Combinatorics

Phillip M Alday

April 2011

1 Basic Idea

As we saw in our brief exploration of set theory, counting is a non trivial problem. One thing that is particularly important to count are arrangements of elements of a set. We call such arrangements PERMUTATIONS. Together with COMBINATIONS, they provide an important tool in experimental design and evaluation as well as in understanding the range of possibilities for a given linguistic structure.

For the purposes of the following discussion, we will assume all sets are finite, unless explicitly stated otherwise. The sole exception to this is the set of natural numbers, which we will often use as the superset of the sets we are manipulating.

2 Definitions

There are several ways to think of permutations that are relevant for linguistic use.[1] Each of the following definitions is more or less equivalent.

> A PERMUTATION is:
>
> - an ordering
>
> - a way of arranging a certain (sub)set of elements
>
> - way of putting together elements from different collections in a particular order

For example, the permutations of $S = \{a, b, c\}$ are:

$$abc, acb, bac, bca, cab, cba$$

We can also take »shorter« permutations. The permutations of S taken two at a time are:

$$ab, ba, ac, ca, ba, cb$$

---

[1]The concept »permutation« has a few different, albeit closely related, meanings in abstract mathematics. Broadly speaking, it is a mapping of a finite set onto itself. We will consider this idea in more detail later: it actually provides the basis for a relatively »new« theory in Chomskyan linguistics.

Now, unlike in »regular«[2] sets, we allow for repeated elements in the sets we wish to permute.[3] The permutations of $R = \{a, a, b\}$ are:

$$aab, aba, baa$$

Notice that the number of distinct permutations has decreased compared to a set with three distinct elements.

Sometimes we wish to consider the possibilities where order doesn't matter: ab and ba don't count as separate items. At this point, we could be using a combintation.

> A COMBINATION is:
>
> - a way of »combining« elements from a set of several sets to make a new set, where order does not matter
>
> - a way to select a subset of a given size from a set
>
> - a way of mixing distinct (sub)sets together

For example, $S = \{a, b, c\}$ only has one combination when taken in its entirety: abc. When we take combinations of two elements of S, there are only three: ab, bc, ac. We note right away that there are fewer combinations of S than permutations. In fact, this holds generally:

Lemma 2.1. *The number of permutations is always greater than or equal to the number of combinations.[4]*

3 Computation

In many linguistic applications, we care more about the number of possible permutations or combinations than actually enumerating them. Since enumeration quickly becomes problematic for all but the most trivial cases, it would be convenient to be able to calculate the number of possibilities without enumeration.

Before we begin calculating how many permutations or combinations are possible given a particular input set, we need to state the Fundamental Counting Principle. Given a choice of X, consisting of n subchoices, the total number of options available is the product of the number of options available in each subchoice.

Postulate 3.1 (Fundamental Counting Principle)**.** *Given a ways of performing an action A and b ways of performing an action B, there are $a \cdot b$ ways of performing both actions. More generally, given r actions and n_i ways for performing an action i, there are $n_1 \cdot n_2 \cdot \ldots \cdot n_{r-2} \cdot n_{r-1} \cdot n_r$ all of the actions together.*

> **Comprehension Check:** In the cafeteria, there are three types of sauce for schnitzel, three different non-salad side dishes and four different salads. You are allowed schnitzel with one sauce, one side dish and one salad for a »Linguist's Special«. How many different meals are possible as part of the »Linguist's Special«?

---

[2]'üblich'

[3]Such sets are also called MULTISETS and an element that appears m times is said to have MULTIPLICITY m.

[4]The »or equal to« part is important here: the trivial example of a one-element set has only one possible permutation and one possible combination.

Up until now, we've have only allowed each element in the input set to occur once in the output permutation or combination. That is, we've largely viewed combinations as a subset of the input set, and permutations as an ordering on that set. Now, there are also instances where we wish to allow input elements to be repeated in the output. An example of this is drawing colored balls from a bag. If I put the balls back in the bag after each draw, then repetitions can occur. If I don't replace the balls, then they can't occur (unless there are duplicates of some elements in the bag like our R example earlier – but more on that later). The formulae for calculating the number of permutations of n objects taken k at a time ($P(n, k)$) both with and without repetition, as well as the analogous formulae for the number of combinations ($C(, n, k)$) are presented in Table 1. The derivations for these formulae from the counting principle were presented in class and provide good examples of the application of the counting principle!

Table 1: Formulae for calculating the number of permutation or combinations of n objects taken k at a time. Remember, $n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1$ and we define $0! = 1$. (There is only one way to arrange an empty set.)

