CS/Math 240: Introduction to Discrete Mathematics

Reading 14 : Counting

Author: Dieter van Melkebeek (updates by Beck Hasti and Gautam Prakriya)

In this reading we discuss counting. Often, we are interested in the cardinality of some finite set. For example in discrete probability theory, which we won't get to in this course, we want to count the size of the set of "good" outcomes and then related to the size of the set of all possible outcomes.

We begin with some basic counting techniques which we illustrate on multiple examples. After that, we generalize some of the basic techniques and give examples of their applications.

14.1 Basic Counting Techniques

We describe some basic counting rules and apply them in nontrivial ways.

14.1.1 Bijection Rule

We saw in the reading on relations that if a function $f : T \to S$ is a total bijection, then $|S| = |T|$. Thus, in order to find the cardinality of S, we can, instead, find a bijection from T to S and find the cardinality of T.

We often use the set of sequences over some alphabet that satisfy some property as the set T.

14.1.2 Sum Rule

Suppose S is the *disjoint union* of sets $S_1, S_2, \ldots, S_s$. Being a disjoint union means that the sets $S_1, S_2, \ldots, S_s$ form a partition of S. When the union is disjoint, we use the symbol $\dot\cup$ instead of $\cup$, and we write $S = \dot\bigcup_{i=1}^{s} S_i$. With this notation at hand, we see that

$$|S| = \sum_{i=1}^{s} |S_i| . \tag{14.1}$$

14.1.3 Product Rule

Now suppose S is the Cartesian product of sets $S_1, S_2, \ldots, S_s$, i.e. $S = S_1 \times S_2 \ldots \times S_n$. Then every element of S is an ordered sequence of elements $s_1, s_2, \ldots, s_s$ where $s_i \in S_i$. Since the choices for each position in the sequence are independent, we have

$$|S| = |S_1| \cdot |S_2| \ldots |S_n| . \tag{14.2}$$

That is the number of elements in S is the product of the number of elements in each set.

14.1.4 Examples of Basic Counting Techniques

Through examples, we now show some applications of the three techniques we described.

Example 14.1: You are in a store that sells five different kinds of bagels: plain, poppy seed, sesame seed, onion, and with all three toppings. You want to buy a dozen bagels and can combine the different kinds in any way you want. You want to know how many different combinations of bagels you can pick.

Let S be the set of all possible combinations of 12 bagels, and let T be the set of binary strings of length 16 that contain exactly four 1s. We demonstrate a bijection from S to T in a moment. First let's see it on an example. Suppose you buy 2 plain bagels, 0 poppy seed bagels, 6 sesame seed bagels, 3 onion bagels, and 1 bagel with all toppings. These correspond to strings of zeros of lengths 2, 0, 6, 3, and 1, respectively. Now put ones between these strings of zeros and concatenate everything to get the string 0011000000100010. The number of zeros between two consecutive ones (or between the beginning of the string and the first one, or between the last one and the end of the string) indicates the number of bagels of one kind.

In general, suppose we buy x_1 plain bagels, x_2 poppy seed bagels, x_3 sesame seed bagels, x_4 onion bagels, and x_5 bagels with everything. Let's denote this choice using the tuple $(x_1, x_2, \ldots, x_5)$. Note that $\sum_{i=1}^{5} x_i = 12$. We map this choice of bagels to the string $0^{x_1}10^{x_2}10^{x_3}10^{x_4}10^{x_5}$, i.e, the function f is defined by $f(x_1, \ldots, x_5) = 0^{x_1}10^{x_2}10^{x_3}10^{x_4}10^{x_5}$. This map is a total function because we can map any choice of bagels to a string of length 16 using the strategy above. To see that this map is surjective, consider any string of the form $0^{x_1}10^{x_2}10^{x_3}10^{x_4}10^{x_5}$ of length 16 with four ones. This string corresponds to the choice of x_1 plain bagels, x_2 poppy seed bagels, and so on. Finally, the map is injective. To see this, suppose $f(x_1, \ldots, x_5) = f(y_1, \ldots, y_5)$, so $0^{x_1}10^{x_2}10^{x_3}10^{x_4}10^{x_5} = 0^{y_1}10^{y_2}10^{y_3}10^{y_4}10^{y_5}$. But then $x_i = y_i$ for all i.

