November 14, 2017

2 - Strings and Binomial Coefficients

William T. Trotter

trotter@math.gatech.edu

Basic Definition

Let n be a positive integer and let $[n] = \{1, 2, \ldots, n\}$. A sequence of length n such as $(a_1, a_2, \ldots, a_n)$ is called a **string** (also a **word**, an **array** or a **vector**).

The entries in a string are called **characters**, **letters**, **coordinates**, etc. The set of possible entries is called the **alphabet**.

Examples

01001010001011001101 - a bit string

201002211001020 – a ternary string

abcacbaccbbaaccbabaddbbadcabbd – a word from a four letter alphabet.

NHZ 4235 – a Georgia auto license plate

I love mathematics (really)!! - a word from an alphabet with 59 letters – upper and lower cases, spaces and punctuation.

Notation for Strings

When displaying a string, commas are often used to avoid ambiguity. For example
345334354647 is a string of length 12 from the alphabet [9].

34, 53, 3, 43, 54, 64, 7 is a string of length 7 from the alphabet [99]

And many people like to enclose a string in parentheses or brackets. For example,

(34, 533, 4354, 647) is a string of length 4 from the alphabet [9999].

Notation for Strings (2)

But keep in mind that a string is a function, so the string (2, 5, 8, 11, 14) is the function $f:[5] \rightarrow [22]$ defined by the rule $f(n) = 3n - 1$.

Often a function of this type is written with subscripts, so this same sequence could be written as $(a_1, a_2, a_3, a_4, a_5)$ with $a_n = 3n - 1$.

Arrays in Computer Languages

Our string (2, 5, 8, 11, 14) could be defined as an array. Here's a code snippet to accomplish this.

```
int  a[6];
for (i = 1; i <= 5; i++) {
  a[i] = 3· i - 1;
}
```

Note In most computer languages, arrays begin with coordinate 0 so declaring the array as a[6] creates storage for entries a[0] through a[5]. We just don't use a[0] in our example.

A Basic Principle of Enumeration

Observation If a project can be considered as a sequence of n tasks which are carried out in order, and for each i, the number of ways to do Task i is m_i, then the total number of ways the project can be done is the product:

$$m_1 \cdot m_2 \cdot m_3 \cdot \ldots \cdot m_n$$

Consequences

Fact The number of bit strings of length n is 2^n.

Fact The number of words of length n from an m letter alphabet is m^n.

Fact The number of Georgia license auto license plates is $26^3 10^4$.

Permutations – Repetition Not Allowed

Examples

12 7 8 6 4 9 11	Yes
X y a A D 7 B E 9	Yes
5 b 7 2 4 9 A 7 6 X	No

Fact The number of permutations of length n from an m letter alphabet is: $P(m, n) = m\,(m-1)\,(m-2)\,\ldots\,(m-n+1)$.

Language $P(m, n)$ is the number of permutations of m objects taken n at a time.

How to Answer a Question

Question How many permutations of 68 objects taken 23 at a time?

Answer P(68, 23)

Comment In almost all situations, I want you to stop right there and leave it to the dedicated reader to determine exactly what the value of P(68, 23) turns out to be. After all, this is just arithmetic. However, if you're really curious, P(68, 23) turns out to be:

20732231223375515741894286164203929600000

Permutations and Combinations

Contrasting Problems

Problem 1 A group of 250 students holds elections to identify a class president, a vice-president, and a treasurer. How many different outcomes are possible.

Problem 2 A group of 250 students holds elections to select a leadership committee consisting of three persons. How many different outcomes are possible?

Permutations and Combinations (2)

Solutions

Problem 1 A group of 250 students holds elections to identify a class president, a vice-president, and a treasurer. How many different outcomes are possible.

Answer $P(250, 3) = 250 \cdot 249 \cdot 248$

Permutations and Combinations (3)

Solutions

Problem 2 A group of 250 students holds elections to identify a leadership committee consisting of three persons. How many different outcomes are possible?

