Math 3U03 Combinatorics

Martin Bays

April 17, 2015

Notes from a course at McMaster University given by Martin Bays in Winter 2015.

The textbook for the course was "Introductory Combinatorics", 5th edition, by Richard A. Brualdi.

1 Nim

Nim: finitely many piles of coins; a move comprises removing a positive number of coins from a single pile; a player loses if they can't move.

Remark:
For any nim position P, either it can be won by the player with the move, or it can be won by the player without the move.

i.e. one of the two players has a "winning strategy", a way to play which guarantees a win.

The "nim sum", $n \oplus m$, of natural numbers n and m is the result of writing the binary expansions of n and m and "adding without carrying". (In computer science, this is called "XORing the bitstrings"; in many programming languages, it's written as "`n^m`".)

Theorem:
The player without the move can win from the Nim position with piles of sizes $n_1, ..., n_k$ iff $n_1 \oplus n_2 \oplus ... \oplus n_k = 0$

Proof:
Suppose inductively that this is true for all nim positions with fewer coins involved.

First, suppose

$n_1 \oplus n_2 \oplus ... \oplus n_k = b \neq 0.$

We show that we can win if we have the move.

Consider binary expansions.
Some n_i has a 1 in the same position as the leading 1 of b,
so

$n_i \oplus b < n_i.$

So we can move by taking coins from the ith pile so as to leave $n_i(+)b$ coins in that pile.

Then in the new position, the nim sum of the pile sizes is

$n_1 \oplus ... \oplus n_{i-1} \oplus n_i \oplus b \oplus n_{i+1} \oplus ... \oplus n_k$
$= b \oplus b$
$= 0$

So by the induction hypothesis, the player without the move wins from here.
But that's us!

Now suppose

$n_1 \oplus n_2 \oplus ... \oplus n_k = 0$

and we don't have the move.

If our opponent can't move, we've won.
Else, suppose they move by taking coins from the ith pile, leaving $m < n_i$.
But then $m \oplus n_i \neq 0$, so

$n_1 \oplus ... \oplus m \oplus ... \oplus n_k \neq n_1 \oplus ... \oplus n_i \oplus ...n_k = 0,$

so by the induction hypothesis, we're left with a position won by the player with the move, which is us.

2 Counting

Example:
If we draw 37 circles in the plane, such that every two circles intersect in two points and such that no three circles intersect in the same point ("mutually overlapping circles in general position"), how many regions in the plane do we get (counting the region outside all the circles as one of the regions)?

Solution:
Let h_n be the number of regions with n circles.
So $h_1 = 2$,
and for $n > 1$: when we add the nth circle, it intersects the existing $n - 1$ circles in $2(n - 1)$ points; so the new circle is divided into $2(n - 1)$ arcs; each arc divides a pre-existing region into two. So we get $2(n - 1)$ extra regions compared to how many we had before, i.e. compared to h_{n-1}. So

$h_n = h_{n-1} + 2(n - 1).$

Applying this recursively, we have

$$\begin{aligned} h_n &= h_{n-1} + 2(n-1) \\ &= h_{n-2} + 2(n-2) + 2(n-1) \end{aligned}$$

$$
\begin{aligned}
&= h_{n-3} + 2(n-3) + 2(n-2) + 2(n-1) \\
&= ... \\
&= h_1 + 2(1) + 2(2) + ... + 2(n-1) \\
&= 2 + 2(1 + 2 + ... + n - 1) \\
&= 2 + 2(n(n-1)/2) \\
&= 2 + n(n-1) \\
&= n^2 - n + 2
\end{aligned}
$$

(We can check this satisfies the recurrence relation:

$$
\begin{aligned}
&((n-1)^2 - (n-1) + 2) + 2(n-1) \\
&\quad = n^2 - 2n + 1 - n - 1 + 2 + 2n - 1 \\
&\quad = n^2 - n + 2,
\end{aligned}
$$

as required.)

So $h_{37} = 37^2 - 37 + 2 = 1334$.

Four principles of counting

We find ways to use descriptions of a finite set to determine its size in terms of the sizes of the sets appearing in the description.

Addition principle

Suppose a set S is partitioned by subsets $S_1, ..., S_n$;
this means that the sets cover S,
i.e. $S = S_1 \cup ... \cup S_n$,
and are disjoint,
i.e. $S_i \cap S_j = \emptyset$ if $i \neq j$.
In other words, suppose S is the disjoint union of the S_i.
Then $|S| = |S_1| + ... + |S_n|$.

Example:
In a nim game with three piles of sizes 3, 7, and 27,
how many possible first moves are there?

Solution:
The first move has to be in one of the three piles, so the answer is the sum of the numbers of moves available when the move is in each of the three piles.
The number of possible moves from a pile with n coins is n.
So the answer is $3 + 7 + 27 = 37$.

Multiplication principle

Let S be a set of ordered pairs (a, b) where the first is picked from a set of size p, and where for each such a there are q possible choices for b.
Then $|S| = pq$

Example:
How many two-digit numbers have different digits?

Solution:
Think of a two-digit number as an ordered pair, e.g. $37 \leftrightarrow (3, 7)$.
Then the set of two-digit numbers with different digits corresponds to the set of ordered pairs of numbers where the first number is from $\{1, ..., 9\}$, and the second number is then one of the 9 digits which is not equal to the first digit. So there are $9 * 9 = 81$ possibilities.

Proof of the multiplication principle:
Partition S according to the first entry of the pair. Then there are p sets each of size q, so by the addition principle

$$\begin{aligned} |S| &= q + ... + q \text{ (} p \text{ times)} \\ &= pq \end{aligned}$$

Of course it generalises to ordered triples and so on:

Example:
How many three digit numbers have no repeated digits?

Solution:
9 choices for the first digit, then 9 for the second, then 8 for the third; so $9 * 9 * 8 = 648$ possibilities.

Example:
How many factors does 12600 have?

Solution:
Find prime factorisation:

$$12600 = 2^3 * 3^2 * 5^2 * 7$$

So any factor has prime factorisation

$$2^a * 3^b * 5^c * 7^d$$

with $0 \leq a \leq 3, 0 \leq b \leq 2, 0 \leq c \leq 2, 0 \leq d \leq 1$.

By uniqueness of prime factorisation, each choice of (a, b, c, d) within these bounds yields a unique factor of 12600.

So the number of factors is the number of such (a, b, c, d), which is

$$4 * 3 * 3 * 2 = 72$$

Example:
How many two digit numbers have the sum of their digits odd, and don't end with 7?

Solution:
We need one digit even and the other odd; the first digit can't be 0, the second can't be 7.

If we pick the first digit first, the number of possibilities for the second digit will depend on the parity of the first digit.

There are 5 possibilities for an odd first digit. The second then has to be even: 5 possibilities.

With an even first digit, we have 4 possibilities for the first, and 4 for the second.

So the answer is $5 * 5 + 4 * 4 = 41$.

Subtraction principle

If S is a subset of U, then $|S| = |U| - |U \setminus S|$.

Proof: S and $U \setminus S$ partition U, so $|U| = |S| + |U \setminus S|$.

Example:
If we toss a coin 10 times, we get a sequence of Heads and Tails,
e.g. HTTHHTTHTH.
How many such sequences contain at least two Heads?

Solution:
The number of such sequences which **don't** contain at least two Heads is much easier to work out; it's 11: 1 for TTTTTTTTTT, and 10 for each of the ten positions a single Head could be.

The total number of sequences is, by the multiplication principle,
$2 * 2 * ... * 2$ (10 times)
$= 2^1 0 = 1024$.

So the answer is
$1024 - 11 = 1013$.

Division principle

If S is partitioned into k sets each with n elements, then $k = |S|/n$.
(Proof: by the multiplication principle, $|S| = k * n$.)

Example:
740 pigeons are nesting in some pigeonholes. If there are 5 pigeons in each pigeonhole, how many pigeonholes are there?

Solution:
$740/5 = 148$.

Finite probability

If we toss a coin 10 times, what's the probability of getting at least two Heads? Answer: there are $2^{10} = 1024$ possible outcomes (sequences of Heads and Tails), each equally likely. 1013 of them end up with at least two heads, so the probability is 1013/1024.

General setup:
we have a <u>sample space</u>, a finite set S of possible outcomes, each equally likely to occur. Then an <u>event</u> is a subset E of S, and we define the probability that E occurs to be

$$\mathrm{Prob}(E) := |E|/|S|$$

Example: Suppose someone offers to play the following game with you: three 6-sided dice will be rolled, and you'll set up a nim game with three piles, the sizes of the piles being given by the dice roll. You will play first. Assuming you both know the winning strategy for nim, what are your chances of winning?

Solution:
The number of possible dice rolls is $6^3 = 216$.

So we want to determine the number of triples (a, b, c) of numbers between 1 and 6 such that the nim sum is zero,

$$a \oplus b \oplus c = 0.$$

Nim-adding c to both sides, this is equivalent to

$$a \oplus b = c.$$

So once a and b are determined, there's at most one choice for c. So we want to see for how many (a, b),

$$1 \leq a \oplus b \leq 6.$$

This can only fail if, in binary,

$$a \oplus b = 000 \text{ or } 111$$

i.e. if $a = b$ or $a + b = 7$.
So this precisely rules out two choices for b for each choice of a.

So there are $6 * 4 = 24$ bad rolls for us, so the probability we win is

$$216 - 24/216 = 192/216 = 8/9.$$

3 Permutations

A <u>permutation</u> of a finite set S is an ordered list of its elements.

An <u>r-permutation</u> of S is an ordered list of r of its elements.

Warning:
there is another, related, meaning of 'permutation': an element of the group of bijections of S. We won't use that meaning in this course.

$P(n, r) :=$ number of r-permutations of a set of size n.

e.g. $P(26, 5)$ = number of strings of 5 distinct letters from the Roman alphabet.

By the multiplication principle,
$P(n, r) = n * (n-1) * ... * (n-(r-1))$ (n choices for first, $n-1$ for second...)

$P(n, r) = n!/(n-r)!$

$P(n, n) = n!$

Remark:
Can interpret "$P(n, r) = n!/(n-r)!$" as follows:
We can obtain an r-permutation of S by taking the first r elements of a permutation of S.
Partition the permutations of S according to the r-permutation which results from this: we see that the elements of each set of the partition correspond to the permutations of the left-over $n-r$ elements, so we recover the formula by the division principle.

A circular r-permutation of a set is a way of putting r of its elements around a circle, with two such considered equal if one can be rotated to the other.

We can obtain a circular r-permutation from an r-permutation by "joining the ends into a circle". Each circular r-permutation is obtained from r different r-permutations, so by the division principle:
number of circular r-permutations of n elements
$= P(n, r)/r$
$= n!/r(n-r)!$

Example:
How many different kinds of necklace can be made from 7 spherical beads of different colours? Consider two necklaces to be of the same kind when they can be non-destructively manipulated to look the same.

Solution:
There are $7!/7 = 6!$ circular permutations of the 13 colours. Each kind of necklace is obtained from exactly **two** circular permutations, because flipping the necklace in space doesn't change the kind. So the answer is
$6!/2 = 360$.

Example:
How many ways can 13 people be sat around a round table, if Professor Q is not to be sat next to his arch-nemesis Inspector P?

Solution:
Without the restriction, there would be 12! seating arrangements.
Consider seating everyone but P; each such arrangement yields two forbidden

arrangements of all 13, one by placing P to Q's right and one by placing P to Q's left. We count each forbidden arrangement once in this way.

So the answer is $12! - 2 * 11! = 10 * 11! = 399168000$

Subsets ("Combinations")

An <u>r-subset</u>, or <u>r-combination</u>, of a set S is a subset of size r.

$C(n,r) = \binom{n}{r}$ = number of r-subsets of a set of size n.

e.g. $\binom{26}{5}$ = number of unordered selections of 5 letters from the roman alphabet

Theorem:
$\binom{n}{r} = n!/r!(n-r)!$

Proof:
The r-permutations of a set are precisely the permutations of the r-subsets. Each r-subset has $r!$ permutations, so

$$P(n,r) = r! * \binom{n}{r}.$$

So

$$\binom{n}{r} = P(n,r)/r! = n!/r!(n-r)!.$$

$\binom{n}{r}$ is also called a "binomial coefficient".

Example:
If we expand out $(x+y)^n$ and collect terms to obtain

$$a_0x^n + a_1x^{n-1}y + ... + a_{n-1}xy^{n-1} + a_ny^n,$$

what are the coefficients a_k?

Solution:
a_k is the number of ways of choosing y k times when we have to choose either x or y from each factor of the product

$$(x+y)(x+y)...(x+y) \ (n \text{ times}),$$

which is the number of subsets of this set of n factors.

So $a_k = \binom{n}{k}$.

Theorem [Pascal's Formula]:
If $0 < k < n$,

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

Proof:
$|S| = n$.
Fix $x \in S$; let $S' := S \setminus \{x\}$.
Partition the k-subsets of S according to whether they contain x.
Those which don't correspond to k-subsets of S',
those which do correspond to $(k-1)$-subsets of S'.

Theorem:

$$\Sigma_{k=0}^{n} \binom{n}{k} = 2^n$$

Proof:
$|S| = n$.
$\Sigma_{k=0}^{n} \binom{n}{k}$ = number of subsets of S.
But to choose a subset of S is to choose for each element of S whether it should or should not go in to the subset. That's two choices for each of the n elements, so by the multiplication principle there are $2 * 2 * ... * 2 = 2^n$ subsets of S.

4 Multisets

Example:
A bag of Scrabble tiles contains 100 tiles: 10 A's, 2 B's, 2 C's, 5 D's and so on.

When you start a game, you take 7 letters from the bag, and put them on a rack. How many possible hands can you get, if we say that the order of the tiles on the rack matters? How about if it doesn't?

A multiset is a "set with multiplicity".