Thus, f is a bijection, and it follows that the number of ways we can pick 12 bagels is the same as the number of binary strings of length 16 with exactly four ones by the bijection rule. We will show later in this reading that this number is 1820. ⊠

Example 14.2: Now let's count the number of subsets of a domain D of size n. Let's call this set S. We give a bijection from the set T of all binary strings of length n to the set S.

Fix an enumeration $d_1, d_2, \ldots, d_n$ of D, and construct the map $f : T \to S$ by

$$f(x_1 x_2 \ldots x_n) = \{d_i \mid x_i = 1\}.$$

This is a total bijection because we map every binary string to some subset of D. It is surjective because a subset A is the image of the string $x_1 x_2 \ldots x_i$ where $x_i = 1$ if $d_i \in S$ and $x_i = 0$ otherwise. The function is injective too. Thus, $|S| = |T|$.

Because $T = \{0,1\}^n = \{0,1\} \times \{0,1\} \times \cdots \times \{0,1\}$, i.e., T is the n-fold Cartesian product of the set $\{0,1\}$ with itself,. it is possible to find $|T|$ using the product rule. Since we can pick each position in a binary string of length n independently of all the other positions, and there are two choices for each position, we get $|T| = 2^n$, and this is also the number of subsets of D. ⊠

Example 14.3: Say some website requires users to have passwords that have between 6 and 8 characters. Furthermore, the first character should be a letter, either uppercase or lowercase, and the remaining characters can be uppercase or lowercase letters or one of the digits 0 through 9.

Let F be the set of all possible characters for the first character in the password, and let N be the set of all possible characters for the symbols that follow. We have $|F| = 52$ and $|N| = 62$.

Let P_i be the set of all possible passwords of length i, and note that the set of all possible passwords is the disjoint union $P_6 \dot\cup P_7 \dot\cup P_8$. Thus, we can use the union rule to count the number of possible passwords.

Before we apply the union rule, we need to determine $|P_i|$. For that, we use the product rule. The choices of the individual characters are independent, and we have $P_i = F \times N^{i-1}$ where N^{i-1} stands for the $(i-1)$-fold Cartesian product of N with itself. Thus, we have $|P_i| = |F||N|^{i-1}$.

Finally, the set of all possible passwords has cardinality

$$|P_6| + |P_7| + |P_8| = 52 \cdot 62^5 + 52 \cdot 62^6 + 52 \cdot 62^7 = 186125210680448,$$

which is roughly $1.9 \cdot 10^{14}$. ⊠

14.2 Generalized Counting Techniques

Now that we have seen some basic applications of counting, it's time to look at generalizations of the techniques we have used so far.

14.2.1 Generalized Product Rule

In the applications of the product rule we have seen, we assumed that the choices for each component of a Cartesian product were independent of each other. We now relax this independence condition.

Let S be the set of sequences of length k. Suppose there are n_1 choices for the first term of the sequence. After the first term is fixed, there are n_2 choices for the second term. After the first two terms are fixed, there are n_3 choices for the third term. This continues until the end of the sequence, where n_k gives the number of choices for the k-th term given that the first $k-1$ terms have been fixed. In this case,

$$|S| = \prod_{i=1}^{i} n_i. \tag{14.3}$$

Example 14.4: Given a chess board (a board with 8 rows and 8 columns), we want to find the number of ways to place three different pieces, say a pawn, a bishop, and a queen, so that no two pieces are in the same row, and no two pieces are in the same column.

