Let Us Count the Ways: Some Useful Combinatorics

Adrian She

In our analysis of the stable matching problem (and variants), we were asked to count the number of possible matchings (disregrading potential instabilities) between the possible "men" and "women" given as input. Mathematical tools from **combinatorics**, which have a deep connection with algorithm design and analysis, enable us to complete this task. Tools from combinatorics can often tell us whether or not a brute force algorithm for a problem is possible or not. In this article, we will overview a number of formulae and principles from combinatorics which enable us to count the size of a solution space.

1 Permutations and the Rule of Product

The first rule we will introduce is known as the **rule of product**, also sometimes called the **fundamental counting principle**.

Principle 1 (Rule of Product)**.** Let A, B be sets. If there are x ways to select an element from A, and y ways to select an element from B, then there are xy ways to select an element from $A \times B$.

For instance, suppose you are ordering dinner from a restaurant where you can only have one main dish and one dessert. Suppose there are 3 possible main dishes and 2 possible desserts, then will have $3 \times 2 = 6$ choices for your dinner. A **possibility tree** is a visual representation of such choices:

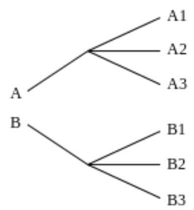


Figure 1: Illustration of a Possibility Tree, Source: Wikipedia

Question 1. If S is a set of size n, how many subsets of S are possible?

Solution 1. There are a number of ways to approach this problem. Let $S = \{x_1, ..., x_n\}$

- The first approach is to apply the product rule. To construct a subset of S, we make n choices, one for each element. The first choice C_1 is including or excluding x_1. The second choice C_2 is including or excluding x_2, and so on until we reach the n^{th} choice C_n, which is including or excluding element x_n. Each subset of S corresponds to an element of $C_1 \times ... \times C_n$. Since each set C_i has 2 elements (include or exclude), then by the rule of product, $|C_1 \times ... \times C_n| = \underbrace{2 \times ... \times 2}_{n \text{ times}} = 2^n$. Hence there are 2^n subsets of S.
- The second approach is proving by induction that there are 2^n possible subsets.
- Another approach is "encoding" every subset of S by a binary string of length n, and applying a product rule argument to show that there are 2^n subsets.

Next, we consider r-permutations, which count **orderings** of sets.

Definition 1. An **r-permutation** is an **ordering** of r elements in a set $S = \{x_1, ..., x_n\}$ **without repetition**.

Ordering means that arrangements (x_1, x_2) and (x_2, x_1) are distinct permutations of a set since the elements x_1, x_2 are put in a different order in each case.

For instance, suppose we have a race where there are 10 runners. Suppose we want to select a first, second, and third place finishers in the race. A choice of finishers is a 3-permutation of the set of 10 runners, since ordering matters in this case.

The rule of product be used to compute the number of r-permutations of a set.

Definition 2 (Factorial)**.** The number $n!$ (n **factorial**) is defined as $n! = n \times n-1 \times ... \times 1$. By convention, $0! = 1$.

For instance, $4! = 4 \times 3 \times 2 \times 1 = 24$.

Formula 1 (Number of r-permutations)**.** The number of r-permutations of a set S of n elements is $\frac{n!}{(n-r)!}$

Proof. An r-permutation of a set S can be thought of a list L of r distinct elements from S. Hence, to construct the list L, we can select among n elements of the set S for its first element, among $n-1$ elements for its second element, and so on, until we select among $n-(r-1)$ elements for its r^{th} element. Hence by the rule of product, the number of r-permutations of S is:

$$n(n-1)...(n-(r-1)) = \frac{n!}{(n-r)(n-(r+1))...(1)} = \frac{n!}{(n-r)!}$$

□

The following examples illustration how to apply Formula 1 in computing the number of r-permutations of a set.

Question 2. If we have n "men" and n "women", how many possible matchings between the men and the women are there?

Solution 2. Let $\{m_1, ..., m_n\}$ denote the set of "men". "Man" m_1 may be paired with one of n "women", m_2 may be paired with one of $n-1$ "women", and so on until m_n is paired with the only remaining "woman". Hence the number of possible matchings is $n(n-1)...1 = n!$.

Question 3. If we have n_m "men" and n_w "women" where $n_m \geq n_w$, how many possible matchings between the men and the women are there wherein each woman has a partner?

