

Combinatorics

Key Topics:

- * Sum Rule and Product Rule Review
- * The Pigeonhole Principle
- * Permutations and Combinations
- * Binomial Coefficients and the Binomial Theorem
- * Permutations and Combinations with Repetition
- * Summary of Theorems

Sum Rule and Product Rule Review

The Sum Rule: If there are $n(A)$ ways to do A, and $n(B)$ ways to do B, and these tasks cannot be done at the same time, then the number of ways to do either task is $n(A) + n(B)$. This can be expanded to any number of terms.

The Product Rule: Suppose that a procedure can be broken down into two tasks. If there are $n(A)$ ways to do A, and $n(B)$ ways to do B after the first task has been done, then there are $n(A) * n(B)$ ways to do the procedure. This can be expanded to any number of terms.

There are 18 math majors and 200 CS majors at Stanford. How many ways are there to pick two representatives, so that one is a math major and one is a CS major? How many ways are there to pick one representative who is either a math major or a CS major?

How many different three-letter initials are there (with repetition and without)?

How many different three-letter initials are there that begin with the letter A?

How many bit strings of length 10 begin and end with a 1?

How many strings are there of lowercase letters of length four or less?

You may be wondering why we care about counting at all as computer scientists. Well, take a look at the following:

What is the value of k after the following code has been executed?

```
k = 0;
for (i1 = 1; i1 <= n1; i1++)
    k = k + 1;
for (i2 = 1; i2 <= n2; i2++)
    k = k + 1;
.
.
.
for (im = 1; im <= nm; im++)
    k = k + 1;
```

What is the value of k after the following code has been executed?

```
k = 0;
for (i1 = 1; i1 <= n1; i1++) {
    for (i2 = 1; i2 <= n2; i2++) {
        . . . // m nested loops
        for (im = 1; im <= nm; im++) {
            k = k + 1; // in the innermost loop
        }
    }
}
```

Now that you are all warmed up, we can look at some more complex problems using the sum and product rule together.

Example

In one version of BASIC, the name of a variable is a string of one or two alphanumeric chars, where uppercase and lowercase are not distinguished. So much for meaningful variable names. Incidentally, alphanumeric means either one of the 26 English letters or one of the 10 digits. In addition, all variables must begin with a letter and must be different from the five reserved words. How many different variable names are possible in this (very simplistic) version of BASIC?

Solution: Let V equal the number of different variable names. Let V_1 be the number of variable names one-char long, and V_2 be the number of variable names two-chars long. So, by the Sum Rule, $V = V_1 + V_2$. V_1 must equal 26 since we can't start with a digit. By the Product Rule, $V_2 = 26 * 36$. But five of these strings must be excluded so we get: $V_2 = 26 * 36 - 5 = 931$. $V = V_1 + V_2$ so $V = 26 + 931 = 957$ different names.

The Pigeonhole Principle

Suppose a bunch of pigeons fly into a bunch of pigeonholes to roost. The **Pigeonhole Principle** states that if there are more pigeons than pigeonholes, then there must be at least one pigeonhole with at least two pigeons in it. Seems obvious, and fortunately for us, we can apply this observation to objects besides just pigeons.

The Pigeonhole Principle: If $k+1$ or more objects are placed in k boxes, then there is at least one box containing two or more of the objects.

Proof by Contradiction: Suppose that none of the k boxes has more than one object. Then the total number of objects would be k . This is a contradiction since we have $k+1$ or more objects.

Here are some applications of this principle:

- 1) Among any group of 367 people, there must be at least two with the same birthday since there are only 366 possible birthdays.
- 2) In any group of 27 English words, there must be at least two that start with the same letter, since there are 26 letters in the alphabet.

All of this may seem painfully obvious to you, but this really is a useful tool once we generalize it:

The Generalized Pigeonhole Principle: If N objects are placed in k boxes, then there is at least one box containing at least $\text{ceil}(N/k)$ objects.

Note: The ceiling function $\text{ceil}(x)$ assigns to the real number x the smallest integer that is greater than or equal to x .

Proof by Contradiction: Suppose that none of the k boxes contains more than $\text{ceil}(N/k) - 1$ objects. Then, the total number of objects is at most:

$$k(\text{ceil}(N/k) - 1) < k(((N/k) + 1) - 1) = N$$

Note that we have used the inequality $\text{ceil}(N/k) < (N/k) + 1$. This is a contradiction since there are a total of N objects.

