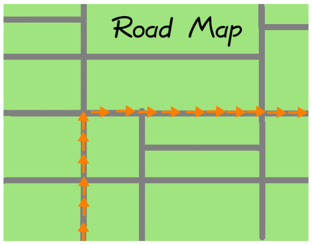
Module 7.7: The Multiplication and Exponent Principles

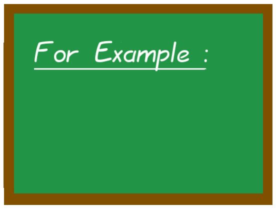


This is the first of three modules that will introduce the rudiments of *combinatorics*, the science of counting the number of ways that something can be done. We study combinatorics not because we desire to count things, but because we can use it in probability. Two important principles of combinatorics will be taught: the "multiplication principle" and the "exponent principle." We will also see another side of the complement principle.

The next two modules will introduce the remaining principles (namely "factorials," "permutations," and "combinations") to cover the six basic building blocks of combinatorics. There is a huge diversity of problems in both combinatorics and in probability that can be solved with these powerful fundamental tools.

In this module, we will also see our first use of tree diagrams, here presented to permit you to count the number of outcomes possible in some simple situations. Later in Module 7, which starts on Page 1097, we will learn to use trees in a different way, to solve very complex probability problems rather directly and simply.

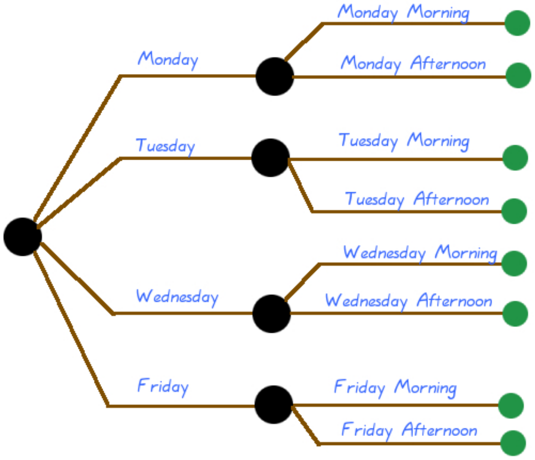
Like many topics in mathematics, the first few problems are going to appear insanely simple. This is not meant to insult your intelligence. As our journey into combinatorics evolves, you will find the problems eventually do get difficult. Sadly, a lot of students mentally dismiss this topic as being too easy to merit much attention. . . until it is too late.



7-7-1

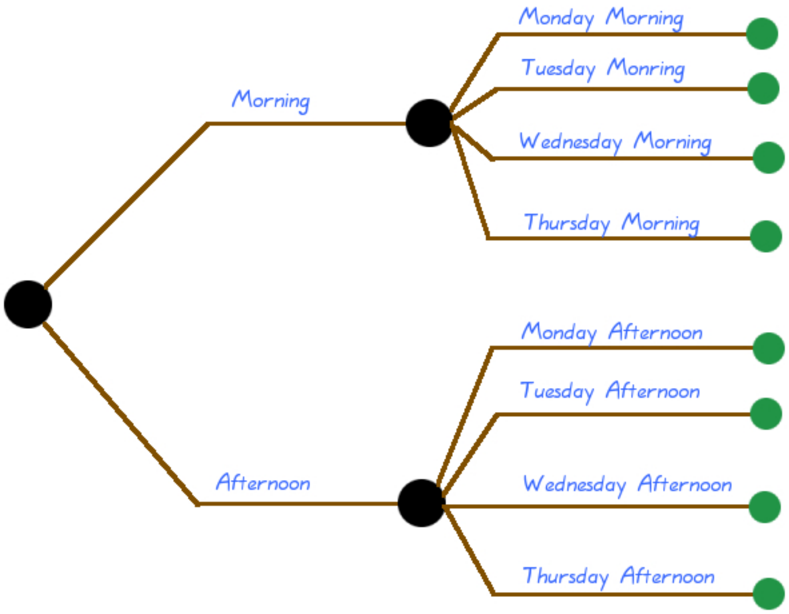
Suppose you are going to get your eyes lasered. The lasering facility has either morning or evening appointments available, and they do surgeries on Monday, Tuesday, Wednesday, or Friday (but on Thursday the surgeon plays golf). How many possible choices of appointment are there?

Consider the following diagram:

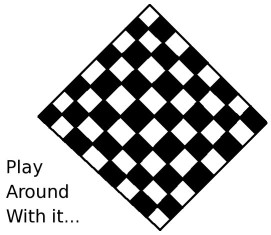


As you can see, there are four choices of day, and then two choices: either morning or evening. The first four lines (which are on the left) represent the choice of day; the eight lines on the right come from the fact that each of the four lines on the left produces two more lines, and so there are $4 \times 2 = 8$ of them. Thus it is clear there are 8 choices.

Alternatively, you could have considered the choice of morning or afternoon first, and then afterward, decided on what day to go. Then you'd have the following diagram:



While the tree diagram is a bit different, there are still 8 "leaves" at the end, which represent 8 final outcomes. This is because $2 \times 4 = 8 = 4 \times 2$.



# 7-7-2

Let's suppose that a computer is randomly assigning the appointments to patients, in such a way as to render each appointment equally likely. An exciting speaker is coming to Alice's corporation on Wednesday afternoon to talk about a new technology, and she really wants to go. Referring to the tree diagram for the laser surgeon, what is the probability that the time she is given for her appointment is a Wednesday afternoon?

[Answer: 1/8.]

Note that in the previous box, because we said "equally likely" and "chooses randomly", then we know that the equally likely assumption (see Page 930) is in effect. In turn, because that assumption is in effect, we are allowed to say that each of the 8 leaves (each of the 8 outcomes) has probability 1/8.

Applications to Science !

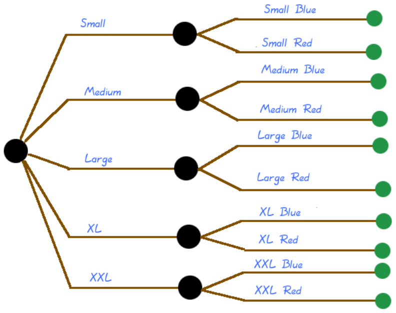
I should mention something about the problem in the previous checkerboard box. You might wonder if appointments are ever actually scheduled at random.

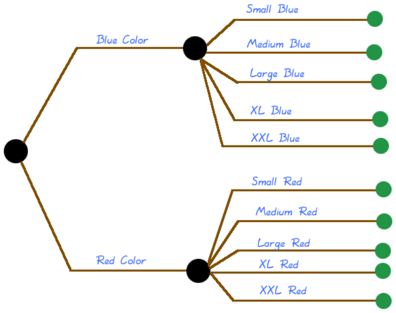
In reality, appointments are not usually scheduled randomly, except in very rare circumstances such as illegal drug tests, where deliberately surprising the patient has an advantage. Randomly scheduled appointments are good at that.

For Example :

# 7-7-3

Now consider that a local honor society at your college is selling T-shirts. They can be red or blue in background, and come in five sizes: `S`, `M`, `L`, `XL`, and `XXL`. If you consider size first, you'd have the diagram on the left; if you consider color first, you'd have the diagram on the right.





Either way, it is quite clear that there are $5 \times 2 = 10 = 2 \times 5$ choices. Notice, once again, that the order in which we considered the properties of the T-shirt did not matter.

Play Around With it...

# 7-7-4

Consider the problem of the previous box. Suppose Alice wears Medium Red and Bob wears Large Blue. What is the probability that a T-shirt (drawn at random from the available categories) will fit either of them, and why? [Answer: $2/10 = 1/5$.]

The "reason why" will be given in the box.

. . . but why?

Looking at the previous box, why is the answer 2/10 or 1/5? Because there are 10 shirt color-size combinations in the sample space, and 2 of them (Medium/Red and Large/Blue) meet the criterion of fitting either Alice or Bob. Thus the answer is 2/10, which reduces to 1/5.

We have used the "equally likely assumption" in this problem. Is it justified? Since the person selecting a shirt is drawing it "at random from the available categories" then there are simply 10 categories. Unless the person selecting a shirt has been bribed or is cheating, then the assumption is entirely justified.

For Example :

# 7-7-5

Let's return to the two color case of the T-shirt problem. Refer back to the original diagrams two boxes ago. If we were to decide that we don't really need the `XXL` size, then we would take a chainsaw to the tree on the left and hack off the bottom left branch for `XXL`. This would cause the two branches attached to it to come off as well, namely "`XXL` Blue" and "`XXL` Red".

If you prefer the diagram on the right then we would take a chainsaw and hack off the bottom branch of the first cluster of five branches, namely "`XXL` Red", and also the second cluster of five branches, namely "`XXL` Blue". Either way, at this point we have $4 \times 2 = 8 = 2 \times 4$ choices, fewer than when we started.

