Massachusetts Institute of Technology
6.042J/18.062J, Fall '05: Mathematics for Computer Science
Prof. Albert R. Meyer and Prof. Ronitt Rubinfeld

Course Notes, Week 9
November 2
revised November 3, 2005, 1270 minutes

Counting, I.

20480135385502964448038	31710048321735013941113017	5763257331083479647409398	8247331000042995311646021
48944599186691567624 0992	32082344215973686470 19265	5800949123548989122628663	8496243997123475922766310
10826620324303796513 70981	34372546563551578648 69113	6042900801199280218026001	8518399140676002660747477
11784808947697061789 94993	35748833930586539237 11365	6116171789137737896701405	8543691283470191452333763
12531273516832396938 51327	36449099460404801899 69149	6144868973001582369723512	8675309258374137092461352
13015051292340778110 69011	37900441327370840944 17246	6247314593851169234746152	8694321112363996867296665
13115671111438664338 82194	38703321274379713553 22815	6814428944266874963488274	8772321203608477245851154
14700294527212035876 86214	40805058045778014513 63100	6870852945543886849147881	8791422161722582546341091
15782710472862574994 33886	41672834610257023481 24920	6914955508120950093732397	9062628024592126283973285
16382439218521762431 92354	42359968311237777882 11249	6949632451365987152423541	9137845566925526349897794
17635802191319859631 02365	46709394457494390421 11220	7128211143613619828415650	9153762966680318929 1934419
18262277956018422310 29694	48153793518653842796 13427	7173920083651862307925394	9270880194077636406984249
18439718626751020372 01420	48370529482129226044 42190	7215654874211755676220587	9324301480722103490379204
23969511937221345261 77237	51063894238550185506 71530	7256932847164391040233050	9436090832146695147140581
27813945682685998010 96354	51423681920047692180 69910	7332822657075235431620317	9475308159734538249013238
27966051967136104054 08019	51812340961301440840 41856	7426441829541573444964139	9492376623917486974923202
29310163947619752631 90347	51982673981256179943 91348	7632198126531809327186321	9511972558779880288252979
29334580582944051551 97296	53175929403162312197 58372	7712154432211912882310511	9602413424619187112552264
30755144104909759203 15348	53843581267717941283 56947	7858918664240262356610010	9631217114906129219461111
31114749852527934528 60017	54392117122489019954 23441	7898156786763212963178679	9908189853102753335981319
31456215879361201184 38701	56103798260092838192 760458	8147591017037573337848616	9913237476341764299813987
31489012556288811031 98549	56323175554652286776 76044	8149436716871371161932035	
31576931053251112843 21993	56921683746370196174 23712	8176063831682536571306791	

Two different subsets of the ninety 25-digit numbers shown above have the same sum. For example, maybe the sum of the numbers in the first column is equal to the sum of the numbers in the second column. Can you find two such subsets? This is a very difficult computational problem. But we'll prove that such subsets must exist! This is the sort of weird conclusion one can reach by tricky use of counting, the topic of this chapter.

Counting seems easy enough: 1, 2, 3, 4, etc. This explicit approach works well for counting simple things, like your toes, and for extremely complicated things for which there's no identifiable structure. However, subtler methods can help you count many things in the vast middle ground, such as:

- The number of different ways to select a dozen doughnuts when there are five varieties available.

- The number of 16-bit numbers with exactly 4 ones.

Counting is useful in computer science for several reasons:

- Determining the time and storage required to solve a computational problem— a central objective in computer science— often comes down to solving a counting problem.

- Counting is the basis of probability theory, which in turn is perhaps the most important topic this term.

- Two remarkable proof techniques, the "pigeonhole principle" and "combinatorial proof", rely on counting. These lead to a variety of interesting and useful insights.

We're going to present a lot of rules for counting. These rules are actually theorems, but we're generally not going to prove them. Our objective is to teach you counting as a practical skill, like integration. And most of the rules seem "obvious" anyway.

1   Counting One Thing by Counting Another

How do you count the number of people in a crowded room? We could count heads, since for each person there is exactly one head. Alternatively, we could count ears and divide by two. Of course, we might have to adjust the calculation if someone lost an ear in a pirate raid or someone was born with three ears. The point here is that we can often *count one thing by counting another*, though some fudge factors may be required. This is the central theme of counting, from the easiest problems to the hardest.

In more formal terms, every counting problem comes down to determining the size of some set. The *size* or *cardinality* of a set S is the number of elements in S and is denoted $|S|$. In these terms, we're claiming that we can often *find the size of one set S by finding the size of a related set T*. We already have a mathematical tool for relating one set to another: relations. Not surprisingly, a particular kind of relation is at the heart of counting.

1.1   The Bijection Rule

If we can pair up all the girls at a dance with all the boys, then there must be an equal number of each. This simple observation generalizes to a powerful counting rule:

Rule 1 (Bijection Rule). If there exists a bijection $f : A \to B$, then $|A| = |B|$.

In the example, A is the set of boys, B is the set of girls, and the function f defines how they are paired.