Here are some more applications of the generalized version:

- 1) Among 100 people there are at least $\text{ceil}(100/12) = 9$ who were born in the same month.
- 2) What is the minimum number of students required in a class to be sure that at least six will receive the same grade, given the five possible grades A, B, C, D, NP?

Now that you have the basic idea down, we can look at a more elegant application of this principle. This example illustrates an important area of combinatorics called Ramsey Theory, which deals with the distribution of subsets of elements of sets. The solution to this problem is in your textbook.

Assume that in a group of six people, each pair of individuals consists of two friends or two enemies. Show that there are either three mutual friends or three mutual enemies in the group.

Permutations and Combinations

A **permutation** of a set of objects is an ordering of the objects in a row. For example, the set of elements {a b c} can be written as follows:

abc acb cba bac bca cab

giving us six possible permutations. In general, given a set of n objects, how many permutations does the set have? Imagine forming a permutation as an n -step process:

- step 1: choose an element to put in position 1
- step 2: choose an element to put in position 2
- ...
- step n : choose an element to put in position n

This might look vaguely familiar to our application of the product rule when certain constraints (like no repetition) are applied. There are n ways to do the first step, $n-1$ ways to do the second step (since one element was used in the first step), $n-2$ ways to do the third step, etc. By the time we get to the n th step, there is only one element left. So, by the product rule we get:

$$n * (n-1) * (n-2) * \dots * 2 * 1 = n!$$

For any integer n with $n \geq 1$, the number of permutations of a set with n elements is $n!$.

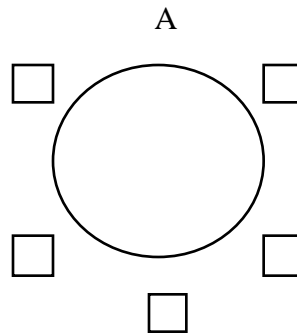
How many ways can the letters in the word "boinga" be arranged in a row?

How many ways can the letters in the word "boinga" be arranged if the letters "bo" must remain next to each other (in order) as a unit?

Now, what if we introduce the idea of ordering objects into a circular arrangement. This adds a little twist to things. Say we have six representatives of six countries at war. An old trick in the diplomacy racket is to use a circular table so there are no ends of the table

which might confer a particular status. How many different ways can we seat these representatives?

To avoid political statements, we will name our representatives A, B, C, D, E and F. Since only relative position matters, start with one of them, say A, and place that person anywhere, say in the top seat in the following diagram.



B through F can be arranged in the seats around A in all possible orders. So there are $5! = 120$ ways to seat the group.

Another twist to this permutation idea is say we want to determine the number of ways to select some number of ordered elements from a set. For example, given the set $\{a \ b \ c\}$, how many ways are there to select two letters from the set and write them in order?

ab ac ba ca cb bc

Each such ordering of 2 elements is called a 2-permutation of the set $\{a \ b \ c\}$.

An r -permutation of a set of n elements is an ordered selection of r elements taken from the set of n elements. The number of r -permutations of a set of n elements is denoted $P(n, r)$. If n and r are integers and $1 \leq r \leq n$, then $P(n, r) = n! / (n - r)!$.

The proof of this formula is fairly straight-forward. Here is the basic idea: Suppose a set of n elements is given. Formation of an r -permutation can be thought of as an r -step process:

step 1: choose an element to be first	(n ways to do this)
step 2: choose an element to be second	($n-1$ ways to do this)
step 3: choose an element to be third	($n-2$ ways to do this)
...	
step r : choose an element to be n th	($n-(r-1)$ or $n-r+1$ ways to do this)

and thus, we have finished forming an r -permutation. It follows from the product rule that the number of ways to form an r -permutation is $n * (n-1) * (n-2) * \dots * (n - r + 1)$.

How do we get $n! / (n - r)!$ from this result?

How many different ways can three of the letters of the word "blurp" be chosen and written in a row?

How many different ways can this be done if "b" must be the first letter?

Combinations

Consider the following:

Suppose 5 members of a group of 12 are to be chosen as a team to work on a project. How many distinct 5-person teams can be selected?

Or in general:

Given a set S with n elements, how many subsets of size r , can be chosen from S ?