Answer $C(250, 3) = (250 \cdot 249 \cdot 248)/(1 \cdot 2 \cdot 3)$

Note We read $C(250, 3)$ as the number of combinations of 250 objects, taken 3 at a time.

Binomial Coefficients (1)

In Line Notation

$C(38, 17) = P(38, 17)/17! = 38!/(21!\ 17!)$

Graphic Notation

$$\binom{38}{17}$$

Note We read this as "38 choose 17"

Binomial Coefficients (2)

Basic Definition

$$\binom{38}{17} = \frac{38!}{17!21!}$$

Note To compute this binomial coefficient, you have to do a lot of multiplication and some division. Maybe there is an alternative way??!!

Beware the dot, dot, dot notation!!!

Question What is the next term: 1, 4, 9, 16, 25 ?

Question What is the next term: 1, 1, 2, 3, 5, 8, 13?

Question What is the sum $1 + 2 + 3 + \ldots + 6$?

Question What is really meant by the definitions:

$$n! = n \cdot (n - 1) \cdot (n - 2) \cdot \ldots 3 \cdot 2 \cdot 1$$

$$P(m, n) = m \cdot (m - 1) \cdot (m - 2) \cdot \ldots \cdot (m - n + 1)$$

A Better Way

Observation Rather than writing 1, 4, 9, 16, 25, …
be explicit and write:

$$a_n = n^2$$

Observation Rather than writing 1, 1, 2, 3, 5, 8, 13, …
be explicit and write:

$$a_1 = 1; \quad a_2 = 1; \quad \text{and when } n \geq 3,$$
$$a_n = a_{n-2} + a_{n-1}.$$

A Better Way (2)

Observation Rather than writing $1 + 2 + \ldots + 6$, say "the sum of the first six positive integers."

Observation An even better way:

Define $S_0 = 0$ and when $n \geq 1$, set $S_n = n + S_{n-1}$.
Then reference S_6.

Note The second alternative reflects a concept that we will study in depth.

A Better Way (3)

Definition $0! = 1$ and when $n > 1$,
$n! = n \cdot (n-1)!$

Example
$5! = 5 \cdot 4!$
$4! = 4 \cdot 3!$
$3! = 3 \cdot 2!$
$2! = 2 \cdot 1!$
$1! = 1 \cdot 0!$

Backtracking We obtain
$1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$ and $5! = 120$

A Better Way (4)

Definition $P(m, 1) = m$ and when $1 < n \leq m$,

$P(m, n) = (m - n + 1) \cdot P(m, n - 1)$.

Example

$P(7, 4) = (7 - 4 + 1) \cdot P(7, 3) = 4 \cdot P(7, 3)$

$P(7, 3) = (7 - 3 + 1) \cdot P(7, 2) = 5 \cdot P(7, 2)$

$P(7, 2) = (7 - 2 + 1) \cdot P(7, 1) = 6 \cdot P(7, 1)$

$P(7, 1) = 7$

Backtracking We obtain

$P(7, 2) = 6 \cdot 7 = 42$

$P(7, 3) = 5 \cdot 42 = 210$

$P(7, 4) = 4 \cdot 210 = 840$

Coding Basics

Declaration

```
int factorial ( int n);
```

Definition

```
int factorial { int n) {
   if (n == 0) return 1;
   else return (n) · factorial (n – 1);
}
```

Note In many languages, multiplication is written using *.

Coding Basics (2)

Declaration

```
int  permutation (int  m, int  n);
```

Definition

```
int  permutation {int  m, int  n) {
  if (n == 1) return  m;
  else return (m –n + 1) · permutation(m, n –1);
}
```

Coding Basics (3)

Observation

As a general rule, programmers prefer loops to a recursive call. Here's an example:

```
int permutation {int m, int n) {
    answer = 1;
    for (i = 1; i <= n; i++) {
        answer = answer* (m + 1 - i);
    }
    return answer;
}
```

Bit Strings and Subsets

Equivalent Problems

Problem 1 How many bit strings of length 38 have exactly 17 ones?