Let S be the number of ways to place the pieces according to our rules, and let T be the set of sequences $(r_p, c_p, r_b, c_b, r_q, c_q) \in \{1, \ldots, 8\}^6$ such that r_p, r_b and r_q are all different, and also c_p, c_b and c_q are all different. The meanings of the six coordinates are the row and column of the pawn, the row and column of the bishop, and the row and column of the queen. This gives us a bijection between S and T.

We have 8 options for the values of r_p and c_p, so we have $n_1 = n_2 = 8$ in the sense of (14.3). Now we have the constraint $r_b \neq r_p$, so the number of options for r_b once r_p and c_p were picked is 7. Similarly, there are 7 options for the value of c_b, so $n_3 = n_4 = 7$. Finally, $r_q \notin \{r_p, r_b\}$ and $c_q \notin \{c_p, c_b\}$, so $n_5 = n_6 = 6$. Finally, using the generalized product rule yields $|S| = |T| = 8 \cdot 8 \cdot 7 \cdot 7 \cdot 6 \cdot 6 = 112896$. ⊠

The next example has a more mathematical flavor to it. We need the following definition.

Definition 14.1. *A* permutation *of a domain D is a sequence π consisting of elements of D such that each element of D appears exactly once in it.*

Example 14.5: We count the number of permutations of a domain D with $|D| = n$. Let S be the set of all permutations of D. Since each element of D appears exactly once in the permutation, the sequence that represents a permutation has length n.

Once we pick the first term in the sequence, we have $n-1$ options for the second term. After that, two terms of the sequence have been chosen. Since those two terms are different and each element of D can only appear once in the permutation, there are $n-2$ options for the third term. This continues until the very last term. In that case, $n-1$ terms have been determined and are all distinct, so there is exactly one choice for the last term. Thus, by the generalized product rule, we have $|S| = n(n-1)(n-2)\cdots 2 \cdot 1$. ⊠

The quantity from Example 14.5 is widely used, so we give it a name.

Definition 14.2. *The* factorial *of n, denoted $n!$, and read "n factorial", is the product of the first n positive integers. That is,*

$$n! = \prod_{i=1}^{n} i.$$

For convenience, we also define $0! = 1$.

If you consider Example 14.5, defining $0! = 1$ makes sense. There is exactly one way to write down the list of the elements of an empty set. Just write down the empty sequence.

Let's derive some bounds for the value of $n!$. First, $n!$ is the product of n integers, each of which is at most n, so $n! \leq n^n = 2^{n \log n}$ (where the logarithm is base 2). Second, the terms $n, n-1, \ldots, n/2+1, \ldots, n - \lfloor n/2 \rfloor + 1$ are all at least $n/2$, and there are at least $n/2$ of them, so $n! \geq (n/2)^{n/2} = 2^{(n/2)(\log n - 1)}$.

The estimates from the previous paragraph are not exactly tight. There is a much better approximation of $n!$ called *Stirling's approximation* which says that

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n. \tag{14.4}$$

Recall that $\sim$ means asymptotic equivalence. We do not prove (14.4) in this course.

14.2.2 Generalized Bijection Rule

Now let's generalize the bijection rule. Suppose $f : T \to S$ is a total function that onto and k-to-1. (A function f is k-to-1 if for each $s \in S$, there are exactly k elements $t \in T$ such that $f(t) = s$). Then $|S| = |T|/k$.

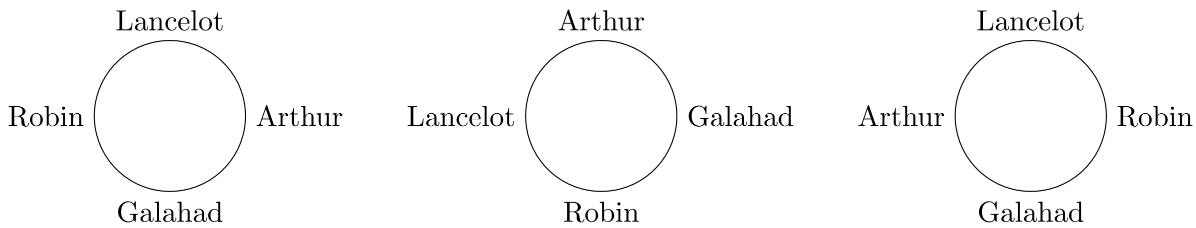
Example 14.6: Consider a chess board. We want to place two pawns on this chess board so that they are in different rows and different columns. How many ways can we do this? One would be tempted to use the approach we took in Example 14.4 without change. Unfortunately, this does not work. Unlike different chess pieces, two pawns are indistinguishable. In other words, placing the first pawn on square a and the second pawn on square b yields the same configuration as placing the first pawn on square b and the second pawn on square a. But we only need a slight modification of the approach from Example 14.4 to get the right answer.