Situation	$P(n, k)$	$C(n, k)$
no repetitions	$\dfrac{n!}{(n-k)!}$	$\dbinom{n}{k} = \dfrac{n!}{k!(n-k)!}$
repetitions	n^k	$\dbinom{n+k-1}{k} = \dfrac{(n+k-1)!}{k!(n-1)!}$

The number of choices of k objects from n without repetition, that is $C(n, k) = \binom{n}{k}$, is a calculation that comes up in many areas of math, science and linguistics. It is called the BINOMIAL COEFFICIENT for reasons we will not go into here. It is a very important formula to learn and recognize! (Expect an exam question!)

Occasionally, we must also deal with the permutations of sets with repeated elements, like $R = \{a, a, b\}$ from above. Not allowing for repetitions here (in the sense of »returning« an element to the set for future selection), we still encounter the issue that there is a repeated element in the set. So, we have repeats of a sort built in. The permutation aab is the same, regardless of whether the left most a corresponds to the first a in R or the second.

A more concrete example is: how many different distinct strings can you build from the letters of »MISSISSIPPI«? We note that 'S' and 'I' are each repeated four times and 'P' twice. We can use this fact to modify our previous calculation for permutations. A letter l repeated r times, represents $r!$ different ways of rearranging the order of the duplicates. That is, we have $r!$ times too many permutations when we use the base $n!$ calculation for total number of permutations. So, for each repeated letter l, we divide by $r!$.

More generally, suppose there is a string S of length n, such that there are r_1 identical letters of one type, r_2 of another, r_i of the i-th repeated type. Then the

number of distinct permutations of S is:

$$(1) \qquad \frac{n!}{r_1! \cdot r_2! \cdot \ldots \cdot r_{i-1}! \cdot r_i!}$$

So for our »MISSISSIPPI« example, there are

$$\frac{11!}{4!4!2!} = 34650$$

distinct permutations, certainly not a number we would want to have to enumerate.

> **Comprehension Check:** Why doesn't it make sense to discuss this as a special case for combinations?[5]

4 Permutations as Mappings

Sometimes, we wish to look at permutations more abstractly as a reordering of any sequence (ordered set). Without loss of generality, we assume that the base sequence is the first n natural numbers: $1, 2, 3, \ldots, n$.[6] We can express a permutation on this sequence so:

$$(2) \qquad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & \cdots & n \\ p_1 & p_2 & p_3 & p_4 & p_5 & \cdots & p_n \end{pmatrix}$$

where the p_i's are the position that each element is moved to. More concretely,

$$(3) \qquad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$$

would take the sequence (1234) and map it to (2413). Now, what happens if we take this same permutation (3) and apply to (2413)? We get yet another sequence: (4321). Repeating the process yet again, we get (3142). Doing it one more time, gives us our original sequence: (1234). So, we have something like this:

$$(4) \qquad (1234) \mapsto (2413) \mapsto (4321) \mapsto (3142) \mapsto (1234)$$

This type of circular path is called a CYCLE. We can express permutations in terms of cycles; in this case, we only need one: $(1, 3, 4, 2)$. That notation reads »one goes to three, three goes to four, four goes to two, two goes to one« and expresses both the individual movements we need at each step as well as the path that each element follows during the cycle. Cycle notation is equivalent to matrix notation; from here on out, we'll be using cycle notation.

So, we're examining permutations as mappings, and like most mappings, we can compose them (apply them successively), as we saw when applied the permutation in (3) repeatedly to its own output. We write the permutations to be applied so: $\pi_1 \pi_2 \ldots \pi_n$, that is, like multiplication. There is one caveat though: we read the

---

[5]More precisely: As long as we're choosing taking a permutation of the entire set and not choosing some $k < n$. In that case, we have to place the additional restriction that $k >= r_i$ for all i, or there is a potential special case. But don't worry about that for now.

[6]If you're having trouble understanding why we can make this assumption without loss of generality, then think of these as subscripts: $x_1, x_2, \ldots, x_n$.

product from right to left, like with function composition. So, given the product $(1,3)(1,4)(1,2)$, we first move one to two and two to one, **then** the new one (what was two at the beginning) to four and four to one, etc. Using the base sequence (1234) again, we do the process like so:

$$(5) \qquad\qquad (1234) \overset{(1,2)}{\mapsto} (2134) \overset{(1,4)}{\mapsto} (4132) \overset{(1,3)}{\mapsto} (3142)$$

We see that this is actually our the permutation from (3)! We performed a complex permutation using only transposition (swapping of two elements). This is an example of an important fact in abstract algebra:

Theorem 4.1. *Every permutation can be written as the product of transpositions.*

Every mathematics student learns this in his first year, yet this idea is what is forms the basis of Merge in Chomsky's Minimalist Program [**?**], where it was considered a major innovation.

Further reading on groups: »Group Think« by Steven Strogatz.[**?**] (In ILIAS!)

References