For Example :

# 7-7-6

Consider the problem of the previous two boxes, but with a third color (white T-shirts) added into the mix.

- Suppose we were to offer a white-backgrounded shirt as well. Then if you consider the diagram on the left, each of the 5 left-most branches of the tree would have 3 and not 2 branches on coming out, resulting in $5 \times 3 = 15$ final leaves or outcomes.
- Likewise, if you consider the diagram on the right, there would be three colors, and not two, yet each of these would still come in five sizes. So we would have $3 \times 5 = 15$ leaves or outcomes.
- Notice that the answer is the same in either case.

Valuable Information

At this point, we have discovered the following idea:

If you present someone with a series of decisions, and the first decision has x choices and the second decision has y choices, then there are a total of xy choices. This is the *multiplication principle*.

Likewise, the principle can be extended. If you present someone with three decisions, the first having x choices, the second having y, and the third having z, then there are a total of xyz choices.

This works for any number of decisions. Now we'll explore applications with three or more decisions, whereas before there were only two decisions.

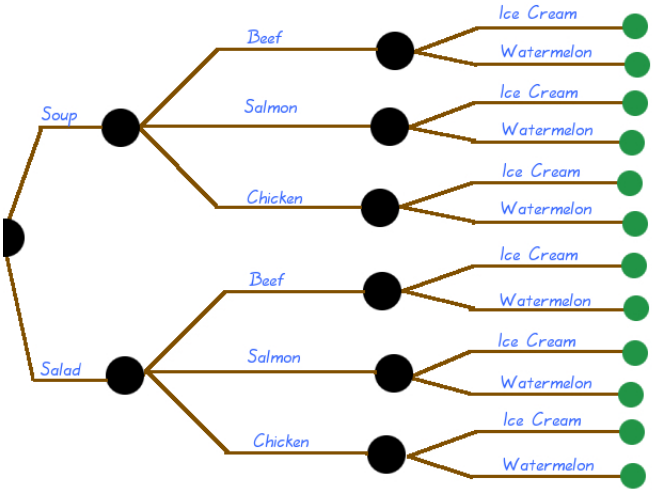
For Example :

# 7-7-7

Suppose someone is at a fancy restaurant for New Year's Eve, and they have a *prix fixe* menu. (That's pronounced like "prefix" except that the "e" before the "f" is long, as in "glee" or "bee;" you definitely do not pronounce the "x" in the middle.) On such a menu, there is usually a small number of choices for each course. Suppose there are two possible appetizers: soup or salad (which you chose first). Further suppose there are three possible main courses: beef, salmon, or chicken (which you chose second). Finally suppose there are two possible desserts: ice cream or watermelon (which you chose third).

The tree diagram is given in the next box.

Here is the tree diagram for the previous example:



As you can see, there are 2 × 3 × 2 = 12 choices.

Play Around With it...

# 7-7-8

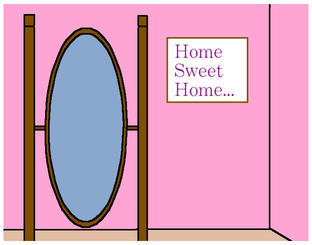
Looking at the example of the previous box, suppose there is a customer who is from overseas, and who cannot understand the menu at all. He orders by choosing randomly.

- What is the probability for each and every outcome (leaf) on the tree? [Answer: 1/12.]
- Let's say that the waiter's favorites are "salad-salmon-watermelon." What is the probability that our randomly guessing foreigner picks the waiter's favorite, by coincidence, for all three courses? [Answer: 1/12.]
- Remembering that the waiter's favorites are "salad-salmon-watermelon," what is the probability that our randomly guessing foreigner picks the waiter's favorite, by coincidence, for at least one course? [Answer: $10/12 = 5/6$.]
- If each path through the tree costs the restaurant a different amount of money in ingredients, time, and wages, then what is the probability that our randomly-ordering customer happens, by coincidence, to order the most expensive possible selection? [Answer: 1/12.]

. . . but why?

The multiplication principle is sometimes called the *restaurant principle*. The last two boxes have shown you why.

That last question in the previous box is a bit of a trick question. We don't know which path through the tree is the most expensive. However, whichever path is the most expensive, it must be probability 1/12, because all the paths have probability 1/12.



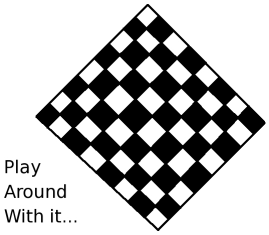
A Pause for Reflection...
We were able to use the "equally likely assumption" in the previous example and previous checkerboard, because the unfortunate customer, who cannot understand the menu, is ordering at random.

Otherwise, we should not assume that each particular choice is equally likely. Some dishes might be more popular than others. Similar arguments can be made for T-shirt sizes. Common sizes like `M` or `L` are much more likely to be ordered than uncommon sizes like `XS` or `XXL`. A common error is that students will use the "equally likely assumption" when it is not warranted.

In order for us to understand when the "equally likely assumption" makes sense and should be used, we should understand when it fails to make sense and should not be used.

Now that we know the multiplication principle, we can stop drawing the tree diagrams. They take up space and time, and do not *yet* serve a purpose.

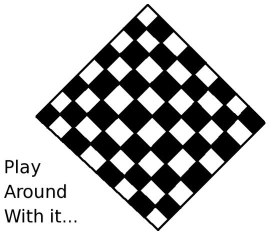
However, later in this chapter, (e.g. on Page 1097) we'll use tree diagrams again, in a very different way. At that time, the branches and leaves will carry much more information than just a label. They'll have probabilities written upon them!



# 7-7-9

Continuing with the New Year's Eve problem of the previous three boxes, try to answer the following questions, but without drawing the probability tree.

- Suppose they add a vegetarian entree, what is the number of choices possible for the entire meal? [Answer: 16 choices.]
- Suppose they add, in addition to the vegetarian entree, a third dessert? [Answer: 24 choices.]



# 7-7-10

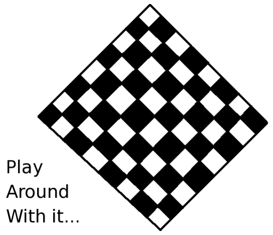
The math club is selling some sweatshirts with a cool student-designed graphic on them, with the goal of raising some funds. There are five possible sizes, `S`, `M`, `L`, `XL`, and `XXL`; the sweatshirts come in four colors.

- How many different possible sweatshirt orders are there? [Answer: $5 \times 4 = 20$.]
- How does the problem change if an `XS` size is introduced, perhaps for the young children of faculty? [Answer: $6 \times 4 = 24$.]

Note: The number rose from 20 to 24, which means 4 new possible orders: namely the `XS` in each of the four colors.

- What if a fifth color is added, in addition to introducing the `XS` size, then how many possible orders will there be? [Answer: $6 \times 5 = 30$.]

Note: Again, the number rose from 24 to 30, which means 6 new possible orders: namely the new color in each of the six sizes, `XS` through `XXL`.

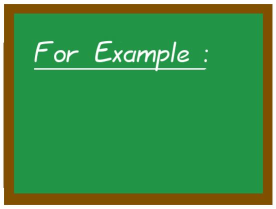


Suppose a restaurant (on its early-bird-specials menu) has a choice of 2 salads and 3 soups for an appetizer, and then 4 entrees. How many possible orders of one appetizer and one entree are there? [Answer: $5 \times 4 = 20$.]

# 7-7-11

Did you say $2 \times 3 \times 4 = 24$ for the problem? Combinatorics problems are notorious for tricky wording. You must read very carefully, and pick out the exact meaning of words.

The tiniest variation of phrasing could result in an entirely different problem. This is not a "big picture" or "main ideas" situation, but rather an example of a situation where we might say "the devil is in the details."



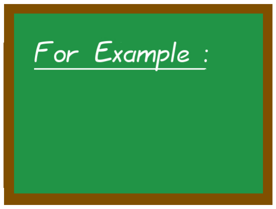
Consider a briefcase with a combination lock. The lock has three wheels, and each wheel is numbered 0–9. How many possible combinations are there, if the choice of a combination is completely unrestricted—the owner can choose any combination desired.

Well, the owner of the case can choose 10 digits for the first wheel, and 10 digits for the second wheel, followed by 10 digits for the third wheel. Thus, there are

$$10 \times 10 \times 10 = 10^3 = 1000$$

possible combinations.

# 7-7-12



Suppose a thief can try one combination every two seconds, and he has 15 minutes while the owner is distracted. What is the possibility that he gets the case open? Assume all combinations are equally likely.

