Counting II

We realize everyone has been working pretty hard this term[1], and we're considering awarding some prizes for *truly exceptional* coursework. Here are some possible categories:

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."

Best Collaboration Statement Inspired by the student who wrote "I worked alone" on quiz 1.

Olfactory Fixation Award A surprisingly competitive category this term, this goes to the student who comes up with the greatest number of odor-related mathematical examples.

We also considered some less flattering categories such as Proof Most Likely Readable from the Surface of the Moon, Solution Most Closely Resembling a Football Play Diagram with Good Yardage Potential, etc. But then we realized that you all might think up similar "awards" for the course staff and decided to turn the whole matter into a counting problem. In how many ways can, say, three different prizes be awarded to n people?

Remember our basic strategy for counting:

1. Learn to count sequences.
2. Translate everything else into a sequence-counting problem via bijections.

We'll flesh out this strategy considerably today, but the rough outline above is good enough for now.

So we first need to find a bijection that translates 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 $P \times P \times P$. 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). All that remains is to count these sequences. By the Product Rule, we have:

$$\begin{aligned} |P \times P \times P| &= |P| \cdot |P| \cdot |P| \\ &= n^3 \end{aligned}$$

Thus, there are n^3 ways to award the prizes to a class of n people.

[1] Actually, these notes were written last fall, but the problem sets are no easier this term. :-)

1 The Generalized Product Rule

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 $(x, y, z) \in P \times P \times P$. 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:

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

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 1 (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.

1.1 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 Generalized Product Rule. (Alternatively, you could conclude this using the ordinary Product Rule; however, the Generalized Product Rule is strictly more powerful. So you might as well forget the orignial Product Rule now and free up some brain space for 6.002.)

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 there are

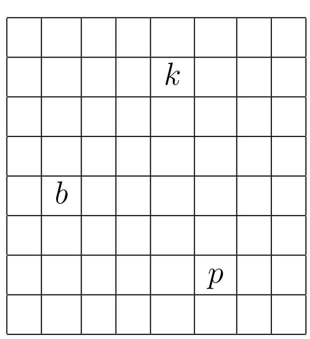
$$\begin{aligned} 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 &= \frac{10!}{2} \\ &= 1,814,400 \end{aligned}$$

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

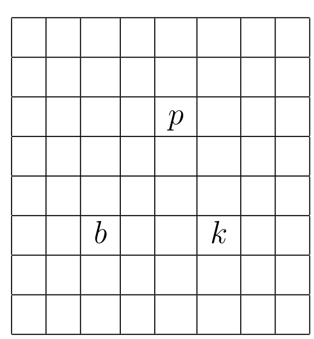
$$\begin{aligned} \text{fraction dollars that are nondefective} &= \frac{1,814,400}{100,000,000} \\ &= 1.8144\% \end{aligned}$$

1.2 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 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 $(8 \cdot 7 \cdot 6)^2$.

1.3 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 $\{a, b, c\}$:

$$\begin{array}{ccc} (a, b, c) & (a, c, b) & (b, a, c) \\ (b, c, a) & (c, a, b) & (c, b, a) \end{array}$$

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.

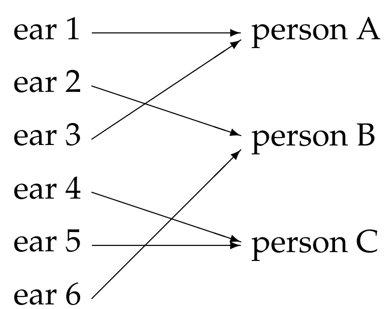
Permutations will come up again in this course approximately 1.6 bazillion times. In fact, permutations are the reason why factorial comes up so often and why we taught you Stirling's approximation:

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

2 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:



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:

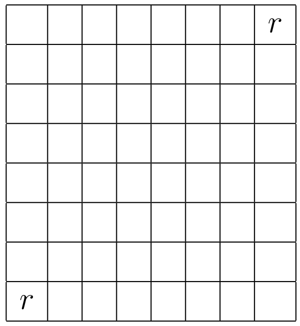
Rule 2 (Division Rule). *If $f : A \to B$ is k-to-1, then $|A| = k \cdot |B|$.*

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.

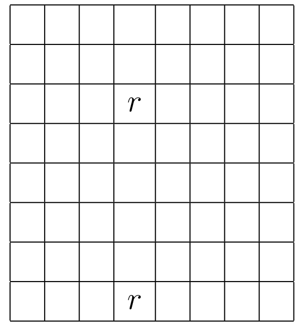
Now this might seem like a stupid way to count people. But, surprisingly, many counting problems are made much easier by initially counting every item multiple times and then correcting the answer using the Division Rule. Let's look at some examples.

2.1 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:

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

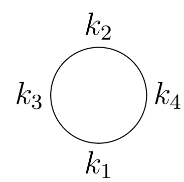
On the second line, we've computed the size of A using the General Product Rule just as in the earlier chess problem.

2.2 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:

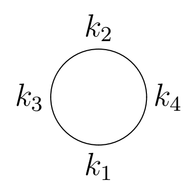
Let A be all the permutations of the knights, and let B be the set of all possible seating arrangements at the round table. We can map each permutaton in set A to a circular seating arrangement in set B by seating the first knight in the permutation anywhere, putting the second knight to his left, the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

$(k_2, k_4, k_1, k_3) \quad \Rightarrow$



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:

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



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.

3 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.

3.1 Union of Two Sets

For two sets, S_1 and S_2, the size of the union is given by the following equation:

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

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

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

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

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

3.2 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:

$$
\begin{aligned}
|M| &= 60 \\
|E| &= 200 \\
|P| &= 40
\end{aligned}
$$

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

$$
\begin{aligned}
|S_1 \cup S_2 \cup S_3| = {} & |S_1| + |S_2| + |S_3| \\
& - |S_1 \cap S_2| - |S_1 \cap S_3| - |S_2 \cap S_3| \\
& + |S_1 \cap S_2 \cap S_3|
\end{aligned}
$$

Remarkably, the expression on the right accounts for each element in the 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:

$$
\begin{aligned}
|M \cup E \cup P| &= |M| + |E| + |P| - |M \cap E| - |M \cap P| - |E \cap P| + |M \cap E \cap P| \\
&= 60 + 200 + 40 - 6 - 5 - 13 + 2 \\
&= 278
\end{aligned}
$$

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:

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

3.3 Union of *n* Sets

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

Rule 3 (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.

4 The Grand Scheme for Counting

The rules and techniques we've covered to this point snap together into an overall scheme for solving elementary counting problems. Here it is:

Grand Scheme for Counting

1. Learn to count sequences using two techniques:
   - the General Product Rule
   - the BOOKKEEPER formula
2. Translate everything else to a sequence-counting problem via:
   - bijections
   - k-to-1 functions
3. But for unions of sets, use Inclusion-Exclusion.

Everything here should be familiar to you by now, except for the BOOKKEEPER formula, which you'll see in recitation tomorrow.