Let S be the set of all possible configurations we can obtain by placing two pawns on the chess board so that they are in different rows and different columns. Consider the set of strings $T = \{(r_1, c_1, r_2, c_2) \in \{1, \ldots, 8\}^4 \mid r_1 \neq r_2 \wedge c_1 \neq c_2\}$. This is an onto function that is 2-to-1 because (a, b, c, d) and (c, d, a, b) define the same configuration, and no other 4-tuple describes the same configuration as those two. Thus, $|S| = |T|/2$ by the generalized bijection rule.

Finally, we find $|T|$ the same way as in Example 14.4 and get $|S| = 8 \cdot 8 \cdot 7 \cdot 7/2 = 1568$. ⊠

Example 14.7: At king Arthur's court, people are seated at a round table. Say there are n people seated at a table. Two seating arrangements are different if there is some person who has a different neighbor on his right (i.e., in the counterclockwise direction) in the two arrangements. For example, the arrangement in Figure 14.1b is the same as the arrangement in Figure 14.1a. On the other hand, the assignment in Figure 14.1c is different from the other two seating assignments in Figure 14.1.

Let S be the set of all possible seating arrangements of the n people, and let T be the set of all permutations of those n people. Let $\pi \in T$, and fix one chair at the table. Define $f(\pi)$ to be the seating arrangement obtained by placing the first person in π on the fixed chair, the second person in π to the first person's right, and keep going around the table. For example, if we fix the leftmost



(a) One seating arrangement.

(b) Same seating arrangement as the first one.

(c) This seating arrangement is different from the previous two. For example, the right neighbor of Galahad is Robin, whereas his right neighbor in the other two arrangements is Arthur.

Figure 14.1: Three ways of seating four people at a round table. The first two are equivalent seating arrangements, and the third one is different from the previous two.

seat (the one where Arthur is) in Figure 14.1a and take the permutation (Arthur, Lancelot, Robin, Galahad), we obtain the seating arrangement in Figure 14.1a.

We claim that the map is n-to-1. Let the people be p_1 through p_n, and assume without loss of generality that $\pi = (p_1, p_2, \ldots, p_n)$. Then person p_i has person p_{i+1} on his right for $i \in \{1, \ldots, n-1\}$, and person p_n has person p_1 on his right. Let's see how many other permutations yield this seating assignment. Once person p_1 is seated, the seats for the other people are determined by the rules for who the neighbor of p_i is. Since there are n locations where we can seat person p_1, there are n permutations of the people that produce the same seating arrangement as the permutation $(p_1, p_2, \ldots, p_n)$ (the other $n-1$ permutations are cyclic shifts of this permutation).

Also note that our function f is surjective because given an arrangement, we can just list an arbitrary person as the first element of the permutation and then go around the table in a counterclockwise direction and add people to subsequent terms in the permutation in the same order in which we see them when we go around the table.

Thus, f is an n-to-1 function from T to S. Since $T = n!$, the generalized bijection rule implies $|S| = |T|/n = n!/n = (n-1)!$. ⊠

Our last example finally solves the counting problem of Example 14.1.

Example 14.8: Consider a domain D with n elements. We would like to know how many s-element subsets D has.

For example, the 2-element subsets of the set $D = \{a, b, c, d\}$ are the six sets $\{a, b\}$, $\{a, c\}$, $\{a, d\}$, $\{b, c\}$, $\{b, d\}$, and $\{c, d\}$.

