COMBINATORICS

"How to count without actually counting"

My notes taken from a series of lectures given by

Prof. Rochelle Liebowitz

Wheaton College

- The Product Rule
- The Addition Rule
- Samples, Permutations, Combinations, Selections
- More on Permutations/Combinations
- Binomial Theorem, Pascal's Triangle
- Principle of Inclusion–Exclusion
- Derangements
- Set Theory
- Pigeonhole Principle and its generalization
- Catalan Numbers

The Product Rule

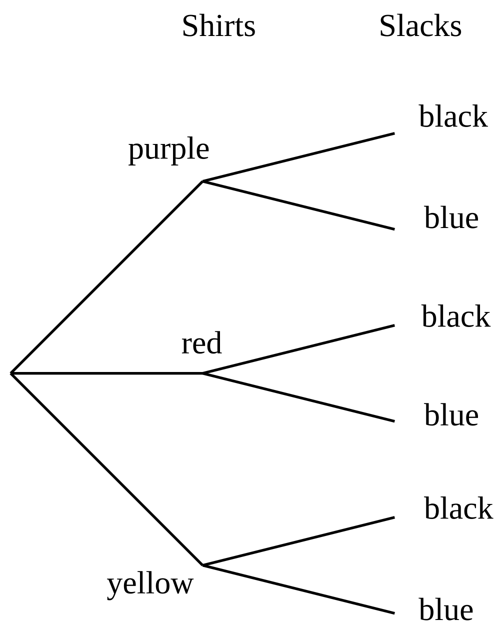
Example: 3 shirts — purple, red, yellow

2 slacks — black, blue

Outfit = shirt and slacks

How many possible outfits are possible?

Solution 1: Possibility (Probability) Tree



Answer:

6 outfits

Solution 2:

Task 1: choose a shirt (3 ways)

AND THEN

Task 2: choose slacks (2 ways)

$$3 \cdot 2 = 6$$

<u>THE PRODUCT RULE</u>: Suppose that a procedure can be broken down into two tasks. If there are m ways to do the first task and n ways to do the second task after the first task has been done, then there are $m \cdot n$ ways to do the procedure.

Example: "two all beef patties, special sauce, lettuce, cheese, pickles, onions, on a sesame seed bun"

"Have it your way!"

The basic burger is two all beef patties on a sesame seed bun. The other five choices are all optional. How many different ways can you place an order?

Task 1: special sauce (yes or no)

Task 2: lettuce (yes or no)

etc.

Total number is $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 2^5 = 32$ ways.

The Addition (Sum) Rule

<u>THE ADDITION (SUM) RULE</u>: If there are m ways of performing task A and there are n ways of performing task B, and A and B are mutually exclusive, then there are $m + n$ ways of performing either task A or task B.

Example: In the original version of BASIC, numeric variable names must be a single letter or a letter followed by a digit. How many variable names are possible?

Solution: Case 1: single letter — 26

Case 2: letter followed by digit — $26 \cdot 10$

Total number = $26 + 26 \cdot 10 = 286$

Samples, Permutations, Combinations, Selections

	order important	order not important
repetition allowed	r–sample (licence plates)	r–selection (balloons)
no repetition	r–permutation (standing in line)	r–combination (committee membership)

(1) r–sample

example: License plates are made up of 5 letters

$$26 \cdot 26 \cdot 26 \cdot 26 \cdot 26 = 26^5$$

THEOREM: The number of r–samples of n distinct objects is $n \cdot n \cdot n \cdot \ldots \cdot n = \boxed{n^r}$

(In the above example $n = 26$ and $r = 5$.)

Notation: $n!$ (n factorial)

Def: For $n \geq 1$, $n! = n(n-1)(n-2)\ldots3\cdot2\cdot1$.

$0! = 1$

Example: $8! = 8\cdot7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1 = 40{,}320$

(2) <u>r–permutation</u>

Example: Pick 3 people out of 8 to stand in line at a ticket booth (that is, to be first, second and third in line).

Solution: 8 ways to fill first position, 7 ways to fill second position, 6 ways to fill third position.

$$8 \cdot 7 \cdot 6 = 336$$

Note that $8 \cdot 7 \cdot 6 = \frac{8!}{5!}$

THEOREM: The number of r–permutations of n distinct objects, denoted $P(n,r)$ is

$$n(n-1)(n-2)\ldots(n-r+1) = \boxed{\frac{n!}{(n-r)!}}$$

(3) <u>r–combination</u>

Example: Choose 3 people from Ann, Beth, Chuck, Dan, and Ed to serve on a committee.

