CS 30: Discrete Math in CS (Winter 2020): Lecture 9

Date: 22nd January, 2020 (Wednesday)
Topic: Combinatorics: Product and Sum Principles
Disclaimer: These notes have not gone through scrutiny and in all probability contain errors.
Please discuss in Piazza/email errors to deeparnab@dartmouth.edu

---

The area of combinatorics is all about counting. It is an important tool in various areas; in CS, this is relevant to, say, figuring out if we have enough IP addresses for the websites, or if we have enough time to brute-force a certain problem. In the coming 3-4 classes, we will see some methods to count objects.

1. **Counting Sequences: Product Principle.**

   Often we need to count the number of length k sequences satisfying certain properties; call such sequences valid. Suppose the properties governing these sequences lead to the following choice rule for "building a valid sequence character-by-character". There are n_1 possibilities for the first character of the sequence. That is, if we look at all the allowed sequences, the first character has n_1 possibilities. So we can pick one of them to be the first character. *Given the choice of the first character*, no matter what the choice was, suppose the second character has n_2 choices. That is, if we look at all the valid sequences whose first character is fixed to some character, the second character has n_2 choices. Similarly, *Given the choice of the first and the second character*, no matter what these choices are, suppose the third character has n_3 choices. More generally,

   For any $1 \le i < k$, *Given the choices of the first i characters*, no matter what they are, suppose the $(i+1)$th character has n_{i+1} choices.

   Then, the total number of such sequences are $n_1 \cdot n_2 \cdot \dots \cdot n_k$.

   Examples:

   (a) *Number of Bit Strings.* How many length k-bit strings are there? Think of the bit-string as a sequence of length n where each character is 0 or 1. Imagine trying to build one such bit-string bit-by-bit. There are, therefore, 2 ways to choose the first bit. So $n_1 = 2$. Now, given any choice of the first bit, 0 or 1, the second bit of a valid bit string could be 0 or 1. Therefore, $n_2 = 2$. Similarly, no matter how we choose the first i bits, there are 2 choices for the $(i+1)$th bit. The product rule therefore implies that the number of bit strings is $2 \times 2 \times \cdots 2 = 2^k$.

   A notation: the set of bit strings of length n is often denoted by $\{0,1\}^n$.

   (b) *Number of Constrained Bit Strings.* How many length n-bit strings are there whose first two bits are the same? Again, let's try building a valid bit-string bit-by-bit. The first bit has 2 choices – it could be 0 or 1. But once the first bit has been fixed, say to 0, then the second bit has only one choice; it has to be fixed to 0. But once the first two bits are fixed, either to 00 or 11, the third bit is back to having 2 choices; no constraint on the third bit. Same for the fourth, fifth, and ... nth bit. Thus, the number of bit strings with the first two bits same are $2 \cdot 1 \cdot 2 \cdots 2 = 2^{n-1}$. ✍

Exercise: *How many length n-bit strings are there whose first 5 bits are the same? Assume $n \geq 5$.*

(c) *Number of Permutations.* How many permutations of $\{1, 2, \ldots, n\}$ are there? A permutation is an ordering of the elements of the set with no repetitions.

Again, the first entry of the order can be chosen in n ways. Given the first, the second can be chosen in $n - 1$ ways. So on and so forth. Thus, the number of permutations of $\{1, 2, \ldots, n\}$ is

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

✍

Exercise: *How many anagrams (not necessarily in the dictionary) of the word TABLE are there?*

(d) *Patterned Strings.* How many 4 digit numbers are there whose first two digits must be even, and the last two digits must be odd? Once again, let us build such 4 digit numbers, digit-by-digit.

The first digit must be even. Also note it can't be 0. So it has 4 choices: $\{2, 4, 6, 8\}$. The second digit must be even too. There is no other constraint on it. It can be 0 too. So it has 5 choices: $\{0, 2, 4, 6, 8\}$. The third digit needs to be odd. So it has 5 choices: $\{1, 3, 5, 7, 9\}$, and finally, the fourth digit has 5 choices too: $\{1, 3, 5, 7, 9\}$. So the final answer is $4 \cdot 5 \cdot 5 \cdot 5 = 500$. ✍

Exercise: *How many 10 digit numbers are there whose first 5 digits are odd and last 5 digits are even?*

(e) *Slightly more complicated rules.* How many 4 digit numbers are there whose first two digits sum to 9? Let's try again the idea of forming such a number digit-by-digit. How many choices are there for the first digit? The answer is 9 : it has to come from the set $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Given any such choice, how many choices are there for the second digit? Here we inspect that the answer is 1. No matter, what the first digit is, the second digit is 9 minus that. In other words, the first digit *fixes* the second digit. How about the third digit? Well, that has no constraint, and has 10 possible choices : $\{0, 1, \ldots, 9\}$. Same for the fourth digit. Thus all in all, the number of such numbers is $9 \cdot 1 \cdot 10 \cdot 10 = 900$. ✍

Exercise: *How many 4 digit numbers are there whose second and fourth digit sum to 8?*

(f) *Failure of the rule.* The rule is not a panacea and will not solve all your problems. In particular, if the number of choices of a certain ith character *changes* depending on the choices for the first $(i - 1)$ characters, then the product rule doesn't apply.