Problem 2 How many subsets of size 17 in a set of size 38?

Answer

$$C(38, 17) = P(38, 17)/17! = 38!/(21!\ 17!)$$

Basic Identities for Binomial Coefficients

Complements

$$C(m, n) = C(m, m - n) \text{ when } 0 \leq n \leq m.$$

Basis for Recursion

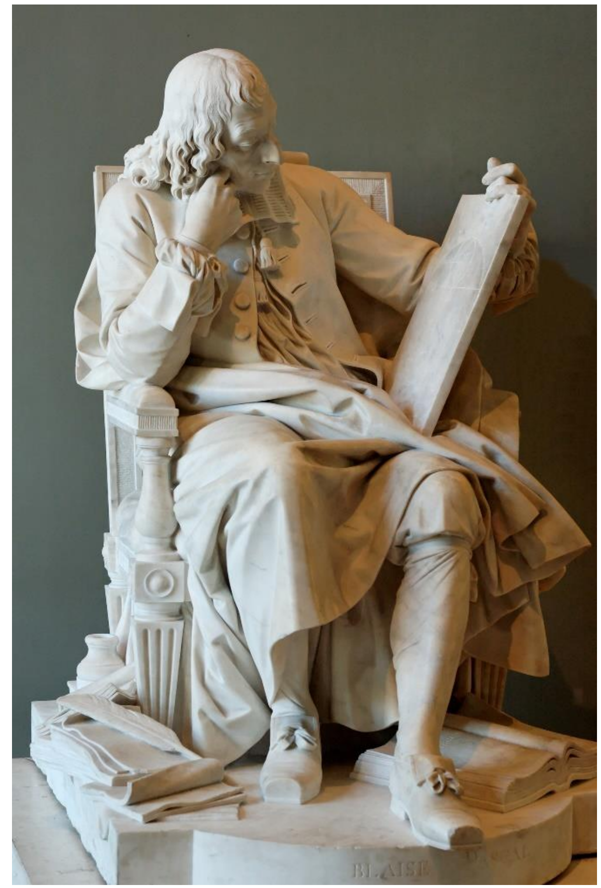
$$C(m, n) = C(m - 1, n) + C(m - 1, n - 1)$$
$$\text{when } 0 < n < m.$$

Pascal's Triangle

Observation Using $C(m, n) = C(m - 1, n) + C(m - 1, n - 1)$, the binomial coefficients can be displayed as follows:

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1
1 8 28 56 70 56 28 8 1

WTT and Pascal in the Louvre



Combinatorial Identities

Identity 1

$C(n, 0) + C(n, 1) + \ldots + C(n, n) = 2^n$ for all $n \geq 1$.

Identity 2

$C(m, n) = C(m - 1, n) + C(m - 1, n-1)$ when $0 < n < m$.

Identity 3

$C(n, 0)2^0 + C(n, 1)2^1 + \ldots + C(n, n)2^n = 3^n$ for all $n \geq 1$.

Remark In each case, you should try to explain why the identity holds by showing that the two sides count the same thing … only two different ways.

Enumerating Distributions

Basic Enumeration Problem Given a set of m objects and n cells (boxes, bins, etc.), how many ways can they be distributed?

Side Constraints

1. Distinct/non-distinct objects
2. Distinct/non-distinct cells
3. Empty cells allowed/not allowed.
4. Upper and lower bounds on number of objects in a cell.

Binomial Coefficients Everywhere

Foundational Enumeration Problem

Given a set of m identical objects and n distinct cells, the number of ways they can be distributed with the requirement that no cell is empty is

$$\binom{m-1}{n-1}$$

Explanation

$$A\ A\ A\ A\ A\ A \mid A\ A \mid A\ A\ A\ A \mid A\ A\ A\ A\ A\ A\ A \mid A \mid A\ A\ A$$

m objects, $m - 1$ gaps. Choose $n - 1$ of them. In this example, there are 23 objects and 6 cells. We have illustrated the distribution (6, 2, 4, 7, 1, 3).