The Bijection Rule acts as a magnifier of counting ability; if you figure out the size of one set, then you can immediately determine the sizes of many other sets via bijections. For example, let's return to two sets mentioned earlier:

A = all ways to select a dozen doughnuts when five varieties are available

B = all 16-bit sequences with exactly 4 ones

Let's consider a particular element of set A:

$$\underbrace{0\,0}_{\text{chocolate}} \quad \underbrace{}_{\text{lemon-filled}} \quad \underbrace{0\,0\,0\,0\,0}_{\text{sugar}} \quad \underbrace{0\,0}_{\text{glazed}} \quad \underbrace{0\,0}_{\text{plain}}$$

We've depicted each doughnut with a 0 and left a gap between the different varieties. Thus, the selection above contains two chocolate doughnuts, no lemon-filled, six sugar, two glazed, and two plain. Now let's put a 1 into each of the four gaps:

$$\underbrace{0\;0}_{\text{chocolate}} \quad 1 \quad \underbrace{}_{\text{lemon-filled}} \quad 1 \quad \underbrace{0\,0\,0\,0\,0\,0}_{\text{sugar}} \quad 1 \quad \underbrace{0\;0}_{\text{glazed}} \quad 1 \quad \underbrace{0\;0}_{\text{plain}}$$

We've just formed a 16-bit number with exactly 4 ones— an element of B!

This example suggests a bijection from set A to set B: map a dozen doughnuts consisting of:

$$c \text{ chocolate, } l \text{ lemon-filled, } s \text{ sugar, } g \text{ glazed, and } p \text{ plain}$$

to the sequence:

$$\underbrace{0\ldots0}_{c} \quad 1 \quad \underbrace{0\ldots0}_{l} \quad 1 \quad \underbrace{0\ldots0}_{s} \quad 1 \quad \underbrace{0\ldots0}_{g} \quad 1 \quad \underbrace{0\ldots0}_{p}$$

The resulting sequence always has 16 bits and exactly 4 ones, and thus is an element of B. Moreover, the mapping is a bijection; every such bit sequence is mapped to by exactly one order of a dozen doughnuts. Therefore, $|A| = |B|$ by the Bijection Rule!

This demonstrates the magnifying power of the bijection rule. We managed to prove that two very different sets are actually the same size— even though we don't know exactly how big either one is. But as soon as we figure out the size of one set, we'll immediately know the size of the other.

This particular bijection might seem frighteningly ingenious if you've not seen it before. But you'll use essentially this same argument over and over, and soon you'll consider it boringly routine.

1.2   Sequences

The Bijection Rule lets us count one thing by counting another. This suggests a general strategy: get really good at counting just a *few* things and then use bijections to count *everything else*. This is the strategy we'll follow. In particular, we'll get really good at counting *sequences*. When we want to determine the size of some other set T, we'll find a bijection from T to a set of sequences S. Then we'll use our super-ninja sequence-counting skills to determine $|S|$, which immediately gives us $|T|$. We'll need to hone this idea somewhat as we go along, but that's pretty much the plan!

2   Two Basic Counting Rules

We'll harvest our first crop of counting problems with two basic rules.

2.1   The Sum Rule

Linus allocates his big sister Lucy a quota of 20 crabby days, 40 irritable days, and 60 generally surly days. On how many days can Lucy be out-of-sorts one way or another? Let set C be her crabby days, I be her irritable days, and S be the generally surly. In these terms, the answer to the question is $|C \cup I \cup S|$. Now assuming that she is permitted at most one bad quality each day, the size of this union of sets is given by the Sum Rule:

Rule 2 (Sum Rule). If $A_1, A_2, \ldots, A_n$ are disjoint sets, then:

$$|A_1 \cup A_2 \cup \ldots \cup A_n| = |A_1| + |A_2| + \ldots + |A_n|$$

Thus, according to Linus' budget, Lucy can be out-of-sorts for:

$$\begin{aligned}
|C \cup I \cup S| &= |C| \cup |I| \cup |S| \\
&= 20 + 40 + 60 \\
&= 120 \text{ days}
\end{aligned}$$

Notice that the Sum Rule holds only for a union of *disjoint* sets. Finding the size of a union of intersecting sets is a more complicated problem that we'll take up later.

2.2   The Product Rule

The product rule gives the size of a product of sets. Recall that if $P_1, P_2, \ldots, P_n$ are sets, then

$$P_1 \times P_2 \times \ldots \times P_n$$

is the set of all sequences whose first term is drawn from P_1, second term is drawn from P_2 and so forth.

Rule 3 (Product Rule). If $P_1, P_2, \ldots P_n$ are sets, then:

$$|P_1 \times P_2 \times \ldots \times P_n| = |P_1| \cdot |P_2| \cdots |P_n|$$

Unlike the sum rule, the product rule does not require the sets $P_1, \ldots, P_n$ to be disjoint. For example, suppose a *daily diet* consists of a breakfast selected from set B, a lunch from set L, and a dinner from set D:

$$\begin{aligned}
B &= \{\text{pancakes, bacon and eggs, bagel, Doritos}\} \\
L &= \{\text{burger and fries, garden salad, Doritos}\} \\
D &= \{\text{macaroni, pizza, frozen burrito, pasta, Doritos}\}
\end{aligned}$$

Then $B \times L \times D$ is the set of all possible daily diets. Here are some sample elements:

$$\begin{aligned}
&(\text{pancakes, burger and fries, pizza}) \\
&(\text{bacon and eggs, garden salad, pasta}) \\
&(\text{Doritos, Doritos, frozen burrito})
\end{aligned}$$

The Product Rule tells us how many different daily diets are possible:

$$\begin{aligned}
|B \times L \times D| &= |B| \cdot |L| \cdot |D| \\
&= 4 \cdot 3 \cdot 5 \\
&= 60
\end{aligned}$$

2.3 Putting Rules Together

Few counting problems can be solved with a single rule. More often, a solution is a flury of sums, products, bijections, and other methods. Let's look at some examples that bring more than one rule into play.

Passwords

The sum and product rules together are useful for solving problems involving passwords, telephone numbers, and license plates. For example, on a certain computer system, a valid password is a sequence of between six and eight symbols. The first symbol must be a letter (which can be lowercase or uppercase), and the remaining symbols must be either letters or digits. How many different passwords are possible?