There are 1000 outcomes in the sample space, as we noted in the previous box. The thief will try 30 of these during each minute while the owner is distracted, because there are 60 seconds in a minute. Over the 15 minutes, there are $(15)(30) = 450$ combinations he will try.

Therefore, the thief has probability

$$\frac{450}{1000} = \frac{9}{20} = 0.45 = 45\%$$

of succeeding. Surely that is too high of a risk to bear.

# 7-7-13

Play
Around
With it...

# 7-7-14

Continuing with the briefcase problem of the previous two boxes,

- What if there were four digits, how many combinations are there? [Answer: 10,000.]
- What if there were five digits, how many combinations are there? [Answer: 100,000.]
- What if there were four digits, what is the probability the thief gets the case open? [Answer: 9/200 or 4.50%.]
- What if there were five digits, what is the probability the thief gets the case open? [Answer: 9/2000 or 0.45%.]

Play
Around
With it...

# 7-7-15

Cipher locks are locks on a door where you have to press a certain code on a keypad to get in. They were very popular at the National Security Agency, where I used to work, until it was discovered that they spread disease such as colds and flus, because so many people have to touch them so frequently.

In any case, how many combinations are there for

- A lock with numbers 1–6 as keys, there are no restrictions on the code, and a code is 5 digits? [Answer: $6^5 = 7776$.]
- A lock with numbers 1–5 as keys, there are no restrictions on the code, and a code is 6 digits? [Answer: $5^6 = 15,625$.]
- A lock with numbers 1–7 as keys, there are no restrictions on the code, and a code is 4 digits? [Answer: $7^4 = 2401$.]

Note, this is one of those situations where combinatorics can be surprising. It was not immediately obvious to me at all that 5-key 6-digit combinations would be more than double the number of possibilities compared to 6-key 5-digit combinations.

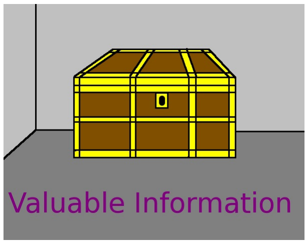
Play
Around
With it...

# 7-7-16

Examine each of the cases of the previous box, and consider a burglar who can try 30 combinations per minute, and who has an hour alone at the cipher lock.

- How many combinations will he try? [Answer: 1800 combinations.]
- What is the probability that he will get through the first cipher lock of the previous box? [Answer: $23.1481\cdots\%$.]
- What is the probability that he will get through the second cipher lock of the previous box? [Answer: 11.52%.]
- What is the probability that he will get through the third cipher lock of the previous box? [Answer: $74.9687\cdots\%$.]
- If we give the burglar two hours, what is the probability that he will get through the third cipher lock of the previous box? [Answer: 100%.]

By the way, please understand that we do not write probabilities above 100%. The probability of 100% represents absolute certainty.

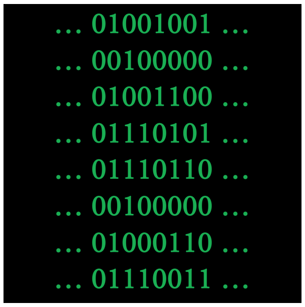


If you examine closely the problem of the previous two boxes, as well as the briefcase problem from Page 997, you probably can see a pattern. Our answer was always of the form $n \times n \times n \times \ldots \times n$.

When one must make c selections, from a set of n options, where order matters, and repeats are allowed, then the number of possibilities is n^c. This is the exponent principle.

This is our first member of a four-entry chart. This chart is phenomenally useful in solving combinatorial problems. We will fill it in during later modules of this chapter.

	Repeats OK	No Repeats
Order Matters	Exponent Princ.	??
Order Doesn't Matter	??	??



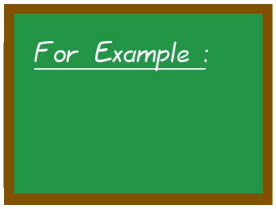
We all learn in elementary school that computers work with 1s and 0s, a system called binary. Sequences of 1s and 0s such as

```
0100 1000          0100 1001
```

are called *binary strings* or *bit strings*. Each 1 or 0 is called a "bit," which stands for "binary digit."

Bit strings are excellent ways of encoding numbers, text, images and sounds. The two 8-bit strings above can represent 72 and 73 as numbers, or the word "HI" in the American Standard Code for Information Interchange (ASCII).

To encode pictures, the color of each pixel must be recorded. We will now explore that in the next box.



# 7-7-17

You've probably heard expressions like "16-bit color" or "24-bit color," either in reference to digital cameras or video games. You might wonder what that means. Our basic question is to figure out how many possible bit strings exist with length 16. In other words, how many ways can we write down a sequence of 16 bits, each of which is 0 or 1?

The first thing to know is that order matters. It is not obvious at all, but order matters for bit strings. If I were to rewrite the bit strings from the previous box in a different order, I might get

```
0100 0001          0100 0011
```

which represents the numbers 65 and 67, or the letters "AC" in ASCII. While order matters, we must now determine if repeats are allowed.

If you think about it, it is very clear that repeats are allowed. After all, I just repeated 0 and 1 several times in the above examples. Since repeats are allowed and order matters, we are using the exponent principle. There are two possible bits (0 and 1), and we need 16 of them. Therefore, there are

$$2^{16} = 65,536$$

possible bit strings of length 16. This means that 16-bit color allows for 65,536 colors.

Play
Around
With it...

# 7-7-18

Using the technique of the previous box, compute the following:

- How many colors are there in 15-bit color? [Answer: 32,768.]
- How many colors are there in 24-bit color? [Answer: 16,777,216.]
- How many colors are there in 8-bit color? [Answer: 256.]
- How many colors are there in 4-bit color? [Answer: 16.]

I remember in the 1980s when really old computers had graphics cards that could be rated as VGA, EGA, or CGA. The worst was CGA, and you only could see 16 colors on your computer screen.

For Example :

# 7-7-19

The secret key of certain types of encryption systems (called “block ciphers”) is just a string of bits. If you know the string of bits (called “the secret key”) that was used to encrypt a message, then you can read the message; if you do not, then you cannot—unless the cipher has been broken. If a cipher has not been broken, one technique that hackers can use is that they can have a computer, or several, try all possible bit strings. This can take a long time, depending on the cipher. Suppose that someone is using a block cipher with 64-bit keys. How many possible keys are there? If a hacker has access to technology that can test 100 million keys per second on each core, and he has access to 100 cores, then how long will it take (in years and days) to try every key? (Use a 365-day year.)

In a bit string, order matters, and of course repetitions are allowed. Therefore, we are using the exponent principle. That means there are

$$2^{64} \approx 1.84467\cdots 10^{19}$$

possible keys. It is a number that is so large, that the human mind often has trouble comprehending it.

In the next box, we will see how to convert this information into a more useful form.

In the previous box, we had computed that there are 2^{64} possible keys to this cipher. You might be wondering how we would use that information to compute how long it would take for the hacker's computers to try all of the keys.

With 100 cores, the hacker's computer system can try

$$(100,000,000)(100) = 10^{10}$$

keys per second. This means that the entire operation will take

$$\frac{1.84467\cdots 10^{19}}{10^{10}} = 1.84467\cdots \times 10^{9}$$

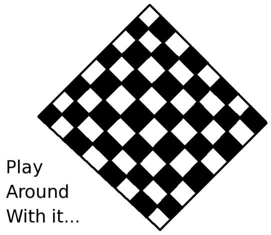
seconds. That turns out to be 58 years and $180.398\cdots$ days.

Play
Around
With it...

# 7-7-20

The previous box said that the hacker had access to 100 cores. This can be 25 laptops, with 4 cores each, as is common among laptops produced in the year 2015. The rate of 100 million keys per second is generous, but only slightly—it is approximately correct. However, some machines used in scientific computing have 64 cores, or many more. Suppose that the Russian Mafia is interested in recovering the secret key used by the Moscow police department. Suppose that they have access to 50 such scientific computing machines with 64 cores.

Using a 365-day year, how many, years, days, and hours will it take to try all the possible secret keys? [Answer: 1 year, 302 days, and $4.79\cdots$ hours.]



# 7-7-21

Two block ciphers used by the US Government are very famous. One of them, the Data Encryption Standard (DES), was proposed in 1977 and is out of date. It used 56-bit keys. The other, the Advanced Encryption Standard (AES), uses 128-bit keys (in one particular standard configuration). Using the same 50 machines as the previous box, tell me how long it will take to try all the keys for these ciphers.

- For the 56-bit case, in days, hours, and minutes?
  [Answer: 2 days, 14 hours, 32.9996 minutes.]
- For the 128-bit case, in years and days?
  [Answer: $3.37196\cdots 10^{19}$ years.]

Note: for comparison, the universe is very approximately 2×10^{10} years old. The length of time indicated, "$3.37196\cdots 10^{19}$ years" is over one billion times the age of the universe.

