unimportant. Thus the set of all $4 \cdot 3 \cdot 2$ lists was a union of some number n of equivalence classes, each of size 6. By the product principle, if we have a union of n disjoint sets, each of size 6, the union has $6n$ elements. But we already knew that the union was the set of all 24 lists of three distinct letters chosen from our four letters. Thus we have $6n = 24$, or $n = 4$ equivalence classes.

In Problem 6 there is a subtle change. In the language we adopted for seating people around a round table, if we choose the flavors V, C, and S, and arrange them in the dish with C to the right of V and S to the right of C, then the scoops are in different relative positions than if we arrange them instead with S to the right of V and C to the right of S. Thus the order in which the scoops go into the dish is somewhat important—somewhat, because putting in V first, then C to its right and S to its right is the same as putting in S first, then V to its right and C to its right. In this case, each list of three flavors is equivalent to only three lists, including itself, and so if there are n equivalence classes, we have $3n = 24$, so there are $24/3 = 8$ equivalence classes.

17. If we have an equivalence relation that divides a set with k elements up into equivalence classes each of size m, what is the number n of equivalence classes? Explain why.

18. In Problem 12, what is the number of equivalence classes? Explain in words the relationship between this problem and the Problem 33.

19. Describe explicitly what makes two lists of beads equivalent in Problem 37 and how Problem 17 can be used to compute the number of different necklaces.

20. What are the equivalence classes (write them out as sets of lists) in Problem 38, and why can't we use Problem 17 to compute the number of equivalence classes?

In Problem 17 you proved our next theorem. In Chapter 1 (Problem 36) we discovered and stated this theorem in the context of partitions and called it the *Quotient Principle*

Theorem 7 *If an equivalence relation on a set S size k has n equivalence classes each of size m, then the number of equivalence classes is k/m.*

Appendix B

Mathematical Induction

B.1 The Principle of Mathematical Induction

B.1.1 The ideas behind mathematical induction

There is a variant of the bijection we used to prove the Pascal Equation that comes up in counting the subsets of a set. In the next problem it will help us compute the total number of subsets of a set, regardless of their size. Our main goal in this problem, however, is to introduce some ideas that will lead us to one of the most powerful proof techniques in combinatorics (and many other branches of mathematics), the principle of mathematical induction.

21. (a) Write down a list of the subsets of $\{1, 2\}$. Don't forget the empty set! Group the sets containing containing 2 separately from the others.

    (b) Write down a list of the subsets of $\{1, 2, 3\}$. Group the sets containing 3 separately from the others.

    (c) Look for a natural way to match up the subsets containing 2 in Part (a) with those not containing 2. Look for a way to match up the subsets containing 3 in Part (b) containing 3 with those not containing 3.

    (d) On the basis of the previous part, you should be able to find a bijection between the collection of subsets of $\{1, 2, \ldots, n\}$ containing n and those not containing n. (If you are having difficulty figuring out the bijection, try repeating Parts (a) and (b).) Describe the bijection (unless you are very familiar with the notation of sets, it is probably easier to describe to describe the function in words rather than symbols) and explain why it is a bijection. Explain why the number of subsets of $\{1, 2, \ldots, n\}$ containing n equals the number of subsets of $\{1, 2, \ldots, n-1\}$.

(e) Parts (a) and (b) suggest strongly that the number of subsets of a n-element set is 2^n . In particular, the empty set has 2^0 subsets, a one-element set has 2^1 subsets, itself and the empty set, and in Parts a and b we saw that two-element and three-element sets have 2^2 and 2^3 subsets respectively. So there are certainly some values of n for which an n-element set has 2^n subsets. One way to prove that an n-element set has 2^n subsets for all values of n is to argue by contradiction. For this purpose, suppose there is a nonnegative integer n such that an n-element set doesn't have exactly 2^n subsets. In that case there may be more than one such n. Choose k to be the smallest such n. Notice that $k-1$ is still a positive integer, because k can't be 0, 1, 2, or 3. Since k was the smallest value of n we could choose to make the statement "An n-element set has $2n$ subsets" false, what do you know about the number of subsets of a $(k-1)$-element set? What do you know about the number of subsets of the k-element set $\{1, 2, \ldots, k\}$ that don't contain k? What do you know about the number of subsets of $\{1, 2, \ldots, k\}$ that do contain k? What does the sum principle tell you about the number of subsets of $\{1, 2, \ldots, k\}$? Notice that this contradicts the way in which we chose k, and the only assumption that went into our choice of k was that "there is a nonnegative integer n such that an n-element set doesn't have exactly 2^n subsets." Since this assumption has led us to a contradiction, it must be false. What can you now conclude about the statement "for every nonnegative integer n, an n-element set has exactly 2^n subsets?"