Let's define two sets, corresponding to valid symbols in the first and subsequent positions in the password.

$$F = \{a, b, \ldots, z, A, B, \ldots, Z\}$$
$$S = \{a, b, \ldots, z, A, B, \ldots, Z, 0, 1, \ldots, 9\}$$

In these terms, the set of all possible passwords is:

$$(F \times S^5) \cup (F \times S^6) \cup (F \times S^7)$$

Thus, the length-six passwords are in set $F \times S^5$, the length-seven passwords are in $F \times S^6$, and the length-eight passwords are in $F \times S^7$. Since these sets are disjoint, we can apply the Sum Rule and count the total number of possible passwords as follows:

$$
\begin{aligned}
\left|(F \times S^5) \cup (F \times S^6) \cup (F \times S^7)\right| &= \left|F \times S^5\right| + \left|F \times S^6\right| + \left|F \times S^7\right| && \text{Sum Rule} \\
&= |F| \cdot |S|^5 + |F| \cdot |S|^6 + |F| \cdot |S|^7 && \text{Product Rule} \\
&= 52 \cdot 62^5 + 52 \cdot 62^6 + 52 \cdot 62^7 \\
&\approx 1.8 \cdot 10^{14} \text{ different passwords}
\end{aligned}
$$

Subsets of an n-element Set

How many different subsets of an n element set X are there? For example, the set $X = \{x_1, x_2, x_3\}$ has eight different subsets:

$$\{\} \qquad \{x_1\} \qquad \{x_2\} \qquad \{x_1, x_2\}$$
$$\{x_3\} \quad \{x_1, x_3\} \quad \{x_2, x_3\} \quad \{x_1, x_2, x_3\}$$

There is a natural bijection from subsets of X to n-bit sequences. Let $x_1, x_2, \ldots, x_n$ be the elements of X. Then a particular subset of X maps to the sequence $(b_1, \ldots, b_n)$ where $b_i = 1$ if and only if x_i is in that subset. For example, if $n = 10$, then the subset $\{x_2, x_3, x_5, x_7, x_{10}\}$ maps to a 10-bit sequence as follows:

$$
\begin{array}{rccccccccccc}
\text{subset:} & \{ & & x_2, & x_3, & & x_5, & & x_7, & & & x_{10} & \} \\
\text{sequence:} & (& 0, & 1, & 1, & 0, & 1, & 0, & 1, & 0, & 0, & 1 &)
\end{array}
$$

We just used a bijection to transform the original problem into a question about sequences— *exactly according to plan!* Now if we answer the sequence question, then we've solved our original problem as well.

But how many different n-bit sequences are there? For example, there are 8 different 3-bit sequences:

$$
\begin{array}{cccc}
(0,0,0) & (0,0,1) & (0,1,0) & (0,1,1) \\
(1,0,0) & (1,0,1) & (1,1,0) & (1,1,1)
\end{array}
$$

Well, we can write the set of all n-bit sequences as a product of sets:

$$
\underbrace{\{0,1\} \times \{0,1\} \times \ldots \times \{0,1\}}_{n \text{ terms}} = \{0,1\}^n
$$

Then Product Rule gives the answer:

$$
\begin{aligned}
|\{0,1\}^n| &= |\{0,1\}|^n \\
&= 2^n
\end{aligned}
$$

This means that the number of subsets of an n-element set X is also 2^n. We'll put this answer to use shortly.

3   More Functions: Injections and Surjections

Bijections are both injective and surjective, which makes them a powerful tool for exact counting. We've observed in earlier Notes that surjections and injections by themselves imply certain size relationships between sets:

Rule 4 (Mapping Rule).

1. If $f : X \to Y$ is surjective, then $|X| \geq |Y|$.

2. If $f : X \to Y$ is injective, then $|X| \leq |Y|$.

3. If $f : X \to Y$ is bijective, then $|X| = |Y|$.

3.1   The Pigeonhole Principle

Here is an old puzzle:

> A drawer in a dark room contains red socks, green socks, and blue socks. How many socks must you withdraw to be sure that you have a matching pair?

For example, picking out three socks is not enough; you might end up with one red, one green, and one blue. The solution relies on the Pigeonhole Principle, which is a friendly name for the contrapositive of part (2) of the Mapping Rule. Let's write it down:

If $|X| > |Y|$, then no function $f : X \to Y$ is injective.

And now rewrite it again to eliminate the word "injective."

Rule 5 (Pigeonhole Principle). If $|X| > |Y|$, then for every function $f : X \to Y$, there exist two different elements of X that are mapped to the same element of Y.

Perhaps the relevance of this abstract mathematical statement to selecting footwear under poor lighting conditions is not obvious. However, let A be the set of socks you pick out, let B be the set of colors available, and let f map each sock to its color. The Pigeonhole Principle says that if $|A| > |B| = 3$, then at least two elements of A (that is, at least two socks) must be mapped to the same element of B (that is, the same color). For example, one possible mapping of four socks to three colors is shown below.

$$A \qquad\qquad f \qquad\qquad B$$

1st sock ⟶ red

2nd sock ⟶ green

3rd sock ⟶ blue

4th sock
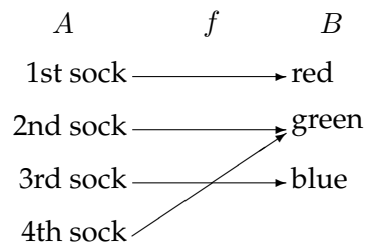
Therefore, four socks are enough to ensure a matched pair.