Let's reduce the problem to the problem of counting sequences of a certain type. Let S be the set of s-element subsets of D, and let T be the set of all permutations of the domain D. Define $f : T \to S$ by setting $f(\pi)$ to be the set of the first s terms in π. Since all terms in a permutation are different, $f(\pi)$ contains s elements.

This function is onto because to get π such that $f(\pi) = A$, we just list the s elements of A in an arbitrary order as the first s elements of π, and then list the remaining elements of the domain in the last $n-s$ positions in an arbitrary order. Permuting the first s elements of π doesn't change the function value, and neither does permuting the last $n-s$ elements. There are $s!$ ways to permute the first s elements, and $(n-s)!$ ways to permute the last $n-s$ elements. Furthermore, we can pick those two permutations independently, so the product rule applies, and we get that f

is $s!(n-s)!$-to-1. Since there are $n!$ permutations of a set of n elements, we get

$$|S| = \frac{n!}{s!(n-s)!}. \tag{14.5}$$

$\boxtimes$

The result of example 14.8 is also a very commonly used quantity, and is called a **binomial coefficient**. We denote the right-hand side of (14.5) by $\binom{n}{s}$ and read it as "n choose s".

Let's return to the situation of Example 14.1. There, we wanted to count the number of strings of length 16 that contain exactly four ones. Note that the positions of ones in such a string in a one-to-one correspondence with the four-elements subsets of $\{1, \ldots, 16\}$. Thus, the number of such strings is $\binom{16}{4} = 1820$, and this is also the number of ways we can pick a dozen bagels that come in five different varieties.

14.2.3 Generalized Sum Rule

The basic sum rule tells us that when we have sets $S_1, S_2, \ldots, S_s$ that are pairwise disjoint.

$$\left|\bigcup_{i=1}^{s} S_i\right| = \sum_{i=1}^{s} |S_i| .$$

This rule fails when there are non-empty intersections. We saw this in the second homework, where you were asked to prove that

$$|S_1 \cup S_2| = |S_1| + |S_2| - |S_1 \cap S_2| \tag{14.6}$$

$$\tag{14.7}$$

This equality is called the *inclusion-exclusion* formula for two sets.
For three sets, we have:

$$|S_1 \cup S_2 \cup S_3| = |S_1| + |S_2| + |S_3| - |S_1 \cap S_2| - |S_1 \cap S_3| - |S_2 \cap S_3| + |S_1 \cap S_2 \cap S_3| \tag{14.8}$$

You can check that this is true using Venn-diagrams. One can generalize this to an arbitrary number of sets sets, but we will not do so in this course.

14.3 Permutations and Combinations

We summarize some useful consequences of the basic counting principles. But first we introduce some terms.

Definition 14.3.
- *An r-permutation of n elements, is an ordered arrangement of r of the n elements.*
- *An r-combination of n elements is a set made up of r of the n elements.*

Example 14.9: Consider the set $S = \{A, B, C\}$. The set of 2-permutations of the elements of S is $\{AB, BA, AC, CA, BC, CB\}$. The set of 2-combinations of S is $\{\{A, B\}, \{A, C\}, \{B, C\}\}$. ⊠

The number of r-combinations of a set of n elements is denoted by $C(n, r)$ or $\binom{n}{r}$. We saw in example 14.8 that

$$C(n, r) = \frac{n!}{r!(n-r)!}.$$

This is the number of ways to pick r items from a set of n items.

To count the number of r-permutations of n elements ($P(n, r)$), note that an r-permutation can be obtained by first picking r of the n elements, (i.e. an r-combination) and then picking an ordering of the r elements picked.

The number of ways to pick r of the n elements is $\binom{n}{r}$ and the number of ways to pick an ordering of these elements is $r!$. So by the product rule,

$$P(n, r) = \binom{n}{r} \times r! = \frac{n!r!}{r!(n-r)!} = \frac{n!}{(n-r)!}.$$

Below is a basic identity involving binomial coefficients: We give two proofs of the identity.