Solution 3. Let $\{w_1, ..., w_{n_w}\}$ denote the set of "women". "Woman" w_1 may be paired with one of n_m "men", w_2 may be paired with one $n_m - 1$ "men", and so on until w_{n_w} is paired with one of $n_m - n_w + 1$ men since $n_w - 1$ men have to taken by this point. Hence, the number of possible matchings is

$$n_m(n_m-1)...(n_m-(n_w-1)) = \frac{n_m!}{(n_m-n_w)!}$$

Note that when $n_m = n_w$, the formula reduces to $n_m!$ since $0! = 1$.

2 Combinations and Bi/Multinomial Coefficients

In permutations, the ordering matters. Now we consider the case where ordering does not matter.

Definition 3. An **r-combination** is a **subset** of r elements of a set $S = \{x_1, ..., x_n\}$.

Crucially, arrangements (x_1, x_2) and (x_2, x_1) are considered **the same combination** of elements in a set since the elements which they contain are the same, **regardless of ordering**.

For instance, suppose we have an election where there are 10 candidates. Suppose we want to elect 3 people to a council. (For simplicity, suppose there are no special positions such as the chair of the council.) A choice of councils is a 3-combination of the set of 10 candidates, since ordering does not matter in this case and that all matters is the candidates elected.

There is a special symbol to denote the number of r-combinations of a set of size n.

Definition 4. The **binomial coefficient** $\binom{n}{r}$ denotes the number of r-combinations of a set of size n.

For small n, r, we can compute the number of r combinations by listing out all the possibilities. For instance, the following are all possible 3-combinations of a set of size 4: $\{A, B, C, D\}$:

$$ABC \quad ABD \quad ACD \quad BCD$$

Hence $\binom{4}{3} = 4$. However, we would prefer to have a general formula.

Formula 2 (Number of r-combinations)**.** The number of r-combinations of a set S of n elements is $\dfrac{n!}{r!(n-r)!}$

Proof. Each r-combination of a set S $\{x_1, ..., x_r\}$ has $r!$ possible r-permutations since there are $r!$ possible ways to reorder the elements in $\{x_1, ..., x_r\}$. Because there are $\frac{n!}{(n-r)!}$ possible r-permutations, then:

$$\frac{n!}{(n-r)!} = r!\binom{n}{r}$$

Now dividing both sides by $r!$ yields

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

□

Combinations come up quite naturally in algorithmic, among other, problems.

Question 4. How many edges does an undirected, complete graph $G = (V, E)$ of n vertices have?

Proof. Let V denote the set of vertices in the graph and by assumption, $|V| = n$ since there are n vertices. Every subset of two vertices in V corresponds to an edge in G. Hence, there are $\binom{n}{2} = \frac{n(n-1)}{2}$ edges in total. □

Question 5. Suppose we draw a poker hand, which includes 5 cards, from a deck of 52. How many hands are possible?

Proof. We will need to compute the number of 5-combinations in a set of 52, which is $\binom{52}{5} = 2,598,960$. □

Question 6. If we have n_m "men" and n_w "women" where $n_m \geq n_w$, how many possible matchings between the men and the women are there wherein each woman has a partner?

Proof. Choose a subset of n_w men who we can match to women. Then there are $\binom{n_m}{n_w}$ possible such subsets, and $n_w!$ possible matchings for each subset. By the rule of product, we have

$$\binom{n_m}{n_w} n_w! = \frac{n_w! n_m!}{(n_m - n_w)! n_w!} = \frac{n_m!}{(n_m - n_w)!}$$

possible matchings. This agrees with our previous computation. □

2.1 Properties of Binomial Coefficients

Binomial coefficients satisfy some interesting properties. Firstly, they satisfy a recursive formula.

Formula 3 (Recurrence Relation for Binomial Coefficients)**.** $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$

We can prove this formula without resorting to direct calculation. This is known as the method of "bijective proof".

Proof. To construct a subset of size k in $S = \{x_1, ..., x_n\}$, we can either include the first element x_1 or exclude it.

In the case we include x_1, we can adjoin it to a subset of size $k-1$ from the $n-1$ elements we haven't looked at yet. This yields $\binom{n-1}{k-1}$ possible subsets.

In the case we exclude x_1, we can construct a subset of size k from the remaining $n-1$ elements, yielding $\binom{n-1}{k}$ subsets in total.

Since all $\binom{n}{k}$ possible subsets of size k can be formed by either including x_1 or excluding it, we have:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

□

In the above proof, we have implicitly applied the rule of sum, which we will state below for completeness:

Principle 2 (Rule of Sum)**.** Let X and Y be disjoint sets. If there are x ways to select an element from X and y ways to select an element from Y, then there are $x + y$ ways to select an element from $X \cup Y$ (X or Y).