Equivalent Problem

Restatement

How many solutions in positive integers to the equation:

$$x_1 + x_2 + x_3 + \ldots x_n = m$$

Given a set of m identical objects and n distinct cells, the number of ways they can be distributed with the requirement that no cell is empty is

$$\binom{m-1}{n-1}$$

Building on What We Know

Restatement

How many solutions in non-negative integers to the equation:

$$x_1 + x_2 + x_3 + \ldots x_n = m$$

Answer

$$\binom{m+n-1}{n-1}$$

Explanation Add n artificial elements, one for each variable.

Mixed Problems

Problem How many integer solutions in non-negative integers to the equation:

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 142$$

Subject to the constraints:

$$x_1, x_2, x_5, x_7 \geq 0; \quad x_3 \geq 8; \quad x_4 > 0; \quad x_6 > 19$$

Answer

$$\binom{\mathbf{119}}{\mathbf{6}}$$

Good = All - Bad

Problem How many integer solutions in non-negative integers to the equation:

$$x_1 + x_2 + x_3 + x_4 = 63$$

Subject to the constraints:

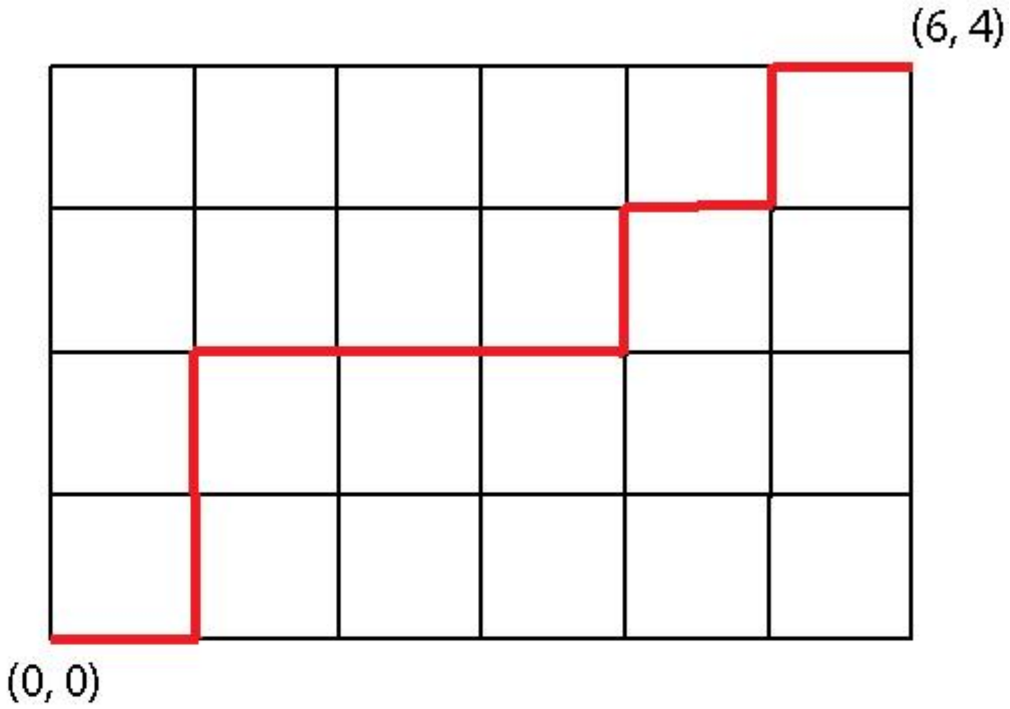
$$x_1, x_2 \geq 0; \quad 2 \leq x_3 \leq 5; \quad x_4 > 0$$