Proposition 14.4. *For all integers n and k, $\binom{n}{k} = \binom{n}{n-k}$.*

First proof. By definition, $\binom{n}{k} = \frac{n!}{(n-k)! \cdot k!}$, and $\binom{n}{n-k} = \frac{n!}{(n-(n-k))!(n-k)!} = \frac{n!}{k! \cdot (n-k)!}$. We see that the expressions for the two binomial coefficients are the same. □

Second proof. Let S be the set of all k-element subsets of a set of n elements. We can describe this set by saying which k elements belong to it, and also by saying which $n - k$ elements are the ones that do not belong to it. For example, if $n = 5$ and $k = 2$, the statements "elements 1 and 3 belong to the subset" and "elements 2, 4 and 5 do not belong to the subset" describe the same subset of $\{1, \ldots, 5\}$.

We have $\binom{n}{k}$ ways to pick the k elements that are in the subset, and we have $\binom{n}{n-k}$ ways to pick the $n - k$ elements that are not in the subset. This means that $\binom{n}{k} = \binom{n}{n-k}$. □

14.3.1 Example: Counting Poker Hands

Let's get some more practice with our counting rules. Consider a deck of cards that's used for playing poker. The cards have 4 different suits: spades, clubs, hearts, and diamonds. For each suit, there are 13 cards, one for each of the values 2 through 10, jack (J), queen (Q), king (K), and ace (A). This means that there are a total of $4 \cdot 13 = 52$ cards.

A poker hand consists of 5 of those 52 cards. Since there is only one card for each suit-value pair, the dealer gives a player those 5 cards by drawing from the deck without replacement. The order in which the cards are drawn does not matter, so there are $\binom{52}{5} = 2598960$ different possible hands.

Example 14.10: A *four-of-a-kind* is a hand with four cards with the same value, one for each suit, and an additional card.

There are 13 ways to pick which value will be present in all suits. Since order doesn't matter, this fully determines four of the five cards in the hand. We can pick the last card any way we want. Bu the generalized product rule, there are now 12 choices for the value and 4 values for the suit. Thus, the total number of four-of-a-kind hands is $13 \cdot 12 \cdot 4 = 624$. ⊠

We are not going to discuss discrete probability in this course, but let's at least mention it now. The discrete probability of an event with a certain property happening is the number of events that have the property divided by the total number of possible events. There are 624 events (hands) of the type four-of-a-kind, and there are a total of 2598960 possible events (hands). Thus, the probability of a four-of-a-kind is $624/2598960 \approx 0.0002$ (where the symbol $\approx$ means "approximately equal to"). This says that about 1 out of 5000 hands is a four-of-a-kind.

Example 14.11: A *full house* consists of three cards with the same value and different suits and two other cards with the same value (that is different from the first one) and two different suits. We use the generalized product rule to count the number of full houses.

There are 13 ways to pick the value shared by 3 cards. The 3 cards can each have 1 possible suit out of 4 and the suits are drawn with replacement, which means there are $\binom{4}{3} = 4$ ways to pick the suits for the three cards. Another way to think about this is that we pick which suit is not present among the three cards, and there are four ways of picking that suit. Thus, the 3 cards can be picked $13 \cdot 4 = 52$ ways.

The two remaining cards must have the same value, and there are 12 options for the value because we cannot pick the value the three other cards have. The two cards will have different suits, and we can pick the two suits $\binom{4}{2} = 6$ ways. Thus, given the first three cards, there are $12 \cdot 6 = 72$ ways to pick the last two cards, and it follows that the total number of different full houses is $52 \cdot 72 = 3744$. ⊠

Example 14.12: Now let's count the number of hands with *two pairs*. That is, there are two cards with the same value and different suit, two other cards with the same value (but different from the first two cards) and different suit, and one additional card.