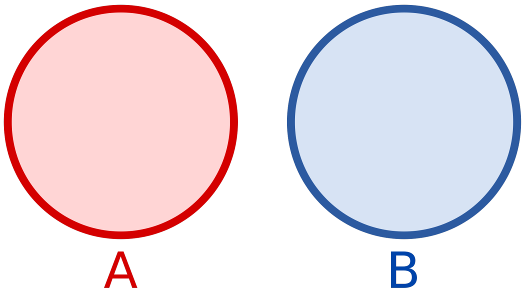


Figure 2: Illustration of a Disjoint Sets, Source: Wikipedia

The rule of sum differs from the rule of product, although both involve combining two sets A, B in some way. For the rule of sum, we select from a single set $A \cup B$ (eg. one dish which is either a main dish or a dessert), whereas for the rule of product, we select one element from each of A and B, which corresponds to an element of $A \times B$ (eg. a main dish and a dessert.)

Next, the following are additional properties of binomial coefficients. As we have done above, think about why these properties are true using the definition as binomial coefficients as counting subsets of a set, instead of direct calculation.

- $\binom{n}{0} = \binom{n}{n} = 1$
- $\sum_{k=0}^{n} \binom{n}{k} = 2^n$
- $\binom{n}{k} = \binom{n}{n-k}$. This is a symmetry property.

The name binomial coefficient comes from the use of binomial coefficients in the **binomial theorem** and the **binomial distribution** of probabilities. For instance, a binomial distribution tells you the chance of seeing k heads in n independent coin flips. Finally, binomial coefficients can be arranged nicely in **Pascal's Triangle**.

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1

Figure 3: Illustration of a Pascal's Triangle, Source: Wikipedia

2.2 Multinomial Coefficients

Binomial coefficients can also be generalized into **multinomial coefficients**.

Definition 5. The **multinomial coefficient** $\binom{n}{k_1, k_2, ..., k_m}$ denotes the ways to seperate n objects into m bins where k_1 objects are first in the first bin, k_2 objects in the second, and so on until the m^{th} bin. Note that we should have $k_1 + k_2 ... + k_m = n$.

Formula 4. $\binom{n}{k_1, k_2, ..., k_m} = \frac{n!}{k_1!k_2!...k_m!}$

Proof. Choose k_1 of the n objects to be labelled in the first bin. Then choose k_2 of the remaining $n - k_1$ objects to be put in the second bin. By the rule of product, we have

$$\begin{aligned}\binom{n}{k_1, k_2, ..., k_m} &= \binom{n}{k_1}\binom{n-k_1}{k_2}...\binom{n-k_1-...-k_{m-1}}{k_m} \\ &= \frac{n!}{k_1!(n-k_1)!}\frac{(n-k_1)!}{k_2!(n-k_1-k_2)!}...\frac{(n-k_1-...-k_{m-1})!}{k_m!0!} \\ &= \binom{n!}{k_1!k_2!...k_m!}\end{aligned}$$

□

Note that $\binom{n}{k} = \binom{n}{k, n-k}$, so the multinomial coefficient is indeed a generalization of the binomial coefficient. The following are some examples of things we can do with multinomial coefficients:

Question 7. How many possible "words" can be formed by ordering letters in "REARRANGEMENT" are there?

Solution 4. REARRANGEMENT is a 13 letter word with 3 Rs, 3 Es, 2 As, 2 Ns, and one G, M, T each. Since we can choose 3 of 13 positions where the Rs are placed. Next we can choose 3 of 10 remaining positions where an E is placed, and so on. Next we can choose 2 of 7 remaining positions where A is placed, 2 of 5 remaining positions where N is placed, until we have 3 remaining positions for G, M, T. This yields the multinomial coefficient:

$$\binom{13}{3,3,2,2,1,1,1} = \frac{13!}{3!3!2!2!} = 43243200$$

Question 8. If we have n_H hospitals, each with $s_1, ..., s_{n_H}$ residency slots which accomodate $s_1 + ... + s_{n_H} = n_R$ residents, how many possible ways are there to assign residents to hospitals? (Assume we only care about which hospital we assign a resident to.)

Solution 5. This is exactly what the multinomial coefficient was built for! In this case, our "objects" are our residents. Next, we have n_H "bins" to which we can assign positions for n_R residents satisfying $s_1 + ... + s_{n_H} = n_R$. This means that the number of arrangements [1] is

$$\binom{n_R}{s_1, ..., s_{n_H}} = \frac{n_R!}{s_1!...s_{n_H}!} = \frac{n_R!}{\prod_{i=1}^{n_H} s_i!}$$

3 Summary

We have only scratched the surface of important combinatorial tools which can be used in analysis of algorithms. In this document, we have covered:

- The rule of product and the rule of sum.
- Different types of permutations and combinations, and ways to compute them.

