Module 2.2: Counting Words and Strings

MCIT Online - CIT592 - Professor Val Tannen

Lecture Notes

Counting words

Recall that the English alphabet uses 26 letters: 5 vowels and 21 consonants.

Problem. How many words of length 3 can be formed with letters of the English alphabet?

Answer. Such a word has 3 positions. Example: $\underline{h}\ \underline{a}\ \underline{h}$.

Therefore each such word can be constructed in 3 steps:

(1) Put a letter in the first position: can be done in 26 ways.

(2) Put a letter in the second position: 26 ways.

(3) Put a letter in the third position: 26 ways.

By the *multiplication rule*, the answer is $26 \times 26 \times 26 = 26^3$.

Counting strings of bits

Recall that there are two **bits** (*bi*nary digi*ts*): 0 and 1.

Problem. How many strings of bits of length n are there?

Answer. Such a string has n positions. Example: $\underline{1}\ \underline{0}\ \cdots \underline{1}\ \underline{1}$.

Therefore each such string can be constructed in n steps:

(1) Put a bit in the first position: can be done in 2 ways.

. . .

(n) Put a bit in the n'th position: 2 ways.

Again, by the multiplication rule the answer is $2 \cdot 2 \cdots 2 = 2^n$.

Counting sequences, in general

We saw that there are 26^3 **words** of length 3 and 2^n **strings** of bits.

Similarly, let's count **words** of length 6 made only of consonants: 21^6.

Also, let's count **words** of length m made only of vowels: 5^m.

In general, using the multiplication rule, we count

n^ℓ **sequences** of length ℓ made of elements from a set of size n.

ACTIVITY : All Distinct Subsets

We have seen in the lecture segment "Counting subsets" how to apply the multiplication rule to counting the total number of distinct subsets of a set. Let us revisit that here.

Given a set $\{a_1, \ldots, a_n\}$, we construct a subset S by considering, for each element in $\{a_1, \ldots, a_n\}$, whether or not to include that element in S. This can be done in n steps by deciding, in step i, whether or not to include the element a_i in S.

ACTIVITY : All Distinct Subsets (Continued)

For example, suppose that $n = 4$ and we are constructing a subset of $\{a_1, a_2, a_3, a_4\}$. One possible way to do this is as follows.

Step 1 : Decide not to include a_1.

Step 2 : Decide to include a_2.

Step 3 : Decide to include a_3.

Step 4 : Decide not to include a_4.

The resulting subset is $S = \{a_2, a_3\}$.

ACTIVITY : All Distinct Subsets (Continued)

Constructing a subset in this way is analogous to constructing a binary string of length n. Putting a 1 in the i^{th} position corresponds to including the element a_i in S, and putting a 0 in the i^{th} position corresponds to *not* including the element a_i in S.

In this way, the subset $\{a_2, a_3\}$ of $\{a_1, a_2, a_3, a_4\}$ that we constructed above corresponds to the 4-bit binary string 0110.

This way of representating subsets gives a *one-to-one correspondence* between subsets of $\{a_1, \ldots, a_n\}$ and binary strings of length n.

Activity : All Distinct Subsets (Continued)

What does this correspondence tell us about the two problems of counting subsets of a set of cardinality n and counting binary strings of length n?

It tells us that these two counting problems are precisely equivalent!

There are exactly 2^n distinct subsets of $\{a_1, \ldots, a_n\}$, and there are exactly 2^n distinct binary strings of length n.

Question:

What subset of $\{a_1, a_2, a_3, a_4\}$ corresponds to the string 1001?

In the video, there is a box here for learners to put in an answer to the question above. As you read these notes, try it yourself using pen and paper!

ACTIVITY : All Distinct Subsets (Continued)

What subset of $\{a_1, a_2, a_3, a_4\}$ is represented by the string 1001?

Answer:

Since the first and fourth bits in the string are 1, and the other bits are 0, the subset is $\{a_1, a_4\}$.