Solution: We simply list off all 10 possible committees.

ABC, ABD, ABE, ACD, ACE, ADE, BCD, BCE, BDE, CDE

NOTATION: the number of r–combinations of n distinct objects is denoted by $C(n,r)$.

Question: How are permutations and combinations related in this problem?

Think of choosing 3 people out of the 5 people to stand in line at a ticket booth.

Choosing 3 people out of 5 to be first, second, third in line = task 1 (choosing 3 out of 5) AND THEN task 2 (ordering the 3 chosen people).

That is, $P(5,3) = C(5,3) \cdot 3!$ So, $C(5,3) = \frac{P(5,3)}{3!} = 10$

In general, $C(n,r) = \boxed{\frac{n!}{(n-r)! \cdot r!}}$

(4) *r*–selection

Example: Balloons come in 5 colors: red, blue, silver, purple, green. How many ways can we buy a dozen balloons?

(How many 12–selections of 5 distinct types?)

Solution: Fill out a menu.

Red	Blue	Silver	Purple	Green	
XXX	X	XXX		XX	Place 12 X's in
		XXX			the menu

What is shown is 3 red, 1 blue, 6 silver, 0 purple, and 2 green.

We can represent this symbolically as

XXXô Xô XXXXXXô ô XX

A 12–selection of 5 distinct types corresponds to an arrangement of 12 X's and 4, (5–1) ô's.

We have 16 slots to fill. Choose 12 of these slots for the X's in C(16,12) ways. Thus the number of 12 selections of 5 distinct types is C(16,12).

In general, the number of *r*–selections of *n* distinct types is $C(r+n-1, r)$. Note that some books define this to be $C(r+n-1, n-1)$. Are these equal?

NOTES:

1. For *r*–permutations and *r*–combinations of *n* distinct objects, $r \le n$.
2. For *r*–samples and *r*–selections of *n* distinct objects (for selections, think of types), *r* may be greater than *n*.

Example:

License plates consist of 5 letters chosen from A, B, C. Number of license plates = number of 5-samples of 3 different objects = 3^5.

Example:

Previous balloon problem ($r = 12$, $n = 5$)

More on Permutations/Combinations

Ordering a set with repeated elements

ex: ordering with no repeated elements

How many ways can you arrange the 8 letters of the word COMPUTER?

Task 1: choose the first letter (8 choices)

Task 2: choose the second letter (7 choices)

etc.

Total number of ways is thus P(8,8) = 8! = 8 · 7 · 6 · 5 · 4 · 3 · 2 · 1 = 40320

ex: ordering with repeated elements

How many ways can you arrange the 7 letters of the word SEEKERS?

Previous method does not work since the 3 E's and 2 S's are indistinguishable. Instead…

Task 1: choose 3 slots for the E's = C(7,3)

Task 2: choose 2 slots for the S's = C(4,2)

Task 3: choose 1 slot for the R = C(2,1)

Task 4: choose last slot for the K = C(1,1)

Total number of ways is C(7,3) · C(4,2) · C(2,1) · C(1,1)

$$= \frac{7!}{4!\cdot 3!} \cdot \frac{4!}{2!\cdot 2!} \cdot \frac{2!}{1!\cdot 1!} \cdot 1 = \frac{7!}{3!\cdot 2!\cdot 1!\cdot 1!} = 420$$

The number of ways to order n objects where:

n_1 are of type 1 and indistinguishable from each other,

n_2 are of type 2 and indistinguishable from each other,

…

n_k are of type k and indistinguishable from each other,

and $n = n_1 + n_2 + \ldots + n_k$ is

$$\frac{n!}{n_1!\cdot n_2!\cdot \ldots \cdot n_k!}$$

This is sometimes called a multinomial coefficient.

Set Theory

ex: $S = \{a,b,c,d,e\}$

List of 2–subsets of S (10 of them)

$\{a,b\}$ $\{a,c\}$ $\{a,d\}$ $\{a,e\}$

$\{b,c\}$ $\{b,d\}$ $\{b,e\}$

$\{c,d\}$ $\{c,e\}$

$\{d,e\}$

Note: Number of 2–subsets of S = C(5,2)

In general, number of r–subsets of an n–set is C(n,r)

Definition: The power set of a set S, denoted by P(S) = the set of all subsets of S.