Examples:

i. *How many 4 digit numbers are there whose first two digits sum to* less than or equal to *9?* Again, let's try to form such a number digit-by-digit. The first digit indeed has 9 choices as before. However, the *number* of choices of the second digit *depends* on the choice of the first digit. For instance, if the first digit is 1, then the second digit has 9 choices: $\{0, 1, 2, 3, \ldots, 8\}$. But if the first digit is 9, then the second digit is forced to be 0. Since the number of choices for the second digit *depends* on what we choose in the first digit, this rule doesn't apply.

ii. *How many* 10*-length bit strings are there with no* three *consecutive* 0*s?* Again, let's try to form such a bit string bit-by-bit. The first bit has two choices: $\{0,1\}$. The second bit, also, has two choices no matter what we choose for the first bit. It is the third bit where the rule breaks down. Because, if the first two bits were 01, or 10, or 11, the third bit has two choices: $\{0,1\}$. However, if the first two bits were 00, then the third bit has only one choice: $\{1\}$. Therefore, since the number of choices of the third bit depends on the choice of the first two bits, we cannot apply the product rule.

2. **Inclusion Exclusion: The Sum Principle.** In all the examples above, we wanted to count the number of "objects" satisfying certain conditions. We saw the product principle as a way to get many answers, but also saw how it can't solve everything we want. The next principle is another very important principle which solves some of the problems we face.

   The first step is to think of the "objects satisfying certain conditions" to form a set S. We want to find out $|S|$. Then the second step is think of this set S as a *union* of two or more sets. The simplest case is when S is a union of two *disjoint* sets A and B, where it is *easier* to figure out the cardinalities $|A|$ and $|B|$. Since $S = A \cup B$ and $A \cap B = \varnothing$, we have $|S| = |A| + |B|$.

   Examples.

   (a) (Counting Passwords: a trivial example) A particular password system allows you to either have a 3 digit number (first number can be 0) as your password, or a 4-letter lower-case string as your password? How many passwords are possible?
   If S is the set of passwords, we can partition into two sets A and B, where A is the set of 3 digit numbers and B is the set of 4-letter strings. Clearly $S = A \cup B$ and $A \cap B = \varnothing$. Furthermore, each of $|A|$ and $|B|$ can be evaluated easily using the product principle. In particular, $|A| = 10^3$ and $|B| = 26^4$. Thus, the answer is $|S| = |A| + |B| = 10^3 + 26^4$.

   (b) *How many four digit numbers have the sum of first two digits strictly less than* 3*?* As we saw earlier, the product principle doesn't quite help since if the first digit is 1, then the second digit has two choices: $\{0,1\}$, but if the first digit is 2, then the second digit has only one choice: $\{0\}$.
   But, as you may already see, this is actually a very simple problem. To illustrate the above idea though, let us call the set of four digit numbers whose sum of first two digits is < 3 to be S. Let us *partition* S into two parts:

   $$A \subseteq S = \{\text{all four digit numbers in } S \text{ whose first digit is } 1\}$$

   and

   $$B \subseteq S = \{\text{all four digit numbers in } S \text{ whose first digit is } 2\}$$

   Note, $S = A \cup B$ since every four digit number whose sum of first two digits is < 3 has to begin with either 1 or 2. Furthermore, $A \cap B = \varnothing$; since a number which starts with 1 cannot start with 2. Finally, both $|A|$ and $|B|$ can be evaluated very easily using the product principle.
   To calculate $|A|$, let's construct a number in A digit by digit. The first digit has only one choice, it is 1. The second digit now has two choices $\{0,1\}$. The third and fourth digits have 10 choices each. Thus, $|A| = 2 \cdot 10 \cdot 10 = 200$.

To calculate $|B|$, let's construct a number in B digit-by-digit. The first digit has only one choice, it is the number 2. The second digit now has only one choice $\{0\}$. The third and fourth digit has 10 choices each. Thus, $|B| = 1 \cdot 10 \cdot 10 = 100$.

Thus, $|S| = |A| + |B| = 300$.