Not surprisingly, the pigeonhole principle is often described in terms of pigeons:

> *If every pigeon flies into a pigeonhole, and there are more pigeons than holes, then at least two pigeons must fly into some hole.*

In this case, the pigeons form set A, the pigeonholes are set B, and f describes which hole each pigeon flies into.

Mathematicians have come up with many ingenious applications for the pigeonhole principle. If there were a cookbook procedure for generating such arguments, we'd give it to you. Unfortunately, there isn't one. One helpful tip, though: when you try to solve a problem with the pigeonhole principle, the key is to clearly identify three things:

1. The set A (the pigeons).

2. The set B (the pigeonholes).

3. The function f (the rule for assigning pigeons to pigeonholes).

Hairs on Heads

There are a number of generalizations of the pigeonhole principle. For example:

Rule 6 (Generalized Pigeonhole Principle). If $|X| > k \cdot |Y|$, then every function $f : X \to Y$ maps at least $k + 1$ different elements of X to the same element of Y.

For example, if you pick two people at random, surely they are extremely unlikely to have *exactly* the same number of hairs on their heads. However, in the remarkable city of Boston, Massachusetts there are actually *three* people who have exactly the same number of hairs! Of course, there are many bald people in Boston, and they all have zero hairs. But we're talking about non-bald people.

Boston has about 500,000 non-bald peole, and the number of hairs on a person's head is at most 200,000. Let A be the set of non-bald people in Boston, let $B = \{1, \ldots, 200,000\}$, and let f map a person to the number of hairs on his or her head. Since $|A| > 2\,|B|$, the Generalized Pigeonhole Principle implies that at least three people have exactly the same number of hairs. We don't know who they are, but we know they exist!

Subsets with the Same Sum

We asserted that two different subsets of the ninety 25-digit numbers listed on the first page have the same sum. This actually follows from the Pigeonhole Principle. Let A be the collection of all subsets of the 90 numbers in the list. Now the sum of any subset of numbers is at most $90 \cdot 10^{25}$, since there are only 90 numbers and every 25-digit number is less than 10^{25}. So let B be the set of integers $\left\{0, 1, \ldots, 90 \cdot 10^{25}\right\}$, and let f map each subset of numbers (in A) to its sum (in B).

We proved that an n-element set has 2^n different subsets. Therefore:

$$|A| = 2^{90}$$
$$\geq 1.237 \times 10^{27}$$

On the other hand:

$$|B| = 90 \cdot 10^{25} + 1$$
$$\leq 0.901 \times 10^{27}$$

Both quantities are enormous, but $|A|$ is a bit greater than $|B|$. This means that f maps at least two elements of A to the same element of B. In other words, by the Pigeonhole Principle, two different subsets must have the same sum!

Notice that this proof gives no indication *which* two sets of numbers have the same sum. This frustrating variety of argument is called a **nonconstructive proof**.

4   The Generalized Product Rule

We realize everyone has been working pretty hard this term, and we're considering awarding some prizes for *truly exceptional* coursework. Here are some possible categories:

Sets with Distinct Subset Sums

How can we construct a set of n positive integers such that all its subsets have *distinct* sums? One way is to use powers of two:

$$\{1, 2, 4, 8, 16\}$$

This approach is so natural that one suspects all other such sets must involve larger numbers. (For example, we could safely replace 16 by 17, but not by 15.) Remarkably, there are examples involving *smaller* numbers. Here is one:

$$\{6, 9, 11, 12, 13\}$$

One of the top mathematicans of the century, Paul Erdős, conjectured in 1931 that there are no such sets involving *significantly* smaller numbers. More precisely, he conjectured that the largest number must be $> c2^n$ for some constant $c > 0$. He offered \$500 to anyone who could prove or disprove his conjecture, but the problem remains unsolved.

Best Administrative Critique We asserted that the quiz was closed-book. On the cover page, one strong candidate for this award wrote, "There is no book."

Awkward Question Award "Okay, the left sock, right sock, and pants are in an antichain, but how— even with assistance— could I put on all three at once?"

Best Collaboration Statement Inspired by a student who wrote "I worked alone" on Quiz 1.

In how many ways can, say, three different prizes be awarded to n people? This is easy to answer using our startegy of translating the problem about awards into a problem about sequences. Let P be the set of n people in 6.042. Then there is a bijection from ways of awarding the three prizes to the set $P^3 ::= P \times P \times P$. In particular, the assignment:

"person x wins prize #1, y wins prize #2, and z wins prize #3"

maps to the sequence (x, y, z). By the Product Rule, we have $\left|P^3\right| = |P|^3 = n^3$, so there are n^3 ways to award the prizes to a class of n people.

But what if the three prizes must be awarded to *different* students? As before, we could map the assignment

"person x wins prize #1, y wins prize #2, and z wins prize #3"

to the triple $(x, y, z) \in P^3$. But this function is *no longer a bijection*. For example, no valid assignment maps to the triple (Dave, Dave, Becky) because Dave is not allowed to receive two awards. However, there *is* a bijection from prize assignments to the set:

$$S = \left\{(x, y, z) \in P^3 \mid x, y, \text{ and } z \text{ are different people}\right\}$$

This reduces the original problem to a problem of counting sequences. Unfortunately, the Product Rule is of no help in counting sequences of this type because the entries depend on one another; in particular, they must all be different. However, a slightly sharper tool does the trick.