Note that for all sets S, we have

$S \subseteq S$, and {} = $\subseteq S$.

That is, both all of S and the empty set are both subsets of S.

Examples:

1) $S = \{a\}$ P(S) = { $,S\}$

2) $S = \{a,b\}$ P(S) = { $,\{a\},\{b\},S\}$

3) $S = \{a,b,c\}$ P(S) = { $,\{a\},\{b\},\{c\},\{a,b\},\{a,c\},\{b,c\},S\}$

Theorem: If $|S| = n$, then $|P(S)| = 2^n$

Proof:

Task 1: choose or do not choose a_1 (2 choices)

Task 2: choose or do not choose a_2 (2 choices)

…

Task n: choose or do not choose a_n (2 choices)

Total number of subsets = $2 \cdot 2 \cdot \ldots \cdot 2 = 2^n$.

Q.E.D.

Binomial Theorem, Pascal's Triangle

Notation:

$$C(n,r) = \binom{n}{r} = C_{n,r} = {}_nC_r = \frac{n!}{r!(n-r)!}$$

We say "*n* choose *r*".

Build a triangular table called Pascal's triangle:

$$\begin{array}{ccccccccc} & & & & \binom{0}{0} & & & & \\ & & & \binom{1}{0} & & \binom{1}{1} & & & \\ & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & \\ & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & \\ \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} \end{array}$$

....

That is:

$$\begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1 \end{array}$$

....

Do you see any patterns?

Do they continue if you extend the triangle?

Can you prove them?

Each row is symmetric

Theorem: $C(n,r) = C(n,n–r)$

Proof 1 (algebraic):

$$C(n,r) = \frac{n!}{r!\cdot(n-r)!} \qquad C(n,n-r) = \frac{n!}{(n-r)!\cdot(n-(n-r))!} = \frac{n!}{(n-r)!\cdot r!}$$

Q.E.D.

Proof 2 (combinatorial):

Consider n–set $S = \{1,2,3,\ldots,n\}$

Create a 1 – 1 correspondence between the r–subsets of S and the $n–r$ subsets of S by pairing each set with its complement.

List of r–subsets of S		List of $(n–r)$–subsets of S
$\{1,2,3,\ldots,r\}$	↔	$\{r+1,r+2,\ldots,n\}$
$\{1,3,4,\ldots,r,r+1\}$	↔	$\{2,r+2,\ldots,n\}$
	…	
A	↔	$\overline{A}$

The number of r–subsets of S is $C(n,r)$ and the number of $(n–r)$–subsets of S is $C(n,n–r)$.

Q.E.D.

Theorem: The sum of the numbers in the n^{th} row of Pascal's triangle is 2^n $(n = 0,1,2,\ldots)$

Proof (combinatorial):

Let $S = \{1,2,3,\ldots,n\}$. Thus $|S| = n$ and $|P(S)| = 2^n$. But this number is made up of:

the number of 0–subsets of $S = C(n,0)$

the number of 1–subsets of $S = C(n,1)$

the number of 2–subsets of $S = C(n,2)$

…

the number of $(n–1)$–subsets of $S = C(n,n–1)$

the number of n–subsets of $S = C(n,n)$.

The sum of all of these is exactly the n^{th} row of Pascal's triangle.

Q.E.D.

Each number is the sum of the two numbers diagonally above it.

Theorem: $C(n+1,r) = C(n,r) + C(n,r-1)$

Proof (combinatorial):

Let $S = \{1,2,3,\ldots,n,n+1\}$.

The total number of r–subsets of S is $C(n+1,r)$.

This total can be divided into two cases:

Case 1: r–subsets that contain the element 1. In this case, choose $r-1$ elements from the set $\{2,3,4,\ldots,n+1\}$ and there are $C(n,r-1)$ of these.

Case 2: r–subsets that do not contain the element 1. In this case, choose r elements from the set $\{2,3,4,\ldots,n+1\}$ and there are $C(n,r)$ of these.

This exhausts all cases.

Q.E.D.

We now have all machinery to develop the Binomial theorem:

$$(a+b)^n = \sum_{k=0}^{n} C(n,k)\cdot a^{n-k}\cdot b^k$$

Inclusion – Exclusion Principle

Let S be a set with ï Sï = N. Let $c_1, c_2, \ldots, c_k$ be a collection of properties (or conditions) satisfied by some (or all, or none) of the elements of S.