Each individual subset of S of size r , is called an **r-combination**.

n and r are non-negative integers with $r \leq n$. An **r-combination** of a set of n elements is a subset of r of the n elements. The symbol

$$\binom{n}{r}$$

$(n \text{ choose } r)$ denotes the number of subsets of size r that can be chosen from the n elements. This is denoted $C(n, r)$.

If we are going to select elements out of a set, we have two methods to apply. We could do an **ordered selection**, where we are interested not only in the elements chosen, but also in the order in which the elements are chosen. You probably noticed that this is the same as our definition of an r -permutation.

Or, we could do an **unordered selection**, where we are interested only in the elements chosen, and we don't care about the order. This is what we mean by an r -combination.

How many unordered selections of 2 elements can be made from $\{0 \ 1 \ 2 \ 3\}$? In other words, what is $C(4, 2)$?

So how do we calculate $C(n, r)$? In order to answer this, we need to analyze the relationship between r -permutations and r -combinations. We'll do this with a simple example:

Write all 2-permutations of the set $\{0 \ 1 \ 2 \ 3\}$. Then, find an equation relating $P(4, 2)$ and $C(4, 2)$.

The reasoning we applied in this example, works in the general case. To form an r -permutation of a set of n elements, first choose a subset of r of the n elements (there are $C(n, r)$ ways to do this). Then, choose an ordering for the r elements (there are $r!$ ways to do this). Then, the number of r -permutations is $P(n, r) = C(n, r) * r!$. If we solve for $C(n, r)$ we get $C(n, r) = P(n, r) / r!$. We know that $P(n, r) = n! / (n - r)!$, so substitution gives us:

$$C(n, r) = n! / (r! * (n - r)!), \text{ assuming } n \text{ and } r \text{ are nonnegative and } r \leq n.$$

Now, we can find the answer to the question that began this section:

Suppose 5 members of a group of 12 are to be chosen as a team to work on a project. How many distinct 5-person teams can be selected?

We need to calculate $C(12, 5) = 12! / (5! * (12 - 5)!)$. The best way to solve this is not by multiplying all the factorials out, even though it's pretty easy to punch these into a calculator. Instead we do this:

$$12 * 11 * 10 * 9 * 8 * 7! / (5 * 4 * 3 * 2 * 1) * 7!$$

The $7!$ terms cancel; the $5 * 2$ in the denominator cancel the 10 on top; the $4 * 3$ in the denominator cancel the 12 on top and we are left with: $11 * 9 * 8 = 792$.

As usual in the combinatorical universe, we can come up with all kinds of special situations. So, let's say that Fred and Ginger (2 of the 12 people in the above example) simply must work together or they will make everyone else's lives miserable. Thus, any team of 5 that we form, must either include both of them or neither of them (the latter being preferable to the other 10 people). Now how many 5-person teams can be formed?

Here is a diagram of the situation:

All possible 5-person teams

teams with both Fred and Ginger	teams with neither Fred nor Ginger
------------------------------------	---------------------------------------

A team with Fred and Ginger contains exactly three other people from the remaining ten. So there are as many such teams as there are 3-person subsets: $C(10, 3) = 120$. The other set of teams is just $C(10, 5) = 252$. Add them together and we get 372 possible teams.

Now suppose Fred and Ginger have a terrible fight, and simply refuse to work on the same team. How many 5-person teams can be formed?

Binomial Coefficients and the Binomial Theorem

Recall that we can denote r-combinations as $C(n, r)$ or:

$$\binom{n}{r}$$

This value (no matter how we denote it) is called a **binomial coefficient**. We use this term because the numbers occur as coefficients in the expansion of powers of binomial expressions such as $(a + b)^n$. There are some important properties of binomial coefficients which we will present. The first comes from a very important mathematician of the 17th century by the name of Pascal:

Pascal's Identity: Let n and k be positive integers with $n \geq k$. Then:

$$C(n+1, k) = C(n, k-1) + C(n, k)$$

How would you prove this?

Pascal's Identity is the basis for an interesting geometric arrangement of the binomial coefficients into a triangle. The n th row of the triangle consists of the binomial coefficients, $C(n, k)$ for k from 0 to n .

$$\begin{array}{ccccccc}
 & & & & C(0,0) & & \\
 & & & C(1,0) & & C(1,1) & \\
 & & C(2,0) & & C(2,1) & & C(2,2) \\
 C(3,0) & & C(3,1) & & C(3,2) & & C(3,3) \\
 \dots & & & & & &
 \end{array}$$

This is known as Pascal's Triangle. Pascal's Identity shows that when two adjacent binomial coefficients are added, we get the one in the next row that lies between them. For

example, $C(2,0) + C(2,1) = C(3,1)$ ($1 + 2 = 3$). This triangle turns out to be a useful little calculator for binomial coefficients.