The results of the previous box can be hard to believe. Let's check them. From the laws of exponents, we know that

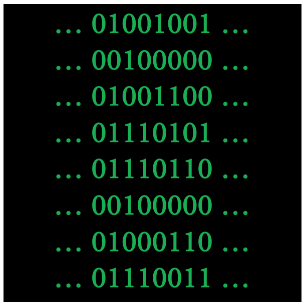
$$2^{64} \div 2^{56} = 2^8 = 256$$

therefore we expect that trying all possible keys for the 56-bit cipher should be 256× faster than finding all the keys for the 64-bit cipher. Meanwhile, the 64-bit cipher requires 1 year and 302 days, which is $365 + 302 = 667$ days.

Therefore, we expect

$$667 \div 256 = 2.60546\cdots \text{ days } = 2 \text{ days } 14.5312\cdots \text{ hours}$$

for the 56-bit cipher, and that's what we had computed.



What is the take-away from the previous two boxes? What is the true message for tomorrow's business executive?

At the dawn of wide-spread e-commerce back in the late 1990s, when computer scientists were trying to encourage businessmen to move away from 56-bit ciphers like DES, there were business people who said "56 bits, 64 bits, what's the difference? It is almost the same, right?" It should be noted that changing encryption systems is expensive, which might add to the hesitation of an executive.

Clearly, that kind of thinking is wrong. In one case, you have 667 days, and in the other, you have roughly 2.5 days. That's not the same thing at all. Executives should listen to the computer scientists that they hire.

Likewise, there is a movement today (in 2015), to get all businesses to use 128-bit keys for all block ciphers. In that case, there would be no possibility of trying all possible bit strings, no matter how many computers the hackers purchase to help them. That kind of security is very useful. Perhaps executives should learn to listen.

Play
Around
With it...

# 7-7-22

You're probably aware that computers on the internet have something called an "IP address." This is just like a phone number—it identifies a device on the network, so that the signals intended for it can actually reach the intended device.

In the protocol that has been in primary use since 1983, called "IPv4" (Internet Protocol Version 4), the IP address is a 32-bit string. How many possible IP addresses are there? [Answer: 4,294,967,296 addresses.]

```
... 01001001 ...
... 00100000 ...
... 01001100 ...
... 01110101 ...
... 01110110 ...
... 00100000 ...
... 01000110 ...
... 01110011 ...
```

Looking at the previous box, it is clear that 4,294,967,296 is not enough, because in September of 2015, the world's population was about 7.36 billion people.

Also, many IP addresses are used for routers and other communications devices that make the internet actually function. Furthermore, many people have multiple internet-capable devices and those devices might have multiple IP addresses each. As if that were not enough, the IP addresses are allocated in a very awkward way, called "subnet masking."

Subnet masking is a bit complicated, so we won't talk about it in this textbook.

Play
Around
With it...

# 7-7-23

Continuing with the previous box, a later version of the protocol is called "IPv6" and uses 128-bit strings.

- How many possible IP addresses are there? [Answer: 3.40282×10^{38} addresses.]

Note: We're now going to compute how many IP addresses this will allow per square inch of the earth's surface. A quick internet search reveals that the formula for the surface area of a sphere is $A = 4\pi r^2$, where r is the sphere's radius. The radius of the earth is $3959.60\cdots$ miles on average.

- What is the surface area of the earth, in square miles?
  [Answer: $1.97021\cdots \times 10^8$ sq mi.]

We will continue with this problem in the next box.

Play
Around
With it...

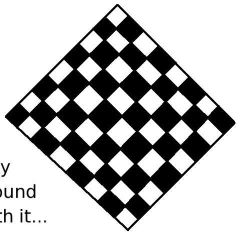
# 7-7-24

Before we continue with the previous box, it is useful to note for our foreign students, that there are 5280 feet are in a mile, and 12 inches in a foot.

- What is the surface area of the earth, in square feet?
  [Answer: $5.49263\cdots \times 10^{15}$ sq ft.]
- What is the surface area of the earth, in square inches?
  [Answer: $7.90939\cdots \times 10^{17}$ sq in.]
- How many IP addresses can IPv6 give to each square inch of the earth's surface?
  [Answer: $4.30225\cdots \times 10^{20}$ addresses per square inch.]

As you can see, this is more than enough, by leaps and bounds. That's more than 430 million trillion IP addresses per square inch (including all the oceans and the ice caps). Unfortunately, IPv6 isn't used very much at all. However, there are other methods of addressing the shortage of IPv4 addresses, including something called "network address translation" (NAT).

The previous two boxes that explained IPv4 and IPv6 via the surface area of the earth were suggested to me by Dr. Dan Drake, at a conference in Washington State, in the Summer of 2015.



# 7-7-25

While the idea of dividing up the world's surface by IP addresses is only a thought experiment, there is a company that has tried to divide up the world into tiny squares, each identified by three ordinary words from the English language. This could be very good for arranging medical evacuations from remote locations, meeting someone in a busy place like Grand Central Station, or even delivering Chinese food. The company claims that three words in order, such as "`usage ample soup`" will be sufficient to identify a small area.

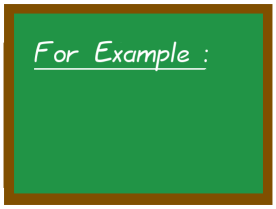
- They claim that there are 57 trillion such squares. If that is true, then how many words are in their dictionary? [Answer: 38,485 words.]

Hint: For the next one, look back at the previous checkerboard box.

- What is the surface area of each square, measured in sq ft? [Answer: $96.3619\cdots$ sq ft.]

This is neat, because you could specify destinations down to rooms in a house. You could even identify in which parking space of a parking lot an object was found. Of course, this can be done with latitude and longitude, but dictating those coordinates via a phone call or even a text message is much more prone to a typographical error than three ordinary words. (Of course, the company has coordinates available in many languages, for use by non-English speakers.)

The above problem was suggested by my husband, Patrick Studdard. For further information, you can read the article "This Program Divided the World Into 57 Trillion Squares and Gave Them Names Like usage.ample.soup," by Joshua Keating, published on `slate.com` on June 27th, 2016.



# 7-7-26

Canadian postal codes work as follows: there are six symbols: a letter, a digit, another letter, another digit, yet another letter, and finally a digit. For example, `A1A 1A1` is St. John's Newfoundland. How many possible postal codes are there?

There are two different ways to approach this. First, we could use the multiplication principle directly:

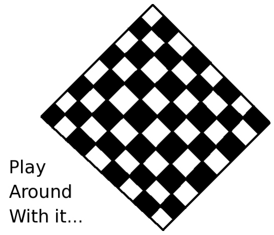
$$26 \times 10 \times 26 \times 10 \times 26 \times 10 = 17,576,000$$

Alternatively you could realize that the three letters have $26^3 = 17,576$ choices, and the three digits have $10^3 = 1000$ choices. Then the multiplication principle allows for

$$17,576 \times 1000 = 17,576,000$$

Often if you see two roads to a solution, a great way to check your work is to follow both roads to completion, and be sure you get the same answer in each case.

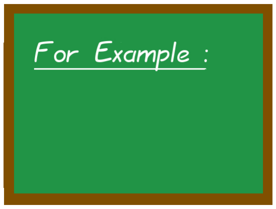
Just as an example, we saw that both routes in the previous box brought us to the same answer, even though the routes were significantly different.



# 7-7-27

What would happen if Canadian postal codes were not in strict alternation? What if you could have any sequence of six symbols, where each symbol could either be a letter or a digit? How many possible postal codes are there?

[Answer: $36^6 = 2,176,782,336$.]

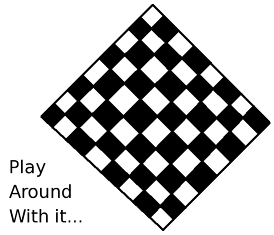


# 7-7-28

Using the information of the previous example and previous checkerboard, we can compute an interesting probability. We can compute the probability that a sequence of six symbols, chosen uniformly at random, where each symbol is either a letter or a digit, just happens to be of the type letter-digit-letter-digit-letter-digit, by coincidence.

Since there are $(26^3)(10^3)$ postal codes of the letter-digit-letter-digit-letter-digit style, and there are 36^6 in general, we can compute

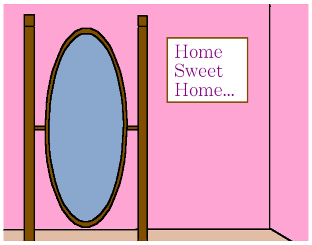
$$\frac{26^3 \times 10^3}{36^6} = \frac{17,576,000}{2,176,782,336} = 8.07430 \times 10^{-3} \approx 0.80\%$$



# 7-7-29

Suppose a charity selects a board of 7 directors. A chairperson and deputy chair must be chosen. The secretary prepares a ballot, with a box to check for any possible pair of names that anyone could want to vote for.