For $1 \le i \le k$, let $N(c_i)$ = the number of elements of S that satisfy property c_i. These elements may or may not satisfy other properties c_j where $i \ne j$.

Let $N(c_i c_j)$ = the number of elements of S that satisfy both c_i and c_j. For multiplication, read “and” or “intersection”. These elements may or may not satisfy other properties c_k, $k \ne i,j$.

Let $N(\overline{c_i}) = N - N(c_i)$ = the number of elements that do <u>not</u> satisfy c_i.

Let $N(\overline{c_i}\ \overline{c_j})$ = the number of elements of S that do <u>not</u> satisfy c_i <u>and</u> do <u>not</u> satisfy c_j.

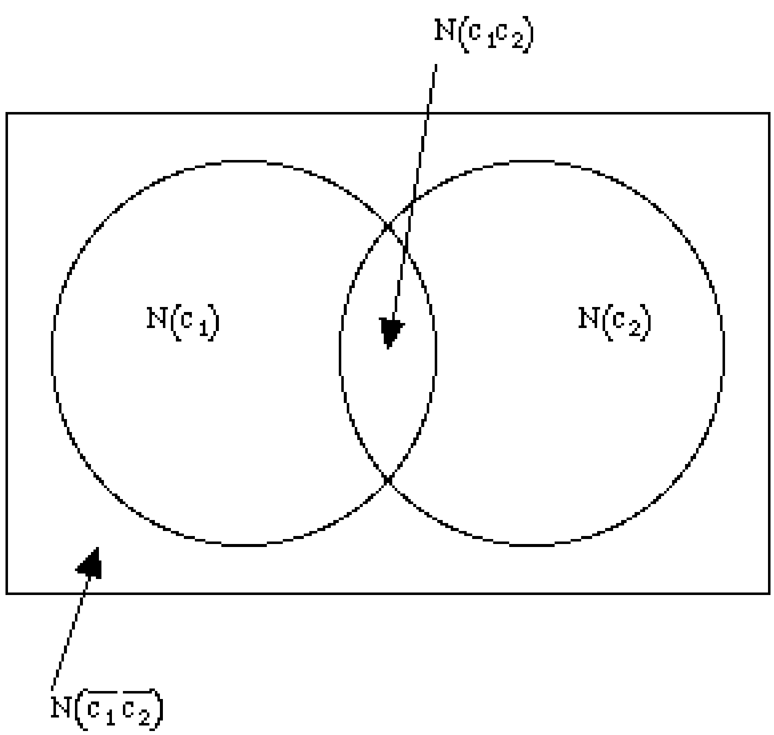
example:

S = people in this class = 25

c_1 = people in this class with blue eyes = 10

c_2 = people in this class who live north of route 9 = 14

c_1c_2 = people in this class who both live north of route 9 and have blue eyes = 5



$$N(\overline{c_i}\ \overline{c_j}) = N - N(c_1) - N(c_2) + N(c_1c_2)$$

So in our example, the number of people in this class who do not have blue eyes and do not live north of route 9 is 25 – 10 – 14 + 5 = 6.