22. The expression

$$1 + 3 + 5 + \cdots + 2n - 1$$

is the sum of the first n odd integers. Experiment a bit with the sum for the first few positive integers and guess its value in terms of n. Now apply the technique of Problem 21 to prove that you are right.

In Problems 21 and 22 our proofs had several distinct elements. We had a statement involving an integer n. We knew the statement was true for the first few nonnegative integers in Problem 21 and for the first few positive integers in problem 22. We wanted to prove that the statement was true for all nonnegative integers in Problem 21 and for all positive integers in Problem 22. In both cases we used the method of proof by contradiction; for that purpose we assumed that there was a value of n for which our formula wasn't true. We then chose k to be the smallest value of n for which our formula wasn't true. This meant that when n was $k-1$, our formula was true, (or else that $k-1$ wasn't a nonnegative integer in Problem 21 or that $k-1$ wasn't a positive integer in Problem 22). What we did next was the crux of the proof. We showed that the truth of our statement for $n = k-1$ implied the truth of our statement for $n = k$. This gave us a contradiction to the assumption that there was an n that made the statement false. In fact, we will see that we can bypass entirely the use of proof by contradiction. We used it to help you discover the central ideas of the technique of proof by mathematical induction.

The central core of mathematical induction is the proof that the truth of a statement about the integer n for $n = k-1$ implies the truth of the statement for $n = k$. For example, once we know that a set of size 0 has 2^0 subsets, if we have proved our implication, we can then conclude that a set of size 1 has 2^1 subsets, from which we can conclude that a set of size 2 has 2^2 subsets, from which we can conclude that a set of size 3 has 2^3 subsets, and so on up to a set of size n having 2^n subsets for any nonnegative integer n we choose. In other words, although it was the idea of proof by contradiction that led us to think about such an implication, we can now do without the contradiction at all. What we need to prove a statement about n by this method is a place to start, that is a value b of n for which we know the statement to be true, and then a proof that the truth of our statement for $n = k-1$ implies the truth of the statement for $n = k$ whenever $k > b$.

B.1.2 Mathematical induction

The **principle of mathematical induction** states that

> In order to prove a statement about an integer n, if we can
>
> 1. Prove the statement when $n = b$, for some fixed integer b

2. Show that the truth of the statement for $n = k - 1$ implies the truth of the statement for $n = k$ whenever $k > b$,

then we can conclude the statement is true for all integers $n \geq b$.

As an example, let us return to Problem 21. The statement we wish to prove is the statement that "A set of size n has 2^n subsets."

> Our statement is true when $n = 0$, because a set of size 0 is the empty set and the empty set has $1 = 2^0$ subsets. (This step of our proof is called a *base step*.)
>
> Now suppose that $k > 0$ and every set with $k-1$ elements has 2^{k-1} subsets. Suppose $S = \{a_1, a_2, \ldots a_k\}$ is a set with k elements. We partition the subsets of S into two blocks. Block B_1 consists of the subsets that do not contain a_n and block B_2 consists of the subsets that do contain a_n. Each set in B_1 is a subset of $\{a_1, a_2, \ldots a_{k-1}\}$, and each subset of $\{a_1, a_2, \ldots a_{k-1}\}$ is in B_1. Thus B_1 is the set of all subsets of $\{a_1, a_2, \ldots a_{k-1}\}$. Therefore by our assumption in the first sentence of this paragraph, the size of B_1 is 2^{k-1}. Consider the function from B_2 to B_1 which takes a subset of S including a_n and removes a_n from it. This function is defined on B_2, because every set in B_2 contains a_n. This function is onto, because if T is a set in B_1, then $T \cup \{a_k\}$ is a set in B_2 which the function sends to T. This function is one-to-one because if V and W are two different sets in B_2, then removing a_k from them gives two different sets in B_1. Thus we have a bijection between B_1 and B_2, so B_1 and B_2 have the same size. Therefore by the sum principle the size of $B_1 \cup B_2$ is $2^{k-1} + 2^{k-1} = 2^k$. Therefore S has 2^k subsets. This shows that if a set of size $k - 1$ has 2^{k-1} subsets, then a set of size k has 2^k subsets. Therefore by the principle of mathematical induction, a set of size n has 2^n subsets for every nonnegative integer n.

The first sentence of the last paragraph is called the *inductive hypothesis*. In an inductive proof we always make an inductive hypothesis as part of proving that the truth of our statement when $n = k - 1$ implies the truth of our statement when $n = k$. The last paragraph itself is called the *inductive step* of our proof. In an inductive step we derive the statement for $n = k$ from the statement for $n = k - 1$, thus proving that the truth of our statement when $n = k - 1$ implies the truth of our statement when $n = k$. The last sentence in the last paragraph is called the *inductive conclusion*. All inductive proofs should have a base step, an inductive hypothesis, an inductive step, and an inductive conclusion.