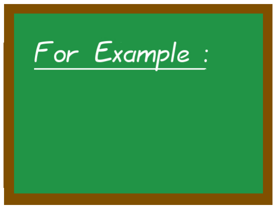
- How many check boxes will there be? [Answer: 42.]
- Suppose that a year later, the board is enlarged to be 19 members. Again, there is an election for a chairperson and deputy. How many checkboxes must there be? [Answer: 342.]



A Pause for Reflection...

Let's re-examine the previous box. The 342 checkboxes would clearly require a large piece of paper for a ballot. Combinatorial problems often defy intuition in the sense that small changes in the parameters of the problem can cause enormous changes in the number or count that you are trying to compute.

For example, in the previous box, the board size only grew by 12 people, but the number of checkboxes grew by 300. This is all the more reason that we should proceed slowly, with diligence and attention. A parallel example of such rapid change can be found on Page 1019 in Module 7, and we've seen an example before on Page 980 in Module 7.



# 7-7-30

Suppose a particular US State has license plates formed of six symbols—three letters followed by three numbers. Suppose further that the numerical sequences 666 and 000 are prohibited. How many possible license plates are there?

For the first, second, or third symbol, any letter from the 26 letters of the English alphabet can be chosen, so there are

$$26 \times 26 \times 26 = 17,576$$

of them. For the numerical part, normally we would say

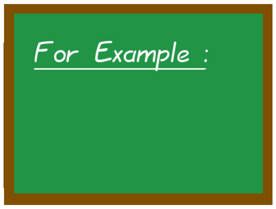
$$10 \times 10 \times 10 = 1000$$

but we must exclude 666 and 000, so we say 998. Then we have

$$17,576 \times 998 = 17,540,848$$

possible license plates.

Of course, this would eventually be a problem, if the state has a large population.



# 7-7-31

Now we're going to reconsider the problem in the previous box, and look at how the problem would change if we add the restriction that no two consecutive symbols can be the same. That means that it is not permitted to have two letters in a row to be equal, nor two numbers in a row to be equal.

For the first letter, we can choose any one of 26 choices. For the second letter, we can choose one of 25—any of the 26 excepting the one we chose to be the first letter. For the third letter, we can still choose 25—the 26 letters of the alphabet excluding the one we chose for the middle letter, but not excluding the letter we chose for the first letter. That comes to

$$26 \times 25 \times 25 = 16,250$$

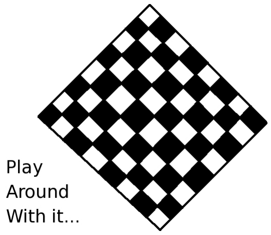
and by similar reasoning, for the numbers we obtain

$$10 \times 9 \times 9 = 810$$

for a total of

$$810 \times 16,250 = 13,162,500$$

Note that we do not subtract anything for 666 nor 000, because those are already prohibited by the prohibition on repeated symbols.



# 7-7-32

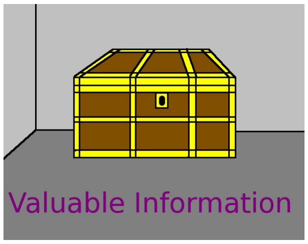
We return now to the situation of the license plates being three letters followed by three numbers, with no restrictions on repetition, except that 666 and 000 are prohibited.

- How many license plates have either a repeated letter or a repeated number, by coincidence?

  [Answer: $17,540,848 - 13,162,500 = 4,378,348$.]

- What is the probability that a random license plate has such a repetition?

  $$\left[\text{Answer: } \frac{4,378,348}{17,540,848} = 0.249608\cdots\right] \approx 24.96\%$$



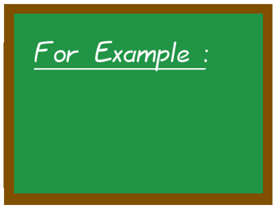
In the previous box, we used an indirect method for finding our answer. If we are seeking to know the answer to the question "What is the probability that a random license plate has a repetition?" then we might be flummoxed.

It is not obvious at all how to directly compute the number of license plates with a repetition. However, we can calculate the number of license plates WITHOUT a repetition (13,162,500) and the number in general (17,540,848), then surely the number of plates WITH a repetition will be

$$17,540,848 - 13,162,500 = 4,378,348$$

This technique is called the *complement principle*. To count the number of objects from a set $\mathcal{S}$ with a certain property, you instead compute the number of objects from a set $\mathcal{S}$ without that property, and the number of objects in the set $\mathcal{S}$ in general, and then subtract.

While this looks childishly easy, it is a very powerful tool. This is one of those things that we learn not because it is hard, but because it is useful. We will explore the complement principle more in this module and the next few modules.



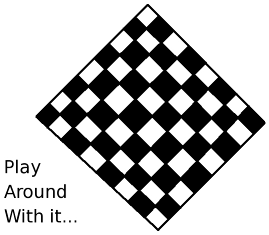
# 7-7-33

Let us consider now a combination lock such as you would see in a locker room. For now, consider a combination of three digits, where each digit can be 0–9. What if we require that no digit can be repeated? Then how many possible combinations are there?

Then there are 10 choices for the first digit, 9 choices for the second digit, and 8 choices for the third digit. This comes to

$$10 \times 9 \times 8 = 720$$

choices.

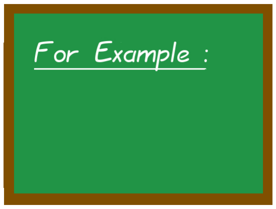


# 7-7-34

How about if no digit can be repeated, and ...

- ... there are four digits? [Answer: $10 \times 9 \times 8 \times 7 = 5040$.]
- ... there are five digits? [Answer: $10 \times 9 \times 8 \times 7 \times 6 = 30,240$.]
- ... there are six digits? [Answer: $10 \times 9 \times 8 \times 7 \times 6 \times 5 = 151,200$.]

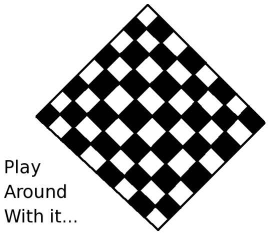
Why would we care about the case when no digits are repeated? In certain circumstances, the manufacture of a locker is much easier, and therefore cheaper, if the acceptable combinations never have repeated digits.



# 7-7-35

As we mentioned in the previous box, locks are easier to design (in certain circumstances) if combinations are forbidden to have repeated numbers. As you know, some locks come with an initial combination chosen by the manufacturer; in most cases, this combination can be changed by the owner of the lock, but in some cases it cannot be. If we choose an initial combination for the lock (during the manufacturing process) uniformly at random (meaning that all combinations are equally likely) then what is the probability we get one without any repeated digits, by coincidence? (As before, the digits can be 0–9.)

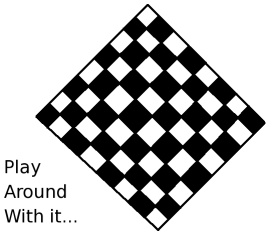
- Three-Digit Locks: Well, there are 1000 combinations in general for three digits, and as we figured out moments ago, there are 720 combinations with no repeated digit, so $720/1000 = 72.00\%$.
- Four-Digit Locks: Likewise, there are 10,000 combinations in general for four digits, from previous our work, and there are 5040 with no repeated digit, so $5040/10,000 = 50.40\%$.



# 7-7-36

What is the probability that a combination for a lock (chosen uniformly at random) will have no repeated digits? As before, the digits can be 0–9. Consider the following combination lengths:

- Five-Digit Locks: [Answer: 30.24%.]
- Six-Digit Locks: [Answer: 15.12%.]



# 7-7-37

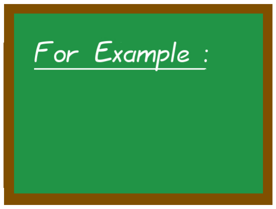
For the following lengths of combinations, compute the probability that a combination for a cipher lock (chosen uniformly at random) will have one or more repeated digits. Hint: use the complement principle.

- Three-Digit Locks: [Answer: $280/1000 = 28.00\%$.]
- Four-Digit Locks: [Answer: $4960/10,000 = 49.60\%$.]
- Five-Digit Locks: [Answer: $69,760/100,000 = 69.76\%$.]
- Six-Digit Locks: [Answer: $848,800/1,000,000 = 84.88\%$.]

Here's a question: suppose I was making combinations for a lock, where each combination was an 8-digit number. How often would I expect that the combination has two adjacent digits being equal, by coincidence?