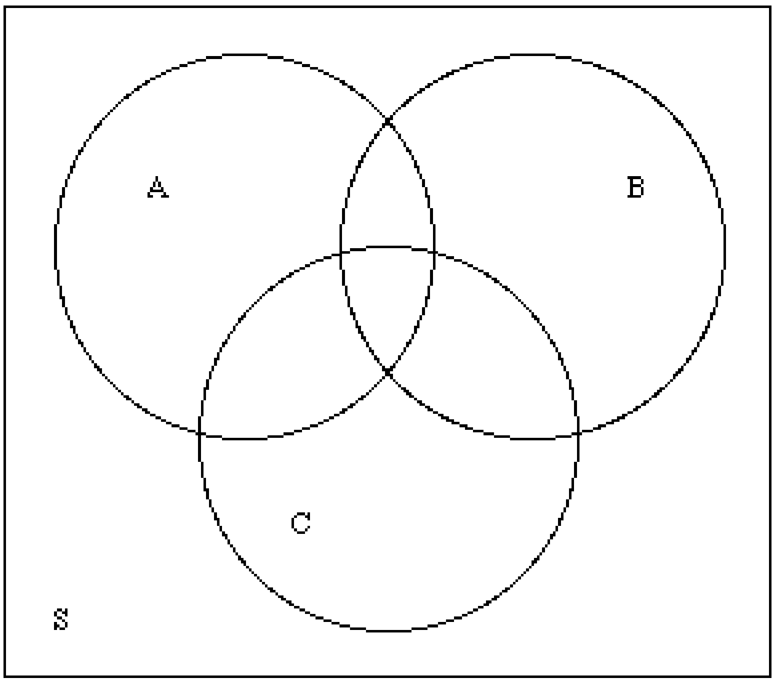
Example: S = people in this class

A = people in this class who live north of route 9

B = people in this class who have blue eyes

C = people in this class who play a musical instrument

Given: $|S| = 40$, $|A| = 10$, $|B| = 15$, $|C| = 12$, $|A \cap B| = 4$, $|A \cap C| = 6$, $|B \cap C| = 8$ and $|A \cap B \cap C| = 3$, fill in all the regions in the Venn diagram below.



This process now generalizes to the Principle of Inclusion–Exclusion.

Theorem: Let S be a set with $|S| = N$, and properties c_i, $1 \le i \le t$, satisfied by some of the elements of S. Then the number of elements of S that satisfy <u>none</u> of the conditions c_i, $1 \le i \le t$, is:

$$N\left(\overline{c_1}\,\overline{c_2} \ldots \overline{c_t}\right) = N - \sum_{i=1}^{t} N(c_i) + \sum_{1 \le i \le j \le t} N(c_1 c_2) - \sum_{1 \le i \le j \le k \le t} N(c_1 c_2 c_3) + \ldots + (-1)^t N(c_1 c_2 \ldots c_t)$$

Corollary: Same hypothesis as the previous theorem. The number of elements of S that satisfy <u>at least one</u> of the properties c_i, $1 \le i \le t$, is:

$$N(c_1+c_2+\ldots+c_t) = \sum_{i=1}^{t} N(c_i) - \sum_{1 \le i \le j \le t} N(c_1c_2) + \sum_{1 \le i \le j \le k \le t} N(c_1c_2c_3) - \ldots + (-1)^{t-1}N(c_1c_2\ldots c_t)$$

For “+”, read “or” or “union”.

Example:

How many (positive) numbers ≤ 396 are relatively prime to 396?

Solution:

Since $396 = 2^2 \cdot 3^2 \cdot 11$, we can rephrase the problem: how many (positive) numbers ≤ 396 are not divisible by 2, 3, and 11.

Let c_1 = property that a (positive) number ≤ 396 is divisible by 2.

Let c_2 = property that a (positive) number ≤ 396 is divisible by 3.

Let c_3 = property that a (positive) number ≤ 396 is divisible by 11.

What we need is $N(\overline{c_1}\ \overline{c_2}\ \overline{c_3})$. By the inclusion–exclusion principle this equals

$$N - N(c_1) - N(c_2) - N(c_3) + N(c_1c_2) + N(c_1c_3) + N(c_2c_3) - N(c_1c_2c_3)$$

$$= 396 - 396/2 - 396/3 - 396/11 + 396/6 + 396/22 + 396/33 - 396/66$$

$$= 396 - 198 - 132 - 36 + 66 + 18 + 12 - 6 = 120.$$

Derangements

Example:

In how many ways can we arrange the numbers 1,2,3,4,5 so that 1 is <u>not</u> in the first position, 2 is <u>not</u> in the second position, … , 5 is <u>not</u> in the fifth position?

Such an arrangement is called a <u>derangement</u> of 1,2,3,4,5. The total number of derangements of 1,2,3,4,5 is symbolized by D_5. So, for example:

2 5 1 3 4 is a derangement,

2 5 1 <u>4</u> 3 is not a derangement since 4 is in the fourth position.