There are a couple details worth noticing. First, in this problem, our base step was the case $n = 0$, or in other words, we had $b = 0$. However, in other proofs, b could be any integer, positive, negative, or 0. Second, our proof that the truth of our statement for $n = k - 1$ implies the truth of our statement for $n = k$ required that k be at least 1, so that there would be an element a_k we could take away in order to describe our bijection. However, condition (2) of the principle of mathematical induction only requires that we be able to prove the implication for $k > 0$, so we were allowed to assume $k > 0$.

23. Use mathematical induction to prove your formula from Problem 22.

B.1.3 Proving algebraic statements by induction

24. Use mathematical induction to prove the well-known formula that for all positive integers n,
$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

25. Experiment with various values of n in the sum
$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{n \cdot (n+1)} = \sum_{i=1}^{n} \frac{1}{i \cdot (i+1)}.$$
Guess a formula for this sum and prove your guess is correct by induction.

26. For large values of n, which is larger, n^2 or 2^n? Use mathematical induction to prove that you are correct.

27. What is wrong with the following attempt at an inductive proof that all integers in any consecutive set of n integers are equal for every positive integer n? For an arbitrary integer i, all integers from i to i are equal, so our statement is true when $n = 1$. Now suppose $k > 1$ and all

integers in any consecutive set of $k-1$ integers are equal. Let S be a set of k consecutive integers. By the inductive hypothesis, the first $k-1$ elements of S are equal and the last $k-1$ elements of S are equal. Therefore all the elements in the set S are equal. Thus by the principle of mathematical induction, for every positive n, every n consecutive integers are equal.

B.1.4 Strong Induction

One way of looking at the principle of mathematical induction is that it tells us that if we know the "first" case of a theorem and we can derive each other case of the theorem from a smaller case, then the theorem is true in all cases. However the particular way in which we stated the theorem is rather restrictive in that it requires us to derive each case form the immediately preceding case. This restriction is not necessary, and removing it leads us to a more general statement of the principal of mathematical induction which people often call the **strong principle of mathematical induction**. It states:

> In order to prove a statement about an integer n if we can
>
> 1. prove our statement when $n = b$ and
> 2. prove that the statements we get with $n = b$, $n = b+1$, $\ldots n = k-1$ imply the statement with $n = k$,
>
> then our statement is true for all integers $n \geq b$.

28. What postage do you think we can make with five and six cent stamps? Is there a number N such that if $n \geq N$, then we can make n cents worth of postage?

You probably see that we can make n cents worth of postage as long as n is at least 20. However you didn't try to make 26 cents in postage by working with 25 cents; rather you saw that you could get 20 cents and then add six cents to that to get 26 cents. Thus if we want to prove by induction that we are right that if $n \geq 20$, then we can make n cents worth of postage, we are going to have to use the strong version of the principle of mathematical induction.

We know that we can make 20 cents with four five-cent stamps. Now we let k be a number greater than 20, and assume that it is possible to make any amount between 20 and $k-1$ cents in postage with five and six cent stamps. Now if k is less than 25, it is 21, 22, 23, or 24. We can make 21 with three fives and one six. We can make 22 with two fives and two sixes, 23 with one five and three sixes, and 24 with four sixes. Otherwise $k-5$ is between 20 and $k-1$ (inclusive) and so by our inductive hypothesis, we know that $k-5$ cents can be made with five and six cent stamps, so with one more five cent stamp, so can k cents. Thus by the (strong) principle of mathematical induction, we can make n cents in stamps with five and six cent stamps for each $n \geq 20$.

Someone might want to argue that we really didn't prove the implication that the truth of our statement for 20 though $k-1$ implies its truth for k, because we didn't use the hypothesis that our statement was true for 20 through $k-1$ in the cases that $k = 21$, 22, 23, and 24. However to prove that one statement implies another, it suffices to show that the second statement is true, whether we use the first statement or not. Thus we really were using induction appropriately. Another point of view would be that we really had five base cases, and once we had proved those five consecutive base cases, then we could reduce any other case to one of these base cases by successively subtracting 5. That view is appropriate too.

29. A number greater than one is called prime if it has no factors other than itself and one. Show that each positive number is either a prime or a power of a prime or a product of powers of prime numbers.

30. Show that the number of prime factors of a positive number $n \geq 2$ is less than or equal to $\log_2 n$. (If a prime occurs to the kth power in a factorization of n, you can consider that power as k prime factors.) (There is a way to do this by induction and a way to do it without induction. It would be ideal to find both ways.)

Index