Notation: $\{2 * a, 3 * b, 1 * c\}$

(Think of a "bag" with 2 a's, 3 b's and a c in it.)

We also allow "infinite multiplicity", denoted $\{\infty * a\}$.

Multiplicities are also called "repetition numbers".

The size of a multiset is the sum of the multiplicities (may be ∞).

An r-permutation of a multiset is an ordered list of r elements from the multiset;
e.g. the 2-permutations of $\{2 * a, 1 * b\}$ are
aa, ab, ba;
the 3-permutations of $\{\infty * a, 2 * b\}$ are
$aaa, aab, aba, baa, abb, bab, bba$.

A permutation of a multiset of size n is an n-permutation.

Example: how many permutations are there of the unfortunate scrabble hand $\{4 * U, 1 * J, 2 * K\}$?

Theorem:
Let S be a multiset with k types with finite multiplicities $n_1, ..., n_k$.
Let $n = \Sigma_i n_i$ be the size of S.
Then the number of permutations of S is
$n!/n_1! * n_2! * ... * n_k!$.

Proof:
Label the elements $1, ..., n$. Each permutation of $\{1, ..., n\}$ yields a permutation of S, and two yield the same permutation precisely when we can get one from another by permuting the labels on elements of the same type.
So there are $n_1! * n_2! * ... * n_k!$ permutations of $\{1, ..., n\}$ per permutation of S. We conclude by the division principle.

Example:
We have 4 black rooks and 4 white rooks. How many ways are there of putting them on a chess board such that no two are attacking (/defending) each other? e.g.

```
.......R  8
....r...  5
..R.....  3
......R.  7
...r....  4
.R......  2
.....r..  6
r.......  1
```

Solution:
First, just choose the 8 squares for them to occupy.

By listing off the filled columns row-by-row, a choice corresponds to a permutation of the columns, so there are 8!.

For a given such choice, a choice of colours corresponds to a permutation of the multiset $\{4 * r, 4 * R\}$.

So the answer is
$8! * (8!/4! * 4!) = 2822400$

An <u>r-submultiset</u>, or <u>r-combination</u>, of a multiset S is a multiset S' of size r such that for all x,
the multiplicity of x in S' is at most the multiplicity of x in S.

e.g. the 2-submultisets of $\{3 * a, b\}$ are $\{2 * a\}$, $\{1 * a, 1 * b\}$.

Theorem:
The number of r-combinations of a multiset with k types each with multiplicity at least r is
$\binom{r+k-1}{r}$

Example:
If we have a bag containing red, green and blue marbles, with many of each, and we draw 5 marbles from the bag, how many possible results (numbers of each colour drawn) are there?
Answer: $\binom{5+3-1}{5} = \binom{7}{5} = 7!/2!5! = 21$

Proof:
We can identify an r-submultiset with an arrangement of $k-1$ partitions interspersed among r identical objects, by counting the numbers of objects between the partitions; e.g. with $k = 6$ and $r = 8$

oo|ooo||o|oo|

corresponds to
$\{2 * a_1, 3 * a_2, 0 * a_3, 1 * a_4, 2 * a_5, 0 * a_6\}$.

These arrangements correspond to choosing r of the $r + k - 1$ characters to be 'o's, so the number of such arrangements is $\binom{r+k-1}{r}$.

5 Pigeonhole Principles

Pigeonhole Principle (PP):
If some pigeons are in some pigeonholes,
and there are fewer pigeonholes than there are pigeons,
then some pigeonhole must contain at least two pigeons.

// The "pigeons" and "pigeonholes" can be abstract!

Example:
If there are 367 people in a room,
there must be two who share a common birthday.

Interlude: maps and numbers

$f : X \to Y$ map between finite sets.

For $y \in Y$, $f^{-1}(y) =$ "fibre of f over y" $= \{x \mid f(x) = y\}$.

Recall:

- f is surjective aka onto,
  written $f : X \twoheadrightarrow Y$,
  if for all $y \in Y$, $|f^{-1}(y)| \geq 1$
- f is injective aka 1-1,
  written $f : X \hookrightarrow Y$,
  if for all $y \in Y$, $|f^{-1}(y)| \leq 1$
- f is bijective aka a (1-1) correspondence aka invertible,
  written $f : X \xrightarrow{\equiv} Y$,
  if f is both injective and surjective,
  i.e. if for all $y \in Y$, $|f^{-1}(y)| = 1$

Remark:
If f is

- injective then $|X| \leq |Y|$ (Pigeonhole principle)
- surjective then $|X| \geq |Y|$
- bijective then $|X| = |Y|$

Applications of the Pigeonhole principle

Example:
If I take 13 coins,
divide them into 9 piles, placed in a row,
then there will be a group of neighbouring piles within the row such that there are exactly 4 coins in the group.
(Generally: n coins, m piles; must be k coins in a contiguous group if

$$n + k < 2 * m$$

(this isn't sharp))

Proof:
Let $a_i :=$ number of coins in first i piles, $1 \leq i \leq 9$.
Consider the 18 numbers

$$a_1, a_2, ..., a_9, a_1 + 4, a_2 + 4, ..., a_9 + 4.$$

Since $1 \leq a_i \leq 13$, these numbers are all between 1 and 17.
So by the PP, two must be equal.

Since no two a_i are equal (since the piles are non-empty),
and similarly no two $a_i + 4$ are equal,
we must have $a_i = a_j + 4$ for some i,j.

So $a_i - a_j = 4$,
so 4 is the sum of the sizes of the piles after i and up to j,
namely piles $i + 1, ..., j$.

Example:
Using as many coins as I want,

I make a row of k piles.
Then there is a group of neighbouring piles such that the number of coins in the group is divisible by k.

Proof:
Let $a_1, ..., a_k$ be as above.

Let r_i be the remainder on dividing a_i by k.
If any $r_i = 0$, we're done.
Else, $0 < r_i < k$,
so by the PP, two remainders are equal,
$r_i = r_j$.
But then $a_i - a_j$ is divisible by 7, and we conclude as in the previous example.

Packed Pigeonhole Principle

Packed Pigeonhole Principle (PPP):
If there are more than $k * n$ pigeons in n pigeonholes,
then some pigeonhole contains more than k pigeons.
(Note: "Packed" is not standard terminology. This principle is commonly referred to as the pigeonhole principle. Brualdi calls something slightly more general (but less pleasing) the "strong pigeonhole principle", but I don't think we need to cover it)

Example:
If $a_1, ..., a_{n^2+1}$ is a sequence of $n^2 + 1$ real numbers,
there is a subsequence of length $n + 1$ which is monotonic, i.e. is either (nonstrictly) increasing or (nonstrictly) decreasing.

Proof:
Suppose there is no increasing subsequence of length $n + 1$.

Let l_i be the length of the longest increasing subsequence starting with a_i.
So $1 \leq l_i \leq n$.

So by the PPP, $n + 1$ of these $n^2 + 1$ numbers are equal;
say $l_{i_1} = ... = l_{i_{n+1}}$.

Now suppose $a_{i_j} < a_{i_{j+1}}$.
Then we can extend the longest increasing subsequence starting with $a_{i_{j+1}}$ to a longer one starting with a_{i_j}, by prepending a_{i_j}.
This contradicts $l_{i_j} = l_{i_{j+1}}$.

So $(a_{i_j})_j$ is a decreasing sequence of length $n + 1$.

Abstract version:
If $f : X \twoheadrightarrow Y$ is a surjection,
and if all fibres are of size at most k,

i.e. $|f^{-1}(y)| \leq k$ for all y,
then $|X| \leq k|Y|$.

Remark (Division principle, map form):
If $f : X \twoheadrightarrow Y$ is a surjection,
and if all fibres have size exactly k,
i.e. $|f^{-1}(y)| = k$ for all y,
then $|X| = k|Y|$.

(Proof: partition X according to the value of f, apply division principle)

Averaging principle:
Given integers $a_1, ..., a_n$,
some a_i is at least the average,
$a_i \geq (a_1 + ... + a_n)/n$
(Note the average might not be an integer!)

Packed Pigeonhole follows from averaging:
if there are more than $k * n$ pigeons,
then the average number of pigeons per pigeonhole is more than k;
some pigeonhole has at least the average number of pigeons,
so has more than k pigeons.

Example:
Discs (p.75)

6 Ramsey Theory

Example:
Given 6 people,
either there are 3 who all like each other,
or there are 3 no two of whom like each other.

Abstract version:
$K_n :=$ "complete graph on n vertices"
$= n$ points with an edge between each pair.

Colour the edges of K_6 each either red or blue,
then there's a red copy of K_3 or there's a blue copy of K_3;
i.e. there is a monochromatic triangle.

Denote this fact
$K_6 \to K_3, K_3$

Proof:
Pick a vertex v_0.
Consider the 5 edges from it.
3 of them are red or 3 of them are blue, since $5 > (3-1) + (3-1)$.
Say 3 are red, and consider the 3 other vertices of these red edges.
If the edges between them are all blue, they form a blue triangle and we're done.
Else, some edge is red; but then it along with the edges from v_0 form a red triangle, and we're done.

Ramsey's Theorem for 2-coloured graphs:
Given n and m positive integers,
there exists r such that for any red-blue colouring of the edges of K_r,
there are n vertices all edges between which are red or there are m vertices all edges between which are blue.

Notation:
We write
$$K_r \to K_n, K_m$$
to mean that r has this property,
and we let $r(m, n)$ ("the (m,n)th Ramsey number") be the least such r.

Remarks:
We saw that $K_6 \to K_3, K_3$;
it's easy to see that $K_5 \not\to K_3, K_3$,
so $r(3, 3) = 6$.

It has been shown that
$$r(3, 4) = 9$$
$$r(3, 5) = 14$$
$$r(4, 4) = 18$$

$r(5, 5)$ is unknown! All we know is
$$43 \leq r(5, 5) \leq 49.$$

Erdös:
"Suppose aliens invade the earth and threaten to obliterate it in a year's time unless human beings can find the Ramsey number for red five and blue five. We could marshal the world's best minds and fastest computers, and within a year we could probably calculate the value. If the aliens demanded the Ramsey number for red six and blue six, however, we would have no choice but to launch a preemptive attack."

Proof of Theorem:
Suppose inductively that
$$K_b \to K_{n-1}, K_m$$
and
$$K_c \to K_n, K_{m-1}.$$

We show that

$$K_{b+c} \to K_n, K_m.$$

So colour K_{b+c}, and suppose there's no red K_n and no blue K_m.

Pick a vertex v_0; consider the $b + c - 1$ edges from it.
Since $b + c - 1 > (b - 1) + (c - 1)$,
b of the edges are red or c of the edges are blue.

Say b are red.
Consider the K_b formed by the vertices these edges connect to v_0.
By the inductive hypothesis, it contains a red K_{n-1} or a blue K_m.
If it contains a red K_{n-1}, adjoining v_0 yields a red K_n;
contradiction.
If it contains a blue K_m, then so does our original K_{b+c};
contradiction.

A symmetrical argument applies in the case that c of the edges from v_0 are blue.

Remark:
This proof yields a recursive upper bound on the Ramsey numbers:

$$r(m, n) \leq r(n - 1, m) + r(n, m - 1)$$

(but this is far from sharp).

7 Binomial coefficients

Miscellaneous Curiosities

Recall:

- For n a non-negative integer and r an integer,
  $\binom{n}{r}$ = number of subsets of size r of a set of size n
  $= \frac{n!}{n!(n-r)!}$ if $0 \leq r \leq n$
  $= 0$ else.
- Pascal's triangle
- Pascal's Formula: $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$
- $\binom{n}{r} = \binom{n}{n-r}$
- $\sum_{r=0}^{n} \binom{n}{r} = 2^n$
- $(x + y)^n = \sum_{r=0}^{n} \binom{n}{r} x^{n-r} y^r$

Remark:
$\binom{n}{r}$ = number of paths from root of Pascal's triangle to the (n, r) position.

Further identities:

- $k\binom{n}{k} = n\binom{n-1}{k-1}$ (immediate from $\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k(k-1)\dots 1}$)
- $(x+1)^n = \sum_{r=0}^{n} \binom{n}{r} x^r$
- $0 = ((-1)+1)^n = \sum_{r=0}^{n} \binom{n}{r} (-1)^r$;
  so alternating sum of binomial coefficients is 0;
  so sum of even coefficients = sum of odd coefficients = 2^{n-1}.

Yet further identities:

(i) $\binom{n+1}{r+1} = \sum_{s=0}^{n} \binom{s}{r}$

(ii) $\sum_{r=0}^{n} \binom{n}{r}^2 = \binom{2n}{n}$

(iii) $\sum_{r=0}^{n} r\binom{n}{r} = n2^{n-1}$

Proofs:

(i) Iteratively apply Pascal's formula:

$$\begin{aligned}\binom{n+1}{r+1} &= \binom{n}{r+1} + \binom{n}{r} \\ &= \binom{n-1}{r+1} + \binom{n-1}{r} \\ &= \dots \\ &= \binom{0}{r+1} + \binom{0}{r} + \dots + \binom{n-1}{r} \\ &= \binom{0}{r} + \dots + \binom{n-1}{r}\end{aligned}$$

Alternative inductive proof:
Easily holds for $n = r = 0$.
Suppose inductively it holds for smaller $n + r$.
Then using Pascal's formula, we have:

$$\begin{aligned}\binom{n+1}{r+1} &= \binom{n}{r+1} + \binom{n}{r} \\ &= \sum_{k=0}^{n-1} \binom{k}{r} + \sum_{k=0}^{n-1} \binom{k}{r-1} \\ &= \sum_{k=0}^{n-1} \left(\binom{k}{r} + \binom{k}{r-1}\right) \\ &= \sum_{k=0}^{n-1} \binom{k+1}{r} \\ &= \sum_{k=1}^{n} \binom{k}{r} \\ &= \sum_{k=0}^{n} \binom{k}{r}\end{aligned}$$

(ii) Consider diamonds in Pascal's triangle.