Rule 7 (Generalized Product Rule). Let S be a set of length-k sequences. If there are:

- n_1 possible first entries,

- n_2 possible second entries for each first entry,

- n_3 possible third entries for each combination of first and second entries, etc.

then:
$$|S| = n_1 \cdot n_2 \cdot n_3 \cdots n_k$$

In the awards example, S consists of sequences (x, y, z). There are n ways to choose x, the recipient of prize #1. For each of these, there are $n - 1$ ways to choose y, the recipient of prize #2, since everyone except for person x is eligible. For each combination of x and y, there are $n - 2$ ways to choose z, the recipient of prize #3, because everyone except for x and y is eligible. Thus, according to the Generalized Product Rule, there are
$$|S| = n \cdot (n - 1) \cdot (n - 2)$$

ways to award the 3 prizes to different people.

4.1   Defective Dollars

A dollar is *defective* some digit appears more than once in the 8-digit serial number. If you check your wallet, you'll be sad to discover that defective dollars are all-too-common. In fact, how common are *nondefective* dollars? Assuming that the digit portions of serial numbers all occur equally often, we could answer this question by computing:

$$\text{fraction dollars that are nondefective} = \frac{\text{\# of serial \#'s with all digits different}}{\text{total \# of serial \#'s}}$$

Let's first consider the denominator. Here there are no restrictions; there are are 10 possible first digits, 10 possible second digits, 10 third digits, and so on. Thus, the total number of 8-digit serial numbers is 10^8 by the Product Rule.

Next, let's turn to the numerator. Now we're not permitted to use any digit twice. So there are still 10 possible first digits, but only 9 possible second digits, 8 possible third digits, and so forth. Thus, by the Generalized Product Rule, there are

$$10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 = \frac{10!}{2}$$
$$= 1,814,400$$

serial numbers with all digits different. Plugging these results into the equation above, we find:

$$\text{fraction dollars that are nondefective} = \frac{1,814,400}{100,000,000}$$
$$= 1.8144\%$$

4.2 A Chess Problem

In how many different ways can we place a pawn (p), a knight (k), and a bishop (b) on a chessboard so that no two pieces share a row or a column? A valid configuration is shown below on the left, and an invalid configuration is shown on the right.

valid   invalid

First, we map this problem about chess pieces to a question about sequences. There is a bijection from configurations to sequences

$$(r_p, c_p, r_k, c_k, r_b, c_b)$$

where r_p, r_k, and r_b are distinct rows and c_p, c_k, and c_b are distinct columns. In particular, r_p is the pawn's row, c_p is the pawn's column, r_k is the knight's row, etc. Now we can count the number of such sequences using the Generalized Product Rule:

- r_p is one of 8 rows
- c_p is one of 8 columns
- r_k is one of 7 rows (any one but r_p)
- c_k is one of 7 columns (any one but c_p)
- r_b is one of 6 rows (any one but r_p or r_k)
- c_b is one of 6 columns (any one but c_p or c_k)

Thus, the total number of configurations is $(8 \cdot 7 \cdot 6)^2$.

4.3 Permutations

A ***permutation*** of a set S is a sequence that contains every element of S exactly once. For example, here are all the permutations of the set $\{a, b, c\}$:

$$(a, b, c) \quad (a, c, b) \quad (b, a, c)$$
$$(b, c, a) \quad (c, a, b) \quad (c, b, a)$$

How many permutations of an n-element set are there? Well, there are n choices for the first element. For each of these, there are $n - 1$ remaining choices for the second element. For every combination of the first two elements, there are $n - 2$ ways to choose the third element, and so forth. Thus, there are a total of

$$n \cdot (n - 1) \cdot (n - 2) \cdots 3 \cdot 2 \cdot 1 = n!$$

permutations of an n-element set. In particular, this formula says that there are $3! = 6$ permuations of the 3-element set $\{a, b, c\}$, which is the number we found above.
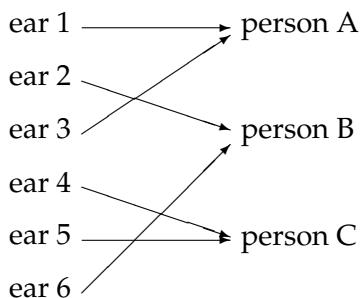
Permutations will come up again in this course approximately 1.6 bazillion times. In fact, permutations are the reason why factorial comes up so often and why we taught you Stirling's approximation:

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

5   The Division Rule

We can count the number of people in a room by counting ears and dividing by two. Or we could count the number of fingers and divide by 10. Or we could count the number of fingers and toes and divide by 20. (Someone is probably short a finger or has an extra ear, but let's not worry about that right now.) These observations lead to an important counting rule.

A k-**to-1 function** maps exactly k elements of the domain to every element of the range. For example, the function mapping each ear to its owner is 2-to-1:

```
ear 1 ─────────────→ person A
ear 2 ╲          ╱
ear 3 ╱────╲────╱── person B
ear 4 ╲      ╲
ear 5 ─────╳────→ person C
ear 6 ╱
```

Similarly, the function mapping each finger to its owner is 10-to-1. And the function maping each each finger or toe to its owner is 20-to-1. Now just as a bijection implies two sets are the same size, a k-to-1 function implies that the domain is k times larger than the domain:

Rule 8 (Division Rule). If $f : A \rightarrow B$ is k-to-1, then $|A| = k \cdot |B|$.

Suppose A is the set of ears in the room and B is the set of people. Since we know there is a 2-to-1 mapping from ears to people, the Division Rule says that $|A| = 2 \cdot |B|$ or, equivalently, $|B| = |A| /2$. Thus, the number of people is half the number of ears.