Answer

$$\binom{63}{3} - \binom{59}{3}$$

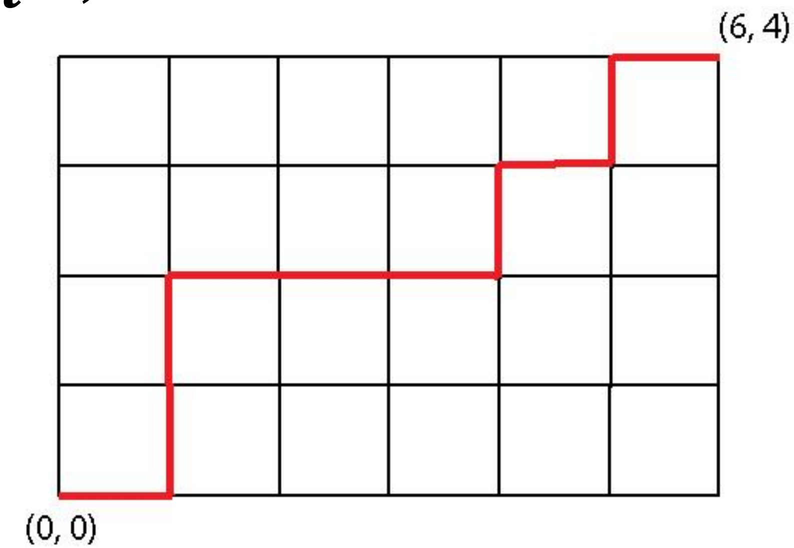
Lattice Paths (1)

Restriction Walk on edges of a grid. Only allowable moves are R (right) and U (up), i.e., no L (left) and no D (down) moves are allowed.



Lattice Paths (2)

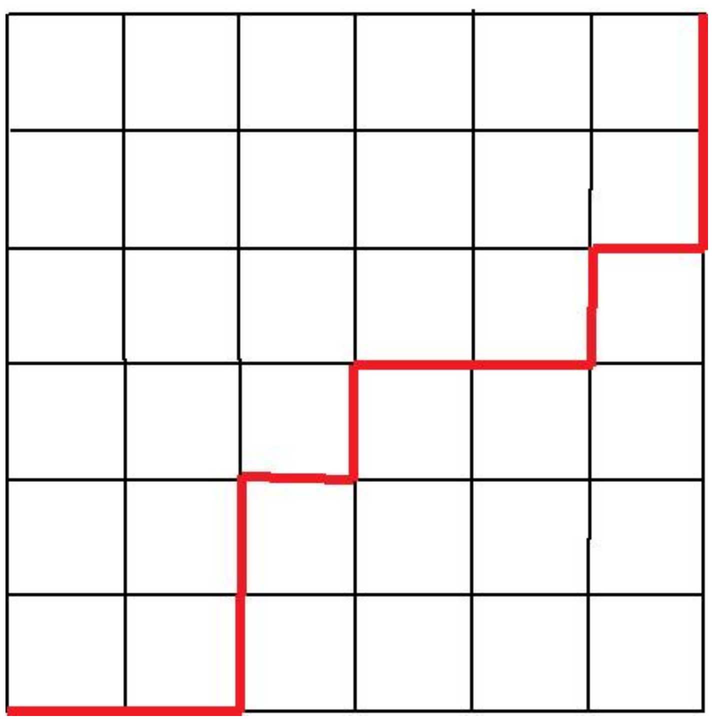
Observation The number of lattice paths from (0, 0) to (m, n) is $\binom{m+n}{m}$.