OR: Given a set S of size $2n$, arbitrarily split it into two sets S_1, S_2 of size n.
Then an n-subset S' of S corresponds to the pair $(S' \cap S_1, S' \cap S_2)$.
The pairs of subsets arising in this way are precisely those of sizes summing to n,
so

$$\binom{2n}{n} = \sum_{r=0}^{n} \binom{n}{r}\binom{n}{n-r} = \sum_{r=0}^{n} \binom{n}{r}^2$$

(iii) Neat algebraic proof:

$$\begin{aligned}n(x+1)^{n-1} &= \frac{d}{dx}(x+1)^n = \frac{d}{dx} \sum_{r=0}^{n} \binom{n}{r} x^r \\ &= \sum_{r=0}^{n} r\binom{n}{r} x^r.\end{aligned}$$

This holds for all x; taking $x = 1$ gives the result.

Examples of (i):

- $\binom{n}{1} = \sum_{s=0}^{n-1} \binom{s}{0} = \sum_{s=0}^{n-1} 1 = n$
- $\binom{n}{2} = \sum_{s=0}^{n-1} \binom{s}{1} = \sum_{s=0}^{n-1} s = n$th triangular number
- $\binom{n}{3} = \sum_{s=0}^{n-1} \binom{s}{2} = n$th pyrimidal number

Multinomial theorem

What is the coefficient $a_{r,s,t}$ of $x^r y^s z^t$ in the expansion of $(x+y+z)^n$?

Clearly $a_{r,s,t} \neq 0$ only if $r+s+t=n$.

$a_{r,s,t}$ is the number of ways of choosing r x's, s y's, and t z's from the n factors $(x+y+z)$;
i.e. the number of strings like "xyzzyxyzzy" with this many of each letter;
i.e. the number of permutations of the multiset $\{r*x, s*y, t*z\}$.

So as we saw before,
$$a_{r,s,t} = \frac{n!}{r!s!t!}.$$
Write $\binom{n}{r\ s\ t}$ for this number.

Generalising to arbitrarily many variables, we have

Theorem:
$(x_1 + ... + x_t)^n = \sum_{n_i \geq 0, n_1+...+n_t=n} \binom{n}{n_1\ n_2\ ...\ n_t} x_1^{n_1} x_2^{n_2} ... x_t^{n_t}$

Here, $\binom{n}{n_1\ n_2\ ...\ n_t} = \frac{n!}{n_1!...n_t!}$ are the <u>multinomial coefficients</u> (only defined if $n_1 + ... + n_t = n$).

Note:
$\binom{n}{r} = \binom{n}{r\ n-r}$.

The number of terms in the multinomial expansion of $(x_1 + ... + x_t)^n$ is the number of n-combinations with t types in unlimited supply, which we saw is
$$\binom{n+t-1}{n}.$$

Unnatural exponents: $(x+y)^\alpha$

Theorem [Newton's Binomial Theorem]:
Let α be a real (or even complex) number.
Suppose $0 \leq |x| < |y|$.
Then
$$(x+y)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k y^{\alpha-k}$$

where
$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)...(\alpha-k+1)}{k!}.$$

Note that for α natural, this agrees with our previous definition.

Proof (not on syllabus):
Dividing through by y^α, sufficient to show that for $|z| < 1$,

$$(1+z)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} z^k.$$

We show this for complex z with $|z| < 1$.
$(1+z)^\alpha = \exp(\alpha \log(1+z))$ for any choice of branch.
This is holomorphic on the domain $|z| < 1$,
so the Taylor series at 0 converges to the value of the function on this domain.

Since $\frac{d}{dz}(1+z)^\alpha = \alpha(1+z)^{\alpha-1}$ and $(1+0)^\alpha = 1$,
this gives

$$\begin{aligned}(1+z)^\alpha &= ((1+z)^\alpha|_{z=0})\frac{z^0}{0!} + \\ &\quad (\alpha(1+z)^{\alpha-1}|_{z=0})\frac{z^1}{1!} + \\ &\quad (\alpha(\alpha-1)(1+z)^{\alpha-2}|_{z=0})\frac{z^2}{2!} + \ldots \\ &= \sum_{k=0}^{\infty}(\alpha(\alpha-1)\ldots(\alpha-k+1))\frac{z^k}{k!} \\ &= \sum_{k=0}^{\infty}\binom{\alpha}{k} z^k\end{aligned}$$

□

Examples:

- $$\begin{aligned}\frac{1}{1+z} &= (1+z)^{-1} \\ &= \sum_{k=0}^{\infty}\binom{-1}{k} z^k \\ &= \sum_{k=0}^{\infty}\frac{(-1)*(-2)*\ldots*(-k)}{k*(k-1)*\ldots*1} z^k \\ &= \sum_{k=0}^{\infty}(-1)^k z^k \\ &= 1 - z + z^2 - z^3 + \ldots\end{aligned}$$

- $$\begin{aligned}\sqrt{37} = \sqrt{6^2+1} = 6\sqrt{1+1/36} &= 6(1+1/36)^{1/2} \\ &= 6\left(\sum_{k=0}^{\infty}\binom{1/2}{k}(1/36)^k\right)\end{aligned}$$

  Now for $k > 0$,

  $$\begin{aligned}\binom{1/2}{k} &= \frac{\frac{1}{2}\frac{1-2}{2}\ldots\frac{1-2(k-1)}{2}}{k!} \\ &= \frac{(-1)^{k-1}1*3*5*\ldots*(2k-3)}{2^k k!} \\ &= \frac{(-1)^{k-1}(2k-2)!}{2^k(2*4*\ldots*(2k-2))k!} \\ &= \frac{(-1)^{k-1}(2k-2)!}{2^{2k-1}(k-1)!k!} \\ &= \frac{(-1)^{k-1}}{k2^{2k-1}}\binom{2k-2}{k-1}\end{aligned}$$

  So

  $$\begin{aligned}\sqrt{1+z} &= 1 + \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k2^{2k-1}}\binom{2k-2}{k-1} z^k \\ &= 1 + \frac{1}{2}z - \frac{1}{8}z^2 + \frac{1}{16}z^3 - \frac{1}{25}z^4 + \ldots\end{aligned}$$

  So $\sqrt{37} = 6(1+1/36)^{1/2}$

  $$\begin{aligned}&\approx 6(1 + 1/(2*36) - 1/(8*36^2) + 1/(16*36^3) - 1/(25*36^4)) \\ &= 6.0828\end{aligned}$$

  (Error is very small: $(6*(1+1/(2*36) - 1/(8*36^2) + 1/(16*36^3) - 1/(25*36^4)))^2 = 36.99999992692224$)

8 Partial Orders

Basics

A partial order on a set X is a binary relation $\leq$ which is

- reflexive ($x = y => x \leq y$)
- transitive (if $x \leq y$ and $y \leq z$ then $x \leq z$)
- antisymmetric (if $x \leq y$ and $y \leq x$ then $x = y$).

A set X equipped with a partial order $\leq$, denoted $(X; \leq)$, is called a partially ordered set or a poset.

"$x \geq y$" means "$y \leq x$".
"$x < y$" means "$x \leq y$ and $x \neq y$".

Examples:

(i) The usual order $\leq$ on the integers.

(ii) The relation of divisibility is a partial order on the natural numbers; $(\mathbb{N}; |)$ is the corresponding poset.

(iii) If A is a set, the set of subsets of A is partially ordered by inclusion, $\subseteq$.

Hasse diagrams:
x covers y if $x > y$ and there is no z such that $x > z > y$.

The Hasse diagram of a finite poset $(X, \leq)$ consists of points for the elements of X and a line drawn upwards from y to x whenever x covers y.

(We will see below that every finite poset has a Hasse diagram.)

x is minimal if $x > y$ for no y.
x is maximal if $x < y$ for no y.

Lemma:
$<$ is transitive: if $x < y$ and $y < z$ then $x < z$.

Proof:
$x \leq z$ by transitivity of $\leq$.
Suppose $x = z$.
Then $y \leq x$ and $x \leq y$, so $x = y$ by antisymmetry, contradicting $x < y$.
So $x \neq z$. $\square$

Lemma:
Any finite poset has at least one minimal element, and at least one maximal element.

Proof:
Suppose $(X; \leq)$ has no minimal element.
Then there exist arbitrarily long chains $x_1 > x_2 > x_3 > ... > x_n$.
By transitivity of $>$, the x_i are distinct, so we contradict finiteness. $\square$

A partial order $\leq$ on a set X is <u>total</u> (aka <u>linear</u>) if for all x and y in X, either $x \leq y$ or $y \leq x$.

Lemma:
Any finite total order $(X; \leq)$ can be enumerated as $X = \{x_1, ..., x_n\}$ with $x_i \leq x_j$ iff $i \leq j$.

(i.e. $(X; \leq)$ is isomorphic to $\{1, ..., n\}$ with the usual order.)

Proof:
If x is minimal in a total order, then $x \leq y$ for any y.

Let x_1 be minimal in X, then let x_2 be minimal in $X \setminus \{x_1\}$, and so on.

Then $x_i \leq x_j$ for $i \leq j$.
By antisymmetry, $x_i \not\leq x_j$ for $i \not\leq j$. $\square$

Lemma:
Any finite poset $(X; \leq)$ can be linearised,
i.e. there exists a total order $\leq'$ such that $x \leq y => x \leq' y$.

Proof:
Let $x_1, ..., x_n$ be the minimal elements of $(X; \leq)$.
Let $X' := X \setminus \{x_1, ..., x_n\}$.
By induction, $(X'; \leq)$ can be linearised, say to $\leq'$.
Extend $\leq'$ to X by defining

- $x_i \leq' x_j$ iff $i \leq j$
- $x_i \leq' y$ for any $y \in X'$

This is total. $\square$

Consequence:
Any finite poset has a Hasse diagram:
draw the points with heights ordered according to a linearisation of the partial order,
then draw a line whenever x covers y, which implies that x is above y.
(Nudge the points horizontally if there are any overlapping lines).

Chains and antichains

Definition:
Let $(X; \leq)$ be a poset.

$\emptyset \neq C \subseteq X$ is a chain if $(C; \leq)$ is a total order.

$\emptyset \neq A \subseteq X$ is an antichain if $a_1 \leq a_2 \Rightarrow a_1 = a_2$ for $a_i \in A$.

A chain partition is a partition $X = C_1 \cup ... \cup C_n$ by disjoint chains.

An antichain partition is a partition $X = A_1 \cup ... \cup A_n$ by disjoint antichains.

Theorem:
In a finite poset $(X; \leq)$,

(i) The maximal size of a chain is equal to the minimal size of an antichain partition.

(ii) [Dilworth's theorem] The maximal size of an antichain is equal to the minimal size of a chain partition.

Proof:
First observe that a chain and an antichain can have no more than 1 point in common,

$$|C \cap A| \leq 1.$$

So given an antichain partition and a chain,
each element of the chain is in precisely one of the antichains,
and no two elements of the chain are in the same antichain,
so

size of any chain $\leq$ size of any antichain partition

so

maximal size of a chain $\leq$ minimal size of an antichain partition.

Similarly for (ii): given a chain partition, each element of an antichain is in exactly one of the chains in the partition,
and no two elements of the antichain are in the same chain, so

maximal size of an antichain $\leq$ minimal size of a chain partition.

So it remains to see

(i) There exists an antichain partition of size the maximal size of a chain;

(ii) There exists a chain partition of size the maximal size of an antichain.

These require separate arguments.

(i) Let C be a chain of maximal length.
Say $C = \{c_1, ..., c_n\}$ with $c_i \leq c_j$ iff $i \leq j$.

For $i = 1, ..., n$, recursively define

$A_i :=$ the set of minimal elements of $X \setminus (A_1 \cup ... \cup A_{i-1})$.

Then A_i is an antichain,
and $c_i \in A_i$ so no A_i is empty.

If $x \in X \setminus \cup_i A_i$ then $x \geq c_i$ for all i,
so we could extend C to a larger chain by adjoining x,
contradicting maximality of C.
So $X = \cup_i A_i$.

So the A_i form an antichain partition of size $n = |C|$ as required.

(ii) By induction on the size of X.

First, suppose some antichain A of maximal size m is not the set of maximal elements and is not the set of minimal elements.

Let
$$A^+ := \bigcup_{a \in A} \{x \mid x \geq a\}$$
$$A^- := \bigcup_{a \in A} \{x \mid x \leq a\}.$$

Then $A^+ \cup A^- = X$ by maximality of A,
and $A^+ \cap A^- = A$.

Now A is a maximal-size antichain in A^+,
and $A^+ \neq X$ since A is not the set of maximal elements,
so by induction, A^+ has a chain partition of size $|A|$.

Similarly, we have a chain partition of A^-.

For each element $a \in A$, a is in one of the chains of A^+ and one of the chains of A^-, and the union of these two chains is a chain C_a in X.

Then $\{C_a \mid a \in A\}$ is a chain partition of X of size $m = |A|$.

For the remaining case,
suppose every maximal-size antichain is either the set of maximal elements or the set of minimal elements.

Let x be minimal and y be maximal, with $x \leq y$.

(To see that such that such x and y exist:
we proved above that maximal and minimal elements always exist,
so we only need to see that some minimal element is comparable with (and hence $\leq$) some maximal element.
Otherwise, the minimal elements and the maximal elements together form an antichain,
which contradicts our assumption unless the set of minimal elements is equal to the set of maximal elements,
in which case we can take $x = y$.)

Then $X \setminus \{x, y\}$ has no antichains of size m but has an antichain of size $m - 1$,
so by induction it has a chain partition of size $m - 1$.
Adjoining the chain $\{x, y\}$, we obtain a chain partition of X of size m.