> **Remark:** *If instead of $S = A \cup B$ where A and B were disjoint, if we have $S = A \cup B \cup C$ where all three were pairwise disjoint, then $|S| = |A| + |B| + |C|$. More generally, if we had $S = A_1 \cup A_2 \cup \cdots \cup A_k$ where $A_i \cap A_j = \varnothing$ whenever $i \neq j$, then $|S| = \sum_{i=1}^{k} |A_i|$.*

✍

> **Exercise:** *Using the above remark, can you now figure out the number of four digit numbers whose first two digits sum up to less than or equal to 9? Hint: divide the set S into 9 disjoint sets depending on what the first digit is.*

(c) *How many bit strings of length L are there with exactly 1 one?* This is another example where the product rule fails. Do you see it? How will we solve this one?

We think a bit about the condition. How do these bit strings look like? Well, there is one position, let us call it a, where the string takes value 1, and is 0 everywhere else.

Therefore, we can partition the set S which we are trying to count into $S_1 \cup S_2 \cup \cdots \cup S_L$ where

$$S_i := \{ \text{ Length } L\text{-bit strings which have exactly 1 one } \textbf{and} \text{ the } i\text{th bit is } 1\}$$

Note that $S_i \cap S_j = \varnothing$ whenever $i \neq j$. Why? Because if a string is in S_i it has one in the ith position and 0 everywhere else. Thus it cannot have a one in the jth position. Therefore, it cannot be in S_j.

Thus, $|S| = |S_1| + |S_2| + \cdots + |S_L|$.

How large is $|S_i|$? It is 1 – there is only one string which has 1 in position i and 0 everywhere else. Therefore, the size of S is L.

(d) *How many 8-bit strings are there with at least 2 ones?* This is an example of a question where looking at the *negation* is much more helpful. Let U be the universe of all 8-bit strings. We are trying to figure out the size of the subset $S \subseteq U$ which have at least two zeros. Looking at $U \setminus S$ is much easier.

Indeed, let A be the set of 8-bit strings with exactly 1 one, and let B be the set of 8 bit strings with exactly 0 ones. Then, note $U = A \cup B \cup S$ and futhermore, A, B and S are pairwise disjoint subsets of U.

Therefore, by the sum principle, $|U| = |A| + |B| + |S|$, which in turn implies

$$|S| = |U| - |A| - |B|$$

But we already know $|U|$, it is 2^8. We just calculated $|A|$, it is 8 (substituting $L = 8$). And $|B| = 1$; there is only one string with exactly 0 ones; it is the all zero string. Thus, $|S| = 2^8 - 8 - 1$.

(e) *How many* 8*-bit strings are there with* **exactly** 2 *ones?* This is an example of a problem where we will use the sum-rule to reduce to subsets whose answers we already know. Once again, let S be the set we are interested in counting. The key idea is the following definition

$$S_i = \text{Set of all } 8\text{-bit strings with } \textbf{exactly } 2 \text{ ones } \textbf{and} \text{ the } \textit{first} \text{ one is in position } i$$

Note:

- $S_8 = \varnothing$; if an 8-bit string has two ones then the first one has to be in position 1 to 7.
- The same reason implies $S = S_1 \cup S_2 \cup \cdots S_7$.
- Finally, $S_i \cap S_j = \varnothing$ whenever $i \neq j$. Why? Suppose not. Suppose there is a bit string x in both A_i and A_j. Since $i \neq j$, without loss of generality, we may assume $i < j$. Now, $x \in A_j$ implies $x_j = 1$ but $x_i = 0$ since j is the first time we see a 1. However, $x \in A_i$ implies $x_i = 1$. Contradiction. Thus the sets are disjoint.

Therefore,

$$|S| = |S_1| + |S_2| + \cdots + |S_7|$$

Now how big is $|S_i|$? The first i bits each has one choice – the first $(i-1)$ bits have to be all zero and the ith bit is one. What remains? Well, in the remaining $(8-i)$ bits, we must have exactly one 1. How many ways can I have an $(8-i)$-bit string with exactly one 1? We just did it two bullet points ago; the answer is $|S_i| = (8-i)$. Thus,

$$|S| = \sum_{i=1}^{7} (8-i) = 7 + 6 + \cdots + 2 + 1 = \frac{7 \times 8}{2} = 28$$

Exercise: *How many L-bit strings have* **exactly two** 1*s?*

Exercise: *After doing the above, try answering how many* 8*-bit strings have* **exactly three** 1*s? Apply the same principle as we did for exactly two* 1*s.*

What if we can split S into two sets A and B, but they are not disjoint? No worries – we already know how to handle this. Answer: Inclusion-Exclusion. So, if we can find two sets A and B such that (a) $S = A \cup B$, and (b) $|A|$, $|B|$, and $|A \cap B|$ are *easier* to figure out, then we can find $|S|$ by using the *baby version of the inclusion-exclusion formula*.

$$|S| = |A \cup B| = |A| + |B| - |A \cap B|$$

Let's look at some examples.