Using intuition alone, it seems to me as though this would be at least somewhat rare. A repeated digit in the 8-digits is one thing. Yet, when you add the word "adjacent," it seems to me that most combinations would not have that. However, intuition is never a substitute for computation.

You will compute the answer to this question yourself, three boxes from now!



# 7-7-38

Let's look at another situation with a 4-digit combination lock. What would happen if we disallow consecutive digits from being the same, but repeats in general are allowed? For example, 8989 is permitted, but 8998 is not permitted.

For three digits, there are 10 choices for the first digit, 9 choices for the second (it cannot equal the first), and 9 choices for the third (it cannot equal the second, but is otherwise unrestricted). This comes to

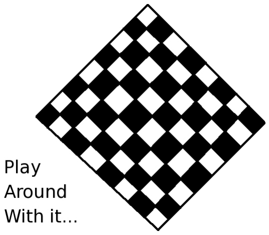
$$10 \times 9 \times 9 = 810$$

choices for three digits, and similarly

$$10 \times 9 \times 9 \times 9 = 7290$$

choices for four digits.

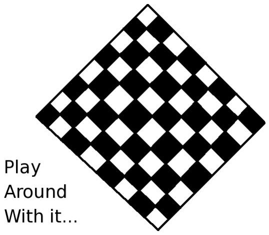
This question can be important because the difficulty, and therefore cost, of manufacturing a lock is affected (under certain circumstances) by whether or not adjacent numbers can be the same.



# 7-7-39

What is the probability that a combination for a cipher lock (chosen uniformly at random) will have no consecutive digits the same, for the following lengths? As before, the digits can be 0–9.

- Four Digits? [Answer: 7290/10,000 = 72.90%.]
- Five Digits? [Answer: 65,610/100,000 = 65.61%.]
- Six Digits? [Answer: 590,490/1,000,000 = 59.04%.]
- Eight Digits? [Answer: $47,829,690/100,000,000 \approx 47.82\%$.]



# 7-7-40

What is the probability that a combination for a cipher lock (chosen uniformly at random) will have some consecutive digits the same, for the following lengths? As before, the digits can be 0–9.

- Four Digits? [Answer: 2710/10,000 = 27.10%.]
- Five Digits? [Answer: 34,390/100,000 = 34.39%.]
- Six Digits? [Answer: $409,510/1,000,000 \approx 40.95\%$.]
- Eight Digits? [Answer: $52,170,310/100,000,000 \approx 52.17\%$.]

As you can see, my intuition was wrong. A slim majority of 8-digit combinations will have two adjacent digits coming out equal. We should always remember that intuition is no substitute for computation.

Play
Around
With it...

# 7-7-41

Have you ever seen an 8-digit lock? I haven't. Let's find out why. Suppose a thief can try 25 combinations a minute, on a lock where no consecutive digits can be the same. How long will it take him (in minutes) to check all the combinations for the following lengths?

- Six Digit Lock: [Answer: 23,619.6 minutes.]
- What is that in days, hours, minutes, and seconds?
  [Answer: 16 days, 9 hours, 39 minutes, and 36 seconds.]
- Eight Digit Lock: [Answer: $1,913,187.\cdots$ minutes.]
- What is that in years, days, and hours, using a 365-day year?
  [Answer: 3 years, 233 days, and 14.46 hours.]

I think it is unrealistic to imagine a thief trying combinations for 16 days straight, with no sleep, no meals, no bathroom breaks, and no stopping for any reason. Therefore, six digits is the most that you ever see in practice—for a reason.

Play
Around
With it...

# 7-7-42

Let's imagine that your boss is discussing the manufacturing of the locks with you. He understands that the cost of manufacturing the lock will be lower if repeated digits are banned. However, he's worried about the security implications. Obviously, the number of combinations to be guessed will be smaller in the case when no digits can be repeated. To what extent does this affect the probability that a thief can guess the combinations? Let's consider a thief who has an hour, and who can try 20 combinations per minute. He is trying to guess the combination to a 6-digit lock. We will assume that the thief is aware of the restrictions on the combinations.

What is the probability that a thief will guess the correct combination...

- ...if there are no restrictions? [Answer: $0.0012 = 0.12\%$.]
- ...if no two digits can be the same? [Answer: $0.00793650\cdots \approx 0.79\%$.]
- ...if no two consecutive digits can be the same? [Answer: $0.00203221\cdots \approx 0.20\%$.]

As you can see, the probability is affected, but remains less than 1% in each case..

... 01001001 ...
... 00100000 ...
... 01001100 ...
... 01110101 ...
... 01110110 ...
... 00100000 ...
... 01000110 ...
... 01110011 ...

Now I'd like to talk to you about something extremely serious. Password security is vital to any business that uses computers in any way whatsoever—that is to say, every business in today's world. Have you ever wondered why websites try to encourage or force you to use a password that contains a numeral or a special symbol? Why are they so picky about the structure of your password?

```
... 01001001 ...
... 00100000 ...
... 01001100 ...
... 01110101 ...
... 01110110 ...
... 00100000 ...
... 01000110 ...
... 01110011 ...
```

Of course, your computer does not store your login password. That would be very dangerous. Instead, when you first set the password, it goes into a special kind of function, called a "one-way hash" and the output of that function is stored. When you next want to log in, whatever you type is also put into that same function. If the outputs match, then that means you've typed the right password and you are logged in. If the outputs do not match, then that means you've typed the wrong thing, and you are not permitted to log in. Because the mathematics of typical "one-way hash functions" are complicated, we're going to skip describing their details.

```
... 01001001 ...
... 00100000 ...
... 01001100 ...
... 01110101 ...
... 01110110 ...
... 00100000 ...
... 01000110 ...
... 01110011 ...
```

The list of usernames and the outputs from the "one-way hash function" are stored in a file called the shadow password file. The shadow password file does not contain the passwords of the users, but instead it stores the output of the "one-way hash function" when the password is the input. Every server needs a copy of this file, so that users can log in. That's a lot of servers: the print server, the file server, the backup server, etc...

There are several ways that the password file can end up in the hands of a hacker. One of those servers can be stolen; there could be a disgruntled employee who decides to punish his employer by using the file, or giving it to someone who can use it; when the server is backed up regularly, some of the old backup media might end up in the trash; when the server is upgraded, old hard-drives can end up in the trash; there could be a vulnerability in some software component; one of the members of the IT staff could be bribed or blackmailed, and so forth.

```
... 01001001 ...
... 00100000 ...
... 01001100 ...
... 01110101 ...
... 01110110 ...
... 00100000 ...
... 01000110 ...
... 01110011 ...
```

Now that we know how easy it is for the password file to end up in the hands of hackers, we can fully understand why the password file does not contain the actual passwords. That would be a complete disaster. The use of the "one-way hash function" makes sense now.

So what do hackers do, once they have the file? One approach is to have a computer "try" every possible password, continuously feeding all of them into that "one-way hash function," until they get a match. This is called "the brute force attack."

In the next box, we're going to calculate how much time it will take a hacker to check all possible passwords, for two cases. In the first case, the company forces people to use a digit or a symbol inside their password; in the second case, the company does not force people to use a digit or a symbol, and the hacker is betting that at least one user has a password entirely composed of lower-case and capital letters. For now, let us consider passwords of length 8, which is the typical length.

For Example :

# 7-7-43

With the previous few boxes in mind, we must consider passwords that contain upper and lower case letters, numerals, and symbols (such as punctuation marks) on your computer keyboard—there are 32 of those. We want to know how many passwords of length 8, made up those choices, are possible. Next, if a hacker's computer can check one hundred million possibilities per second (which is about right), then how long will it be to check all possible passwords?

Of course, there are 52 letters of the alphabet when we consider capital letters and lower case letters separately. To this we add the 10 numerals, and the 32 possible symbols. All in all, we have

$$26 + 26 + 10 + 32 = 94$$

possible "things" that can go into any spot of a user's password. Each password has 8 spots in it. Of course, order matters because "`18United`" and "`18Untied`" are different passwords. It is equally obvious that repeats are allowed, because of passwords such as "`Beaver56`" where the letter "e" has been repeated. Since repeats are allowed and order matters, we are to use the exponent principle. We conclude that there are

$$94^8 = 6,095,689,385,410,816 \approx 6.09568 \times 10^{15}$$

possible passwords.

In the next box, we will make sense of this huge number, by converting it into time.

Now we can convert the large number in the previous box, 94^8, into a length of time. Since the hacker's computer can check 10^8 passwords per second, we have

$$\frac{6.09568 \times 10^{15} \text{ passwords}}{10^8 \text{ passwords/second}} = 6.09568 \times 10^7 \text{seconds}$$

but that's hard to comprehend. Instead, that is

$$705.518 \cdots \text{days}$$

just a hair short of two years (730 days excluding leap years).