Another useful identity concerning binomial coefficients:

$$\sum_{k=0}^n C(n, k) = 2^n$$

How would you prove this identity? (this is in your textbook)

The **binomial theorem** gives the coefficients of the expansion of powers of binomial expressions. What is a binomial expression, you ask? It is an expression that is the sum of two terms, like $x + y$ (these terms can be products of constants and variables, but that is not relevant here).

First, let's see why binomial coefficients are even involved in this area. Think about the expansion of $(x + y)^3$. We could just sit down and multiply it all out, or we could be clever little combinatoric wizards and make the following observation: When $(x+y)(x+y)(x+y)$ is expanded all the products of a term in the first sum, a term in the second sum, and a term in the third sum are added. Terms of the form x^3 , x^2y , xy^2 , and y^3 arise. To obtain a term of the form x^3 , an x must be chosen from the three sums and this can be done in only one way. Thus, x^3 must have a coefficient of 1. To obtain a term of x^2y , we need two x 's chosen from two of the three sums (and one y from the remaining sum). The number of such terms must be $C(3,2)$ since it is the number of 2-combinations of 3 objects. To obtain a term of xy^2 , we need one x chosen from one of the three sums (and two y 's from the remaining two sums). The number of such terms must be $C(3,1)$. The reasoning for y^3 is the same as that for x^3 , and we get:

$$(x + y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$$

The Binomial Theorem:

$$(x + y)^n = \sum_{j=0}^n C(n, j) x^{n-j} y^j$$

What is the coefficient of $x^{12}y^{13}$ in the expansion $(x + y)^{25}$?

Permutations and Combinations with Repetition

All the permutation and combination problems we have seen thus far did not have any repeated elements. This is a serious constraint since in many counting problems, elements may be used repeatedly. For example, if we could not re-use letters and numbers on license plates, we would have much less of a traffic problem than we do now.

Let's start with permutations:

How many strings of length n can be formed from the English alphabet? By the product rule, since there are 26 letters and since each letter can be used repeatedly, we see there are 26^n strings of length n . This simple example can be generalized to a formula:

The number of r -permutations of a set of n objects with repetition allowed is n^r .

As for combinations, consider the following example:

How many ways are there to select four pieces of fruit from a bowl containing apples, oranges, and pears if the order in which the fruit is selected does not matter, only the type of fruit and not the individual piece matters, and there are at least four of each type of fruit in the bowl.

Well, one way to solve this is to just go diving for fruit:

4 apples	4 oranges	4 pears
3 apples, 1 orange	3 apples, 1 pear	3 oranges, 1 apple
3 oranges, 1 pear	3 pears, 1 apple	3 pears, 1 orange
2 apples, 2 oranges	2 apples, 2 pears	2 oranges, 2 pears
2 apples, 1 orange, 1 pear	2 oranges, 1 apple, 1 pear	2 pears, 1 apple, 1 orange

There are 15 ways, and it turns out that the solution is actually the number of 4-combinations with repetition, allowed from a three-element set {apple, pear, orange}.

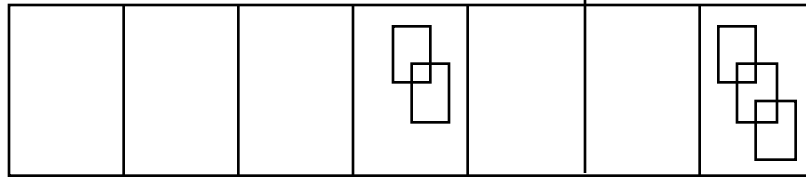
How can we generalize this to come up with a formula? Another example will show us the way:

How many ways are there to select five bills from a money bag containing \$1, \$2, \$5, \$10, \$20, \$50, and \$100 bills. Assume that the order in which the bills are chosen does not matter, that the bills of each denomination are indistinguishable (unfortunately), and that there are at least five bills of each type.