There are many interesting classes of combinatorial numbers, including numbers which describe partitions of sets such as the Stirling and Bell numbers. There are also other interesting asymptotic approximations for combinatorial numbers which can be explored.

Overall, techniques in combinatorics are varied and having a solid grasp of the material covered in this document will aid your abilities in analyzing algorithms.

[1] Recall that $\sum_{i=1}^{n} a_i$ is shorthand for the sum $a_1 + a_2 + ... + a_{n-1} + a_n$. Similarly, $\prod_{i=1}^{n} b_i$ is shorthand for the product $b_1 \times b_2 \times ... \times b_{n-1} \times b_n$

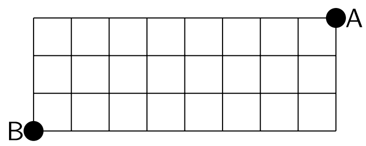
4 Practice Problems

1. Let X be a set with n elements. How many possible divisions of X into sets A, B are possible such that $A \subset X, B \subset X$ and $A \cap B = \emptyset$? (Note: $A \cup B = X$ does not have to be true!)

2. Footblog asks a user to create an alphanumeric password (using upper and lower case English letters and digits 0-9) when a user registers on the system. If the password must be between 5 and 8 characters long, how many passwords are possible?

3. In the card game of bridge, a deck of 52 cards is divided such that each player gets 13 cards in their hand. How many possible ways are there to distribute the cards?

4. Consider the following grid graph, which perhaps represents the street layout of a particular neighbourhood:



A person at "A" must travel left or down at each intersection of the grid graph. How many paths from A to B are possible? (The grid is 8 units long and 3 units wide.)

5. A "multiset" is a set which can contain more than one copy of the same element. For instance $\{A, B, B\}$ is a multiset, but $\{A, B, B\}$ and $\{B, A, B\}$ are the same multiset.

   (a) Suppose we have a set S of size n. How many multisets of size r can be made from S? (Hint: Find a way to relate this quantity to combinations.)

   (b) Using the formula derived in (a), how many non-negative solutions are possible to $x_1 + x_2 + x_3 + x_4 = 10$? (That is, solutions where $x_1, x_2, x_3, x_4 \geq 0$)

5 Solutions

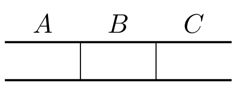
1. For each element of X, we can either choose to include it in A, include it in B, or include it in neither set. Hence, we have 3 choices for what to do with each element of X. Since there are n elements, there are 3^n possible divisions of X into A, B satisfying the given properties.

2. Our character set has 62 elements. Hence a password of length n from the given character set has 62^n elements. Hence, if our password is between 5 and 8 letters then $62^5 + 62^6 + 62^7 + 62^8 \approx 2.22 \times 10^{14}$ passwords are possible.

3. Separating the cards into 4 decks of 13 yields the multinomial coefficient $\binom{52}{13,13,13,13}$. This can be calculated as
$$\binom{52}{13}\binom{39}{13}\binom{26}{13}\binom{13}{13} = \frac{52!}{(13!)^4} \approx 5.36 \times 10^{28}$$

4. Each path through the graph can be thought of as a string of "W"s and "S"s indicating whether A has gone south or west at a particular grid corner. Since A must go west 8 times and south 3 times, we need to count the number of distinct rearrangements of WWWWWWWWSSS. This is $\binom{11}{3} = \binom{11}{8} = 165$.

5. (a) A multiset of size r can be encoded by "balls and bars". For concreteness' sake, suppose $n = 3$. Then distributing r indistinguishable balls among the partitions defined below yields a multiset of size r:



Since we have $n - 1$ possible "bars" (because our original set has n elements) and r possible "balls", we have $r + n - 1$ objects in total we need to arrange. As each ball and each partition is indistinguishable, we have

$$\frac{(r+n-1)!}{r!(n-1)!} = \binom{n+r-1}{r} = \binom{n+r-1}{n-1}$$

total possible multisets. You may refer to `https://en.wikipedia.org/wiki/Multiset#Counting_multisets` for more information about this formula.

(b) Each solution to $x_1+x_2+x_3+x_4 = 10$ places 10 "balls" in some way into 4 partitions (where some partitions can be empty.) By the previous formula, we have $\binom{13}{10} = \binom{13}{4} = 715$ total solutions.

References

[1] Susanna Epp. Discrete Mathematics and Applications.

[2] Nick Harvey. CPSC 221 Lecture Slides.

[3] George Knill et al. Mathpower 12, Western Edition.