Cambridge University Press
978-1-108-47654-6 — An Invitation to Combinatorics
Shahriar Shahriari
Excerpt
More Information

# Introduction

*Accurate reckoning. The entrance into the knowledge of all existing things and all obscure secrets.*

*–The Ahmes–Rhind Papyrus*

## What is Combinatorics?

Combinatorics is a collection of techniques and a language for the study of (finite or countably infinite) discrete structures. Given a set of elements (and possibly some structure on that set), typical questions in combinatorics are:

- Does a specific arrangement of the elements exist?
- How many such arrangements are there?
- What properties do these arrangements have?
- Which one of the arrangements is maximal, minimal, or optimal according to some criterion?

Unlike many other areas of mathematics – e.g., analysis, algebra, topology – the core of combinatorics is neither its subject matter nor a set of "fundamental" theorems. More than anything else, combinatorics is a collection – some may say a hodgepodge – of techniques, attitudes, and general principles for solving problems about discrete structures. For any given problem, a combinatorist combines some of these techniques and principles – e.g., the pigeonhole principle, the inclusion–exclusion principle, the marriage theorem, various counting techniques, induction, recurrence relations, generating functions, probabilistic arguments, asymptotic analysis – with (often clever) ad hoc arguments. The result is a fun and difficult subject.

In today's mathematical world, in no small part due to the power of digital computers, most mathematicians find much use for the tool box of combinatorics. In problems of pure mathematics, often, after deciphering the layers of theory, you find a combinatorics problem at the core. Outside of mathematics, and as an example, combinatorial problems abound in computer science.

## Typical Problems

To whet your appetite, here is a preliminary sample of problems that we will encounter in the course of this text.

- How many sequences $a_1, a_2, \ldots, a_{12}$ are there consisting of four 0's and eight 1's, if no two consecutive terms are both 0's?
- A bakery has eight kinds of donuts, and a box holds one dozen donuts. How many different boxes can you buy? How many different boxes are there that contain at least one of each kind?
- A bakery sells seven kinds of donuts. How many ways are there to choose one dozen donuts if no more than three donuts of any kind are used?
- Determine the number of $n$-digit numbers with all digits odd, such that 1 and 3 each occur a *positive* even number of times.
- We are trying to reconstruct a word that is made from the letters $A$, $B$, $C$, $D$, and $R$. We are given a frequency table that shows the number of times a specific triple occurs in the word:

| triple | frequency |
|---|---|
| $ABR$ | 2 |
| $ACA$ | 1 |
| $ADA$ | 1 |
| $BRA$ | 2 |
| $CAD$ | 1 |
| $DAB$ | 1 |
| $RAC$ | 1 |

  For example, $ABR$ occurs twice while $ACA$ appears once. We want to know all words with the same triples and with the same frequency table. The answer may be that there are no such words. Note that by a word we mean an ordered collection of letters and we are not concerned with meaning.
- A particular signaling network consists of six pieces of communications equipment:

$$x_1, x_2, y_1, y_2, z_1, z_2$$

  We can choose various pairs of these and link each pair through an intermediate facility (e.g., microwave towers, trunk groups). Intermediate facilities are expensive to build but if one fails, then all the links through them become inoperative. So, if we build just one intermediate facility and route all of our connections through it, then its failure will disconnect everything. However, the design specification requires that if one intermediate facility fails, then there will remain at least one link between at least one of the $x$'s and one of the $y$'s and between one of the $x$'s and one of the $z$'s, as well as between one of the $y$'s and one of the $z$'s. What is the minimum number of facilities that we need and how should the connections be designed?
- A soccer ball is usually tiled with 12 pentagons and 20 hexagons. Are any other combinations of pentagons and hexagons possible?

## How Do We "Count"?

Counting the number of configurations of a certain type is an important part of combinatorics. In all of the examples in the previous section, it is clear what kind of an answer we are looking for. We want a specific numerical answer or an example of a specific configuration.

However, in many problems, it may be possible to present a solution that is satisfactory in many ways but is not quite a direct answer. We look at several examples.

(a) Let $[n] = \{1, 2, \ldots, n\}$, and let $f(n)$ be the number of subsets of $[n]$. Then $f(n) = 2^n$.

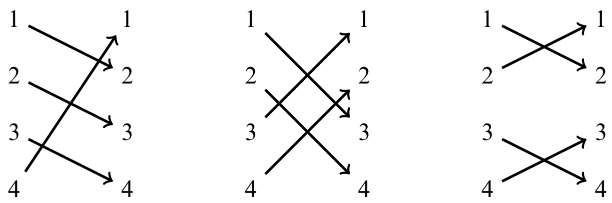
**Proof.** For any particular subset of $[n]$, each element of $[n]$ is either in that subset or not. Thus, to construct a typical subset, we have to make one of two choices for each element of $[n]$. Furthermore, these choices are independent of each other. Hence, the total number of choices – and consequently the total number of subsets – is

$$\underbrace{2 \times 2 \times \cdots \times 2}_{n} = 2^n.$$

□

This proof gives a closed formula for the answer, and, in fact, gives more than was asked. It also tells us how to construct all the subsets.

(b) Assume $n$ people give their $n$ hats to a hat-check person. Let $f(n)$ be the number of ways that the hats can be returned, so that everyone has one hat, but no one has their own hat. If we list the hats on the left and their owners on the right, then the figure below shows three ways of returning hats to four people so that none of them gets their own hat back.