We can pick the value for the first pair 13 ways, and the suit $\binom{4}{2}$ ways. When the first pair is picked, we can pick the value for the second pair 12 ways, and we have $\binom{4}{2}$ options for the suits of the two cards with that value. Finally, there are 11 values to choose from for the last card, and any one out of the four suits is fair game. This gives a total of $13 \cdot \binom{4}{2} \cdot 12 \cdot \binom{4}{2} \cdot 11 \cdot 4 = 247104$.

But we are overcounting. For example, suppose we pick the hand $4\heartsuit, 4\diamondsuit, A\spadesuit, A\heartsuit, J\clubsuit$. That is, the first pair is a pair of fours (hearts and diamonds), the second pair is a pair of aces (spades and hearts), and the last card is a jack of clubs. But we could pick the aces first and the fours second. That is, the hand $A\spadesuit, A\heartsuit, 4\heartsuit, 4\diamondsuit, J\clubsuit$ is the same hand, but we count it as a separate hand. The order in which we pick the pairs does not matter, so there is a 2-to-1 correspondence between the ways of picking a the hand using our method and the set of hands that consist of two pairs. By the generalized bijection rule, this means that the number of hands that have two pairs is actually $247104/2 = 123552$.

Note that this was not an issue when counting the number of possible ways to pick a full house, because $3 \neq 2$. So we didn't overcount there. ⊠

There is another approach one can use to count the number of hands with two pairs. First, pick the two values for the pairs. This is done by picking 2 out of the 13 possible values, and there are $\binom{13}{2}$ ways for this. Now that the values for the pairs have been picked, we can pick the suits for each of the pairs. The suits for each pair are independent of each other, and there are $\binom{4}{2}$ ways to pick the suits for each pair. Finally, we can pick the remaining card $11 \cdot 4$ ways like before. Thus, the total number of ways to pick a hand with two pairs is $\binom{13}{2} \cdot \binom{4}{2}^2 \cdot 11 \cdot 4$. This looks exactly the same as our expression in Example 14.12, except the division by two is now accounted for by the binomial coefficient $\binom{13}{2} = 13 \cdot 12/2$.

Example 14.13: Now let's discuss a hand that is mostly not useful in poker. Suppose we have at least one card from each suit.

We can pick one value of each suit, and the choices are independent. Thus, we can do this 13^4 ways by the product rule. Afterwards, we can pick one suit, and for each suit there are 12 cards left to choose from, which gives us 48 options for the last card. Using the generalized product rule, we multiply the two values together to get $13^4 \cdot 48 = 1370928$.

But now we are overcounting. For example, suppose we pick the hand $2\spadesuit, 3\clubsuit, 4\heartsuit, 4\diamondsuit, 2\heartsuit$. We get the same hand if we pick $2\spadesuit, 3\clubsuit, 2\heartsuit, 4\diamondsuit, 4\heartsuit$, that is, we pick the 2 of hearts as the one heart card and the 4 of hearts as the additional fifth card, instead of picking them the other way around. There are no other ways to get this hand, so there is a 2-to-1 mapping between the ways to pick hands and the actual hands, which means we need to divide the number we obtained in the previous paragraph by 2. It follows that the number of hands that contain at least one card of every suit is $1370928/2 = 685464$. ⊠

14.4 Binomial Theorem

Now we look at a basic theorem that has many applications in counting. Let's start with a motivating example.

Example 14.14: We want to evaluate $(x+y)^4$ by turning it into a sum of products of different powers of x and y. That is, we want to turn $(x+y)^4 = (x+y)(x+y)(x+y)(x+y)$ into a sum of the form $a_4x^4 + a_3x^3y + a_2x^2y^2 + a_1xy^3 + a_0y^4$. We can evaluate the product as

$$
\begin{aligned}
(x+y)^4 &= xxxx + \\
&+ xxxy + xxyx + xyxx + yxxx + \\
&+ xxyy + xyxy + xyyx + yxxy + yxyx + yyxx + \\
&+ xyyy + yxyy + yyxy + yyyx + \\
&+ yyyy \\
&= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4
\end{aligned}
$$