□

Bonus: Sperner's Theorem

Example:
Inductively define "symmetric" chain partitions S_n of the set of subsets of $\{1, ..., n\}$:

Let S_1 be the partition with only one chain,
$\emptyset \subsetneq \{1\}$.

Given a chain partition S_n of $\{1, ..., n\}$,
let S_{n+1} have, for each chain $A_1 \subsetneq ... \subsetneq A_k$ of S_n,

- the chain $A_1 \subsetneq ... \subsetneq A_k \subsetneq A_k \cup \{n+1\}$,
- and, if $k > 1$, the chain $A_1 \cup \{n+1\} \subsetneq ... \subsetneq A_{k-1} \cup \{n+1\}$.

Each chain $A_1 \subsetneq ... \subsetneq A_k$ in S_n has subsequent elements differing in size by one,
and $|A_1| + |A_k| = n$.

Hence each chain contains a subset of size $\lfloor \frac{n}{2} \rfloor$,
so $|S_n| = \binom{n}{\lfloor \frac{n}{2} \rfloor}$.

($\lfloor \frac{n}{2} \rfloor = n/2$ "rounded down")

So by Dilworth, we obtain "Sperner's Theorem":
The set of subsets of $\{1, ..., n\}$ of size $\lfloor \frac{n}{2} \rfloor$ is a maximal-size antichain.

9 Inclusion-Exclusion

Let A_1 and A_2 be finite subsets of a set X.
If A_1 and A_2 are disjoint, the addition principle tells us $|A_1 \cup A_2| = |A_1| + |A_2|$.

If they're not disjoint, this "double-counts" the elements of the intersection; we can fix this by subtracting the size of the intersection,
yielding the general formula $|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$.

For three finite subsets A_1,A_2,A_3 of some set X, similar reasoning yields
$$|A_1 \cup A_2 \cup A_3| =$$
$$|A_1| + |A_2| + |A_3| - (|A_1 \cap A_2| + |A_2 \cap A_3| + |A_1 \cap A_3|) + |A_1 \cap A_2 \cap A_3|.$$

Example:
How many integers in [1,100] are divisible by 2, 5, or 7?

Let $D_n := \{k \in \mathbb{Z} \cap [1, 100] : n \mid k\}$.

Note $|D_n| = \lfloor \frac{100}{n} \rfloor$.

By inclusion-exclusion,
$$|D_2 \cup D_5 \cup D_7| = |D_2| + |D_5| + |D_7|$$
$$-(|D_2 \cap D_5| + |D_5 \cap D_7| + |D_2 \cap D_7|)$$
$$+|D_2 \cap D_5 \cap D_7|$$
$$= |D_2| + |D_5| + |D_7| - (D_{10} + D_{35} + D_{14}) + D_{70}$$

$$= 50 + 20 + 14 - (10 + 2 + 7) + 1$$
$$= 66.$$

Theorem [Inclusion-Exclusion Principle]:
Let $A_1, ..., A_n$ be finite subsets of a set X.
Then

$$\begin{aligned} |A_1 \cup ... \cup A_n| = \\ & \sum_i |A_i| \\ & - \sum_{i<j} |A_i \cap A_j| \\ & + \sum_{i<j<k} |A_i \cap A_j \cap A_k| \\ & - ... \\ & + (-1)^{n-1} |A_1 \cap ... \cap A_n| \end{aligned}$$

Proof:
Let $x \in A_1 \cup ... \cup A_n$.
We show that x "contributes 1" to the right hand side.
Say x is in $m \geq 1$ of the n sets.
Then x contributes 1 to $m = \binom{m}{1}$ of the $|A_i|$,
to $\binom{m}{2}$ of the $|A_i \cap A_j|$,
to $\binom{m}{3}$ of the $|A_i \cap A_j \cap A_k|$,
and so on.

So x contributes

$$\begin{aligned} & \sum_{k>0} (-1)^{k-1} \binom{m}{k} \\ & = - \sum_{k>0} (-1)^k \binom{m}{k} \\ & = 1 - \sum_{k \geq 0} (-1)^k \binom{m}{k} \\ & = 1 - 0 \\ & = 1 \end{aligned}$$

Remark:
Neat alternative expression:

$$|\bigcup_i A_i| = \sum_{\emptyset \neq I \subseteq \{1,...,n\}} (-1)^{|I|-1} \left|\bigcap_{i \in I} A_i\right|.$$

Bonus:
Version of the proof using this notation:

$$\left|\bigcup_i A_i\right| = \sum_{x\subseteq\bigcup_i A_i} 1$$

$$= \sum_{x\subseteq\bigcup_i A_i} \left(-\sum_{k>0} (-1)^k \binom{\#\{i \mid x\in A_i\}}{k} \right)$$

$$= \sum_{x\subseteq\bigcup_i A_i, I\subseteq\{i \mid x\in A_i\}} (-1)^{|I|-1}$$

$$= \sum_{\{(x,I) \mid x\in\bigcup_i A_i, I\subseteq\{1,\ldots,n\}, x\subseteq\bigcap_{i\in I} A_i\}} (-1)^{|I|-1}$$

$$= \sum_{\emptyset\neq I\subseteq\{1,\ldots,n\}} (-1)^{|I|-1} \left|\bigcap_{i\in I} A_i\right|$$

Example:
How many strings of 8 letters from the Roman alphabet contain 'j', 'q', 'x', 'y' and 'z'?

We could do this positively, but it would be fiddly.

Instead, let's count the number of strings which **don't** contain all of these letters, i.e. which omit 'j' or omit 'q' or... .

Let O_j be the strings which omit 'j', O_{jq} the strings which omit 'j' **and** 'q', and so on.

Then by inclusion-exclusion,

$$\begin{aligned} |O_j \cup O_q \cup O_x \cup O_y \cup O_z| = |O_j| + |O_q| + \ldots \\ -(|O_{jq}| + |O_{jx}| + \ldots) \\ +(|O_{jqx}| + |O_{jqy}| + \ldots) \\ -(|O_{jqxy}| + |O_{jqxz}| + \ldots) \\ +|O_{jqxyz}| \end{aligned}$$

Now $|O_j| = |O_q| = \ldots = 25^8$,
and $|O_{jq}| = |O_{jx}| = \ldots = 24^8$,
and so on.

So

$$|O_j \cup O_q \cup O_x \cup O_y \cup O_z| = \binom{5}{1}25^8 - \binom{5}{2}24^8 + \binom{5}{3}23^8 - \binom{5}{4}22^8 + \binom{5}{5}21^8$$

and the answer to the original question is $26^8 - |O_j \cup O_q \cup O_x \cup O_y \cup O_z|$, which comes to 87408720.

Combinations of multisets, revisited

Recall:
The number of r-combinations of a multiset with at least r of each of its t types is $\binom{r+t-1}{t-1}$.

If there are fewer than r of some of the types, we can use inclusion-exclusion.

This is clearest if we transform the problem.

An r-combination of a multiset
$\{c_1 * a_1, ..., c_t * a_t\}$
corresponds to a solution in non-negative integers of the equation
$x_1 + ... + x_t = r$
subject to the constraints
$x_1 \le c_1, ..., x_t \le c_t$.

Concrete example:
I take 8 marbles from a bag containing 3 red marbles, 2 blue marbles, and 10 green marbles. How many possibilities are there for the numbers of each colour I get? Equivalently, what is

$$|\{(r, b, g) \mid r + b + g = 8, 0 \le r \le 3, 0 \le b \le 2\}|?$$

So we want to count the number of such solutions,
and we know that the answer is $\binom{r+t-1}{t-1}$ if there are no constraints.

By the subtraction principle,
the number of solutions in the constrained case is the number in the unconstrained case minus the number which fail at least one constraint,
$\binom{r+t-1}{t-1} - |F_1 \cup ... \cup F_t|$,
where $F_i := \{(x_1, ..., x_t) \mid x_1 + ... + x_t = r, x_i > c_i\}$.

So we can use the inclusion-exclusion principle if we can determine the sizes of the intersections of the F_i.

If $c_i \ge r$, then $F_i = \emptyset$.
Otherwise, subtracting c_i+1 from x_i puts F_i in correspondence with $\{(y_1, ..., y_t) \mid y_1 + ... + y_t = r - (c_i + 1), y_i \ge 0\}$,
so $|F_i| = \binom{r-(c_i+1)+t-1}{t-1}$.

Similarly, $|F_i \cap F_j| = \binom{r-(c_i+1)-(c_j+1)+t-1}{t-1}$, and so on.

So inclusion-exclusion yields
$\sum_{I \subseteq \{1,...,t\}} (-1)^{|I|} \binom{r-(\sum_{i\in I}(c_i+1))+t-1}{t-1}$.

Marble example:
$\binom{8+2}{2} - \binom{8-(3+1)+2}{2} - \binom{8-(2+1)+2}{2} + \binom{8-(3+1)-(2+1)+2}{2}$
$= \binom{10}{2} - \binom{6}{2} - \binom{7}{2} + \binom{3}{2}$
$= 12$

Scrabble example:
How many 7-tile hands can be drawn from a standard 100-tile bag of scrabble tiles?
Using the above formula, my computer calculates it as 3199724.

(for the curious, here's the Haskell code I used to calculate this:

```
import Math.Combinatorics.Binomial (choose)
combs :: Int -> [Int] -> Int
combs r cs =
   let
       t = length cs
       subs = subs' [] cs
       -- subs': returns relevant subsequences of cs, omitting those which
       -- will contribute 0 to the final sum (without this, the algorithm
       -- would have complexity exponential in t)
       subs' sub [] = [sub]
       subs' sub _ | (sum (map (+1) sub) > r) = []
       subs' sub (c:cs) = subs' (c:sub) cs ++ subs' sub cs
  in sum [ (-1)^(length sub) *
        choose (r - sum (map (+1) sub) + t-1) (t-1) | sub <- subs ]
scrabbleBag :: [Int]
scrabbleBag = concat [ replicate n c | (n,c) <-
        [(5,1), (10,2), (1,3), (4,4), (3,6), (1,8), (2,9), (1,12)] ]
main :: IO ()
main = print $ combs 7 scrabbleBag
```

)

Derangements

A derangement is a permutation which leaves nothing in its original position.

e.g.

`(5,3,4,2,1)` is a derangement of
`(1,2,3,4,5)`, and
"`endgreatmen`" is a derangement of
"`derangement`".

$D_n :=$ the number of derangements of a sequence of length n,
$\quad = $ number of derangements of $(1, 2, ..., n)$.

We can use inclusion-exclusion to determine D_n.

A derangement of $(1, ..., n)$ is a permutation $(a_1, ..., a_n)$ which satisfies the conditions $a_1 \neq 1, ..., a_n \neq n$.

Let P_i be the set of permutations which fail the ith of these conditions, i.e. such that $a_i = i$.

Easily, for $I \subseteq \{1, ..., n\}$,
$$|\bigcap_{i\in I} P_i| = (n - |I|)!.$$

So by inclusion-exclusion,
$$\begin{aligned} D_n &= n! - |\bigcup_i P_i| \\ &= n! - \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|-1} |\bigcap_{i\in I} P_i| \\ &= n! - \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|-1} (n - |I|)! \\ &= n! - \sum_{i=1}^{n} (-1)^{i-1} \binom{n}{i} (n-i)! \\ &= \sum_{i=0}^{n} (-1)^{i} \binom{n}{i} (n-i)! \\ &= \sum_{i=0}^{n} (-1)^{i} \frac{n!}{i!} \\ &= n! \sum_{i=0}^{n} \frac{(-1)^i}{i!} \end{aligned}$$

Note then that the probability that a random n-permutation is a derangement is
$$Pr_n = \frac{|D_n|}{n!} = \sum_{i=0}^{n} \frac{(-1)^i}{i!},$$
so
$$\begin{aligned} \lim_{n\to\infty} Pr_n &= \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} \\ &= e^{-1} \approx 0.368 \end{aligned}$$

The convergence is very fast;
e.g. $Pr_n \approx 0.368$ to 3 significant figures for $n \geq 6$.

Example:
A deranged scientist removes the heads from a large number of different animals and re-attaches them at random. What is the probability that every resulting creature is a chimera, i.e. that no head is reattached to its own body?

Answer:
About e^{-1}.

10 Number sequences

A <u>number sequence</u> is simply an infinite sequence $h_0, h_1, h_2, \ldots$ of numbers. For us, h_i will typically be an integer.

Examples:
1,2,3,4,5,...
2,4,8,16,32,...
2,3,5,7,13,...
1,1,2,3,5,8,13,...
1,5,10,10,5,1,0,0,0,0,...

Generating functions

The generating function of a number sequence $h_0, h_1, ...$ is the formal power series $g(x) = \sum_{n=0}^{\infty} h_n x^n$.

Technical remark:
Despite the notation and terminology, we do not assume any convergence;
we do not need $g(a)$ to make sense for a a real number,
so g doesn't really have to be a function in the usual sense.
for example, $\sum_{n=0}^{\infty} n^n x^n$ doesn't converge for $x \neq 0$,
but it's a perfectly good generating function.

We use the usual algebraic notation for generating functions. We can make sense of algebraic operations as follows:

Given formal power series $g(x) = \sum_{n=0}^{\infty} h_n x^n$ and $g'(x) = \sum_{n=0}^{\infty} h'_n x^n$, and a number c, we define

$$g(x) + g'(x) := \sum_{n=0}^{\infty}(h_n + h'_n)x^n$$
$$cg(x) := \sum_{n=0}^{\infty} ch_n x^n$$
$$g(x)g'(x) := \sum_{n=0}^{\infty}(\sum_{j=0}^{n} h_j h'_{n-j})x^n.$$

We also write $c(x) = \frac{a(x)}{b(x)}$ to mean that $a(x) = b(x)c(x)$ (this is well-defined).