No way we are going to enumerate all the possibilities... What we really want is to count the 5-combinations with repetition allowed from a 7-element set. Suppose we have a cash box with seven compartments, one for each type of bill:

--	--	--	--	--	--	--

As we choose the five bills, we place them in the appropriate compartment. A shorthand for these compartments would be to have six bars (|) for the dividers and 5 stars (*) for the bill selection. So for example, two \$10's and three \$1's look like this in the compartments:



and like this in our shorthand: |||**|||***. One \$100, one \$50, two \$20's and one \$5 is represented as: *|*|**||*|. The number of ways to select the five bills corresponds to the number of ways to arrange six bars and five stars. Or, in other words, it's the number of ways to select the positions of five stars from eleven possible positions. This corresponds to the number of unordered selections of five objects from a set of 11 objects of $C(11, 5)$. The following theorem generalizes these ideas.

There are $C(n+r-1, r)$ r -combinations from a set with n elements when repetition of elements is allowed.

How would you prove this theorem?

What is the value of k after the following code has been executed?

```
k = 0;
for (i1 = 1; i1 <= n; i1++) {
    for (i2 = 1; i2 <= i1; i2++) {
        ...
        for (i_m = 1; i_m <= i_{m-1}; i_m++) {
            k = k + 1;
        }
    }
}
```

Summary of Theorems

The Sum Rule: If there are $n(A)$ ways to do A, and $n(B)$ ways to do B, and these tasks cannot be done at the same time, then the number of ways to do either task is $n(A) + n(B)$. This can be expanded to any number of terms.

The Product Rule: Suppose that a procedure can be broken down into two tasks. If there are $n(A)$ ways to do A, and $n(B)$ ways to do B after the first task has been done, then there are $n(A) * n(B)$ ways to do the procedure. This can be expanded to any number of terms.

The Pigeonhole Principle: If $k+1$ or more objects are placed in k boxes, then there is at least one box containing two or more of the objects.

The Generalized Pigeonhole Principle: If N objects are placed in k boxes, then there is at least one box containing at least $\text{ceil}(N/k)$ objects.

For any integer n with $n \geq 1$, the number of permutations of a set with n elements is $n!$.

An r -permutation of a set of n elements is an ordered selection of r elements taken from the set of n elements. The number of r -permutations of a set of n elements is denoted $P(n, r)$. If n and r are integers and $1 \leq r \leq n$, then $P(n, r) = n! / (n - r)!$.

n and r are non-negative integers with $r \leq n$. An **r -combination** of a set of n elements is a subset of r of the n elements. The symbol

$$\binom{n}{r}$$

$(n \text{ choose } r)$ denotes the number of subsets of size r that can be chosen from the n elements. Sometimes this is denoted $C(n, r)$.

$$C(n, r) = n! / (r! * (n - r)!), \text{ assuming } n \text{ and } r \text{ are nonnegative and } r \leq n.$$

Pascal's Identity: Let n and k be positive integers with $n \geq k$. Then:

$$C(n+1, k) = C(n, k-1) + C(n, k)$$

Another useful identity concerning binomial coefficients:

$$\sum_{k=0}^n C(n, k) = 2^n$$

The Binomial Theorem:

$$(x + y)^n = \sum_{j=0}^n C(n, j) x^{n-j} y^j$$

The number of r-permutations of a set of n objects with repetition allowed is n^r .

There are $C(n+r-1, r)$ r-combinations from a set with n elements when repetition of elements is allowed.

Bibliography

D. Cohen, *Basic Techniques of Combinatorial Theory*, New York: Wiley, 1978.

A. Tucker, *Applied Combinatorics*, New York: Wiley, 1985.

Historical Notes

Ever since the early days of mathematical reasoning (back with Plato and Aristotle), a common definition of mathematics was: “the science of quantity”. But not until the 17th century do we find a formal treatment of combinatorics. Much of the foundation for this theory (as well as probability) was laid by Blaise Pascal (1623-1662). It was while working in the area of probability that Pascal defined the Pascal triangle. In 1654, Pascal abandoned his mathematical pursuits to devote himself to theology. As the story goes, he returned to mathematics just once one night when he had a terrible toothache. He sought comfort by studying and documenting the mathematical properties of the cycloid. Miraculously, his toothache disappeared which he took as a divine sign of approval for his interest in mathematics.

G. Lejeune Dirichlet (1805-1859) was another important contributor to the study of combinatorics. He was one who defined the Pigeonhole Principle. Finally, Frank Ramsey who developed “Ramsey Theory” also made important contributions to combinatoric theory, as well as the mathematical theory of economics. Unfortunately, he died at the age of 26.