Now this might seem like a stupid way to count people. But, surprisingly, many counting problems are made much easier by initially counting every item multiple times and then correcting the answer using the Division Rule. Let's look at some examples.
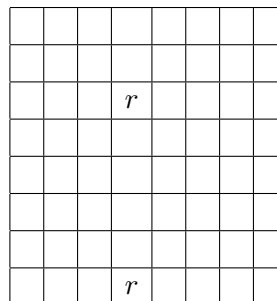
5.1   Another Chess Problem

In how many different ways can you place two identical rooks on a chessboard so that they do not share a row or column? A valid configuration is shown below on the left, and an invalid

configuration is shown on the right.
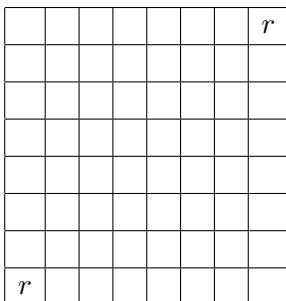
valid   invalid

Let A be the set of all sequences

$$(r_1, c_1, r_2, c_2)$$

where r_1 and r_2 are distinct rows and c_1 and c_2 are distinct columns. Let B be the set of all valid rook configurations. There is a natural function f from set A to set B; in particular, f maps the sequence (r_1, c_1, r_2, c_2) to a configuration with one rook in row r_1, column c_1 and the other rook in row r_2, column c_2.

But now there's a snag. Consider the sequences:

$$(1, 1, 8, 8) \qquad \text{and} \qquad (8, 8, 1, 1)$$

The first sequence maps to a configuration with a rook in the lower-left corner and a rook in the upper-right corner. The second sequence maps to a configuration with a rook in the upper-right corner and a rook in the lower-left corner. The problem is that those are two different ways of describing the *same* configuration! In fact, this arrangement is shown on the left side in the diagram above.

More generally, the function f map exactly two sequences to *every* board configuration; that is f is a 2-to-1 function. Thus, by the quotient rule, $|A| = 2 \cdot |B|$. Rearranging terms gives:

$$|B| = \frac{|A|}{2}$$
$$= \frac{(8 \cdot 7)^2}{2}$$

On the second line, we've computed the size of A using the General Product Rule just as in the earlier chess problem.

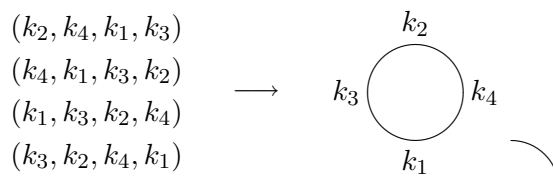## 5.2   Knights of the Round Table

In how many ways can King Arthur seat n different knights at his round table? Two seatings are considered equivalent if one can be obtained from the other by rotation. For example, the following two arrangements are equivalent:

Let A be all the permutations of the knights, and let B be the set of all possible seating arrangements at the round table. We can map each permutaton in set A to a circular seating arrangement in set B by seating the first knight in the permutation anywhere, putting the second knight to his left, the third knight to the left of the second, and so forth all the way around the table. For example:

$$(k_2, k_4, k_1, k_3) \longrightarrow$$
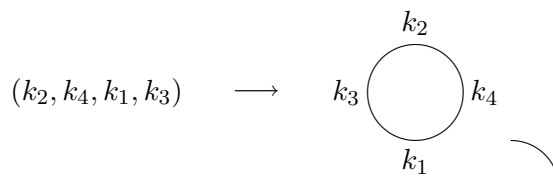
This mapping is actually an n-to-1 function from A to B, since all n cyclic shifts of the original sequence map to the same seating arrangement. In the example, $n = 4$ different sequences map to the same seating arranement:

$$
\begin{array}{l}
(k_2, k_4, k_1, k_3) \\
(k_4, k_1, k_3, k_2) \\
(k_1, k_3, k_2, k_4) \\
(k_3, k_2, k_4, k_1)
\end{array}
\longrightarrow
$$

Therefore, by the division rule, the number of circular seating arrangements is:

$$
\begin{aligned}
|B| &= \frac{|A|}{n} \\
&= \frac{n!}{n} \\
&= (n-1)!
\end{aligned}
$$

Note that $|A| = n!$ since there are $n!$ permutations of n knights.

6   Inclusion-Exclusion

How big is a union of sets? For example, suppose there are 60 Math majors, 200 EECS majors, and 40 Physics majors. How many students are there in these three departments? Let M be the set

of Math majors, E be the set of EECS majors, and P be the set of Physics majors. In these terms, we're asking for $|M \cup E \cup P|$.

The Sum Rule says that the size of union of *disjoint* sets is the sum of their sizes:

$$|M \cup E \cup P| = |M| + |E| + |P| \qquad \text{(if } M, E, \text{ and } P \text{ are disjoint)}$$

However, the sets M, E, and P might *not* be disjoint. For example, there might be a student majoring in both Math and Physics. Such a student would be counted twice on the right sides of this equation, once as an element of M and once as an element of P. Worse, there might be a triple-major counting *three* times on the right side!

Our last counting rule determines the size of a union of sets that are not necessarily disjoint. Before we state the rule, let's build some intuition by considering some easier special cases: unions of just two or three sets.

6.1   Union of Two Sets

For two sets, S_1 and S_2, the Inclusion-Exclusion rule is that the size of their union is:

$$|S_1 \cup S_2| = |S_1| + |S_2| - |S_1 \cap S_2| \tag{1}$$

Intuitively, each element of S_1 is accounted for in the first term, and each element of S_2 is accounted for in the second term. Elements in *both* S_1 and S_2 are counted *twice*— once in the first term and once in the second. This double-counting is corrected by the final term.