We can often use this algebraic structure to write generating functions compactly.

Example 1:
Consider the binomial coefficients $\binom{m}{n}$ for a fixed m.
This is a finite number sequence, but we can make it infinite by appending 0s,

$$\binom{m}{0}, \binom{m}{1}, ..., \binom{m}{m}, 0, 0, 0, ...$$

So the generating function is

$$\binom{m}{0} + \binom{m}{1}x + \binom{m}{2}x^2 ... + \binom{m}{m}x^m$$
$$= (x+1)^m$$

Example 2:
The generating function of the number sequence

$$1, 1, 1, ...$$

is $g(x) = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + ...$

Now, multiplying out,

$$(1 + x + x^2 + ...)(1 - x) = 1 + (-1 + 1)x + (-1 + 1)x^2 + ...,$$

so $g(x) = 1/(1 - x)$.
□

Generating functions provide an efficient notation for describing and manipulating classes of combinatorial problems.

Example 3:
Given t, let h_n be the number of n-combinations of a multiset with t types and infinite multiplicity for each type.

We know $h_n = \binom{n+t-1}{t-1}$, so the generating function is
$$g(x) = \sum_{n=0}^{\infty} h_n x^n = \sum_{n=0}^{\infty} \binom{n+t-1}{t-1} x^n.$$

But note also that
$$g(x) = (1 + x + x^2 + ...)^t,$$
since when we multiply the right hand side out,
the coefficient of x^n is precisely the number of ways of obtaining x^n as $x^{e_1}x^{e_2}...x^{e_t}$,
which is the number of solutions in non-negative integers to $e_1 + ... + e_k = n$, which (as we've seen before) is h_n.

So as in the previous example,
$$g(x) = (1 + x + x^2 + ...)^t = \left(\frac{1}{1-x}\right)^t = \frac{1}{(1-x)^t}.$$
$\square$

Note we found here the power series expansion of $\frac{1}{(1-x)^t}$, which will come in handy later.

Lemma 1:
$$\frac{1}{(1-x)^t} = \sum_{n=0}^{\infty} \binom{n+t-1}{t-1} x^n$$

If we have restrictions on how many of each type we're allowed to take in a combination, we can incorporate these into an algebraic expression for the generating function.

Example 4:
Find the generating function for the number h_n of bags of n marbles consisting of an even number of red marbles, at least 1 green marble, at most 36 blue marbles, and an odd number of yellow marbles.

Arguing as in the previous example, the generating function is
$$g(x) = (1+x^2+x^4+...)(x+x^2+x^3+...)(1+x+x^2+...+x^36)(x+x^3+x^5+...)$$
$$= \frac{1}{1-x^2}\frac{x}{1-x}\frac{1-x^{37}}{1-x}\frac{x}{1-x^2}.$$
$$= \frac{x^2(1-x^{37})}{(1-x^2)^2(1-x)^2}.$$

Example 5:
Find the generating function for the number h_n of bags of n marbles consisting of an even number of red marbles, a multiple of 3 of green marbles, at most 2 blue marbles, and at most one yellow marble. Hence explicitly determine h_n.

$$g(x) = (1 + x^2 + x^4 + ...)(1 + x^3 + x^6 + ...)(1 + x + x^2)(1 + x)$$
$$= \frac{1}{1-x^2}\frac{1}{1-x^3}\frac{1-x^3}{1-x}(1 + x)$$
$$= \frac{1+x}{(1-x^2)(1-x)}$$

$$= \frac{1+x}{(1+x)(1-x)(1-x)}$$
$$= \frac{1}{(1-x)^2}$$
$$= \sum_{n=0}^{\infty} \binom{n+2-1}{2-1} x^n \text{ (by Lemma 1)}$$
$$= \sum_{n=0}^{\infty} (n+1)x^n.$$

So there are $n+1$ such bags of n marbles!

(Exercise: find a direct proof of this, without going via generating functions.)

Example 6:
Find the generating function for the number h_n of ways of making n cents out of Canadian coins.

The coins in current circulation are worth 5,10,25,100, and 200 cents each.

So h_n is the number of solutions in non-negative integers to

$$5N + 10D + 25Q + 100L + 200T = n.$$

Equivalently, h_n is the number of solutions to

$$e_1 + e_2 + e_3 + e_4 + e_5 = n$$

where e_1 is a multiple of 5, e_2 is a multiple of 10, etc.

So as above,

$$g(x) = (x^5 + x^{10} + x^{15} + ...)(x^{10} + x^{20} + ...)...(x^{200} + x^{400} + ...)$$
$$= \frac{1}{(1-x^5)(1-x^{10})...(1-x^{200})}.$$

Exponential Generating Functions

The <u>exponential generating function</u> of a number sequence $h_0, h_1, ...$ is the formal power series

$$g^{(e)}(x) = \sum_{n=0}^{\infty} h_n \frac{x^n}{n!}.$$

While ordinary generating functions are useful for counting combinations, exponential generating functions are useful for counting permutations.

Example 7:
The exponential generating function of

$$(m, 0), P(m, 1), ..., P(m, m), 0, 0, 0, ...$$

is

$$g^{(e)} = \sum_{n=0}^{m} \frac{m!}{(m-n)!} \frac{x^n}{n!}$$
$$= \sum_{n=0}^{m} \binom{m}{n} x^n$$
$$= (1+x)^m$$

Example 8:
Let h_n be the number of n-permutations of a multiset with k different types, each with infinite multiplicity,

$$\{\infty \cdot a_1, ..., \infty \cdot a_k\}.$$

So $h_n = k^n$.
Then the exponential generating function is : $g^{(e)}(x) = \sum_{n=0}^{\infty} \frac{k^n x^n}{n!} = e^{kx}$.

(remark for anyone who might worry what exactly we mean by this last equality: we can just define e^{ax} to be the formal power series $\sum_{n=0}^{\infty} \frac{a^n}{n!} x^n$. This obeys the usual law $e^{ax}e^{bx} = e^{(a+b)x}$. We could define more, but this will suffice for our purposes.)

Theorem:
Let h_n be the number of n-permutations of the multiset
$$S := \{n_1 \cdot a_1, ..., n_k \cdot a_k\},$$
with $n_i \in \mathbb{N} \cup \{\infty\}$.

Then the exponential generating function is
$$g^{(e)} = f_{n_1}(x) f_{n_2}(x) ... f_{n_k}(x)$$
where
$$f_n(x) = \sum_{i=0}^{n} \frac{x^i}{i!}$$
and in particular, $f_\infty(x) = e^x$.

Proof:
$$h_n = \sum_{S' \text{ an } n\text{-combination of } S} (\text{number of permutations of } S')$$
$$= \sum_{\{(m_1,...,m_k) \mid m_1+...+m_k=n,\ 0 \le m_i \le n_i\}} (\text{number of permutations of } \{m_1 * a_1, ..., m_k * a_k\}$$
$$= \sum_{\{(m_1,...,m_k) \mid m_1+...+m_k=n,\ 0 \le m_i \le n_i\}} \frac{n!}{m_1!...m_k!}$$
$$= n! \sum_{\{(m_1,...,m_k) \mid m_1+...+m_k=n,\ 0 \le m_i \le n_i\}} \frac{1}{m_1!...m_k!}$$

Meanwhile, if we multiply out $f_{n_1}(x) f_{n_2}(x) ... f_{n_k}(x)$, we find the coefficient of x^n is
$$= \sum_{\{(m_1,...,m_k) \mid m_1+...+m_k=n,\ 0 \le m_i \le n_i\}} \frac{1}{m_1!...m_k!}.$$
So this is indeed the exponential generating function.
□

Just as we saw with ordinary generating functions, if we have restrictions on how many of each type we are allowed in a permutation, we can incorporate these restrictions into the factors in the above expression for the exponential generating function, by only including the appropriate powers of x.

Often, expanding out the resulting power series will give us a solution to the combinatorial problem, as the following example demonstrates.

Example 9:
How many n-digit numbers can be written using only the digits '1','2', and '3', using an even number of '2's and at least 1 '3'?

The exponential generating function is
$$g^{(e)}(x) = \left(\sum_{n=0}^{\infty} \frac{x^n}{n!}\right)\left(\sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}\right)\left(\sum_{n=1}^{\infty} \frac{x^n}{n!}\right)$$
$$= (e^x)\left(\frac{e^x + e^{-x}}{2}\right)(e^x - 1)$$
$$= \frac{1}{2}(e^{3x} + e^x - e^{2x} - 1)$$
$$= \sum_{n=1}^{\infty} \frac{3^n + 1 - 2^n}{2n!}$$
So the answer is $\frac{3^n+1-2^n}{2}$

11 Recurrence relations

Warm-up: The Fibonacci sequence

The **Fibonacci sequence** is the sequence f_n satisfying

$f_0 = 0, f_1 = 1$

$f_{n+2} = f_n + f_{n+1}$

so

0,1,1,2,3,5,8,13,21,34,55,...

Such an expression for a term in a sequence as a function of previous terms is called a **recurrence relation**.

Other examples:

$h_0 = 1$

$h_{n+1} = h_n + 3$

$h_0 = 1$

$h_{n+1} = 3h_n$

In these cases, we can easily find an expression for h_n in terms of n.

Can we do this for f_n?

To do so, we should consider the more general problem where we vary the initial values f_0 and f_1, and just consider sequences f_n satisfying the recurrence relation $f_{n+2} = f_n + f_{n+1}$.

If f_n and f'_n are two such sequences, then so is $c_1 f_n + c_2 f'_n$ for any c_1, c_2 (i.e. the solutions form a vector space).

So if we can find **some** solutions to the Fibonacci recurrence relation, we can easily generate more - perhaps including the Fibonacci sequence itself.

(In fact, if you recall your linear algebra, you should be able to see that we only need to find two linearly independent sequences to generate all of them)

Let's look for solutions of the particularly simple form

$f_n = q^n$

with $q \neq 0$. Then the recurrence relation becomes

$q^{n+2} = q^n + q^{n+1}$

$\leftrightarrow q^n(q^2 - q - 1) = 0$

$\leftrightarrow (q^2 - q - 1) = 0$ (since $q \neq 0$)

This is a quadratic equation, so it has two solutions.

They are

$\phi = \frac{1+\sqrt{5}}{2}, \phi' = \frac{1-\sqrt{5}}{2}$.

(ϕ is known as the Golden Ratio; it is the unique positive real satisfying $\frac{1+\phi}{\phi} = \phi$)

So for any c_1 and c_2,

$c_1\phi^n + c_2\phi'^n$

satisfies the Fibonacci recurrence relation.

Let's find such a sequence which satisfies the initial conditions of the Fibonacci sequence, $f_0 = f_1 = 1$; then it must be the Fibonacci sequence.

$c_1 + c_2 = 0$
$c_1\phi + c_2\phi' = 1$

We can solve this system of simultaneous equations

$$\begin{pmatrix} 1 & 1 \\ \phi & \phi' \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \frac{1}{\phi' - \phi} \begin{pmatrix} \phi' & -1 \\ -\phi & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

So we obtain

Theorem:
The Fibonacci numbers are

$f_n = \frac{\phi^n - \phi'^n}{\sqrt{5}}$

where

$\phi = \frac{1+\sqrt{5}}{2}, \phi' = \frac{1-\sqrt{5}}{2}$.

Note that $\phi' = 1 - \phi$, so we could also write this as

$f_n = \frac{\phi^n - (1-\phi)^n}{\sqrt{5}}$.

Homogeneous Linear Recurrence Relations with Constant Coefficients

A <u>homogeneous linear recurrence relation with constant coefficients</u> is an equation

$h_{n+k} = a_0 h_n + a_1 h_{n+1} + ... + a_{k-1} h_{n+k-1}$,

with a_i complex numbers.

k is the <u>order</u> of the recurrence relation.
A number sequence h_n satisfying the recurrence relation is called a <u>solution</u> to the recurrence relation.

If we add <u>initial conditions</u>

$h_0 = c_0, ..., h_{k-1} = c_{k-1}$,

this clearly uniquely determines a solution.

Examples:

(i) The Fibonacci sequence.

(ii) Geometric sequences, $h_{n+1} = a h_n$.

(iii) The life-cycle of *inventioni exemplicus* is as follows:
a new hatchling remains in the larval stage until its first summer,
then spends a year maturing, and in the subsequent summer lays a clutch of 7 eggs (which quickly hatch into larvae),
then in the summer after lays a second clutch of 6 eggs, then dies.

All *excemplicus* are female (they reproduce parthenogenetically).

Suppose no *exemplicus* die except at the end of their life cycle.

If 100 exemplicus hatchlings are introduced one summer, how many exemplicus larvae will there be at the end of the nth summer thereafter?

Solution:
At the end of the nth summer, there are 7 larvae born from each 2-year-old *exemplicus*, and 6 from each 3-year-old.

So $h_n = 6h_{n-3} + 7h_{n-2}$ for $n \geq 3$,
i.e. $h_{n+3} = 6h_n + 7h_{n+1}$ for $n \geq 0$.

We also have the initial conditions
$h_0 = 100, h_1 = 0, h_2 = 700.$

So we get
$h_3 = 600, h_4 = 4900, h_5 = 8400, h_6 = 37900, ...$

We proceed to generalise the solution to the Fibonacci recurrence relation to solve general homogeneous linear recurrence relation with constant coefficients.

Given a recurrence relation
$h_{n+k} = a_0 h_n + a_1 h_{n+1} + ... + a_{k-1} h_{n+k-1}$,
i.e. $h_{n+k} = \sum_{j=0}^{k-1} a_j h_{n+j}$,
we again look for solutions $h_n = q^n$.

Clearly $h_n = q^n$ is a solution iff
$q^k = a_0 + a_1 q + ... + a_{k-1} q^{k-1}$,
i.e. $q^k - a_{k-1} q^{k-1} - ... - a_1 q - a_0 = 0$.