This is why serious businesses (and other organizations like universities) will force employees to reset their password every year, twice a year, or even every quarter.

Play Around With it...

# 7-7-44

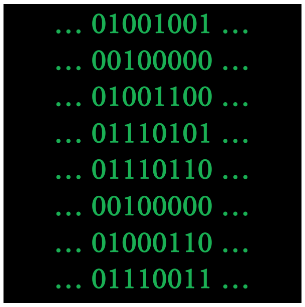
Now repeat the work of the previous box, but imagine a hacker who is only going to try all the passwords that are made entirely from upper-case and lower-case letters.

- How many possible passwords of length 8 are there, in this case?
  [Answer: $52^8 = 53,459,728,531,456 \approx 5.34597 \times 10^{13}$.]
- How many seconds will that require? (Our hacker's computer can try one hundred million passwords per second.)
  [Answer: $5.34597 \times 10^5 = 534,597.\cdots$ seconds.]
- What is that in days, hours, minutes, and seconds?
  [Answer: 6 days, 4 hours, 29 minutes, and 57 seconds.]

Of course, if you agree with me on the number of days, hours, and minutes, but disagree on the number of seconds, then that's just due to rounding error inside of your calculator.

Let's check our conversion of 534,597 seconds into 6 days, 4 hours, 29 minutes, and 57 seconds.

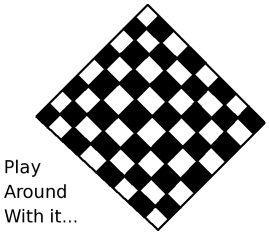
$$\begin{aligned} 6 \times 24 \times 60 \times 60 &= 518,400 \\ 4 \times 60 \times 60 &= 14,400 \\ 29 \times 60 &= 1740 \\ & +57 \\ \hline \text{Total} &= 534,597 \end{aligned}$$



Let's review what we learned in the previous three boxes. If the company requires employees to use passwords that contain numerals and symbols, it will take a hacker slightly more than 705 days, i.e. almost two years, to check every possible password.

Alternatively, if the company does not require employees to use passwords that contain numerals and symbols, then the hacker can "bet" that at least one employee has a password without numerals and without symbols—i.e. a password composed entirely of capital and lower case letters. In this case, the hacker can check every possible password of that form in less than a week.

I'm sure that we can all agree, at this point, that it really does make sense to force people to include a numeral or a symbol in their password.

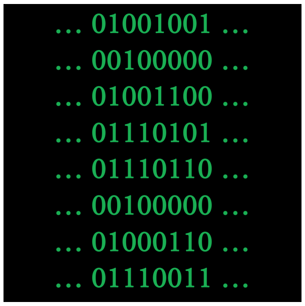


# 7-7-45

Before we continue, let's verify if you understood the previous calculations. Let's consider the possibility of passwords that contain letters (both capital and lower case) and numerals, but not symbols. As before, the passwords are length 8 and our hacker's computer can try one hundred million passwords per second.

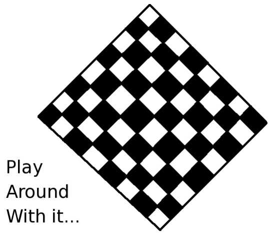
- How many possible passwords of length 8 are there, in this case?
  [Answer: $62^8 = 218,340,105,584,896 \approx 2.18340 \times 10^{14}$.]
- How many seconds will that require?
  [Answer: 2.18340×10^6 seconds.]
- What is that in days, hours, minutes, and seconds?
  [Answer: 25 days, 6 hours, 30 minutes, and 1 second.]

As before, if you agree with me on the number of days, hours, and minutes, but disagree on the number of seconds, then that's just due to rounding error inside of your calculator.



As you already know, passwords for most systems are case-sensitive. That means that typing `NotStr8!`, `Notstr8!`, `notStr8!`, `NOTstr8!`, and `notSTR8!` are five different passwords. In the past, some executives objected to this, because they would frequently get capital and lower-case letters confused. Luckily, that generation is now retiring.

In the next box, we will suppose that your CEO is just old-fashioned and severely dislikes this distinction between capital and lower-case. There will be security-related consequences.

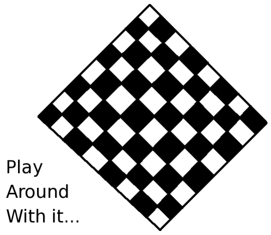


# 7-7-46

Suppose that your CEO dislikes the distinction between capital and lower-case. The CEO proposes that from now on, the password systems should treat capital and lower-case letters identically, and while numerals should be allowed, symbols should not be allowed. We will imagine that a disgruntled IT employee, who knows the rules but not the passwords of the senior leadership, is trying to guess the password for a senior executive.

- Hint: Since the capital letters and lower case letters are treated identically, we have 26 letters and 10 numerals, for a baseline set of 36 things for each spot in the password.
- How many possible passwords of length 8 are there, in this case? [Answer: $36^8 = 2,821,109,907,456 \approx 2.82110 \cdots \times 10^{12}$.]

We will continue in the next box.



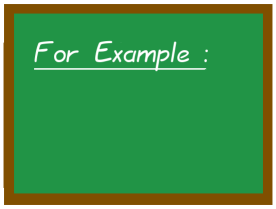
# 7-7-47

Continuing with the previous box. . .

- How many seconds will that require? (Our hacker's computer can try one hundred million passwords per second.) [Answer: $2.82110 \times 10^4 = 28,211.0 \cdots$ seconds.]
- What is that in days, hours, minutes, and seconds? [Answer: 7 hours, 50 minutes, and $11.09 \cdots$ seconds.]

As always, if you agree with me on the number of days, hours, and minutes, but disagree on the number of seconds, then that's just due to rounding error inside of your calculator.

So far, we've gotten a lot of mileage out of using the exponent principle. In particular, we've used the exponent principle to quickly compute answers that would be obtained more slowly with the multiplication principle. In contrast, it is interesting to look at a few problems which can only be solved with the multiplication principle, because they cannot be reduced to the exponent principle (nor any of the principles that are coming in the next few modules).



# 7-7-48

Suppose a college is having a debate about the lowering of the drinking age. There are to be nine speakers, 5 in favor, and 4 against. In how many ways can the ordering of the speakers be made? (You may assume that no one speaks twice.) How about if the speakers must alternate between pro and con? What is the probability that an ordering determined by drawing names out of a hat would result, by coincidence, with the speakers alternating pro and con?

With no restrictions, there are 9 potential speakers for the first position, and 8 for the second position; likewise 7 for the third, 6 for the fourth and so on. This comes to

$$9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 362,880$$

in general, with no guarantee of the speakers alternating.

In the next box, we will examine the case of alternating speakers.

Continuing with the previous box, we have computed the number of possibilities for non-alternating speakers. Now let's compute the number of possibilities for alternating speakers.

Obviously if the speakers alternate, the first speaker must be in favor—there are five choices; the next would be against—there are four choices; the next would be in favor—at this time four choices remain; next comes someone against—three speakers are left; this is followed by someone in favor—three choices exist now; and so forth. This gives us:

$$\underbrace{5}_{\text{pro}} \times \underbrace{4}_{\text{con}} \times \underbrace{4}_{\text{pro}} \times \underbrace{3}_{\text{con}} \times \underbrace{3}_{\text{pro}} \times \underbrace{2}_{\text{con}} \times \underbrace{2}_{\text{pro}} \times \underbrace{1}_{\text{con}} \times \underbrace{1}_{\text{pro}} = 2880$$

possible arrangements.

Last but not least, we should take the ratio of these two numbers. That will give us the probability of alternation happening by coincidence. That's because 2880 schedules of speakers have the alternation out of a possible 362,880. We now obtain

$$\frac{2880}{362,880} = 0.00793650\cdots \approx 0.79\%$$

for the probability that the speakers alternate by coincidence. This is rather unlikely.

Play Around With it...

# 7-7-49

There are to be seven speakers at graduation, 4 women and 3 men.

- How many possible sequences of speakers are there? [Answer: 5040.]
- How many if the speakers must alternate gender? [Answer: 144.]
- What is the probability that a random ordering alternates gender by coincidence? [Answer: $2.85714\cdots\%$.]

For Example :

# 7-7-50

Suppose we want to know how many three-digit odd numbers exist. Let's consider each of the three digits separately. The leftmost digit cannot be a zero, but it can be anything else. After all, if it were a zero, then we'd be looking at a two-digit number, not a three digit number. The middle digit can be anything at all. Last but not least, the rightmost digit must be 1, 3, 5, 7, or 9. Why? Because if it were instead 0, 2, 4, 6, or 8, then we'd have an even number—whereas we want an odd number.

Therefore, there are

$$9 \times 10 \times 5 = 450$$

possible odd three-digit numbers.