We can prove equation (1) rigorously by applying the Sum Rule to some disjoint subsets of $S_1 \cup S_2$. As a first step, we observe that given any two sets, S, T, we can decompose S into the disjoint sets consisting of those elements in S but not T, and those elements in S and also in T. That is, S is the union of the disjoint sets $S - T$ and $S \cap T$. So by the Sum Rule we have

$$|S| = |S - T| + |S \cap T|, \qquad\qquad \text{and so}$$
$$|S - T| = |S| - |S \cap T|. \tag{2}$$

Now we decompose $S_1 \cup S_2$ into three disjoint sets:

$$S_1 \cup S_2 = (S_1 - S_2) \cup (S_2 - S_1) \cup (S_1 \cap S_2). \tag{3}$$

Now we have

$$
\begin{aligned}
|S_1 \cup S_2| &= |(S_1 - S_2) \cup (S_2 - S_1) \cup (S_1 \cap S_2)| && \text{(by (3))} \\
&= |S_1 - S_2| + |S_2 - S_1| + |S_1 \cap S_2| && \text{(Sum Rule)} \\
&= (|S_1| - |S_1 \cap S_2|) + (|S_2| - |S_1 \cap S_2|) + |S_1 \cap S_2| && \text{(by (2))} \\
&= |S_1| + |S_2| - |S_1 \cap S_2| && \text{(algebra)}
\end{aligned}
$$

6.2   Union of Three Sets

So how many students are there in the Math, EECS, and Physics departments? In other words, what is $|M \cup E \cup P|$ if:

$$|M| = 60$$
$$|E| = 200$$
$$|P| = 40$$

The size of a union of three sets is given by a more complicated Inclusion-Exclusion formula:

$$|S_1 \cup S_2 \cup S_3| = |S_1| + |S_2| + |S_3|$$
$$- |S_1 \cap S_2| - |S_1 \cap S_3| - |S_2 \cap S_3|$$
$$+ |S_1 \cap S_2 \cap S_3|$$

Remarkably, the expression on the right accounts for each element in the union of S_1, S_2, and S_3 exactly once. For example, suppose that x is an element of all three sets. Then x is counted three times (by the $|S_1|$, $|S_2|$, and $|S_3|$ terms), subtracted off three times (by the $|S_1 \cap S_2|$, $|S_1 \cap S_3|$, and $|S_2 \cap S_3|$ terms), and then counted once more (by the $|S_1 \cap S_2 \cap S_3|$ term). The net effect is that x is counted just once.

So we can't answer the original question without knowing the sizes of the various intersections. Let's suppose that there are:

> 4   Math - EECS double majors
> 3   Math - Physics double majors
> 11   EECS - Physics double majors
> 2   triple majors

Then $|M \cap E| = 4 + 2$, $|M \cap P| = 3 + 2$, $|E \cap P| = 11 + 2$, and $|M \cap E \cap P| = 2$. Plugging all this into the formula gives:

$$|M \cup E \cup P| = |M| + |E| + |P| - |M \cap E| - |M \cap P| - |E \cap P| + |M \cap E \cap P|$$
$$= 60 + 200 + 40 - 6 - 5 - 13 + 2$$
$$= 278$$

6.2.1   Sequences with 42, 04, or 60

In how many permutations of the set $\{0, 1, 2, \ldots, 9\}$ do either 4 and 2, 0 and 4, or 6 and 0 appear consecutively? For example, none of these pairs appears in:

$$(7, 2, 9, 5, 4, 1, 3, 8, 0, 6)$$

The 06 at the end doesn't count; we need 60. On the other hand, both 04 and 60 appear consecutively in this permutation:

$$(7, 2, 5, \underline{6}, \underline{0}, \underline{4}, 3, 8, 1, 9)$$

Let P_{42} be the set of all permutations in which 42 appears; define P_{60} and P_{04} similarly. Thus, for example, the permutation above is contained in both P_{60} and P_{04}. In these terms, we're looking for the size of the set $P_{42} \cup P_{04} \cup P_{60}$.

First, we must determine the sizes of the individual sets, such as P_{60}. We can use a trick: group the 6 and 0 together as a single symbol. Then there is a natural bijection between permutations of $\{0, 1, 2, \ldots 9\}$ containing 6 and 0 consecutively and permutations of:

$$\{60, 1, 2, 3, 4, 5, 7, 8, 9\}$$

For example, the following two sequences correspond:

$$(7, 2, 5, \underline{6}, \underline{0}, 4, 3, 8, 1, 9) \qquad \leftrightarrow \qquad (7, 2, 5, \underline{60}, 4, 3, 8, 1, 9)$$

There are 9! permutations of the set containing 60, so $|P_{60}| = 9!$ by the Bijection Rule. Similarly, $|P_{04}| = |P_{42}| = 9!$ as well.

Next, we must determine the sizes of the two-way intersections, such as $P_{42} \cap P_{60}$. Using the grouping trick again, there is a bijection with permutations of the set:

$$\{42, 60, 1, 3, 5, 7, 8, 9\}$$

Thus, $|P_{42} \cap P_{60}| = 8!$. Similarly, $|P_{60} \cap P_{04}| = 8!$ by a bijection with the set:

$$\{604, 1, 2, 3, 5, 7, 8, 9\}$$

And $|P_{42} \cap P_{04}| = 8!$ as well by a similar argument. Finally, note that $|P_{60} \cap P_{04} \cap P_{42}| = 7!$ by a bijection with the set:

$$\{6042, 1, 3, 5, 7, 8, 9\}$$

Plugging all this into the formula gives:

$$|P_{42} \cup P_{04} \cup P_{60}| = 9! + 9! + 9! - 8! - 8! - 8! + 7!$$

6.3   Union of n Sets

The size of a union of n sets is given by the following rule.