For example, observe that the terms $xxxy$ and $xyxx$ are the same, and both contribute to the term with x^3y in it. We get a contribution to the term x^3y if we pick x in three out of the four terms in the product $(x+y)(x+y)(x+y)(x+y)$, and y from the remaining term. Thus, there are $\binom{4}{3} = 4$ ways to get a contribution towards the term with x^3y. Similarly, to find the coefficient in front of the term x^2y^2 we count the number of ways we can pick x and y from the four terms so that both x and y are picked twice. This number is $\binom{4}{2} = 6$ because once we choose which terms to pick x from, the terms which we pick y from are determined. ⊠

In fact, the observation we made in Example 14.14 holds in general. We want to express $(x+y)^n$ as a sum of *monomials* (a monomial is a constant times some product of variables; for example $6x^2y^2$ is a monomial). The monomials have the form $a_kx^ky^{n-k}$ for $k \in \{0, \dots n\}$, where $x^0 = y^0 = 1$ as usual. To obtain x^ky^{n-k}, we pick x from k of the terms in the product $(x+y)^n$, and pick y from all the remaining terms. This can be done $\binom{n}{k}$ ways. Thus, we get the following.

Theorem 14.5 (Binomial theorem)**.** *For all x, y, and n,*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

You can give a formal proof by induction. We leave it to you as an exercise. The binomial theorem also explains why we call the values $\binom{n}{k}$ binomial coefficients. The word binomial comes from the fact that we take a sum of two terms (in the case of the binomial theorem the two terms are x and y).

Below is a simple application of the Binomial theorem:

Proposition 14.6. $2^n = \sum_{k=0}^{n} \binom{n}{k}$

Proof. By the binomial theorem, $(1+1)^n = \sum_{k=0}^{n} \binom{n}{k} 1^k 1^{n-k}$, which is equal to $\sum_{k=0}^{n} \binom{n}{k}$. □

14.5 Pigeonhole Principle

The pigeonhole principle is not a counting technique, but it is a powerful tool, so it deserves mentioning now. It often gets used after we find the cardinality of some set.

Theorem 14.7 (Pigeonhole principle)**.** *Let S and T be sets such that $|T| > |S|$. Then for every total function $f : T \to S$, there are at least two elements in T that map to the same element of s under f.*

Before we see an application, let's discuss why this is called the pigeonhole principle. In the theorem above, think of the elements of T as pigeons, and of elements of S as pigeonholes. Since $|T| > |S|$, there are more pigeons than pigeonholes, so after all pigeons enter some pigeonhole, there has to be at least one hole that contains at least two pigeons.

Example 14.15: The UW-Madison student ID-numbers are 10 digits long, and each digit is an integer between 0 and 9. The smallest the sum of those digits can be is 0 (when all digits are zero), and the largest the sum can be is 90 (when all digits are nine). Thus, there are 91 possibilities for the sum of the digits in a UW-Madison student ID number.

Let S be the set of all possible digit sums of student ID numbers. We have just seen that $|S| = 91$. There are more than 200 students enrolled in CS/Math 240 right now. Let T be the set of all students enrolled in CS/Math 240. Since $|T| > |S|$, the pigeonhole principle implies that there are at least two students in CS/Math 240 whose student ID numbers have the same digit sum. ⊠

We remark that regardless of its simplicity, the pigeonhole principle is quite a powerful tool that is used throughout mathematics. Its power stems from the fact that it makes no assumptions whatsoever about the sets S and T, and that all we need in order to apply it is to know $|S|$ and $|T|$.

The drawback is that since we don't know anything about the sets S and T besides their cardinalities, we cannot find two elements of T that map to the same element of S. The pigeonhole principle merely asserts their existence, but does not give us a way of finding them. Thus, arguments using the pigeonhole principle are *nonconstructive*.

This is contrary to, say, the proof that every connected graph has a spanning tree, where we described procedure that always results in a spanning tree.