The polynomial $x^k - a_{k-1} x^{k-1} - ... - a_1 x - a_0$ is called the characteristic polynomial of the recurrence relation.
It is a degree k polynomial, so has k roots in the complex numbers (counting multiplicities).

Suppose that it has k **distinct** roots, $q_1, ..., q_k$.
(See the "Bonus" section for what happens when we have repeated roots.)

Claim:
the k vectors $((q_1^0, ..., q_1^{k-1}), ..., (q_k^0, ..., q_k^{k-1}))$ are linearly independent in $\mathbb{C}^k$,

Proof:
Otherwise, considering the columns of the k-by-k matrix whose rows are these vectors,

the k vectors $((q_1^0, ..., q_k^0), ..., (q_1^{k-1}, ..., q_k^{k-1}))$ are linearly dependent.
i.e. there are $b_0, ..., b_{k-1} \in \mathbb{C}$ not all 0, such that the polynomial

$$b_0 + b_1 x + ... + b_{k-1} x^{k-1}$$

has roots $q_1, ..., q_k$.

But a degree $k-1$ polynomial can't have k distinct roots;
contradiction.

□

So any given initial conditions $h_0 = c_0, ..., h_{k-1} = c_{k-1}$ can be written as a linear combination

$$h_n = b_1 q_1^n + ... + b_k q_k^n = \sum_{i=1}^k b_i q_i^n.$$

Taking this as a definition of h_n for all n,
we see that not only does it satisfy the initial conditions by choice of b_i,
but it satisfies the recurrence relation; indeed

$$\begin{aligned} h_{n+k} &= \sum_{i=1}^k b_i q_i^{n+k} \\ &= \sum_{i=1}^k b_i q_i^n q_i^k \\ &= \sum_{i=1}^k b_i q_i^n (\sum_{j=0}^{k-1} a_j q_i^j) \\ &= \sum_{j=0}^{k-1} \sum_{i=1}^k b_i q_i^n a_j q_i^j \\ &= \sum_{j=0}^{k-1} a_j \sum_{i=1}^k b_i q_i^{n+j} \\ &= \sum_{j=0}^{k-1} a_j h_{n+j} \end{aligned}$$

Example:
Let's solve the *exemplicus* example.

$$h_{n+3} = 6h_n + 7h_{n+1},$$
$$h_0 = 100, h_1 = 0, h_2 = 700.$$

The characteristic polynomial is

$$x^3 - 7x - 6 = (x-3)(x+2)(x+1)$$

so the solutions are of the form

$$h_n = b_1 3^n + b_2(-2)^n + b_3(-1)^n.$$

Solving

$$100 = h_0 = b_1 + b_2 + b_3$$
$$0 = h_1 = 3b_1 - 2b_2 - b_3$$
$$700 = h_2 = 9b_1 + 4b_2 + b_3$$

gives

$$b_1 = 45, b_2 = 80, b_3 = 25.$$

So the solution is

$$h_n = 45 * 3^n + 80 * (-2)^n - 25 * (-1)^n.$$

□

Bonus: Solving recurrence relations with generating functions

Generating functions provide a convenient device for solving recurrence relations (although in theoretical terms, they only provide a different way to package the same linear algebra).

If $g(x)$ is the generating function for the sequence h_n,
i.e. the coefficient of x^n in $g(x)$ is h_n,
then the coefficient of x^{n+1} in $xg(x)$ is h_n.

So if h_n satisfy a recurrence relation
$h_{n+k} = a_0 h_n + a_1 h_{n+1} + ... + a_{k-1} h_{n+k-1}$
then in
$g(x) - a_0 x^k g(x) - a_1 x^{k-1} g(x) - ... - a_{k-1} x g(x)$,
x^{n+k} has coefficient 0 for $n \geq 0$,
i.e. this is a polynomial of order $k-1$.

Using initial conditions, we can find this polynomial, and so express $g(x)$ as a rational function.

For example, consider the Fibonacci relations $f_{n+2} = f_n + f_{n+1}$, $f_0 = 0$, $f_1 = 1$.

If $g(x)$ is the generating function, then
$g(x) - xg(x) - x^2 g(x) = f_0 + f_1 x + f_2 x^2 + f_3 x^3 + ... - f_0 x - f_1 x^2 - f_2 x^3 - ... - f_0 x^2 - f_1 x^3 - ...$
$= f_0 + (f_1 - f_0)x + (f_2 - f_1 - f_0)x^2 + (f_3 - f_2 - f_1)x^3 + ...$
$= f_0 + (f_1 - f_0)x$
$= 0 + (1 - 0)x$
$= x$,
so
$(1 - x - x^2)g(x) = x$
so
$g(x) = \frac{x}{1-x-x^2}$.

Furthermore, by factoring the denominator and finding partial fractions, we can expand this as a power series and so solve the recurrence equations.
In this case, the solutions to $1 - x - x^2 = 0$ are the reciprocals of the solutions ϕ, ϕ' to $x^2 - x - 1$, so
$g(x) = \frac{x}{1-x-x^2}$
$= \frac{x}{(x-\phi^{-1})(x-\phi'^{-1})}$
$= \frac{\phi\phi' x}{(1-\phi x)(1-\phi' x)}$
$= \frac{a}{1-\phi x} + \frac{b}{1-\phi' x}$
where
$0 = a + b$
$\phi\phi' = -\phi' a - \phi b$
$=> b = -a$
$=> \phi\phi' = a(\phi - \phi')$
$=> a = \frac{1}{\sqrt{5}}$ (using the definitions of ϕ, ϕ')
$=> b = \frac{-1}{\sqrt{5}}$
so

$$\begin{aligned} g(x) &= \tfrac{1}{\sqrt{5}}((1-\phi x)^{-1} - (1-\phi' x)^{-1}) \\ &= \tfrac{1}{\sqrt{5}}((\textstyle\sum_{n=0}^{\infty} \phi^n) - (\sum_{n=0}^{\infty} \phi'^n)) \\ &= \tfrac{1}{\sqrt{5}} \textstyle\sum_{n=0}^{\infty} (\phi^n - \phi'^n) \end{aligned}$$

so we reclaim the formula we found before,

$$f_n = \frac{\phi^n - \phi'^n}{\sqrt{5}},$$

and we had to do the same algebra to get there.

Bonus: abstract reformulation, and handling repeated roots

Consider the (infinite dimensional) complex vector space of complex number sequences $h = h_0, h_1, \ldots$. Let σ be the downshift operator, which from a sequence h obtains the new sequence $(\sigma h)_n = h_{n+1}$. Note that this is a linear map.

Then a linear homogeneous constant coefficient recurrence relation, which we can write as $\sum_{i=0}^{k} a_i h_{n+i} = 0$, with $a_k \neq 0$, can be rewritten as

$$(\textstyle\sum_{i=0}^{k} a_i \sigma^i) h = 0.$$

The subspace of solutions to this is then the kernel of the linear operator $\sum_{i=0}^{k} a_i \sigma^i$. This is a finite dimensional vector space. The solution method described above is a matter of finding a basis of eigenvectors of σ on this space. Note that eigenvectors are precisely geometric series cq^n.

In general of course, the eigenvectors of σ won't span the space. But its generalised eigenvectors will. One can check inductively that the kth generalised q-eigenspace of σ, i.e. the kernel of $(\sigma - q)^k$,
is the space of sequences $f(n)q^n$ where f is a polynomial of degree at most $k-1$. So if the characteristic polynomial factors as

$$\textstyle\sum_{i=0}^{k} a_i \sigma^i = \Pi(\sigma - q_i)^{k_i},$$

the space of solutions has a basis of eigensequences

$$n^j q_i^n \text{ where } j < k_i,$$

so any solution can be expressed as a linear combination of these.

We might as well note that this abstract formulation also applies to homogeneous linear differential equations over the constants: replace the space of sequences with, say, the space of complex analytic functions in one variable, and replace σ with the differentiation operator. The kth generalised λ-eigenspace consists of $f(x)e^{\lambda x}$ for $f(x)$ a polynomial of degree at most $k-1$.

12 Difference sequences, sums of powers, and Stirling numbers

Difference sequences

Notation:
If $h_0, h_1, \ldots$ is a number sequence, we will sometimes refer to the sequence

just as h.

Definition:
Δ is the operator on number sequences of taking successive differences;
for a number sequence h, the number sequence Δh is defined by

$$\Delta h_n = h_{n+1} - h_n.$$

$\Delta^2 h_n = \Delta\Delta h_n$, etc.

We write out a sequence and its iterated differences as an infinite triangle;
e.g. if $h_n = n^2$, the iterated differences are as follows:

```
1    4    9    16   25 ...
   3    5    7    9 ...
      2    2    2 ...
         0    0 ...
            0 ...
```

Remark:
Δ is a linear operator, i.e. for sequences h and h' and numbers c and c',

$$\Delta(ch + c'h')_n = c\Delta h_n + c'\Delta h'_n.$$

Hence the powers Δ^k are also linear.

Lemma:
Let f be a polynomial of degree at most d, and let $h_n = f(n)$ be the sequence of its values on natural numbers.

Then $\Delta^{d+1} h_n = 0$ for all n.

Proof:
By linearity, it suffices to show this for monomials $f(x) = x^d$.

So let $h_n = n^d$, and suppose inductively that the lemma holds for polynomials of degree less than d.

Then

$$\begin{aligned}\Delta h_n &= h_{n+1} - h_n = (n+1)^d - n^d \\ &= n^d + dn^{d-1} + \tbinom{d}{2} n^{d-2} + ... + 1 - n^d \\ &= \tbinom{d}{1} n^{d-1} + \tbinom{d}{2} n^{d-2} + ... + 1\end{aligned}$$

which has degree $d - 1$.

So by the inductive hypothesis,
$0 = \Delta^d \Delta h_n = \Delta^{d+1} h_n$. □

Now suppose $h_n = f(n)$ with f a polynomial of degree d.
By the above lemma, the numbers $h_0, \Delta h_0, ..., \Delta^d h_0$ determine the whole sequence h,
since we can generate the whole triangle from the initial diagonal $h_0, \Delta h_0, ..., \Delta^d h_0, 0, 0, ...$

Let's find a formula for h_n in terms of $h_0, \Delta h_0, ..., \Delta^d h_0$.

Generating the triangle is a linear process,
so if we can find a formula for h_n generated from an initial diagonal $0, 0, ..., 0, 1, 0, 0, ...$,
with $\Delta^k h_0 = 1$ and all other $\Delta^i h_0 = 0$,
we can then take a linear combination.

We get a "twisted Pascal's triangle", e.g.:

```
0   0   0   0   1   5   15  35  70 ...
  0   0   0   1   4   10  20  35 ...
    0   0   1   3   6   10  15 ...
      0   1   2   3   4   5 ...
        1   1   1   1   1 ...
          0   0   0   0 ...
```

and so we see that $h_n = \binom{n}{k}$.

To prove this: let $f(x) := \frac{x(x-1)(x-2)...(x-(k-1))}{k!}$;
then $f(0) = f(1) = ... = f(k-1) = 0$ and $f(k) = 1$,
so the difference triangle of $f(n)$ also starts with

```
0   0   0   0   1
  0   0   0   1
    0   0   1
      0   1
        1
```

, and since by the lemma it also has 0s thereafter,
we must have $h_n = f(n)$.

Then we directly calculate that $f(n) = \binom{n}{k}$.

So, taking linear combinations, we conclude :

Theorem:
If the initial diagonal of the difference triangle of h_n is $c_0, c_1, ..., c_d, 0, 0, ...$
(i.e. if $\Delta^k h_0 = c_i$ for $k \leq d$, and $\Delta^k h_0 = 0$ for $k > d$),
then
$$h_n = \sum_{k=0}^{d} c_k \binom{n}{k}.$$

Sums of powers

We can use this theorem to give neat formulae for sums of powers $\sum_{n=0}^{k} n^d$,
generalising the formulae you know and love for $d = 1$ and $d = 2$ (and maybe even $d = 3$, if you're that generous with your affections),
and more generally to give formulae for $\sum_{n=0}^{k} f(n)$ where f is any polynomial.

First recall the formula (from the section on Binomial Coefficients)
$$\binom{k+1}{r+1} = \sum_{n=0}^{k} \binom{n}{r}.$$

So to find $\sum_{n=0}^{k} f(n)$, we can first use the above theorem to find an expression for $f(n)$ in terms of binomial coefficients, then use this formula to sum them.

Example:
Let's find a formula for $\sum_{n=0}^{k} n^4$.

Drawing the start of the difference triangle,

```
0   1    16   81   256
  1    15   65   175
    14   50   110
       36   60
         24
```

, and recalling that all further rows are 0 since n^4 has degree 4, we see that the initial diagonal is $0, 1, 14, 36, 24, 0, 0, \ldots$.

So by the above theorem,
$$n^4 = \binom{n}{1} + 14\binom{n}{2} + 36\binom{n}{3} + 24\binom{n}{4}.$$

So using the formula
$$\binom{k+1}{r+1} = \sum_{n=0}^{k} \binom{n}{r},$$
we find
$$\sum_{n=0}^{k} n^4$$
$$= \sum_{n=0}^{k} \left(\binom{n}{1} + 14\binom{n}{2} + 36\binom{n}{3} + 24\binom{n}{4}\right)$$
$$= \binom{k+1}{2} + 14\binom{k+1}{3} + 36\binom{k+1}{4} + 24\binom{k+1}{5}$$

Exercise:
Repeat this procedure for n^1, n^2 and n^3, and check that the answers you get agree with the standard formulae.