Let c_i (i = 1,2,3,4,5) be the property that an arrangement of 1,2,3,4,5 has i in the i^{th} position. So what we want is $D_5 = N(\overline{c_1}\ \overline{c_2}\ \overline{c_3}\ \overline{c_4}\ \overline{c_5})$. We first calculate N = 5! Next we note that $N(c_1)$ =

4! This is so since we can place 1 in the first position and place the other four in any order. In a similar way $N(c_1c_2) = 3!$, $N(c_1c_2c_3) = 2!$, $N(c_1c_2c_3c_4) = 1!$, and $N(c_1c_2c_3c_4c_5) = 0!$ Considering all possible combinations we have

$$D_5 = C(5,0)\cdot 5! + C(5,1)\cdot 4! + C(5,2)\cdot 3! + C(5,3)\cdot 2! + C(5,4)\cdot 1! + C(5,5)\cdot 0!$$

that is,

$$5!\,(1/2! - 1/3! + 1/4! - 1/5!) = 44$$

Note that the probability of getting a derangement of 1,2,3,4,5 is 44/5!

In general, $D_n = n!\,(1/2! - 1/3! + 1/4! - 1/5! + \ldots + (-1)^n/n!)$ and the probability of getting a derangement of 1,2,3,…,n is $D_n/n!$ or approximately $1/e \approx .36788\ldots$

Pigeonhole Principle

If n pigeons fly into m pigeonholes and n > m, then some hole contains 2 or more pigeons.

Examples:

Among any group of 367 people, there must be at least two people with the same birthday.

Every 5–card poker hand has at least two cards of the same suit.

In any set of 4 integers, 2 of these integers must have the same remainder when divided by 3.

Generalization of the Pigeonhole Principle:

If n pigeons fly into m pigeonholes and n > km for some positive integer k, then some pigeonhole contains at least k + 1 pigeons.

Examples:

In a class of 29 students, if the only grades given were A,B,C,D,F, then at least six students got the same grade.

At a conference with 2500 participants, at least 7 people will have the same birthday.

If there are 78 students in the auditorium, then at least 3 will have last names beginning with the same letter.

<u>EXAMPLE</u>:

While on a 4 week vacation, Bob will play at least one set of tennis each day, but he plans to not play more than 40 sets total during the vacation. No matter how he distributes his sets during the 4 weeks, there will be a consecutive span of days during which he plays exactly 15 sets.

Proof:

Let a_i = total number of sets Bob has played from the start of vacation to the end of the i^{th} day, i = 1,2,3,…,28. Then,

$$1 \leq a_1 < a_2 < \ldots < a_{28} \leq 40 \quad (*)$$

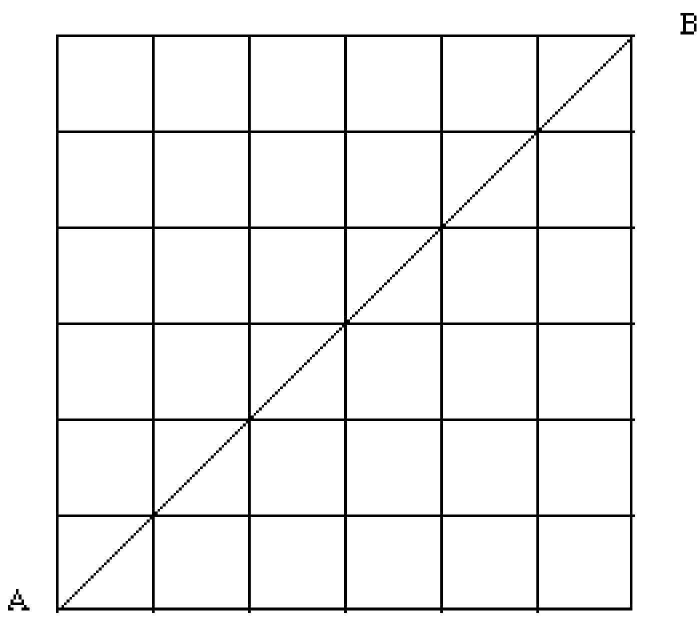
and

$$a_1 + 15 < a_2 + 15 < \ldots < a_{28} + 15 \leq 55 \quad (**)$$

(*) and (**) show 56 numbers (28 + 28) which are less than or equal to 55. Therefore, by the Pigeonhole Principle, at least two of them must be equal. So $a_j = a_i + 15$. Q.E.D.

Catalan Numbers

Problem 1:



The previous picture represents a 6 × 6 grid. You want to trace out a path from A to B using only the horizontal and vertical lines. In addition, you may not go above the diagonal line. You may touch it but not cross it. How many paths are possible in an n × n grid?

Problem 2:

DEFINITION: A balanced n – string is a string of n left parentheses and n right parentheses satisfying the property:

(*) As you read from left to right the number of ('s so far is ≥ the number of)'s so far.