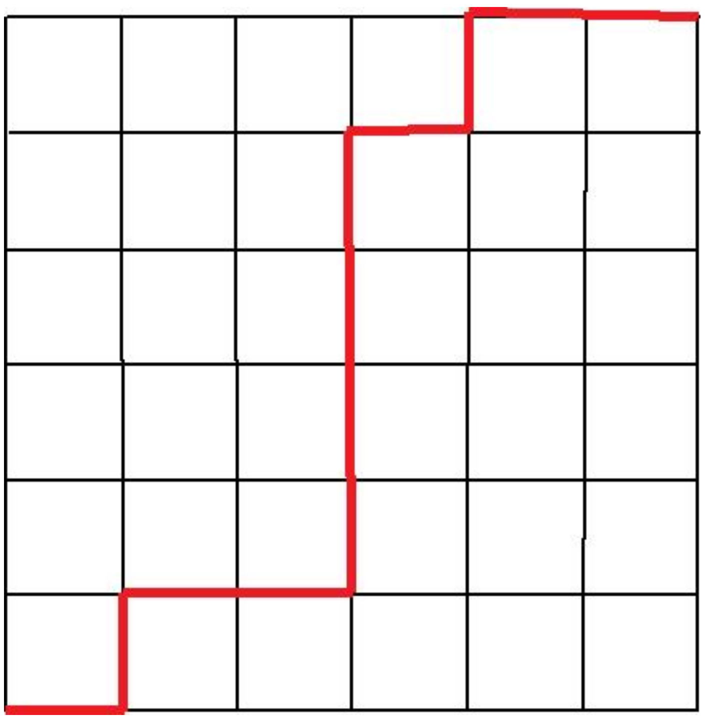
Explanation A lattice path corresponds to a choice of m horizontal moves in a sequence of m + n moves.
Here the choices are: RUURRRURUR

Lattice Paths – Not Above Diagonal

Question How many lattice paths from (0, 0) to (n, n) never go above the diagonal?



Good



Bad

Lattice Paths – Not Above Diagonal

Solution The number of lattice paths from (0, 0) to (n, n) which never go above the diagonal is the Catalan Number:

$$\frac{\binom{2n}{n}}{n+1}$$

Observation The first few Catalan numbers are:

1, 1, 2, 5, 14. What is the next one?

Parentheses and Catalan Numbers

Basic Problem How many ways to parenthesize an expression like:

$$x_1 * x_2 * x_3 * x_4 * \ldots * x_n$$

For example, when $n = 4$, we have 5 ways:

$$x_1 * (x_2 * (x_3 * x_4))$$
$$x_1 * ((x_2 * x_3) * x_4))$$
$$(x_1 * x_2) * (x_3 * x_4)$$
$$((x_1 * x_2) * x_3) * x_4$$
$$(x_1 * (x_2 * x_3)) * x_4$$

Can you verify that there are 14 ways when $n = 5$?

Multinomial Coefficients

Problem How many different arrangements of

AABBBCCCCCCDEEEEEEFFFFFFFF ?

Answer

$$\binom{26}{2,3,1,6,6,8} = \frac{26!}{2!\,3!\,1!\,6!\,6!\,8!}$$

Note Informally, this is known as the "MISSISSIPPI" problem.

Binomial and Multinomial Coefficients

Observation

When there are only two parts, a multinomial coefficient is just a binomial coefficient. So for example,

$$\binom{26}{7,19} = \binom{26}{7}$$

However You should only use the binomial notation in this case.

The Binomial Theorem

Theorem

$$(x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

Explanation

$$(x + y)^n = (x + y)(x + y)(x + y)(x + y) \dots (x + y)$$

From each of n terms, you either take x or y, so if k is the number of times you take y, then you take x exactly $n - k$ times.

Applying the Binomial Theorem

Theorem

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

Problem What is the coefficient of $a^{14}b^{18}$ in

$$(3a^2 - 5b)^{25}$$

Answer

$$\binom{25}{7} 3^7 (-5)^{18}$$

The Multinomial Theorem

Theorem

$$(x_1 + x_2 + x_3 + x_4)^n = \sum_{k_1+k_2+k_3+k_4=n} \binom{n}{k_1, k_2, k_3, k_4} {x_1}^{k_1} {x_2}^{k_2} {x_3}^{k_3} {x_4}^{k_4}$$

Problem What is the coefficient of $a^6 b^8 c^6 d^6$ in

$$(4a^3 - 5b + 9c^2 + 7d)^{19}$$

Answer

$$\binom{19}{2,8,3,6} 4^2 (-5)^8 9^3 7^6$$