Stirling numbers

We would like to understand the mysterious numbers which appear in the formula for $\sum_{n=0}^{k} n^p$,
i.e the numbers $c(p, k)$ defined by
$$c(p, k) := \Delta^k h_0 \text{ where } h_n = n^p.$$

So as we saw, these are the numbers $c(p, k)$ such that
$$n^p = \sum_{k=0}^{p} c(p, k)\binom{n}{k}.$$

We observe (and will eventually prove) that $c(p, k)$ seems to be divisible by $k!$, so set
$$S(p, k) := c(p, k)/k!.$$

So, introducing the notation $[n]_k := P(n, k) = k!\binom{n}{k}$,
$$n^p = \sum_{k=0}^{p} S(p, k)[n]_k.$$

These numbers $S(p, k)$ are the Stirling numbers of the second kind.

Here's a table, written in Pascal triangle format with k going across and p going down, and starting with $S(1,1) = 1$:

```
          1
        1   1
      1   3   1
    1   7   6   1
  1  15  25  10   1
1  31  90  65  15   1
```

This corresponds to the formulae

$n^1 = [n]_1$
$n^2 = [n]_1 + [n]_2$
$n^3 = [n]_1 + 3[n]_2 + [n]_1$
...

All values of $S(p,k)$ not shown in the triangle are 0, except $S(0,0) = 1$.

Lemma:
For all $p > 0$, and all k,
$S(p,k) = S(p-1,k-1) + kS(p-1,k)$.

Proof:
First, note that $S(p,k) = 0$ when $k > p$, by considering degrees of polynomials.
Also $S(p,k) = 0$ when $k < 0$, by definition.

Now

$$\begin{aligned} n^p = nn^{p-1} &= n\sum_{k=0}^{p-1} S(p-1,k)[n]_k \\ &= \sum_{k=0}^{p-1} S(p-1,k)((n-k)+k)[n]_k \\ &= \sum_{k=0}^{p-1} S(p-1,k)[n]_{k+1} + \sum_{k=0}^{p-1} kS(p-1,k)[n]_k \\ &= \sum_{k=1}^{p} S(p-1,k-1)[n]_k + \sum_{k=0}^{p-1} kS(p-1,k)[n]_k \\ &= \sum_{k=0}^{p} S(p-1,k-1)[n]_k + \sum_{k=0}^{p} kS(p-1,k)[n]_k \\ &\qquad \textit{(using } S(p-1,-1) = 0 = S(p-1,p)\textit{)} \\ &= \sum_{k=0}^{p}(S(p-1,k-1) + kS(p-1,k))[n]_k, \end{aligned}$$

so we conclude by comparing coefficients with
$n^p = \sum_{k=0}^{p} S(p,k)[n]_k$. □

Theorem:
$S(p,k)$ is the number of partitions of a set of p objects into k indistinguishable boxes in which no box is empty,
i.e. the number of partitions of a set of size p into a set of k non-empty subsets,
i.e. the number of sets of non-empty subsets of $\{1, ..., p\}$ which are disjoint and have union $\{1, ..., p\}$.

Proof:
Write $S'(p,k)$ for this number.

Suppose $p \geq 1$ and $1 \leq k \leq p$.
Consider a partition of $\{1, ..., p\}$ into a set of k non-empty subsets, and consider removing p.
First, suppose the set in the partition which contains p is just $\{p\}$.
Then on removing p, we obtain a partition of $\{1, ..., p-1\}$ into $k-1$ subsets.
Otherwise, on removing p we obtain a partition of $\{1, ..., p-1\}$ into k subsets.
In the first case, the map is bijective, but in the second case there are k ways of obtaining the same partition of $\{1, ..., p-1\}$, since p could have been removed from any of the k sets in that partition.

So

$$S'(p, k) = S'(p-1, k-1) + kS'(p-1, k).$$

Clearly $S'(p, k) = 0$ for $k < 0$ or $k > p$ or $p < 0$, and $S(0, 0) = 1$.
So by induction on p, $S(p, k) = S'(p, k)$ for all p and k.
□

So now we know that $S(p, k)$ is an integer.

Moreover, we can now reason combinatorially to find a formula for $S(p, k)$:

Theorem:
For $p \geq 0$ and $0 \leq k \leq p$,

$$S(p, k) = \sum_{i=0}^{k} (-1)^i \frac{(k-i)^p}{i!(k-i)!}$$

Proof:
Fix p and k.

Let P be the number of partitions of $\{1, ..., p\}$ into an **ordered sequence** of k non-empty subsets.
So $P = k!S(p, k)$.

A partition of $\{1, ..., p\}$ into an ordered sequence of k subsets,
with no restrictions on the subsets being non-empty,
just corresponds to a k-colouring of $\{1, ..., p\}$,
i.e. a choice of which of the k sets in the partition each element should go in,
so there are k^p such partitions.

Let A_i be the partitions of $\{1, ..., p\}$ into an ordered sequence of k subsets, where the ith is empty.

Such a partition corresponds to a partition into $k-1$ possibly empty subsets, by ignoring the one which is required to be empty.
So $|A_i| = (k-1)^p$.

Similarly, $|A_i \cap A_j| = (k-2)^p$ for $i \neq j$, and generally $|\bigcap_{i \in I} A_i| = (k - |I|)^p$.

So by inclusion-exclusion,

$$S(p, k) = \tfrac{1}{k!} P$$

$$
\begin{aligned}
&= \tfrac{1}{k!}(k^p - |\bigcup_i A_i|) \\
&= \tfrac{1}{k!}(k^p - \sum_{\emptyset \neq I \subseteq \{1,\dots,k\}} (-1)^{|I|-1} |\bigcap_{i\in I} A_i|) \\
&= \tfrac{1}{k!}(k^p - \sum_{i=1}^{k} (-1)^{i-1} \tbinom{k}{i} (k-i)^p) \\
&= \tfrac{1}{k!}(\sum_{i=0}^{k} (-1)^{i} \tbinom{k}{i} (k-i)^p) \\
&= \sum_{i=0}^{k} (-1)^{i} \tfrac{(k-i)^p}{i!(k-i)!})
\end{aligned}
$$

□

13 Heterosexuality, and Hall's Marriage Theorem

Basic "marriage" problem

A set A of men;
a set B of women;
certain pairings (a, b) are "compatible".

A **matching** is a choice of some compatible pairings ("marriages"), such that no man is paired to multiple women and no woman to multiple men.

Can we find a matching in which every woman is paired with some man?

Graphical formulation

(You can skip this if you don't know what a graph is.)

Bipartite graph: two collections of A and B of vertices; all edges have one vertex in A and the other in B.

A matching is a set of edges such that no vertex is incident to more than one of the edges.

Can we find a matching such that every $b \in B$ is incident to one of the edges?

SDR formulation

Given a family of subsets $(A_i)_{i\in B}$ of a set A,
can we find elements $a_i \in A_i$ which are distinct, i.e. $a_i \neq a_j$ when $i \neq j$?

Such a selection of $a_i \in A_i$ for $i \in B$ is called a system of distinct representatives (SDR).

Other instances

Example:
Given an $m \times n$ board with certain squares missing,

can we place a rook on each row such that no row or column contains two rooks?

Here we marry rows and columns, consecrating the marriage with the symbolic placing of a rook.

Example:
Certain jobs are to performed by certain people, one job by each;
not all people are suitable for all jobs.
Can we assign each job to some person?

Example:
Various classes are to be scheduled at various times in various rooms;
some classes require certain rooms,
and some times are unsuitable for some classes.

Suppose all classes are to start on the hour.

Can all classes be scheduled?

Here, we marry classes and hour-room pairs.

Hall's Marriage Theorem

Definition:
A family $(A_i)_{i\in B}$ of sets satisfies the marriage condition if for any k, the union of any k of the sets in the family has size at least k;
i.e. for every $B' \subseteq B$,

$$\left|\bigcup_{i\in B'} A_i\right| \geq |B'|.$$

Theorem [Philip Hall, 1935]:
A family $(A_i)_{i\in B}$ of finite sets has a system of distinct representatives iff it satisfies the marriage condition.

Proof:
The marriage condition is necessary, since if $a_i \in A_i$ is an SDR and $B' \subseteq B$

$\bigcup_{j\in B'} A_j \supseteq \{a_j \mid j \in B'\}$

so, by distinctness,

$\left|\bigcup_{j\in B'} A_j\right| \geq |\{a_j \mid j \in B'\}| = |B'|.$

Now suppose the marriage condition holds, and suppose inductively that the theorem holds when B is smaller.

Suppose first that for all $B' \subsetneq B$,

$$\left|\bigcup_{i \in B'} A_i\right| \geq |B'| + 1.$$

Let $i_0 \in B$,
and let $a_0 \in A_{i_0}$. We proceed by deleting i_0 and a.
Let $A'_i := A_i \setminus \{a_0\}$;
then $(A'_i)_{i \in B \setminus \{i_0\}}$ also satisfies the marriage condition.
Indeed, for any $B' \subseteq B \setminus \{i_0\}$,

$$\begin{aligned} |\textstyle\bigcup_{j \in B'} A'_i| &\geq |\textstyle\bigcup_{j \in B'} A_i| - 1 \\ &\geq (|B'| + 1) - 1 \\ &= |B'|. \end{aligned}$$

So by the inductive hypothesis, this family has an SDR,
adjoining a_0 to which yields an SDR for $(A_i)_{i \in B}$.

Otherwise, say $B^0 \subsetneq B$ with

$$\left|\bigcup_{i \in B^0} A_i\right| = |B^0|.$$

Let $A^0 := \bigcup_{i \in B^0} A_i$.
We proceed by deleting B^0 and A^0.
Let $A'_i := A_i \setminus A^0$, and consider $(A'_i)_{i \in (B \setminus B^0)}$.
Then for $B' \subseteq B \setminus B^0$,

$$\begin{aligned} |\textstyle\bigcup_{j \in B'} A'_i| &\geq |\textstyle\bigcup_{j \in B' \cup B^0} A_i| - |B^0| \\ &\geq |B' \cup B^0| - |B^0| \\ &= |B'| \end{aligned}$$

So by the inductive hypothesis, this family has an SDR.

Now $(A_i)_{i \in B^0}$ also satisfies MC,
and $B^0 \subsetneq B$,
so by the inductive hypothesis,
this family also has an SDR.

Since $A'_i \cap A_j = \emptyset$ for $j \in B^0$,
the union of these two SDRs is an SDR for $(A_i)_{i \in B}$.
□

Dominoes on a chessboard

Consider an $m \times n$ chequered board,
meaning that each square is either white or black, and no two neighbouring squares are of the same colour,
and suppose we delete some squares.

e.g.

, · , ,

```
, .   . , .
.     , .
,     . , .
```

(here I use dots for white squares, commas for black squares, and spaces for deleted squares.).

Can we put 2×1 dominoes on the non-deleted squares of the board, such that the dominoes don't overlap and every non-deleted square is covered?

e.g. here's a solution in the case of the above example, using 8 dominoes denoted by the corresponding numbers:

```
  3 3 4   5
2 2   4 6 5
1     7 6
1     7 8 8
```

.

We can view this as a marriage problem: we want to marry white squares to black squares, and only adjacent squares are compatible.

So we can solve the problem iff there are as many white squares as black squares and the family of sets of (say) black squares adjacent to the white squares satisfies MC.

Exercise:
Draw the corresponding bipartite graph for the example board above.

Find some boards which don't satisfy the marriage condition,
then try (and fail) to cover them with dominoes.

Latin squares

Definition:
A Latin rectangle is an $m \times c$ array with each entry an element of $\{1, ..., n\}$, such that no number appears twice in any row or column.

Theorem:
Any $m \times n$ Latin rectangle with $m < n$ can be completed by adding rows to form an $n \times n$ Latin square.

Proof:
By induction, it suffices to show that we can add a row to form an $m \times n + 1$ Latin rectangle.

Let A_i, for $i = 1, ..., n$, be the set of numbers in $\{1, ..., n\}$ **not** appearing in the ith column.

Then $(A_i)_{i\in\{1,...,n\}}$ satisfies the marriage condition.
Indeed suppose $I \subseteq \{1, ..., n\}$.
Each number in $\{1, ..., n\}$ occurs in each of the m rows,
and so occurs in at most $n - m$ of the A_i.
Each A_i has size $n - m$.
So

$$|I|(n - m) = \sum_{i\in I} |A_i| \leq (n - m)|\bigcup_{i\in I} A_i|,$$

so

$$|\bigcup_{i\in I} A_i| \geq |I|.$$

So by Hall's Marriage Theorem, an SDR $(a_i \in A_i)_{i\in\{1,...,n\}}$ exists;
by distinctness and the definition of A_i, adding this as a row yields an $m \times n + 1$ Latin rectangle. □

14 Impartial Games

An impartial game is a two-player game in which players take turns to make moves, and where the moves available from a given position don't depend on whose turn it is.

A player loses if they can't make a move on their turn (i.e. a player wins if they move to a position from which no move is possible).

In games which are not impartial, the two players take on different roles (e.g. one controls white pieces and the other black), so the moves available in a given position depend on whose turn it is. Before we treat these more complicated games, we first consider the special case of impartial games.

Any nim position is an impartial game. Write $*n$ for the nim position with a single heap of size n.

In particular, $*0$ is the "zero game", written 0: the player to move immediately loses.

We consider a position in a game as a game in itself.
We can specify a game by giving the set of its options - the games to which the first player can move.
e.g. $0 = \emptyset$, $*1 = \{0\}$, $*n = \{*0, *1, *2, ..., *(n - 1)\}$.

For impartial games G and H, $G + H$ is the game where the two games are played side-by-side, with a player on their turn getting to decide which of the two to move in.

So e.g. $*2 + *2 + *3$ is the nim position with two heaps of size 2 and one of size 3.