Examples.

(a) *ATM PIN machine numbers.* A four-digit ATM pin (note the first digit can be 0) doesn't allow the first three numbers or the last three numbers to be the same. *How many PINs are disallowed?*

Let us denote the set of disallowed PINs as S. So a four digit number is in S if its first three numbers are the same or the last three numbers are the same. By the way the question is designed, we can see what the two sets should be of which S is the union of.

Indeed, define A to the set of four digit numbers whose first three digits are the same. Define B to the set of four digit numbers whose last three digits are the same.

Let's first see both $|A|$ and $|B|$ are easy to calculate by the product principle. Indeed, construct an element of A digit-by-digit. The first digit has 10 choices. But once this is fixed, the second and third digit have 1 choice each. The last digit again has 10 choices. So, we get $|A| = 100$. Arguing similarly, we get $|B| = 100$ as well.

How about the intersection $A \cap B$? Are A and B disjoint sets? No. The intersection is the set of four digit numbers whose first three numbers are the same ***and*** the last three numbers are the same. That is .... $A \cap B$ is the set of four digit numbers whose ***all*** digits are the same.

Thus, $|A \cap B|$ is also easy to calculate using the product principle. The first digit has 10 choices. But then it fixes everything else. Thus, $|A \cap B| = 10$.

The number of disallowed passwords, therefore, is

$$|S| = |A| + |B| - |A \cap B| = 100 + 100 - 10 = 190$$

✍

Exercise: *What is the number of four digit numbers which either have the first two digits the same or the last two digits the same?*

(b) *Divisibility. How many numbers between* 1 *and* 100 *(both inclusive) are divisible by either* 2 *or* 3 *or both?*

Let us denote as S the set of numbers between 1 and 100 (both inclusive) which are divisible by either 2 or 3.

Let A be the set of numbers between 1 and 100 (both inclusive) which are divisible by 2. Let B be the set of numbers between 1 and 100 (both inclusive) which are divisible by 3. We see that $S = A \cup B$, and thus,

$$|S| = |A| + |B| - |A \cap B|$$

What is $|A|$? It is the number of even numbers, and this number is $\lfloor 100/2 \rfloor = 50$ (do you see why?) What is $|B|$? It is given by $\lfloor 100/3 \rfloor = 33$.

What about $A \cap B$? What is this set? It is the set of all numbers between 1 and 100 (both inclusive) which are divisible by both 2 and 3. That is, these numbers are divisible by 6. There are $\lfloor 100/6 \rfloor = 16$ of them. Thus, the answer we are looking for is

$$|S| = |A| + |B| - |A \cap B| = 50 + 33 - 16 = 67$$

✍

Exercise: *How many numbers between* 1 *and* 100 *(both inclusive) are divisible by* 4 *or by* 6 *or both?*

✍

Exercise: *How many numbers between* 1 *and* 100 *(both inclusive) are divisible by* 2 *or by* 3 *but* **not** *both?*

Why should we stop at describing S as a union of only two sets? We can express S as a union of 3 sets A, B, C and we can apply the "toddler version" of the inclusion-exclusion formula. Which remember, states if $S = A \cup B \cup C$, then

$$|S| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

Using this, you can solve the following examples.

Examples.

(a) *How many numbers between* 1 *and* 100 *(both inclusive) are divisible by either* $2, 3$ *or* 5*?*

Again, let A be the set of numbers between 1 and 100 (both inclusive) which are divisible by 2; B be the set of numbers between 1 and 100 (both inclusive) which are divisible by 3, and C be the set of numbers between 1 and 100 (both inclusive) which are divisible by 5.

We want to figure out $|S|$ where $S = A \cup B \cup C$. We also know,

$$|A| = 50, \quad |B| = 33, \quad |C| = 20$$

The set $A \cap B$ is the set of numbers between 1 and 100 (both inclusive) which are divisible by 6. The set $A \cap C$ is the set of numbers between 1 and 100 (both inclusive) which are divisible by 10. The set $B \cap C$ is the set of numbers between 1 and 100 (both inclusive) which are divisible by 15. The set $A \cap B \cap C$ is the set of numbers between 1 and 100 (both inclusive) which are divisible by 30.

Thus,

$$|A \cap B| = 16, \quad |A \cap C| = 10, \quad |B \cap C| = 6, \quad |A \cap B \cap C| = 3$$

Thus,

$$|S| = 50 + 33 + 20 - 16 - 10 - 6 + 3 = 74$$