Play Around With it...

# 7-7-51

Now consider the previous question:

- How many two-digit odd numbers are there? [Answer: 45.]
- How many five-digit even numbers are there? [Answer: 45,000.]

For Example :

# 7-7-52

Suppose that we are interested in numbers that look like 1551 and 45,654 or even 189,981. These numbers are the same when being read forwards and backwards, and are called *palindromes*. They can be short (like 11 or 787) or long (like 87,055,078 or 1,234,321). How many three-digit palindromes are there? How many four digit-palindromes are there?

Let's consider the three-digit case first. There are nine choices for the first digit, because it cannot be zero. The second digit can be anything, so there are ten choices. The third digit must match the first digit—that means only one possible digit can take the third position. By the multiplication principle, there are

$$9 \times 10 \times 1 = 90$$

possible palindromes that are three digits in length.

Next, let's consider the four-digit case. There are nine choices for the first digit, because it cannot be zero. The second digit can be anything, so there are ten choices. The third digit must match the second digit—that means only one possible digit can take the third position. The fourth digit must match the first digit—that means only one possible digit can take the fourth position. By the multiplication principle, there are

$$9 \times 10 \times 1 \times 1 = 90$$

possible palindromes that are four digits in length.

Play Around With it...

# 7-7-53

- How many four-digit palindromes are there? [Answer: 90.]
- How many four-digit numbers exist? [Answer: 9000.]
- What is the probability that a random four-digit number is a palindrome? [Answer: 1/100.]
- How many five-digit palindromes are there? [Answer: 900.]
- How many five-digit numbers exist? [Answer: 90,000.]
- What is the probability that a random five-digit number is a palindrome? [Answer: 1/100.]

Play Around With it...

# 7-7-54

- How many six-digit palindromes are there? [Answer: 900.]
- How many six-digit numbers exist? [Answer: 900,000.]
- What is the probability that a random six-digit number is a palindrome? [Answer: 1/1000.]
- How many seven-digit palindromes are there? [Answer: 9000.]
- How many seven-digit numbers exist? [Answer: 9,000,000.]
- What is the probability that a random seven-digit number is a palindrome? [Answer: 1/1000.]

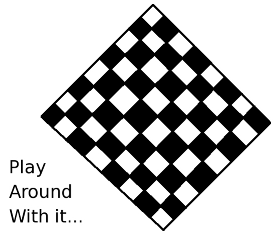
You can also make palindromes out of letters and words. For example,

$$\underbrace{\texttt{DO GEESE SEE GOD}}_{\text{forwards}} \rightarrow \texttt{DOGEESESEEGOD} \rightarrow \underbrace{\texttt{DOG EES ESEEG OD}}_{\text{backwards}}$$

$$\underbrace{\texttt{WAS IT A CAR OR A CAT I SAW}}_{\text{forwards}} \rightarrow \texttt{WASITACARORACATISAW} \rightarrow \underbrace{\texttt{WAS I TAC A RO RAC A TI SAW}}_{\text{backwards}}$$

$$\underbrace{\texttt{A MAN A PLAN A CANAL PANAMA}}_{\text{forwards}} \rightarrow \texttt{AMANAPLANACANALPANAMA} \rightarrow \underbrace{\texttt{AMANAP LANAC A NALP A NAM A}}_{\text{backwards}}$$

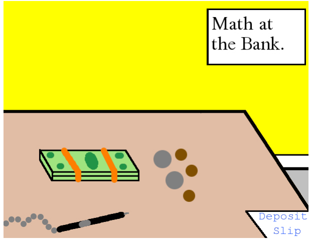
While palindromes made from words (or from numbers) are merely a silly form of recreation, nucleic acids sometimes form palindromes in DNA and RNA. These can lead to the creation of a “hairpin loop” or “stem loop” in the RNA, which in turn can be a building block of larger RNA structures.



# 7-7-55

A company that runs a collection of 24-hour gyms in international airports uses a codeword for each customer, made up of letters of the alphabet. The customers punch their codeword into a machine to access the gym late at night, or on holidays, when no one is on staff. They have 49,107 customers world-wide already, with 160 new customers everyday. If the codewords are four letters, with no repeated letters being allowed, how long will it be before they run out of codewords?

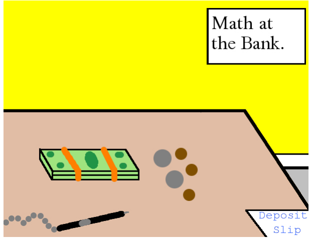
- How many codewords are possible? [Answer: 358,800.]
- How many additional customers can be accommodated? [Answer: 309,693.]
- How many days until that threshold is reached? [Answer: $1935.58\cdots$ days.]
- Using a 365-day year, how many years is that? [Answer: $5.30296\cdots$ years.]



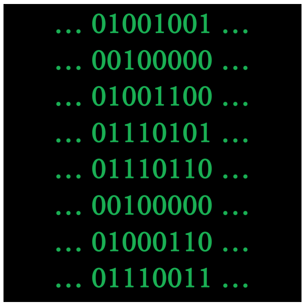
Thinking about the previous box from a business perspective, if all the machines world-wide would have to be reprogrammed to accommodate larger codewords, it might be rather expensive and take a few months. That's why business managers need to keep an eye out for this sort of thing.

On the other hand, they have time, as it will be $5.30296\cdots$ years, or 5 years and 110.58 days (excluding leap years), until this happens.

Let's say that a contract is signed on July 12th, 2013. In most of Europe, including the UK, you would write that date as 12/7/13. That's because they use a day-month-year system of notation. However, in the USA, we almost always would write 7/12/13. That's because we use a month-day-year system of notation. At times, this can be the source of great confusion and frustration. To a European, 7/12/13 indicates the 7th of December, 2013, and not the 12th of July.

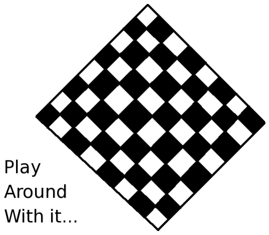


For example, imagine trying to deposit a check in early August, that was signed 7/12/13, and having the check rejected, because the date of the check appears to be the 7th of December, 2013. The wise business executive would therefore write “July 12th, 2013” or use the military style “12 July 2013,” in all cases. That notation removes all doubt, since “July” is written in letters instead of numbers. However, it is important that you, as a business person, be aware of this potential confusion, so that you can resolve it smoothly if a conflict should arise.



To make matters even more confusing, most large-scale computing services run over UNIX, and July 12th, 2013 would be recorded as 2013-07-12. The reason that this notation is used in UNIX is to facilitate sorting. If you pretend that the digits 20130712 are representing the number 20,130,712, then sorting by number is functionally equivalent to sorting by date.

It might take time to see this, but if you consider other date formats, the trick of the previous paragraph doesn’t work. Consider July 12th, 2013 and June 1st, 2014. If I write them as 07-12-2013 and 06-01-2014, treating them as 7,122,013 and as 6,012,014, then 6,012,014 is the smaller number, even though it is later in time. Likewise 08-02-2012 being 8,022,012 is a still larger number, but it is earlier in time than those two other dates. The UNIX trick really is rather clever.

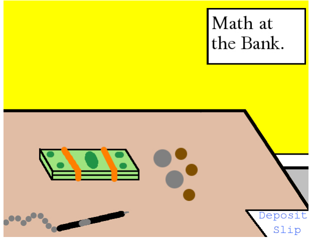


# 7-7-56

As we learned in the previous box, some dates can be confused. For example, 4-2-15 can indicate April 2nd, 2015 to an American, and February 4th, 2015 to a European. Similarly, 5-6-15 can indicate May 6th, 2015 to an American, and June 5th, 2015 to a European. On the other hand, 4-18-15 and 5-19-15 cannot be misinterpreted, because there is no 18th month nor 19th month.

As you can see, a date is “confusable” if the day of the month is 12 or smaller, and “not confusable” if it is 13 or larger. Now answer the following questions:

- If I choose a random day of a 365-day year, what is the probability that the date is confusable? [Answer: $0.394520\cdots \approx 39.45\%$.]
- How about during a leap year? [Answer: $0.393442\cdots \approx 39.34\%$.]



In summary, it is wiser to write

- April 2nd, 2015
- 2-Apr-2015
- Apr 2, 2015

or something similar, because those are totally unambiguous. It is unwise to write 4-2-15 (American notation) or 2-4-15 (European notation), because those can be confused. In fact, over the years, I’ve noticed many senior academic and business leaders, with large numbers of contacts on both sides of the Atlantic, are careful to write their dates in an unmistakable fashion by annotating the month with letters and not a number.

This is the end of our module on the multiplication principle and the exponent principle. I hope that everything has been clear, and that you have also learned something about information security along the way.