You should check that $f(1) = 0$, $f(2) = 1$, and $f(3) = 2$. You should also try to find $f(4)$.

We will show in Section 8.3 that

$$f(n) = n! \sum_{i=0}^{n} \frac{(-1)^i}{i!}.$$

This is a formula – and you should check your answers for $n = 1, \ldots, 4$ using it – but we would have preferred a "nice" closed formula. While you may not find this formula pleasing, it does work, and will become even more meaningful when we understand the significance of each term in the sum.

It is also possible to show that $f(n)$ is the nearest integer to $\frac{n!}{e}$. This is, of course, easier to use. But this formula may not have a combinatorial significance, and hence, it may be argued, that it gives us less insight. It is, however, fascinating that in answering a question about hats, the number $e$ would make an appearance.

(c) Let $[n] = \{1, \ldots, n\}$. Let $f(n)$ be the number of subsets of $[n]$ that do not contain two consecutive integers. For example, if $n = 4$, then the subsets of $\{1, 2, 3, 4\}$ that do not contain two consecutive integers are

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

Thus $f(4) = 8$. You should also check that $f(1) = 2$, $f(2) = 3$, and $f(3) = 5$.

We will show in Example 1.18 (and again in Example 9.34) that

$$f(n) = \frac{1}{\sqrt{5}}(\tau^{n+2} - \overline{\tau}^{n+2}),\ \text{where } \tau = \frac{1}{2}(1+\sqrt{5}),\ \overline{\tau} = \frac{1}{2}(1-\sqrt{5}).$$

Again, it is unclear how irrational numbers got involved in counting a discrete phenomenon. This formula can actually be used but seems to give little insight into the problem. Sometimes, there are alternatives to finding a closed formula. For this problem, we can prove the following recurrence relation:

CLAIM: $f(n) = f(n-1) + f(n-2)$.

**Proof.** All "good" subsets of $[n]$ either have $n$ or don't have $n$. The ones that don't have $n$ are exactly the "good" subsets of $[n-1]$. The "good" subsets of $[n]$ that include $n$ are exactly the "good" subsets of $[n-2]$ together with $n$. Thus $f(n) = f(n-1) + f(n-2)$. □

Seeing the recurrence relation, we know that the sequence $f(1), f(2), \ldots$ is the Piṅgala–Fibonacci sequence (see Section 1.2.1), and we can use the recurrence relation to generate as many values of $f$ as we want. In fact, the closed formula quoted above can be derived from this recurrence relation.

(d) We have a sequence $a_0 = 1, a_1, a_2, \ldots$ such that, for all $n \geq 1$,

$$\sum_{k=0}^{n} a_k a_{n-k} = 1.$$

We want to find $a_{47}$.

For example,

$$\begin{aligned}
&a_0 = 1 && && \\
&a_0a_1 + a_1a_0 = 1 &&\Rightarrow\quad a_1 + a_1 = 1 &&\Rightarrow\quad a_1 = 1/2\\
&a_0a_2 + a_1a_1 + a_2a_0 = 1 &&\Rightarrow\quad a_2 + \left(\tfrac{1}{2}\right)^2 + a_2 = 1 &&\Rightarrow\quad a_2 = 3/8.
\end{aligned}$$

We see that we could continue and, step by step, calculate the terms of the sequence. This is quite tedious. An alternative way to approach this problem is through generating functions. If we want to find a formula for some function $f(n)$ where $n$ is a natural number, then we can form the generating function of $f(n)$:

$$\sum_{n\geq 0} f(n)x^n = f(0) + f(1)x + f(2)x^2 + \cdots + f(n)x^n + \cdots.$$

Here we are concerned with formal power series, and questions of convergence do not come up (at least for elementary applications). Sometimes this power series has a nice closed form and then we can manipulate this function and get information about $f(n)$.

So for our problem, let

$$F(x) = a_0 + a_1x + a_2x^2 + \cdots.$$

Now

$$\begin{aligned} F(x)F(x) &= (a_0 + a_1x + a_2x^2 + \cdots)(a_0 + a_1x + a_2x^2 + \cdots) \\ &= a_0^2 + (a_0a_1 + a_1a_0)x + (a_0a_2 + a_1a_1 + a_2a_0)x^2 + \cdots \\ &= 1 + x + x^2 + \cdots \\ &= \frac{1}{1-x}. \end{aligned}$$

We have $(F(x))^2 = \frac{1}{1-x}$ and so

$$F(x) = \frac{1}{\sqrt{1-x}}.$$

Now $a_n$ is the coefficient of $x^n$ in the Taylor series expansion of $F(x)$. So, we can use a symbolic algebra software such as SageMath®, Maple®, or Mathematica® to find any desired value of $a_n$.

As an example, in Maple, we first define the function by > $F := \frac{1}{\sqrt{1-x}}$, and then get the coefficient of $x^{47}$ in the Taylor polynomial expansion of $F$ at $x=0$ by `>coeftayl(F,x=0,47)`. We get

$$a_{47} = \frac{5080316063578657 0329644235}{618970019642690137449562112}.$$

It is amazing that calculus can help in solving such a discrete problem. In fact, the generating function $F(x)$ can actually be used to get

$$a_n = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2^n n!}.$$

However, it is not clear that this formula is any better than the generating function.

As the examples show, we will not only use a myriad of techniques for solving counting problems, but we will also refine our sense of what a good solution should look like. This all will (hopefully) become clear as we get our hands dirty and start solving problems.