Rule 9 (Inclusion-Exclusion).
$$|S_1 \cup S_2 \cup \cdots \cup S_n| =$$

the sum of the sizes of the individual sets
minus the sizes of all two-way intersections
plus the sizes of all three-way intersections
minus the sizes of all four-way intersections
plus the sizes of all five-way intersections, etc.

There are various ways to write the Inclusion-Exclusion formula in mathematical symbols, but none are particularly clear, so we've just used words. The formulas for unions of two and three sets are special cases of this general rule.

6.3.1   Counting Primes

How many of the numbers $1, 2, \ldots, 100$ are prime? One way to answer this question is to test each number up to 100 for primality and keep a count. This requires considerable effort. (Is 57 prime? How about 67?)

Another approach is to use the Inclusion-Exclusion Principle. This requires one trick: to determine the number of primes, we will first count the number of *non-primes*. By the Sum Rule, we can then find the number of primes by subtraction from 100. This trick of "counting the complement" is a good one to remember.

Reduction to a Union of Four Sets

The set of non-primes in the range $1, \ldots, 100$ consists of the set, C, of composite numbers in this range: $4, 6, 8, 9, \ldots, 99, 100$ and the number 1, which is neither prime nor composite. The main job is to determine the size of the set C of composite numbers. For this purpose, define A_m to be the set of numbers in the range $m + 1, \ldots, 100$ that are divisible by m:

$$A_m ::= \{x \le 100 \mid x > m \text{ and } (m \mid x)\}$$

For example, A_2 is all the even numbers from 4 to 100. The following Lemma will now allow us to compute the cardinality of C by using Inclusion-Exclusion for the union of four sets:

Lemma 6.1.
$$C = A_2 \cup A_3 \cup A_5 \cup A_7.$$

Proof. We prove the two sets equal by showing that each contains the other.

To show that $A_2 \cup A_3 \cup A_5 \cup A_7 \subseteq C$, let n be an element of $A_2 \cup A_3 \cup A_5 \cup A_7$. Then $n \in A_m$ for $m = 2, 3, 5$ or 7. This implies that n is in the range $1, \ldots, 100$ and is composite because it has m as a factor. That is, $n \in C$.

Conversely, to show that $C \subseteq A_2 \cup A_3 \cup A_5 \cup A_7$, let n be an element of C. Then n is a composite number in the range $1, \ldots, 100$. This means that n has at least two prime factors. Now if both prime factors are > 10, then their product would be a number > 100 which divided n, contradicting the fact that $n < 100$. So n must have a prime factor ≤ 10. But 2, 3, 5, and 7 are the only primes ≤ 10. This means that n is an element of A_2, A_3, A_5, or A_7, and so $n \in A_2 \cup A_3 \cup A_5 \cup A_7$. $\square$

Computing the Cardinality of the Union

Now it's easy to find the cardinality of each set A_m: every mth integer is divisible by m, so the number of integers in the range $1, \ldots, 100$ that are divisible by m is simply $\lfloor 100/m \rfloor$. So

$$|A_m| = \left\lfloor \frac{100}{m} \right\rfloor - 1,$$

where the -1 arises because we defined A_m to exclude m itself. This formula gives:

$$
\begin{aligned}
|A_2| &= \lfloor \tfrac{100}{2} \rfloor - 1 = 49 \\
|A_3| &= \lfloor \tfrac{100}{3} \rfloor - 1 = 32 \\
|A_5| &= \lfloor \tfrac{100}{5} \rfloor - 1 = 19 \\
|A_7| &= \lfloor \tfrac{100}{7} \rfloor - 1 = 13
\end{aligned}
$$

Notice that these sets A_2, A_3, A_5, and A_7 are not disjoint. For example, 6 is in both A_2 and A_3. Since the sets intersect, we must use the Inclusion-Exclusion Principle:

$$
\begin{aligned}
|C| = {}& |A_2 \cup A_3 \cup A_5 \cup A_7| \\
= {}& |A_2| + |A_3| + |A_5| + |A_7| \\
& - |A_2 \cap A_3| - |A_2 \cap A_5| - |A_2 \cap A_7| - |A_3 \cap A_5| - |A_3 \cap A_7| - |A_5 \cap A_7| \\
& + |A_2 \cap A_3 \cap A_5| + |A_2 \cap A_3 \cap A_7| + |A_2 \cap A_5 \cap A_7| + |A_3 \cap A_5 \cap A_7| \\
& - |A_2 \cap A_3 \cap A_5 \cap A_7|
\end{aligned}
$$

There are a lot of terms here! Fortunately, all of them are easy to evaluate. For example, $|A_3 \cap A_7|$ is the number of multiples of $3 \cdot 7 = 21$ in the range 1 to 100, which is $\lfloor 100/21 \rfloor = 4$. Substituting such values for all of the terms above gives:

$$
\begin{aligned}
|C| = {}& 49 + 32 + 19 + 13 \\
& - 16 - 10 - 7 - 6 - 4 - 2 \\
& + 3 + 2 + 1 + 0 \\
& - 0 \\
= {}& 74
\end{aligned}
$$

This calculation shows that there are 74 composite numbers in the range 1 to 100. Since the number 1 is neither composite nor prime, there are $100 - 74 - 1 = 25$ primes in this range.

At this point it may seem that checking each number from 1 to 100 for primality and keeping a count of primes might have been easier than using Inclusion-Exclusion. However, the Inclusion-Exclusion approach used here is asymptotically faster as the range of numbers grows large.