If $G = \{G_1, ..., G_n\}$ and $H = \{H_1, ..., H_n\}$, then

$$G + H = \{G_1 + H, ..., G_n + H, G + H_1, ..., G + H_n\}.$$

Remark:

- $G + H = H + G$
- $(G + H) + K = G + (H + K)$

We refer to the player who is to move first as "Player 1", and their opponent as "Player 2". So after a move, Player 1 in the new game was Player 2 in the original game, and vice versa.

The finite (or short) games are defined recursively by:

- 0 is finite
- if $G_1, ..., G_n$ are finite, then $\{G_1, ..., G_n\}$ is finite.

Lemma:
In any finite impartial game G, either Player 1 has a winning strategy, or Player 2 has.

We say that G is a "first/second-player win" accordingly, and that the outcome of G is a win for Player 1/2.

Proof:
Suppose inductively that the lemma holds for all options of G.

If some option of G is a win for Player 2, then Player 1 can win in G by making that move.

Else, Player 2 can win in G, since they have a winning strategy in whatever game Player 1 moves to.

Green Hackenbush

A green hackenbush game consists of some dots joined by lines, with some dots "on the ground"; a move comprises deleting a line, and then deleting all lines which are no longer connected to the ground.

Example:

```
    *           *         *         *         *
    |           |         |         |         |
    |           |         |         |         |
    *           *         *         *         *         *
   / \           \       /         / \       / \       / \
  /   \           \     /         /   \     /   \     /   \
 *     * = { *     * , *     * , *     * , *     * , *     * }
  \   /       \   /     \   /         /     \         \   /
   \ /         \ /       \ /         /       \         \ /
----*----   ----*---- ----*---- ----*---- ----*---- ----*----
```

```
    *             *
    |             |
    |             |
    *             *           *
   / \             \         / \
  /   \             \       /   \
 *     *  = {        *  ,  *     *  ,        *  ,           }
      /             /           /           /
     /             /           /           /
----*----     ----*----   ----*----   ----*----   ----*----
```

$P + Q =$ [picture where we draw P and Q side by side, with no connections between them]

Equivalence of impartial games

Definition:
Two games G and H are equivalent, $G \equiv H$,
if for any game K,
$G + K$ has the same outcome as $H + K$.

Remark:

- If $G \equiv H$ then G has the same outcome as H;
- $0 + G \equiv G$;
- if $G \equiv H$ then $G + K \equiv H + K$ for any K.

Lemma:
$G \equiv 0$ iff G is a second-player win.

Proof:
$\Rightarrow$:
If $G \equiv 0$, then $G = G + 0$ has the same outcome as $0 + 0 = 0$, which is a second-player win.

$\Leftarrow$:
Suppose G is a second player win, and K is any game.
We want to show $G + K$ has the same outcome as $0 + K = K$.
The following is a winning strategy in $G + K$ for whichever player has the winning strategy in K:

On our turn, play the next move in the winning strategy for K,
unless our opponent just played in G - then play the next move in the winning strategy for G as Player 2. $\square$

Example:
$G + G \equiv 0$

Indeed, the second player wins by mirroring any move made in one copy of G in the other copy of G.

Lemma:
$G + H \equiv 0$ iff $G \equiv H$

Proof:
If $G \equiv H$, then $G + H \equiv G + G \equiv 0$ by the above example.

If $G + H \equiv 0$, then
$$G = G + 0 \equiv G + H + H \equiv 0 + H = H. \qquad \square$$

Exercise:
Confirm, by finding a winning strategy for Player 2 in the sum, that

```
    *
    |
    |
    *
   / \      ==   *1
  /   \
 *     *
  \   /
   \ /
----*----
```

Remark:
One easily checks that games with equivalent options are equivalent,
i.e. if $G_i \equiv G_i'$, then $\{G_1, ..., G_n\} \equiv \{G_1', ..., G_n'\}$.

From now on, we will just write $=$ in place of $\equiv$ - we are only interested in games up to equivalence.

Our winning strategy for nim tells us how to add nim heaps:

Proposition:
$*n + *m = *(n \oplus m)$

Proof:
$n \oplus m \oplus (n \oplus m) = 0$,
so as we proved when discussing Nim,
$*n + *m + *(n \oplus m)$ is a second-player win. $\square$

It follows that every nim game is equivalent to a single heap, with size the nim sum of the sizes of the heaps in the original game.

Example:

```
    *      *
    |       \
    |        \
    *         *
   / \       /    =  *1 + *3  =  *(1 (+) 3)  =  *2
  /   \     /
 *     *   *
  \   /    |
   \ /     |
----*------*-
```

Impartial games are secret heaps

Now we show that not only every Nim game, but every impartial game is equivalent to a Nim heap.

Theorem [Sprague-Grundy]:
Any finite impartial game G is equivalent to a nim heap;
$G = *n$ where n is the smallest non-negative integer such that G has no option equivalent to $*n$.

Proof:
Suppose inductively that the theorem holds for all options of G.

We show that we can win in $G + *n$ as the second player.

If Player 1 moves in $*n$, say to $*m$ with $m < n$,
then by definition of n we can move in G to a game equivalent to $*m$,
yielding a game equivalent to $*m + *m = 0$.
So we win.

If Player 1 moves in G, picking an option G_i of G,
then by the inductive hypothesis, G_i is equivalent to some $*m$.

So we have to win as first player in $*m + *n$,
but by definition of n, $m \neq n$, so $*m + *n = *(m \oplus n) \neq *0$.
Again, we win. $\square$

Exercise:
Confirm by using the Sprague-Grundy theorem that

```
    *
    |
    |
    *
   / \       =  *1
  /   \
 *     *
  \   /
   \ /
----*----
```

Exercise:
By using the Sprague-Grundy theorem to find the nim heaps equivalent to each component, and then applying the winning strategy from nim, find a winning move for the first player in the following Green Hackenbush game.

```
          *
          |
          |
    *     *     *    *
   / \     \   /    / \
  /   \     \ /    /   \
 *     *     *    *-----*
  \   /      |     \   /
   \ /       |      \ /
----*--------*-------*----
```

15 Bonus: Games

Now, we consider dropping the impartiality condition.

A (combinatorial) game is played by two players, Left and Right, taking turns.
A player who can't move loses.

A game is given by the moves available to Left (Left-options) and to Right (Right-options).

We write $G = \{L_1, ..., L_n \mid R_1, ..., R_m\}$ for the game with Left-options L_i and Right-options R_i, which are themselves games.

Finiteness is defined as in the impartial case.

A game is impartial iff the set of Right-options is equal to the set of Left-options, and all options are impartial.

Real-life games

Chess nearly fits into this framework - but it's possible for the game to end in a draw, which we haven't allowed in our formalism, and it isn't finite.

Go fits even better, though rare loopy situations make it technically infinite.

Some other games of variable notoriety which fit, at least roughly, our definition of a game: draughts/chequers, pente, gess, khet, dots-and-boxes, sprouts, connect 4, gomoku, tic-tac-toe, hex, Y, shogi, xiàngqí, hnefatafl.
See `https://en.wikipedia.org/wiki/List_of_abstract_strategy_games` for many more.

Red-Blue Hackenbush

Like Green Hackenbush, but each line is either red or blue. Only Right can cut Red lines, only Left can cut bLue lines.

Domineering

Played on an 8x8 (say) chess board; players take turns to place 2x1 dominoes on free squares. Left places her dominoes vertically, Right places his horizontally.

Basic theory

We still have the impartial games $*n$,
$$*n = \{*0, *1, ..., *(n-1) \mid *0, *1, ..., *(n-1)\}.$$
We call these nimbers.
We write $*$ for $*1 = \{0 \mid 0\}$.

$G+H$ is the game where players choose which of G and H to move in; Left plays as Left in both, Right as Right in both.

Outcomes:
Now there are four possibilities:

- Right wins (whoever moves first);
- Left wins (whoever moves first);
- Player 1 (whether that's Left or Right) wins;
- Player 2 (whether that's Left or Right) wins.

Exercise:
Determine the outcomes of these four basic domineering games:

```
                               ***
*        **        **          *
*                  *           *
```

We define equivalence ("$G = H$") as in the impartial case, but with this notion of outcome.

By the same arguments as in the impartial case, we have:

Lemma:
$G = 0$ iff G is a second-player win.

Definition:
The negative of a game $G = \{L_1, ..., L_n \mid R_1, ..., R_m\}$ is the game $-G$ in which Left and Right switch roles, i.e.
$-G := \{-R_1, ..., -R_m \mid -L_1, ..., -L_n\}$.

Note that G is impartial iff $-G = G$.

Write $G - H$ for $G + (-H)$.

$G - G = G + (-G)$ is a second-player win, by mirroring moves made in one component in the other.

Just as in the impartial case, we have

Lemma:
$G - H = 0$ iff $G = H$

Proof:
If $G = H$, then $G - H = H - H = 0$.

If $G - H = 0$, then
$$G = G + 0 = G + H - H = G - H + H = 0 + H = H. \qquad \square$$

Definition:
For games G and H, we say

- $G > H$ if $G - H$ is a win for Left;
- $G < H$ if $G - H$ is a win for Right;
- $G = H$ if $G - H$ is a win for Player 2;
- $G \parallel H$ ("G is confused with H") if $G - H$ is a win for Player 1.

Lemma:
$\leq$ is a partial order, and addition respects it,
i.e. $G \leq H \Rightarrow G + K \leq H + K$.

Numbers

$1 = \{0 \mid \}$,
$2 = \{1 \mid \}$,
$3 = \{2 \mid \}$,
etc.

So $-1 = \{ \mid 0\}$,
$-2 = \{ \mid -1\}$,
$-3 = \{ \mid -2\}$,
etc.

$\frac{1}{2} = \{0 \mid 1\}$.

Exercise:
By e.g. considering the following red-blue hackenbush position (L means bLue, R means Red), prove
$\frac{1}{2} + \frac{1}{2} - 1 = 0$

```
     *       *
    R|      R|
     |       |
     *       *      *
    L|      L|     R|
     |       |      |
 ----*-------*------*----
```

$\frac{1}{4} = \{0 \mid \frac{1}{2}\}$.

Exercise:
Confirm that

```
        *
       R|
        |
        *       1
       R|    =  -
        |       4
        *
       L|
        |
    ----*---
```

Generally, we can define

$\frac{1}{2^{n+1}} = \{0 \mid \frac{1}{2^n}\}.$

and then for $m \in \mathbb{N}$,

$\frac{m}{2^n} = m\frac{1}{2^n} =$ [the sum of m copies of $\frac{1}{2^n}$]

$\frac{-m}{2^n} = -(m\frac{1}{2^n})$

Definition:
A finite game is a <u>number</u> if it is equivalent to one of these games $\frac{m}{2^n}$ (with $m \in \mathbb{Z}$, $n \in \mathbb{N}$).

Lemma:
The usual ordering on numbers agrees with the definition of $<$ for games.

Lemma:
If G is a finite game,
and every Left-option is less than every Right-option,
then G is the simplest number $\frac{m}{2^n}$ which is greater than every Left-option and less than every Right-option,
where $\frac{m}{2^n}$ is <u>simpler</u> than $\frac{m'}{2^{n'}}$ if $n < n'$, or if $n = n'$ and $|m| < |m'|$.

Proof:
We show $\frac{m}{2^n} - G = 0$.

Say Left plays first.

If she plays in $-G$, it is to a number $-k$ with $k > \frac{m}{2^n}$,
so Right wins the resulting game $\frac{m}{2^n} - k$.

If she plays in $\frac{m}{2^n}$, it follows from the definitions (exercise) that it must be to a simpler number $k < \frac{m}{2^n}$.

So since $\frac{m}{2^n}$ is simplest,
k must be less than some Left-option H of G.

So Right can play the Right-option $-H$ of $-G$,
and win the resulting game $k - H$.

Exercise:
Use this lemma to confirm the following values of domineering games:

```
**    1          -1   ***   3
*   = -    *** =  -   * * = -
*     2     *     2   *     4
```

Lemma:
The usual addition of numbers agrees with addition of games.

Exercise:
Confirm, both by using addition and by thinking through strategies, that

```
*** ***
* * * * ** ***  = 0
*   *         *
```

More on Domineering

Exercise:
Confirm

```
               *
**            ***
** = *        *   = { 1 | 1 } = 1 + *
              *
```

There is however much more to games than numbers and nimbers.

Consider

```
**
**  = { 1 | -1 } =: +/- 1
```

Whoever moves in this game gets a free move for their efforts.

Exercise:
Confirm that for a number x,

- $\pm 1 > x$ if $x < 1$,
- $\pm 1 < x$ if $x > 1$,
- $\pm 1 \parallel x$ if $-1 \leq x \leq 1$.

Games like ± 1, where there is an advantage to moving, are called hot games.

For a positive number x, define $\pm x := \{x \mid -x\}$. These games are called switches.

Lemma:
If x and y are numbers with $x \geq y$, then
$\{x \mid y\} = \frac{x+y}{2} \pm \frac{x-y}{2}$.

Example:

```
***
*** = { 2 | -1/2 } = 3/4 +/- 5/4.
```

Roughly, it is sound strategy to play in the "hottest" component first.

Certainly this is true of games which are sums of switches, numbers and nimbers:
on your turn, you should play in the largest switch if there is one,
else if there are impartial components you should use the nim strategy to deal with them,
and finally you should (always) only play in a number if there's nothing else left.

This advice is far from being enough to let you win any domineering game (it's actually a hard game, tournaments have been played), but it's a start!

See pp114-121 of "On Numbers and Games" for much more on Domineering.

Further reading

The classic texts are

- E. Berlekamp, J. Conway, and R. Guy, "Winning Ways for your Mathematical Plays, vol 1";
- John Conway, "On Numbers and Games".

There are also two more recent books,

- M. Albert, R. Nowakowski, D. Wolfe, "Lessons in Play: An Introduction to Combinatorial Game Theory";
- Aaron Siegel, "Combinatorial Game Theory".