Examples:

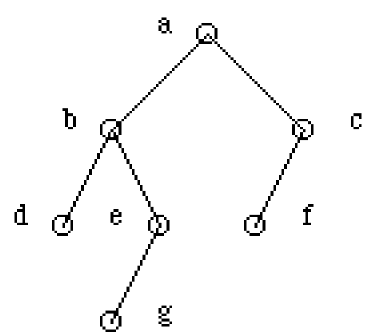
	(( ) ( )) ( )	
left so far	1 2 2 3 3 3 4 4	
right so far	0 0 1 1 2 3 3 4	This is a balanced 4 – string

	(( ) ( ))) (( )	
left so far	1 2 2 3 3 3 3 4 5 5	
right so far	0 0 1 1 2 3 4 4 4 5	
	↑	This is an unbalanced 5 – string since it violates property

(*). Note that it still has 5 (and 5).

The question is how many balanced n – strings are there?

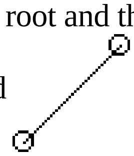
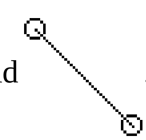
Problem 3:



This is an example of a rooted, ordered, binary tree having 7 nodes.

DEFINITIONS:

rooted: the top node is the root and the tree is built downward.

ordered: having a left child  is different from having a right child  .

binary: every node has at most 2 edges leading down from it (at most 2 children)

How many rooted, ordered, binary trees with n nodes are there?

All of the above problems are equivalent and can be described by using the letters L and R. Let C_n = the number of balanced n – strings, etc. These numbers are called the Catalan numbers. They begin 1, 2, 5, 14, … In order to develop a formula for the Catalan numbers we note that C_n = the number of ways of ordering n L's and n R's with no restrictions – the number of unbalanced n – strings (which we will call U_n). That is,

$$C_n = C(2n,n) - U_n.$$

The problem is now to calculate U_n. Let's take unbalanced n – strings and perform the following algorithm to them:

1) Find the first position where the number of R's is > the number of L's. Call this position r.

2) From position r + 1 to n, switch L's to R's and R's to L's.

Example:

unbalanced 6 – string:

	L	L	R	L	R	R	R	L	L	R	L	R
L so far	1	2	2	3	3	3	3	4	5	5	6	6
R so far	0	0	1	1	2	3	4	4	4	5	5	6
							↑ position r = 7					

Using the algorithm:

L L R L R R R R R L R L

Notice that the output is 5 L's and 7 R's.

Now let's see if this process is reversible.

We will input (for n = 5) 4 L's and 6 R's. Find the first position where the number of R's is > the number of L's and then switch L's and R's.

L L L L R R R R R R becomes L L L L R R R R R L

or perhaps

L L R R L L R R R R becomes L L R R L L R R R L

or perhaps

L R L R R R L L R R becomes L R L R R L R R L L.

Notice that in every case, the randomly constructed string of 4 L's and 6 R's turns into an unbalanced 5 – string. What we have shown is that the number of strings with 4 L's and 6 R's is equal to the number of unbalanced 5 – strings.

In general, the number of unbalanced n – strings, U_n, equals the number of strings with n – 1 L's and n + 1 R's, i.e. C(2n,n – 1).

Therefore,

$$
\begin{aligned}
C_n &= C(2n,n) - C(2n,n-1) \\
&= \frac{(2n)!}{n!n!} - \frac{(2n)!}{(n+1)!(n-1)!} \\
&= \frac{(2n)!}{(n+1)n!n!}
\end{aligned}
$$

which we can write as

$$
\boxed{C_n = \frac{C(2n,n)}{n+1}}
$$

Starting with n = 0, the first few Catalan numbers are:

1, 1, 2, 5, 14, 42, 132, …

Two last examples with Catalan numbers are:

1. Given 2n people, all of different heights, in how many ways can we form 2 lines of n people such that:

    a) as we go from left to right, heights increase

    b) the i^{th} person in line 1 is shorter than the i^{th} person in line 2?

2. DEFINITION: A triangulation of a convex polygon with n + 2 vertices is a set formed from n – 1 diagonals which do not intersect in the interior of the polygon but only at the vertices, and which divide the interior of the polygon into n triangles.

    How many triangulations of a convex polygon are there having n + 2 vertices?